

CCAMP Working Group
IETF Internet draft
Proposed Status: Standard

JP Vasseur (Editor)
Cisco Systems, Inc.
Arthi Ayyangar (Editor)
Juniper Networks
Raymond Zhang
Infonet Service Corporation

Expires: August 2005

February 2005

draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt

A Per-domain path computation method for computing Inter-domain Traffic Engineering (TE) Label Switched Path (LSP)

Status of this Memo

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

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#). Internet-Drafts are Working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

Abstract

This document specifies a per-domain path computation method for computing inter-domain Traffic Engineering (TE) Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Label Switched (LSP) paths. In this document a domain is referred to as a collection of network elements within a common sphere of address management or path computational responsibility such as IGP areas and Autonomous Systems.

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

The principle of per-domain path computation specified in this document comprises of a GMPLS or MPLS node along an inter-domain LSP path, computing a path up to the last reachable hop within its domain. It covers cases where this last hop (domain exit point) may already be specified in the signaling message as well the case where this last hop may need to be determined by the GMPLS node.

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](#)].

Table of content

1. Terminology	3
2. Introduction	4
3. General assumptions	5
4. Per-Domain path computation algorithm	8
4.1 Example with an inter-area TE LSP	9
4.1.1 Case 1: T1 is a contiguous TE LSP	9
4.1.2 Case 2: T1 is a stitched or nested TE LSP	10
4.2 Example with an inter-AS TE LSP	10
4.2.1 Case 1: T1 is a contiguous TE LSP	12
4.2.2 Case 2: T1 is a stitched or nested TE LSP	13
5 Path optimality/diversity	13
6 MPLS Traffic Engineering Fast Reroute for inter-domain TE LSPs	13
6.1 Failure of an internal network element	13
6.2 Failure of an inter-ASBR links (inter-AS TE)	14
6.3 Failure of an ABR or an ASBR node	14
7. Reoptimization of an inter-domain TE LSP	14
7.1 Contiguous TE LSPs	14
7.2 Stitched or nested (non-contiguous) TE LSPs	15
8. Security Considerations	16
9. Intellectual Property Considerations	17
10. Acknowledgments	17
11. References	17
11.1 Normative references	17
11.2 Informative references	18

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

1. Terminology

LSR: Label Switch Router

LSP: MPLS Label Switched Path

PCE: Path Computation Element. An LSR in charge of computing TE LSP path for which it is not the Head-end. For instance, an ABR (inter-area) or an ASBR (Inter-AS) can play the role of PCE.

PCC: Path Computation Client (any head-end LSR) requesting a path computation from the Path Computation Element.

Local Repair: local protection techniques used to repair TE LSPs quickly when a node or link along the LSPs path fails.

Protected LSP: an LSP is said to be protected at a given hop if it has one or multiple associated backup tunnels originating at that hop.

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

PLR: Point of Local Repair. The head-end of a bypass tunnel.

MP: Merge Point. The LSR 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.

Fast Reroutable LSP: any LSP for which the "Local protection desired" bit is set in the Flag field of the SESSION_ATTRIBUTE object of its Path messages or signaled with a FAST-REROUTE object.

CSPF: Constraint-based Shortest Path First.

Inter-AS MPLS TE LSP: A TE LSP whose head-end LSR and tail-end LSR do not reside within the same Autonomous System (AS), or whose head-end LSR and tail-end LSR are both in the same AS but the TE LSP's path may be across different ASes. Note that this definition also applies to TE LSP whose Head-end and Tail-end LSRs reside in different sub-ASes (BGP confederations).

Inter-area MPLS TE LSP: A TE LSP where the head-end LSR and tail-end LSR do not reside in the same area or both the head-end and tail end LSR reside in the same area but the TE LSP transits one or more different areas along the path.

Vasseur, Ayyangar and Zhang

3

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

ABR Routers: routers used to connect two IGP areas (areas in OSPF or L1/L2 in IS-IS)

Interconnect routers or ASBR routers: routers used to connect together ASes of a different or the same Service Provider via one or more Inter-AS links.

Boundary LSR: a boundary LSR is either an ABR in the context of inter-area MPLS TE or an ASBR in the context of inter-AS MPLS TE.

TED: MPLS Traffic Engineering Database

The notion of contiguous, stitched and nested TE LSPs is defined in [[INTER-DOMAIN-SIG](#)] and will not be repeated here.

2. Introduction

The requirements for inter-area and inter-AS MPLS Traffic Engineering have been developed by the Traffic Engineering Working Group and have been stated in [INTER-AREA-REQS] and [INTER-AS-REQS] respectively. The framework for inter-domain MPLS Traffic Engineering has been provided in [INTER-DOMAIN-FRAMEWORK].

The set of mechanisms to establish and maintain inter-domain TE LSPs are specified in [[INTER-DOMAIN-SIG](#)].

This document exclusively focuses on the path computation aspects and defines a method for computing inter-domain TE LSP paths.

When the visibility of an end to end complete path spanning multiple

domains is not available at the head end node, one approach described in the document consists of using a per-domain path computation scheme used during LSP setup to determine the inter-domain LSP path as it traverses multiple domains.

Note that the mechanisms proposed in this document could also be applicable to MPLS TE domains other than areas and ASes.

According to the wide set of requirements defined in [INTER-AS-TE-REQS] and [INTER-AREA-TE-REQS], coming up with a single solution covering all the requirements is certainly possible but may not be desired: indeed, as described in [INTER-AS-TE-REQS] the spectrum of deployment scenarios is quite large and designing a solution addressing the super-set of all the requirements would lead to providing a rich set of mechanisms not required in several cases. Depending on the deployment scenarios of a SP, certain requirements stated above may be strict while certain other requirements may be relaxed.

There are different ways in which inter-domain TE LSP path computation may be performed. For example, if the requirement is to get an end-to-end constraint-based shortest path across multiple domains, then a

Vasseur, Ayyangar and Zhang

4

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

mechanism using one or more distributed PCEs could be used to compute the shortest path across different domains. Alternatively, one could also use some static or discovery mechanisms to determine the next boundary LSR per domain as the inter-domain TE LSP is being signaled. Other offline mechanisms for path computation are not precluded either. Depending on the Service Provider's requirements, one may adopt either of these techniques for inter-domain path computation.

Note that the adequate path computation method may be chosen based upon the TE LSP characteristics and requirements. This document specifies an inter-domain path computation technique based on per-domain path computation and could be used in place or in conjunction with other techniques.

3. General assumptions

In the rest of this document, we make the following set of assumptions:

1) Common assumptions

- Each domain in all the examples below is assumed to be capable of doing Traffic Engineering (i.e. running OSPF-TE or ISIS-TE and RSVP-TE). A domain may itself comprise multiple other domains. E.g. An AS

may itself be composed of several other sub-AS(es) (BGP confederations) or areas/levels.

- The inter-domain LSPs are signaled using RSVP-TE ([\[RSVP-TE\]](#)).
- The path (ERO) for the inter-domain TE LSP traversing multiple areas/ASes may be signaled as a set of (loose and/or strict) hops. The hops may identify:
 - The complete strict path end to end across different areas/ASes
 - The complete strict path in the source area/AS followed by boundary LSRs (and domain identifiers, e.g. AS numbers)
 - The complete list of boundary LSRs along the path
 - The current boundary LSR and the LSP destination

In this case, the set of (loose or strict) hops can either be statically configured on the Head-end LSR or dynamically computed. In the former case, the resulting path is statically configured on the Head-end LSR. In the latter case (dynamic computation), a per-domain path computation method is defined in this document with some Auto-discovery mechanism based on BGP and/or IGP information yielding the next-hop boundary node (domain exit point, say ABR/ASBR) along the path as the LSP is being signaled, along with crankback mechanisms. Note that alternatively next-hop may be statically configured on the Head-end LSR in which case next-hop auto-discovery would not be needed.

- Furthermore, the boundary LSRs are assumed to be capable of performing local path computation for expansion of a loose next-hop in

Vasseur, Ayyangar and Zhang

5

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

the signaled ERO if the path is not signaled by the head-end LSR as a set of strict hops or if the strict hop is an abstract node (e.g. an AS). This can be done by performing a CSPF computation up to that next loose hop as opposed to the TE LSP destination or by making use of some PCEs. In any case, no topology or resource information needs to be distributed between areas/ASes (as mandated per [\[INTER-AREA-REQS\]](#) and [\[INTER-AS-REQS\]](#)), which is critical to preserve IGP/BGP scalability and confidentiality in the case of TE LSPs spanning multiple routing domains.

Note that PCE-based path computation may be mentioned in this document for the sake of reference but such techniques are outside the scope of this document.

- The paths for the intra-domain FA-LSPs or LSP segments or for a contiguous TE LSP within the area/AS, may be pre-configured or computed

dynamically based on the arriving inter-domain LSP setup request; depending on the requirements of the transit area/AS. Note that this capability is explicitly specified as a requirement in [INTER-AS-TE-REQS]. When the paths for the FA-LSPs/LSP segments are pre-configured, the constraints as well as other parameters like local protection scheme for the intra-area/AS FA-LSP/LSP segment are also pre-configured.

- While certain constraints like bandwidth can be used across different areas/ASes, certain other TE constraints like resource affinity, color, metric, etc. as listed in [RFC2702] could be translated at areas/ASes boundaries. If required, it is assumed that, at the area/AS boundary LSRs, there will exist some sort of local mapping based on offline policy agreement, in order to translate such constraints across area/AS boundaries. It is expected that such an assumption particularly applies to inter-AS TE: for example, the local mapping would be similar to the Inter-AS TE Agreement Enforcement Policies stated in [INTER-AS-TE-REQS].

2) Example of topology for the inter-area TE case

The following example will be used for the inter-area TE case in this document.

```
<---area1--><---area0---><---area2----->
-----ABR1-----ABR1-----
|   /   |           |   \   |
R0--X1   |           |   X2--X3--R1
|         |           |   /   |
-----ABR2-----ABR2-----
<===== Inter-area TE LSP =====>
```

Assumptions

- ABR1, ABR2, ABR1 and ABR2 are ABRs,
- X1: an LSR in area 1,

Vasseur, Ayyangar and Zhang

6

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

- X2, X3: LSRs in area 2,
- An inter-area TE LSP T0 originated at R0 in area1 and terminating at R1 in area2,

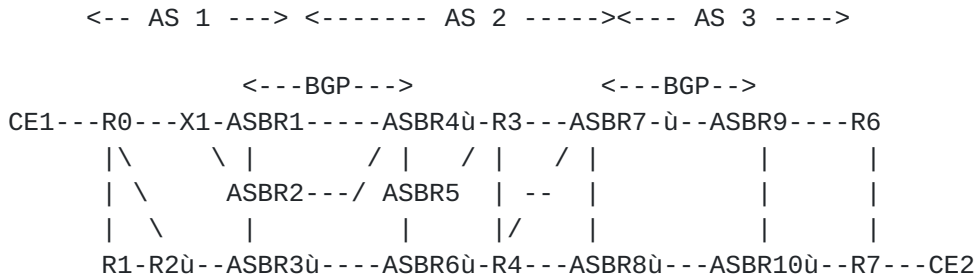
Notes:

- The terminology used in the example above corresponds to OSPF but the path computation methods proposed in this document equally applies to the case of an IS-IS multi-levels network.
- Just a few routers in each area are depicted in the diagram above for

the sake of simplicity.

3) Example of topology for the inter-AS TE case:

We will consider the following general case, built on a superset of the various scenarios defined in [INTER-AS-TE-REQS]:



<===== Inter-AS TE LSP(LSR to LSR)=====>

or

<===== Inter-AS TE LSP (CE to ASBR)=>

or

<===== Inter-AS TE LSP (CE to CE)=====>

The diagram above covers all the inter-AS TE deployment cases described in [INTER-AS-TE-REQS].

Assumptions:

- Three interconnected ASes, respectively AS1, AS2, and AS3. Note that AS3 might be AS1 in some scenarios described in [INTER-AS-TE-REQS],
- The various ASBRs are BGP peers, without any IGP running on the single hop links interconnecting the ASBRs and also referred to as inter-ASBR links,
- Each AS runs an IGP (IS-IS or OSPF) with the required IGP TE extensions (see [[OSPF-TE](#)] and [IS-IS-TE]). In other words, the ASes are TE enabled,

- Each AS can be made of several IGP areas. The path computation techniques described in this document applies to the case of a single AS made of multiple IGP areas, multiples ASes made of a single IGP

areas or any combination of the above. For the sake of simplicity, each routing domain will be considered as single area in this document.

- An inter-AS TE LSP T1 originated at R0 in AS1 and terminating at R6 in AS3.

4. Per-domain path computation algorithm

Regardless of the nature of the inter-domain TE LSP (contiguous, stitched or nested), a similar set of mechanisms for local TE LSP path computation (next hop resolution) can be used.

When an ABR/ASBR receives a Path message with a loose next-hop or an abstract node in the ERO, then it carries out the following actions:

- 1) It checks if the loose next-hop is accessible via the TED. If the loose next-hop is not present in the TED, then it checks if the next-hop at least has IP reachability (via IGP or BGP). If the next-hop is not reachable, then the path computation stops and the LSR sends back a PathErr upstream. If the next-hop is reachable, then it finds an ABR/ASBR to get to the next-hop. In the absence of an auto-discovery mechanism, the ABR in the case of inter-area TE or the ASBR in the next-hop AS in the case of inter-AS TE should be the loose next-hop in the ERO and hence should be accessible via the TED, otherwise the path computation for the inter-domain TE LSP will fail.

- 2) If the next-hop boundary LSR is present in the TED.

- a) Case of a contiguous TE LSP. The ABR/ASBR just performs an ERO expansion (unless not allowed by policy) after having computed the path to the next loose hop (ABR/ASBR) that obeys the set of required constraints. If no path satisfying the set of constraints can be found, the path computation stops and a Path Error MUST be sent for the inter-domain TE LSP.

- b) Case of stitched or nested LSP

- i) if the ABR/ASBR (receiving the LSP setup request) is a candidate LSR for intra-area FA-LSP/LSP segment setup, and if there is no FA-LSP/LSP segment from this LSR to the next-hop boundary LSR (satisfying the constraints) it SHOULD signal a FA-LSP/LSP segment to the next-hop boundary LSR. If pre-configured FA-LSP(s) or LSP segment(s) already exist, then it SHOULD try to select from among those intra-area/AS LSPs. Depending on local policy, it MAY signal a new FA-LSP/LSP segment if this selection fails. If the FA-LSP/LSP segment is successfully signaled or selected, it propagates the

inter-domain Path message to the next-hop following the procedures described in [LSP-HIER]. If, for some reason the dynamic FA-LSP/LSP segment setup to the next-hop boundary LSR fails, the path computation stops and a PathErr is sent upstream for the inter-domain LSP. Similarly, if selection of a preconfigured FA-LSP/LSP segment fails and local policy prevents dynamic FA-LSP/LSP segment setup, then the path computation stops and a PathErr is sent upstream for the inter-domain TE LSP.

ii) If, however, the boundary LSR is not a FA-LSP/LSP segment candidate, then it SHOULD simply compute a CSPF path up to the next-hop boundary LSR carry out an ERO expansion to the next-hop boundary LSR) and propagate the Path message downstream. The outgoing ERO is modified after the ERO expansion to the loose next-hop.

Note that in both cases, path computation may be stopped due to some local policy.

4.1. Example with an inter-area TE LSP

4.1.1. Case 1: T1 is a contiguous TE LSP

When the path message reaches ABR1, it first determines the egress LSR from its area 0 along the LSP path (say ABR#1), either directly from the ERO (if for example the next hop ABR is specified as a loose hop in the ERO) or by using some constraint-aware auto-discovery mechanism. In the former case, every inter-AS TE LSP path is defined as a set of loose and strict hops but at least the ABRs traversed by the inter-area TE LSP MUST be specified as loose hops on the head-End LSR.

- Example 1 (set of strict hops end to end): R0-X1-ABR1-ABR#1-X2-X3-R1
- Example 2 (set of loose hops): R0-ABR1(loose)-ABR#1(loose)-R1(loose)
- Example 3 (mix of strict and loose hops): R0-X1-ASBR1-ABR#1(loose)-X2-X3-R1

At least, the set of ABRs from the TE LSP head-end to the tail-End MUST be present in the ERO as a set of loose hops. Optionally, a set of paths can be configured on the head-end LSR, ordered by priority. Each priority path can be associated with a different set of constraints. Typically, it might be desirable to systematically have a last resort option with no constraint to ensure that the inter-area TE LSP could always be set up if at least a path exists between the inter-area TE LSP source and destination. Note that in case of set up failure or when an RSVP Path Error is received indicating the TE LSP has suffered a

failure, an implementation might support the possibility to retry a particular path option a specific amount of time (optionally with dynamic intervals between each trial) before trying a lower priority

path option. Any path can be defined as a set of loose and strict hops. In other words, in some cases, it might be desirable to rely on the dynamic path computation in some area, and exert a strict control on the path in other areas (defining strict hops).

Once it has computed the path up to the next ABR, ABR1 sends the Path message for the inter-area TE LSP to ABRÆ1. ABRÆ1 then repeats the a similar procedure and the Path message for the inter-area TE LSP will reach the destination R1. If ABRÆ1 cannot find a path obeying the set of constraints for the inter-area TE LSP, the path computation stops and ABRÆ1 MUST send a PathErr message to ABR1. Then ABR1 can in turn triggers a new computation by selecting another egress boundary LSR (ABRÆ2 in the example above) if crankback is allowed for this inter-area TE LSP (see [CRANBACK]). If crankback is not allowed for that inter-area TE LSP or if ABR1 has been configured not to perform crankback, then ABR1 MUST stop any path computation for the TE LSP and MUST forward a PathErr up to the head-end LSR (R0) without trying to select another egress LSR.

4.1.2. Case 2: T1 is a stitched or nested TE LSP

When the path message reaches ABR1, ABR1 first determines the egress LSR from its area 0 along the LSP path (say ABRÆ1), either directly from the ERO or by using some constraint-aware auto-discovery mechanism.

ABR1 will check if it has a FA-LSP or LSP segment to ABRÆ1 matching the constraints carried in the inter-area TE LSP Path message. If not, ABR1 will compute the path for a FA-LSP or LSP segment from ABR1 to ABRÆ1 satisfying the constraint and will set it up accordingly. Note that the FA-LSP or LSP segment could have also been pre-configured.

Once the ABR has selected the FA-LSP/LSP segment for the inter-area LSP, using the signaling procedures described in [[LSP-HIER](#)], ABR1 sends the Path message for inter-area TE LSP to ABRÆ1. Note that irrespective of whether ABR1 does nesting or stitching, the Path message for the inter-area TE LSP is always forwarded to ABRÆ1. ABRÆ1 then repeats the exact same procedures and the Path message for the inter-area TE LSP will reach the destination R1. If ABRÆ1 cannot find a path obeying the set of constraints for the inter-area TE LSP, then ABRÆ1 MUST send a PathErr message to ABR1. Then ABR1 can in turn either select another

FA-LSP/LSP segment to ABRÆ1 if such an LSP exists or select another egress boundary LSR (ABRÆ2 in the example above) if crankback is allowed for this inter-area TE LSP (see [CRANBACK]). If crankback is not allowed for that inter-area TE LSP or if ABR1 has been configured not to perform crankback, then ABR1 MUST forward a PathErr up to the inter-area head-end LSR (R0) without trying to select another egress LSR.

4.2. Example with an inter-AS TE LSP

Vasseur, Ayyangar and Zhang

10

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

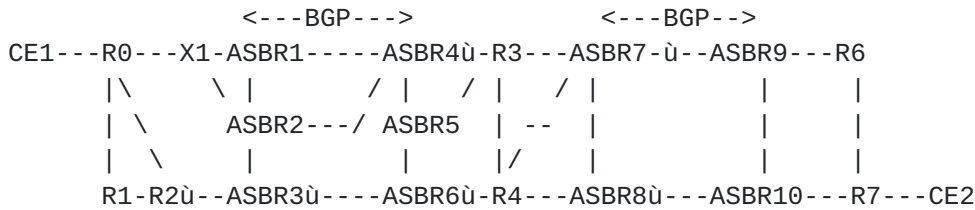
The procedures for establishing an inter-AS TE LSP are very similar to those of an inter-area TE LSP described above. The main difference is related to the presence of inter-ASBRs link(s).

The links interconnecting ASBRs are usually not TE enabled and no IGP is running at the AS boundaries. An implementation supporting inter-AS MPLS TE MUST obviously allow the set up of inter-AS TE LSP over the region interconnecting multiple ASBRs. In other words, an ASBR compliant with this document MUST support the set up of TE LSP over ASBR to ASBR links, performing all the usual operations related to MPLS Traffic Engineering (call admission control, à) as defined in [RSVP-TE].

In term of computation of an inter-AS TE LSP path, an interesting optimization consists of allowing the ASBRs to flood the TE information related to the inter-ASBR link(s) although no IGP TE is enabled over those links (and so there is no IGP adjacency over the inter-ASBR links). This of course implies for the inter-ASBR links to be TE-enabled although no IGP is running on those links. This allows a head-end LSR to make a more appropriate route selection up to the first ASBR in the next hop AS and will significantly reduce the number of signaling steps in route computation. This also allows the entry ASBR in an AS to make a more appropriate route selection up to the entry ASBR in the next hop AS taking into account constraints associated with the ASBR-ASBR links. Moreover, this reduces the risk of call set up failure due to inter-ASBR links not satisfying the inter-AS TE LSP set of constraints. Note that the TE information is only related to the inter-ASBR links: the TE LSA/LSP flooded by the ASBR includes not only the TE-enabled links contained in the AS but also the inter-ASBR links.

Note that no summarized TE information is leaked between ASes which is compliant with the requirements listed in [INTER-AREA-TE-REQS] and [INTER-AS-TE-REQS].

Example:



For instance, in the diagram depicted above, when ASBR1 floods its IGP TE LSA (opaque LSA for OSPF)/LSP (TLV 22 for IS-IS) in its routing domain, it reflects the reservation states and TE properties of the following links: X1-ASBR1, ASBR1-ASBR2 and ASBR1-ASBR4.

Thanks to such an optimization, the inter-ASBRs TE link information corresponding to the links originated by the ASBR is made available in the TED of other LSRs in the same area/AS that the ASBR belongs to.

Vasseur, Ayyangar and Zhang

11

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

Consequently, the CSPF computation for an inter-AS TE LSP path can also take into account the inter-ASBR link(s). This will improve the chance of successful path computation up to the next AS in case of a bottleneck on some inter-ASBR links and it potentially reduces one level of crankback. Note that no topology information is flooded and these links are not used in IGP SPF computations. Only the TE information for the links originated by the ASBR is advertised.

4.2.1. Case 1: T1 is a contiguous TE LSP

The inter-AS TE path may be configured on the head-end LSR as a set of strict hops, loose hops or a combination of both.

- Example 1 (set of strict hops end to end): R0-X1-ASBR1-ASBR4-ASBR5-R3-ASBR7-ASBR9-R6
- Example 2 (set of loose hops): R0-ASBR4(loose)-ASBR9(loose)-R6(loose)
- Example 3 (mix of strict and loose hops): R0-R2-ASBR3-ASBR2-ASBR1-ASBR4(loose)-ASBR10(loose)-ASBR9-R6

When a next hop is a loose hop, a dynamic path calculation (also called ERO expansion) is required taking into account the topology and TE information of its own AS and the set of TE LSP constraints. In the example 1 above, the inter-AS TE LSP path is statically configured as a set of strict hops; thus, in this case, no dynamic computation is required. Conversely, in the example 2, a per-AS path computation is performed, respectively on R0 for AS1, ASBR4 for AS2 and ASBR9 for AS3. Note that when an LSR has to perform an ERO expansion, the next hop

must either belong to the same AS, or must be the ASBR directly connected to the next hops AS. In this later case, the ASBR reachability MUST be announced in the IGP TE LSA/LSP originated by its neighboring ASBR. Indeed, in the example 2 above, the TE LSP path is defined as: R0-ASBR4(loose)-ASBR9(loose)-R6(loose). This implies that R0 must compute the path from R0 to ASBR4, hence the need for R0 to get the TE reservation state related to the ASBR1-ASBR4 link (flooded in AS1 by ASBR1). In addition, ASBR1 MUST also announce the IP address of ASBR4 specified in the T1 path configuration.

If an auto-discovery mechanism is available, every LSR receiving an RSVP Path message, will have to determine automatically the next hop ASBR, based on the IGP/BGP reachability of the TE LSP destination. With such a scheme, the head-end LSR and every downstream ASBR loose hop (except the last loose hop that computes the path to the final destination) automatically computes the path up to the next ASBR, the next loose hop based on the IGP/BGP reachability of the TE LSP destination. If a particular destination is reachable via multiple loose hops (ASBRs), local heuristics may be implemented by the head-end LSR/ASBRs to select the next hop an ASBR among a list of possible choices (closest exit point, metric advertised for the IP destination (ex: OSPF LSA External - Type 2), local policy,...). Once the next ASBR has been determined, an ERO expansion is performed as in the previous

Vasseur, Ayyangar and Zhang

12

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

case.

Once it has computed the path up to the next ASBR, ASBR1 sends the Path message for the inter-area TE LSP to ASBR4 (supposing that ASBR4 is the selected next hop ASBR). ASBR4 then repeats the exact same procedures and the Path message for the inter-AS TE LSP will reach the destination R1. If ASBR4 cannot find a path obeying the set of constraints for the inter-AS TE LSP, then ASBR4 MUST send a PathErr message to ASBR1. Then ASBR1 can in turn either select another ASBR (ASBR5 in the example above) if crankback is allowed for this inter-AS TE LSP (see [CRANBACK]). If crankback is not allowed for that inter-AS TE LSP or if ASBR1 has been configured not to perform crankback, then ABR1 MUST stop the path computation and MUST forward a PathErr up to the head-end LSR (R0) without trying to select another egress LSR. In this case, the head-end LSR can in turn select another sequence of loose hops, if configured. Alternatively, the head-end LSR may decide to retry the same path; this can be useful in case of set up failure due an outdated IGP TE database in some downstream AS. An alternative could also be for the head-end LSR to retry to same sequence of loose hops after having

relaxed some constraint(s).

4.2.2. Case 2: T1 is a stitched or nested TE LSP

The signaling procedures are very similar to the inter-area LSP setup case described earlier. In this case, the FA-LSPs or LSP segments will only be originated by the ASBRs at the entry to the AS.

5. Path optimality/diversity

Since the inter-domain path is computed on a per domain (area, AS) basis, one cannot guarantee that the shortest inter-domain path can be found.

Moreover, computing two diverse paths might not be possible in some topologies (due to the well-known trapping problem).

As already pointed out, the required path computation method can be selected by the Operator on a per LSP basis.

6. MPLS Traffic Engineering Fast Reroute for inter-domain TE LSPs

The signaling aspects of Fast Reroute and related constraints for each TE LSP types in the case of inter-domain TE LSP has been covered in [[INTER-DOMAIN-SIG](#)] and will not be repeated in this document.

There are multiple failure scenarios to consider in the case of Fast Reroute for inter-domain TE LSPs.

6.1. Failure of an internal network element

Vasseur, Ayyangar and Zhang

13

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

The case of a failure of a network element within an area/AS such as a link, SRLG or a node does not differ from Fast Reroute for intra-domain TE LSP.

6.2. Failure of an inter-ASBR links (inter-AS TE)

In order to protect inter-domain TE LSPs from the failure of an inter-ASBR link, this requires the computation of a backup tunnel path that crosses a non IGP TE-enabled region (between two ASes). If the inter-ASBR TE related information is flooded in the IGP, each ASBR is capable of computing the path according to the backup tunnel constraints. Otherwise, the backup tunnel path MUST be statically configured.

6.3. Failure of an ABR or an ASBR node

The constraints to be taken into account during the backup tunnel path computation significantly differs upon the TE LSP type, network element to protect (entry/exit boundary node) and the Fast Reroute method in use (facility backup versus one-to-one). Those constraints have been explored in detail in [[INTER-DOMAIN-SIG](#)] but since the backup tunnel is itself an inter-domain TE LSP, its path computation can be performed according to the two path computation methods described in this document.

7. Reoptimization of an inter-domain TE LSP

The ability to reoptimize an existing inter-domain TE LSP path is of course a requirement. The reoptimization process significantly differs based upon the nature of the TE LSP and the mechanism in use for the TE LSP path computation.

The following mechanisms can be used for re-optimization, which are dependent on the nature of the inter-domain TE LSP.

7.1. Contiguous TE LSPs

After an inter-AS TE LSP has been set up, a more optimal route might appear in the various traversed ASes. Then in this case, it is desirable to get the ability to reroute an inter-AS TE LSP in a non-disruptive fashion (making use of the so called Make Before Break procedure) to follow this more optimal path. This is known as a TE LSP reoptimization procedure.

[LOOSE-REOPT] proposes a mechanisms allowing:

- The head-end LSR to trigger on every LSR whose next hop is a loose hop the re evaluation of the current path in order to detect a potentially more optimal path. This is done via explicit signaling request: the head-end LSR sets the `ERO`

Expansion requestö bit of the SESSION-ATTRIBUTE object carried in the RSVP Path message.

- An LSR whose next hop is a loose-hop to signal to the head-end LSR that a better path exists. This is performed by sending an RSVP Path Error Notify message (ERROR-CODE = 25), sub-code 6

(Better path exists).

This indication may be sent either:

- In response to a query sent by the head-end LSR,
- Spontaneously by any LSR having detected a more optimal path

Such a mechanism allows for the reoptimization of a TE LSP if and only if a better path in some downstream area/AS is detected.

The reoptimization event can either be timer or event-driven based (a link UP event for instance).

Note that the reoptimization MUST always be performed in a non-disruptive fashion.

Once the head-end LSR is informed of the existence of a more optimal path either in its head-end area/AS or in another AS, the inter-AS TE Path computation is triggered using the same set of mechanisms as when the TE LSP is first set up. Then the inter-AS TE LSP is set up following the more optimal path, making use of the make before break procedure. In case of a contiguous LSP, the reoptimization process is strictly controlled by the head-end LSR which triggers the make-before-break procedure, regardless of the location where the more optimal path is.

Note that in the case of loose hop reoptimization, the TE LSP may follow a preferable path within one or more domain(s) whereas in the case of PCE-based path computation techniques, the reoptimization process may lead to following a completely different inter-domain path (including a different set of ABRs/ASBRs) since end-to-end shortest path is computed.

7.2. Stitched or nested (non-contiguous) TE LSPs

In the case of a stitched or nested inter-domain TE LSP, the re-optimization process is treated as a local matter to any Area/AS. The main reason is that the inter-domain TE LSP is a different LSP (and therefore different RSVP session) from the intra-domain LSP segment or FA-LSP in an area or an AS. Therefore, reoptimization in an area/AS is done by locally reoptimizing the intra-domain FA LSP or LSP segment. Since the inter-domain TE LSPs are transported using LSP segments or FA-LSP across each domain, optimality of the inter-domain TE LSP in an area/AS is dependent on the optimality of the corresponding LSP

segments or FA-LSPs. If, after an inter-domain LSP is setup, a more optimal path is available within an area/AS, the corresponding LSP segment(s) or FA-LSP will be re-optimized using "make-before-break" techniques discussed in [[RSVP-TE](#)]. Reoptimization of the FA LSP or LSP segment automatically reoptimizes the inter-domain TE LSPs that the LSP segment transports. Reoptimization parameters like frequency of reoptimization, criteria for reoptimization like metric or bandwidth availability; etc can vary from one area/AS to another and can be configured as required, per intra-area/AS TE LSP segment or FA-LSP if it is preconfigured or based on some global policy within the area/AS.

Hence, in this scheme, since each area/AS takes care of reoptimizing its own LSP segments or FA-LSPs, and therefore the corresponding inter-domain TE LSPs, the make-before-break can happen locally and is not triggered by the head-end LSR for the inter-domain LSP. So, no additional RSVP signaling is required for LSP re-optimization and reoptimization is transparent to the HE LSR of the inter-domain TE LSP.

If, however, an operator desires to manually trigger reoptimization at the head-end LSR for the inter-domain TE LSP, then this solution does not prevent that. A manual trigger for reoptimization at the head-end LSR SHOULD force a reoptimization thereby signaling a "new" path for the same LSP (along the optimal path) making use of the make-before-break procedure. In response to this new setup request, the boundary LSR may either initiate new LSP segment setup, in case the inter-domain TE LSP is being stitched to the intra-area/AS LSP segment or it may select an existing or new FA-LSP in case of nesting. When the LSP setup along the current optimal path is complete, the head end should switchover the traffic onto that path and the old path is eventually torn down. Note that the head-end LSR does not know a priori whether a more optimal path exists. Such a manual trigger from the head-end LSR of the inter-domain TE LSP is, however, not considered to be a frequent occurrence.

Note that stitching or nesting rely on local optimization: the reoptimization process allows to locally reoptimize each TE LSP segment or FA-LSP: hence, the reoptimization is not global and consequently the end to end path may no longer be optimal, should it be optimal when set up.

Procedures described in [[LOOSE-REOPT](#)] MUST be used if the operator does not desire local re-optimization of certain inter-domain LSPs. In this case, any re-optimization event within the domain MUST be reported to the head-end node. This SHOULD be a configurable policy.

8. Security Considerations

When signaling an inter-AS TE, an Operator may make use of the already defined security features related to RSVP (authentication). This may require some coordination between Service Providers to share the keys

(see [RFC 2747](#) and [RFC 3097](#)).

Vasseur, Ayyangar and Zhang

16

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

9. Intellectual Property Considerations

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

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

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

IPR Disclosure Acknowledgement

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

10. Acknowledgments

We would like to acknowledge input and helpful comments from Adrian Farrel.

11 References

11.1. Normative References

[RFC] Bradner, S., "Key words for use in RFCs to indicate requirements levels", [RFC 2119](#), March 1997.

[RFC3667] Bradner, S., "IETF Rights in Contributions", [BCP 78](#), RFC 3667, February 2004.

[RFC3668] Bradner, S., Ed., "Intellectual Property Rights in IETF Technology", [BCP 79](#), [RFC 3668](#), February 2004.

Vasseur, Ayyangar and Zhang

17

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#) February 2005

[RSVP] Braden, et al, " Resource ReSerVation Protocol (RSVP) ù Version 1, Functional Specificationö, [RFC 2205](#), September 1997.

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

[REFRESH-REDUCTION] Berger et al, öRSVP Refresh Overhead Reduction Extensionsö, [RFC2961](#), April 2001.

[FAST-REROUTE] Ping Pan, et al, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", [draft-ietf-mpls-rsvp-lsp-fastreroute-03.txt](#), December 2003.

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

[ISIS-TE] Li, T., Smit, H., "IS-IS extensions for Traffic Engineering", [RFC 3784](#), June 2004.

[11.2.](#) Informative references

[INT-AREA-REQ] Le Roux, J.L., Vasseur, J.P., Boyle, J., "Requirements for inter-area MPLS Traffic Engineering", [draft-ietf-tewg-interarea-mpls-te-req-03.txt](#), work in progress.

[INT-AS-REQ] Zhang, R., Vasseur, J.P., "MPLS Inter-AS Traffic Engineering Requirements", [draft-ietf-tewg-interas-mpls-te-req-09.txt](#), work in progress.

[INT-DOMAIN-FRWK] Farrel, A., Vasseur, J.P., Ayyangar, A., "A Framework for Inter-Domain MPLS Traffic Engineering", [draft-ietf-ccamp-inter-domain-framework-00.txt](#), work in progress.

[FACILITY-BACKUP] Le Roux, J.L., Vasseur, J.P. et al. "Framework for PCE based MPLS Facility Backup Path Computation", [draft-leroux-pce-backup-comp-frwk-00.txt](#), work in progress

[INTER-DOMAIN-SIG] Ayyangar, A., Vasseur, JP. "Inter-domain MPLS Traffic Engineering - RSVP extensions", [draft-ayyengar-ccamp-inter-domain-rsvp-te](#), work in progress.

[LSP-ATTRIBUTE] Farrel A. et al, "Encoding of Attributes for Multiprotocol Label Switching (MPLS) Label Switched Path (LSP) Establishment Using RSVP-TE", [draft-ietf-mpls-rsvpte-attributes-04](#), (work in progress).

[GMPLS-OVERLAY] G. Swallow et al, "GMPLS RSVP Support for the Overlay Model", (work in progress).

[EXCLUDE-ROUTE] Lee et al, "Exclude Routes - Extension to RSVP-TE", [draft-ietf-ccamp-rsvp-te-exclude-route-00.txt](#), work in progress.

Vasseur, Ayyangar and Zhang

18

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

[LSPPING] Kompella, K., Pan, P., Sheth, N., Cooper, D., Swallow, G., Wadhwa, S., Bonica, R., "Detecting Data Plane Liveliness in MPLS", Internet Draft <[draft-ietf-mpls-lsp-ping-02.txt](#)>, October 2002. (Work in Progress)

[MPLS-TTL], Agarwal, et al, "Time to Live (TTL) Processing in MPLS Networks", [RFC 3443](#) Updates [RFC 3032](#) ", January 2003

[LOOSE-PATH-REOPT] Vasseur, Ikejiri and Zhang "Reoptimization of an explicit loosely routed MPLS TE paths", [draft-ietf-ccamp-loose-path-reopt-01.txt](#), July 2004, Work in Progress.

[NODE-ID] Vasseur, Ali and Sivabalan, "Definition of an RRO node-id subobject", [draft-ietf-mpls-nodeid-subobject-03.txt](#), work in progress.

[LSP-HIER] Kompella K., Rekhter Y., "LSP Hierarchy with Generalized MPLS TE", [draft-ietf-mpls-lsp-hierarchy-08.txt](#), March 2002.

[MPLS-TTL], Agarwal, et al, "Time to Live (TTL) Processing in MPLS Networks", [RFC 3443](#) (Updates [RFC 3032](#)) ", January 2003.

Authors' Address:

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

Email: jpv@cisco.com

Arthi Ayyangar (Editor)
Juniper Networks, Inc
[1194 N.Mathilda Ave](#)
Sunnyvale, CA 94089
USA
e-mail: arthi@juniper.net

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

Full Copyright Statement

Vasseur, Ayyangar and Zhang

19

[draft-vasseur-ccamp-inter-domain-pd-path-comp-00.txt](#)

February 2005

Copyright (C) The Internet Society (2005). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

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

