               Hierarchical Stateful Path Computation Element (PCE)

Abstract

   A stateful Path Computation Element (PCE) maintains information on
   the current network state received from the Path Computation Clients
   (PCCs), including computed Label Switched Paths (LSPs), reserved
   resources within the network, and pending path computation requests.
   This information may then be considered when computing the path for a
   new traffic-engineered LSP or for any associated/dependent LSPs. The
   path-computation response from a PCE helps the PCC to gracefully
   establish the computed LSP.

   The Hierarchical Path Computation Element (H-PCE) architecture allows
   the optimum sequence of interconnected 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, but not
   stateless, PCEs using the hierarchical PCE architecture.

Status of This Memo

   This document is not an Internet Standards Track specification; it is
   published for informational purposes.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Not all documents
   approved by the IESG are candidates for any level of Internet
   Standard; see Section 2 of RFC 7841.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   https://www.rfc-editor.org/info/rfc8751.

Table of Contents

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 the requests of Path Computation
   Clients (PCCs).

   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).

   [RFC8051] describes general considerations for a stateful PCE
   deployment; it also 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. For its computations, 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.
   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, make changes
   to the attributes of an existing LSP, or delegate control of specific
   LSPs to a new PCE.

   The ability to compute constrained paths for Traffic Engineering (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 that can be used for computing end-to-end
paths for interdomain MPLS TE and GMPLS Label Switched Paths (LSPs).
Within the H-PCE architecture [RFC6805], the Parent PCE (P-PCE) is
used to compute a multidomain path based on the domain connectivity
information. A Child PCE (C-PCE) may be responsible for a single
domain or multiple domains. The C-PCE is used to compute the
intradomain path based on its domain topology information.

This document presents general considerations for stateful PCEs, and
not stateless PCEs, in the hierarchical PCE architecture. It focuses
on the behavior changes and additions to the existing stateful PCE
mechanisms (including PCE-initiated LSP setup and active stateful 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)
interdomain 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; thus, it
maintains the state of the interdomain LSPs that are reported to it.
The interdomain LSP could also be delegated by the C-PCE to the
P-PCE, so that the P-PCE could update the interdomain 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 interdomain 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 interdomain path recomputation as
described in [RFC6805].

The end-to-end interdomain path computation and setup is described in
[RFC6805]. Additionally, a per-domain stitched-LSP model is also
applicable in a P-PCE initiation model. Sections 3.1, 3.2, and 3.3
describe the end-to-end contiguous LSP setup, whereas Section 3.3.1
describes 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 is well suited for
ACTN deployments, as per the hierarchical PCE architecture described
in Section 3.3.1 of this document and Section 4.1 of [RFC8453].

[RFC8637] examines the applicability of PCE to the ACTN framework.
To support the function of multidomain coordination via hierarchy,
the hierarchy of stateful PCEs plays 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 that the VN be

provisioned. The H-PCE architecture as described in [RFC6805] can
support this function using Path Computation Request and Reply (PCReq
and PCRep, respectively) messages between the P-PCE and C-PCEs. When
the CNC requests VN provisioning, the MDSC decomposes this request
into multiple interdomain LSP provisioning requests, which might be
further decomposed into 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 interdomain path of the LSP.

1.2.2. End-to-End Contiguous LSP

Different signaling options for interdomain RSVP-TE 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 for establishing the
optimum path when the sequence of domains is not known in advance.

That document 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.

A stateful P-PCE has to be aware of the interdomain LSPs for it to
consider them during path computation. For instance, when a domain-
diverse path is required from another LSP, the P-PCE needs to be
aware of the LSP. This is the passive stateful P-PCE, as described
in Section 3.1. Additionally, the interdomain 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 a protection LSP
[RFC8745]) or an unrelated LSP whose resource change leads to
reoptimization 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 interdomain 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 interdomain LSP between the transit/egress
domain C-PCEs and the P-PCE is out of the 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
interdomain 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 reoptimization 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 interdomain 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 delegation if it needs to make
a change to the LSP (refer to [RFC8741]).

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

LSPDB: 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. Requirements 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). Usually, 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 Child PCEs (C-PCEs) [RFC6805] and have a direct relationship with
the P-PCE. The P-PCE builds the domain topology map either via
direct configuration or from learned information received from each
C-PCE. The network policy could be applied while building the domain
topology map. This has been described in detail in [RFC6805].

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 (EC-EP and 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 to indicate the state of an LSP.

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 the 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; at any
   time, the C-PCE may withdraw the delegation or the P-PCE may give
   up the delegation.

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 a
   C-PCE to initiate a TE LSP.

Note that this hierarchy is recursive, so a Label Switching Router
(LSR), as a PCC, could delegate control to a PCE. That PCE may, in
turn, delegate to its parent, which may further delegate to its
parent (if it exists). Similarly, update operations can also be
applied recursively.

[RFC8685] 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. To indicate the support for stateful H-PCE
operations described in this document, a PCEP speaker MUST include
both TLVs in an Open message. It is RECOMMENDED that any
implementation that supports stateful operations [RFC8231] and H-PCE
[RFC8685] also implement the 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 minimize computational loops. The procedures described
in [PCE-STATE-SYNC] facilitate stateful communication between PCEs
for various use cases. The procedures and extensions as described in
Section 3 of [PCE-STATE-SYNC] are also applicable to child and parent
PCE communication. The SPEAKER-IDENTITY-ID TLV (defined in
[RFC8232]) is included in the LSP object to identify the ingress
(PCC). The PCEP-specific identifier for the LSP (PLSP-ID [RFC8231])
used in the forwarded PCRpt by the C-PCE to the P-PCE is the same as
the original one used by the PCC.

3.1. Passive Operations

Procedures 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 interdomain 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 interdomain 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], the 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 the C-PCE may dictate which LSPs are reported to the P-PCE. The
PCRpt message is sent from C-PCE to P-PCE.

State synchronization mechanisms as described in [RFC8231] and

[RFC8232] are applicable to a PCEP session between C-PCE and P-PCE as well.

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

```
    -----------------------------------------------------------------
   |   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      |                   |
   |                              ---------------                    |
   |                                                                 |
    -----------------------------------------------------------------
```
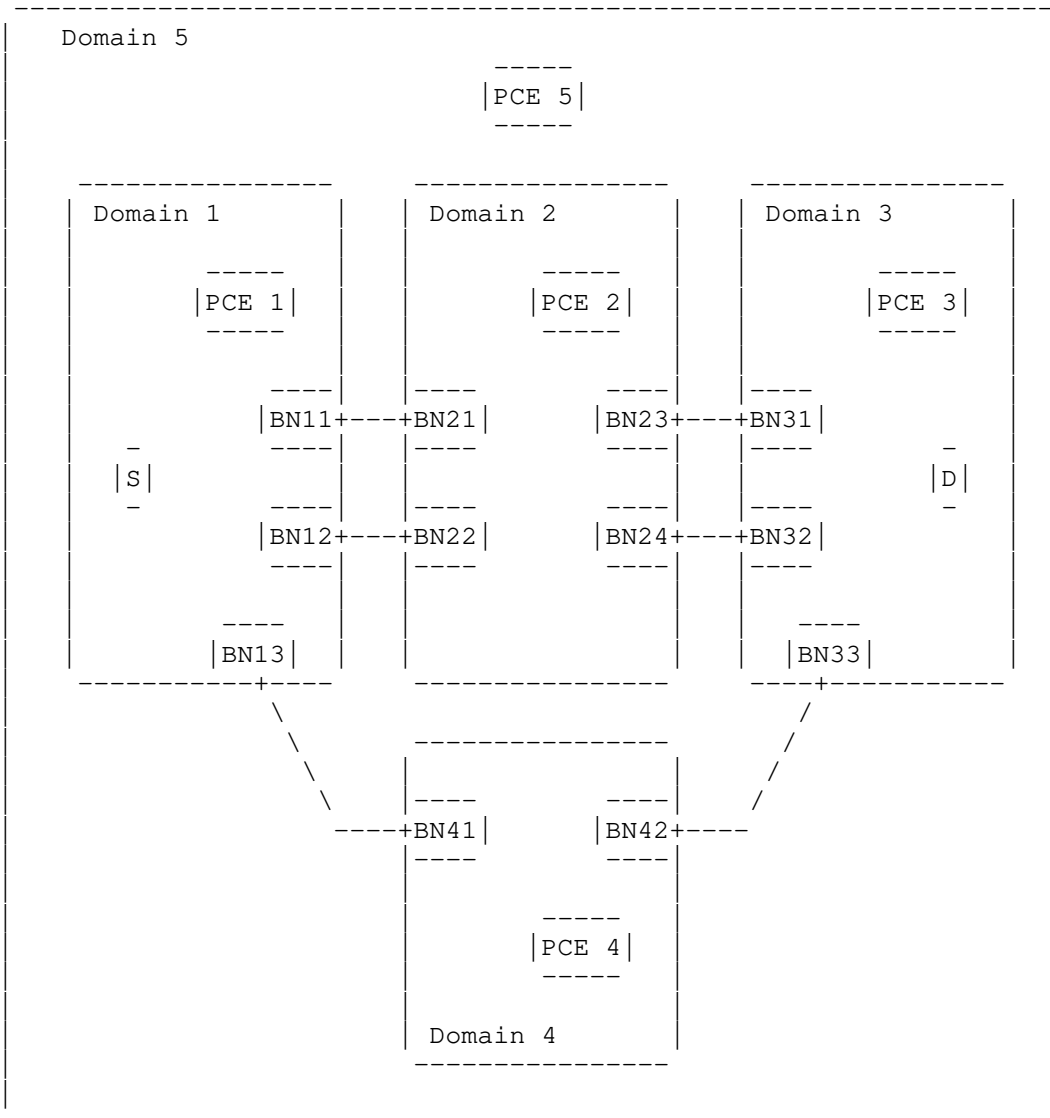
Figure 1: Hierarchical Domain Topology Example

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 the LSP status to PCE1 ("GOING-UP").

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

(C) The ingress LSR notifies PCE1 of the LSP state when the state is "UP".

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

The ingress LSR could trigger path reoptimization 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 an active stateful PCE. The active
PCE functionality uses two specific PCEP messages:

*   Update Request (PCUpd)

*   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 the LSP's state.

As per [RFC8051], delegation is an operation to grant a PCE temporary
rights to modify a subset of LSP parameters on the LSPs of one or
more PCCs. The C-PCE may further choose to delegate to its P-PCE
based on a local policy. The PCRpt message with the "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 an LSP delegated to a P-PCE via the C-PCE, the
P-PCE can use the same PCUpd message to request a change to the C-PCE
(the ingress domain PCE). The C-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 ("Hierarchical Domain Topology Example").

(A)   The ingress LSR delegates the LSP to PCE1 via a PCRpt message
      with D flag set.

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

(C)   Steps 4 to 10 in 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)   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 the LSP status to PCE1 ("GOING-UP").

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

(H)   The ingress LSR notifies PCE1 of the LSP state when the state is
      "UP".

(I)   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 the case of an interdomain LSP
   in hierarchical PCE architecture, the initiation operations can be
   carried out at the P-PCE. In that case, after the P-PCE finishes the
   E2E path computation, it can send the PCInitiate message to the C-PCE
   (the ingress domain PCE), and the C-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 ("Hierarchical
   Domain Topology Example"):

(A) The P-PCE (PCE5) is requested to initiate an LSP. Steps 4 to 10
     in 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) 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 PCE1 the LSP status ("GOING-UP").

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

(F) The ingress LSR notifies PCE1 of the LSP state when the state is
     "UP".

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

The ingress LSR (PCC) generates 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 intradomain LSPs are initiated
in each domain and 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 LSPs are stitched together by some mechanism, which is out of
the scope of this document (Refer to [STATEFUL-INTERDOMAIN]).

The following steps are performed for the per-domain stitched LSP
operation, again using the reference architecture described in
Figure 1 ("Hierarchical Domain Topology Example"):

(A) The P-PCE (PCE5) is requested to initiate an LSP. Steps 4 to 10
     in Section 4.6.2 of [RFC6805] are executed to determine the end-
     to-end path, which is broken into per-domain LSPs. For example:

     * 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):

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

(C) PCE3 further propagates the initiate message to BN33.

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

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

(F) The node BN33 notifies PCE3 of the LSP state when the state is
     "UP".

(G) 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)   PCE4 further propagates the initiate message to BN41.

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

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

   (L)   The node BN41 notifies PCE4 of the LSP state when the state is
         "UP".

   (M)   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 a PCInitiate message for the LSP (S-BN41).

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

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

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

   (R)   The node S notifies PCE1 of the LSP state when the state is
         "UP".

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

   Additionally:

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

   Note that each per-domain LSP can be set up 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
   [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 (i.e., the highest security mechanism available for PCEP)
   when communicating with the parent.

   Any multidomain 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 the confidentiality of their
   domain-path information using path-keys [RFC5520], and the
   hierarchical PCE architecture is specifically designed to enable as
   much isolation of information about domain topology and capabilities
   as is possible. The LSP state in the PCRpt message must continue to
   maintain the internal domain confidentiality when required.

   The security considerations for PCE-initiated LSP in [RFC8281] are
   also applicable from P-PCE to C-PCE.

   Further, Section 6.3 describes the use of a 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 Transport Layer Security (TLS) [RFC8446] (per

the recommendations and best current practices in BCP 195 [RFC7525])
and/or TCP Authentication Option (TCP-AO) [RFC5925]. The guidance
for implementing PCEP with TLS can be found in [RFC8253].

In the case of TLS, due care needs to be taken while exposing the
parameters of the X.509 certificate -- such as
subjectAltName:otherName, which is set to Speaker Entity Identifier
[RFC8232] as per [RFC8253] -- to ensure uniqueness and avoid any
mismatch.

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 a report from an unauthorized child
PCE, the report should be dropped. All mechanisms 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 "ietf-pcep" PCEP
YANG module is specified in [PCE-PCEP-YANG]. This YANG module will
be required to be augmented to also include details for stateful
H-PCE deployment and operation. The exact model and attributes are
out of scope for this document.

5.3. Liveness Detection and Monitoring

Mechanisms defined in this document do not imply any new liveness-
detection or monitoring requirements in addition to those already
listed in [RFC5440].

5.4. Verification of 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
(described in [RFC8051]) to the LSP traversing multiple domains.

As described in Section 3, a PCEP speaker includes both the H-PCE-
CAPABILITY TLV [RFC8685] and STATEFUL-PCE-CAPABILITY TLV [RFC8231] to
indicate support for stateful H-PCE. Note that there is a
possibility of a PCEP speaker that does not support the stateful
H-PCE feature but does provide support for stateful-PCE [RFC8231] and
H-PCE [RFC8685] features. This PCEP speaker will also include both
the TLVs; in this case, a PCEP peer could falsely assume that the
stateful H-PCE feature is also supported. On further PCEP message
exchange, the stateful messages will not be propagated further (as
described in this document), and a stateful H-PCE-based "parent"
control of the LSP will not happen. A PCEP peer should be prepared

for this eventuality as a part of normal procedures.

5.7. Error Handling between PCEs

   Apart from the basic error handling described in this document, an
   implementation could also use the enhanced error and notification
   mechanism for stateful H-PCE operations described in
   [PCE-ENHANCED-ERRORS]. Enhanced features such as error-behavior
   propagation, notification, and error-criticality level are further
   defined in [PCE-ENHANCED-ERRORS].

6. Other Considerations

6.1. Applicability to Interlayer Traffic Engineering

   [RFC5623] describes a framework for applying the PCE-based
   architecture to interlayer (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 Figure 2.

```
                                          +----------+
                                          | Parent   |
                                       /| PCE       |
                                      / +----------+
                                     /    /    Stateful
                                    /    /     P-PCE
                                   /    /
                                  /    /
                                 /    /
                   Stateful+-----+ /    /
                   C-PCE    | PCE |/    /
                   Hi       | Hi  |    /
                            +-----+   /
    +---+     +---+                  /      +---+     +---+
   + LSR +--+ LSR +.......................+ LSR +--+ LSR +
   + H1  +  + H2  +                 /      + H3  +  + H4  +
    +---+     +---+\        +-----+/       /+---+     +---+
                   \        | PCE |       /
                    \       | Lo  |      /
         Stateful    \      +-----+     /
         C-PCE        \                /
         Lo            \+---+    +---+/
                       + LSR +--+ LSR +
                       + L1  +  + L2  +
                        +---+     +---+
```
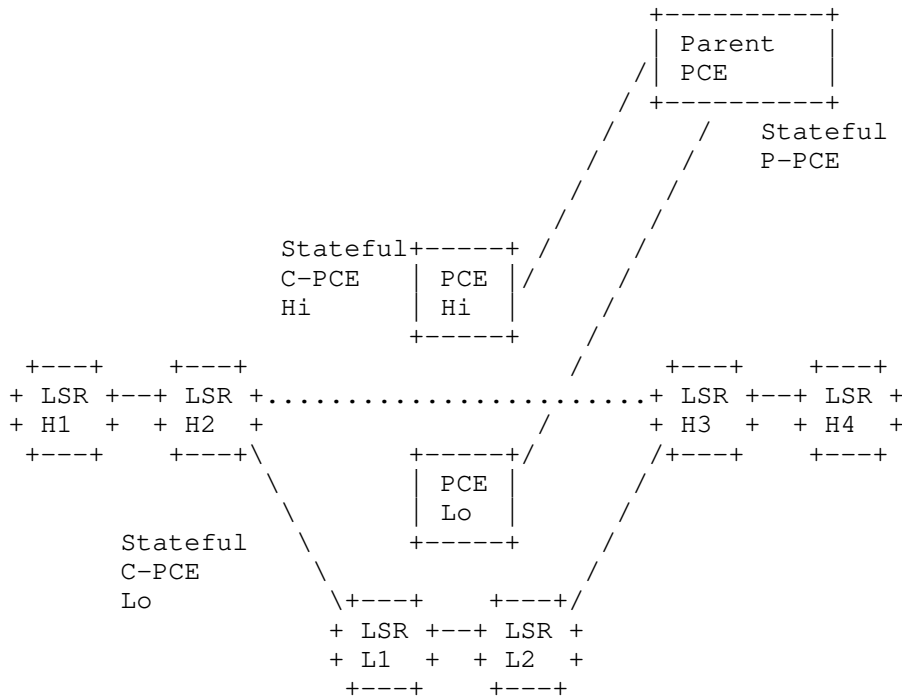
                Figure 2: Sample Interlayer Topology

   All procedures described in Section 3 are also applicable to
   interlayer path setup, and therefore to separate domains.

6.2. Scalability Considerations

   It should be noted that if all the C-PCEs were to 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 that are involved in
   H-PCE -- i.e., the LSPs that 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. The child PCE could also conceal the path
   information during path computation. 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

This document has no IANA actions.

8. References

8.1. 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.-P., 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,
              <https://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>.

   [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>.

   [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,
              <https://www.rfc-editor.org/info/rfc6805>.

   [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>.

   [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>.

   [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>.

   [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>.

   [RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol
              Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018,
              <https://www.rfc-editor.org/info/rfc8446>.

8.2. 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.-P., 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, JL., 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>.

[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>.

[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>.

[RFC8685]   Zhang, F., Zhao, Q., Gonzalez de Dios, O., Casellas, R.,
            and D. King, "Path Computation Element Communication
            Protocol (PCEP) Extensions for the Hierarchical Path
            Computation Element (H-PCE) Architecture", RFC 8685,
            DOI 10.17487/RFC8685, December 2019,
            <https://www.rfc-editor.org/info/rfc8685>.

[RFC8741]   Raghuram, A., Goddard, A., Karthik, J., Sivabalan, S., and
            M. Negi, "Ability for a Stateful Path Computation Element
            (PCE) to Request and Obtain Control of a Label Switched
            Path (LSP)", RFC 8741, DOI 10.17487/RFC8741, March 2020,
            <https://www.rfc-editor.org/info/rfc8741>.

[RFC8745]   Ananthakrishnan, H., Sivabalan, S., Barth, C., Minei, I.,
            and M. Negi, "Path Computation Element Communication
            Protocol (PCEP) Extensions for Associating Working and
            Protection Label Switched Paths (LSPs) with Stateful PCE",
            RFC 8745, DOI 10.17487/RFC8745, March 2020,
            <https://www.rfc-editor.org/info/rfc8745>.

[PCE-ENHANCED-ERRORS]
            Poullyau, H., Theillaud, R., Meuric, J., Zheng, H., and X.
            Zhang, "Extensions to the Path Computation Element
            Communication Protocol for Enhanced Errors and
            Notifications", Work in Progress, Internet-Draft, draft-
            ietf-pce-enhanced-errors-06, 14 August 2019,
            <https://tools.ietf.org/html/draft-ietf-pce-enhanced-
            errors-06>.

[PCE-PCEP-YANG]
          Dhody, D., Hardwick, J., Beeram, V., and J. Tantsura, "A
          YANG Data Model for Path Computation Element
          Communications Protocol (PCEP)", Work in Progress,
          Internet-Draft, draft-ietf-pce-pcep-yang-13, 31 October
          2019,
          <https://tools.ietf.org/html/draft-ietf-pce-pcep-yang-13>.

[PCE-STATE-SYNC]
          Litkowski, S., Sivabalan, S., Li, C., and H. Zheng, "Inter
          Stateful Path Computation Element (PCE) Communication
          Procedures.", Work in Progress, Internet-Draft, draft-
          litkowski-pce-state-sync-07, 11 January 2020,
          <https://tools.ietf.org/html/draft-litkowski-pce-state-
          sync-07>.

[STATEFUL-INTERDOMAIN]
          Dugeon, O., Meuric, J., Lee, Y., and D. Ceccarelli, "PCEP
          Extension for Stateful Inter-Domain Tunnels", Work in
          Progress, Internet-Draft, draft-dugeon-pce-stateful-
          interdomain-02, 4 March 2019,
          <https://tools.ietf.org/html/draft-dugeon-pce-stateful-
          interdomain-02>.

Contributors

   Avantika
   ECI Telecom
   India

   Email: avantika.srm@gmail.com

   Xian Zhang
   Huawei Technologies
   Bantian, Longgang District
   Guangdong
   Shenzhen, 518129
   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
   28045 Madrid
   Spain

   Phone: +34913128832
   Email: oscar.gonzalezdedios@telefonica.com

Authors' Addresses

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

   Email: dhruv.ietf@gmail.com

   Young Lee
   Samsung Electronics

   Email: younglee.tx@gmail.com

   Daniele Ceccarelli
   Ericsson
   Torshamnsgatan, 48
   SE- 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

   Daniel King
   Lancaster University
   United Kingdom

   Email: d.king@lancaster.ac.uk