

PCE Working Group  
Internet-Draft  
Intended status: Informational  
Expires: March 06, 2020

D. Dhody  
Huawei Technologies  
Y. Lee  
Futurewei Technologies  
D. Ceccarelli  
Ericsson  
J. Shin  
SK Telecom  
D. King  
Lancaster University  
September 03, 2019

**Hierarchical Stateful Path Computation Element (PCE).  
draft-ietf-pce-stateful-hpce-12**

**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 Path computation response from a PCE is helpful for the PCC to gracefully establish the computed LSP.

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, and not Stateless, PCEs using the Hierarchical PCE architecture.

**Status of This Memo**

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>. 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."

This Internet-Draft will expire on December 18, 2019.

## Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](https://trustee.ietf.org/license-info) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

<a href="#">1.</a>	<a href="#">Introduction</a>	<a href="#">3</a>
<a href="#">1.1.</a>	<a href="#">Background</a>	<a href="#">3</a>
<a href="#">1.2.</a>	<a href="#">Use-cases and Applicability of Hierarchical Stateful PCE</a>	<a href="#">4</a>
<a href="#">1.2.1.</a>	<a href="#">Applicability to ACTN</a>	<a href="#">5</a>
<a href="#">1.2.2.</a>	<a href="#">End-to-End Contiguous LSP</a>	<a href="#">5</a>
<a href="#">1.2.3.</a>	<a href="#">Applicability of a Stateful P-PCE</a>	<a href="#">6</a>
<a href="#">2.</a>	<a href="#">Terminology</a>	<a href="#">6</a>
<a href="#">2.1.</a>	<a href="#">Requirement Language</a>	<a href="#">7</a>
<a href="#">3.</a>	<a href="#">Hierarchical Stateful PCE</a>	<a href="#">7</a>
<a href="#">3.1.</a>	<a href="#">Passive Operations</a>	<a href="#">9</a>
<a href="#">3.2.</a>	<a href="#">Active Operations</a>	<a href="#">11</a>
<a href="#">3.3.</a>	<a href="#">PCE Initiation of LSPs</a>	<a href="#">12</a>
<a href="#">3.3.1.</a>	<a href="#">Per Domain Stitched LSP</a>	<a href="#">13</a>
<a href="#">4.</a>	<a href="#">Security Considerations</a>	<a href="#">15</a>
<a href="#">5.</a>	<a href="#">Manageability Considerations</a>	<a href="#">16</a>
<a href="#">5.1.</a>	<a href="#">Control of Function and Policy</a>	<a href="#">16</a>
<a href="#">5.2.</a>	<a href="#">Information and Data Models</a>	<a href="#">16</a>
<a href="#">5.3.</a>	<a href="#">Liveness Detection and Monitoring</a>	<a href="#">16</a>
<a href="#">5.4.</a>	<a href="#">Verify Correct Operations</a>	<a href="#">16</a>
<a href="#">5.5.</a>	<a href="#">Requirements On Other Protocols</a>	<a href="#">16</a>
<a href="#">5.6.</a>	<a href="#">Impact On Network Operations</a>	<a href="#">17</a>
<a href="#">5.7.</a>	<a href="#">Error Handling between PCEs</a>	<a href="#">17</a>



<a href="#">6.</a>	<a href="#">Other Considerations . . . . .</a>	<a href="#">17</a>
<a href="#">6.1.</a>	<a href="#">Applicability to Inter-Layer Traffic Engineering . . . . .</a>	<a href="#">17</a>
<a href="#">6.2.</a>	<a href="#">Scalability Considerations . . . . .</a>	<a href="#">18</a>
<a href="#">6.3.</a>	<a href="#">Confidentiality . . . . .</a>	<a href="#">18</a>
<a href="#">7.</a>	<a href="#">IANA Considerations . . . . .</a>	<a href="#">18</a>
<a href="#">8.</a>	<a href="#">Acknowledgments . . . . .</a>	<a href="#">18</a>
<a href="#">9.</a>	<a href="#">References . . . . .</a>	<a href="#">18</a>
<a href="#">9.1.</a>	<a href="#">Normative References . . . . .</a>	<a href="#">18</a>
<a href="#">9.2.</a>	<a href="#">Informative References . . . . .</a>	<a href="#">19</a>
	<a href="#">Contributors . . . . .</a>	<a href="#">21</a>
	<a href="#">Authors' Addresses . . . . .</a>	<a href="#">22</a>

## [1.](#) Introduction

### [1.1.](#) Background

The Path Computation Element communication Protocol (PCEP) [[RFC5440](#)] 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 (LSP-DB).

[RFC8051] 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.

[RFC8231] 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. [[RFC8281](#)] describes the setup, maintenance and teardown of PCE-initiated LSPs under the stateful PCE model.

[RFC8231] 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 PCEs, and not Stateless PCEs, in the 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.

In this document, Sections 3.1 and 3.2 focus on end to end (E2E) inter-domain TE LSP. Section 3.3.1 describes the operations for stitching Per Domain LSPs.

## **1.2. Use-cases and Applicability of Hierarchical Stateful PCE**

As per [RFC6805], in the hierarchical PCE architecture, a P-PCE maintains a domain topology map that contains the child domains and their interconnections. Usually, the P-PCE has no information about the content of the child domains. But if the PCE is applied to the Abstraction and Control of TE Networks (ACTN) [RFC8453] as described in [RFC8637], the Provisioning Network Controller (PNC) can provide an abstract topology to the Multi-Domain Service Coordinator (MDSC). Thus the P-PCE in MDSC could be aware of topology information in much more detail than just the domain topology.

In a PCEP session between a PCC (Ingress) and a C-PCE, the C-PCE acts as per the stateful PCE operations described in [RFC8231] and [RFC8281]. The same C-PCE behaves as a PCC on the PCEP session towards the P-PCE. The P-PCE is stateful in nature and thus maintains the state of the inter-domain LSPs that are reported to it. The inter-domain LSP could also be delegated by the C-PCE to the P-PCE, so that the P-PCE could update the inter-domain path. The trigger for this update could be the LSP state change reported for this LSP or any other LSP. It could also be a change in topology at the P-PCE such as inter-domain link status change. In case of use of stateful H-PCE in ACTN, a change in abstract topology learned by the P-PCE could also trigger the update. Some other external factors (such as a measurement probe) could also be a trigger at the P-PCE. Any such



update would require an inter-domain path recomputation as described in [[RFC6805](#)].

The inter-domain LSP could be set up using the end-to-end signaling as described in [[RFC6805](#)]. Additionally a per-domain stitched LSP model is also applicable in a P-PCE initiation model. [Section 3.1](#), [Section 3.2](#), and [Section 3.3](#) describe the end-to-end Contiguous LSP setup, whereas [Section 3.3.1](#) describe the per-domain stitching.

#### **[1.2.1](#). Applicability to ACTN**

[RFC8453] describes a framework for the Abstraction and Control of TE Networks (ACTN), where each Provisioning Network Controller (PNC) is equivalent to a C-PCE, and the 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 deployments.

[RFC8637] examines the applicability of PCE to the ACTN framework. To support the function of multi domain coordination via hierarchy, the hierarchy of stateful PCEs play a crucial role.

In the ACTN framework, a Customer Network Controller (CNC) can request the MDSC to check whether there is a possibility to meet Virtual Network (VN) requirements before requesting for the VN to be provisioned. The H-PCE architecture as described in [[RFC6805](#)] can support this function using PCReq and PCRep messages between the P-PCE and C-PCEs. When the CNC requests for VN provisioning, the MDSC decompose this request into multiple inter-domain LSP provisioning requests, which might be further decomposed to per-domain path segments. This is described in [Section 3.3.1](#). The MDSC uses the LSP Initiate Request (PCInitiate) message from the P-PCE towards the C-PCE, and the C-PCE reports the state back to the P-PCE via a Path Computation State Report (PCRpt) message. The P-PCE could make changes to the LSP via the use of a Path Computation Update Request (PCUpd) message.

In this case, the P-PCE (as MDSC) interacts with multiple C-PCEs (as PNCs) along the inter-domain path of the LSP.

#### **[1.2.2](#). End-to-End Contiguous LSP**

Different signaling methods for inter-domain RSVP-TE signaling are identified in [[RFC4726](#)]. Contiguous LSPs are achieved using the procedures of [[RFC3209](#)] and [[RFC3473](#)] to create a single end-to-end LSP that spans all domains. [[RFC6805](#)] describes the technique to establish the optimum path when the sequence of domains is not known





in advance. It shows how the PCE architecture can be extended to allow the optimum sequence of domains to be selected, and the optimum end-to-end path to be derived.

In case of a stateful P-PCE, the stateful P-PCE has to be aware of the inter-domain LSPs for it to consider them during path computation. For example, a domain diverse path from another LSP. This is the Passive Stateful P-PCE as described in [Section 3.1](#). Additionally, the inter-domain LSP could be delegated to the P-PCE, so that P-PCE could trigger an update via a PCUpd message. The update could be triggered on receipt of the PCRpt message that indicates a status change of this LSP or some other LSP. The other LSP could be an associated LSP (such as protection) or an unrelated LSP whose resource change leads to re-optimization at the P-PCE. This is the Active Stateful Operation as described in [Section 3.2](#). Further, the P-PCE could be instructed to create an inter-domain LSP on its own using the PCInitiate message for an E2E contiguous LSP. The P-PCE would send the PCInitiate message to the Ingress domain C-PCE, which would further instruct the Ingress PCC.

In this document, for the Contiguous LSP, the above interactions are only between the ingress domain C-PCE and the P-PCE. The use of stateful operations for an inter-domain LSP between the transit/egress domain C-PCEs towards the P-PCE is out of scope of this document.

### **[1.2.3](#). Applicability of a Stateful P-PCE**

[RFC8051] 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. These are also applicable to the stateful P-PCE when used for the inter-domain LSP path computation and setup. It should be noted that though the stateful P-PCE has limited direct visibility inside the child domain, it could still trigger re-optimization with the help of child PCEs based on LSP state changes, abstract topology changes, or some other external factors.

The C-PCE would delegate control of the inter-domain LSP to the P-PCE so that the P-PCE can make changes to it. Note that, if the C-PCE becomes aware of a topology change that is hidden from the P-PCE, it could take back the delegation from the P-PCE to act on it itself. Similarly, a P-PCE could also request for delegation if it needs to make a change to the LSP (refer to [\[I-D.ietf-pce-lsp-control-request\]](#)).

## **[2](#). Terminology**



The terminology is as per [[RFC4655](#)], [[RFC5440](#)], [[RFC6805](#)], [[RFC8051](#)], [[RFC8231](#)], and [[RFC8281](#)].

Some key terms are listed below for easy reference.

ACTN: Abstraction and Control of Traffic Engineering Networks

CNC: Customer Network Controller

C-PCE: Child Path Computation Element

H-PCE: Hierarchical Path Computation Element

IGP: Interior Gateway Protocol

LSP: Label Switched Path

LSP-DB: Label Switched Path Database

LSR: Label Switching Router

MDSC: Multi-Domain Service Coordinator

PCC: Path Computation Client

PCE: Path Computation Element

PCEP: Path Computation Element communication Protocol

PNC: Provisioning Network Controller

P-PCE: Parent Path Computation Element

TED: Traffic Engineering Database

VN: Virtual Network

### **[2.1.](#) Requirement Language**

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [[RFC2119](#)] [[RFC8174](#)] when, and only when, they appear in all capitals, as shown here.

### **[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.

Note that, in the scope of this document, both the C-PCEs and the P-PCE are stateful in nature.

[RFC8231] 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 P-PCE (EC-EP) or from a P-PCE towards C-PCE (EP-EC). All PCE types herein (i.e., EC-EP or EP-EC) are assumed to be "stateful PCE".

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

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

LSP State Synchronization (EC-EP): 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 (EC-EP,EP-EC): 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 (EP-EC): a stateful P-PCE requests modification of attributes on a C-PCE's TE LSP.

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

Note that this hierarchy is recursive and thus a Label Switching Router (LSR), as a PCC 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.

[I-D.ietf-pce-hierarchy-extensions] defines the H-PCE Capability TLV that is used in the Open message to advertise the H-PCE capability. [RFC8231] defines the Stateful PCE Capability TLV used in the Open message to indicate stateful support. The presence of both TLVs in an Open message indicates the support for stateful H-PCE operations as described in this document.

Further consideration may be made for optional procedures for stateful communication coordination between PCEs, including procedures to minimise computational loops. The procedures described in [I-D.litkowski-pce-state-sync] facilitate stateful communication between PCEs for various use-cases. The procedures and extensions as described in Section 3 of [I-D.litkowski-pce-state-sync] are also applicable to Child and Parent PCE communication. The SPEAKER-IDENTITY-TLV (defined in [RFC8232]) is included in the LSP object to identify the Ingress (PCC). The PLSP-ID used in the forwarded PCRpt by the C-PCE to P-PCE is same as the original one used by the PCC.

### **3.1. Passive Operations**

Procedures as described in [RFC6805] are applied, where the ingress PCC triggers a path computation request for the destination towards the C-PCE in the domain where the LSP originates. The C-PCE further forwards the 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 [RFC8231], 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 [RFC8231] and [RFC8232] are applicable to a PCEP session between C-PCE and P-PCE as well.





We use the sample hierarchical domain topology example from [\[RFC6805\]](#) as the reference topology for the entirety of this document. It is shown in Figure 1.

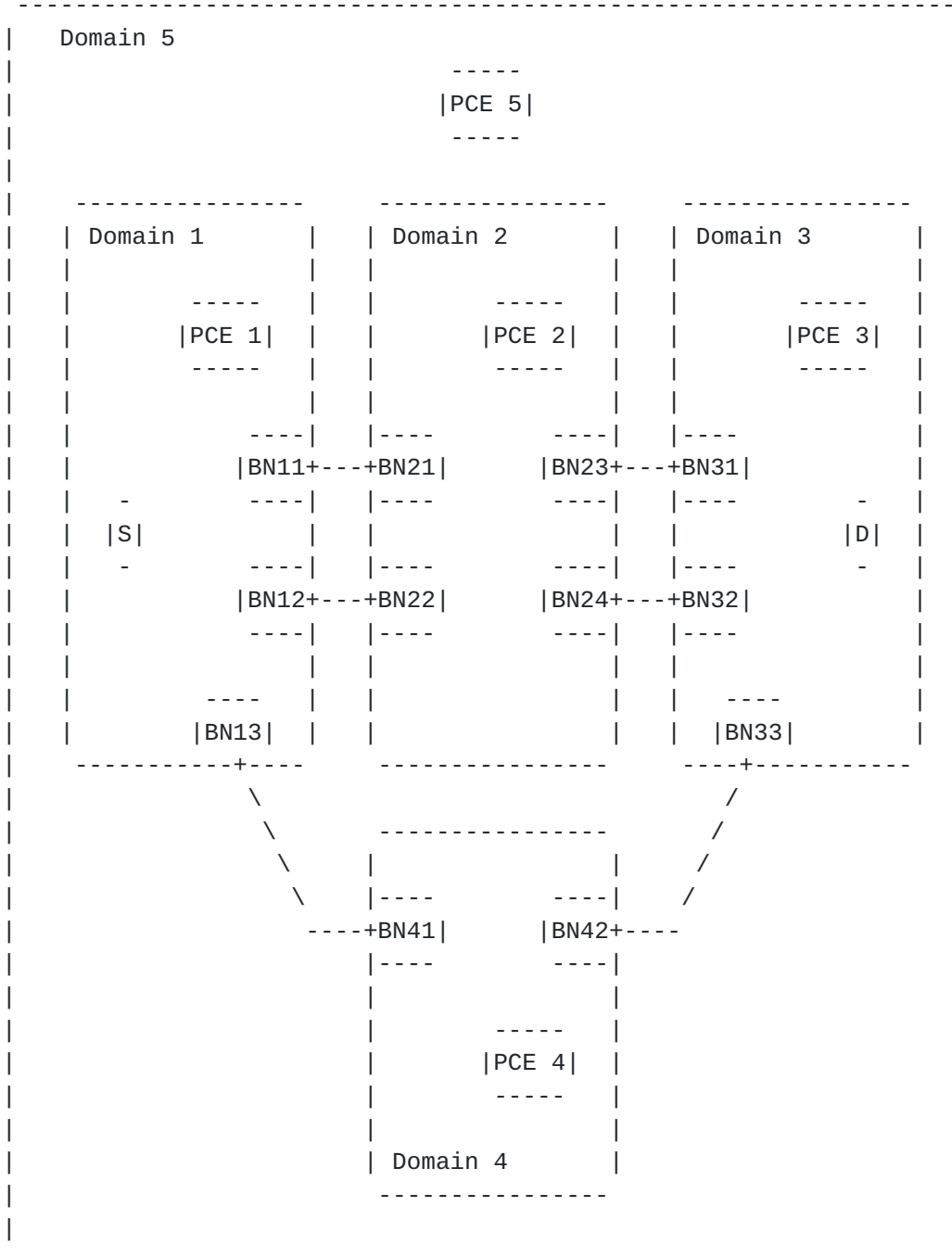


Figure 1: Sample Hierarchical Domain Topology

Steps 1 to 11 are exactly as described in [section 4.6.2 of \[RFC6805\]](#) (Hierarchical PCE End-to-End Path Computation Procedure), the



following additional steps are added for stateful PCE to be executed at the end:

- (A) The Ingress LSR initiates the setup of the LSP as per the path and reports to the PCE1 the LSP status ("GOING-UP").
- (B) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).
- (C) The Ingress LSR notifies the LSP state to PCE1 when the state is "UP".
- (D) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

The Ingress LSR could trigger path re-optimization by sending the path computation request as described in [\[RFC6805\]](#), at this time it can include the LSP object in the PCReq message as described in [\[RFC8231\]](#).

### **[3.2.](#) Active Operations**

[\[RFC8231\]](#) 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 PCC for modifying LSP attributes. The PCC sends back a PCRpt to acknowledge the requested operation or report any change in LSP's state.

As per [\[RFC8051\]](#), 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 sends an LSP Update Request to the PCC 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.



For active operations, the following steps are required when delegating the LSP, again using the reference architecture described in Figure 1 (Sample Hierarchical Domain Topology).

- (A) The Ingress LSR delegates the LSP to the PCE1 via PCRpt message with D flag set.
- (B) The PCE1 further delegates the LSP to the P-PCE (PCE5).
- (C) Steps 4 to 10 of [section 4.6.2 of \[RFC6805\]](#) are executed at P-PCE (PCE5) to determine the end to end path.
- (D) The P-PCE (PCE5) sends the update request to the C-PCE (PCE1) via PCUpd message.
- (E) The PCE1 further updates the LSP to the Ingress LSR (PCC).
- (F) The Ingress LSR initiates the setup of the LSP as per the path and reports to the PCE1 the LSP status ("GOING-UP").
- (G) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).
- (H) The Ingress LSR notifies the LSP state to PCE1 when the state is "UP".
- (I) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

### **3.3. PCE Initiation of LSPs**

[RFC8281] 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 steps are performed, for PCE initiated operations, again using the reference architecture described in Figure 1 (Sample Hierarchical Domain Topology):



- (A) 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.
- (B) The P-PCE (PCE5) sends the initiate request to the child PCE (PCE1) via PCInitiate message.
- (C) The PCE1 further propagates the initiate message to the Ingress LSR (PCC).
- (D) The Ingress LSR initiates the setup of the LSP as per the path and reports to the PCE1 the LSP status ("GOING-UP").
- (E) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).
- (F) The Ingress LSR notifies the LSP state to PCE1 when the state is "UP".
- (G) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

The Ingress LSR (PCC) would generate the PLSP-ID for the LSP and inform the C-PCE, which is propagated to the P-PCE.

### **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 (Refer [\[I-D.dugeon-pce-stateful-interdomain\]](#)).

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

- (A) 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 -
  - o S-BN41
  - o BN41-BN33





- o 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)

- (B) The P-PCE (PCE5) sends the initiate request to the child PCE (PCE3) via PCInitiate message for LSP (BN33-D).
- (C) The PCE3 further propagates the initiate message to BN33.
- (D) BN33 initiates the setup of the LSP as per the path and reports to the PCE3 the LSP status ("GOING-UP").
- (E) The PCE3 further reports the status of the LSP to the P-PCE (PCE5).
- (F) The node BN33 notifies the LSP state to PCE3 when the state is "UP".
- (G) The PCE3 further reports the status of the LSP to the P-PCE (PCE5).

For LSP (BN41-BN33)

- (H) The P-PCE (PCE5) sends the initiate request to the child PCE (PCE4) via PCInitiate message for LSP (BN41-BN33).
- (I) The PCE4 further propagates the initiate message to BN41.
- (J) BN41 initiates the setup of the LSP as per the path and reports to the PCE4 the LSP status ("GOING-UP").
- (K) The PCE4 further reports the status of the LSP to the P-PCE (PCE5).
- (L) The node BN41 notifies the LSP state to PCE4 when the state is "UP".
- (M) The PCE4 further reports the status of the LSP to the P-PCE (PCE5).

For LSP (S-BN41)

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



- (O) The PCE1 further propagates the initiate message to node S.
- (P) S initiates the setup of the LSP as per the path and reports to the PCE1 the LSP status ("GOING-UP").
- (Q) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).
- (R) The node S notifies the LSP state to PCE1 when the state is "UP".
- (S) The PCE1 further reports the status of the LSP to the P-PCE (PCE5).

Additionally:

- (T) Once P-PCE receives report of each per-domain LSP, it should use suitable 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).

Note that each per-domain LSP can be setup in parallel. Further, it is also possible to stitch the per-domain LSP at the same time as the per-domain LSPs are initiated. This option is defined in [\[I-D.dugeon-pce-stateful-interdomain\]](#).

#### 4. Security Considerations

The security considerations listed in [\[RFC8231\]](#), [\[RFC6805\]](#) and [\[RFC5440\]](#) apply to this document as well. As per [\[RFC6805\]](#), it is expected that the parent PCE will require all child PCEs to use full security when communicating with the parent.

Any multi-domain operation necessarily involves the exchange of information across domain boundaries. This is bound to represent a significant security and confidentiality risk especially when the child domains are controlled by different commercial concerns. PCEP allows individual PCEs to maintain confidentiality of their domain path information using path-keys [\[RFC5520\]](#), and the hierarchical PCE architecture is specifically designed to enable as much isolation of domain topology and capabilities information as is possible. The LSP state in the PCRpt message must continue to maintain the internal domain confidentiality when required.

The security consideration for PCE-Initiated LSP as per [\[RFC8281\]](#) is also applicable from P-PCE to C-PCE.



Further, [section 6.3](#) describes the use of path-key [[RFC5520](#)] for confidentiality between C-PCE and P-PCE.

Thus it is RECOMMENDED to secure the PCEP session (between the P-PCE and the C-PCE) using either Transport Layer Security (TLS) [[RFC8253](#)] per the recommendations and best current practices in [[RFC7525](#)] or TCP Authentication Option (TCP-AO) [[RFC5925](#)].

## **5. Manageability Considerations**

All manageability requirements and considerations listed in [[RFC5440](#)], [[RFC6805](#)], [[RFC8231](#)], and [[RFC8281](#)] apply to Stateful H-PCE defined in this document. In addition, requirements and considerations listed in this section apply.

### **5.1. Control of Function and Policy**

Support of the hierarchical procedure will be controlled by the management organization responsible for each child PCE. The parent PCE must only accept path computation requests from authorized child PCEs. If a parent PCE receives report from an unauthorized child PCE, the report should be dropped. All mechanism as described in [[RFC8231](#)] and [[RFC8281](#)] continue to apply.

### **5.2. Information and Data Models**

An implementation should allow the operator to view the stateful and H-PCE capabilities advertised by each peer. The PCEP YANG module [[I-D.ietf-pce-pcep-yang](#)] may be extended to include details for stateful H-PCE deployment and operation, exact attributes to be modeled is out of scope for this document.

### **5.3. Liveness Detection and Monitoring**

Mechanisms defined in this document do not imply any new liveness detection and monitoring requirements in addition to those already listed in [[RFC5440](#)].

### **5.4. Verify Correct Operations**

Mechanisms defined in this document do not imply any new operation verification requirements in addition to those already listed in [[RFC5440](#)] and [[RFC8231](#)].

### **5.5. Requirements On Other Protocols**

Mechanisms defined in this document do not imply any new requirements on other protocols.



## 5.6. Impact On Network Operations

Mechanisms defined in [RFC5440] and [RFC8231] also apply to PCEP extensions defined in this document.

The stateful H-PCE technique brings the applicability of stateful PCE as described in [RFC8051], for the LSP traversing multiple domains.

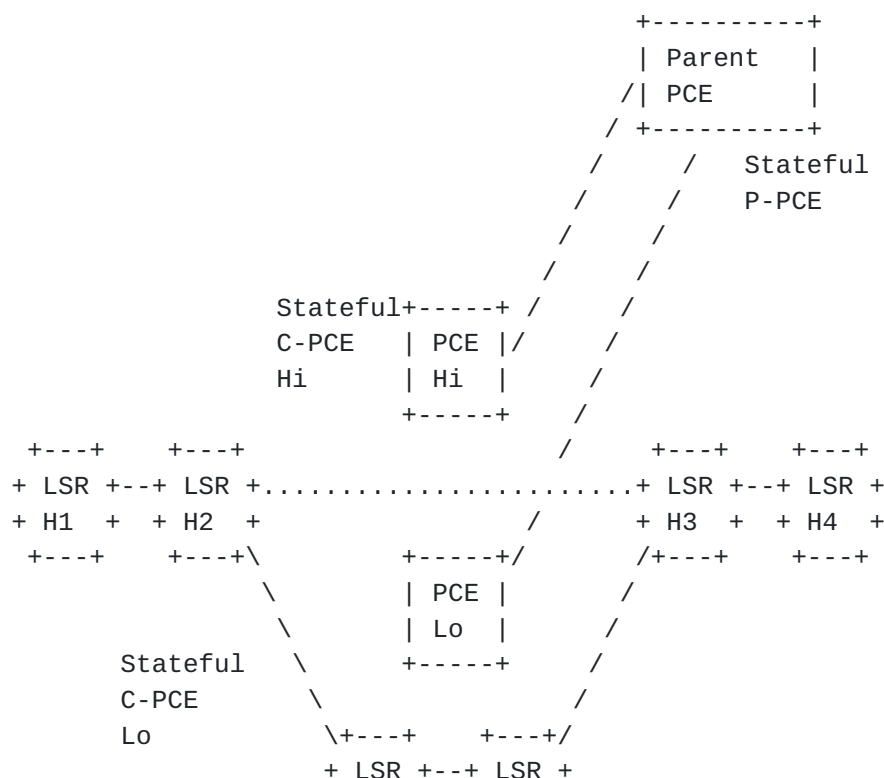
## 5.7. Error Handling between PCEs

Error types and notifications useful for correct PCEP operation may be implemented for managing parent and child PCE interaction. PCEP Error behavior propagation, notification and error criticality level, are further defined in [I-D.ietf-pce-enhanced-errors].

## 6. Other Considerations

### 6.1. Applicability to Inter-Layer Traffic Engineering

[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.







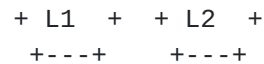


Figure 2: Sample Inter-Layer Topology

All procedures described in [Section 3](#) are applicable to inter-layer (and therefore separate domains) path setup as well.

## 6.2. 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. Scalability considerations for PCEP as per [\[RFC8231\]](#) continue to apply for the PCEP session between child and parent PCE.

## 6.3. Confidentiality

As described in [section 4.2 of \[RFC6805\]](#), information about the content of child domains is not shared for both scaling and confidentiality reasons. Along with the confidentiality during path computation, the child PCE could also conceal the path information, a C-PCE may replace a path segment with a path-key [\[RFC5520\]](#), effectively hiding the content of a segment of a path.

## 7. IANA Considerations

There are no IANA considerations.

## 8. Acknowledgments

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

Thanks to Tal Mazrahi for the RTGDIR review, Paul Kyzivat for the GENART review, and Stephen Farrell for SECDIR review.

## 9. References

### 9.1. 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,



<<https://www.rfc-editor.org/info/rfc2119>>.

- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", [RFC 4655](#), DOI 10.17487/RFC4655, August 2006, <<https://www.rfc-editor.org/info/rfc4655>>.
- [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>>.
- [RFC5520] Bradford, R., Ed., Vasseur, JP., and A. Farrel, "Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Path-Key-Based Mechanism", [RFC 5520](#), DOI 10.17487/RFC5520, April 2009, <<https://www.rfc-editor.org/info/rfc5520>>.
- [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>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", [RFC 8231](#), DOI 10.17487/RFC8231, September 2017, <<https://www.rfc-editor.org/info/rfc8231>>.
- [RFC8281] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for PCE-Initiated LSP Setup in a Stateful PCE Model", [RFC 8281](#), DOI 10.17487/RFC8281, December 2017, <<https://www.rfc-editor.org/info/rfc8281>>.

## 9.2. Informative References

- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), DOI 10.17487/RFC3209, December 2001, <<https://www.rfc-editor.org/info/rfc3209>>.
- [RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label



Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", [RFC 3473](#), DOI 10.17487/RFC3473, January 2003, <<https://www.rfc-editor.org/info/rfc3473>>.

- [RFC4726] Farrel, A., Vasseur, J., and A. Ayyangar, "A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering", [RFC 4726](#), DOI 10.17487/RFC4726, November 2006, <<https://www.rfc-editor.org/info/rfc4726>>.
- [RFC5623] Oki, E., Takeda, T., Le Roux, J.L., and A. Farrel, "Framework for PCE-Based Inter-Layer MPLS and GMPLS Traffic Engineering", [RFC 5623](#), DOI 10.17487/RFC5623, September 2009, <<https://www.rfc-editor.org/info/rfc5623>>.
- [RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", [RFC 5925](#), DOI 10.17487/RFC5925, June 2010, <<https://www.rfc-editor.org/info/rfc5925>>.
- [RFC7525] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", [BCP 195](#), [RFC 7525](#), DOI 10.17487/RFC7525, May 2015, <<https://www.rfc-editor.org/info/rfc7525>>.
- [RFC8051] Zhang, X., Ed. and I. Minei, Ed., "Applicability of a Stateful Path Computation Element (PCE)", [RFC 8051](#), DOI 10.17487/RFC8051, January 2017, <<https://www.rfc-editor.org/info/rfc8051>>.
- [RFC8232] 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", [RFC 8232](#), DOI 10.17487/RFC8232, September 2017, <<https://www.rfc-editor.org/info/rfc8232>>.
- [RFC8253] Lopez, D., Gonzalez de Dios, O., Wu, Q., and D. Dhody, "PCEPS: Usage of TLS to Provide a Secure Transport for the Path Computation Element Communication Protocol (PCEP)", [RFC 8253](#), DOI 10.17487/RFC8253, October 2017, <<https://www.rfc-editor.org/info/rfc8253>>.
- [RFC8453] Ceccarelli, D., Ed. and Y. Lee, Ed., "Framework for Abstraction and Control of TE Networks (ACTN)", [RFC 8453](#), DOI 10.17487/RFC8453, August 2018, <<https://www.rfc-editor.org/info/rfc8453>>.



- [RFC8637] Dhody, D., Lee, Y., and D. Ceccarelli, "Applicability of the Path Computation Element (PCE) to the Abstraction and Control of TE Networks (ACTN)", [RFC 8637](#), DOI 10.17487/RFC8637, July 2019, <<https://www.rfc-editor.org/info/rfc8637>>.
- [I-D.litkowski-pce-state-sync] Litkowski, S., Sivabalan, S., and D. Dhody, "Inter Stateful Path Computation Element communication procedures", [draft-litkowski-pce-state-sync-06](#) (work in progress), July 2019.
- [I-D.ietf-pce-hierarchy-extensions] Zhang, F., Zhao, Q., Dios, O., Casellas, R., and D. King, "Extensions to Path Computation Element Communication Protocol (PCEP) for Hierarchical Path Computation Elements (PCE)", [draft-ietf-pce-hierarchy-extensions-11](#) (work in progress), June 2019.
- [I-D.ietf-pce-pcep-yang] Dhody, D., Hardwick, J., Beeram, V., and J. Tantsura, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", [draft-ietf-pce-pcep-yang-12](#) (work in progress), July 2019.
- [I-D.dugeon-pce-stateful-interdomain] Dugeon, O., Meuric, J., Lee, Y., Dhody, D., and D. Ceccarelli, "PCEP Extension for Stateful Inter-Domain Tunnels", [draft-dugeon-pce-stateful-interdomain-02](#) (work in progress), March 2019.
- [I-D.ietf-pce-lsp-control-request] Raghuram, A., Goddard, A., Yadlapalli, C., Karthik, J., Sivabalan, S., Parker, J., and M. Negi, "Ability for a stateful Path Computation Element (PCE) to request and obtain control of a LSP", [draft-ietf-pce-lsp-control-request-08](#) (work in progress), August 2019.
- [I-D.ietf-pce-enhanced-errors] Pouyllau, et al., "Extensions to PCEP for Enhanced Errors", [draft-ietf-pce-enhanced-errors-06](#) (work in progress), August 2019.

## Contributors

Avantika  
ECI Telecom





India

EMail: avantika.srm@gmail.com

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

EMail: zhang.xian@huawei.com

Udayasree Palle

EMail: udayasreereddy@gmail.com

Oscar Gonzalez de Dios  
Telefonica I+D  
Don Ramon de la Cruz 82-84  
Madrid, 28045  
Spain  
Phone: +34913128832

EMail: oscar.gonzalezdedios@telefonica.com

#### Authors' Addresses

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

EMail: dhruv.ietf@gmail.com

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

EMail: younglee.tx@gmail.com

Daniele Ceccarelli  
Ericsson  
Torshamnsgatan, 48



Stockholm  
Sweden

E-Mail: [daniele.ceccarelli@ericsson.com](mailto:daniele.ceccarelli@ericsson.com)

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

E-Mail: [jongyoon.shin@sk.com](mailto:jongyoon.shin@sk.com)

Daniel King  
Lancaster University  
UK

E-Mail: [d.king@lancaster.ac.uk](mailto:d.king@lancaster.ac.uk)

