Network Working Group
Request for Comments: 5376
Category: Informational

N. Bitar
Verizon
R. Zhang
BT
K. Kumaki
KDDI R&D Labs

November 2008

Inter-AS Requirements for the Path Computation Element Communication Protocol (PCECP)

Status of This Memo

Copyright Notice

Abstract

Multiprotocol Label Switching Traffic Engineered (MPLS TE) Label Switched Paths (LSPs) may be established wholly within an Autonomous System (AS) or may cross AS boundaries.

The Path Computation Element (PCE) is a component that is capable of computing constrained paths for (G)MPLS TE LSPs. The PCE Communication Protocol (PCECP) is defined to allow communication between Path Computation Clients (PCCs) and PCEs, as well as between PCEs. The PCECP is used to request constrained paths and to supply computed paths in response. Generic requirements for the PCECP are set out in "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657. This document extends those requirements to cover the use of PCECP in support of inter-AS MPLS TE.

Table of Contents

1. Introduction

[RFC4216] defines the scenarios motivating the deployment of inter-AS Multiprotocol Label Switching Traffic Engineering (MPLS TE) and specifies the requirements for inter-AS MPLS TE when the ASes are under the administration of one Service Provider (SP) or the administration of different SPs.

Three signaling options are defined for setting up an inter-AS TE Label Switched Path (LSP):

1) contiguous TE LSP as documented in [RFC5151];
2) stitched inter-AS TE LSP discussed in [RFC5150];
3) nested TE LSP as in [RFC4206].

[RFC5152] defines mechanisms for the computation of inter-domain TE LSPs using network elements along the signaling paths to compute per-domain constrained path segments. The mechanisms in [RFC5152] do not guarantee an optimum constrained path across multiple ASes where an optimum path for a TE LSP is one that has the smallest cost, according to a normalized TE metric (based upon a TE metric or Interior Gateway Protocol (IGP) metric adopted in each transit AS) among all possible paths that satisfy the LSP TE constraints.

The Path Computation Element (PCE) [RFC4655] is a component that is capable of computing paths for MPLS TE and Generalized Multiprotocol Label Switching Protocol ((G)MPLS TE) LSPs. The requirements for a PCE have come from SP demands to compute optimum constrained paths across multiple areas and/or domains, and to be able to separate the path computation elements from the forwarding elements.

The PCE Communication Protocol (PCECP) is defined to allow communication between Path Computation Clients (PCCs) and PCEs, and between PCEs. The PCECP is used to request (G)MPLS TE paths and to supply computed paths in response. Generic requirements for the PCECP are discussed in [RFC4657]. This document provides a set of PCECP requirements that are specific to inter-AS (G)MPLS TE path computation.

2. Terminology

This document adopts the definitions and acronyms defined in Section 3 of [RFC4216] and Section 2 of [RFC4655]. In addition, we use the following terminology:

ASBR: Autonomous System Border Router (see section 3 of RFC 4216)

PCECP: PCE Communication Protocol

(G)MPLS TE: MPLS or Generalized MPLS Traffic Engineering

Inter-AS (G)MPLS TE path: An MPLS TE or Generalized MPLS (GMPLS) path that traverses two or more ASes.

Intra-AS (G)MPLS TE path: An MPLS TE or GMPLS path that is confined to a single AS. It may traverse one or more IGP areas.

Intra-AS PCE: A PCE responsible for computing (G)MPLS TE paths remaining within a single AS.

Inter-AS PCE: A PCE responsible for computing inter-AS (G)MPLS paths or path segments, possibly by cooperating with intra-AS PCEs.

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.

3. Reference Model

Figure 1 depicts the reference model for PCEs in an inter-AS application. We refer to two types of PCE functions in this document: inter-AS PCEs and intra-AS PCEs. Inter-AS PCEs perform the procedures needed for inter-AS (G)MPLS TE path computation while intra-AS PCEs perform the functions needed for intra-AS (G)MPLS TE path computation.

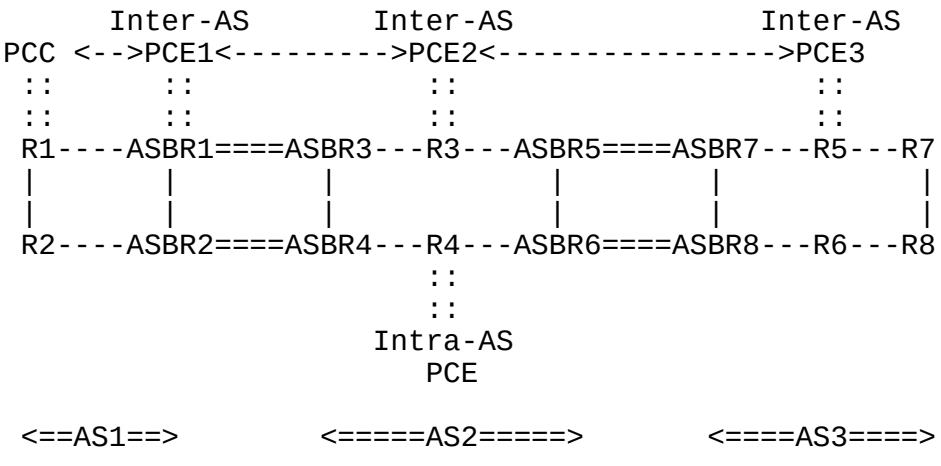

Figure 1: Inter- and Intra-AS PCE Reference Model

Let's follow a scenario that illustrates the interaction among PCCs, inter-AS PCEs, and intra-AS PCEs, as shown in Figure 1. R1 in AS1 wants to setup a (G)MPLS TE path, call it LSP1, with certain constraints to R7 in AS3. R1 determines, using mechanisms out of the

scope of this document, that R7 is an inter-AS route and that R1 (itself) needs to contact its Inter-AS PCE1 to compute the path. R1, as a PCC, sends a PCECP path computation request to PCE1. PCE1 determines that R7 is reachable via AS2 and that PCE2 is the PCE to ask for path computation across AS2. PCE1 sends a PCECP path computation request to PCE2. Inter-AS PCE2, in turn, sends a PCECP path computation request to Intra-AS PCE R4 to compute a path within AS2 (in certain cases, the same router such as R3 can assume both inter-AS and intra-AS path computation functions). R4 may for instance return a PCECP path computation response to PCE2 with ASBR3 as the entry point to AS2 from AS1 and ASBR7 as the exit point to AS3. PCE2 then sends a PCECP path computation request to PCE3 to compute the path segment across AS3, starting at ASBR7 and terminating at R7. PCE3 returns a PCECP path computation response to PCE2 with the path segment ASBR7-R7. PCE2 then returns path ASBR3-ASBR5-ASBR7-R7 to PCE1, which, in turn, returns path ASBR1-ASBR3-ASBR5-ASBR7-R7 to PCC R1.

As described in the above scenario, in general, a PCC may contact an inter-AS PCE to request the computation of an inter-AS path. That PCE may supply the path itself or may solicit the services of other PCEs, which may themselves be inter-AS PCEs, or may be intra-AS PCEs with the responsibility for computing path segments within just one AS.

This document describes the PCE Communication Protocol requirements for inter-AS path computation, i.e., for PCCs to communicate path computation requests for inter-AS (G)MPLS TE paths to PCEs, and for the PCEs to respond. It also includes the requirements for PCEs to communicate inter-AS path computation requests and responses.

3.1. Scope of Deployment Model

All attempts to predict future deployment scopes within the Internet have proven fruitless. Nevertheless, it may be helpful to provide some discussion of the scope of the inter-AS deployment model as envisioned at the time of writing.

It is expected that most, if not all, inter-AS PCECP-based communications will be between PCEs operating in the cooperative PCE model described in [RFC4655]. Clearly, in this model, the requesting PCE acts as a PCC for the purpose of issuing a path computation request, but nevertheless, the requesting node fills the wider role of a PCE in its own AS. It is currently considered unlikely that a PCC (for example, a normal Label Switching Router) will make a path computation request to a PCE outside its own AS. This means that the PCECP relationships between ASes are limited to at most n squared (n^2), where n is the number of peering PCEs in the various ASes

(considered to be no greater than 100 in [RFC4657]). In practice, however, it is likely that only a few PCEs in one AS will be designated for PCECP communications with a PCE in an adjacent AS, and each of these will only have a few PCEs in the adjacent AS to choose from. A deployment model might place the PCEs as co-resident with the ASBRs, resulting in a manageable scaling of the PCE-PCE relationships. Scaling considerations (Section 4.2), manageability considerations (Section 4.3), and security considerations (Section 5) should be examined in the light of these deployment expectations.

4. Detailed PCECP Requirements for Inter-AS G(MPLS) TE Path Computation

This section discusses detailed PCECP requirements for inter-AS (G)MPLS TE LSPs. Depending on the deployment environment, some or all of the requirements described here may be utilized. Specifically, some requirements are more applicable to inter-provider inter-AS (G)MPLS TE operations than to intra-provider operations.

4.1. PCE Communication Protocol Requirements

Requirements specific to inter-AS PCECP path computation requests and responses are discussed in the following sections.

4.1.1. Requirements for Path Computation Requests

The following are inter-AS specific requirements for PCECP requests for path computation:

1. [RFC4657] states the requirement for a priority level to be associated with each path computation request. This document does not change that requirement. However, PCECP should include a mechanism that enables an inter-AS PCE to inform the requesting inter-AS PCE of a change in the request priority level that may have resulted from the application of a local policy.

2. A path computation request by an inter-AS PCE or a PCC to another inter-AS PCE MUST be able to specify the sequence of ASes and/or ASBRs across the network by providing ASBRs and/or ASes as hops in the desired path of the TE LSP to the destination. For instance, an inter-AS PCE MUST be able to specify to the inter-AS PCE serving the neighboring AS a preferred ASBR for exiting to that AS and reach the destination. That is, where multiple ASBRs exist, the requester MUST be able to indicate a preference for one of them. The PCE must be able to indicate whether the specified ASBR or AS is mandatory or non-mandatory on the (G)MPLS TE path.

3. PCECP MUST allow a requester to provide a list of ASes and/or ASBRs to be excluded from the computed path.

4. A PCECP path computation request from one inter-AS PCE to another MUST include the AS number of the requesting AS to enable the correct application of local policy at the second inter-AS PCE.

5. A path computation request from a PCC to an inter-AS PCE or an inter-AS PCE to another MUST be able to specify the need for protection against node, link, or Shared Risk Link Group (SRLG) failure using 1:1 detours or facility backup. It MUST be possible to request protection across all ASes or across specific ASes.

6. PCECP MUST support the disjoint path requirements as specified in [RFC4657]. In addition, it MUST allow the specification of AS-diversity for the computation of a set of two or more paths.

7. A PCECP path computation request message MUST be able to identify the scope of diversified path computation to be end-to-end (i.e., between the endpoints of the (G)MPLS TE tunnel) or to be limited to a specific AS.

4.1.2. Requirements for Path Computation Responses

The following are inter-AS specific requirements for PCECP responses for path computation:

1. A PCECP path computation response from one inter-AS PCE to another MUST be able to include both ASBRs and ASes in the computed path while preserving path segment and topology confidentiality.

2. A PCECP path computation response from one inter-AS PCE to the requesting inter-AS PCE MUST be able to carry an identifier for a path segment it computes to preserve path segment and topology confidentiality. The objective of the identifier is to be included in the TE LSP signaling, whose mechanism is out of scope of this document, to be used for path expansion during LSP signaling.

3. If a constraint for a desired ASBR (see Section 4.1.1, requirement 2) cannot be satisfied by a PCE, PCECP SHOULD allow the PCE to notify the requester of that fact as an error in a path computation response.

4. A PCECP path computation response from an inter-AS PCE to a requesting inter-AS PCE or a PCC MUST be able to carry a cumulative inter-AS path cost. Path cost normalization across ASes is out of scope of this document.

5. A PCECP path computation response from an inter-AS PCE to a PCC SHOULD be able to carry the intra-AS cost of the path segment within the PCC AS.

6. A PCECP path computation response MUST be able to identify diversified paths for the same (G)MPLS TE LSP. End-to-end (i.e., between the two endpoints of the (G)MPLS TE tunnel) disjoint paths are paths that do not share nodes, links, or SRLGs except for the LSP head-end and tail-end. In cases where diversified path segments are desired within one or more ASes, the disjoint path segments may share only the ASBRs of the first AS and the ASBR of the last AS across these ASes.

4.2. Scalability and Performance Considerations

PCECP design for use in the inter-AS case SHOULD consider the following criteria:

- PCE message processing load.
- Scalability as a function of the following parameters:
  - o number of PCCs within the scope of an inter-AS PCE
  - o number of intra-AS PCEs within the scope of an inter-AS PCE
  - o number of peering inter-AS PCEs per inter-AS PCE
- Added complexity caused by inter-AS features.

4.3. Management Considerations

[RFC4657] specifies generic requirements for PCECP management. This document specifies new requirements that apply to inter-AS operations.

The PCECP MIB module MUST provide objects to control the behavior of PCECP in inter-AS applications. These objects include the ASes within the scope of an inter-AS PCE, inter-AS PCEs in neighboring ASes to which the requesting PCE will or will not communicate, confidentiality, and policies.

The built-in diagnostic tools MUST enable failure detection and status checking of PCC/PCE-PCE PCECP. Diagnostic tools include statistics collection on the historical behavior of PCECP as specified in [RFC4657], but additionally it MUST be possible to analyze these statistics on a neighboring AS basis (i.e., across the inter-AS PCEs that belong to a neighboring AS).

The MIB module MUST support trap functions when thresholds are crossed or when important events occur as stated in [RFC4657]. These thresholds SHOULD be specifiable per neighbor AS as well as per peer inter-AS PCE, and traps should be accordingly generated.

Basic liveliness detection for PCC/PCE-PCE PCECP is described in [RFC4657]. The PCECP MIB module SHOULD allow control of liveliness check behavior by providing a liveliness message frequency MIB object, and this frequency object SHOULD be specified per inter-AS PCE peer. In addition, there SHOULD be a MIB object that specifies the dead-interval as a multiplier of the liveliness message frequency so that if no liveliness message is received within that time from an inter-AS PCE, the inter-AS PCE is declared unreachable.

4.4. Confidentiality

Confidentiality mainly applies to inter-provider (inter-AS) PCE communication. It is about protecting the information exchanged between PCEs and about protecting the topology information within an SP's network. Confidentiality rules may also apply among ASes owned by a single SP. Each SP will in most cases designate some PCEs for inter-AS (G)MPLS TE path computation within its own administrative domain and some other PCEs for inter-provider inter-AS (G)MPLS TE path computation. Among the inter-provider-scoped inter-AS PCEs in each SP domain, there may also be a subset of the PCEs specifically enabled for path computation across a specific set of ASes of different peer SPs.

PCECP MUST allow an SP to hide from other SPs the set of hops within its own ASes that are traversed by an inter-AS inter-provider TE LSP (c.f., Section 5.2.1 of [RFC4216]). In a multi-SP administrative domain environment, SPs may want to hide their network topologies for security or commercial reasons. Thus, for each inter-AS TE LSP path segment an inter-AS PCE computes, it may return to the requesting inter-AS PCE an inter-AS TE LSP path segment from its own ASes without detailing the explicit intra-AS hops. As stated earlier, PCECP responses SHOULD be able to carry path-segment identifiers that replace the details of that path segment. The potential use of that identifier for path expansion, for instance, during LSP signaling is out of scope of this document.

4.5. Policy Controls Affecting Inter-AS PCECP

Section 5.2.2 of [RFC4216] discusses the policy control requirements for inter-AS RSVP-TE signaling at the AS boundaries for the enforcement of interconnect agreements, attribute/parameter translation and security hardening.

This section discusses those policy control requirements that are similar to what are discussed in section 5.2.2 of [RFC4216] for PCECP. Please note that SPs may still require policy controls during

signaling of TE LSPs to enforce their bilateral or multilateral agreements at AS boundaries, but signaling is out of scope for this document.

4.5.1. Inter-AS PCE Peering Policy Controls

An inter-AS PCE sends path computation requests to its neighboring inter-AS PCEs, and an inter-AS PCE that receives such a request enforces policies applicable to the sender of the request. These policies may include rewriting some of the parameters or rejecting requests based on parameter values. Such policies may be applied for PCEs belonging to different SPs or to PCEs responsible for ASes within a single SP administrative domain. Parameters that might be subject to policy include bandwidth, setup/holding priority, Fast Reroute request, Differentiated Services Traffic Engineering (DS-TE) Class Type (CT), and others as specified in section 5.2.2.1 of [RFC4216].

For path computation requests that are not compliant with locally configured policies, PCECP SHOULD enable a PCE to send an error message to the requesting PCC or PCE indicating that the request has been rejected because a specific parameter did not satisfy the local policy.

4.5.2. Inter-AS PCE Re-Interpretation Policies

Each SP may have different definitions in its use of, for example, DS-TE TE classes. An inter-AS PCE receiving a path computation request needs to interpret the parameters and constraints and adapt them to the local environment. Specifically, a request constructed by a PCC or PCE in one AS may have parameters and constraints that should be interpreted differently or translated by the receiving PCE that is in a different AS. A list of signaling parameters subject to policy re-interpretation at AS borders can be found in section 5.2.2.2 of [RFC4216], and the list for path computation request parameters and constraints is the same. In addition, the transit SPs along the inter-AS TE path may be GMPLS transport providers, which may require re-interpretation of MPLS-specific PCECP path computation request objects in order to enable path computation over a GMPLS network or vice versa.

5. Security Considerations

The PCECP is a communications protocol for use between potentially remote entities (PCCs and PCEs) over an IP network. Security concerns arise in order to protect the PCC, PCE, and the information they exchange. [RFC4758] specifies requirements on the PCECP to protect against spoofing, snooping, and DoS attacks. That document

is concerned with general protocol requirements applicable to the basic use of the PCECP. This document is specific to the application of the PCE architecture in an inter-AS environment, and so it is appropriate to highlight the security considerations that apply in that environment.

Security requirements that exist within a single administrative domain become critical in the multi-AS case when the control of IP traffic and access to the network may leave the authority of a single administration.

5.1. Use and Distribution of Keys

How the participants in a PCECP session discover each other and the need for the session is out of scope of this document. It may be through configuration or automatic discovery. However, when a PCECP session is established, the PCECP speakers MUST have mechanisms to authenticate each other's identities and validate the data they exchange. They also SHOULD have mechanisms to protect the data that they exchange via encryption. Such mechanisms usually require the use of keys, and so the PCECP MUST describe techniques for the exchange and use of security keys. Where inter-AS PCE discovery is used, and PCECP security is required, automated key distribution mechanisms MUST also be used. Since such key exchange must (necessarily) operate over an AS boundary, proper consideration needs to be given to how inter-AS key exchanges may be carried out and how the key exchange, itself, may be secured. Key distribution mechanisms MUST be defined with consideration of [RFC4107]. Where a PCECP session is configured between a pair of inter-AS PCEs, a security key may be manually set for that session.

5.2. Application of Policy

Policy forms an important part of the operation of PCEs in an inter-AS environment as described in Section 4.5, especially when ASes are administrated by different SPs. A wider discussion of the application of policy to the PCE architecture can be found in [PCE-POLICY].

Policy may also form part of the security model for the PCECP and may be particularly applicable to inter-AS path computation requests. A fundamental element of the application of policy at a PCE is the identity of the requesting PCC/PCE. This makes the use of authentication described in Section 5.1 particularly important. Where policy information is exchanged as part of the computation request and/or response, the policy object is transparent to the PCECP being delivered un-inspected and unmodified to the policy component of a PCE or PCC. Therefore, the policy components are

responsible for protecting (for example, encrypting) the policy information and using additional identification and authentication if a higher level of validation is required than is provided by the base protocol elements of the PCECP.

5.3. Confidentiality

The PCECP MUST provide a mechanism to preserve the confidentiality of path segments computed by a PCE in one AS and provided in a computation response to another AS.

Furthermore, a PCE SHOULD be provided with a mechanism to mask its identity such that its presence in the path computation chain in a cooperative PCE model (such as described in [BRPC]) cannot be derived from the computed path. This will help to protect the PCE from targeted attacks. Clearly, such confidentiality does not extend to the PCECP peer (i.e., a PCC or another PCE) that invokes the PCE with a path computation request.

5.4. Falsification of Information

In the PCE architecture, when PCEs cooperate, one PCE may return a path computation result that is composed of multiple path segments, each computed by a different PCE. In the inter-AS case, each PCE may belong to a different administrative domain, and the source PCC might not know about the downstream PCEs, nor fully trust them. Although it is possible and RECOMMENDED to establish a chain of trust between PCEs, this might not always be possible. In this case, it becomes necessary to guard against a PCE changing the information provided by another downstream PCE. Some mechanism MUST be available in the PCECP, and echoed in the corresponding signaling, that allows an AS to verify that the signaled path conforms to the path segment computed by the local PCE and returned on the path computation request.

6. Acknowledgments

We would like to thank Adrian Farrel, Jean-Philippe Vasseur, and Jean Louis Le Roux for their useful comments and suggestions. Pasi Eronen and Sandy Murphy provided valuable early Security Directorate reviews. Adrian Farrel re-wrote the Security Considerations section.

7. Normative References

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

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

[BRPC] Vasseur, JP., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward Recursive PCE-based Computation (BRPC) Procedure To Compute Shortest Constrained Inter-domain Traffic Engineering Label Switched paths", Work in Progress, April 2008.

[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.

[RFC4758] Nystroem, M., "Cryptographic Token Key Initialization Protocol (CT-KIP) Version 1.0 Revision 1", RFC 4758, November 2006.

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

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

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

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

Authors' Addresses

Nabil Bitar
Verizon
117 West Street
Waltham, MA 02451 USA
EMail: nabil.n.bitar@verizon.com

Kenji Kumaki
KDDI R&D Laboratories, Inc.
2-1-15 Ohara Fujimino
Saitama 356-8502, JAPAN
EMail: ke-kumaki@kddi.com

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