PCE Working Group D. Dhody
Internet-Draft Y. Lee
Intended status: Informational Huawei Technologies
Expires: May 02, 2017 D. Ceccarelli
Ericsson
J. Shin
SK Telecom
D. King
Lancaster University
O. Gonzalez de Dios
Telefonica I+D
October 29, 2016

Hierarchical Stateful Path Computation Element (PCE).
draft-dhodylee-pce-stateful-hpce-02

Abstract

A Stateful Path Computation Element (PCE) maintains information on the current network state, including: computed Label Switched Path (LSPs), reserved resources within the network, and pending path computation requests. This information may then be considered when computing new traffic engineered LSPs, and for associated and dependent LSPs, received from Path Computation Clients (PCCs).

The Hierarchical Path Computation Element (H-PCE) architecture, provides an architecture to allow the optimum sequence of inter-connected domains to be selected, and network policy to be applied if applicable, via the use of a hierarchical relationship between PCEs.

Combining the capabilities of Stateful PCE and the Hierarchical PCE would be advantageous. This document describes general considerations and use cases for the deployment of Stateful PCE(s) using the Hierarchical PCE architecture.

Status of This Memo

Table of Contents

1. Introduction

The Path Computation Element communication Protocol (PCEP) provides mechanisms for Path Computation Elements (PCEs) to perform path computations in response to Path Computation Clients' (PCCs) requests.

A stateful PCE is capable of considering, for the purposes of path computation, not only the network state in terms of links and nodes (referred to as the Traffic Engineering Database or TED) but also the status of active services (previously computed paths, and currently reserved resources, stored in the Label Switched Paths Database (LSPDB).

[I-D.ietf-pce-stateful-pce-app] describes general considerations for a stateful PCE deployment and examines its applicability and benefits, as well as its challenges and limitations through a number of use cases.

[I-D.ietf-pce-stateful-pce] describes a set of extensions to PCEP to provide stateful control. A stateful PCE has access to not only the information carried by the network's Interior Gateway Protocol (IGP), but also the set of active paths and their reserved resources for its computations. The additional state allows the PCE to compute constrained paths while considering individual LSPs and their interactions. [I-D.ietf-pce-pce-initiated-lsp] describes the setup, maintenance and teardown of PCE-initiated LSPs under the stateful PCE model.

[I-D.ietf-pce-stateful-pce] also describes the active stateful PCE. The active PCE functionality allows a PCE to reroute an existing LSP or make changes to the attributes of an existing LSP, or delegate control of specific LSPs to a new PCE.

The ability to compute shortest constrained TE LSPs in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks across multiple domains has been identified as a key motivation for PCE development. [RFC6805] describes a Hierarchical PCE (H-PCE) architecture which can be used for computing end-to-end paths for inter-domain MPLS Traffic Engineering (TE) and GMPLS Label Switched Paths (LSPs). Within the Hierarchical PCE (H-PCE) architecture [RFC6805], the Parent PCE (P-PCE) is used to compute a multi-domain path based on the domain connectivity information. A Child PCE (C-PCE) may be responsible for a single domain or multiple domains, it is used to compute the intra-domain path based on its domain topology information.

This document presents general considerations for stateful PCE(s) in hierarchical PCE architecture. In particular, the behavior changes and additions to the existing stateful PCE mechanisms (including PCE-initiated LSP setup and active PCE usage) in the context of networks using the H-PCE architecture.

1.1. Requirements Language

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].

2. Terminology

The terminology is as per [RFC4655], [RFC5440], [RFC6805], and [I-D.ietf-pce-stateful-pce].

3. Hierarchical Stateful PCE

As described in [RFC6805], in the hierarchical PCE architecture, a P-PCE maintains a domain topology map that contains the child domains (seen as vertices in the topology) and their interconnections (links in the topology). The P-PCE has no information about the content of the child domains. Each child domain has at least one PCE capable of computing paths across the domain. These PCEs are known as C-PCEs and have a direct relationship with the P-PCE. The P-PCE builds the domain topology map either via direct configuration (allowing network policy to also be applied) or from learned information received from each C-PCE.

[I-D.ietf-pce-stateful-pce] specifies new functions to support a stateful PCE. It also specifies that a function can be initiated either from a PCC towards a PCE (C-E) or from a PCE towards a PCC (E-C).

This document extends these functions to support H-PCE Architecture from a C-PCE towards a P-PCE (CE-PE) or from a P-PCE towards a C-PCE (PE-CE). All PCE types herein (i.e., PE or CE) are assumed to be 'stateful PCE'.

A number of interactions are expected in the Hierarchical Stateful PCE architecture, these include:

LSP State Report (CE-PE): a child stateful PCE sends an LSP state report to a Parent Stateful PCE whenever the state of a LSP changes.

LSP State Synchronization (CE-PE): after the session between the Child and Parent stateful PCEs is initialized, the P-PCE must learn the state of C-PCE's TE LSPs.

LSP Control Delegation (CE-PE,PE-CE): a C-PCE grants to the P-PCE the right to update LSP attributes on one or more LSPs; the C-PCE may withdraw the delegation or the P-PCE may give up the delegation at any time.

LSP Update Request (PE-CE): a stateful P-PCE requests modification of attributes on a C-PCE's TE LSP.

PCE LSP Initiation Request (PE-CE): a stateful P-PCE requests C-PCE to initiate a TE LSP.

Note that this hierarchy is recursive and thus a LSR could delegate the control to a PCE, which may delegate to its parent, which may further delegate it to its parent (if it exist or needed). Similarly update operations could also be applied recursively.

3.1. Passive Operations

Procedures as described in [RFC6805] are applied, where the ingress C-PCE sends a request to the P-PCE. The P-PCE selects a set of candidate domain paths based on the domain topology and the state of the inter-domain links. It then sends computation requests to the C-PCEs responsible for each of the domains on the candidate domain paths. Each C-PCE computes a set of candidate path segments across its domain and sends the results to the P-PCE. The P-PCE uses this information to select path segments and concatenate them to derive the optimal end-to-end inter-domain path. The end-to-end path is then sent to the C-PCE that received the initial path request, and this C-PCE passes the path on to the PCC that issued the original request.

As per [I-D.ietf-pce-stateful-pce], PCC sends an LSP State Report carried on a PCRpt message to the C-PCE, indicating the LSP's status. The C-PCE MAY further propagate the State Report to the P-PCE. A local policy at C-PCE MAY dictate which LSPs to be reported to the P-PCE. The PCRpt message is sent from C-PCE to P-PCE.

State synchronization mechanism as described in [I-D.ietf-pce-stateful-pce] and [I-D.ietf-pce-stateful-sync-optimizations] are applicable to PCEP session between C-PCE and P-PCE as well.

Taking the sample hierarchical domain topology example from [RFC6805] as the reference topology for the entirety of this document.

```
      -----------------------------------------------------------------
     |   Domain 5                                                      |
     |                              -----                              |
     |                             |PCE 5|                             |
     |                              -----                              |
     |                                                                 |
     |    ----------------     ----------------     ----------------   |
     |   | Domain 1       |   | Domain 2       |   | Domain 3       |  |
     |   |                |   |                |   |                |  |
     |   |        -----   |   |        -----   |   |        -----   |  |
     |   |       |PCE 1|  |   |       |PCE 2|  |   |       |PCE 3|  |  |
     |   |        -----   |   |        -----   |   |        -----   |  |
     |   |                |   |                |   |                |  |
     |   |            ----|   |----        ----|   |----            |  |
     |   |           |BN11+---+BN21|      |BN23+---+BN31|           |  |
     |   |   -        ----|   |----        ----|   |----        -   |  |
     |   |  |S|           |   |                |   |           |D|  |  |
     |   |   -        ----|   |----        ----|   |----        -   |  |
     |   |           |BN12+---+BN22|      |BN24+---+BN32|           |  |
     |   |            ----|   |----        ----|   |----            |  |
     |   |                |   |                |   |                |  |
     |   |         ----   |   |                |   |   ----         |  |
     |   |        |BN13|  |   |                |   |  |BN33|        |  |
     |    -----------+----     ----------------     ----+-----------   |
     |                \                                /               |
     |                 \       ----------------       /                |
     |                  \     |                |     /                 |
     |                   \    |----        ----|    /                  |
     |                    ----+BN41|      |BN42+----                   |
     |                        |----        ----|                       |
     |                        |                |                       |
     |                        |        -----   |                       |
     |                        |       |PCE 4|  |                       |
     |                        |        -----   |                       |
     |                        |                |                       |
     |                        | Domain 4       |                       |
     |                         ----------------                        |
     |                                                                 |
      -----------------------------------------------------------------
```

Figure 1: Sample Hierarchical Domain Topology

Steps 1 to 11 are exactly as described in section 4.6.2 (Hierarchical PCE End-to-End Path Computation Procedure) of [RFC6805], the following additional steps are added for stateful PCE:

(1) The Ingress LSR initiates the setup of the LSP as per the path and reports to the PCE1 the LSP status ("GOING-UP").

(2) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

(3) The Ingress LSR notifies the LSP state to PCE1 when the state is "UP".

(4) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

3.2. Active Operations

[I-D.ietf-pce-stateful-pce] describes the case of active stateful PCE. The active PCE functionality uses two specific PCEP messages:

o Update Request (PCUpd)
o State Report (PCRpt)

The first is sent by the PCE to a Path Computation Client (PCC) for modifying LSP attributes. The PCC sends back a PCRpt to acknowledge the requested operation. PCRpt has the same structure of PCNtf message.

As per [I-D.ietf-pce-stateful-pce], Delegation is an operation to grant a PCE, temporary rights to modify a subset of LSP parameters on one or more PCC's LSPs. The C-PCE may further choose to delegate to P-PCE based on a local policy. The PCRpt message with "D" (delegate) flag is sent from C-PCE to P-PCE.

To update an LSP, a PCE send to the PCC, an LSP Update Request using a PCUpd message. For LSP delegated to the P-PCE via the child PCE, the P-PCE can use the same PCUpd message to request change to the C-PCE (the Ingress domain PCE), the PCE further propagates the update request to the PCC.

The P-PCE uses the same mechanism described in Section 3.1 to compute the end to end path using PCReq and PCRep messages.

The following additional steps are also initially performed, for active operations, again using the reference architecture described in Figure 1 (Sample Hierarchical Domain Topology).

(1) The Ingress LSR delegates the LSP to the PCE1 via PCRpt message with D flag set.

(2) The PCE1 further delegates the LSP to the P-PCE (PCE5).

Steps 4 to 10 of section 4.6.2 of [RFC6805] are executed to determine the end to end path.

(3) The P-PCE (PCE5) sends the update request to the C-PCE (PCE1) via PCUpd message.

(4) The PCE1 further updates the LSP to the Ingress LSR (PCC).

(5) The Ingress LSR initiates the setup of the LSP as per the path and reports to the PCE1 the LSP status ("GOING-UP").

(6) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

(7) The Ingress LSR notifies the LSP state to PCE1 when the state is "UP".

(8) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

3.3. PCE Initiation Operation

[I-D.ietf-pce-pce-initiated-lsp] describes the setup, maintenance and teardown of PCE-initiated LSPs under the stateful PCE model, without the need for local configuration on the PCC, thus allowing for a dynamic network that is centrally controlled and deployed. To instantiate or delete an LSP, the PCE sends the Path Computation LSP Initiate Request (PCInitiate) message to the PCC. In case of inter-domain LSP in Hierarchical PCE architecture, the initiation operations can be carried out at the P-PCE. In which case after P-PCE finishes the E2E path computation, it can send the PCInitiate message to the C-PCE (the Ingress domain PCE), the PCE further propagates the initiate request to the PCC.

The following additional steps are also initially performed, for PCE initiated operations, again using the reference architecture described in Figure 1 (Sample Hierarchical Domain Topology):

(1) The P-PCE (PCE5) is requested to initiate a LSP.

Steps 4 to 10 of section 4.6.2 of [RFC6805] are executed to determine the end to end path.

(2) The P-PCE (PCE5) sends the initiate request to the child PCE (PCE1) via PCInitiate message.

(3) The PCE1 further propagates the initiate message to the Ingress LSR (PCC).

(4) The Ingress LSR initiates the setup of the LSP as per the path and reports to the PCE1 the LSP status ("GOING-UP").

(5) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

(6) The Ingress LSR notifies the LSP state to PCE1 when the state is "UP".

(7) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

3.3.1. Per Domain Stitched LSP

The hierarchical PCE architecture as per [RFC6805] is primarily used for E2E LSP. With PCE-Initiated capability, another mode of operation is possible, where multiple intra-domain LSPs are initiated in each domain which are further stitched to form an E2E LSP. The P-PCE sends PCInitiate message to each C-PCE separately to initiate individual LSP segments along the domain path. These individual per domain LSP are stitched together by some mechanism, which is out of scope of this document. The P-PCE may also send the PCInitiate message to the ingress C-PCE to initiate the E2E LSP separately.

The following additional steps are also initially performed, for the Per Domain stiched LSP operation, again using the reference architecture described in Figure 1 (Sample Hierarchical Domain Topology):

(1) The P-PCE (PCE5) is requested to initiate a LSP.

Steps 4 to 10 of section 4.6.2 of [RFC6805] are executed to determine the end to end path, which are broken into per-domain LSPs say -

- S-BN41

- BN41-BN33

- BN33-D

It should be noted that the P-PCE MAY use other mechanisms to determine the suitable per-domain LSPs (apart from [RFC6805]).

For LSP (BN33-D)

(2) The P-PCE (PCE5) sends the initiate request to the child PCE (PCE3) via PCInitiate message for LSP (BN33-D).

(3) The PCE3 further propagates the initiate message to BN33.

(4) BN33 initiates the setup of the LSP as per the path and reports to the PCE3 the LSP status ("GOING-UP").

(5) The PCE3 further reports the status of the LSP to the P-PCE (PCE5).

(6) The node BN33 notifies the LSP state to PCE3 when the state is "UP".

(7) The PCE3 further reports the status of the LSP to the P-PCE (PCE5).

For LSP (BN41-BN33)

(8) The P-PCE (PCE5) sends the initiate request to the child PCE (PCE4) via PCInitiate message for LSP (BN41-BN33).

(9) The PCE4 further propagates the initiate message to BN41.

(10) BN41 initiates the setup of the LSP as per the path and reports to the PCE4 the LSP status ("GOING-UP").

(11) The PCE4 further reports the status of the LSP to the P-PCE (PCE5).

(l2) The node BN41 notifies the LSP state to PCE4 when the state is "UP".

(13) The PCE4 further reports the status of the LSP to the P-PCE (PCE5).

For LSP (S-BN41)

(14) The P-PCE (PCE5) sends the initiate request to the child PCE (PCE1) via PCInitiate message for LSP (S-BN41).

(15) The PCE1 further propagates the initiate message to node S.

(16) S initiates the setup of the LSP as per the path and reports to the PCE1 the LSP status ("GOING-UP").

(17) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

(18) The node S notifies the LSP state to PCE1 when the state is "UP".

(19) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

Additionally:

(20) Once P-PCE receives report of each per-domain LSP, it should use some stitching mechanism, which is out of scope of this document. In this step, P-PCE (PCE5) could also initiate an E2E LSP (S-D) by sending the PCInitiate message to Ingress C-PCE (PCE1).

4. Other Considerations

4.1. Applicability to Inter-Layer

[RFC5623] describes a framework for applying the PCE-based architecture to inter-layer (G)MPLS traffic engineering. The H-PCE Stateful architecture with stateful P-PCE coordinating with the stateful C-PCEs of higher and lower layer is shown in the figure below.

```
                                                     +----------+
                                                    /| Parent   |
                                                  /  | PCE      |
                                                /    +----------+
                                              /        / Stateful
                                             /       /
                                           /        /
                                          /       /
                          Stateful +---+/        /
                          Child   + PCE +      /
                          PCE Hi  + Hi  +     /
                                   +---+    /
          +---+    +---+                   /      +---+    +---+
         + LSR +--+ LSR +........................+ LSR +--+ LSR +
         + H1  +  + H2  +                 /      + H3  +  + H4  +
          +---+    +---+\          +---+/        /+---+    +---+
                         \        + PCE +       /
                          \       + Lo  +      /
                Stateful   \       +---+      /
                C-PCE   \                /
                Lo           \+---+    +---+/
                             + LSR +--+ LSR +
                             + L1  +  + L2  +
                              +---+    +---+
```

Figure 2: Sample Inter-Layer Topology

All procedures described in Section 3 are applicable to inter-layer path setup as well.

4.2. Applicability to ACTN

[I-D.ietf-teas-actn-framework] describes framework for Abstraction and Control of TE Networks (ACTN), where each Physical Network Controller (PNC) is equivalent to C-PCE and P-PCE is the Multi-Domain Service Coordinator (MDSC). The Per domain stitched LSP as per the Hierarchical PCE architecture described in Section 3.3.1 and Section 4.1 is well suited for ACTN.

[I-D.dhody-pce-applicability-actn] examines the applicability of PCE to the ACTN framework. To support the function of multi domain coordination via hierarchy, the stateful hierarchy of PCEs plays a crucial role.

In ACTN framework, Customer Network Controller (CNC) can request the MDSC to check if there is a possibility to meet Virtual Network (VN) requirements (before requesting for VN provision). The H-PCE architecture as described in [RFC6805] can supports via the use of PCReq and PCRep messages between the P-PCE and C-PCEs.

5. Scalability Considerations

It should be noted that if all the C-PCEs would report all the LSPs in their domain, it could lead to scalability issues for the P-PCE. Thus it is recommended to only report the LSPs which are involved in H-PCE, i.e. the LSPs which are either delegated to the P-PCE or initiated by the P-PCE.

6. Security Considerations

TBD.

7. Manageability Considerations

7.1. Control of Function and Policy

TBD.

7.2. Information and Data Models

TBD.

7.3. Liveness Detection and Monitoring

TBD.

7.4. Verify Correct Operations

TBD.

7.5. Requirements On Other Protocols

TBD.

7.6. Impact On Network Operations

TBD.

8. IANA Considerations

There are no IANA considerations.

9. Acknowledgments

Thanks to Manuela Scarella, Haomian Zheng, Sergio Marmo, Stefano Parodi, Giacomo Agostini, Jeff Tantsura and Rajan Rao for suggestions.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <http://www.rfc-editor.org/info/rfc5440>.

[RFC6805] King, D., Ed. and A. Farrel, Ed., "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS", RFC 6805, DOI 10.17487/RFC6805, November 2012, <http://www.rfc-editor.org/info/rfc6805>.

[I-D.ietf-pce-stateful-pce] Crabbe, E., Minei, I., Medved, J., and R. Varga, "PCEP Extensions for Stateful PCE", draft-ietf-pce-stateful-pce-16 (work in progress), September 2016.

[I-D.ietf-pce-pce-initiated-lsp]
Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", draft-ietf-pce-pce-initiated-lsp-07 (work in progress), July 2016.

10.2. Informative References

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <http://www.rfc-editor.org/info/rfc4655>.

[RFC5623] Oki, E., Takeda, T., Le Roux, JL., and A. Farrel, "Framework for PCE-Based Inter-Layer MPLS and GMPLS Traffic Engineering", RFC 5623, DOI 10.17487/RFC5623, September 2009, <http://www.rfc-editor.org/info/rfc5623>.

[I-D.ietf-pce-stateful-pce-app]
Zhang, X. and I. Minei, "Applicability of a Stateful Path Computation Element (PCE)", draft-ietf-pce-stateful-pce-app-07 (work in progress), September 2016.

[I-D.ietf-pce-stateful-sync-optimizations]
Crabbe, E., Minei, I., Medved, J., Varga, R., Zhang, X., and D. Dhody, "Optimizations of Label Switched Path State Synchronization Procedures for a Stateful PCE", draft-ietf-pce-stateful-sync-optimizations-06 (work in progress), October 2016.

[I-D.ietf-teas-actn-framework]
Ceccarelli D. and Y. Lee, "Framework for Abstraction and Control of Transport Networks", draft-ietf-teas-actn-framework-01 (work in progress), October 2016.

[I-D.dhody-pce-applicability-actn]
Dhody, D., Lee, Y., and D. Ceccarelli, "Applicability of Path Computation Element (PCE) for Abstraction and Control of TE Networks (ACTN)", draft-dhody-pce-applicability-actn-01 (work in progress), October 2016.

Appendix A. Contributor Addresses

Avantika
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

EMail: avantika.sushilkumar@huawei.com

Xian Zhang
Huawei Technologies
Bantian, Longgang District
Shenzhen, Guangdong 518129
P.R.China

EMail: zhang.xian@huawei.com

Udayasree Palle
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

EMail: udayasree.palle@huawei.com

Authors' Addresses

Dhruv Dhody
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

EMail: dhruv.ietf@gmail.com

Young Lee
Huawei Technologies
5340 Legacy Drive, Building 3
Plano, TX 75023
USA

EMail: leeyoung@huawei.com

Daniele Ceccarelli
Ericsson

Torshamnsgatan,48
Stockholm
Sweden

EMail: daniele.ceccarelli@ericsson.com

Jongyoon Shin
SK Telecom
6 Hwangsaeul-ro, 258 beon-gil, Bundang-gu, Seongnam-si,
Gyeonggi-do 463-784
Republic of Korea

EMail: jongyoon.shin@sk.com

Dan King
Lancaster University
UK

EMail: d.king@lancaster.ac.uk

Oscar Gonzalez de Dios
Telefonica I+D
Don Ramon de la Cruz 82-84
Madrid, 28045
Spain

Phone: +34913128832
Email: ogondio@tid.es