Network Working Group
Request for Comments: 4927
Category: Informational

J.-L. Le Roux, Ed.
France Telecom

June 2007

# Path Computation Element Communication Protocol (PCECP) Specific Requirements for Inter-Area MPLS and GMPLS Traffic Engineering

## Status of This Memo



## Copyright Notice



## Abstract

For scalability purposes, a network may comprise multiple Interior Gateway Protocol (IGP) areas. An inter-area Traffic Engineered Label Switched Path (TE-LSP) is an LSP that transits through at least two IGP areas. In a multi-area network, topology visibility remains local to a given area, and a head-end Label Switching Router (LSR) cannot compute an inter-area shortest constrained path. One key application of the Path Computation Element (PCE)-based architecture is the computation of inter-area TE-LSP paths. The PCE Communication Protocol (PCECP) is used to communicate computation requests from Path Computation Clients (PCCs) to PCEs, and to return computed paths in responses. This document lists a detailed set of PCECP-specific requirements for support of inter-area TE-LSP path computation. It complements the generic requirements for a PCE Communication Protocol.

Table of Contents



## 1. Introduction

[RFC4105] lists a set of motivations and requirements for setting up TE-LSPs across IGP area boundaries. These LSPs are called inter-area TE-LSPs. These requirements include the computation of inter-area shortest constrained paths with a key guideline being to respect the IGP hierarchy concept, and particularly the containment of topology information. The main challenge with inter-area MPLS-TE lies in path computation. Indeed, the head-end LSR cannot compute an explicit path across areas, as its topology visibility is limited to its own area.

Inter-area path computation is one of the key applications of the PCE-based architecture [RFC4655]. The computation of optimal inter-area paths may be achieved using the services of one or more PCEs.

Such PCE-based inter-area path computation could rely for instance on a single multi-area PCE that has the TE database of all the areas in the IGP domain and can directly compute an end-to-end constrained shortest path. Alternatively, this could rely on the cooperation between PCEs whereby each PCE covers one or more IGP areas and the full set of PCEs covers all areas.

The generic requirements for a PCE Communication Protocol (PCECP), which allows a PCC to send path computation requests to a PCE and the PCE to send path computation responses to a PCC, are set forth in [RFC4657]. The use of a PCE-based approach for inter-area path computation implies specific requirements on a PCE Communication Protocol, in addition to the generic requirements already listed in [RFC4657]. This document complements these generic requirements by listing a detailed set of PCECP requirements specific to inter-area path computation.

It is expected that PCECP procedures be defined to satisfy these requirements.

Note that PCE-based inter-area path computation may require a mechanism for automatic PCE discovery across areas, which is out of the scope of this document. Detailed requirements for such a mechanism are discussed in [RFC4674].

## 2. Terminology

LSR: Label Switching Router.

LSP: MPLS Label Switched Path.

TE-LSP: Traffic Engineered Label Switched Path.

IGP area: OSPF area or IS-IS level.

ABR: IGP Area Border Router, a router that is attached to more than one IGP area (ABR in OSPF or L1/L2 router in IS-IS).

Inter-Area TE-LSP: TE-LSP that traverses more than one IGP area.

CSPF: Constrained Shortest Path First.

SRLG: Shared Risk Link Group.

PCE: Path Computation Element: an entity (component, application or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints.

PCC: Path Computation Client, any application that request path computation to be performed by a PCE.

PCECP: PCE Communication Protocol, a protocol for communication between PCCs and PCEs, and between PCEs.

ERO: Resource Reservation Protocol (RSVP)-TE Explicit Route Object. It encodes the explicit path followed by a TE-LSP.

## 2.1. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

# 3. Motivations for PCE-Based Inter-Area Path Computation

IGP hierarchy enables improved IGP scalability by dividing the IGP domain into areas and limiting the flooding scope of topology information to within area boundaries. A router in an area has full topology information for its own area, but only information about reachability to destinations in other areas. Thus, a head-end LSR cannot compute an end-to-end path that crosses the boundary of its IGP area(s).

A current solution for computing inter-area TE-LSP path relies on a per-domain path computation [PD-COMP]. It is based on loose hop routing with an ERO expansion on each ABR. This allows an LSP to be set up following a constrained path, but faces two major limitations:

- This does guarantee the use of an optimal constrained path.

- This may lead to several crankback signaling messages and hence delay the LSP setup, and may also invoke possible alternate routing activities.

Note that, here, by optimal constrained path we mean the shortest constrained path across multiple areas, taking into account either the IGP or TE metric [RFC3785]. In other words, such a path is the path that would have been computed by making use of some CSPF algorithm in the absence of multiple IGP areas.

The PCE-based architecture [RFC4655] is well suited to inter-area path computation. It allows the path computation limitations resulting from the limited topology visibility to be overcome by introducing path computation entities with more topology visibility, or by allowing cooperation between path computation entities in each area.

There are two main approaches for the computation of an inter-area optimal path:

- Single-PCE computation: The path is computed by a single PCE that has topology visibility in all areas and can compute an end-to-end optimal constrained path on its own.

- Multiple-PCE computation with inter-PCE communication: The path computation is distributed on multiple PCEs, which have partial topology visibility. They compute path segments in their domains of visibility and collaborate with each other so as to arrive at an end-to-end optimal constrained path. Such collaboration is ensured thanks to inter-PCE communication.

Note that the use of a PCE-based approach to perform inter-area path computation implies specific functional requirements in a PCECP, in addition to the generic requirements listed in [RFC4657]. These specific requirements are discussed in the next section.

## 4. Detailed Inter-Area Specific Requirements on PCECP

This section lists a set of additional requirements for the PCECP that complement requirements listed in [RFC4657] and are specific to inter-area (G)MPLS-TE path computation.

### 4.1. Control and Recording of Area Crossing

In addition to the path constraints specified in [RFC4657], the request message MUST allow indicating whether or not area crossing is permitted. Indeed, when the source and destination reside in the same IGP area, there may be intra-area and inter-area feasible paths. As set forth in [RFC4105], if the shortest path is an inter-area path, an operator either may want to avoid, as far as possible, crossing areas and thus may prefer selecting a sub-optimal intra-area path or, conversely, may prefer to use a shortest path, even if it crosses areas.

Also, when the source and destination reside in the same area it may be useful to know whether the returned path is an inter-area path. Hence, the response message MUST allow indicating whether the computed path is crossing areas.

### 4.2. Area Recording

It may be useful for the PCC to know the set of areas crossed by an inter-area path and the corresponding path segments. Hence, the response message MUST allow identifying the crossed areas. Also, the response message MUST allow segmenting the returned path and marking each segment so that it is possible to tell which piece of the path lies within which area.

### 4.3. Strict Explicit Path and Loose Path

A Strict Explicit Path is defined as a set of strict hops, while a Loose Path is defined as a set of at least one loose hop and zero, one or more strict hops. An inter-area path may be strictly explicit or loose (e.g., a list of ABRs as loose hops). It may be useful to indicate to the PCE if a Strict Explicit path is required or not. Hence, the PCECP request message MUST allow indicating whether a Strict Explicit Path is required/desired.

### 4.4. PCE List Enforcement and Recording in Multiple-PCE Computation

In case of multiple-PCE inter-area path computation, a PCC may want to indicate a preferred list of PCEs to be used, one per area. In each area, the preferred PCE should be tried before another PCE is selected. Note that if there is no preferred PCE indicated for an area, any PCE in that area may be used.

Hence, the PCECP request message MUST support the inclusion of a list of preferred PCEs per area. Note that this requires that a PCC in one area has knowledge of PCEs in other areas. This could rely on configuration or on a PCE discovery mechanism, allowing discovery across area boundaries (see [RFC4674]).

Also, it would be useful to know the list of PCEs that effectively participated in the computation. Hence, the request message MUST support a request for PCE recording, and the response message MUST support the recording of the set of one or more PCEs that took part in the computation.

It may also be useful to know the path segments computed by each PCE. Hence, the request message SHOULD allow a request for the identification of path segments computed by a PCE, and the response message SHOULD allow identifying the path segments computed by each PCE.

### 4.5. Inclusion of Area IDs in Request

Knowledge of the areas in which the source and destination lie would allow a PCE to select an appropriate downstream PCE. This would be useful when the area ID(s) of a PCE (i.e., the area(s) where it has visibility) is/are known, which can be achieved by the PCE Discovery Protocol (see [RFC4674]) or by any other means.

A PCE may not have any visibility of the source/destination area and hence may not be able to determine the area of the source/destination. In such a situation, it would be useful for a PCC to indicate the source and destination area IDs in its request message.

For that purpose, the request message MUST support the inclusion of the source and destination area IDs. Note that this information could be learned by the PCC through configuration.

### 4.6. Area Inclusion/Exclusion

In some situations, it may be useful for the request message to indicate one or more area(s) that must be followed by the path to be computed. It may also be useful for the request message to indicate one or more area(s) that must be avoided by the path to be computed (e.g., request for a path between LSRs in two stub areas connected to the same ABR(s), which must not cross the backbone area). Hence, the request message MUST allow indicating a set of one or more area(s) that must be explicitly included in the path, and a set of one or more area(s) that must be explicitly excluded from the path.

### 4.7. Inter-Area Diverse Path Computation

For various reasons, including protection and load balancing, the computation of diverse inter-area paths may be required. There are various levels of diversity in an inter-area context:

- Per-area diversity (intra-area path segments are link, node, or SRLG disjoint)

- Inter-area diversity (end-to-end inter-area paths are link, node, or SRLG disjoint)

Note that two paths may be disjoint in the backbone area but non-disjoint in peripheral areas. Also two paths may be node-disjoint within areas but may share ABRs, in which case path segments within an area are node-disjoint, but end-to-end paths are not node disjoint.

The request message MUST allow requesting the computation of a set of inter-area diverse paths between the same node pair or between distinct node pairs. It MUST allow indicating the required level of diversity of a set of inter-area paths (link, node, and SRLG diversity), as well as the required level of diversity of a set of intra-area segments of inter-area paths (link, node, and SRLG diversity) on a per-area basis.

The response message MUST allow indicating the level of diversity of a set of computed inter-area loose paths (link, node, and SRLG diversity), globally, and on a per-area basis (link, node, and SRLG diversity of intra-area path segments).

Note that, in order to ensure SRLG consistency, SRLG identifiers within the IGP domain should be assigned and allocated by the same entity.

Note that specific objective functions may be requested for diverse path computation, such as minimizing the cumulated cost of a set of diverse paths as set forth in [RFC4657].

### 4.8. Inter-Area Policies

In addition to the policy requirements discussed in [RFC4657], the application of inter-area path computation policies requires some additional information to be carried in the PCECP request messages. The request message MUST allow for the inclusion of the address of the originating PCC. This may be useful in a multiple-PCE computation, so as to apply policies not only based on the PCECP peer but also based on the originating PCC.

Note that work on supported policy models and the corresponding requirements/implications is being undertaken as a separate work item in the PCE working group [PCE-POL-FMWK].

### 4.9. Loop Avoidance

In case of multiple-PCE inter-area path computation, there may be risks of PCECP request loops. A mechanism MUST be defined to detect and correct PCECP request message loops. This may rely, for instance, on the recording, in the request message, of the set of traversed PCEs.

Also, the returned path in a response message MUST be loop free.

## 5. Manageability Considerations

The inter-area application implies some new manageability requirements in addition to those already listed in [RFC4657]. The PCECP PCC and PCE MIB modules MUST allow recording the proportion of inter-area requests and the success rate of inter-area requests. The PCECP PCC MIB module MUST also allow recording the performances of a PCE chain (minimum, maximum, and average response times), in case of multiple-PCE inter-area path computation.

It is really important, for diagnostic and troubleshooting reasons, to monitor the availability and performances of each PCE of a PCE chain used for inter-area path computation. Particularly, it is really important to identify the PCE(s) responsible for a delayed reply.

Hence, a mechanism MUST be defined to monitor the performances of a PCE chain. It MUST allow determining the availability of each PCE of the chain as well as its minimum, maximum, and average response times.

## 6. Security Considerations

IGP areas are administrated by the same entity. Hence, the inter-area application does not imply a new trust model or new security issues beyond those already defined in [RFC4657].

## 7. Acknowledgments

We would also like to thank Adrian Farrel, Jean-Philippe Vasseur, Bruno Decraene, Yannick Le Louedec, Dimitri Papadimitriou, and Lou Berger for their useful comments and suggestions. Thanks also to Ross Callon, Catherine Meadow, and Dan Romascanu for their review during the final stages of publication.

## 8. References

### 8.1. Normative References

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

[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.

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

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

## 8.2. Informative References

[RFC4674] Le Roux, J., Ed., "Requirements for Path Computation Element (PCE) Discovery", RFC 4674, October 2006.

[PD-COMP] Vasseur, J.P., Ed., Ayyangar, A., Ed., and R. Zhang, "A Per-domain path computation method for computing Inter-domain Traffic Engineering (TE) Label Switched Path (LSP)", Work in Progress, April 2007.

[PCE-POL-FMWK] Bryskin, I., Papadimitriou, D., Berger L., and J. Ash, "Policy-Enabled Path Computation Framework", Work in Progress, March 2007.

[RFC3785] Le Faucheur, F., Uppili, R., Vedrenne, A., Merckx, P., and T. Telkamp, "Use of Interior Gateway Protocol (IGP) Metric as a second MPLS Traffic Engineering (TE) Metric", BCP 87, RFC 3785, May 2004.

## 9. Contributors

Jerry Ash
AT&T
Room MT D5-2A01
200 Laurel Avenue
Middletown, NJ 07748, USA
Phone: +1-(732)-420-4578
EMail: gash5107@yahoo.com

Nabil Bitar
Verizon
40 Sylvan Road
Waltham, MA 02145
EMail: nabil.n.bitar@verizon.com

Dean Cheng
Cisco Systems Inc.
3700 Cisco Way
San Jose, CA 95134 USA
Phone: +1 408 527 0677
EMail: dcheng@cisco.com

Kenji Kumaki
KDDI Corporation
Garden Air Tower
Iidabashi, Chiyoda-ku,
Tokyo 102-8460, JAPAN
Phone: +81-3-6678-3103
EMail: ke-kumaki@kddi.com

Eiji Oki
NTT
Midori-cho 3-9-11
Musashino-shi, Tokyo 180-8585, JAPAN
EMail: oki.eiji@lab.ntt.co.jp

Raymond Zhang
BT
2160 E. Grand Ave.
El Segundo, CA 90245
USA
EMail: raymond.zhang@bt.com

Renhai Zhang
Huawei Technologies
No. 3 Xinxi Road, Shangdi,
Haidian District,
Beijing City,
P. R. China
EMail: zhangrenhai@huawei.com

Editor's Address

Jean-Louis Le Roux
France Telecom
2, avenue Pierre-Marzin
22307 Lannion Cedex
FRANCE
EMail: jeanlouis.leroux@orange-ftgroup.com

Acknowledgement

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