

PCE Working Group
Internet-Draft
Intended status: Informational
Expires: December 19, 2018

D. Dhody
Y. Lee
Huawei Technologies
D. Ceccarelli
Ericsson
June 17, 2018

Applicability of Path Computation Element (PCE) for Abstraction and
Control of TE Networks (ACTN)
draft-ietf-pce-applicability-actn-06

Abstract

Abstraction and Control of TE Networks (ACTN) refers to the set of virtual network (VN) operations needed to orchestrate, control and manage large-scale multi-domain TE networks so as to facilitate network programmability, automation, efficient resource sharing, and end-to-end virtual service aware connectivity and network function virtualization services.

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.

This document examines the applicability of PCE to the ACTN framework.

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 19, 2018.

Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 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

1. Introduction	2
1.1. Path Computation Element (PCE)	2
1.1.1. Role of PCE in SDN	3
1.1.2. PCE in multi-domain and multi-layer deployments	4
1.1.3. Relationship to PCE based central control	4
1.2. Abstraction and Control of TE Networks (ACTN)	5
1.3. PCE and ACTN	7
2. Architectural Considerations	7
2.1. Multi domain coordination via Hierarchy	7
2.2. Abstraction function	8
2.3. Customer mapping function	9
2.4. Virtual Service Coordination	10
3. Interface Considerations	10
4. Realizing ACTN with PCE (and PCEP)	11
5. IANA Considerations	15
6. Security Considerations	15
7. Acknowledgments	15
8. References	15
8.1. Normative References	15
8.2. Informative References	16
Authors' Addresses	19

1. Introduction

1.1. Path Computation Element (PCE)

The Path Computation Element Communication Protocol (PCEP) [RFC5440] provides mechanisms for Path Computation Elements (PCEs) [RFC4655] to perform path computations in response to Path Computation Clients (PCCs) requests.

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.

A stateful PCE [RFC8231] 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 a PCC to delegate control of specific LSPs to a new PCE.

1.1.1. Role of PCE in SDN

Software-Defined Networking (SDN) [RFC7149] refers to a separation between the control elements and the forwarding components so that software running in a centralized system called a controller, can act to program the devices in the network to behave in specific ways. A required element in an SDN architecture is a component that plans how the network resources will be used and how the devices will be programmed. It is possible to view this component as performing specific computations to place flows within the network given knowledge of the availability of network resources, how other forwarding devices are programmed, and the way that other flows are routed. It is concluded in [RFC7399], that this is the same function that a PCE might offer in a network operated using a dynamic control plane. This is the function and purpose of a PCE, and the way that a PCE integrates into a wider network control system including SDN is presented in Application-Based Network Operation (ABNO) [RFC7491].

1.1.2. PCE in multi-domain and multi-layer deployments

Computing paths across large multi-domain environments require special computational components and cooperation between entities in different domains capable of complex path computation. The PCE provides an architecture and a set of functional components to address this problem space. A PCE may be used to compute end-to-end paths across multi-domain environments using a per-domain path computation technique [RFC5152]. The Backward recursive PCE based path computation (BRPC) mechanism [RFC5441] defines a PCE-based path computation procedure to compute inter-domain constrained MPLS and GMPLS TE networks. However, both per-domain and BRPC techniques assume that the sequence of domains to be crossed from source to destination is known, either fixed by the network operator or obtained by other means.

[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) when the domain sequence is not known. Within the Hierarchical PCE (H-PCE) architecture, 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.

[I-D.ietf-pce-stateful-hpce] state the 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.

[RFC5623] describes a framework for applying the PCE-based architecture to inter-layer to (G)MPLS TE. It provides suggestions for the deployment of PCE in support of multi-layer networks. It also describes the relationship between PCE and a functional component in charge of the control and management of the Virtual Network Topology (VNT) [RFC5212], called the VNT Manager (VNTM).

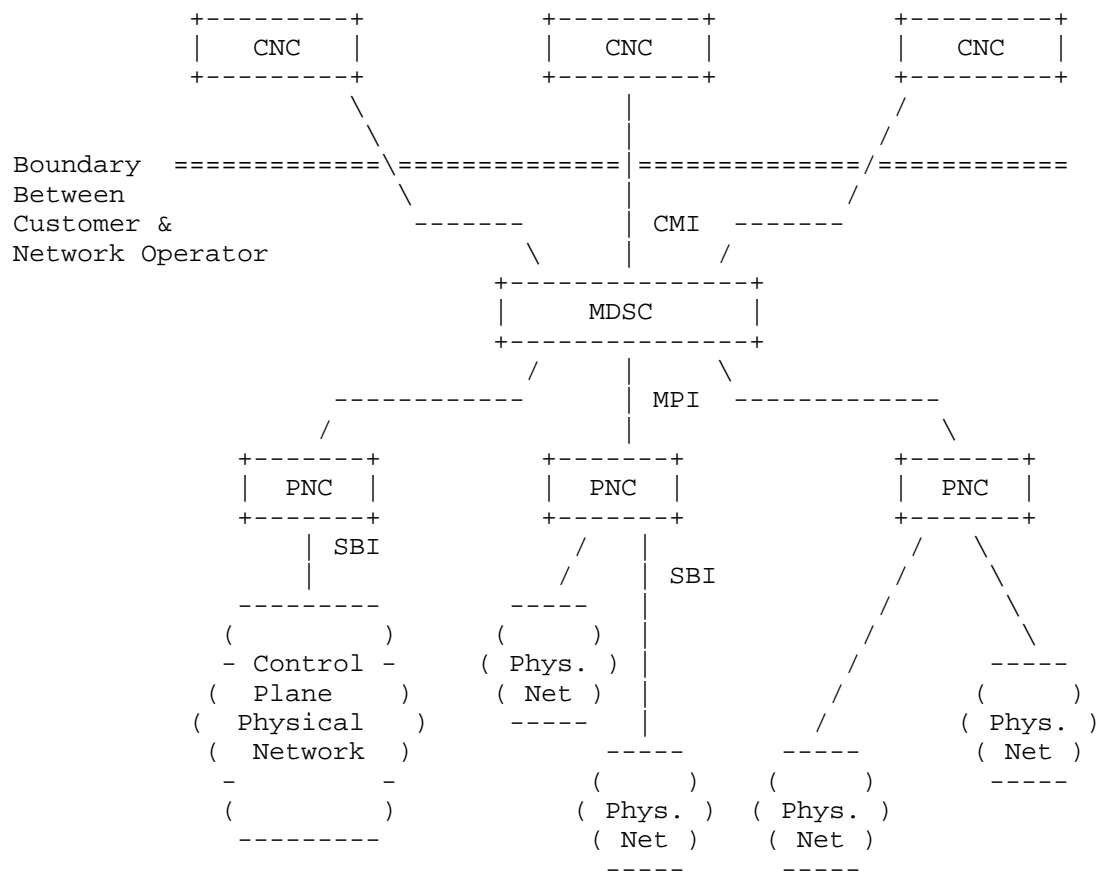
1.1.3. Relationship to PCE based central control

[RFC8283] introduces the architecture for PCE as a central controller (PCECC), it further examines the motivations and applicability for PCEP as a southbound interface, and introduces the implications for the protocol. The section 2.1.3 of [RFC8283] describe an hierarchy of PCE-based controller as per the Hierarchy of PCE framework defined in [RFC6805].

1.2. Abstraction and Control of TE Networks (ACTN)

[I-D.ietf-teas-actn-requirements] describes the high-level ACTN requirements. [I-D.ietf-teas-actn-framework] describes the architecture model for ACTN including the entities (Customer Network Controller(CNC), Multi-domain Service Coordinator(MDSC), and Provisioning Network Controller (PNC) and their interfaces.

The ACTN reference architecture identified a three-tier control hierarchy as depicted in Figure 1:



CMI - (CNC-MDSC Interface)

MPI - (MDSC-PNC Interface)

Figure 1: ACTN Hierarchy

The two interfaces with respect to the MDSC, one north of the MDSC (CMI CNC-MDSC Interface) and one south (MPI MDSC-PNC Interface). A hierarchy of MDSC is possible with a recursive MPI interface.

[I-D.ietf-teas-actn-info-model] provides an information model for ACTN interfaces.

1.3. PCE and ACTN

This document examines the PCE and ACTN architecture and describes how the PCE architecture is applicable to ACTN. It also lists the PCEP extensions that are needed to use PCEP as an ACTN interface. This document also identifies any gaps in PCEP, that exist at the time of publication of this document.

Further, ACTN, Stateful H-PCE, and PCECC are based on the same basic hierarchy framework and thus compatible with each other.

2. Architectural Considerations

ACTN [I-D.ietf-teas-actn-framework] architecture is based on hierarchy and recursiveness of controllers. It defines three types of controllers (depending on the functionalities they implement). The main functionalities are -

- o Multi domain coordination function
- o Abstraction function
- o Customer mapping/translation function
- o Virtual service coordination function

Section 3 of [I-D.ietf-teas-actn-framework] describes these functions.

It should be noted that, this document lists all possible ways in which PCEP could be used for each of the above functions, but all functions are not required to be implemented via PCEP. Operator may choose to use the PCEP for multi domain coordination via stateful H-PCE but use RESTCONF [RFC8040] or BGP-LS [RFC7752] to get the topology and support abstraction function.

2.1. Multi domain coordination via Hierarchy

With the definition of domain being "everything that is under the control of the single logical controller", as per [I-D.ietf-teas-actn-framework], it is needed to have a control entity that oversees the specific aspects of the different domains and to build a single abstracted end-to-end network topology in order to coordinate end-to-end path computation and path/service provisioning.

The MDSC in ACTN framework realizes this function by coordinating the per-domain PNCs in a hierarchy of controllers. It also needs to

detach from the underlying network technology and express customer concerns by business needs.

[RFC6805] and [I-D.ietf-pce-stateful-hpce] describes a hierarchy of PCE with Parent PCE coordinating multi-domain path computation function between Child PCE(s). It is easy to see how these principles align, and thus how stateful H-PCE architecture can be used to realize ACTN.

The Per domain stitched LSP in the Hierarchical stateful PCE architecture, described in Section 3.3.1 of [I-D.ietf-pce-stateful-hpce] is well suited for multi-domain coordination function. This includes domain sequence selection; E2E path computation; Controller (PCE) initiated path setup and reporting. This is also applicable to multi-layer coordination in case of IP+optical networks.

[I-D.litkowski-pce-state-sync]" describes the procedures to allow a stateful communication between PCEs for various use-cases. The procedures and extensions are also applicable to Child and Parent PCE communication and thus useful for ACTN as well.

2.2. Abstraction function

To realize ACTN, an abstracted view of the underlying network resources needs to be built. This includes global network-wide abstracted topology based on the underlying network resources of each domain. This also include abstract topology created as per the customer service connectivity requests and represented as a network slice allocated to each customer.

In order to compute and provide optimal paths, PCEs require an accurate and timely Traffic Engineering Database (TED). Traditionally this TED has been obtained from a link state (LS) routing protocol supporting traffic engineering extensions. PCE may construct its TED by participating in the IGP ([RFC3630] and [RFC5305] for MPLS-TE; [RFC4203] and [RFC5307] for GMPLS). An alternative is offered by BGP-LS [RFC7752].

In case of H-PCE [RFC6805], the parent PCE needs to build the domain topology map of the child domains and their interconnectivity. [RFC6805] and [I-D.ietf-pce-inter-area-as-applicability] suggest that BGP-LS could be used as a "northbound" TE advertisement from the child PCE to the parent PCE.

[I-D.dhodylee-pce-pcep-ls] proposes another approaches for learning and maintaining the Link-State and TE information as an alternative to IGPs and BGP flooding, using PCEP itself. The child PCE can use

this mechanism to transport Link-State and TE information from child PCE to a Parent PCE using PCEP.

In ACTN, there is a need to control the level of abstraction based on the deployment scenario and business relationship between the controllers. The mechanism used to disseminate information from PNC (child PCE) to MDSC (parent PCE) should support abstraction. [I-D.ietf-teas-actn-framework] describes a few alternative approaches of abstraction. The resulting abstracted topology can be encoded using the PCEP-LS mechanisms [I-D.dhodylee-pce-pcep-ls] and its optical network extension [I-D.lee-pce-pcep-ls-optical]. PCEP-LS is an attractive option when the operator would wish to have a single control plane protocol (PCEP) to achieve ACTN functions.

[I-D.ietf-teas-actn-framework] discusses two ways to build abstract topology from an MDSC standpoint with interaction with PNCs. The primary method is called automatic generation of abstract topology by configuration. With this method, automatic generation is based on the abstraction/summarization of the whole domain by the PNC and its advertisement on the MPI. The secondary method is called on-demand generation of supplementary topology via Path Compute Request/Reply. This method may be needed to obtain further complementary information such as potential connectivity from child PCEs in order to facilitate an end-to-end path provisioning. PCEP is well suited to support both methods.

2.3. Customer mapping function

In ACTN, there is a need to map customer virtual network (VN) requirements into network provisioning request to the PNC. That is, the customer requests/commands are mapped into network provisioning requests that can be sent to the PNC. Specifically, it provides mapping and translation of a customer's service request into a set of parameters that are specific to a network type and technology such that network configuration process is made possible.

[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. As described in [I-D.ietf-pce-stateful-hpce], for inter-domain LSP in Hierarchical PCE architecture, the initiation operations can be carried out at the parent PCE. In which case after parent PCE finishes the E2E path computation, it can send the PCInitiate message to the child PCE, the child PCE further propagates the initiate request to the LSR. The customer request is received by the MDSC (parent PCE) and based on

the business logic, global abstracted topology, network conditions and local policy, the MDSC (parent PCE) translates this into per domain LSP initiation request that a PNC (child PCE) can understand and act on. This can be done via the PCInitiate message.

PCEP extensions for associating opaque policy between PCEP peer [I-D.ietf-pce-association-policy] can be used.

2.4. Virtual Service Coordination

Virtual service coordination function in ACTN incorporates customer service-related information into the virtual network service operations in order to seamlessly operate virtual networks while meeting customer's service requirements.

[I-D.leedhody-pce-vn-association] describes the need for associating a set of LSPs with a VN "construct" to facilitate VN operations in PCE architecture. This association allows the PCEs to identify which LSPs belong to a certain VN.

This association based on VN is useful for various optimizations at the VN level which can be applied to all the LSPs that are part of the VN slice. During path computation, the impact of a path for an LSP is compared against the paths of other LSPs in the VN. This is to make sure that the overall optimization and SLA of the VN rather than of a single LSP. Similarly, during re-optimization, advanced path computation algorithm and optimization technique can be considered for all the LSPs belonging to a VN/customer and optimize them all together.

3. Interface Considerations

As per [I-D.ietf-teas-actn-framework], to allow virtualization and multi domain coordination, the network has to provide open, programmable interfaces, in which customer applications can create, replace and modify virtual network resources and services in an interactive, flexible and dynamic fashion while having no impact on other customers. The two ACTN interfaces are -

- o The CNC-MDSC Interface (CMI) is an interface between a Customer Network Controller and a Multi Domain Service Coordinator. It requests the creation of the network resources, topology or services for the applications. The MDSC may also report potential network topology availability if queried for current capability from the Customer Network Controller.
- o The MDSC-PNC Interface (MPI) is an interface between a Multi Domain Service Coordinator and a Provisioning Network Controller.

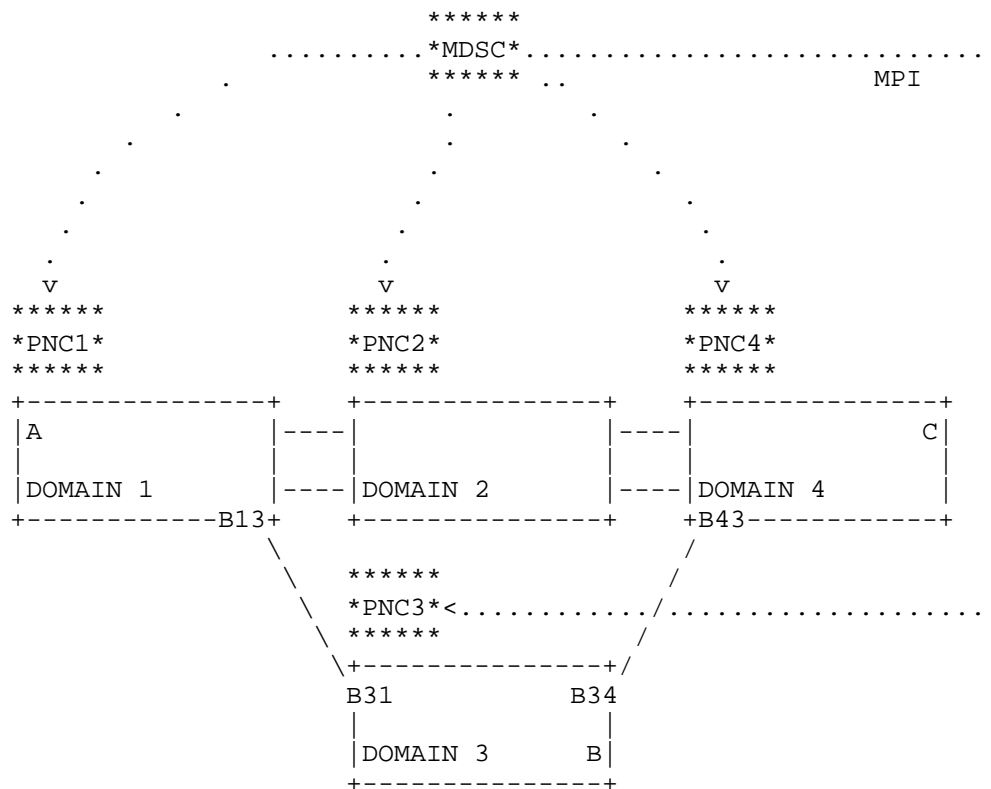
It communicates the creation request, if required, of new connectivity of bandwidth changes in the physical network, via the PNC. In multi-domain environments, the MDSC needs to establish multiple MPIs, one for each PNC, as there are multiple PNCs responsible for its domain control.

- o In case of hierarchy of MDSC, the MPI is applied recursively. From an abstraction point of view, the top level MDSC which interfaces the CNC operates on a higher level of abstraction (i.e., less granular level) than the lower level MSDCs.

PCEP is especially suitable on the MPI as it meets the requirement and the functions as set out in the ACTN framework [I-D.ietf-teas-actn-framework]. Its recursive nature is well suited via the multi-level hierarchy of PCE. PCEP can also be applied to the CMI as the CNC can be a path computation client while the MDSC can be a path computation server. The Section 4 describe how PCE and PCEP could help realize ACTN on the MPI.

4. Realizing ACTN with PCE (and PCEP)

As per the example in the Figure 2, there are 4 domains, each with its own PNC and a MDSC at top. The PNC and MDSC need PCE as a important function. The PNC (or child PCE) already uses PCEP to communicate to the network device. It can utilize the PCEP as the MPI to communicate between controllers too.



```
MDSC -> Parent PCE
PNC   -> Child  PCE
MPI   -> PCEP
```

Figure 2: ACTN with PCE

- o Building Domain Topology at MDSC: PNC (or child PCE) needs to have the TED to compute path in its domain. As described in Section 2.2, it can learn the topology via IGP or BGP-LS. PCEP-LS is also a proposed mechanism to carry link state and traffic engineering information within PCEP. A mechanism to carry abstracted topology while hiding technology specific information between PNC and MDSC is described in [I-D.dhodylee-pce-pcep-ls]. At the end of this step the MDSC (or parent PCE) has the abstracted topology from each of its PNC (or child PCE). This could be as simple as a domain topology map as described in [RFC6805] or it can have full topology information of all domains. The latter is not scalable and thus an abstracted topology of each

domain interconnected by inter-domain links is the most common case.

- * Topology Change: When the PNC learns of any topology change, the PNC needs to decide if the change needs to be notified to the MDSC. This is dependent on the level of abstraction between the MDSC and the PNC.
- o VN Instantiate: MDSC is requested to instantiate a VN, the minimal information that is required would be a VN identifier and a set of end points. Various path computation, setup constraints and objective functions may also be provided. In PCE terms, a VN Instantiate can be considered as a set of paths belonging to the same VN. As described in Section 2.4 and [I-D.leedhody-pce-vn-association] the VN association can help in identifying the set of paths that belong to a VN. The rest of the information like the endpoints, constraints and objective function (OF) is already defined in PCEP in terms of a single path.
- * Path Computation: As per the example in the Figure 2, the VN instantiate requires two end to end paths between (A in Domain 1 to B in Domain 3) and (A in Domain 1 to C in Domain 4). The MDSC (or parent PCE) triggers the end to end path computation for these two paths. MDSC can do path computation based on the abstracted domain topology that it already has or it may use the H-PCE procedures (Section 2.1) using the PCReq and PCRep messages to get the end to end path with the help of the child PCEs (PNC). Either way, the resulted E2E paths may be broken into per-domain paths.
- * A-B: (A-B13,B13-B31,B31-B)
- * A-C: (A-B13,B13-B31,B34-B43,B43-C)
- * Per Domain Path Instantiation: Based on the above path computation, MDSC can issue the path instantiation request to each PNC via PCInitiate message (see [I-D.ietf-pce-stateful-hpce] and [I-D.leedhody-pce-vn-association]). A suitable stitching mechanism would be used to stitch these per domain LSPs. One such mechanism is described in [I-D.lee-pce-lsp-stitching-hpce], where PCEP is extended to support stitching in stateful H-PCE context.
- * Per Domain Path Report: Each PNC should report the status of the per-domain LSP to the MDSC via PCRpt message, as per the Hierarchy of stateful PCE ([I-D.ietf-pce-stateful-hpce]). The

status of the end to end LSP (A-B and A-C) is made up when all the per domain LSP are reported up by the PNCs.

- * Delegation: It is suggested that the per domain LSPs are delegated to respective PNC, so that they can control the path and attributes based on each domain network conditions.
- * State Synchronization: The state needs to be synchronized between the parent PCE and child PCE. The mechanism described in [I-D.litkowski-pce-state-sync] can be used.
- o VN Modify: MDSC is requested to modify a VN, for example the bandwidth for VN is increased. This may trigger path computation at MDSC as described in the previous step and can trigger an update to existing per-intra-domain path (via PCUpd message) or creation (or deletion) of a per-domain path (via PCInitiate message). As described in [I-D.ietf-pce-stateful-hpce], this should be done in make-before-break fashion.
- o VN Delete: MDSC is requested to delete a VN, in this case, based on the E2E paths and the resulting per-domain paths need to be removed (via PCInitiate message).
- o VN Update (based on network changes): Any change in the per-domain LSP are reported to the MDSC (via PCRpt message) as per [I-D.ietf-pce-stateful-hpce]. This may result in changes in the E2E path or VN status. This may also trigger a re-optimization leading to a new per-domain path, update to existing path, or deletion of the path.
- o VN Protection: The VN protection/restoration requirements, need to be applied to each E2E path as well as each per domain path. The MDSC needs to play a crucial role in coordinating the right protection/restoration policy across each PNC. The existing protection/restoration mechanism of PCEP can be applied on each path.
- o In case PNC generates an abstract topology to the MDSC, the PCInitiate/PCUpd messages from the MDSC to a PNC will contain a path with abstract nodes and links. PNC would need to take that as an input for path computation to get a path with physical nodes and links. Similarly PNC would convert the path received from the device (with physical nodes and links) into abstract path (based on the abstract topology generated before with abstract nodes and links) and reported to the MDSC.

5. IANA Considerations

This is an informational document and thus does not have any IANA allocations to be made.

6. Security Considerations

The ACTN framework described in [I-D.ietf-teas-actn-framework] defines key components and interfaces for managed traffic engineered networks. It also lists various security considerations such as request and control of resources, confidentiality of the information, and availability of function which should be taken into consideration.

When PCEP is used on the MPI, this interface needs to be secured, use of [RFC8253] is RECOMMENDED. Each PCEP extension listed in this document, presents its individual security considerations, which continue to apply.

7. Acknowledgments

The authors would like to thank Jonathan Hardwick for the inspiration behind this document. Further thanks to Avantika for her comments with suggested text.

8. References

8.1. Normative References

- [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>>.
- [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>>.
- [I-D.ietf-teas-actn-framework] Ceccarelli, D. and Y. Lee, "Framework for Abstraction and Control of Traffic Engineered Networks", draft-ietf-teas-actn-framework-15 (work in progress), May 2018.

8.2. Informative References

- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <<https://www.rfc-editor.org/info/rfc3630>>.
- [RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <<https://www.rfc-editor.org/info/rfc4203>>.
- [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>>.
- [RFC5152] Vasseur, JP., Ed., Ayyangar, A., Ed., and R. Zhang, "A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)", RFC 5152, DOI 10.17487/RFC5152, February 2008, <<https://www.rfc-editor.org/info/rfc5152>>.
- [RFC5212] Shiimoto, K., Papadimitriou, D., Le Roux, JL., Vigoureux, M., and D. Brungard, "Requirements for GMPLS-Based Multi-Region and Multi-Layer Networks (MRN/MLN)", RFC 5212, DOI 10.17487/RFC5212, July 2008, <<https://www.rfc-editor.org/info/rfc5212>>.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <<https://www.rfc-editor.org/info/rfc5305>>.
- [RFC5307] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 5307, DOI 10.17487/RFC5307, October 2008, <<https://www.rfc-editor.org/info/rfc5307>>.
- [RFC5441] Vasseur, JP., Ed., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", RFC 5441, DOI 10.17487/RFC5441, April 2009, <<https://www.rfc-editor.org/info/rfc5441>>.

- [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>>.
- [RFC7149] Boucadair, M. and C. Jacquenet, "Software-Defined Networking: A Perspective from within a Service Provider Environment", RFC 7149, DOI 10.17487/RFC7149, March 2014, <<https://www.rfc-editor.org/info/rfc7149>>.
- [RFC7399] Farrel, A. and D. King, "Unanswered Questions in the Path Computation Element Architecture", RFC 7399, DOI 10.17487/RFC7399, October 2014, <<https://www.rfc-editor.org/info/rfc7399>>.
- [RFC7491] King, D. and A. Farrel, "A PCE-Based Architecture for Application-Based Network Operations", RFC 7491, DOI 10.17487/RFC7491, March 2015, <<https://www.rfc-editor.org/info/rfc7491>>.
- [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <<https://www.rfc-editor.org/info/rfc7752>>.
- [RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <<https://www.rfc-editor.org/info/rfc8040>>.
- [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>>.
- [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>>.
- [RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <<https://www.rfc-editor.org/info/rfc8283>>.
- [I-D.ietf-pce-stateful-hpce]
Dhody, D., Lee, Y., Ceccarelli, D., Shin, J., King, D., and O. Dios, "Hierarchical Stateful Path Computation Element (PCE).", draft-ietf-pce-stateful-hpce-04 (work in progress), March 2018.
- [I-D.ietf-teas-actn-requirements]
Lee, Y., Ceccarelli, D., Miyasaka, T., Shin, J., and K. Lee, "Requirements for Abstraction and Control of TE Networks", draft-ietf-teas-actn-requirements-09 (work in progress), March 2018.
- [I-D.ietf-teas-actn-info-model]
Lee, Y., Belotti, S., Dhody, D., Ceccarelli, D., and B. Yoon, "Information Model for Abstraction and Control of TE Networks (ACTN)", draft-ietf-teas-actn-info-model-09 (work in progress), June 2018.
- [I-D.ietf-pce-inter-area-as-applicability]
King, D., Meuric, J., Dugeon, O., Zhao, Q., Dhody, D., and O. Dios, "Applicability of the Path Computation Element to Inter-Area and Inter-AS MPLS and GMPLS Traffic Engineering", draft-ietf-pce-inter-area-as-applicability-06 (work in progress), July 2016.
- [I-D.dhodylee-pce-pcep-ls]
Dhody, D., Lee, Y., and D. Ceccarelli, "PCEP Extension for Distribution of Link-State and TE Information.", draft-dhodylee-pce-pcep-ls-10 (work in progress), March 2018.
- [I-D.lee-pce-pcep-ls-optical]
Lee, Y., zhenghaomian@huawei.com, z., Ceccarelli, D., weiw@bupt.edu.cn, w., Park, P., and B. Yoon, "PCEP Extension for Distribution of Link-State and TE information for Optical Networks", draft-lee-pce-pcep-ls-optical-04 (work in progress), February 2018.

[I-D.leedhody-pce-vn-association]

Lee, Y., Dhody, D., Zhang, X., and D. Ceccarelli, "PCEP Extensions for Establishing Relationships Between Sets of LSPs and Virtual Networks", draft-leedhody-pce-vn-association-04 (work in progress), February 2018.

[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-03 (work in progress), April 2018.

[I-D.ietf-pce-association-policy]

Dhody, D., Sivabalan, S., Litkowski, S., Tantsura, J., and J. Hardwick, "Path Computation Element communication Protocol extension for associating Policies and LSPs", draft-ietf-pce-association-policy-02 (work in progress), February 2018.

[I-D.lee-pce-lsp-stitching-hpce]

Lee, Y., Dhody, D., and D. Ceccarelli, "PCEP Extensions for Stitching LSPs in Hierarchical Stateful PCE Model", draft-lee-pce-lsp-stitching-hpce-01 (work in progress), December 2017.

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

PCE Working Group
Internet-Draft
Intended status: Informational
Expires: November 17, 2019

D. Dhody
Y. Lee
Huawei Technologies
D. Ceccarelli
Ericsson
May 16, 2019

Applicability of the Path Computation Element (PCE) to the Abstraction
and Control of TE Networks (ACTN)
draft-ietf-pce-applicability-actn-12

Abstract

Abstraction and Control of TE Networks (ACTN) refers to the set of virtual network (VN) operations needed to orchestrate, control and manage large-scale multi-domain TE networks so as to facilitate network programmability, automation, efficient resource sharing, and end-to-end virtual service aware connectivity and network function virtualization services.

The Path Computation Element (PCE) is a component, application, or network node that is capable of computing a network path or route based on a network graph and applying computational constraints. The PCE serves requests from Path Computation Clients (PCCs) that communicate with it over a local API or using the Path Computation Element Communication Protocol (PCEP).

This document examines the applicability of PCE to the ACTN framework.

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 November 17, 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 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

1. Introduction	2
2. Background Information	3
2.1. Path Computation Element (PCE)	3
2.1.1. Role of PCE in SDN	4
2.1.2. PCE in Multi-domain and Multi-layer Deployments	4
2.1.3. Relationship to PCE Based Central Control	5
2.2. Abstraction and Control of TE Networks (ACTN)	5
3. Architectural Considerations	7
3.1. Multi-Domain Coordination via Hierarchy	7
3.2. Abstraction	8
3.3. Customer Mapping	9
3.4. Virtual Service Coordination	10
4. Interface Considerations	10
5. Realizing ACTN with PCE (and PCEP)	11
6. IANA Considerations	15
7. Security Considerations	15
8. Acknowledgments	16
9. References	16
9.1. Normative References	16
9.2. Informative References	17
Appendix A. Additional Information	21
Authors' Addresses	21

1. Introduction

Abstraction and Control of TE Networks (ACTN) [RFC8453] refers to the set of virtual network (VN) operations needed to orchestrate, control and manage large-scale multi-domain TE networks so as to facilitate network programmability, automation, efficient resource sharing, and end-to-end virtual service aware connectivity and network function virtualization services.

The Path Computation Element (PCE) [RFC4655] is a component, application, or network node that is capable of computing a network path or route based on a network graph and applying computational constraints. The PCE serves requests from Path Computation Clients (PCCs) that communicate with it over a local API or using the Path Computation Element Communication Protocol (PCEP).

This document examines the PCE and ACTN architecture and describes how PCE architecture is applicable to ACTN. It also lists the PCEP extensions that are needed to use PCEP as an ACTN interface. This document also identifies any gaps in PCEP, that exist at the time of publication of this document.

Further, ACTN, stateful H-PCE [I-D.ietf-pce-stateful-hpce], and PCE as a central controller (PCECC) [RFC8283] are based on the same basic hierarchy framework and thus compatible with each other.

2. Background Information

2.1. Path Computation Element (PCE)

The Path Computation Element Communication Protocol (PCEP) [RFC5440] provides mechanisms for Path Computation Clients (PCCs) to request a Path Computation Element (PCE) [RFC4655] to perform path computations.

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.

A stateful PCE [RFC8231] 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 a PCC to delegate control of specific LSPs to a new PCE.

2.1.1. Role of PCE in SDN

Software-Defined Networking (SDN) [RFC7149] refers to a separation between the control elements and the forwarding components so that software running in a centralized system called a controller, can act to program the devices in the network to behave in specific ways. A required element in an SDN architecture is a component that plans how the network resources will be used and how the devices will be programmed. It is possible to view this component as performing specific computations to place flows within the network given knowledge of the availability of network resources, how other forwarding devices are programmed, and the way that other flows are routed. It is concluded in [RFC7399], that this is the same function that a PCE might offer in a network operated using a dynamic control plane. This is the function and purpose of a PCE, and the way that a PCE integrates into a wider network control system including SDN is presented in Application-Based Network Operation (ABNO) [RFC7491].

2.1.2. PCE in Multi-domain and Multi-layer Deployments

Computing paths across large multi-domain environments requires special computational components and cooperation between entities in different domains capable of complex path computation. The PCE provides an architecture and a set of functional components to address this problem space. A PCE may be used to compute end-to-end paths across multi-domain environments using a per-domain path computation technique [RFC5152]. The Backward Recursive PCE based path computation (BRPC) mechanism [RFC5441] defines a PCE-based path computation procedure to compute inter-domain constrained MPLS and GMPLS TE networks. However, per-domain technique assumes that the sequence of domains to be crossed from source to destination is known, either fixed by the network operator or obtained by other means. BRPC can work best with a known domain sequence, and it will also work nicely with a small set of interconnected domains. However, it doesn't work well for is a large set of interconnected domains.

[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) when the

domain sequence is not known. Within the Hierarchical PCE (H-PCE) architecture, 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.

[I-D.ietf-pce-stateful-hpce] state the considerations for stateful PCEs 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.

[RFC5623] describes a framework for applying the PCE-based architecture to inter-layer to (G)MPLS TE. It provides suggestions for the deployment of PCE in support of multi-layer networks. It also describes the relationship between PCE and a functional component in charge of the control and management of the Virtual Network Topology (VNT) [RFC5212], called the VNT Manager (VNTM).

2.1.3. Relationship to PCE Based Central Control

[RFC8283] introduces the architecture for PCE as a central controller (PCECC), it further examines the motivations and applicability for PCEP as a southbound interface, and introduces the implications for the protocol. Section 2.1.3 of [RFC8283] describe a hierarchy of PCE-based controller as per the Hierarchy of PCE framework defined in [RFC6805].

2.2. Abstraction and Control of TE Networks (ACTN)

[RFC8453] describes the high-level ACTN requirements and the architecture model for ACTN including the entities Customer Network Controller (CNC), Multi-domain Service Coordinator (MDSC), and Provisioning Network Controller (PNC) and their interfaces.

The ACTN reference architecture is shown in Figure 1 which is reproduced here from [RFC8453] for convenience. [RFC8453] remains the definitive reference for the ACTN architecture. As depicted in Figure 1, the ACTN architecture identifies a three-tier hierarchy.

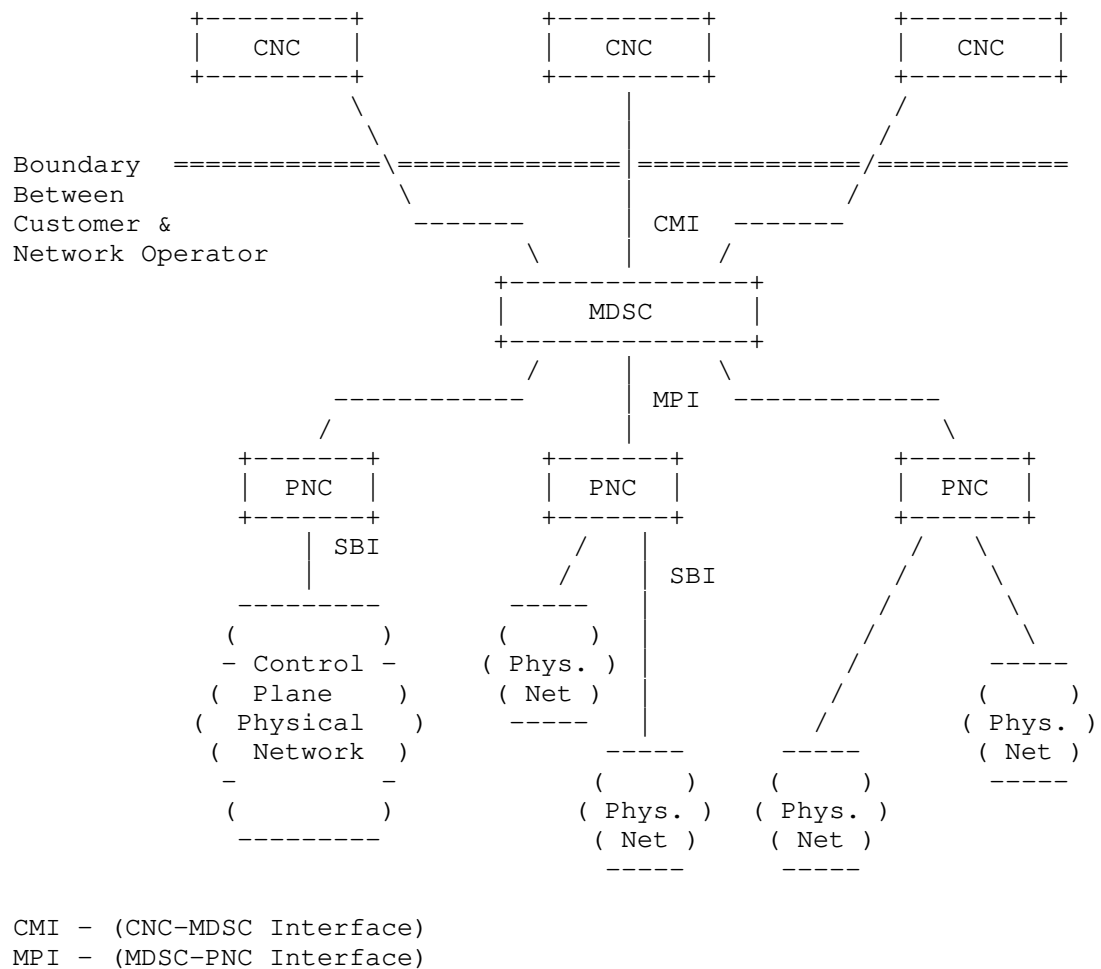


Figure 1: ACTN Hierarchy

There are two interfaces with respect to the MDSC: one north of the MDSC (the CNC-MDSC Interface : CMI), and one south (the MDSC-PNC Interface : MPI). A hierarchy of MDSCs is possible with a recursive MPI interface.

[RFC8454] provides an information model for ACTN interfaces.

3. Architectural Considerations

The ACTN architecture [RFC8453] is based on hierarchy and recursiveness of controllers. It defines three types of controllers (depending on the functionalities they implement). The main functionalities are -

- o Multi-domain coordination
- o Abstraction
- o Customer mapping/translation
- o Virtual service coordination

Section 3 of [RFC8453] describes these functions.

It should be noted that this document lists all possible ways in which PCE could be used for each of the above functions, but all functions are not required to be implemented via PCE. Similarly, this document presents the ways in which PCEP could be used as the communications medium between functional components. Operators may choose to use the PCEP for multi-domain coordination via stateful H-PCE, but alternatively use Network Configuration Protocol (NETCONF) [RFC6241], RESTCONF [RFC8040], or BGP - Link State (BGP-LS) [RFC7752] to get access to the topology and support abstraction function.

3.1. Multi-Domain Coordination via Hierarchy

With the definition of domain being "everything that is under the control of the single logical controller", as per [RFC8453], it is needed to have a control entity that oversees the specific aspects of the different domains and to build a single abstracted end-to-end network topology in order to coordinate end-to-end path computation and path/service provisioning.

The MDSC in ACTN framework realizes this function by coordinating the per-domain PNCs in a hierarchy of controllers. It also needs to detach from the underlying network technology and express customer concerns by business needs.

[RFC6805] and [I-D.ietf-pce-stateful-hpce] describe a hierarchy of PCEs with the Parent PCE coordinating multi-domain path computation function between Child PCEs. It is easy to see how these principles align, and thus how the stateful H-PCE architecture can be used to realize ACTN.

The per domain stitched LSP in the Hierarchical stateful PCE architecture, described in Section 3.3.1 of [I-D.ietf-pce-stateful-hpce] is well suited for multi-domain coordination function. This includes domain sequence selection; End-to-End (E2E) path computation; Controller (PCE) initiated path setup and reporting. This is also applicable to multi-layer coordination in case of IP+optical networks.

[I-D.litkowski-pce-state-sync] describes the procedures to allow a stateful communication between PCEs for various use-cases. The procedures and extensions are also applicable to Child and Parent PCE communication and thus useful for ACTN as well.

3.2. Abstraction

To realize ACTN, an abstracted view of the underlying network resources needs to be built. This includes global network-wide abstracted topology based on the underlying network resources of each domain. This also includes abstract topology created as per the customer service connectivity requests and represented as a VN slice allocated to each customer.

In order to compute and provide optimal paths, PCEs require an accurate and timely Traffic Engineering Database (TED). Traditionally this TED has been obtained from a link state (LS) routing protocol supporting traffic engineering extensions. PCE may construct its TED by participating in the IGP ([RFC3630] and [RFC5305] for MPLS-TE; [RFC4203] and [RFC5307] for GMPLS). An alternative is offered by BGP-LS [RFC7752].

In case of H-PCE [RFC6805], the Parent PCE needs to build the domain topology map of the child domains and their interconnectivity. [RFC6805] and [I-D.ietf-pce-inter-area-as-applicability] suggest that BGP-LS could be used as a "northbound" TE advertisement from the Child PCE to the Parent PCE.

[I-D.dhodylee-pce-pcep-ls] proposes another approach for learning and maintaining the Link-State and TE information as an alternative to IGPs and BGP flooding, using PCEP itself. The Child PCE can use this mechanism to transport Link-State and TE information from Child PCE to a Parent PCE using PCEP.

In ACTN, there is a need to control the level of abstraction based on the deployment scenario and business relationship between the controllers. The mechanism used to disseminate information from PNC (Child PCE) to MDSC (Parent PCE) should support abstraction. [RFC8453] describes a few alternative approaches of abstraction. The resulting abstracted topology can be encoded using the PCEP-LS

mechanisms [I-D.dhodylee-pce-pcep-ls] and its optical network extension [I-D.lee-pce-pcep-ls-optical]. PCEP-LS is an attractive option when the operator would wish to have a single control plane protocol (PCEP) to achieve ACTN functions.

[RFC8453] discusses two ways to build abstract topology from an MDSC standpoint with interaction with PNCs. The primary method is called automatic generation of abstract topology by configuration. With this method, automatic generation is based on the abstraction/summarization of the whole domain by the PNC and its advertisement on the MPI. The secondary method is called on-demand generation of supplementary topology via Path Compute Request/Reply. This method may be needed to obtain further complementary information such as potential connectivity from Child PCEs in order to facilitate an end-to-end path provisioning. PCEP is well suited to support both methods.

3.3. Customer Mapping

In ACTN, there is a need to map customer virtual network (VN) requirements into a network provisioning request to the PNC. That is, the customer requests/commands are mapped by the MDSC into network provisioning requests that can be sent to the PNC. Specifically, the MDSC provides mapping and translation of a customer's service request into a set of parameters that are specific to a network type and technology such that network configuration process is made possible.

[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. As described in [I-D.ietf-pce-stateful-hpce], for inter-domain LSP in Hierarchical PCE architecture, the initiation operations can be carried out at the Parent PCE. In which case, after Parent PCE finishes the E2E path computation, it can send the PCInitiate message to the Child PCE, the Child PCE further propagates the initiate request to the Label Switching Router (LSR). The customer request is received by the MDSC (Parent PCE) and based on the business logic, global abstracted topology, network conditions and local policy, the MDSC (Parent PCE) translates this into per domain LSP initiation request that a PNC (Child PCE) can understand and act on. This can be done via the PCInitiate message.

PCEP extensions for associating opaque policy between PCEP peer [I-D.ietf-pce-association-policy] can be used.

3.4. Virtual Service Coordination

Virtual service coordination function in ACTN incorporates customer service-related information into the virtual network service operations in order to seamlessly operate virtual networks while meeting customer's service requirements.

[I-D.leedhody-pce-vn-association] describes the need for associating a set of LSPs with a VN "construct" to facilitate VN operations in PCE architecture. This association allows the PCEs to identify which LSPs belong to a certain VN.

This association based on VN is useful for various optimizations at the VN level which can be applied to all the LSPs that are part of the VN slice. During path computation, the impact of a path for an LSP is compared against the paths of other LSPs in the VN. This is to make sure that the overall optimization and SLA of the VN rather than of a single LSP. Similarly, during re-optimization, advanced path computation algorithm and optimization technique can be considered for all the LSPs belonging to a VN/customer and optimize them all together.

4. Interface Considerations

As per [RFC8453], to allow virtualization and multi-domain coordination, the network has to provide open, programmable interfaces, in which customer applications can create, replace and modify virtual network resources and services in an interactive, flexible and dynamic fashion while having no impact on other customers. The two ACTN interfaces are -

- o The CNC-MDSC Interface (CMI) is an interface between a Customer Network Controller and a Multi-Domain Service Coordinator. It requests the creation of the network resources, topology or services for the applications. The MDSC may also report potential network topology availability if queried for current capability from the Customer Network Controller.
- o The MDSC-PNC Interface (MPI) is an interface between a Multi-Domain Service Coordinator and a Provisioning Network Controller. It communicates the creation request, if required, of new connectivity of bandwidth changes in the physical network, via the PNC. In multi-domain environments, the MDSC needs to establish multiple MPIs, one for each PNC, as there are multiple PNCs responsible for its domain control.

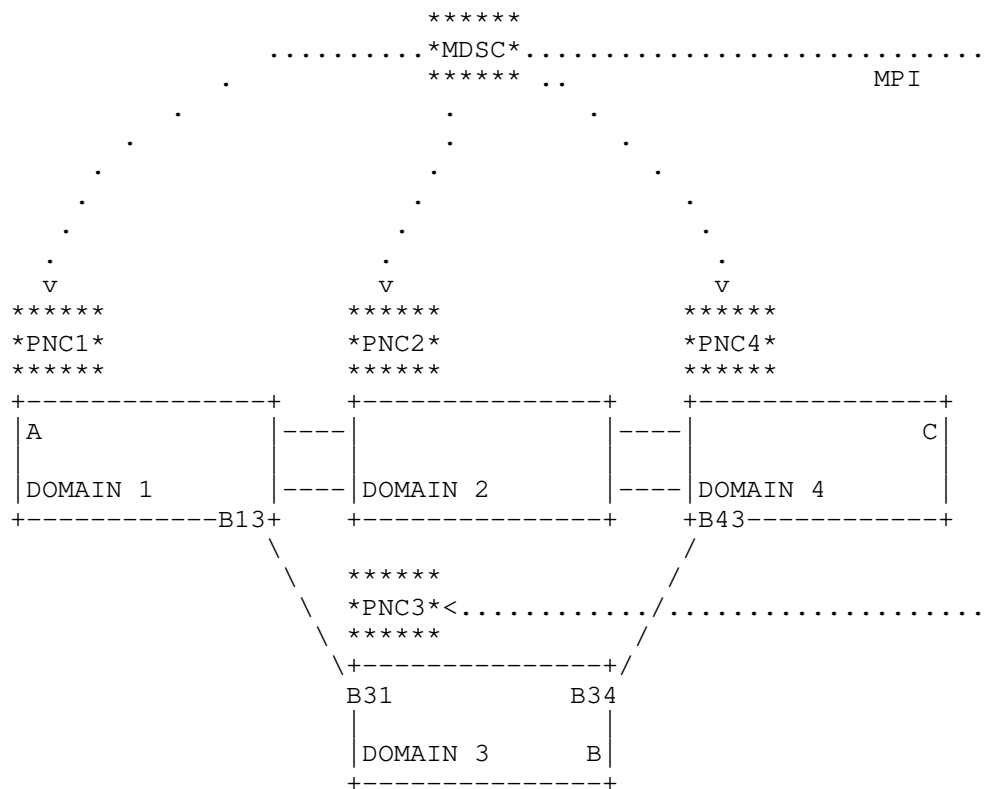
In the case of hierarchy MDSCs, the MPI is applied recursively. From an abstraction point of view, the top level MDSC which interfaces the

CNC operates on a higher level of abstraction (i.e., less granular level) than the lower level MSDCs.

PCEP is especially suitable on the MPI as it meets the requirement and the functions as set out in the ACTN framework [RFC8453]. Its recursive nature is well suited via the multi-level hierarchy of PCE. PCEP can also be applied to the CMI as the CNC can be a path computation client while the MDSC can be a path computation server. Section 5 describes how PCE and PCEP could help realize ACTN on the MPI.

5. Realizing ACTN with PCE (and PCEP)

As per the example in Figure 2, there are 4 domains, each with its own PNC and an MDSC on top. The PNC and MDSC need PCE as an important function. The PNC (or Child PCE) already uses PCEP to communicate to the network device. It can utilize the PCEP as the MPI to communicate between controllers too.



MDSC -> Parent PCE
PNC -> Child PCE
MPI -> PCEP

Figure 2: ACTN with PCE

- o Building Domain Topology at MDSC: PNC (or Child PCE) needs to have the TED to compute path in its domain. As described in Section 3.2, it can learn the topology via IGP or BGP-LS. PCEP-LS is also a proposed mechanism to carry link state and traffic engineering information within PCEP. A mechanism to carry abstracted topology while hiding technology specific information between PNC and MDSC is described in [I-D.dhodylee-pce-pcep-ls]. At the end of this step the MDSC (or Parent PCE) has the abstracted topology from each of its PNC (or Child PCE). This could be as simple as a domain topology map as described in [RFC6805] or it can have full topology information of all domains. The latter is not scalable and thus an abstracted topology of each

domain interconnected by inter-domain links is the most common case.

- * **Topology Change:** When the PNC learns of any topology change, the PNC needs to decide if the change needs to be notified to the MDSC. This is dependent on the level of abstraction between the MDSC and the PNC.
- o **VN Instantiate:** When an MDSC is requested to instantiate a VN, the minimal information that is required would be a VN identifier and a set of end points. Various path computation, setup constraints and objective functions may also be provided. In PCE terms, a VN Instantiate can be considered as a set of paths belonging to the same VN. As described in Section 3.4 and [I-D.leedhody-pce-vn-association] the VN association can help in identifying the set of paths that belong to a VN. The rest of the information like the endpoints, constraints and objective function (OF) is already defined in PCEP in terms of a single path.
- * **Path Computation:** As per the example in Figure 2, the VN instantiate requires two end to end paths between (A in Domain 1 to B in Domain 3) and (A in Domain 1 to C in Domain 4). The MDSC (or Parent PCE) triggers the end to end path computation for these two paths. MDSC can do path computation based on the abstracted domain topology that it already has or it may use the H-PCE procedures (Section 3.1) using the PCReq and PCRep messages to get the end to end path with the help of the Child PCEs (PNC). Either way, the resultant E2E paths may be broken into per-domain paths.
- * **A-B:** (A-B13,B13-B31,B31-B)
- * **A-C:** (A-B13,B13-B31,B31-B34,B34-B43,B43-C)
- * **Per Domain Path Instantiation:** Based on the above path computation, MDSC can issue the path instantiation request to each PNC via PCInitiate message (see [I-D.ietf-pce-stateful-hpce] and [I-D.leedhody-pce-vn-association]). A suitable stitching mechanism would be used to stitch these per domain LSPs. One such mechanism is described in [I-D.dugeon-pce-stateful-interdomain], where PCEP is extended to support stitching in stateful H-PCE context.
- * **Per Domain Path Report:** Each PNC should report the status of the per-domain LSP to the MDSC via PCRpt message, as per the Hierarchy of stateful PCE ([I-D.ietf-pce-stateful-hpce]). The

status of the end to end LSP (A-B and A-C) is made up when all the per domain LSP are reported up by the PNCs.

- * **Delegation:** It is suggested that the per domain LSPs are delegated to respective PNC, so that they can control the path and attributes based on each domain network conditions.
- * **State Synchronization:** The state needs to be synchronized between the Parent PCE and Child PCE. The mechanism described in [I-D.litkowski-pce-state-sync] can be used.
- o **VN Modify:** MDSC is requested to modify a VN, for example the bandwidth for VN is increased. This may trigger path computation at MDSC as described in the previous step and can trigger an update to existing per-intra-domain path (via PCUpd message) or creation (or deletion) of a per-domain path (via PCInitiate message). As described in [I-D.ietf-pce-stateful-hpce], this should be done in make-before-break fashion.
- o **VN Delete:** MDSC is requested to delete a VN, in this case, based on the E2E paths and the resulting per-domain paths need to be removed (via PCInitiate message).
- o **VN Update (based on network changes):** Any change in the per-domain LSP is reported to the MDSC (via PCRpt message) as per [I-D.ietf-pce-stateful-hpce]. This may result in changes in the E2E path or VN status. This may also trigger a re-optimization leading to a new per-domain path, update to existing path, or deletion of the path.
- o **VN Protection:** The VN protection/restoration requirements, need to be applied to each E2E path as well as each per domain path. The MDSC needs to play a crucial role in coordinating the right protection/restoration policy across each PNC. The existing protection/restoration mechanism of PCEP can be applied on each path.
- o **In case a PNC generates an abstract topology towards the MDSC,** the PCInitiate/PCUpd messages from the MDSC to a PNC will contain a path with abstract nodes and links. A PNC would need to take that as an input for path computation to get a path with physical nodes and links. Similarly, a PNC would convert the path received from the device (with physical nodes and links) into an abstract path (based on the abstract topology generated before with abstract nodes and links) and report it to the MDSC.

6. IANA Considerations

This document makes no requests for IANA action.

7. Security Considerations

Various security considerations for PCEP are described in [RFC5440] and [RFC8253]. Security considerations as stated in Section 10.1, Section 10.6, and Section 10.7 of [RFC5440] continue to apply on PCEP when used as ACTN interface. Further, this document lists various extensions of PCEP that are applicable, each of them specify various security considerations which continue to apply here.

The ACTN framework described in [RFC8453] defines key components and interfaces for managed traffic engineered networks. It also lists various security considerations such as request and control of resources, confidentiality of the information, and availability of function which should be taken into consideration.

As per [RFC8453], securing the request and control of resources, confidentiality of the information, and availability of function should all be critical security considerations when deploying and operating ACTN platforms. From a security and reliability perspective, ACTN may encounter many risks such as malicious attack and rogue elements attempting to connect to various ACTN components (with PCE being one of them). Furthermore, some ACTN components represent a single point of failure and threat vector and must also manage policy conflicts and eavesdropping of communication between different ACTN components. [RFC8453] further states that all protocols used to realize the ACTN framework should have rich security features, and customer, application and network data should be stored in encrypted data stores. When PCEP is used as an ACTN interface, the security of PCEP provided by Transport Layer Security (TLS) [RFC8253], as per the recommendations and best current practices in [RFC7525] (unless explicitly set aside in [RFC8253]), is used.

As per [RFC8453], regarding the MPI, a PKI- based mechanism is suggested, such as building a TLS or HTTPS connection between the MDSC and PNCs, to ensure trust between the physical network layer control components and the MDSC. Which MDSC the PNC exports topology information to, and the level of detail (full or abstracted), should also be authenticated, and specific access restrictions and topology views should be configurable and/or policy based. When PCEP is used in MPI, the security functions as per [RFC8253] are used to fulfill these requirements.

As per [RFC8453], regarding the CMI, suitable authentication and authorization of each CNC connecting to the MDSC will be required. If PCEP is used in CMI, the security functions as per [RFC8253] can be used to support peer authentication, message encryption, and integrity checks.

8. Acknowledgments

The authors would like to thank Jonathan Hardwick for the inspiration behind this document. Further thanks to Avantika for her comments with suggested text.

Thanks to Adrian Farrel and Daniel King for their substantial reviews.

Thanks to Yingzhen Qu for RTGDIR review.

Thanks to Rifaat Shekh-Yusef for SECDIR review.

Thanks to Meral Shirazipour for GENART review.

Thanks to Roman Danyliw and Barry Leiba for IESG review comments.

Thanks to Deborah Brungard for being the responsible AD.

9. References

9.1. Normative References

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

- [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>>.

9.2. Informative References

- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <<https://www.rfc-editor.org/info/rfc3630>>.
- [RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <<https://www.rfc-editor.org/info/rfc4203>>.
- [RFC5152] Vasseur, JP., Ed., Ayyangar, A., Ed., and R. Zhang, "A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)", RFC 5152, DOI 10.17487/RFC5152, February 2008, <<https://www.rfc-editor.org/info/rfc5152>>.
- [RFC5212] Shiimoto, K., Papadimitriou, D., Le Roux, JL., Vigoureux, M., and D. Brungard, "Requirements for GMPLS-Based Multi-Region and Multi-Layer Networks (MRN/MLN)", RFC 5212, DOI 10.17487/RFC5212, July 2008, <<https://www.rfc-editor.org/info/rfc5212>>.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <<https://www.rfc-editor.org/info/rfc5305>>.
- [RFC5307] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 5307, DOI 10.17487/RFC5307, October 2008, <<https://www.rfc-editor.org/info/rfc5307>>.
- [RFC5441] Vasseur, JP., Ed., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", RFC 5441, DOI 10.17487/RFC5441, April 2009, <<https://www.rfc-editor.org/info/rfc5441>>.

- [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>>.
- [RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <<https://www.rfc-editor.org/info/rfc6241>>.
- [RFC7149] Boucadair, M. and C. Jacquenet, "Software-Defined Networking: A Perspective from within a Service Provider Environment", RFC 7149, DOI 10.17487/RFC7149, March 2014, <<https://www.rfc-editor.org/info/rfc7149>>.
- [RFC7399] Farrel, A. and D. King, "Unanswered Questions in the Path Computation Element Architecture", RFC 7399, DOI 10.17487/RFC7399, October 2014, <<https://www.rfc-editor.org/info/rfc7399>>.
- [RFC7491] King, D. and A. Farrel, "A PCE-Based Architecture for Application-Based Network Operations", RFC 7491, DOI 10.17487/RFC7491, March 2015, <<https://www.rfc-editor.org/info/rfc7491>>.
- [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>>.
- [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <<https://www.rfc-editor.org/info/rfc7752>>.
- [RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <<https://www.rfc-editor.org/info/rfc8040>>.
- [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>>.

- [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>>.
- [RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <<https://www.rfc-editor.org/info/rfc8283>>.
- [RFC8454] Lee, Y., Belotti, S., Dhody, D., Ceccarelli, D., and B. Yoon, "Information Model for Abstraction and Control of TE Networks (ACTN)", RFC 8454, DOI 10.17487/RFC8454, September 2018, <<https://www.rfc-editor.org/info/rfc8454>>.
- [I-D.ietf-pce-stateful-hpce]
Dhody, D., Lee, Y., Ceccarelli, D., Shin, J., King, D., and O. Dios, "Hierarchical Stateful Path Computation Element (PCE).", draft-ietf-pce-stateful-hpce-07 (work in progress), April 2019.
- [I-D.ietf-pce-inter-area-as-applicability]
King, D. and H. Zheng, "Applicability of the Path Computation Element to Inter-Area and Inter-AS MPLS and GMPLS Traffic Engineering", draft-ietf-pce-inter-area-as-applicability-07 (work in progress), December 2018.
- [I-D.dhodylee-pce-pcep-ls]
Dhody, D., Lee, Y., and D. Ceccarelli, "PCEP Extension for Distribution of Link-State and TE Information.", draft-dhodylee-pce-pcep-ls-13 (work in progress), February 2019.

- [I-D.lee-pce-pcep-ls-optical]
Lee, Y., Zheng, H., Ceccarelli, D., weiw@bupt.edu.cn, w., Park, P., and B. Yoon, "PCEP Extension for Distribution of Link-State and TE information for Optical Networks", draft-lee-pce-pcep-ls-optical-07 (work in progress), March 2019.
- [I-D.leedhody-pce-vn-association]
Lee, Y., Zhang, X., and D. Ceccarelli, "PCEP Extensions for Establishing Relationships Between Sets of LSPs and Virtual Networks", draft-leedhody-pce-vn-association-07 (work in progress), February 2019.
- [I-D.litkowski-pce-state-sync]
Litkowski, S., Sivabalan, S., Li, C., and H. Zheng, "Inter Stateful Path Computation Element (PCE) Communication Procedures.", draft-litkowski-pce-state-sync-05 (work in progress), March 2019.
- [I-D.ietf-pce-association-policy]
Litkowski, S., Sivabalan, S., Tantsura, J., Hardwick, J., and M. Negi, "Path Computation Element communication Protocol extension for associating Policies and LSPs", draft-ietf-pce-association-policy-05 (work in progress), February 2019.
- [I-D.dugeon-pce-stateful-interdomain]
Dugeon, O., Meuric, J., Lee, Y., and D. Ceccarelli, "PCEP Extension for Stateful Inter-Domain Tunnels", draft-dugeon-pce-stateful-interdomain-02 (work in progress), March 2019.
- [EXP]
Casellas, R., Vilalta, R., Martinez, R., Munoz, R., Zheng, H., and Y. Lee, "Experimental Validation of the ACTN architecture for flexi-grid optical networks using Active Stateful Hierarchical PCEs", 19th International Conference on Transparent Optical Networks (ICTON) , July 2017, <<http://www.cttc.es/publication/experimental-validation-of-the-actn-architecture-for-flexi-grid-optical-networks-using-active-stateful-hierarchical-pces/>>.

Appendix A. Additional Information

In the paper [EXP], the application of the ACTN architecture is presented to demonstrate the control of a multi-domain flexi-grid optical network, by proposing, adopting and extending -

- o the Hierarchical active stateful PCE architectures and protocols
- o the PCEP protocol to support efficient and incremental link state topological reporting, known as PCEP-LS
- o the per link partitioning of the optical spectrum based on variable-sized allocated frequency slots enabling network sharing and virtualization
- o the use of a model-based interface to dynamically request the instantiation of virtual networks for specific clients / tenants.

The design and the implementation of the testbed are reported in order to validate the approach.

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

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 16, 2019

F. Zhang
Q. Zhao
Huawei
O. Gonzalez de Dios
Telefonica I+D
R. Casellas
CTTC
D. King
Old Dog Consulting
July 15, 2018

Extensions to Path Computation Element Communication Protocol (PCEP) for
Hierarchical Path Computation Elements (PCE)
draft-ietf-pce-hierarchy-extensions-05

Abstract

The Hierarchical Path Computation Element (H-PCE) architecture RFC 6805, provides a mechanism to allow the optimum sequence of domains to be selected, and the optimum end-to-end path to be derived through the use of a hierarchical relationship between domains.

This document defines the Path Computation Element Protocol (PCEP) extensions for the purpose of implementing necessary Hierarchical PCE procedures and protocol extensions.

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 January 16, 2019.

Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 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

1. Introduction	3
1.1. Scope	4
1.2. Terminology	5
1.3. Requirements Language	5
2. Requirements for H-PCE	5
2.1. Path Computation Request	5
2.1.1. Qualification of PCEP Requests	6
2.1.2. Multi-domain Objective Functions	6
2.1.3. Multi-domain Metrics	7
2.2. Parent PCE Capability Advertisement	7
2.3. PCE Domain Discovery	7
2.4. Domain Diversity	8
3. PCEP Extensions	8
3.1. OPEN object	8
3.1.1. H-PCE capability TLV	8
3.1.2. Domain-ID TLV	9
3.2. RP object	10
3.2.1. H-PCE-FLAG TLV	10
3.2.2. Domain-ID TLV	11
3.3. Objective Functions	11
3.3.1. OF Codes	11
3.3.2. OF Object	13
3.4. Metric Object	13
3.5. SVEC Object	14
3.6. PCEP-ERROR object	14
3.6.1. Hierarchy PCE Error-Type	14
3.7. NO-PATH Object	15
4. H-PCE Procedures	15
4.1. OPEN Procedure between Child PCE and Parent PCE	15
4.2. Procedure to obtain Domain Sequence	16
5. Error Handling	16

6.	Manageability Considerations	17
6.1.	Control of Function and Policy	17
6.1.1.	Child PCE	17
6.1.2.	Parent PCE	18
6.1.3.	Policy Control	18
6.2.	Information and Data Models	18
6.3.	Liveness Detection and Monitoring	18
6.4.	Verify Correct Operations	19
6.5.	Requirements On Other Protocols	19
6.6.	Impact On Network Operations	19
7.	IANA Considerations	19
7.1.	PCEP TLV Type Indicators	19
7.2.	H-PCE-CAPABILITY TLV Flags	20
7.3.	Domain-ID TLV Domain type	20
7.4.	H-PCE-FLAG TLV Flags	20
7.5.	OF Codes	21
7.6.	METRIC Types	21
7.7.	New PCEP Error-Types and Values	22
7.8.	New NO-PATH-VECTOR TLV Bit Flag	22
7.9.	SVEC Flag	22
8.	Security Considerations	23
9.	Implementation Status	23
9.1.	Inter-layer traffic engineering with H-PCE	23
9.2.	Telefonica Netphony (Open Source PCE)	25
9.3.	Implementation 3: H-PCE Proof of Concept developed by Huawei	26
10.	Contributing Authors	27
11.	References	27
11.1.	Normative References	27
11.2.	Informative References	28
	Authors' Addresses	30

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.

The capability to compute the routes of end-to-end inter-domain MPLS Traffic Engineering (MPLS-TE) and GMPLS Label Switched Paths (LSPs) is expressed as requirements in [RFC4105] and [RFC4216]. This capability may be realized by a PCE [RFC4655]. The methods for establishing and controlling inter-domain MPLS-TE and GMPLS LSPs are documented in [RFC4726].

[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 architecture, the parent PCE is used to compute a multi-domain path based on the domain connectivity information. A child PCE may be responsible for a single domain or multiple domains, it is used to compute the intra-domain path based on its own domain topology information.

The H-PCE end-to-end domain path computation procedure is described below:

- o A path computation client (PCC) sends the inter-domain path computation requests to the child PCE responsible for its domain;
- o The child PCE forwards the request to the parent PCE;
- o The parent PCE computes the likely domain paths from the ingress domain to the egress domain;
- o The parent PCE sends the intra-domain path computation requests (between the domain border nodes) to the child PCEs which are responsible for the domains along the domain path;
- o The child PCEs return the intra-domain paths to the parent PCE;
- o The parent PCE constructs the end-to-end inter-domain path based on the intra-domain paths;
- o The parent PCE returns the inter-domain path to the child PCE;
- o The child PCE forwards the inter-domain path to the PCC.

In addition, the parent PCE may be requested to provide only the sequence of domains to a child PCE so that alternative inter-domain path computation procedures, including Per Domain (PD) [RFC5152] and Backwards Recursive Path Computation (BRPC) [RFC5441] may be used.

This document defines the PCEP extensions for the purpose of implementing Hierarchical PCE procedures, which are described in [RFC6805].

1.1. Scope

The following functions are out of scope of this document.

- o Determination of Destination Domain (section 4.5 of [RFC6805])

- * via collection of reachability information from child domain;
 - * via requests to the child PCEs to discover if they contain the destination node;
 - * or any other methods.
- o Parent Traffic Engineering Database (TED) methods (section 4.4 of [RFC6805])
 - o Learning of Domain connectivity and boundary nodes (BN) addresses.
 - o Stateful PCE Operations. (Refer [I-D.ietf-pce-stateful-hpce])

1.2. Terminology

This document uses the terminology defined in [RFC4655], [RFC5440] and the additional terms defined in section 1.4 of [RFC6805].

1.3. 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.

2. Requirements for H-PCE

This section compiles the set of requirements of the PCEP protocol to support the H-PCE architecture and procedures.

[RFC6805] identifies high-level requirements of PCEP extensions required to support the hierarchical PCE model.

2.1. Path Computation Request

The Path Computation Request (PCReq) messages are used by a PCC or PCE to make a path computation request to a PCE. In order to achieve the full functionality of the H-PCE procedures, the PCReq message needs to include:

- o Qualification of PCE Requests;
- o Multi-domain Objective Functions (OF);
- o Multi-domain Metrics.

2.1.1.1. Qualification of PCEP Requests

As described in section 4.8.1 of [RFC6805], the H-PCE architecture introduces new request qualifications, which are:

- o It MUST be possible for a child PCE to indicate that a path computation request sent to a parent PCE should be satisfied by a domain sequence only, that is, not by a full end-to-end path. This allows the child PCE to initiate a per-domain (PD) [RFC5152] or a backward recursive path computation (BRPC) [RFC5441].
- o As stated in [RFC6805], section 4.5, if a PCC knows the egress domain, it can supply this information as the path computation request. It SHOULD be possible to specify the destination domain information in a PCEP request, if it is known.
- o It MAY be possible to indicate that the inter domain path computed by parent PCE should disallow domain re-entry.

2.1.1.2. Multi-domain Objective Functions

For inter-domain path computation, there is one new objective Function which is defined in section 1.3.1 and 4.1 of [RFC6805]:

- o Minimize the number of domains crossed. A domain can be either an Autonomous System (AS) or an Internal Gateway Protocol (IGP) area depending on the type of multi-domain network hierarchical PCE is applied to.

Another objective Function to minimize the number of border nodes is also defined in this document.

During the PCEP session establishment procedure, the parent PCE needs to be capable of indicating the Objective Functions (OF) [RFC5541] capability in the Open message. This capability information may then be announced by child PCEs, and used for selecting the PCE when a PCC wants a path that satisfies one or multiple inter-domain objective functions.

When a PCC requests a PCE to compute an inter-domain path, the PCC needs to be capable of indicating the new objective functions for inter-domain path. Note that a given child PCE may also act as a parent PCE (for some other child PCE).

For the reasons described previously, new OF codes need to be defined for the new inter-domain objective functions. Then the PCE can notify its new inter-domain objective functions to the PCC by carrying them in the OF-list TLV which is carried in the OPEN object.

The PCC can specify which objective function code to use, which is carried in the OF object when requesting a PCE to compute an inter-domain path.

A parent PCE MUST be capable of ensuring homogeneity, across domains, when applying OF codes for strict OF intra-domain requests.

2.1.3. Multi-domain Metrics

For inter-domain path computation, there are several path metrics of interest.

- o Domain count (number of domains crossed);
- o Border Node count.

A PCC may be able to limit the number of domains crossed by applying a limit on these metrics. Details in Section 3.4.

2.2. Parent PCE Capability Advertisement

Parent and child PCE relationships are likely to be configured. However, as mentioned in [RFC6805], it would assist network operators if the child and parent PCEs could indicate their H-PCE capabilities.

During the PCEP session establishment procedure, the child PCE needs to be capable of indicating to the parent PCE whether it requests the parent PCE capability or not. Also, during the PCEP session establishment procedure, the parent PCE needs to be capable of indicating whether its parent capability can be provided or not.

A PCEP Speaker (Parent PCE or Child PCE or PCC) includes the "H-PCE Capability" TLV, described in Section 3.1.1, in the OPEN Object to advertise its support for PCEP extensions for H-PCE Capability.

2.3. PCE Domain Discovery

A PCE domain is a single domain with an associated PCE. Although it is possible for a PCE to manage multiple domains simultaneously. The PCE domain could be an IGP area or AS.

The PCE domain identifiers MAY be provided during the PCEP session establishment procedure.

2.4. Domain Diversity

In a multi-domain environment, Domain Diversity is defined in [RFC6805]. A pair of paths are domain-diverse if they do not traverse any of the same transit domains. Domain diversity may be maximized for a pair of paths by selecting paths that have the smallest number of shared domains. Path computation should facilitate the selection of domain diverse paths as a way to reduce the risk of shared failure and automatically helps to ensure path diversity for most of the route of a pair of LSPs.

The main motivation behind domain diversity is to avoid fate sharing, but it can also be because of some geo-political reasons and commercial relationships that would require domain diversity. for example, a pair of paths should choose different transit Autonomous System (AS) because of some policy considerations.

In case when full domain diversity could not be achieved, it is helpful to minimize the common shared domains. Also it is interesting to note that other scope of diversity (node, link, SRLG etc) can still be applied inside the common shared domains.

3. PCEP Extensions

This section defines PCEP extensions to ([RFC5440]) so as to support the H-PCE procedures.

3.1. OPEN object

Two new TLVs are defined in this document to be carried within an OPEN object. This way, during PCEP session establishment, the H-PCE capability and Domain information can be advertised.

3.1.1. H-PCE capability TLV

The H-PCE-CAPABILITY TLV is an optional TLV associated with the OPEN Object [RFC5440] to exchange H-PCE capability of PCEP speakers.

Its format is shown in the following figure:

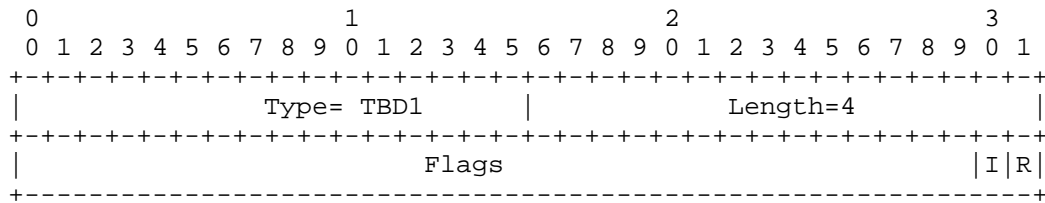


Figure 1: H-PCE-CAPABILITY TLV format

The type of the TLV is TBD1 (to be assigned by IANA) and it has a fixed length of 4 octets.

The value comprises a single field - Flags (32 bits):

R (Parent PCE Request bit): if set, will signal that the child PCE wishes to use the peer PCE as a parent PCE.

I (Parent PCE Indication bit): if set, will signal that the PCE can be used as a parent PCE by the peer PCE.

The inclusion of this TLV in an OPEN object indicate that the H-PCE extensions are supported by the PCEP speaker. The PCC MAY include this TLV to indicate that it understands the H-PCE extensions. The Child PCE MUST include this TLV and set the R flag (and unset the I flag) on the PCEP session towards the Parent PCE. The Parent PCE MUST include this TLV and set the I flag and unset the R flag on the PCEP session towards the child PCE. The parent-child PCEP session is set to be established only when this capability is advertised.

If such capability is not exchanged and the parent PCE receive a "H-PCE path computation request", it MUST send a PCErr message with Error-Type=TBD8 (H-PCE error) and Error-Value=1 (Parent PCE Capability not advertised).

3.1.2. Domain-ID TLV

The Domain-ID TLV when used in OPEN object identify the domain(s) served by the PCE. The child PCE uses this mechanism to inform the domain information to the parent PCE.

The Domain-ID TLV is defined below:

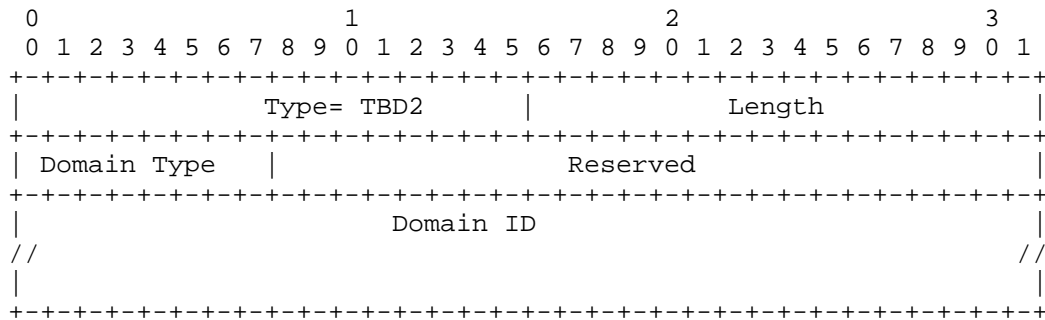


Figure 2: Domain-ID TLV format

The type of the TLV is TBD2 (to be assigned by IANA) and it has a variable Length of the value portion. The value part comprises of -

Domain Type (8 bits): Indicates the domain type. Four types of domain are currently defined:

- * Type=1: the Domain ID field carries a 2-byte AS number. Padded with trailing zeros to a 4-byte boundary.
- * Type=2: the Domain ID field carries a 4-byte AS number.
- * Type=3: the Domain ID field carries an 4-byte OSPF area ID.
- * Type=4: the Domain ID field carries (2-byte Area-Len, variable length IS-IS area ID). Padded with trailing zeros to a 4-byte boundary.

Reserved: Zero at transmission; ignored at receipt.

Domain ID (variable): Indicates an IGP Area ID or AS number. It can be 2 bytes, 4 bytes or variable length depending on the domain identifier used. It is padded with trailing zeros to a 4-byte boundary.

In case a PCE serves more than one domain, multiple Domain-ID TLV is included for each domain it serves.

3.2. RP object

3.2.1. H-PCE-FLAG TLV

The H-PCE-FLAG TLV is an optional TLV associated with the RP Object [RFC5440] to indicate the H-PCE path computation request and options.

Its format is shown in the following figure:

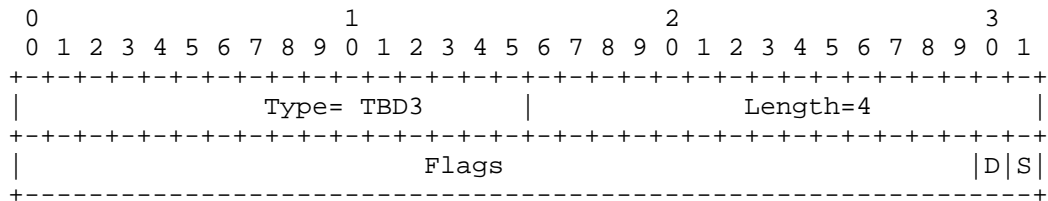


Figure 3: H-PCE-FLAG TLV format

The type of the TLV is TBD3 (to be assigned by IANA) and it has a fixed length of 4 octets.

The value comprises a single field - Flags (32 bits):

S (Domain Sequence bit): if set, will signal that the child PCE wishes to get only the domain sequence in the path computation reply. Refer section 3.7 of [RFC7897] for details.

D (Disallow Domain Re-entry bit): if set, will signal that the computed path does not enter a domain more than once.

3.2.2. Domain-ID TLV

The usage of Domain-ID TLV carried in an OPEN object is used to indicate a (list of) managed domains and is described in Section 3.1.2. This TLV when carried in a RP object, indicates the destination domain ID. If a PCC knows the egress domain, it can supply this information in the PCReq message. The format and procedure of this TLV is defined in Section 3.1.2.

3.3. Objective Functions

3.3.1. OF Codes

[RFC5541] defines a mechanism to specify an objective function that is used by a PCE when it computes a path. Two new objective functions are defined for the H-PCE experiment.

o MTD

* Name: Minimize the number of Transit Domains (MTD)

* Objective Function Code - TBD4 (to be assigned by IANA)

- * Description: Find a path P such that it passes through the least number of transit domains.
- * Objective functions are formulated using the following terminology:
 - + A network comprises a set of N domains $\{D_i, (i=1\dots N)\}$.
 - + A path P passes through K domains $\{D_{pi}, (i=1\dots K)\}$.
 - + Find a path P such that the value of K is minimized.
- o MBN
 - * Name: Minimize the number of border nodes.
 - * Objective Function Code - TBD5 (to be assigned by IANA)
 - * Description: Find a path P such that it passes through the least number of border nodes.
 - * Objective functions are formulated using the following terminology:
 - + A network comprises a set of N nodes $\{N_i, (i=1\dots N)\}$.
 - + A path P is a list of K nodes $\{N_{pi}, (i=1\dots K)\}$.
 - + $B(N)$ is a function that determine if the node is a border node. $B(N_i) = 1$ if N_i is border node; $B(N_k) = 0$ if N_k is not a border node.
 - + The number of border node in a path P is denoted by $B(P)$, where $B(P) = \sum\{B(N_{pi}), (i=1\dots K)\}$.
 - + Find a path P such that $B(P)$ is minimized.

MCTD

- o Name: Minimize the number of Common Transit Domains.
- o Objective Function Code: TBD13
- o Description: Find a set of paths such that it passes through the least number of common transit domains.

3.3.2. OF Object

The OF (Objective Function) object [RFC5541] is carried within a PCReq message so as to indicate the desired/required objective function to be applied by the PCE during path computation. As per section 3.2 of [RFC5541] a single OF object may be included in a path computation request.

The new OF code described in Section 3.3.1 are applicable at the inter-domain level (parent), it is also necessary to specify the OF code that may be applied at the intra-domain (child) path computation level. To accommodate this, the OF-List TLV (described in section 2.1. of [RFC5541]) is included in the OF object as an optional TLV.

OF-List TLV allow encoding of multiple OF codes. When this TLV is included inside the OF object, only the first OF-code in the OF-LIST TLV is considered. The parent PCE MUST use this OF code in the OF object when sending the intra domain path computation request to the child PCE.

If the objective functions defined in this document are unknown/unsupported by a PCE, then the procedure as defined in [RFC5541] is followed.

3.4. Metric Object

The METRIC object is defined in section 7.8 of [RFC5440], comprising metric-value, metric-type (T field) and flags. This document defines the following types for the METRIC object for H-PCE:

- o T=TBDD6: Domain count metric (number of domains crossed);
- o T=TBDD7: Border Node count metric (number of border nodes crossed).

The domain count metric type of the METRIC object encodes the number of domain crossed in the path. The border node count metric type of the METRIC object encodes the number of border nodes in the path.

A PCC or child PCE MAY use these metric in PCReq message an inter-domain path meeting the number of domain or border nodes requirement. As per [RFC5440], in this case, the B bit is set to suggest a bound (a maximum) for the metric that must not be exceeded for the PCC to consider the computed path as acceptable.

A PCC or child PCE MAY also use this metric to ask the PCE to optimize the metric during inter-domain path computation. In this case, the B flag is cleared.

The Parent PCE MAY use these metric in a PCRep message along with a NO-PATH object in the case where the PCE cannot compute a path meeting this constraint. A PCE MAY also use this metric to send the computed end to end metric in a reply message.

3.5. SVEC Object

[RFC5440] defines SVEC object which includes flags for the potential dependency between the set of path computation requests (Link, Node and SRLG diverse). This document proposes a new flag O for domain diversity.

Following new bit is added in the Flags field:

- o O (Domain diverse) bit - TBD12 : when set, this indicates that the computed paths corresponding to the requests specified by the following RP objects MUST NOT have any transit domain(s) in common.

The Domain Diverse O-bit can be used in Hierarchical PCE path computation to compute synchronized domain diverse end to end path or diverse domain sequences.

When domain diverse O bit is set, it is applied to the transit domains. The other bit in SVEC object (N, L, S etc) MAY be set and MUST still be applied in the ingress and egress shared domain.

3.6. PCEP-ERROR object

3.6.1. Hierarchy PCE Error-Type

A new PCEP Error-Type [RFC5440] is used for the H-PCE extension as defined below:

Error-Type	Meaning
TBD8	H-PCE error Error-value=1: parent PCE capability was not advertised Error-value=2: parent PCE capability cannot be provided

Figure 4: H-PCE error

3.7. NO-PATH Object

To communicate the reason(s) for not being able to find a multi-domain path or domain sequence, the NO-PATH object can be used in the PCRep message. [RFC5440] defines the format of the NO-PATH object. The object may contain a NO-PATH-VECTOR TLV to provide additional information about why a path computation has failed.

Three new bit flags are defined to be carried in the Flags field in the NO-PATH-VECTOR TLV carried in the NO-PATH Object.

- o Bit number TBD9: When set, the parent PCE indicates that destination domain unknown;
- o Bit number TBD10: When set, the parent PCE indicates unresponsive child PCE(s);
- o Bit number TBD11: When set, the parent PCE indicates no available resource available in one or more domain(s).

4. H-PCE Procedures

4.1. OPEN Procedure between Child PCE and Parent PCE

If a child PCE wants to use the peer PCE as a parent, it MUST set the R (parent PCE request flag) in the H-PCE-CAPABILITY TLV inside the OPEN object carried in the Open message during the PCEP session initialization procedure.

If the parent PCE can provide the parent function to the peer PCE, it MUST set the I (parent PCE indication flag) in the H-PCE-CAPABILITY TLV inside the OPEN object carried in the Open message during the PCEP session creation procedure.

The child PCE MAY also report its list of domain IDs to the parent PCE by specifying them in the Domain-ID TLVs in the OPEN object carried in the Open message during the PCEP session initialization procedure.

The OF codes defined in this document can be carried in the OF-list TLV of the OPEN object. If the OF-list TLV carries the OF codes, it means that the PCE is capable of implementing the corresponding objective functions. This information can be used for selecting a proper parent PCE when a child PCE wants to get a path that satisfies a certain objective function.

When a specific child PCE sends a PCReq to a peer PCE that requires parental activity and H-PCE capability flags were not set in the

session establishment procedure as described above, the peer PCE should send a PCErr message to the child PCE and specify the error-type=TBD (H-PCE error) and error-value=1 (parent PCE capability was not advertised) in the PCEP-ERROR object.

When a specific child PCE sends a PCReq to a peer PCE that requires parental activity and the peer PCE does not want to act as the parent for it, the peer PCE should send a PCErr message to the child PCE and specify the error-type=TBD (H-PCE error) and error-value=2 (parent PCE capability cannot be provided) in the PCEP-ERROR object.

4.2. Procedure to obtain Domain Sequence

If a child PCE only wants to get the domain sequence for a multi-domain path computation from a parent PCE, it can set the Domain Path Request bit in the H-PCE-FLAG TLV in the RP object carried in a PCReq message. The parent PCE which receives the PCReq message tries to compute a domain sequence for it (instead for E2E path). If the domain path computation succeeds the parent PCE sends a PCRep message which carries the domain sequence in the ERO to the child PCE. Refer [RFC7897] for more details about domain sub-objects in the ERO. Otherwise it sends a PCReq message which carries the NO-PATH object to the child PCE.

5. Error Handling

A PCE that is capable of acting as a parent PCE might not be configured or willing to act as the parent for a specific child PCE. This fact could be determined when the child sends a PCReq that requires parental activity, and could result in a negative response in a PCEP Error (PCErr) message and indicate the hierarchy PCE error-type=TBD8 (H-PCE error) and suitable error-value. (Section 3.6)

Additionally, the parent PCE may fail to find the multi-domain path or domain sequence due to one or more of the following reasons:

- o A child PCE cannot find a suitable path to the egress;
- o The parent PCE do not hear from a child PCE for a specified time;
- o The objective functions specified in the path request cannot be met.

In this case, the parent PCE MAY need to send a negative path computation reply specifying the reason. This can be achieved by including NO-PATH object in the PCRep message. Extension to NO-PATH object is needed to include the aforementioned reasons described in Section 3.7.

6. Manageability Considerations

General PCE and PCEP management considerations are discussed in [RFC4655] and [RFC5440]. There are additional management considerations for H-PCE which are described in [RFC6805], and repeated in this section.

The administrative entity responsible for the management of the parent PCEs must be determined for the following cases:

- o multi-domains (e.g., IGP areas or multiple ASes) within a single service provider network, the management responsibility for the parent PCE would most likely be handled by the service provider,
- o multiple ASes within different service provider networks, it may be necessary for a third party to manage the parent PCEs according to commercial and policy agreements from each of the participating service providers.

6.1. Control of Function and Policy

Control and function will need to be carefully managed in a H-PCE network. A child PCE will need to be configured with the address of its parent PCE. It is expected that there will only be one or two parents of any child.

The parent PCE also needs to be aware of the child PCEs for all child domains that it can see. This information is most likely to be configured (as part of the administrative definition of each domain).

Discovery of the relationships between parent PCEs and child PCEs does not form part of the hierarchical PCE architecture. Mechanisms that rely on advertising or querying PCE locations across domain or provider boundaries are undesirable for security, scaling, commercial, and confidentiality reasons. Specific behavior of the child and parent PCE are described in the following sub-sections.

6.1.1. Child PCE

Support of the hierarchical procedure will be controlled by the management organization responsible for each child PCE. A child PCE must be configured with the address of its parent PCE in order for it to interact with its parent PCE. The child PCE must also be authorized to peer with the parent PCE.

6.1.2. Parent PCE

The parent PCE must only accept path computation requests from authorized child PCEs. If a parent PCE receives requests from an unauthorized child PCE, the request should be dropped. This means that a parent PCE must be configured with the identities and security credentials of all of its child PCEs, or there must be some form of shared secret that allows an unknown child PCE to be authorized by the parent PCE.

6.1.3. Policy Control

It may be necessary to maintain a policy module on the parent PCE [RFC5394]. This would allow the parent PCE to apply commercially relevant constraints such as SLAs, security, peering preferences, and monetary costs.

It may also be necessary for the parent PCE to limit end-to-end path selection by including or excluding specific domains based on commercial relationships, security implications, and reliability.

6.2. Information and Data Models

A MIB module for PCEP was published as RFC 7420 [RFC7420] that describes managed objects for modeling of PCEP communication. A YANG module for PCEP has also been proposed [I-D.ietf-pce-pcep-yang].

A H-PCE MIB module, or additional data model, will be required to report parent PCE and child PCE information, including:

- o parent PCE configuration and status,
- o child PCE configuration and information,
- o notifications to indicate session changes between parent PCEs and child PCEs, and
- o notification of parent PCE TED updates and changes.

6.3. Liveness Detection and Monitoring

The hierarchical procedure requires interaction with multiple PCEs. Once a child PCE requests an end-to-end path, a sequence of events occurs that requires interaction between the parent PCE and each child PCE. If a child PCE is not operational, and an alternate transit domain is not available, then a failure must be reported.

6.4. Verify Correct Operations

Verifying the correct operation of a parent PCE can be performed by monitoring a set of parameters. The parent PCE implementation should provide the following parameters monitored by the parent PCE:

- o number of child PCE requests,
- o number of successful hierarchical PCE procedures completions on a per-PCE-peer basis,
- o number of hierarchical PCE procedure completion failures on a per-PCE-peer basis, and
- o number of hierarchical PCE procedure requests from unauthorized child PCEs.

6.5. Requirements On Other Protocols

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

6.6. Impact On Network Operations

The hierarchical PCE procedure is a multiple-PCE path computation scheme. Subsequent requests to and from the child and parent PCEs do not differ from other path computation requests and should not have any significant impact on network operations.

7. IANA Considerations

7.1. PCEP TLV Type Indicators

IANA Manages the PCEP TLV code point registry (see [RFC5440]). This is maintained as the "PCEP TLV Type Indicators" sub-registry of the "Path Computation Element Protocol (PCEP) Numbers" registry.

This document defines three new PCEP TLVs. IANA is requested to make the following allocation:

Type	TLV name	References

TBD1	H-PCE-CAPABILITY TLV	This I-D
TBD2	Domain-ID TLV	This I-D
TBD3	H-PCE-FLAG TLV	This I-D

7.2. H-PCE-CAPABILITY TLV Flags

This document requests that a new sub-registry, named " H-PCE-CAPABILITY TLV Flag Field", is created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the Flag field in the H-PCE-CAPABILITY TLV of the PCEP OPEN object.

New values are to be assigned by Standards Action [RFC5226]. Each bit should be tracked with the following qualities:

- o Bit number (counting from bit 0 as the most significant bit)
- o Capability description
- o Defining RFC

The following values are defined in this document:

Bit	Description	Reference
31	R (Parent PCE Request bit)	This I.D.
30	I (Parent PCE Indication bit)	This I.D.

7.3. Domain-ID TLV Domain type

This document requests that a new sub-registry, named " Domain-ID TLV Domain type", is created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the Domain-Type field of the Domain-ID TLV.

Value	Meaning
1	2-byte AS number
2	4-byte AS number
3	4-byte OSPF area ID
4	Variable length IS-IS area ID

7.4. H-PCE-FLAG TLV Flags

This document requests that a new sub-registry, named "H-PCE-FLAGS TLV Flag Field", is created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the Flag field in the H-PCE-FLAGS TLV of the PCEP RP object. New values are to be assigned by Standards Action [RFC5226]. Each bit should be tracked with the following qualities:

- o Bit number (counting from bit 0 as the most significant bit)

- o Capability description
- o Defining RFC

The following values are defined in this document:

Bit	Description	Reference
31	S (Domain Sequence bit)	This I.D.
30	D (Disallow Domain Re-entry bit)	This I.D.

7.5. OF Codes

IANA maintains registry of Objective Function (described in [RFC5541]) at the sub-registry "Objective Function". Two new Objective Functions have been defined in this document.

IANA is requested to make the following allocations:

Code Point	Name	Reference
TBD4	Minimum number of Transit Domains (MTD)	This I.D.
TBD5	Minimize number of Border Nodes (MBN)	This I.D.
TBD13	Minimize the number of Common Transit Domains. (MCTD)	This I.D.

7.6. METRIC Types

IANA maintains one sub-registry for "METRIC object T field". Two new metric types are defined in this document for the METRIC object (specified in [RFC5440]).

IANA is requested to make the following allocations:

Value	Description	Reference
TBD6	Domain Count metric	This I.D.
TBD7	Border Node Count metric	This I.D.

7.7. New PCEP Error-Types and Values

IANA maintains a registry of Error-Types and Error-values for use in PCEP messages. This is maintained as the "PCEP-ERROR Object Error Types and Values" sub-registry of the "Path Computation Element Protocol (PCEP) Numbers" registry.

IANA is requested to make the following allocations:

Error-Type	Meaning and error values	Reference

TBD8	H-PCE Error	This I.D.
	Error-value=1 Parent PCE Capability not advertised	
	Error-value=2 Parent PCE Capability not supported	

7.8. New NO-PATH-VECTOR TLV Bit Flag

IANA maintains a registry of bit flags carried in the PCEP NO-PATH-VECTOR TLV in the PCEP NO-PATH object as defined in [RFC5440]. IANA Is requested to assign three new bit flag as follows:

Bit Number	Name Flag	Reference

TBD9	Destination Domain unknown	This I.D.
TBD10	Unresponsive child PCE(s)	This I.D.
TBD11	No available resource in one or more domain	This I.D.

7.9. SVEC Flag

IANA maintains a registry of bit flags carried in the PCEP SVEC object as defined in [RFC5440]. IANA Is requested to assign one new bit flag as follows:

Bit Number	Name Flag	Reference

TBD13	Domain Diverse	This I.D.

8. Security Considerations

The hierarchical PCE procedure relies on PCEP and inherits the security requirements defined in [RFC5440]. As PCEP operates over TCP, it may also make use of TCP security mechanisms, such as TCP-AO or [RFC8253].

H-PCE operation also relies on information used to build the TED. Attacks on a parent or child PCE may be achieved by falsifying or impeding this flow of information. If the child PCE listens to the IGP or BGP-LS for populating the TED, then normal IGP or BGP-LS security measures may be applied, and it should be noted that an IGP routing system is generally assumed to be a trusted domain such that router subversion is not a risk. The parent PCE TED is constructed as described in this document and may involve:

- o multiple parent-child relationships using PCEP
- o the parent PCE listening to child domain IGPs (with the same security features as a child PCE listening to its IGP)
- o an external mechanism (such as [RFC7752]), which will need to be authorized and secured.

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 H-PCE architecture is specifically designed to enable as much isolation of domain topology and capabilities information as is possible.

For further considerations of the security issues related to inter-AS path computation, see [RFC5376].

9. Implementation Status

The H-PCE architecture and protocol procedures describe in this I-D were implemented and tested for a variety of optical research applications.

9.1. Inter-layer traffic engineering with H-PCE

This work was led by:

- o Ramon Casellas [ramon.casellas@cttc.es]

- o Centre Tecnologic de Telecomunicacions de Catalunya (CTTC)

The H-PCE instances (parent and child) were multi-threaded asynchronous processes. Implemented in C++11, using C++ Boost Libraries. The targeted system used to deploy and run H-PCE applications was a POSIX system (Debian GNU/Linux operating system).

Some parts of the software may require a Linux Kernel, the availability of a Routing Controller running collocated in the same host and the usage of libnetfilter / libipq and GNU/Linux firewalling capabilities. Most of the functionality, including algorithms is done by means of plugins (e.g., as shared libraries or .so files in Unix systems).

The CTTC PCE supports the H-PCE architecture, but also supports stateful PCE with active capabilities, as an OpenFlow controller, and has dedicated plugins to support monitoring, BRPC, P2MP, path keys, back end PCEs. Management of the H-PCE entities was supported via HTTP and CLI via Telnet.

Further details of the H-PCE prototyping and experimentation can be found in the following scientific papers:

R. Casellas, R. Martinez, R. Munoz, L. Liu, T. Tsuritani, I. Morita, "Inter-layer traffic engineering with hierarchical-PCE in MPLS-TP over wavelength switched optical networks" , Optics Express, Vol. 20, No. 28, December 2012.

R. Casellas, R. Martinez, R. Munoz, L. Liu, T. Tsuritani, I. Morita, M. Msurusawa, "Dynamic virtual link mesh topology aggregation in multi-domain translucent WSON with hierarchical-PCE", Optics Express Journal, Vol. 19, No. 26, December 2011.

R. Casellas, R. Munoz, R. Martinez, R. Vilalta, L. Liu, T. Tsuritani, I. Morita, V. Lopez, O. Gonzalez de Dios, J. P. Fernandez-Palacios, "SDN based Provisioning Orchestration of OpenFlow/GMPLS Flexi-grid Networks with a Stateful Hierarchical PCE", in Proceedings of Optical Fiber Communication Conference and Exposition (OFC), 9-13 March, 2014, San Francisco (EEUU).
Extended Version to appear in Journal Of Optical Communications and Networking January 2015

F. Paolucci, O. Gonzalez de Dios, R. Casellas, S. Duhovnikov, P. Castoldi, R. Munoz, R. Martinez, "Experimenting Hierarchical PCE Architecture in a Distributed Multi-Platform Control Plane Testbed" , in Proceedings of Optical Fiber Communication Conference and Exposition (OFC) and The National Fiber Optic

Engineers Conference (NFOEC), 4-8 March, 2012, Los Angeles, California (USA).

R. Casellas, R. Martinez, R. Munoz, L. Liu, T. Tsuritani, I. Morita, M. Tsurusawa, "Dynamic Virtual Link Mesh Topology Aggregation in Multi-Domain Translucent WSON with Hierarchical-PCE", in Proceedings of 37th European Conference and Exhibition on Optical Communication (ECOC 2011), 18-22 September 2011, Geneve (Switzerland).

R. Casellas, R. Munoz, R. Martinez, "Lab Trial of Multi-Domain Path Computation in GMPLS Controlled WSON Using a Hierarchical PCE", in Proceedings of OFC/NFOEC Conference (OFC2011), 10 March 2011, Los Angeles (USA).

9.2. Telefonica Netphony (Open Source PCE)

The Telefonica Netphony PCE is an open source Java-based implementation of a Path Computation Element, with several flavours, and a Path Computation Client. The PCE follows a modular architecture and allows to add customized algorithms. The PCE has also stateful and remote initiation capabilities. In current version, three components can be built, a domain PCE (aka child PCE), a parent PCE (ready for the H-PCE architecture) and a PCC (path computation client).

This work was led by:

- o Oscar Gonzalez de Dios [oscar.gonzalezdedios@telefonica.com]
- o Victor Lopez Alvarez [victor.lopezalvarez@telefonica.com]
- o Telefonica I+D, Madrid, Spain

The PCE code is publicly available in a GitHub repository:

- o <https://github.com/telefonicaid/netphony-pce>

The PCEP protocol encodings are located in the following repository:

- o <https://github.com/telefonicaid/netphony-network-protocols>

The traffic engineering database and a BGP-LS speaker to fill the database is located in:

- o <https://github.com/telefonicaid/netphony-topology>

The parent and child PCE are multi-threaded java applications. The path computation uses the jgrapht free Java class library (0.9.1) that provides mathematical graph-theory objects and algorithms. Current version of netphony PCE runs on java 1.7 and 1.8, and has been tested in GNU/Linux, Mac OS-X and Windows environments. The management of the parent and domain PCEs is supported through CLI via Telnet, and configured via XML files.

Further details of the netphony H-PCE prototyping and experimentation can be found in the following research papers:

- o O. Gonzalez de Dios, R. Casellas, F. Paolucci, A. Napoli, L. Gifre, A. Dupas, E. Hugues-Salas, R. Morro, S. Belotti, G. Meloni, T. Rahman, V.P Lopez, R. Martinez, F. Fresi, M. Bohn, S. Yan, L. Velasco, . Layec and J. P. Fernandez-Palacios: Experimental Demonstration of Multivendor and Multidomain EON With Data and Control Interoperability Over a Pan-European Test Bed, in Journal of Lightwave Technology, Dec. 2016, Vol. 34, Issue 7, pp. 1610-1617.
- o O. Gonzalez de Dios, R. Casellas, R. Morro, F. Paolucci, V. Lopez, R. Martinez, R. Munoz, R. Villalta, P. Castoldi: "Multi-partner Demonstration of BGP-LS enabled multi-domain EON, in Journal of Optical Communications and Networking, Dec. 2015, Vol. 7, Issue 12, pp. B153-B162.
- o F. Paolucci, O. Gonzalez de Dios, R. Casellas, S. Duhovnikov, P. Castoldi, R. Munoz, R. Martinez, "Experimenting Hierarchical PCE Architecture in a Distributed Multi-Platform Control Plane Testbed" , in Proceedings of Optical Fiber Communication Conference and Exposition (OFC) and The National Fiber Optic Engineers Conference (NFOEC), 4-8 March, 2012, Los Angeles, California (USA).

9.3. Implementation 3: H-PCE Proof of Concept developed by Huawei

Huawei developed this H-PCE on the Huawei Versatile Routing Platform (VRP) to experiment with the hierarchy of PCE. Both end to end path computation as well as computation for domain-sequence are supported.

This work was led by:

- o Udayasree Pallee [udayasreereddy@gmail.com]
- o Dhruv Dhody [dhruv.ietf@gmail.com]
- o Huawei Technologies, Bangalore, India

Further work on stateful H-PCE [I-D.ietf-pce-stateful-hpce] is being carried out on ONOS.

10. Contributing Authors

Xian Zhang
Huawei
EMail: zhang.xian@huawei.com

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

EMail: dhruv.ietf@gmail.com

11. References

11.1. 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>>.
- [RFC5152] Vasseur, JP., Ed., Ayyangar, A., Ed., and R. Zhang, "A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)", RFC 5152, DOI 10.17487/RFC5152, February 2008, <<https://www.rfc-editor.org/info/rfc5152>>.
- [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>>.
- [RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", RFC 5541, DOI 10.17487/RFC5541, June 2009, <<https://www.rfc-editor.org/info/rfc5541>>.
- [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>>.

11.2. Informative References

- [RFC4105] Le Roux, J., Ed., Vasseur, J., Ed., and J. Boyle, Ed., "Requirements for Inter-Area MPLS Traffic Engineering", RFC 4105, DOI 10.17487/RFC4105, June 2005, <<https://www.rfc-editor.org/info/rfc4105>>.
- [RFC4216] Zhang, R., Ed. and J. Vasseur, Ed., "MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements", RFC 4216, DOI 10.17487/RFC4216, November 2005, <<https://www.rfc-editor.org/info/rfc4216>>.
- [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>>.
- [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>>.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 5226, DOI 10.17487/RFC5226, May 2008, <<https://www.rfc-editor.org/info/rfc5226>>.
- [RFC5376] Bitar, N., Zhang, R., and K. Kumaki, "Inter-AS Requirements for the Path Computation Element Communication Protocol (PCECP)", RFC 5376, DOI 10.17487/RFC5376, November 2008, <<https://www.rfc-editor.org/info/rfc5376>>.
- [RFC5394] Bryskin, I., Papadimitriou, D., Berger, L., and J. Ash, "Policy-Enabled Path Computation Framework", RFC 5394, DOI 10.17487/RFC5394, December 2008, <<https://www.rfc-editor.org/info/rfc5394>>.
- [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>>.

- [RFC5441] Vasseur, JP., Ed., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", RFC 5441, DOI 10.17487/RFC5441, April 2009, <<https://www.rfc-editor.org/info/rfc5441>>.
- [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>>.
- [RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <<https://www.rfc-editor.org/info/rfc7420>>.
- [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <<https://www.rfc-editor.org/info/rfc7752>>.
- [RFC7897] Dhody, D., Palle, U., and R. Casellas, "Domain Subobjects for the Path Computation Element Communication Protocol (PCEP)", RFC 7897, DOI 10.17487/RFC7897, June 2016, <<https://www.rfc-editor.org/info/rfc7897>>.
- [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>>.
- [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-08 (work in progress), June 2018.
- [I-D.ietf-pce-stateful-hpce]
Dhody, D., Lee, Y., Ceccarelli, D., Shin, J., King, D., and O. Dios, "Hierarchical Stateful Path Computation Element (PCE).", draft-ietf-pce-stateful-hpce-05 (work in progress), June 2018.

Authors' Addresses

Fatai Zhang
Huawei
Huawei Base, Bantian, Longgang District
Shenzhen 518129
China

EMail: zhangfatai@huawei.com

Quintin Zhao
Huawei
125 Nagog Technology Park
Acton, MA 01719
USA

EMail: quintin.zhao@huawei.com

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

EMail: ogondio@tid.es

Ramon Casellas
CTTC
Av. Carl Friedrich Gauss n.7
Barcelona, Castelldefels
Spain

EMail: ramon.casellas@cttc.es

Daniel King
Old Dog Consulting
UK

EMail: daniel@olddog.co.uk

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 2, 2019

F. Zhang
Q. Zhao
Huawei
O. Gonzalez de Dios
Telefonica I+D
R. Casellas
CTTC
D. King
Old Dog Consulting
June 1, 2019

Extensions to Path Computation Element Communication Protocol (PCEP) for
Hierarchical Path Computation Elements (PCE)
draft-ietf-pce-hierarchy-extensions-11

Abstract

The Hierarchical Path Computation Element (H-PCE) architecture is defined in RFC 6805. It provides a mechanism to derive an optimum end-to-end path in a multi-domain environment by using a hierarchical relationship between domains to select the optimum sequence of domains and optimum paths across those domains.

This document defines extensions to the Path Computation Element Protocol (PCEP) to support Hierarchical PCE procedures.

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 2, 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 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

1. Introduction3
1.1. Scope4
1.2. Terminology5
1.3. Requirements Language5
2. Requirements for H-PCE5
2.1. Path Computation Request6
2.1.1. Qualification of PCEP Requests6
2.1.2. Multi-domain Objective Functions6
2.1.3. Multi-domain Metrics7
2.2. Parent PCE Capability Advertisement7
2.3. PCE Domain Identification7
2.4. Domain Diversity7
3. PCEP Extensions8
3.1. Applicability to PCC-PCE Communications8
3.2. OPEN Object8
3.2.1. H-PCE Capability TLV8
3.2.1.1 Backwards Compatibility9
3.2.2. Domain-ID TLV10
3.3. RP Object11
3.3.1. H-PCE-FLAG TLV11
3.3.2. Domain-ID TLV12
3.4. Objective Functions12
3.4.1. OF Codes12
3.4.2. OF Object13
3.5. Metric Object14
3.6. SVEC Object15
3.7. PCEP-ERROR Object15
3.7.1. Hierarchy PCE Error-Type15
3.8. NO-PATH Object16
4. H-PCE Procedures16
4.1. OPEN Procedure between Child PCE and Parent PCE16
4.2. Procedure to Obtain Domain Sequence17
5. Error Handling17
6. Manageability Considerations18
6.1. Control of Function and Policy18
6.1.1. Child PCE18
6.1.2. Parent PCE

6.1.3.	Policy Control19
6.2.	Information and Data Models19
6.3.	Liveness Detection and Monitoring20
6.4.	Verify Correct Operations20
6.5.	Requirements On Other Protocols20
6.6.	Impact On Network Operations20
7.	IANA Considerations20
7.1.	PCEP TLV Type Indicators20
7.2.	H-PCE-CAPABILITY TLV Flags21
7.3.	Domain-ID TLV Domain type21
7.4.	H-PCE-FLAG TLV Flags22
7.5.	OF Codes22
7.6.	METRIC Types22
7.7.	New PCEP Error-Types and Values23
7.8.	New NO-PATH-VECTOR TLV Bit Flag23
7.9.	SVEC Flag24
7.10.	NO-PATH VECTOR TLV Bit Flag.24
8.	Security Considerations24
9.	Contributing Authors24
10.	Acknowledgements25
11.	References25
11.1.	Normative References25
11.2.	Informative References25
	Authors' Addresses28
	Appendix28
A1.	Implementation Status28
A1.1.	Inter-layer traffic engineering with H-PCE29
A1.2.	Telefonica Netphony (Open Source PCE)30
A1.3.	H-PCE Proof of Concept developed by Huawei31

1. Introduction

The Path Computation Element communication Protocol (PCEP) provides a mechanism for Path Computation Elements (PCEs) and Path Computation Clients (PCCs) to exchange requests for path computation and responses that provide computed paths.

The capability to compute the routes of end-to-end inter-domain MPLS Traffic Engineering (MPLS-TE) and GMPLS Label Switched Paths (LSPs) is expressed as requirements in [RFC4105] and [RFC4216]. This capability may be realized by a PCE [RFC4655]. The methods for establishing and controlling inter-domain MPLS-TE and GMPLS LSPs are documented in [RFC4726].

[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 architecture, the parent PCE is used to compute a multi-domain path based on the domain connectivity information. A child PCE may be responsible for single or multiple domains and is used to compute the intra-domain path based on its own domain topology information.

The H-PCE end-to-end domain path computation procedure is described below:

- o A path computation client (PCC) sends the inter-domain path computation requests to the child PCE responsible for its domain;
- o The child PCE forwards the request to the parent PCE;
- o The parent PCE computes the likely domain paths from the ingress domain to the egress domain;
- o The parent PCE sends the intra-domain path computation requests (between the domain border nodes) to the child PCEs which are responsible for the domains along the domain path;
- o The child PCEs return the intra-domain paths to the parent PCE;
- o The parent PCE constructs the end-to-end inter-domain path based on the intra-domain paths;
- o The parent PCE returns the inter-domain path to the child PCE;
- o The child PCE forwards the inter-domain path to the PCC.

The parent PCE may be requested to provide only the sequence of domains to a child PCE so that alternative inter-domain path computation procedures, including Per Domain (PD) [RFC5152] and Backwards Recursive Path Computation (BRPC) [RFC5441], may be used.

This document defines the PCEP extensions for the purpose of implementing Hierarchical PCE procedures, which are described in [RFC6805].

1.1. Scope

The following functions are out of scope of this document:

- o Determination of Destination Domain (section 4.5 of [RFC6805]):
 - * via a collection of reachability information from child domain;
 - * via requests to the child PCEs to discover if they contain the

- destination node;
- * or any other methods.
- o Parent Traffic Engineering Database (TED) methods (section 4.4 of [RFC6805]), although suitable mechanisms include:
 - * YANG-based management interfaces;
 - * BGP-LS [RFC7752];
 - * Future extension to PCEP (such as [I-D.dhodylee-pce-pcep-ls]).
- o Learning of Domain connectivity and boundary nodes (BN) addresses, methods to achieve this function include:
 - * YANG-based management interfaces;
 - * BGP-LS [RFC7752];
 - * Future extension to PCEP (such as [I-D.dhodylee-pce-pcep-ls]).
- o Stateful PCE Operations (Refer [I-D.ietf-pce-stateful-hpce])
- o Applicability of hierarchical PCE to large multi-domain environments.
 - * The hierarchical relationship model is described in [RFC6805]. It is applicable to environments with small groups of domains where visibility from the ingress LSRs is limited. As highlighted in [RFC7399] applying the hierarchical PCE model to very large groups of domains, such as the Internet, is not considered feasible or desirable.

1.2. Terminology

This document uses the terminology defined in [RFC4655], [RFC5440] and the additional terms defined in Section 1.4 of [RFC6805].

1.3. 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.

2. Requirements for H-PCE

This section compiles the set of requirements to the PCEP extensions to support the H-PCE architecture and procedures.

[RFC6805] identifies high-level requirements of PCEP extensions required to support the hierarchical PCE model.

2.1. Path Computation Request

The Path Computation Request (PCReq) [RFC5440] messages are used by a PCC or a PCE to make a path computation request to a PCE. In order to achieve the full functionality of the H-PCE procedures, the PCReq message needs to include:

- o Qualification of PCE Requests (Section 4.8.1. of [RFC6805]);
- o Multi-domain Objective Functions (OF);
- o Multi-domain Metrics.

2.1.1. Qualification of PCEP Requests

As described in Section 4.8.1 of [RFC6805], the H-PCE architecture introduces new request qualifications, which are:

- o The ability for a child PCE to indicate that a path computation request sent to a parent PCE should be satisfied by a domain sequence only, that is, not by a full end-to-end path. This allows the child PCE to initiate a per-domain (PD) [RFC5152] or a backward recursive path computation (BRPC) [RFC5441].
- o As stated in [RFC6805], Section 4.5, if a PCC knows the egress domain, it can supply this information as the path computation request. The PCC may also want to specify the destination domain information in a PCEP request, if it is known.
- o An inter domain path computed by parent PCE should be capable of disallowing specific domain re-entry.

2.1.2. Multi-domain Objective Functions

For H-PCE inter-domain path computation, there are three new Objective Functions defined in this document:

- o Minimize the number of Transit Domains (MTD)
- o Minimize the number of border nodes (MBN)
- o Minimize the number of Common Transit Domains (MCTD)

The PCC may specify the multi-domain Objective Function code to use when requesting inter-domain path computation, it may also

include intra-domain OFs, such as Minimum Cost Path (MCP) [RFC5441], which must be considered by participating child PCEs.

2.1.3. Multi-domain Metrics

For inter-domain path computation, there are several path metrics of interest.

- o Domain count (number of domains crossed);
- o Border Node count.

A PCC may be able to limit the number of domains crossed by applying a limit on these metrics. Details in Section 3.4.

2.2. Parent PCE Capability Advertisement

A PCEP Speaker (Parent PCE or Child PCE) that supports and wishes to use the procedures described in this document must advertise the fact and negotiate its role with its PCEP peers. It does this using the "H-PCE Capability" TLV, described in Section 3.2.1, in the OPEN Object to advertise its support for PCEP extensions for H-PCE Capability.

During the PCEP session establishment procedure, the child PCE needs to be capable of indicating to the parent PCE whether it requests the parent PCE capability or not.

2.3. PCE Domain Identification

A PCE domain is a single domain with an associated PCE. Although it is possible for a PCE to manage multiple domains simultaneously. The PCE domain could be an IGP area or AS.

The PCE domain identifiers MAY be provided during the PCEP session establishment procedure.

2.4. Domain Diversity

In a multi-domain environment, Domain Diversity is defined in [RFC6805] and described as "A pair of paths are domain-diverse if they do not transit any of the same domains. A pair of paths that share a common ingress and egress are domain-diverse if they only share the same domains at the ingress and egress (the ingress and egress domains). Domain diversity may be maximized for a pair of paths by selecting paths that have the smallest number of shared domains."

The main motivation behind domain diversity is to avoid fate sharing, but it can also be because of some geo-political reasons and commercial relationships that would require domain diversity. For example, a pair of paths should choose different transit Autonomous System (AS) because of some policy considerations.

In the case when full domain diversity could not be achieved, it is helpful to minimize the commonly shared domains. Also, it is interesting to note that other scope of diversity (node, link, SRLG etc.) can still be applied inside the commonly shared domains.

3. PCEP Extensions

This section defines extensions to PCEP [RFC5440] to support the H-PCE procedures.

3.1 Applicability to PCC-PCE Communications

Although the extensions defined in this document are intended primarily for use between a child PCE and a parent PCE, they are also applicable for communications between a PCC and its PCE.

Thus, the information that may be encoded in a PCReq can be sent from a PCC towards the child PCE. This includes the RP object (Section 3.3) and the Objective Function (OF) codes and objects (Section 3.4). A PCC and a child PCE could also exchange the capability (Section 3.2.1) during its session.

This allows a PCC to request paths that transit multiple domains utilizing the capabilities defined in this document.

3.2. OPEN Object

Two new TLVs are defined in this document to be carried within an OPEN object. This way, during the PCEP session establishment, the H-PCE capability and Domain information can be advertised.

3.2.1. H-PCE Capability TLV

The H-PCE-CAPABILITY TLV is an optional TLV associated with the OPEN Object [RFC5440] to exchange H-PCE capability of PCEP speakers.

Its format is shown in the following figure:

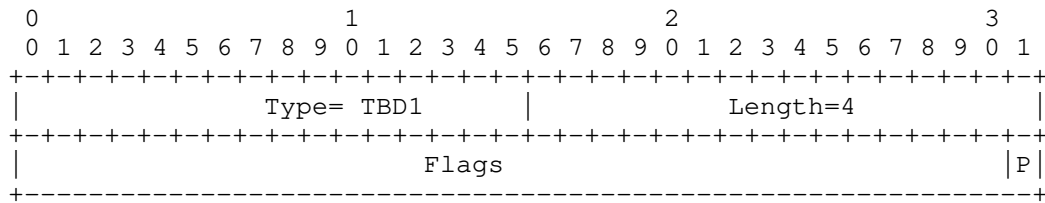


Figure 1: H-PCE-CAPABILITY TLV format

The type of the TLV is TBD1 (to be assigned by IANA), and it has a fixed length of 4 octets.

The value comprises a single field - Flags (32 bits):

P (Parent PCE Request bit): if set, will signal that the child PCE wishes to use the peer PCE as a parent PCE.

Unassigned bits MUST be set to 0 on transmission and MUST be ignored on receipt.

The inclusion of this TLV in an OPEN object indicates that the H-PCE extensions are supported by the PCEP speaker. The child PCE MUST include this TLV and set the P flag. The parent PCE MUST include this TLV and unset the P flag.

The setting of the P flag (parent PCE request bit) would mean that the PCEP speaker wants the peer to be a parent PCE, so in the case of a PCC to Child-PCE relationship, neither entity would set the P flag.

If both peers attempt to set the P flag then the session establishment MUST fail, and the PCEP speaker MUST respond with PCErr message using Error-Type 1: "PCEP Session Establishment Failure" as per [RFC5440].

If the PCE understands the H-PCE path computation request but did not advertise its H-PCE capability, it MUST send a PCErr message with Error-Type=TBD8 ("H-PCE error") and Error-Value=1 ("H-PCE Capability not advertised").

3.2.1.1 Backwards Compatibility

Section 7.1 of [RFC5440] requires that "Unrecognized TLVs MUST be ignored.

That means that a PCE that does not support this document but that receives an Open Message containing an Open Object that includes

an H-PCE-CAPABILITIES TLV will ignore that TLV and will continue to attempt to establish a PCEP session. It will, however, not include the TLV in the Open message that it sends, so the H-PCE relationship will not be created.

If a PCE does not support the extensions defined in this document but receives them in a PCEP message (notwithstanding the fact that the session was not established as supporting a H-PCE relationship), the receiving PCE will ignore the H-PCE related parameters because they are all encoded in TLVs within standard PCEP objects.

3.2.2. Domain-ID TLV

The Domain-ID TLV, when used in the OPEN object, identifies the domains served by the PCE. The child PCE uses this mechanism to inform the domain information to the parent PCE.

The Domain-ID TLV is defined below:

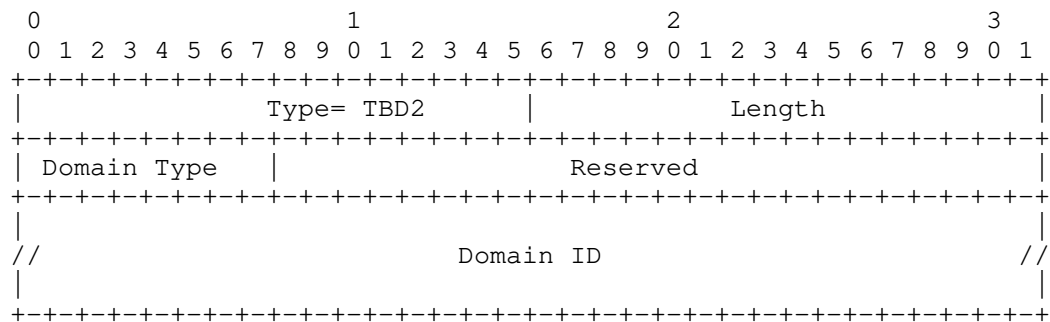


Figure 2: Domain-ID TLV format

The type of the TLV is TBD2 (to be assigned by IANA), and it has a variable Length of the value portion. The value part comprises:

Domain Type (8 bits): Indicates the domain type. Four types of domain are currently defined:

- * Type=1: the Domain ID field carries a 2-byte AS number. Padded with trailing zeros to a 4-byte boundary.
- * Type=2: the Domain ID field carries a 4-byte AS number.
- * Type=3: the Domain ID field carries a 4-byte OSPF area ID.
- * Type=4: the Domain ID field carries (2-byte Area-Len, variable length IS-IS area ID). Padded with trailing zeros to a 4-byte

boundary.

Reserved: Zero at transmission; ignored at the receipt.

Domain ID (variable): Indicates an IGP Area ID or AS number as per the Domain Type field. It can be 2 bytes, 4 bytes or variable length depending on the domain identifier used. It is padded with trailing zeros to a 4-byte boundary. In case of IS-IS it includes the Area-Len as well.

In the case a PCE serves more than one domain, multiple Domain-ID TLVs are included for each domain it serves.

3.3. RP Object

3.3.1. H-PCE-FLAG TLV

The H-PCE-FLAG TLV is an optional TLV associated with the RP Object [RFC5440] to indicate the H-PCE path computation request and options.

Its format is shown in the following figure:

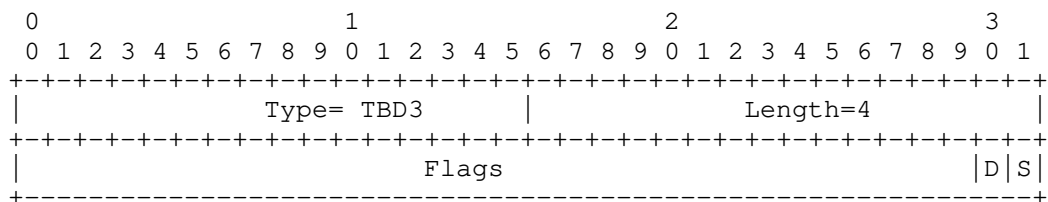


Figure 3: H-PCE-FLAG TLV format

The type of the TLV is TBD3 (to be assigned by IANA), and it has a fixed length of 4 octets.

The value comprises a single field - Flags (32 bits):

S (Domain Sequence bit): if set, will signal that the child PCE wishes to get only the domain sequence in the path computation reply. Refer to Section 3.7 of [RFC7897] for details.

D (Disallow Domain Re-entry bit): if set, will signal that the computed path does not enter a domain more than once.

Unassigned bits MUST be set to 0 on transmission and MUST be ignored on receipt.

The presence of the TLV indicates that the H-PCE based path

computation is requested as per this document.

3.3.2. Domain-ID TLV

The Domain-ID TLV, carried in an OPEN object, is used to indicate a (list of) managed domains and is described in Section 3.3.1. This TLV, when carried in an RP object, indicates the destination domain ID. If a PCC knows the egress domain, it can supply this information in the PCReq message. The format and procedure of this TLV are defined in Section 3.2.2.

If a Domain-id TLV is used in the RP object, and the destination is not actually in the indicated domain, then the parent PCE should respond with a NO-PATH object and NO-PATH VECTOR TLV should be used. A new bit number is assigned to indicate "Destination not found in the indicated domain" (see Section 3.7).

3.4. Objective Functions

3.4.1. OF Codes

[RFC5541] defines a mechanism to specify an Objective Function that is used by a PCE when it computes a path. Three new Objective Functions are defined for H-PCE, these are:

o MTD

- * Name: Minimize the number of Transit Domains (MTD)
- * Objective Function Code - TBD4 (to be assigned by IANA)
- * Description: Find a path P such that it passes through the least number of transit domains.
- * Objective functions are formulated using the following terminology:
 - + A network comprises a set of N domains $\{D_i, (i=1\dots N)\}$.
 - + A path P passes through K unique domains $\{D_{pi}, (i=1\dots K)\}$.
 - + Find a path P such that the value of K is minimized.

o MBN

- * Name: Minimize the number of border nodes.
- * Objective Function Code - TBD5 (to be assigned by IANA)

- * Description: Find a path P such that it passes through the least number of border nodes.
- * Objective functions are formulated using the following terminology:
 - + A network comprises a set of N links $\{L_i, (i=1\dots N)\}$.
 - + A path P is a list of K links $\{L_{pi}, (i=1\dots K)\}$.
 - + $D(L_{pi})$ is a function that determines if the links L_{pi} and L_{pi+1} belong to different domains, $D(L_i) = 1$ if link L_i and L_{i+1} belong to different domains, $D(L_k) = 0$ if link L_k and L_{k+1} belong to the same domain.
 - + The number of border node in a path P is denoted by $B(P)$, where $B(P) = \sum\{D(L_{pi}), (i=1\dots K-1)\}$.
 - + Find a path P such that $B(P)$ is minimized.

There is one objective function that applies to a set of synchronized path computation requests to increase the domain diversity:

- o MCTD

- * Name: Minimize the number of Common Transit Domains
- * Objective Function Code - TBD13 (to be assigned by IANA)
- * Description: Find a set of paths such that it passes through the least number of common transit domains.
 - + A network comprises a set of N domains $\{D_i, (i=1\dots N)\}$.
 - + A path P passes through K unique domains $\{D_{pi}, (i=1\dots K)\}$.
 - + A set of paths $\{P_1\dots P_m\}$ have L transit domains that are common to more than one path $\{D_{pi}, (i=1\dots L)\}$.
 - + Find a set of paths such that the value of L is minimized.

3.4.2. OF Object

The OF (Objective Function) object [RFC5541] is carried within a PCReq message so as to indicate the desired/required objective function to be applied by the PCE during path computation. As per Section 3.2 of [RFC5541] a single OF object may be included in a path

computation request.

The new OF codes described in Section 3.4.1 are applicable at the inter-domain path computation performed by the parent PCE, it is also necessary to specify the OF code that may be applied for the intra-domain path computation performed by the child PCE. To accommodate this, the OF-List TLV (described in Section 2.1. of [RFC5541]) is included in the OF object as an optional TLV.

The OF-List TLV allows encoding of multiple OF codes. When this TLV is included inside the OF object, only the first OF-code in the OF-LIST TLV is considered. The parent PCE MUST use this OF code in the OF object when sending the intra domain path computation request to the child PCE. If the OF list TLV is included in the OF Object, the OF Code inside the OF Object MUST include one of the H-PCE Objective Functions defined in this document, the OF Code inside the OF List TLV MUST NOT include an H-PCE Objective Function. If this condition is not met, the PCEP speaker MUST respond with a PCerr message with Error-Type=10 (Reception of an invalid object) and Error-Value=TBD15 (Incompatible OF codes in H-PCE).

If the Objective Functions defined in this document are unknown or unsupported by a PCE, then the procedure as defined in [RFC5541] is followed.

3.5. Metric Object

The METRIC object is defined in Section 7.8 of [RFC5440], comprising of metric-value, metric-type (T field) and flags. This document defines the following types for the METRIC object for H-PCE:

- o T=TBD6: Domain count metric (number of domains crossed);
- o T=TBD7: Border Node count metric (number of border nodes crossed).

The domain count metric type of the METRIC object encodes the number of domains crossed in the path. The border node count metric type of the METRIC object encodes the number of border nodes in the path. If a domain is re-entered, then domain should be double counted.

A PCC or child PCE MAY use the metric in a PCReq message for an inter-domain path computation, meeting the number of domain or border nodes crossing requirement. As per [RFC5440], in this case, the B bit is set to suggest a bound (a maximum) for the metric that must not be exceeded for the PCC to consider the computed path as acceptable.

A PCC or child PCE MAY also use this metric to ask the PCE to optimize the metric during inter-domain path computation. In this

case, the B flag is cleared, and the C flag is set.

The Parent PCE MAY use the metric in a PCRep message along with a NO-PATH object in the case where the PCE cannot compute a path meeting this constraint. A PCE MAY also use this metric to send the computed end to end metric value in a reply message.

3.6. SVEC Object

[RFC5440] defines SVEC object which includes flags for the potential dependency between the set of path computation requests (Link, Node and SRLG diverse). This document defines a new flag O for domain diversity.

The following new bit is added to the Flags field:

- o Domain Diverse O-bit - TBD14 : when set, this indicates that the computed paths corresponding to the requests specified by the following RP objects MUST NOT have any transit domains in common.

The Domain Diverse O-bit can be used in Hierarchical PCE path computation to compute synchronized domain diverse end to end path or diverse domain sequences.

When domain diverse O bit is set, it is applied to the transit domains. The other bit in SVEC object (N, L, S etc.) MAY be set and MUST still be applied in the ingress and egress shared domain.

3.7. PCEP-ERROR Object

3.7.1. Hierarchy PCE Error-Type

A new PCEP Error-Type [RFC5440] is used for the H-PCE extension as defined below:

Error-Type	Meaning
TBD8	H-PCE error Error-value=1: H-PCE capability was not advertised Error-value=2: parent PCE capability cannot be provided

Figure 4: H-PCE error

3.8. NO-PATH Object

To communicate the reason(s) for not being able to find a multi-domain path or domain sequence, the NO-PATH object can be used in the PCRep message. [RFC5440] defines the format of the NO-PATH object. The object may contain a NO-PATH-VECTOR TLV to provide additional information about why a path computation has failed.

Three new bit flags are defined to be carried in the Flags field in the NO-PATH-VECTOR TLV carried in the NO-PATH Object.

- o Bit number TBD9: When set, the parent PCE indicates that destination domain unknown;
- o Bit number TBD10: When set, the parent PCE indicates unresponsive child PCE(s);
- o Bit number TBD11: When set, the parent PCE indicates no available resource available in one or more domains.
- o Bit number TBD12: When set, the parent PCE indicates that the destination is not found in the indicated domain.

4. H-PCE Procedures

The H-PCE path computation procedure is described in [RFC6805].

4.1. OPEN Procedure between Child PCE and Parent PCE

If a child PCE wants to use the peer PCE as a parent, it MUST set the P (parent PCE request flag) in the H-PCE-CAPABILITY TLV inside the OPEN object carried in the Open message during the PCEP session initialization procedure.

The child PCE MAY also report its list of domain IDs, to the parent PCE, by specifying them in the Domain-ID TLVs in the OPEN object. This object is carried in the OPEN message during the PCEP session initialization procedure

The OF codes defined in this document can be carried in the OF-list TLV of the OPEN object. If the OF-list TLV carries the OF codes, it means that the PCE is capable of implementing the corresponding objective functions. This information can be used for selecting a proper parent PCE when a child PCE wants to get a path that satisfies a certain Objective Function.

When a child PCE sends a PCReq to a peer PCE, which requires parental

activity and H-PCE capability flags TLV but which were not included in the session establishment procedure described above, the peer PCE SHOULD send a PCErr message to the child PCE and MUST specify the error-type=TBD8 (H-PCE error) and error-value=1 (H-PCE capability was not advertised) in the PCEP-ERROR object.

When a specific child PCE sends a PCReq to a peer PCE, that requires parental activity and the peer PCE does not want to act as the parent for it, the peer PCE SHOULD send a PCErr message to the child PCE and MUST specify the error-type=TBD8 (H-PCE error) and error-value=2 (Parent PCE capability cannot be provided) in the PCEP-ERROR object.

4.2. Procedure to Obtain Domain Sequence

If a child PCE only wants to get the domain sequence for a multi-domain path computation from a parent PCE, it can set the Domain Path Request bit in the H-PCE-FLAG TLV in the RP object carried in a PCReq message. The parent PCE which receives the PCReq message tries to compute a domain sequence for it (instead of the E2E path). If the domain path computation succeeds the parent PCE sends a PCRep message which carries the domain sequence in the Explicit Route Object (ERO) to the child PCE. Refer to [RFC7897] for more details about domain sub-objects in the ERO. Otherwise, it sends a PCReq message which carries the NO-PATH object to the child PCE.

5. Error Handling

A PCE that is capable of acting as a parent PCE might not be configured or willing to act as the parent for a specific child PCE. This fact could be determined when the child sends a PCReq that requires parental activity, and could result in a negative response in a PCEP Error (PCErr) message and indicate the hierarchy PCE error-type=TBD8 (H-PCE error) and suitable error-value. (Section 3.7)

Additionally, the parent PCE may fail to find the multi-domain path or domain sequence due to one or more of the following reasons:

- o A child PCE cannot find a suitable path to the egress;
- o The parent PCE does not hear from a child PCE for a specified time;
- o The Objective Functions specified in the path request cannot be met.

In this case, the parent PCE MAY need to send a negative path computation reply specifying the reason. This can be achieved by

including NO-PATH object in the PCRep message. Extension to NO-PATH object is needed to include the aforementioned reasons described in Section 3.7.

6. Manageability Considerations

General PCE and PCEP management considerations are discussed in [RFC4655] and [RFC5440]. There are additional management considerations for H-PCE which are described in [RFC6805], and repeated in this section.

The administrative entity responsible for the management of the parent PCEs must be determined for the following cases:

- o multi-domains (e.g., IGP areas or multiple ASes) within a single service provider network, the management responsibility for the parent PCE would most likely be handled by the service provider,
- o multiple ASes within different service provider networks, it may be necessary for a third party to manage the parent PCEs according to commercial and policy agreements from each of the participating service providers.

6.1. Control of Function and Policy

Control and function will need to be carefully managed in an H-PCE network. A child PCE will need to be configured with the address of its parent PCE. It is expected that there will only be one or two parents of any child.

The parent PCE also needs to be aware of the child PCEs for all child domains that it can see. This information is most likely to be configured (as part of the administrative definition of each domain).

Discovery of the relationships between parent PCEs and child PCEs do not form part of the hierarchical PCE architecture. Mechanisms that rely on advertising or querying PCE locations across domain or provider boundaries are undesirable for security, scaling, commercial, and confidentiality reasons. The specific behaviour of the child and parent PCE are described in the following sub-sections.

6.1.1. Child PCE

Support of the hierarchical procedure will be controlled by the management organization responsible for each child PCE. A child PCE must be configured with the address of its parent PCE in order for it to interact with its parent PCE. The child PCE must also be

authorized to peer with the parent PCE.

6.1.2. Parent PCE

The parent PCE MUST only accept path computation requests from authorized child PCEs. If a parent PCE receives requests from an unauthorized child PCE, the request SHOULD be dropped. This means that a parent PCE MUST be able to cryptographically authenticate requests from child PCEs.

Multi-party shared key authentication schemes are not recommended for inter-domain relationships because of the potential for impersonation and repudiation and for the operational difficulties should revocation be required.

The choice of authentication schemes to employ may be left to implementers of H-PCE and are not discussed further in this document.

6.1.3. Policy Control

It may be necessary to maintain a policy module on the parent PCE [RFC5394]. This would allow the parent PCE to apply commercially relevant constraints such as SLAs, security, peering preferences, and monetary costs.

It may also be necessary for the parent PCE to limit the end-to-end path selection by including or excluding specific domains based on commercial relationships, security implications, and reliability.

6.2. Information and Data Models

A MIB module for PCEP was published as RFC 7420 [RFC7420] that describes managed objects for modelling of PCEP communication. A YANG module for PCEP has also been proposed [I-D.ietf-pce-pcep-yang].

Additionally, H-PCE MIB module, or additional data model, will be required to report parent PCE and child PCE information, including:

- o parent PCE configuration and status,
- o child PCE configuration and information,
- o notifications to indicate session changes between parent PCEs and child PCEs, and
- o notification of parent PCE TED updates and changes.

6.3. Liveness Detection and Monitoring

The hierarchical procedure requires interaction with multiple PCEs. Once a child PCE requests an end-to-end path, a sequence of events occurs that requires interaction between the parent PCE and each child PCE. If a child PCE is not operational, and an alternate transit domain is not available, then the failure must be reported.

6.4. Verify Correct Operations

Verifying the correct operation of a parent PCE can be performed by monitoring a set of parameters. The parent PCE implementation should provide the following parameters monitored at the parent PCE:

- o number of child PCE requests,
- o number of successful hierarchical PCE procedures completions on a per-PCE-peer basis,
- o number of hierarchical PCE procedure completion failures on a per-PCE-peer basis, and
- o number of hierarchical PCE procedure requests from unauthorized child PCEs.

6.5. Requirements On Other Protocols

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

6.6. Impact On Network Operations

The hierarchical PCE procedure is a multiple-PCE path computation scheme. Subsequent requests to and from the child and parent PCEs do not differ from other path computation requests and should not have any significant impact on network operations.

7. IANA Considerations

IANA maintains the "Path Computation Element Protocol (PCEP) Numbers" registry. This document requests IANA actions to allocate code points for the protocol elements defined in this document.

7.1. PCEP TLV Type Indicators

IANA Manages the PCEP TLV code point registry (see [RFC5440]). This is maintained as the "PCEP TLV Type Indicators" sub-registry of the

"Path Computation Element Protocol (PCEP) Numbers" registry.

This document defines three new PCEP TLVs. IANA is requested to make the following allocation:

Type	TLV name	References
TBD1	H-PCE-CAPABILITY TLV	This I-D
TBD2	Domain-ID TLV	This I-D
TBD3	H-PCE-FLAG TLV	This I-D

7.2. H-PCE-CAPABILITY TLV Flags

This document requests that a new sub-registry, named "H-PCE-CAPABILITY TLV Flag Field", is created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the Flag field in the H-PCE-CAPABILITY TLV of the PCEP OPEN object.

New values are to be assigned by Standards Action [RFC8126]. Each bit should be tracked with the following qualities:

- o Bit number (counting from bit 0 as the most significant bit)
- o Capability description
- o Defining RFC

The following values are defined in this document:

Bit	Description	Reference
31	P (Parent PCE Request bit)	This I.D.

7.3. Domain-ID TLV Domain type

This document requests that a new sub-registry, named "Domain-ID TLV Domain type", is created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the Domain-Type field of the Domain-ID TLV. The allocation policy for this new sub-registry is IETF Review [RFC8126].

Value	Meaning
1	2-byte AS number
2	4-byte AS number
3	4-byte OSPF area ID
4	Variable length IS-IS area ID

7.4. H-PCE-FLAG TLV Flags

This document requests that a new sub-registry, named "H-PCE-FLAG TLV Flag Field", is created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the Flag field in the H-PCE-FLAGS TLV of the PCEP RP object. New values are to be assigned by Standards Action [RFC8126]. Each bit should be tracked with the following qualities:

- o Bit number (counting from bit 0 as the most significant bit)
- o Capability description
- o Defining RFC

The following values are defined in this document:

Bit	Description	Reference
31	S (Domain Sequence bit)	This I.D.
30	D (Disallow Domain Re-entry bit)	This I.D.

7.5. OF Codes

IANA maintains a registry of Objective Function (described in [RFC5541]) at the sub-registry "Objective Function". Three new Objective Functions have been defined in this document.

IANA is requested to make the following allocations:

Code Point	Name	Reference
TBD4	Minimum number of Transit Domains (MTD)	This I.D.
TBD5	Minimize number of Border Nodes (MBN)	This I.D.
TBD13	Minimize the number of Common Transit Domains (MCTD)	This I.D.

7.6. METRIC Types

IANA maintains one sub-registry for "METRIC object T field". Two new metric types are defined in this document for the METRIC object

(specified in [RFC5440]).

IANA is requested to make the following allocations:

Value	Description	Reference
TBD6	Domain Count metric	This I.D.
TBD7	Border Node Count metric	This I.D.

7.7. New PCEP Error-Types and Values

IANA maintains a registry of Error-Types and Error-values for use in PCEP messages. This is maintained as the "PCEP-ERROR Object Error Types and Values" sub-registry of the "Path Computation Element Protocol (PCEP) Numbers" registry.

IANA is requested to make the following allocations:

Error-Type	Meaning and error values	Reference
TBD8	H-PCE Error	This I.D.
	Error-value=1 H-PCE Capability not advertised	
	Error-value=2 Parent PCE Capability cannot be provided	
10	Reception of an invalid object [RFC5440]	
	Error-value=TBD15: Incompatible OF codes in H-PCE	This I.D.

7.8. New NO-PATH-VECTOR TLV Bit Flag

IANA maintains a sub-registry "NO-PATH-VECTOR TLV Flag Field" of bit flags carried in the PCEP NO-PATH-VECTOR TLV in the PCEP NO-PATH object as defined in [RFC5440]. IANA is requested to assign three new bit flag as follows:

Bit Number	Name Flag	Reference
TBD9	Destination Domain unknown	This I.D.
TBD10	Unresponsive child PCE(s)	This I.D.
TBD11	No available resource in one or more domain	This I.D.
TBD12	Destination is not found in the indicated domain.	This I.D.

7.9. SVEC Flag

IANA maintains a sub-registry "SVEC Object Flag Field" of bit flags carried in the PCEP SVEC object as defined in [RFC5440]. IANA is requested to assign one new bit flag as follows:

Bit Number	Name Flag	Reference
TBD14	Domain Diverse O-bit	This I.D.

8. Security Considerations

The hierarchical PCE procedure relies on PCEP and inherits the security considerations defined in [RFC5440]. As PCEP operates over TCP, it may also make use of TCP security mechanisms, such as TCP Authentication Option (TCP-AO) [RFC5925] or Transport Layer Security (TLS) [RFC8253].

Any multi-domain operation necessarily involves the exchange of information across domain boundaries. This may 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 H-PCE architecture is specifically designed to enable as much isolation of domain topology and capabilities information as is possible.

For further considerations of the security issues related to inter-AS path computation, see [RFC5376].

9. Contributing Authors

Xian Zhang
Huawei
EMail: zhang.xian@huawei.com

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

EMail: dhruv.ietf@gmail.com

10. Acknowledgements

The authors would like to thank Mike McBride, Kyle Rose, Roni Even for their detailed review, comments and suggestions which helped improve this document.

11. References

11.1. 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>>.
- [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>>.
- [RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", RFC 5541, DOI 10.17487/RFC5541, June 2009, <<https://www.rfc-editor.org/info/rfc5541>>.
- [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>>.

11.2. Informative References

- [RFC4105] Le Roux, J., Ed., Vasseur, J., Ed., and J. Boyle, Ed., "Requirements for Inter-Area MPLS Traffic Engineering", RFC 4105, DOI 10.17487/RFC4105, June 2005, <<https://www.rfc-editor.org/info/rfc4105>>.
- [RFC4216] Zhang, R., Ed. and J. Vasseur, Ed., "MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements", RFC 4216, DOI 10.17487/RFC4216, November 2005, <<https://www.rfc-editor.org/info/rfc4216>>.
- [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>>.

- [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>>.
- [RFC5152] Vasseur, JP., Ed., Ayyangar, A., Ed., and R. Zhang, "A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)", RFC 5152, DOI 10.17487/RFC5152, February 2008, <<https://www.rfc-editor.org/info/rfc5152>>.
- [RFC5376] Bitar, N., Zhang, R., and K. Kumaki, "Inter-AS Requirements for the Path Computation Element Communication Protocol (PCECP)", RFC 5376, DOI 10.17487/RFC5376, November 2008, <<https://www.rfc-editor.org/info/rfc5376>>.
- [RFC5394] Bryskin, I., Papadimitriou, D., Berger, L., and J. Ash, "Policy-Enabled Path Computation Framework", RFC 5394, DOI 10.17487/RFC5394, December 2008, <<https://www.rfc-editor.org/info/rfc5394>>.
- [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>>.
- [RFC5441] Vasseur, JP., Ed., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", RFC 5441, DOI 10.17487/RFC5441, April 2009, <<https://www.rfc-editor.org/info/rfc5441>>.
- [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>>.
- [RFC7399] Farrel, A. and D. King, "Unanswered Questions in the Path Computation Element Architecture", RFC 7399, DOI 10.17487/RFC7399, October 2014,

<<http://www.rfc-editor.org/info/rfc7399>>.

- [RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <<https://www.rfc-editor.org/info/rfc7420>>.
- [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <<https://www.rfc-editor.org/info/rfc7752>>.
- [RFC7897] Dhody, D., Palle, U., and R. Casellas, "Domain Subobjects for the Path Computation Element Communication Protocol (PCEP)", RFC 7897, DOI 10.17487/RFC7897, June 2016, <<https://www.rfc-editor.org/info/rfc7897>>.
- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <<https://www.rfc-editor.org/info/rfc8126>>.
- [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>>.
- [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-11 (work in progress), March 2019.
- [I-D.ietf-pce-stateful-hpce] Dhody, D., Lee, Y., Ceccarelli, D., Shin, J., King, D., and O. Dios, "Hierarchical Stateful Path Computation Element (PCE).", draft-ietf-pce-stateful-hpce-07 (work in progress), April 2019.
- [I-D.dhodylee-pce-pcep-ls] Dhody, D., Lee, Y., and D. Ceccarelli, "PCEP Extension for Distribution of Link-State and TE Information.", draft-dhodylee-pce-pcep-ls-13 (work in progress), February 2019.

Authors' Addresses

Fatai Zhang
Huawei
Huawei Base, Bantian, Longgang District
Shenzhen 518129
China

EMail: zhangfatai@huawei.com

Quintin Zhao
Huawei
125 Nagog Technology Park
Acton, MA 01719
USA

EMail: quintin.zhao@huawei.com

Oscar Gonzalez de Dios
Telefonica
Don Ramon de la Cruz 82-84
Madrid 28045
Spain

EMail: oscar.gonzalezdedios@telefonica.com

Ramon Casellas
CTTC
Av. Carl Friedrich Gauss n.7
Barcelona, Castelldefels
Spain

EMail: ramon.casellas@cttc.es

Daniel King
Old Dog Consulting
UK

EMail: daniel@olddog.co.uk

Appendix

A1. Implementation Status

Zhang, et al.

Expires December, 2019

[Page 28]

The H-PCE architecture and protocol procedures describe in this I-D were implemented and tested for a variety of optical research applications.

The Appendix should be removed before publication.

A1.1. Inter-layer traffic engineering with H-PCE

This work was led by:

- o Ramon Casellas [ramon.casellas@cttc.es]
- o Centre Tecnologic de Telecomunicacions de Catalunya (CTTC)

The H-PCE instances (parent and child) were multi-threaded asynchronous processes. Implemented in C++11, using C++ Boost Libraries. The targeted system used to deploy and run H-PCE applications was a POSIX system (Debian GNU/Linux operating system).

Some parts of the software may require a Linux Kernel, the availability of a Routing Controller running collocated in the same host and the usage of libnetfilter / libipq and GNU/Linux firewalling capabilities. Most of the functionality, including algorithms is done by means of plugins (e.g., as shared libraries or .so files in Unix systems).

The CTTC PCE supports the H-PCE architecture, but also supports stateful PCE with active capabilities, as an OpenFlow controller, and has dedicated plugins to support monitoring, BRPC, P2MP, path keys, back end PCEs. Management of the H-PCE entities was supported via HTTP and CLI via Telnet.

Further details of the H-PCE prototyping and experimentation can be found in the following scientific papers:

R. Casellas, R. Martinez, R. Munoz, L. Liu, T. Tsuritani, I. Morita, "Inter-layer traffic engineering with hierarchical-PCE in MPLS-TP over wavelength switched optical networks" , Optics Express, Vol. 20, No. 28, December 2012.

R. Casellas, R. Martinez, R. Munoz, L. Liu, T. Tsuritani, I. Morita, M. Msurusawa, "Dynamic virtual link mesh topology aggregation in multi-domain translucent WSON with hierarchical-PCE", Optics Express Journal, Vol. 19, No. 26, December 2011.

R. Casellas, R. Munoz, R. Martinez, R. Vilalta, L. Liu, T. Tsuritani, I. Morita, V. Lopez, O. Gonzalez de Dios, J. P. Fernandez-Palacios, "SDN based Provisioning Orchestration of

OpenFlow/GMPLS Flexi-grid Networks with a Stateful Hierarchical PCE", in Proceedings of Optical Fiber Communication Conference and Exposition (OFC), 9-13 March, 2014, San Francisco (EEUU).
Extended Version to appear in Journal Of Optical Communications and Networking January 2015

F. Paolucci, O. Gonzalez de Dios, R. Casellas, S. Duhovnikov, P. Castoldi, R. Munoz, R. Martinez, "Experimenting Hierarchical PCE Architecture in a Distributed Multi-Platform Control Plane Testbed" , in Proceedings of Optical Fiber Communication Conference and Exposition (OFC) and The National Fiber Optic Engineers Conference (NFOEC), 4-8 March, 2012, Los Angeles, California (USA).

R. Casellas, R. Martinez, R. Munoz, L. Liu, T. Tsuritani, I. Morita, M. Tsurusawa, "Dynamic Virtual Link Mesh Topology Aggregation in Multi-Domain Translucent WSON with Hierarchical-PCE", in Proceedings of 37th European Conference and Exhibition on Optical Communication (ECOC 2011), 18-22 September 2011, Geneve (Switzerland).

R. Casellas, R. Munoz, R. Martinez, "Lab Trial of Multi-Domain Path Computation in GMPLS Controlled WSON Using a Hierarchical PCE", in Proceedings of OFC/NFOEC Conference (OFC2011), 10 March 2011, Los Angeles (USA).

A1.2. Telefonica Netphony (Open Source PCE)

The Telefonica Netphony PCE is an open source Java-based implementation of a Path Computation Element, with several flavours, and a Path Computation Client. The PCE follows a modular architecture and allows to add customized algorithms. The PCE has also stateful and remote initiation capabilities. In current version, three components can be built, a domain PCE (aka child PCE), a parent PCE (ready for the H-PCE architecture) and a PCC (path computation client).

This work was led by:

- o Oscar Gonzalez de Dios [oscar.gonzalezdedios@telefonica.com]
- o Victor Lopez Alvarez [victor.lopezalvarez@telefonica.com]
- o Telefonica I+D, Madrid, Spain

The PCE code is publicly available in a GitHub repository:

- o <https://github.com/telefonicaid/netphony-pce>

The PCEP protocol encodings are located in the following repository:

- o <https://github.com/telefonicaid/netphony-network> protocols

The traffic engineering database and a BGP-LS speaker to fill the database is located in:

- o <https://github.com/telefonicaid/netphony-topology>

The parent and child PCE are multi-threaded java applications. The path computation uses the jgrapht free Java class library (0.9.1) that provides mathematical graph-theory objects and algorithms. Current version of netphony PCE runs on java 1.7 and 1.8, and has been tested in GNU/Linux, Mac OS-X and Windows environments. The management of the parent and domain PCEs is supported through CLI via Telnet, and configured via XML files.

Further details of the netphony H-PCE prototyping and experimentation can be found in the following research papers:

- o O. Gonzalez de Dios, R. Casellas, F. Paolucci, A. Napoli, L. Gifre, A. Dupas, E. Hugues-Salas, R. Morro, S. Belotti, G. Meloni, T. Rahman, V.P Lopez, R. Martinez, F. Fresi, M. Bohn, S. Yan, L. Velasco, . Layec and J. P. Fernandez-Palacios: Experimental Demonstration of Multivendor and Multidomain EON With Data and Control Interoperability Over a Pan-European Test Bed, in Journal of Lightwave Technology, Dec. 2016, Vol. 34, Issue 7, pp. 1610-1617.
- o O. Gonzalez de Dios, R. Casellas, R. Morro, F. Paolucci, V. Lopez, R. Martinez, R. Munoz, R. Villalta, P. Castoldi: "Multi-partner Demonstration of BGP-LS enabled multi-domain EON, in Journal of Optical Communications and Networking, Dec. 2015, Vol. 7, Issue 12, pp. B153-B162.
- o F. Paolucci, O. Gonzalez de Dios, R. Casellas, S. Duhovnikov, P. Castoldi, R. Munoz, R. Martinez, "Experimenting Hierarchical PCE Architecture in a Distributed Multi-Platform Control Plane Testbed" , in Proceedings of Optical Fiber Communication Conference and Exposition (OFC) and The National Fiber Optic Engineers Conference (NFOEC), 4-8 March, 2012, Los Angeles, California (USA).

A1.3. H-PCE Proof of Concept developed by Huawei

Huawei developed this H-PCE on the Huawei Versatile Routing Platform (VRP) to experiment with the hierarchy of PCE. Both end to end path computation as well as computation for domain-sequence are supported.

This work was led by:

- o Udayasree Pallee [udayasreereddy@gmail.com]
- o Dhruv Dhody [dhruv.ietf@gmail.com]
- o Huawei Technologies, Bangalore, India

Further work on stateful H-PCE [I-D.ietf-pce-stateful-hpce] is being carried out on ONOS.

PCE Working Group
Internet-Draft
Intended status: Informational
Expires: December 20, 2018

D. Dhody
Y. Lee
Huawei Technologies
D. Ceccarelli
Ericsson
J. Shin
SK Telecom
D. King
Lancaster University
O. Gonzalez de Dios
Telefonica I+D
June 18, 2018

Hierarchical Stateful Path Computation Element (PCE).
draft-ietf-pce-stateful-hpce-05

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

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."

Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 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

1. Introduction	3
1.1. Requirements Language	3
2. Terminology	3
3. Hierarchical Stateful PCE	4
3.1. Passive Operations	4
3.2. Active Operations	7
3.3. PCE Initiation Operation	8
3.3.1. Per Domain Stitched LSP	8
4. Other Considerations	10
4.1. Applicability to Inter-Layer	10
4.2. Applicability to ACTN	11
5. Security Considerations	12
6. Manageability Considerations	12
6.1. Control of Function and Policy	12
6.2. Information and Data Models	12
6.3. Liveness Detection and Monitoring	12
6.4. Verify Correct Operations	12
6.5. Requirements On Other Protocols	12
6.6. Impact On Network Operations	12
7. IANA Considerations	12
8. Acknowledgments	12
9. References	12
9.1. Normative References	12
9.2. Informative References	13
Appendix A. Contributor Addresses	14
Authors' Addresses	14

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 (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 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.

The initial section of the document focuses on end to end (E2E) inter-domain TE LSP. Section 3.3.1 describe the operations for the Per Domain LSP that could be stitched.

1.1. 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.

2. Terminology

The terminology is as per [RFC4655], [RFC5440], [RFC6805], [RFC8231], and [RFC8281].

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.

[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 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 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 should be used in the OPEN message to advertise the H-PCE capability. [RFC8231] defines the stateful PCE capability TLV. The presence of both TLVs represent the support for stateful H-PCE operations as described in this document.

[I-D.litkowski-pce-state-sync] describes the procedures to allow a 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 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 [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 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.

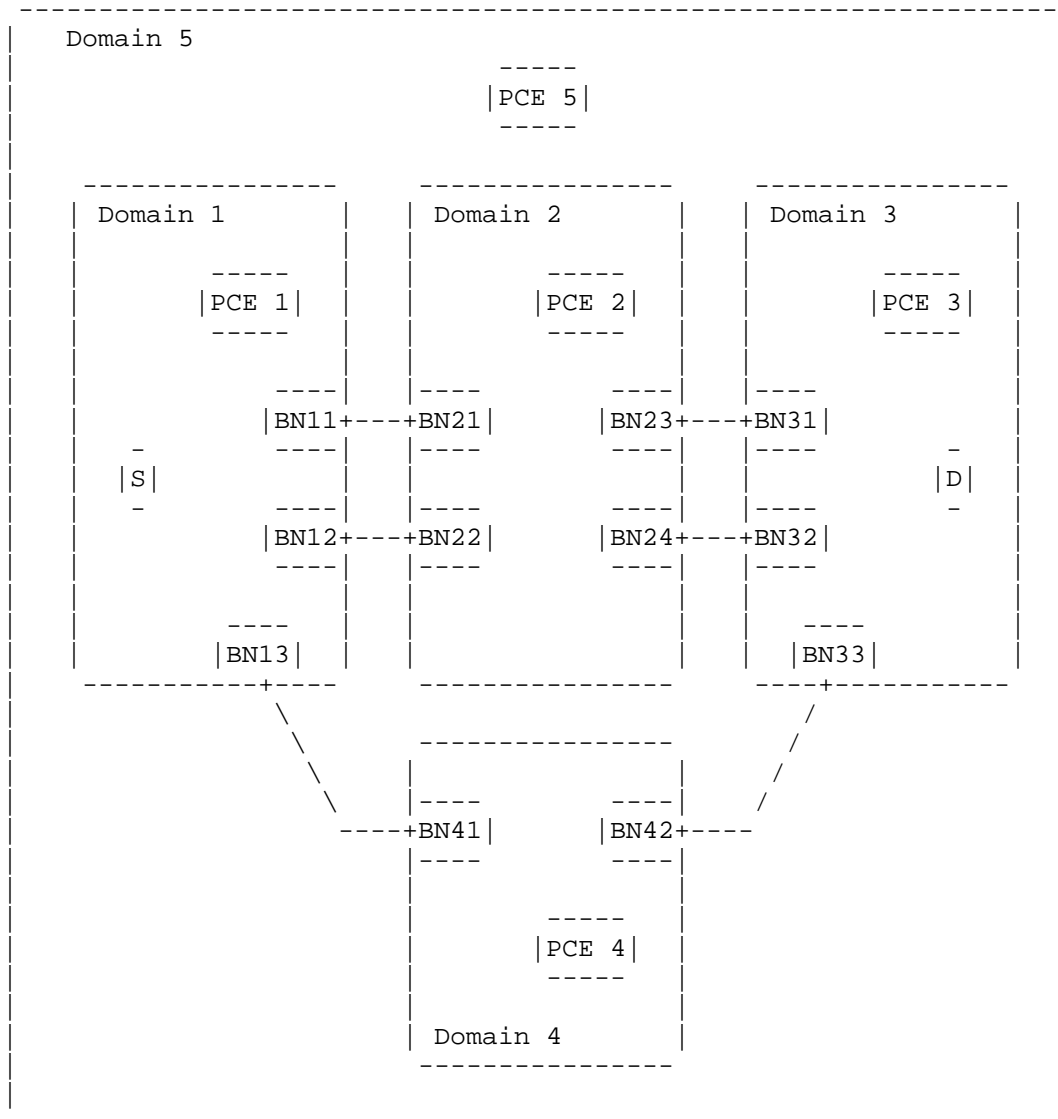


Figure 1: Sample Hierarchical Domain Topology

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).

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 Path Computation Client (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 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

[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 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).

The Ingress LSR (PCC) generates the PLSP-ID for the LSP and inform the C-PCE, which is propagated to the P-PCE as described in [I-D.litkowski-pce-state-sync]

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 additional steps are also initially performed, for the Per Domain stitched 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 -

- 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)

- (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).
- (12) 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).
```

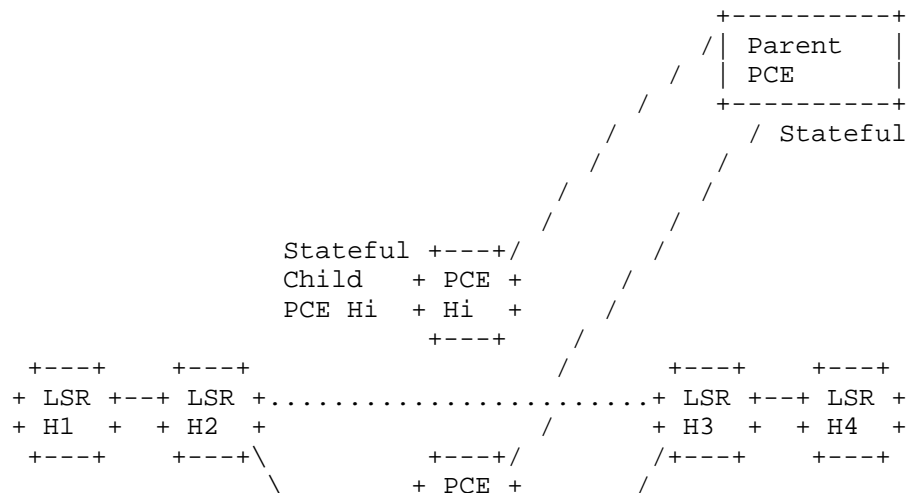
Additionally:

- (20) Once P-PCE receives report of each per-domain LSP, it should use 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). It is also possible to stitch the per-domain LSP at the same time as the per-domain LSPs are initiated as defined in [I-D.dugeon-pce-stateful-interdomain].

#### 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.



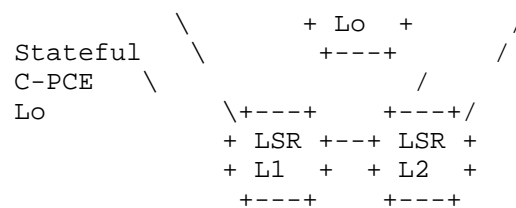


Figure 2: Sample Inter-Layer Topology

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 Provisioning 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.ietf-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. Other Considerations

##### 5.1. 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.

##### 5.2. 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.

## 6. 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 SHOULD continue to use this.

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

Thus securing the PCEP session (between the P-PCE and the C-PCE) using mechanism like TCP Authentication Option (TCP-AO) [RFC5925] or Transport Layer Security (TLS) [RFC8253] is RECOMMENDED.

## 7. 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.

### 7.1. 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.

## 7.2. 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] can be extended to include details stateful H-PCE.

## 7.3. 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].

## 7.4. 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].

## 7.5. Requirements On Other Protocols

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

## 7.6. 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.

## 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>>.
- [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>>.

## 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>>.
- [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, <<http://www.rfc-editor.org/info/rfc5520>>.
- [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>>.

- [RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <<http://www.rfc-editor.org/info/rfc5925>>.
- [RFC8051] Zhang, X., Ed. and I. Minei, Ed., "Applicability of a Stateful Path Computation Element (PCE)", RFC 8051, DOI 10.17487/RFC8051, January 2017, <<http://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>>.
- [I-D.ietf-teas-actn-framework]  
Ceccarelli D. and Y. Lee, "Framework for Abstraction and Control of Transport Networks", draft-ietf-teas-actn-framework-15 (work in progress), May 2018.
- [I-D.ietf-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-ietf-pce-applicability-actn-06 (work in progress), June 2018.
- [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-03 (work in progress), April 2018.
- [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-05 (work in progress), June 2018.
- [I-D.ietf-pce-pcep-yang]

Dhody, D., Hardwick, J., Beeram, V., and j. jeffrant@gmail.com, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", draft-ietf-pce-pcep-yang-08 (work in progress), June 2018.

[I-D.dugeon-pce-stateful-interdomain]

Dugeon, O. and J. Meuric, "PCEP Extension for Stateful Inter-Domain Tunnels", draft-dugeon-pce-stateful-interdomain-00 (work in progress), October 2017.

## Appendix A. Contributor Addresses

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

EMail: s.avantika.avantika@gmail.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: udayasreereddy@gmail.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](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

Email: [jongyoon.shin@sk.com](mailto:jongyoon.shin@sk.com)

Daniel King  
Lancaster University  
UK

Email: [d.king@lancaster.ac.uk](mailto: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](mailto:ogondio@tid.es)

PCE Working Group  
Internet-Draft  
Intended status: Informational  
Expires: April 23, 2020

D. Dhody  
Huawei Technologies  
Y. Lee  
SKKU  
D. Ceccarelli  
Ericsson  
J. Shin  
SK Telecom  
D. King  
Lancaster University  
October 21, 2019

Hierarchical Stateful Path Computation Element (PCE)  
draft-ietf-pce-stateful-hpce-15

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 Path (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 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 April 23, 2020.

## 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 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

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

|      |                                                            |    |
|------|------------------------------------------------------------|----|
| 6.   | Other Considerations . . . . .                             | 17 |
| 6.1. | Applicability to Inter-Layer Traffic Engineering . . . . . | 18 |
| 6.2. | Scalability Considerations . . . . .                       | 18 |
| 6.3. | Confidentiality . . . . .                                  | 19 |
| 7.   | IANA Considerations . . . . .                              | 19 |
| 8.   | Acknowledgments . . . . .                                  | 19 |
| 9.   | References . . . . .                                       | 19 |
| 9.1. | Normative References . . . . .                             | 19 |
| 9.2. | Informative References . . . . .                           | 20 |
|      | Contributors . . . . .                                     | 22 |
|      | Authors' Addresses . . . . .                               | 23 |

## 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 constrained paths for 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. The C-PCE 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. 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) 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 end-to-end inter-domain path computation and setup is 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 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 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 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 PCReq and PCRep messages between the P-PCE and C-PCEs. When the CNC requests VN provisioning, the MDSC decomposes 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 options for inter-domain 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 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.

A stateful P-PCE has to be aware of the inter-domain 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 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 [I-D.ietf-pce-stateful-path-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 and 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 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). 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 a 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 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, so a Label Switching Router (LSR), as a PCC, could delegate control to a PCE, which 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.

[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. 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 [I-D.ietf-pce-hierarchy-extensions] would 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 [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 (PCEP-specific identifier for the LSP, as per [RFC8231]) 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 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.

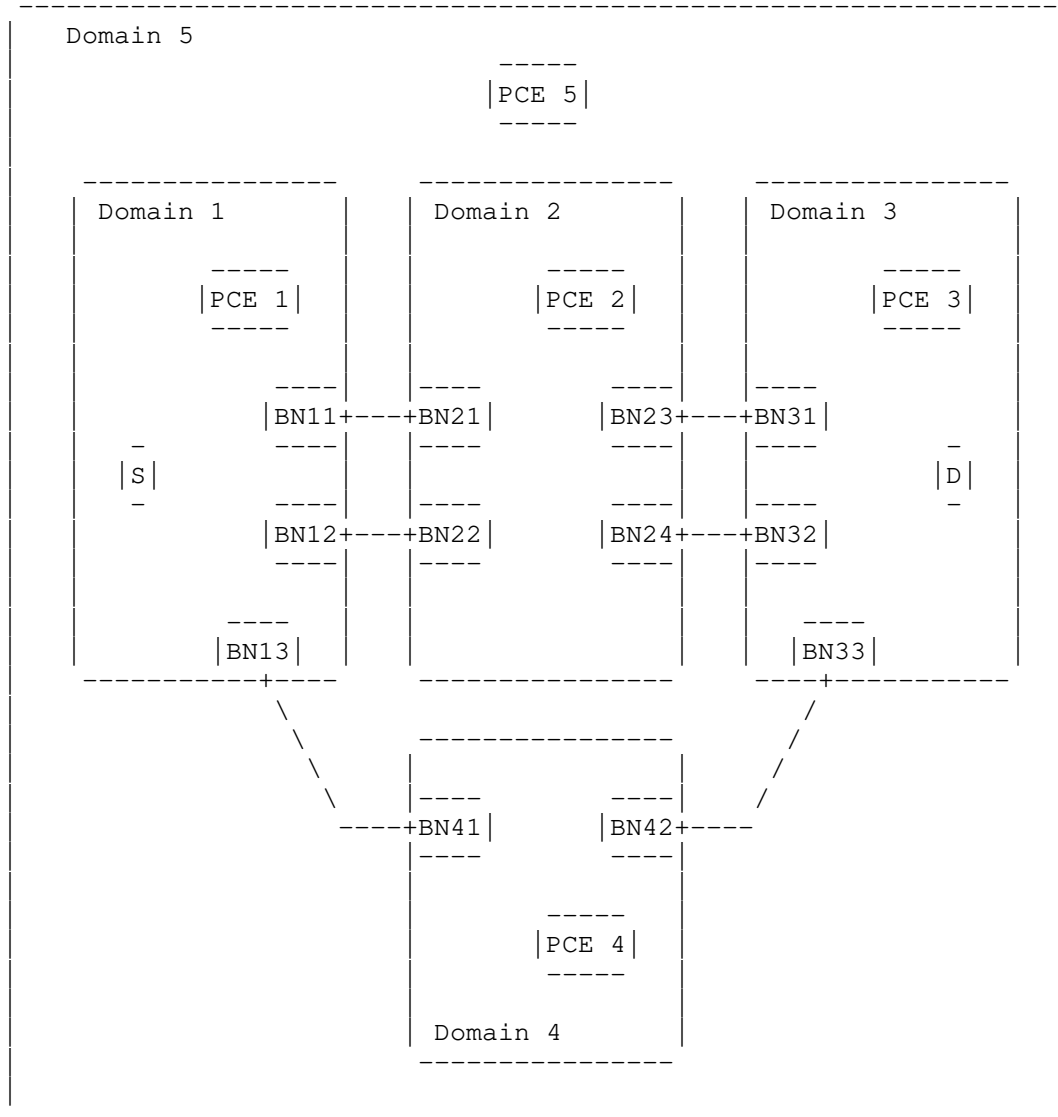


Figure 1: Hierarchical Domain Topology Example

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 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 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 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 an inter-domain LSP in Hierarchical PCE architecture, the initiation operations can be carried out at the P-PCE. In which case after the P-PCE finishes the E2E path computation, it can send the PCInitiate message to the C-PCE (the Ingress domain PCE), 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 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 (Hierarchical Domain Topology Example):

- (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 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 [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 (i.e. the highest security mechanism available for PCEP) 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 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 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 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 "ietf-pcep" PCEP YANG module is specified in [I-D.ietf-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 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.

As described in Section 3, a PCEP speaker includes both the H-PCE Capability TLV [I-D.ietf-pce-hierarchy-extensions] 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 [I-D.ietf-pce-hierarchy-extensions] features. This PCEP speaker will also include both the TLVs and 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 get further propagated (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 as per [I-D.ietf-pce-enhanced-errors]. Enhanced features such as 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.

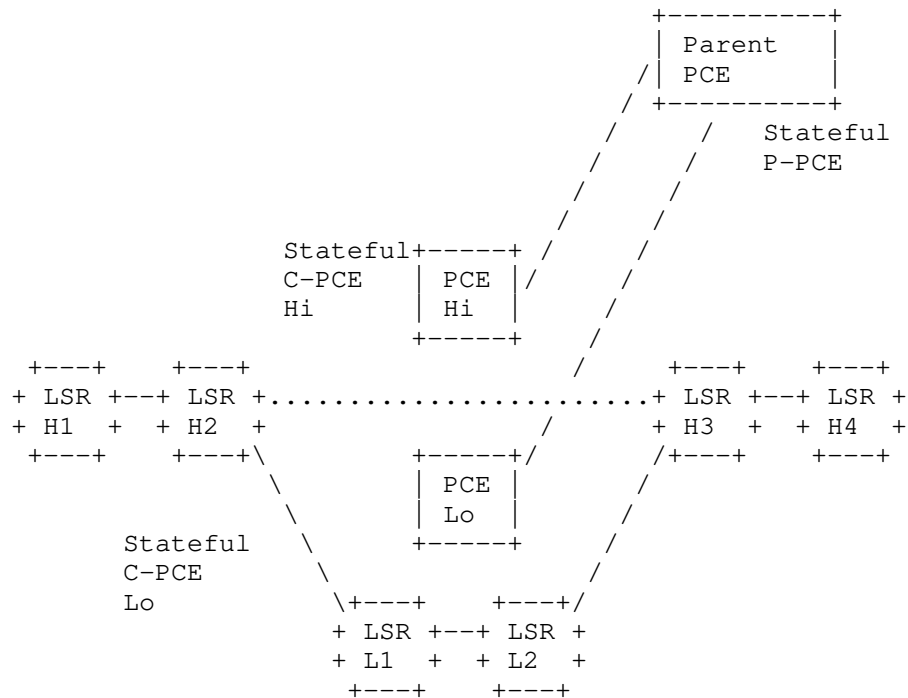


Figure 2: Sample Inter-Layer Topology

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. 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

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.

Thanks to Barry Leiba, Martin Vigoureux, Benjamin Kaduk, and Roman Danyliw for IESG 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>>.

- [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, <<http://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>>.

## 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, 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>>.
- [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-11 (work in progress), October 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.
- [I-D.ietf-pce-stateful-path-protection]  
Ananthakrishnan, H., Sivabalan, S., Barth, C., Minei, I.,  
and M. Negi, "PCEP Extensions for Associating Working and  
Protection LSPs with Stateful PCE", draft-ietf-pce-  
stateful-path-protection-11 (work in progress), September  
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  
SKKU

EMail: younglee.tx@gmail.com

Daniele Ceccarelli  
Ericsson  
Torshamnsgatan, 48  
Stockholm  
Sweden

EMail: 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

EMail: jongyoon.shin@sk.com

Daniel King  
Lancaster University  
UK

EMail: d.king@lancaster.ac.uk

PCE Working Group  
Internet-Draft  
Intended Status: Standards track  
Expires: December 23, 2018

Y. Lee  
D. Dhody  
X. Zhang  
Huawei Technologies  
D. Ceccarelli  
Ericsson  
June 21, 2018

PCEP Extensions for Establishing Relationships Between Sets of LSPs  
and Virtual Networks  
draft-leedhody-pce-vn-association-05

## Abstract

This document describes how to extend Path Computation Element (PCE) Communication Protocol (PCEP) association mechanism introduced by the PCEP Association Group specification, to further associate sets of LSPs with a higher-level structure such as a virtual network (VN) requested by clients or applications. This extended association mechanism can be used to facilitate virtual network control using PCE architecture.

## Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at  
<https://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at  
<https://www.ietf.org/shadow.html>

This Internet-Draft will expire on September 01, 2018.

#### Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://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

|                                             |   |
|---------------------------------------------|---|
| 1. Introduction.....                        | 2 |
| 1.1. Requirements Language.....             | 3 |
| 2. Terminology.....                         | 4 |
| 3. Operation Overview.....                  | 4 |
| 4. Extensions to PCEP.....                  | 4 |
| 5. Applicability to H-PCE architecture..... | 6 |
| 6. Security Considerations.....             | 7 |
| 7. IANA Considerations.....                 | 7 |
| 7.1. Association Object Type Indicator..... | 7 |
| 7.2. PCEP TLV Type Indicator.....           | 8 |
| 7.3. PCEP Error.....                        | 8 |
| 8. References.....                          | 8 |
| 8.1. Normative References.....              | 8 |
| 8.2. Informative References.....            | 9 |
| Author's Addresses.....                     | 9 |

#### 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.

[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.

[I-D.ietf-pce-association-group] introduces a generic mechanism to create a grouping of LSPs. This grouping can then be used to define association between sets of LSPs or between a set of LSPs and a set of attributes.

[ACTN-REQ] describes various Virtual Network (VN) operations initiated by a customer/application. In this context, there is a need for associating a set of LSPs with a VN "construct" to facilitate VN operations in PCE architecture. This association allows the PCEs to identify which LSPs belong to a certain VN. The PCE could then use this association to optimize all LSPs belonging to the VN together. The PCE could further take VN specific actions on the LSPs such as relaxation of constraints, policy actions, setting default behavior etc.

[I-D.ietf-pce-applicability-actn] examines the PCE and ACTN architecture and describes how the PCE architecture is applicable to ACTN. [RFC6805] and [I-D.ietf-pce-stateful-hpce] describes a hierarchy of stateful PCEs with Parent PCE coordinating multi-domain path computation function between Child PCE(s) and thus making it the base for PCE applicability for ACTN. In this text child PCE would be same as Provisioning Network Controller (PNC), and the parent PCE as Multi-domain Service Coordinator (MDSC) [ACTN-FWK].

This document specifies a PCEP extension to associate a set of LSPs based on Virtual Network (VN) (or customer). A Virtual Network (VN) is a customer view of the TE network. Depending on the agreement between client and provider various VN operations and VN views are possible as described in [ACTN-FWK].

## 1.1. 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.

## 2. Terminology

The terminology is as per [RFC4655], [RFC5440], [RFC6805], [RFC8231] and [ACTN-FWK]..

## 3. Operation Overview

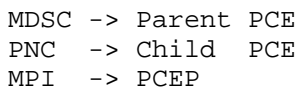
As per [I-D.ietf-pce-association-group], LSPs are associated with other LSPs with which they interact by adding them to a common association group.

An association group based on VN is useful for various optimizations that should be applied by considering all the LSPs in the association. This includes, but not limited to -

- o Path Computation: When computing path for a LSP, the impact of this LSP, on the other LSPs belonging to the same VN is useful to analyze. The aim would be optimize overall VN and all LSPs, rather than a single LSP. Also, the optimization criteria such as minimize the load of the most loaded link (MLL) [RFC5541] and other could be applied for all the LSP belonging to the same VN, identified by the VN association.

- o Path Re-Optimization: The child PCE or the parent PCE would like to use advanced path computation algorithm and optimization technique that consider all the LSPs belonging to a VN/customer and optimize them all together during the re-optimization.

This association is useful in PCEP session between parent PCE (MDSC) and child PCE (PNC).



single VNAG. If an implementation encounters more than one VNAG, it MUST consider the first occurrence and ignore the others.

[I-D.ietf-pce-association-group] specify the mechanism for the capability advertisement of the association types supported by a PCEP speaker by defining a ASSOC-Type-List TLV to be carried within an OPEN object. This capability exchange for the association type described in this document (i.e. VN Association Type) MUST be done before using the policy association. Thus the PCEP speaker MUST include the VN Association Type (TBD1) in the ASSOC-Type-List TLV before using the VNAG in the PCEP messages.

This Association-Type is dynamic in nature and created by the Parent PCE (MDSC) for the LSPs belonging to the same VN or customer. These associations are conveyed via PCEP messages to the PCEP peer. Operator-configured Association Range MUST NOT be set for this association-type and MUST be ignored.

#### 4. Extensions to PCEP

The format of VNAG is as per the ASSOCIATION object [I-D.ietf-pce-association-group].

This document defines one mandatory TLV "VIRTUAL-NETWORK-TLV" and one new optional TLV "VENDOR-INFORMATION-TLV"; apart from this TLV, VENDOR-INFORMATION-TLV can be used to carry arbitrary vendor specific information.

- o VIRTUAL-NETWORK-TLV: Used to communicate the VN Identifier.
- o VENDOR-INFORMATION-TLV: Used to communicate arbitrary vendor specific behavioral information, described in [RFC7470].

The format of VIRTUAL-NETWORK-TLV is as follows.

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| | Length (variable) |
+-----+-----+-----+-----+-----+-----+-----+-----+
| |
// Virtual Network Name //
| |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Figure 1: The VIRTUAL-NETWORK-TLV formats

Type: TBD2 (to be allocated by IANA)

Length: Variable Length

Virtual Network Name (variable): an unique symbolic name for the VN. It SHOULD be a string of printable ASCII characters, without a NULL terminator. The VN name is a human-readable string that identifies a VN. The VN name MUST remain constant throughout an LSP's lifetime, which may span across multiple consecutive PCEP sessions and/or PCC restarts. The VN name MAY be specified by an operator or auto-generated by the PCEP speaker.

The VIRTUAL-NETWORK-TLV MUST be included in VNAG object. If a PCEP speaker receives the VNAG object without the VIRTUAL-NETWORK-TLV, it MUST send a PCErr message with Error-Type=6 (mandatory object missing) and Error-Value=TBD3 (VIRTUAL-NETWORK-TLV missing) and close the session.

The format of VENDOR-INFORMATION-TLV is defined in [RFC7470].

## 5. Applicability to H-PCE architecture

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 architecture, the parent PCE is used to compute a multi-domain path based on the domain connectivity information. A child 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.

[I-D.ietf-pce-stateful-hpce] introduces general considerations for stateful PCE(s) in hierarchical PCE architecture. In particular, the behavior changes and additions to the existing stateful PCE mechanisms in the context of a H-PCE architecture.

In Stateful H-PCE architecture, the Parent PCE receives a virtual network creation request by its client over its Northbound API. This VN is uniquely identified by an Association ID in VNAG as well as the VIRTUAL-NETWORK name. This VN may comprise multiple LSPs in the network in a single domain or across multiple domains.

As the Parent PCE computes the optimum E2E paths for each tunnel in VN, it MUST associate each LSP with the VN to which it belongs. Parent PCE sends a PCInitiate Message with this association

information in the VNAG Object (See Section 4 for details). This in effect binds an LSP that is to be instantiated at the child PCE with the VN.

Whenever changes occur with the instantiated LSP in a domain network, the domain child PCE reports the changes using a PCRpt Message in which the VNAG Object indicates the relationship between the LSP and the VN.

Whenever an update occurs with VNs in the Parent PCE (via the client's request), the parent PCE sends an PCUpd Message to inform each affected child PCE of this change.

The Child PCE could then use this association to optimize all LSPs belonging to the same VN association together. The Child PCE could further take VN specific actions on the LSPs such as relaxation of constraints, policy actions, setting default behavior etc. The parent PCE could also maintain all E2E LSP or per-domain path segments under a single VN association.

## 6. Security Considerations

This document defines one new type for association, which do not add any new security concerns beyond those discussed in [RFC5440], [RFC8231] and [I-D.ietf-pce-association-group] in itself.

Some deployments may find VN associations and their implications as extra sensitive and thus should employ suitable PCEP security mechanisms like TCP-AO or [RFC8253].

## 7. IANA Considerations

### 7.1. Association Object Type Indicator

This document defines a new association type, originally defined in [I-D.ietf-pce-association-group], for path protection. IANA is requested to make the assignment of a new value for the sub-registry "ASSOCIATION Type Field" (request to be created in [I-D.ietf-pce-association-group]), as follows:

| Value | Name                | Reference   |
|-------|---------------------|-------------|
| TBD1  | VN Association Type | [This I.D.] |

### 7.2. PCEP TLV Type Indicator

This document defines a new TLV for carrying additional information of LSPs within a path protection association group. IANA is

requested to make the assignment of a new value for the existing "PCEP TLV Type Indicators" registry as follows:

| Value | Name                | Reference   |
|-------|---------------------|-------------|
| TBD2  | VIRTUAL-NETWORK-TLV | [This I.D.] |

### 7.3. PCEP Error

This document defines new Error-Type and Error-Value related to path protection association. IANA is requested to allocate new error values within the "PCEP-ERROR Object Error Types and Values" sub-registry of the PCEP Numbers registry, as follows:

| Error-Type | Meaning                                                                               |
|------------|---------------------------------------------------------------------------------------|
| 6          | Mandatory Object missing<br>Error-value=TBD3: VIRTUAL-NETWORK TLV missing [This I.D.] |

## 8. Manageability Considerations

### 8.1. Control of Function and Policy

An operator MUST BE allowed to mark LSPs that belong to the same VN. This could also be done automatically based on the VN configuration.

### 8.2. Information and Data Models

The PCEP YANG module [I-D.ietf-pce-pcep-yang] should support the association between LSPs including VN association.

### 8.3. 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].

### 8.4. Verify Correct Operations

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

### 8.5. Requirements On Other Protocols

Mechanisms defined in this document do not imply any new requirements

on other protocols.

## 8.6. Impact On Network Operations

Mechanisms defined in this document do not have any impact on network operations in addition to those already listed in [RFC5440].

## 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.
- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, May 2017.
- [RFC8231] E. Crabbe, I. Minei, J. Medved, and R. Varga, "PCEP Extensions for Stateful PCE", RFC 8231, September 2017.
- [RFC8281] E. Crabbe, et. al., "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", RFC 8281, December 2017.
- [I-D.ietf-pce-association-group] I, Minei, Ed., "PCEP Extensions for Establishing Relationships Between Sets of LSPs", draft-ietf-pce-association-group, work in progress.

### 9.2. Informative References

- [RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.
- [RFC6805] A. Farrel and D. King, "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS", RFC 6805, November 2012.
- [I-D.ietf-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-ietf-pce-applicability-actn, work-in-progress.
- [I-D.ietf-pce-stateful-hpce] Dhody, D. and Lee, Y., "Hierarchical

Stateful Path Computation Element (PCE)",  
draft-ietf-pce-stateful-hpce, work in progress.

- [ACTN-REQ] Y. Lee, D. Dhody, S. Belotti, K. Pithewan, and D. Ceccarelli, "Requirements for Abstraction and Control of TE Networks", draft-ietf-teas-actn-requirements, work in progress.
- [ACTN-FWK] Ceccarelli D. and Y. Lee, "Framework for Abstraction and Control of Transport Networks", draft-ietf-teas-actn-framework (work in progress).
- [RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", RFC 5541, DOI 10.17487/RFC5541, June 2009, <<http://www.rfc-editor.org/info/rfc5541>>.
- [RFC7470] Zhang, F. and A. Farrel, "Conveying Vendor-Specific Constraints in the Path Computation Element Communication Protocol", RFC 7470, DOI 10.17487/RFC7470, March 2015, <<http://www.rfc-editor.org/info/rfc7470>>.
- [RFC8051] Zhang, X., Ed. and I. Minei, Ed., "Applicability of a Stateful Path Computation Element (PCE)", RFC 8051, DOI 10.17487/RFC8051, January 2017, <<http://www.rfc-editor.org/info/rfc8051>>.
- [RFC8253] Lopez, D., Dios, O., Wu, W., and D. Dhody, "Secure Transport for PCEP", RFC 8253, October 2017 .
- [I-D.ietf-pce-pcep-yang]  
Dhody, D., Hardwick, J., Beeram, V., and j. jeffrant@gmail.com, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", draft-ietf-pce-pcep-yang (work in progress).

Author's Addresses

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

Email: [leeyoung@huawei.com](mailto:leeyoung@huawei.com)

Dhruv Dhody (Editor)  
Huawei Technologies  
Divyashree Technopark, Whitefield  
Bangalore, Karnataka 560066  
India

Email: [dhruv.ietf@gmail.com](mailto:dhruv.ietf@gmail.com)

Xian Zhang  
Huawei Technologies  
China

Email: [zhang.xian@huawei.com](mailto:zhang.xian@huawei.com)

Daniele Ceccarelli  
Ericsson  
Torshamnsgatan, 48  
Stockholm,  
Sweden

Email: [daniele.ceccarelli@ericsson.com](mailto:daniele.ceccarelli@ericsson.com)

PCE Working Group  
Internet-Draft  
Intended Status: Standards track  
Expires: December 21, 2019

Y. Lee  
Futurewei  
X. Zhang  
Huawei Technologies  
D. Ceccarelli  
Ericsson

June 19, 2019

Path Computation Element communication Protocol (PCEP) extensions  
for Establishing Relationships between sets of LSPs and Virtual  
Networks  
draft-leedhody-pce-vn-association-08

## Abstract

This document describes how to extend Path Computation Element (PCE) Communication Protocol (PCEP) association mechanism introduced by the PCEP Association Group specification, to further associate sets of LSPs with a higher-level structure such as a virtual network (VN) requested by clients or applications. This extended association mechanism can be used to facilitate virtual network control using PCE architecture.

## Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at  
<https://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at  
<https://www.ietf.org/shadow.html>

This Internet-Draft will expire on December 21, 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 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://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

|                                                        |    |
|--------------------------------------------------------|----|
| 1. Introduction . . . . .                              | 3  |
| 1.1. Requirements Language . . . . .                   | 4  |
| 2. Terminology . . . . .                               | 5  |
| 3. Operation Overview . . . . .                        | 5  |
| 4. Extensions to PCEP . . . . .                        | 7  |
| 5. Applicability to H-PCE architecture . . . . .       | 8  |
| 6. Implementation Status . . . . .                     | 9  |
| 6.1. Huawei's Proof of Concept based on ONOS . . . . . | 9  |
| 7. Security Considerations . . . . .                   | 10 |
| 8. IANA Considerations . . . . .                       | 10 |
| 8.1. Association Object Type Indicator . . . . .       | 10 |
| 8.2. PCEP TLV Type Indicator . . . . .                 | 10 |
| 8.3. PCEP Error . . . . .                              | 11 |
| 9. Manageability Considerations . . . . .              | 11 |
| 9.1. Control of Function and Policy . . . . .          | 11 |
| 9.2. Information and Data Models . . . . .             | 11 |
| 9.3. Liveness Detection and Monitoring . . . . .       | 11 |
| 9.4. Verify Correct Operations . . . . .               | 11 |
| 9.5. Requirements On Other Protocols . . . . .         | 11 |
| 9.6. Impact On Network Operations . . . . .            | 12 |
| 10. References . . . . .                               | 12 |
| 10.1. Normative References . . . . .                   | 12 |

|                                        |    |
|----------------------------------------|----|
| 10.2. Informative References . . . . . | 12 |
|----------------------------------------|----|

## 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.

[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.

[I-D.ietf-pce-association-group] introduces a generic mechanism to create a grouping of LSPs. This grouping can then be used to define association between sets of LSPs or between a set of LSPs and a set of attributes.

[RFC8453] describes various Virtual Network (VN) operations initiated by a customer/application. In this context, there is a need for associating a set of LSPs with a VN "construct" to facilitate VN operations in PCE architecture. This association allows the PCEs to identify which LSPs belong to a certain VN. The PCE could then use this association to optimize all LSPs belonging to the VN together. The PCE could further take VN specific actions on the LSPs such as relaxation of constraints, policy actions, setting default behavior etc.

[I-D.ietf-pce-applicability-actn] examines the PCE and ACTN architecture and describes how the PCE architecture is applicable to ACTN. [RFC6805] and [I-D.ietf-pce-stateful-hpce] describes a hierarchy of stateful PCEs with Parent PCE coordinating multi-domain path computation function between Child PCE(s) and thus making it the base for PCE applicability for ACTN. In this text child PCE would be same as Provisioning Network Controller (PNC), and the parent PCE as

Multi-domain Service Coordinator (MDSC) [RFC8453].

This document specifies a PCEP extension to associate a set of LSPs based on Virtual Network (VN) (or customer). A Virtual Network (VN) is a customer view of the TE network. Depending on the agreement between client and provider various VN operations and VN views are possible as described in [RFC8453].

### 1.1. 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.

## 2. Terminology

The terminology is as per [RFC4655], [RFC5440], [RFC6805], [RFC8231] and [RFC8453].

## 3. Operation Overview

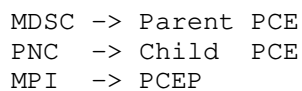
As per [I-D.ietf-pce-association-group], LSPs are associated with other LSPs with which they interact by adding them to a common association group.

An association group based on VN is useful for various optimizations that should be applied by considering all the LSPs in the association. This includes, but not limited to -

- o Path Computation: When computing path for a LSP, the impact of this LSP, on the other LSPs belonging to the same VN is useful to analyze. The aim would be optimize overall VN and all LSPs, rather than a single LSP. Also, the optimization criteria such as minimize the load of the most loaded link (MLL) [RFC5541] and other could be applied for all the LSP belonging to the same VN, identified by the VN association.

- o Path Re-Optimization: The child PCE or the parent PCE would like to use advanced path computation algorithm and optimization technique that consider all the LSPs belonging to a VN/customer and optimize them all together during the re-optimization.

This association is useful in PCEP session between parent PCE (MDSC) and child PCE (PNC). The figure describes a typical VN operations using PCEP for illustration purpose.



Local policies on the PCE MAY define the computational and optimization behavior for the LSPs in the VN. An LSP MUST NOT belong to more than one VNAG. If an implementation encounters more than one VNAG, it MUST consider the first occurrence and ignore the others.

[I-D.ietf-pce-association-group] specify the mechanism for the capability advertisement of the association types supported by a PCEP speaker by defining a ASSOC-Type-List TLV to be carried within an OPEN object. This capability exchange for the association type described in this document (i.e. VN Association Type) MUST be done before using the policy association. Thus the PCEP speaker MUST include the VN Association Type (TBD1) in the ASSOC-Type-List TLV before using the VNAG in the PCEP messages.

This Association-Type is dynamic in nature and created by the Parent PCE (MDSC) for the LSPs belonging to the same VN or customer. These associations are conveyed via PCEP messages to the PCEP peer. Operator-configured Association Range MUST NOT be set for this association-type and MUST be ignored.

#### 4. Extensions to PCEP

The format of VNAG is as per the ASSOCIATION object [I-D.ietf-pce-association-group].

This document defines one mandatory TLV "VIRTUAL-NETWORK-TLV" and one new optional TLV "VENDOR-INFORMATION-TLV"; apart from this TLV, VENDOR-INFORMATION-TLV can be used to carry arbitrary vendor specific information.

- o VIRTUAL-NETWORK-TLV: Used to communicate the VN Identifier.
- o VENDOR-INFORMATION-TLV: Used to communicate arbitrary vendor specific behavioral information, described in [RFC7470].

The format of VIRTUAL-NETWORK-TLV is as follows.

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| | Length (variable) |
+-----+-----+-----+-----+-----+-----+-----+-----+
| | |
// Virtual Network Name //
| | |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Figure 1: The VIRTUAL-NETWORK-TLV formats

Type: TBD2 (to be allocated by IANA)

Length: Variable Length

Virtual Network Name (variable): an unique symbolic name for the VN. It SHOULD be a string of printable ASCII characters, without a NULL terminator. The VN name is a human-readable string that identifies a VN. The VN name MUST remain constant throughout an LSP's lifetime, which may span across multiple consecutive PCEP sessions and/or PCC restarts. The VN name MAY be specified by an operator or auto-generated by the PCEP speaker.

The VIRTUAL-NETWORK-TLV MUST be included in VNAG object. If a PCEP speaker receives the VNAG object without the VIRTUAL-NETWORK-TLV, it MUST send a PCErr message with Error-Type=6 (mandatory object missing) and Error-Value=TBD3 (VIRTUAL-NETWORK-TLV missing) and close the session.

The format of VENDOR-INFORMATION-TLV is defined in [RFC7470].

## 5. Applicability to H-PCE architecture

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 architecture, the parent PCE is used to compute a multi-domain path based on the domain connectivity information. A child 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.

[I-D.ietf-pce-stateful-hpce] introduces general considerations for stateful PCE(s) in hierarchical PCE architecture. In particular, the behavior changes and additions to the existing stateful PCE mechanisms in the context of a H-PCE architecture.

In Stateful H-PCE architecture, the Parent PCE receives a virtual network creation request by its client over its Northbound API. This VN is uniquely identified by an Association ID in VNAG as well as the VIRTUAL-NETWORK name. This VN may comprise multiple LSPs in the network in a single domain or across multiple domains.

As the Parent PCE computes the optimum E2E paths for each tunnel in

VN, it MUST associate each LSP with the VN to which it belongs. Parent PCE sends a PCInitiate Message with this association information in the VNAG Object (See Section 4 for details). This in effect binds an LSP that is to be instantiated at the child PCE with the VN.

Whenever changes occur with the instantiated LSP in a domain network, the domain child PCE reports the changes using a PCRpt Message in which the VNAG Object indicates the relationship between the LSP and the VN.

Whenever an update occurs with VNs in the Parent PCE (via the client's request), the parent PCE sends an PCUpd Message to inform each affected child PCE of this change.

The Child PCE could then use this association to optimize all LSPs belonging to the same VN association together. The Child PCE could further take VN specific actions on the LSPs such as relaxation of constraints, policy actions, setting default behavior etc. The parent PCE could also maintain all E2E LSP or per-domain path segments under a single VN association.

## 6. Implementation Status

[Note to the RFC Editor - remove this section before publication, as well as remove the reference to RFC 7942.]

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in RFC 7942 [RFC7942]. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to RFC 7942, "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature. It is up to the individual working groups to use this information as they see fit".

### 6.1. Huawei's Proof of Concept based on ONOS

The PCE function was developed in the ONOS open source platform. This extension was implemented on a private version as a proof of concept to ACTN.

- o Organization: Huawei
- o Implementation: Huawei's PoC based on ONOS
- o Description: PCEP as a southbound plugin was added to ONOS. To support ACTN, this extension in PCEP is used. Refer <https://wiki.onosproject.org/display/ONOS/PCEP+Protocol>
- o Maturity Level: Prototype
- o Coverage: Full
- o Contact: satishk@huawei.com

## 7. Security Considerations

This document defines one new type for association, which do not add any new security concerns beyond those discussed in [RFC5440], [RFC8231] and [I-D.ietf-pce-association-group] in itself.

Some deployments may find the Virtual Network Name and the VN associations as extra sensitive; and thus should employ suitable PCEP security mechanisms like TCP-AO [RFC5925] or [RFC8253].

## 8. IANA Considerations

### 8.1. Association Object Type Indicator

This document defines a new association type, originally defined in [I-D.ietf-pce-association-group], for path protection. IANA is requested to make the assignment of a new value for the sub-registry "ASSOCIATION Type Field" (request to be created in [I-D.ietf-pce-association-group]), as follows:

| Value | Name                | Reference   |
|-------|---------------------|-------------|
| TBD1  | VN Association Type | [This I.D.] |

### 8.2. PCEP TLV Type Indicator

This document defines a new TLV for carrying additional information of LSPs within a path protection association group. IANA is requested to make the assignment of a new value for the existing "PCEP TLV Type Indicators" registry as follows:

| Value | Name                | Reference   |
|-------|---------------------|-------------|
| TBD2  | VIRTUAL-NETWORK-TLV | [This I.D.] |

### 8.3. PCEP Error

This document defines new Error-Type and Error-Value related to path protection association. IANA is requested to allocate new error values within the "PCEP-ERROR Object Error Types and Values" sub-registry of the PCEP Numbers registry, as follows:

| Error-Type | Meaning |
|------------|---------|
|------------|---------|

|   |                                                                                       |
|---|---------------------------------------------------------------------------------------|
| 6 | Mandatory Object missing<br>Error-value=TBD3: VIRTUAL-NETWORK TLV missing [This I.D.] |
|---|---------------------------------------------------------------------------------------|

## 9. Manageability Considerations

### 9.1. Control of Function and Policy

An operator MUST BE allowed to mark LSPs that belong to the same VN. This could also be done automatically based on the VN configuration.

### 9.2. Information and Data Models

The PCEP YANG module [I-D.ietf-pce-pcep-yang] should support the association between LSPs including VN association.

### 9.3. 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].

### 9.4. Verify Correct Operations

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

### 9.5. Requirements On Other Protocols

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

## 9.6. Impact On Network Operations

Mechanisms defined in this document do not have any impact on network operations in addition to those already listed in [RFC5440].

## 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.
- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, May 2017.
- [RFC8231] E. Crabbe, I. Minei, J. Medved, and R. Varga, "PCEP Extensions for Stateful PCE", RFC 8231, September 2017.
- [RFC8281] E. Crabbe, et. al., "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", RFC 8281, December 2017.
- [I-D.ietf-pce-association-group] I, Minei, Ed., "PCEP Extensions for Establishing Relationships Between Sets of LSPs", draft-ietf-pce-association-group, work in progress.

### 10.2. Informative References

- [RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.
- [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] A. Farrel and D. King, "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS", RFC 6805, November 2012.
- [RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", BCP 205, RFC 7942, DOI 10.17487/RFC7942, July 2016.

- [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>>.
- [I-D.ietf-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-ietf-pce-applicability-actn, work-in-progress.
- [I-D.ietf-pce-stateful-hpce] Dhody, D. and Lee, Y., "Hierarchical Stateful Path Computation Element (PCE)", draft-ietf-pce-stateful-hpce, work in progress.
- [RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", RFC 5541, DOI 10.17487/RFC5541, June 2009, <<http://www.rfc-editor.org/info/rfc5541>>.
- [RFC7470] Zhang, F. and A. Farrel, "Conveying Vendor-Specific Constraints in the Path Computation Element Communication Protocol", RFC 7470, DOI 10.17487/RFC7470, March 2015, <<http://www.rfc-editor.org/info/rfc7470>>.
- [RFC8051] Zhang, X., Ed. and I. Minei, Ed., "Applicability of a Stateful Path Computation Element (PCE)", RFC 8051, DOI 10.17487/RFC8051, January 2017, <<http://www.rfc-editor.org/info/rfc8051>>.
- [RFC8253] Lopez, D., Dios, O., Wu, W., and D. Dhody, "Secure Transport for PCEP", RFC 8253, October 2017 .
- [I-D.ietf-pce-pcep-yang] Dhody, D., Hardwick, J., Beeram, V., and j. jeffrant@gmail.com, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", draft-ietf-pce-pcep-yang (work in progress).

Contributor's Addresses

Dhruv Dhody  
Huawei Technologies  
Divyashree Technopark, Whitefield  
Bangalore, Karnataka 560066  
India

Email: dhruv.ietf@gmail.com

Qin Wu  
Huawei Technologies  
China

Email: bill.wu@huawei.com

Author's Addresses

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

Email: younglee.tx@gmail.com

Xian Zhang  
Huawei Technologies  
China

Email: zhang.xian@huawei.com

Daniele Ceccarelli  
Ericsson  
Torshamnsgatan, 48  
Stockholm,  
Sweden

Email: daniele.ceccarelli@ericsson.com

Network Working Group  
Internet-Draft  
Intended status: Experimental  
Expires: December 22, 2018

C. Li  
M. Chen  
J. Dong  
Z. Li  
D. Dhody  
Huawei Technologies  
June 20, 2018

PCE Controlled ID Space  
draft-li-pce-controlled-id-space-00

Abstract

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. The Stateful PCE extensions allow stateful control of Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Label Switched Paths (LSPs) using PCEP. Furthermore, PCEP can be used for computing paths in SR networks.

Stateful PCE provide active control of MPLS-TE LSPs via PCEP, for a model where the PCC delegates control over one or more locally configured LSPs to the PCE. Further, stateful PCE could also create and delete PCE-initiated LSPs itself. A PCE-based central controller (PCECC) simplify the processing of a distributed control plane by blending it with elements of Software-Defined Networking (SDN) and without necessarily completely replacing it.

In some use cases, such as PCECC, Binding Segment Identifier (SID), SR Path Identification, there is a requirement for a stateful PCE to make allocation of labels, SID, Path-ID respectively. These use cases require for the PCE to be aware of the various identifier space from which to make allocations on behalf of PCC. This documents specify a mechanism for a PCC to inform the PCE of the identifier space under its control via PCEP. The identifier could be MPLS label, SID, Path ID or another future identifier to be allocated by a PCE.

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.

## 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 22, 2018.

## Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 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

|                                                 |    |
|-------------------------------------------------|----|
| 1. Introduction . . . . .                       | 3  |
| 2. Terminology . . . . .                        | 4  |
| 3. Use cases . . . . .                          | 4  |
| 3.1. PCE-based Central Control . . . . .        | 4  |
| 3.2. Binding SID Allocation . . . . .           | 5  |
| 3.3. Path ID Allocation . . . . .               | 5  |
| 4. Overview . . . . .                           | 5  |
| 5. Objects . . . . .                            | 7  |
| 5.1. Open Object . . . . .                      | 7  |
| 5.1.1. LABEL-CONTROL-SPACE TLV . . . . .        | 7  |
| 5.1.2. SRv6-PATH-ID-CONTROL-SPACE TLV . . . . . | 8  |
| 6. Other Considerations . . . . .               | 9  |
| 7. IANA Considerations . . . . .                | 10 |
| 8. Security Considerations . . . . .            | 10 |

|                                        |    |
|----------------------------------------|----|
| 9. Acknowledgements . . . . .          | 10 |
| 10. References . . . . .               | 10 |
| 10.1. Normative References . . . . .   | 10 |
| 10.2. Informative References . . . . . | 11 |
| Authors' Addresses . . . . .           | 12 |

## 1. Introduction

[RFC5440] defines the stateless Path Computation Element communication Protocol (PCEP) for Path Computation Elements (PCEs) to perform path computations in response to Path Computation Clients (PCCs) requests. For supporting stateful operations, [RFC8231] specifies a set of extensions to PCEP to enable stateful control of LSPs within and across PCEP sessions in compliance with [RFC4657]. Furthermore, [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.

[RFC8283] introduces the architecture for PCE as a central controller, it examines the motivations and applicability for PCEP as a control protocol in this environment, and introduces the implications for the protocol. Also, [I-D.zhao-pce-pcep-extension-for-pce-controller] specifies the procedures and PCEP protocol extensions for using the PCE as the central controller, where LSPs are calculated/setup/initiated and label forwarding entries are downloaded through extending PCEP. However, the document assumes that label range to be used by a PCE is known and set on both PCEP peers. This extension adds the capability to advertise the range via a PCEP extension.

Similarly, [I-D.zhao-pce-pcep-extension-pce-controller-sr] specifies the procedures and PCEP protocol extensions when a PCE-based controller is also responsible for configuring the forwarding actions on the routers (SR SID distribution in this case), in addition to computing the paths for packet flows in a segment routing network and telling the edge routers what instructions to attach to packets as they enter the network. However, the document assumes that label range to be used by a PCE is known and set on both PCEP peers. This extension adds the capability to advertise the range (from SRGB or SRLB of the node) via a PCEP extension.

[I-D.li-pce-sr-path-segment] defines a procedure for path ID in PCEP for SR by defining the PATH-ID TLV. The path ID can be a path segment in SR-MPLS [I-D.cheng-spring-mpls-path-segment], or a path ID in SRv6 [I-D.li-spring-passive-pm-for-srv6-np], or other IDs that can identify an SR path. This document specifies the extension to support advertisement of the various ID space to the PCE to control.

The usecase are described in Section 3. The ID space range information can be advertised via the TLVs in the Open message. The detailed procedures will be described in Section 4, and the objects' format will be introduced in Section 5.

## 2. Terminology

This memo makes use of the terms defined in [RFC5440], [RFC8231], [RFC8283] and [I-D.ietf-spring-segment-routing].

## 3. Use cases

### 3.1. PCE-based Central Control

A PCE-based central controller (PCECC) can simplify the processing of a distributed control plane by blending it with elements of SDN and without necessarily completely replacing it. Thus, the LSP can be calculated/setup/initiated and the label forwarding entries can also be downloaded through a centralized PCE server to each network devices along the path while leveraging the existing PCE technologies as much as possible.

[I-D.zhao-pce-pcep-extension-for-pce-controller] describe a mode where LSPs are provisioned as explicit label instructions at each hop on the end-to-end path. Each router along the path must be told what label forwarding instructions to program and what resources to reserve. The controller uses PCEP to communicate with each router along the path of the end-to-end LSP. For this to work, the PCE-based controller will take responsibility for managing some part of the MPLS label space for each of the routers that it controls as described in section 3.1.2. of [RFC8283]. A mechanism for a PCC to inform the PCE of such a label space to control is needed within PCEP.

[I-D.ietf-pce-segment-routing] specifies extensions to PCEP that allow a stateful PCE to compute, update or initiate SR-TE paths. [I-D.zhao-pce-pcep-extension-pce-controller-sr] describes the mechanism for PCECC to allocate and provision the node/prefix/adjacency label (SID) via PCEP. To make such allocation PCE needs to be aware of the label space from Segment Routing Global Block (SRGB) or Segment Routing Local Block (SRLB) of the node that it controls. A mechanism for a PCC to inform the PCE of such a label space to control is needed within PCEP. The full SRGB/SRLB of a node could be learned via existing IGP or BGP-LS mechanism.

### 3.2. Binding SID Allocation

The headend of an SR Policy binds a binding SID to its policy [I-D.ietf-spring-segment-routing]. The instantiation of which may involve a list of SIDs. Currently binding SID are allocated by the node, but there is an inherent advantage in the binding SID to be allocated by a PCE to allow SR policies to be dynamically created, updated according to the network status and operations. Therefore, a PCE needs to obtain the authority and control to allocate binding SID actively from the PCC's label space as described in above use case.

### 3.3. Path ID Allocation

Path identification is needed for several use cases such as performance measurement in Segment Routing (SR) network. For identifying an SR path, [I-D.cheng-spring-mpls-path-segment] introduces a new segment that is referred to as Path Segment, and [I-D.li-spring-passive-pm-for-srv6-np] introduces the path ID in SRv6.

[I-D.li-pce-sr-path-segment] defines a procedure for path ID in PCEP for SR. It describes a mode in which PCE could allocate path ID and inform the ingress and egress PCC. To make such an allocation a PCE needs to be aware of the path ID space under its control. A mechanism for a PCC to inform the PCE of such a path ID space is needed within PCEP.

## 4. Overview

During PCEP Initialization Phase, Open messages are exchanged between PCCs and PCEs. The OPEN object may also contain a set of TLVs used to convey capabilities in the Open message. The ID could be a MPLS label, SRv6 path ID or any other future ID space for PCE to allocate. A PCC can include a corresponding ID-CONTROL-SPACE TLVs, in the OPEN Object to inform the corresponding ID space information that it wants the PCE to control. This TLV MUST NOT be included by the PCE and MUST be ignored on receipt by a PCC. This is an optional TLV, the PCE could be aware of the ID space from some other means outside of PCEP.

For delegating multiple types of ID space, multiple TLVs corresponding to each ID type MUST be included in a Open message. Each TLV (corresponding to each ID type) SHOULD be included only once in a Open Message. On receipt, only the first instance is processed and others MUST be ignored. The ID type can be MPLS label, SRv6 path ID [I-D.li-spring-passive-pm-for-srv6-np] or other ID. The following ID-CONTROL-SPACE TLVs are defined in this document -

- o LABEL-CONTROL-SPACE - for MPLS Labels (including for SR-MPLS)
- o SRv6-PATH-ID-CONTROL-SPACE - for SRv6 Path ID

The procedure of ID space control to PCE is shown below:

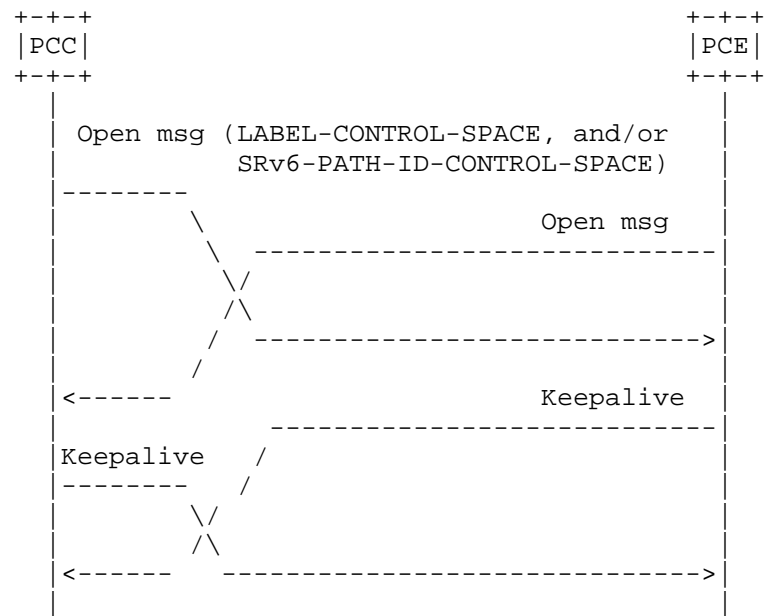


Figure 1: ID space control to PCE

If the ID space control procedure is successful, the PCE will return a KeepAlive message to the PCC. If there is any error in processing the corresponding TLV, an Error (PCErr) message will be sent to the PCC with Error-Type=1 (PCEP session establishment failure) and Error-value=TBD (ID space control failure).

After this process, a stateful PCE can learn the PCE controlled ID spaces of a node (PCC) under its control. A PCE can then allocate IDs within the control ID space. For example, a PCE can actively allocate labels and download forwarding instructions for the PCECC LSP as described in [I-D.zhao-pce-pcep-extension-for-pce-controller]. A PCE can also allocate labels from SRGB/SRLB for PCECC-SR [I-D.zhao-pce-pcep-extension-pce-controller-sr], binding segments, and path segments [I-D.cheng-spring-mpls-path-segment]. The full SRGB/SRLB of a node could be learned via existing IGP or BGP-LS mechanism. Similarly a PCE can allocate SRv6 Path ID

[I-D.li-spring-passive-pm-for-srv6-np] according to the SRv6 Path ID space under its control.

## 5. Objects

### 5.1. Open Object

For advertising the PCE controlled ID space to a PCE, this document defines several TLVs within the Open object.

#### 5.1.1. LABEL-CONTROL-SPACE TLV

For a PCC to inform the label space under the PCE control, this document defines a new LABEL-CONTROL-SPACE TLV.

The LABEL-CONTROL-SPACE TLV is an optional TLV for use in the OPEN object, and its format is shown in the following figure:

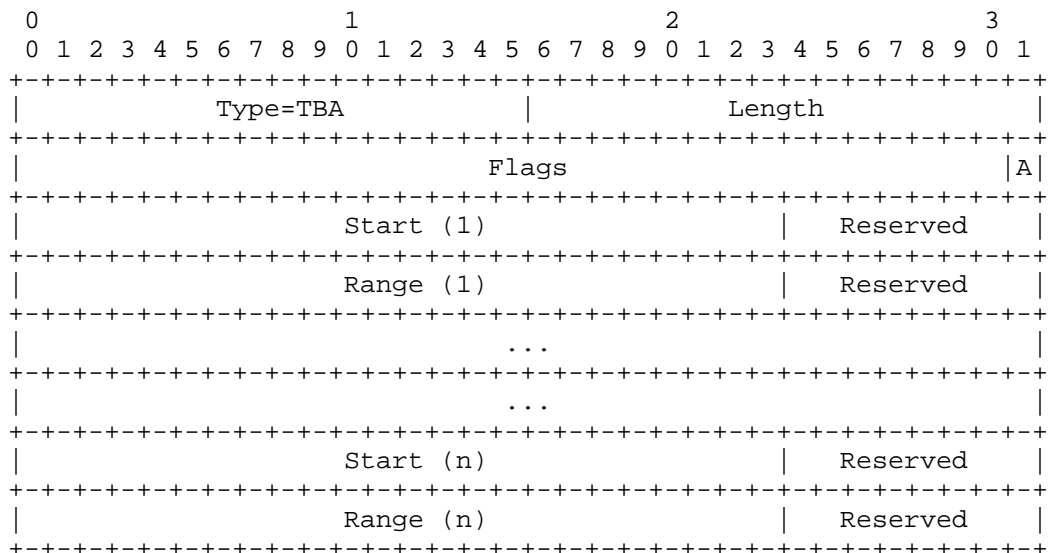


Figure 2: LABEL-CONTROL-SPACE TLV

The type (16 bits) of the TLV is TBA. The length field (16 bits) and has a variable value.

Flags (32 bits): Following flags are currently defined

- o A-flag: All space flag, set when all the label space is delegated to a PCE. When A-flag is set, the pair of Start and End SHOULD

NOT appear unless the PCC needs to notify the entire ID space to a PCE.

The unassigned bits of Flags field MUST be set to 0 on transmission and MUST be ignored on receipt.

Start(i) (24 bits): indicates the beginning of the label block i.

Range(i) (24 bits): indicates the range of the label block i.

Reserved: SHOULD be set to 0 on transmission and MUST be ignored on reception.

The number of label blocks can be calculated according to value of the length field in the TLV.

A stateful PCE can actively allocate labels and download forwarding instructions for the PCECC LSP as described in [I-D.zhao-pce-pcep-extension-for-pce-controller]. A PCE can also allocate labels from SRGB/SRLB for PCECC-SR [I-D.zhao-pce-pcep-extension-pce-controller-sr] and binding segments can be selected for the PCE controlled space. Also, Path segment [I-D.cheng-spring-mpls-path-segment] can be allocated by a stateful PCE in a similar same way as described in [I-D.li-pce-sr-path-segment].

#### 5.1.2. SRv6-PATH-ID-CONTROL-SPACE TLV

For a PCC to inform the SRv6 path ID space under the PCE control, this document defines a new SRv6-PATH-ID-CONTROL-SPACE TLV.

The SRv6-PATH-ID-CONTROL-SPACE TLV is an optional TLV for use in the OPEN object, and its format is shown in the following figure:

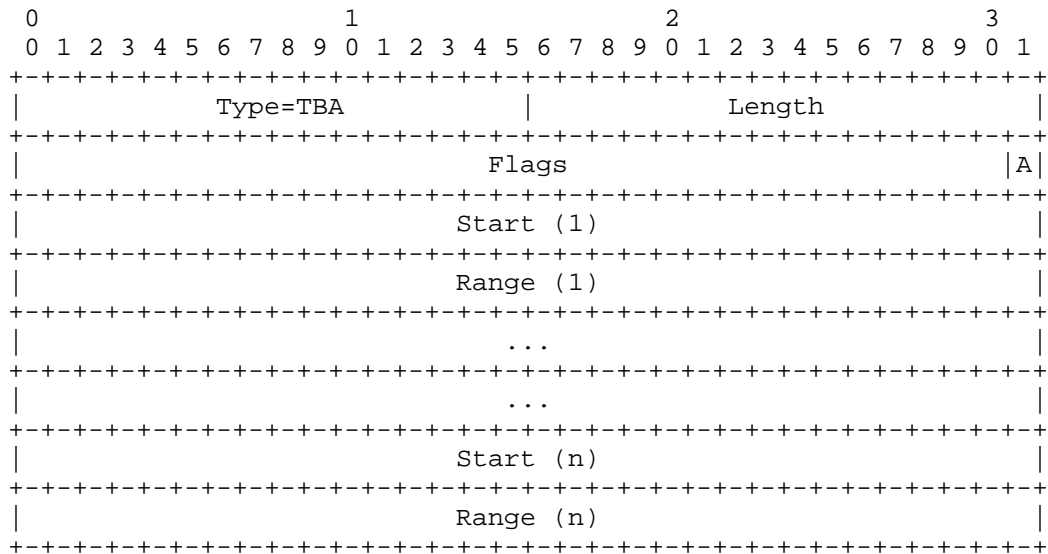


Figure 3: SRv6-PATH-ID-CONTROL-SPACE TLV

The type (16 bits) of the TLV is TBA. The length field (16 bits) and has a variable value.

Flags (32 bits): is of same format as LABEL-CONTROL-SPACE TLV. Any bits assigned in the LABEL-CONTROL-SPACE TLV are also applicable for this.

Start(i) (32 bits): indicates the beginning of the SRv6 Path ID block *i*.

Range(i) (32 bits): indicates the range of the SRv6 Path ID block i.

The number of Path ID blocks can be calculated according to the length field in the TLV. Given the controlled ID spaces, a stateful PCE can actively allocate path IDs to SRv6 paths from the controlled ID spaces as described in [I-D.li-pce-sr-path-segment].

## 6. Other Considerations

In case of multiple PCEs, a PCC MAY decide to give control over different ID space to each instance of the PCE. In case a PCC includes the same ID space to multiple PCEs, the PCE SHOULD use synchronization mechanism (such as [I-D.litkowski-pce-state-sync]) to avoid allocating the same ID.

## 7. IANA Considerations

TBA.

## 8. Security Considerations

TBA.

## 9. Acknowledgements

TBA.

## 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, <<https://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, <<https://www.rfc-editor.org/info/rfc5440>>.
- [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>>.
- [RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <<https://www.rfc-editor.org/info/rfc8283>>.

## 10.2. Informative References

- [RFC4657] Ash, J., Ed. and J. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, DOI 10.17487/RFC4657, September 2006, <<https://www.rfc-editor.org/info/rfc4657>>.
- [I-D.ietf-pce-segment-routing]  
Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "PCEP Extensions for Segment Routing", draft-ietf-pce-segment-routing-11 (work in progress), November 2017.
- [I-D.ietf-spring-segment-routing]  
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-15 (work in progress), January 2018.
- [I-D.zhao-pce-pcep-extension-for-pce-controller]  
Zhao, Q., Li, Z., Dhody, D., Karunanithi, S., Farrel, A., and C. Zhou, "PCEP Procedures and Protocol Extensions for Using PCE as a Central Controller (PCECC) of LSPs", draft-zhao-pce-pcep-extension-for-pce-controller-08 (work in progress), June 2018.
- [I-D.zhao-pce-pcep-extension-pce-controller-sr]  
Zhao, Q., Li, Z., Dhody, D., Karunanithi, S., Farrel, A., and C. Zhou, "PCEP Procedures and Protocol Extensions for Using PCE as a Central Controller (PCECC) of SR-LSPs", draft-zhao-pce-pcep-extension-pce-controller-sr-03 (work in progress), June 2018.
- [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-03 (work in progress), April 2018.
- [I-D.cheng-spring-mpls-path-segment]  
Cheng, W., Wang, L., Li, H., Chen, M., Zigler, R., and S. Zhan, "Path Segment in MPLS Based Segment Routing Network", draft-cheng-spring-mpls-path-segment-01 (work in progress), March 2018.

[I-D.li-spring-passive-pm-for-srv6-np]

Li, C. and M. Chen, "Passive Performance Measurement for SRv6 Network Programming", draft-li-spring-passive-pm-for-srv6-np-00 (work in progress), March 2018.

#### Authors' Addresses

Cheng Li  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

EMail: chenglil3@huawei.com

Mach(Guoyi) Chen  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

EMail: Mach.chen@huawei.com

Jie Dong  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

EMail: jie.dong@huawei.com

Zhenbin Li  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

EMail: lizhenbin@huawei.com

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

EMail: [dhruv.ietf@gmail.com](mailto:dhruv.ietf@gmail.com)

Network Working Group  
Internet-Draft  
Intended status: Experimental  
Expires: 22 September 2022

C. Li  
H. Shi  
Huawei Technologies  
A. Wang  
China Telecom  
W. Cheng  
China Mobile  
C. Zhou  
HPE  
21 March 2022

PCE Controlled ID Space  
draft-li-pce-controlled-id-space-11

Abstract

The Path Computation Element Communication Protocol (PCEP) provides a mechanisms for the Path Computation Elements (PCEs) to perform path computations in response to Path Computation Clients (PCCs) requests. The Stateful PCE extensions allow stateful control of Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Label Switched Paths (LSPs) using PCEP. Furthermore, PCE can be used for computing paths in the SR networks.

Stateful PCE provide active control of MPLS-TE LSPs via PCEP, for a model where the PCC delegates control over one or more locally configured LSPs to the PCE. Further, stateful PCE could also create and remove PCE-initiated LSPs by itself. A PCE-based Central Controller (PCECC) simplify the processing of a distributed control plane by integrating with elements of Software-Defined Networking (SDN).

In some use cases, such as PCECC or Binding Segment Identifier (SID) for Segment Routing (SR), there are requirements for a stateful PCE to make allocation of labels, SIDs, etc. These use cases require PCE aware of various identifier spaces from where to make allocations on behalf of a PCC. This document describes a mechanism for a PCC to inform the PCE of the identifier space set aside for the PCE control via PCEP. The identifier could be an MPLS label, a SID or any other to-be-defined identifier that can be allocated by a PCE.

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 22 September 2022.

#### Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

#### Table of Contents

|                                                        |    |
|--------------------------------------------------------|----|
| 1. Introduction . . . . .                              | 3  |
| 2. Terminology . . . . .                               | 4  |
| 2.1. Requirements Language . . . . .                   | 4  |
| 3. Use cases . . . . .                                 | 4  |
| 3.1. PCE-based Central Control . . . . .               | 4  |
| 3.1.1. PCECC for MPLS/SR-MPLS . . . . .                | 4  |
| 3.1.2. PCECC for SRv6 . . . . .                        | 5  |
| 3.2. Binding SID Allocation . . . . .                  | 5  |
| 4. Overview . . . . .                                  | 6  |
| 5. Objects . . . . .                                   | 7  |
| 5.1. Open Object . . . . .                             | 7  |
| 5.1.1. LABEL-CONTROL-SPACE TLV . . . . .               | 7  |
| 5.1.2. FUNCT-ID-CONTROL-SPACE TLV . . . . .            | 9  |
| 6. Other Considerations . . . . .                      | 11 |
| 7. IANA Considerations . . . . .                       | 11 |
| 7.1. PCEP TLV Type Indicators . . . . .                | 11 |
| 7.2. LABEL-CONTROL-SPACE TLV's Flag field . . . . .    | 11 |
| 7.3. FUNCT-ID-CONTROL-SPACE TLV's Flag field . . . . . | 12 |
| 8. Security Considerations . . . . .                   | 12 |
| 9. Acknowledgements . . . . .                          | 13 |

|                              |    |
|------------------------------|----|
| 10. References               | 13 |
| 10.1. Normative References   | 13 |
| 10.2. Informative References | 14 |
| Appendix A. Contributors     | 16 |
| Authors' Addresses           | 17 |

## 1. Introduction

[RFC5440] defines the stateless Path Computation Element Communication Protocol (PCEP) for the Path Computation Elements (PCEs) to perform path computation in response to Path Computation Clients (PCCs) requests. For supporting stateful operations, [RFC8231] specifies a set of extensions to PCEP to enable stateful control of LSPs within and across PCEP sessions in compliance with [RFC4657]. Furthermore, [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.

[RFC8283] introduces the architecture for PCE as a central controller, it examines the motivations and applicability for PCEP as a control protocol in this environment, and introduces the implications for the protocol. Also, [RFC9050] specifies the procedures and PCEP extensions for using the PCE as a Central Controller (PCECC), where LSPs are calculated/set up/initiated and label forwarding entries are downloaded through extending PCEP. However, the document assumes that label range to be used by a PCE is known and set on both PCEP peers. This extension adds the capability to advertise the label range via a PCEP extension.

Similarly, [RFC9050] specifies the procedures and PCEP extensions when a PCE-based controller is also responsible for configuring the forwarding actions on the routers (SR SID distribution in this case), in addition to computing the paths for packet flows in a segment routing network and telling the edge routers what instructions to attach to packets as they enter the network. However, the document assumes that label range to be used by a PCE is known and set on both PCEP peers. This extension adds the capability to advertise the range (from SRGB or SRLB of the node) via a PCEP extension.

In addition, [I-D.dhody-pce-pcep-extension-pce-controller-srv6] specifies the procedures and PCEP extensions of PCECC for SRv6. An SRv6 SID is represented as LOC:FUNCT ([RFC8986]) where LOC is the L most significant bits and FUNCT is the 128-L least significant bits. The FUNCT part of the SID is an opaque identification of a local function bound to the SID. This extension adds the capability to advertise the range of Function ID (FUNCT part) via a PCEP extension.

Once the PCC/node has given control over an ID space (for example labels), the PCC/node MUST NOT allocate the ID from this ID space. For example, a PCC/node MUST NOT use this labels from the PCE controlled label space to make allocation for VPN Prefix distributed via BGP or labels used for LDP/RSVP-TE signalling. This is done to make sure that the PCE control over ID space does not conflict with the existing node allocation.

The use case are described in Section 3. The ID space range information can be advertised via the TLVs in the Open message. The detailed procedures are described in Section 4, and the TLV format is specified in Section 5.

## 2. Terminology

This memo makes use of the terms defined in [RFC5440], [RFC8231], [RFC8283] and [RFC8402].

### 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. Use cases

### 3.1. PCE-based Central Control

A PCE-based Central Controller (PCECC) can simplify the processing of a distributed control plane by integrating with elements of SDN. Thus, the LSP/SR path can be calculated/set up/initiated and the label/SID forwarding entries can also be downloaded through a centralized PCE server to each network devices along the path while leveraging the existing PCE technologies as much as possible.

#### 3.1.1. PCECC for MPLS/SR-MPLS

[RFC9050] describes a mode where LSPs are provisioned as explicit label instructions at each hop on the end-to-end path. Each router along the path must be told what label forwarding instructions to program and what resources to reserve. The controller uses PCEP to communicate with each router along the path of the end-to-end LSP. For this to work, the PCE-based controller will take responsibility for managing some part of the MPLS label space for each router that it controls as described in section 3.1.2. of [RFC8283]. A mechanism for a PCC to inform the PCE of such a label space to control is

needed within PCEP.

[RFC8664] specifies extensions to PCEP that allow a stateful PCE to compute, update or initiate SR-TE paths. [RFC9050] describes the mechanism for PCECC to allocate and distribute the node/prefix/adjacency label (SID) via PCEP. To make such allocation, PCE needs to be aware of the label space from Segment Routing Global Block (SRGB) or Segment Routing Local Block (SRLB) [RFC8402] of the node that it can control. A mechanism for a PCC to inform the PCE of such label space to control is needed within PCEP. The full SRGB/SRLB of a node could be learned via existing IGP or BGP-LS mechanism.

### 3.1.2. PCECC for SRv6

[I-D.dhody-pce-pcep-extension-pce-controller-srv6] describes the mechanism for PCECC to allocate and provision the SRv6 SID via PCEP. An SRv6 SID is represented as LOC:FUNCT ([RFC8986]) where LOC is the L most significant bits and FUNCT is the 128-L least significant bits. The FUNCT part of the SID is an opaque identification of a local function bound to the SID. To make such allocation, PCE needs to be aware of the Function ID space (FUNCT part) of the node that it controls. A mechanism for a PCC to inform the PCE of such a Function ID space to control is needed within PCEP.

### 3.2. Binding SID Allocation

The headend of an SR Policy binds a Binding SID (BSID) [I-D.ietf-pce-binding-label-sid] to its policy [I-D.ietf-spring-segment-routing-policy]. The instantiation of which may involve a list of SIDs. The Binding SID can be allocated by the node as described in [I-D.ietf-pce-binding-label-sid], but there is an inherent advantage in the Binding SID to be allocated by a PCE to allow SR policies to be dynamically created, updated according to the network status and operations. This is described in [RFC9050]. Therefore, a PCE needs to obtain the authority and control to allocate Binding SID actively from the PCC's label space as described in above use case.

This is applicable for both SR-MPLS and SRv6 BSID.

#### 4. Overview

During PCEP Initialization Phase, Open messages are exchanged between the PCCs and the PCEs. The OPEN object may also contain a set of TLVs used to convey the capabilities in the Open message. The term 'ID' in this document, could be a MPLS label, SRv6 Function ID or any other future ID space for PCE to control and allocate from. A PCC can include a corresponding ID-CONTROL-SPACE TLVs in the OPEN Object to inform the corresponding ID space information that it wants the PCE to control. This TLV MUST NOT be included by the PCE and MUST be ignored on receipt by a PCC. This is an optional TLV, the PCE could be aware of the ID space from some other means outside of PCEP.

For delegating multiple types of ID space, multiple TLVs corresponding to each ID type MUST be included in an Open message. The ID type can be MPLS label or other type of ID. The following ID-CONTROL-SPACE TLV is defined in this document -

- \* LABEL-CONTROL-SPACE TLV - for MPLS Labels (including for SR-MPLS)
- \* FUNCTION-ID-CONTROL-SPACE TLV - for SRv6 SID Function ID

The procedure of ID space control to PCE is shown below:

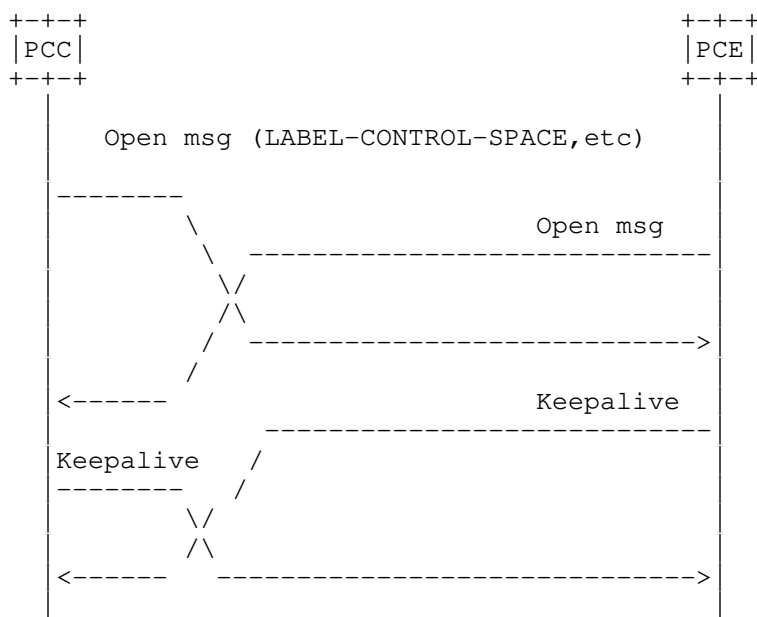


Figure 1: ID space control to PCE

If the ID space control procedure is successful, the PCE will return a KeepAlive message to the PCC. If there is any error in processing the corresponding TLV, an Error (PCErr) message will be sent to the PCC with Error-Type=1 (PCEP session establishment failure) and Error-value=TBD (ID space control failure).

After this process, a stateful PCE can learn the PCE-controlled ID spaces of a node (PCC) under its control. A PCE can then allocate IDs within the controlled-ID space. For example, a PCE can actively allocate labels and download forwarding instructions for the PCECC LSP as described in [RFC9050]. A PCE can also allocate labels from the PCE controlled portion of the SRGB/SRLB for PCECC-SR [RFC9050]. The full SRGB/SRLB of a node could be learned via existing IGP or BGP-LS mechanism.

The procedure for handling the FUNCTION-ID-CONTROL-SPACE TLV is same as above.

## 5. Objects

### 5.1. Open Object

For advertising the PCE-controlled ID space to a PCE, this document defines several TLVs within the OPEN object.

#### 5.1.1. LABEL-CONTROL-SPACE TLV

For a PCC to inform the label space under the PCE control, this document defines a new LABEL-CONTROL-SPACE TLV.

The LABEL-CONTROL-SPACE TLV is an optional TLV in the OPEN object, and its format is shown in the following figure:

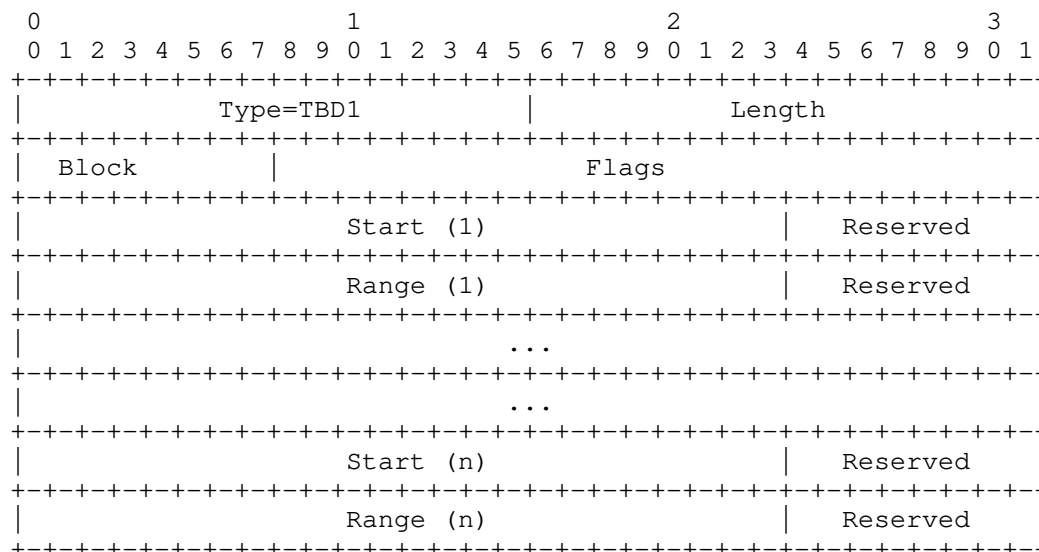


Figure 2: LABEL-CONTROL-SPACE TLV

The type (16 bits) of the TLV is TBD1. The length field (16 bits) and has a variable value.

Block(8 bits): the number of ID blocks. The range of a block is described by a start field and a range field.

Flags (24 bits): No flag is currently defined. The unassigned bits of Flags field MUST be set to 0 on transmission and MUST be ignored on receipt.

Start(i) (24 bits): indicates the beginning of the label block i.

Range(i) (24 bits): indicates the range of the label block i.

Reserved: MUST be set to 0 on transmission and MUST be ignored on reception.

LABEL-CONTROL-SPACE TLV SHOULD be included only once in a Open Message. On receipt, only the first instance is processed and others MUST be ignored.

A stateful PCE can actively allocate labels and download forwarding instructions for the PCECC LSP as described in [RFC9050]. A PCE can also allocate labels from SRGB/SRLB for PCECC-SR [I-D.ietf-pce-pcep-extension-pce-controller-sr]. The Binding Segments can also be selected for the PCE controlled space [RFC9050].

5.1.2. FUNCT-ID-CONTROL-SPACE TLV

For a PCC to inform the SRv6 SID Function ID space under the PCE control, this document defines a new FUNCT-ID-CONTROL-SPACE TLV.

The FUNCT-ID-CONTROL-SPACE TLV is an optional TLV for use in the OPEN object, and its format is shown in the following figure:

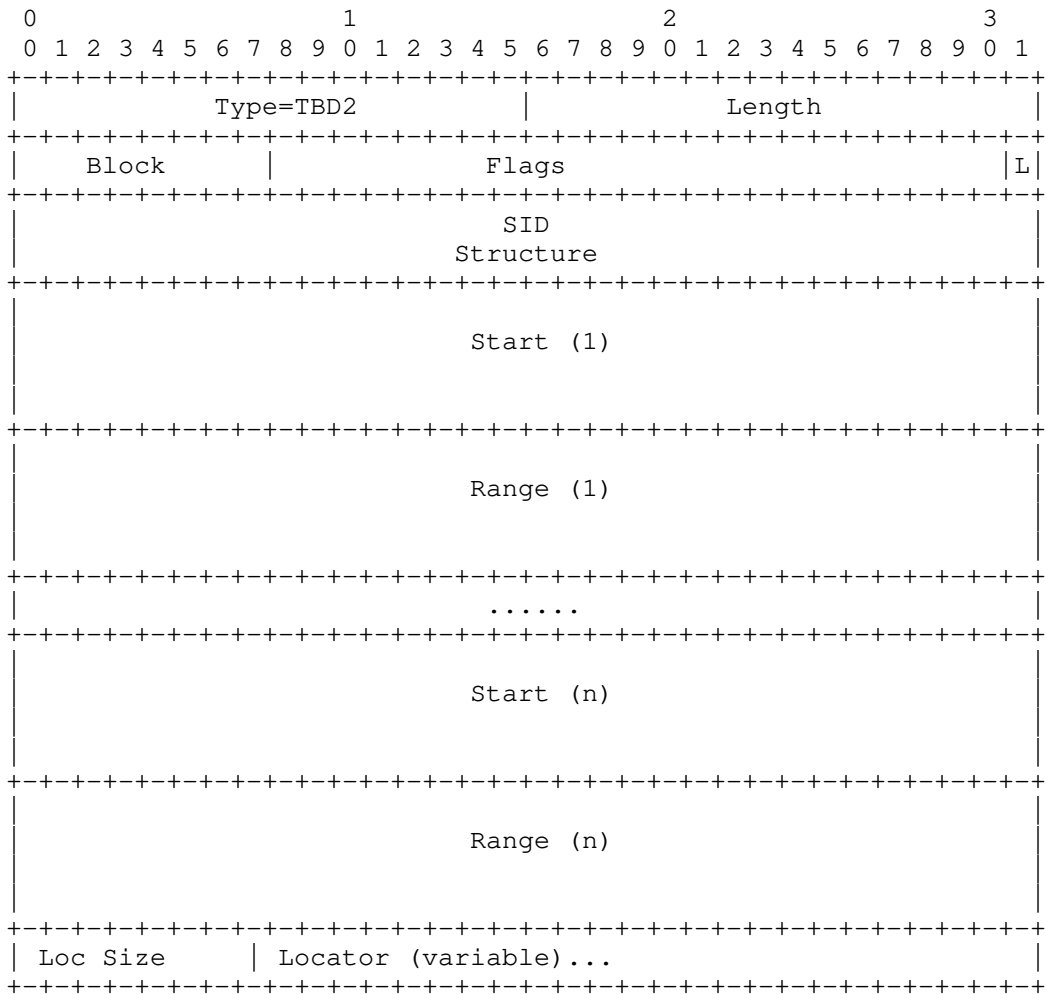


Figure 3: FUNCT-ID-CONTROL-SPACE TLV

The type (16 bits) of the TLV is TBD2. The length field (16 bits) and has a variable value.

Block(8 bits): the number of ID blocks. The range of a block is described by a start field and a range field.

Flags (24 bits): Following flags are currently defined

- \* L-flag: Locator flag, set when the locator information is included in this TLV. If L-flag is unset, Loc Size and variable Locator field MUST NOT be included in this TLV, and the ID spaces are applicable to all Locators.

The unassigned bits of Flags field MUST be set to 0 on transmission and MUST be ignored on receipt.

SID Structure: 64-bit field formatted as per "SID Structure" in [I-D.ietf-pce-segment-routing-ipv6].

Start(i) (128 bits): indicates the beginning of the Function ID block i.

Range(i) (128 bits): indicates the range of the Function ID block i.

Loc size(8 bits): indicates the bit length of a Locator. Appears only when the L-flag is set.

Locator (variable length): the value of a Locator. The Function ID spaces specified in this TLV are associated with this locator.

As per [RFC5440], the value portion of the PCEP TLV needs to be 4-bytes aligned, so a FUNCT-ID-CONTROL-SPACE TLV is padded with trailing zeros to a 4-byte boundary.

Multiple FUNCT-ID-CONTROL-SPACE TLVs MAY be included in a OPEN object to specify Function ID space specific to each locator.

A stateful PCE can actively allocate SRv6 SID and download SIDs for the PCECC-SRv6 as described in [I-D.dhody-pce-pcep-extension-pce-controller-srv6].

Note that SRv6 SID allocation involves LOC:FUNCT; the LOC is assumed to be known at PCE and FUNCT is allocated from the PCE controlled Function ID block.

## 6. Other Considerations

In case of multiple PCEs, a PCC MAY decide to give control over different ID space to each instance of the PCE. In case a PCC includes the same ID space to multiple PCEs, the PCE MUST use synchronization mechanism (such as [I-D.ietf-pce-state-sync]) to avoid allocating the same ID.

The PCE would allocate ID from the PCE controlled ID space. The PCC would not allocate ID by itself from this space as long as it has an active PCEP session to a PCE to which it has given control over the ID space.

Note that if there is any change in the ID space, the PCC MUST bring the session down and re-establish the session with new TLVs. During state synchronization the PCE would need to consider the new ID space into consideration and SHOULD re-establish the LSP/SR-paths if needed.

The PCC can regain control of the ID space by closing the PCEP session and require new session without ID space TLVs specified in this document.

## 7. IANA Considerations

IANA maintains the "Path Computation Element Protocol (PCEP) Numbers" registry. This document requests IANA actions to allocate code points for the protocol elements defined in this document.

### 7.1. PCEP TLV Type Indicators

IANA maintains a subregistry called "PCEP TLV Type Indicators". IANA is requested to make an assignment from this subregistry as follows:

| Value | Meaning                    | Reference  |
|-------|----------------------------|------------|
| TBD1  | LABEL-CONTROL-SPACE TLV    | [This.I-D] |
| TBD2  | FUNCT-ID-CONTROL-SPACE TLV | [This.I-D] |

### 7.2. LABEL-CONTROL-SPACE TLV's Flag field

This document defines the LABEL-CONTROL-SPACE TLV and requests that IANA to create a new sub-registry to manage the value of the LABEL-CONTROL-SPACE TLV's 24-bits Flag field. New values are to be assigned by Standards Action [RFC8126]. Each bit should be tracked with the following qualities:

- \* Bit number (counting from bit 0 as the most significant bit)
- \* Capability description
- \* Defining RFC

Currently, there is no allocation in this registry.

| Bit  | Name       | Reference  |
|------|------------|------------|
| 0-23 | Unassigned | [This.I-D] |

### 7.3. FUNCT-ID-CONTROL-SPACE TLV's Flag field

This document defines the FUNCT-ID-CONTROL-SPACE TLV and requests that IANA to create a new sub-registry to manage the value of the FUNCT-ID-CONTROL-SPACE TLV's 24-bits Flag field. New values are to be assigned by Standards Action [RFC8126]. Each bit should be tracked with the following qualities:

- \* Bit number (counting from bit 0 as the most significant bit)
- \* Capability description
- \* Defining RFC

Currently, there is no allocation in this registry.

| Bit  | Name       | Reference  |
|------|------------|------------|
| 23   | L-Bit      | [This.I-D] |
| 0-22 | Unassigned | [This.I-D] |

## 8. Security Considerations

The security considerations described in [RFC9050], [I-D.ietf-pce-pcep-extension-pce-controller-sr], and [I-D.dhody-pce-pcep-extension-pce-controller-srv6] and apply to the extensions described in this document.

As per [RFC8231], it is RECOMMENDED that these PCEP extensions only be activated on authenticated and encrypted sessions across PCEs and PCCs belonging to the same administrative authority, using Transport Layer Security (TLS) [RFC8253] as per the recommendations and best current practices in [RFC7525] (unless explicitly set aside in [RFC8253]).

## 9. Acknowledgements

TBD.

## 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, <<https://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, <<https://www.rfc-editor.org/info/rfc5440>>.
- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <<https://www.rfc-editor.org/info/rfc8126>>.
- [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>>.

- [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>>.
- [RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <<https://www.rfc-editor.org/info/rfc8283>>.
- [I-D.ietf-pce-segment-routing-ipv6]  
Li, C., Negl, M., Sivabalan, S., Koldychev, M., Kaladharan, P., and Y. Zhu, "PCEP Extensions for Segment Routing leveraging the IPv6 data plane", Work in Progress, Internet-Draft, draft-ietf-pce-segment-routing-ipv6-12, 6 March 2022, <<https://www.ietf.org/internet-drafts/draft-ietf-pce-segment-routing-ipv6-12.txt>>.

## 10.2. Informative References

- [RFC4657] Ash, J., Ed. and J.L. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, DOI 10.17487/RFC4657, September 2006, <<https://www.rfc-editor.org/info/rfc4657>>.
- [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>>.
- [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.
- [RFC8664] Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing", RFC 8664, DOI 10.17487/RFC8664, December 2019, <<https://www.rfc-editor.org/info/rfc8664>>.

- [RFC9050] Li, Z., Peng, S., Negi, M., Zhao, Q., and C. Zhou, "Path Computation Element Communication Protocol (PCEP) Procedures and Extensions for Using the PCE as a Central Controller (PCECC) of LSPs", RFC 9050, DOI 10.17487/RFC9050, July 2021, <<https://www.rfc-editor.org/info/rfc9050>>.
- [I-D.ietf-pce-pcep-extension-pce-controller-sr] Li, Z., Peng, S., Negi, M. S., Zhao, Q., and C. Zhou, "Path Computation Element Communication Protocol (PCEP) Procedures and Extensions for Using PCE as a Central Controller (PCECC) for Segment Routing (SR) MPLS Segment Identifier (SID) Allocation and Distribution.", Work in Progress, Internet-Draft, draft-ietf-pce-pcep-extension-pce-controller-sr-04, 6 March 2022, <<https://www.ietf.org/archive/id/draft-ietf-pce-pcep-extension-pce-controller-sr-04.txt>>.
- [I-D.ietf-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-ietf-pce-state-sync-01, 20 October 2021, <<https://www.ietf.org/archive/id/draft-ietf-pce-state-sync-01.txt>>.
- [I-D.ietf-spring-segment-routing-policy] Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", Work in Progress, Internet-Draft, draft-ietf-spring-segment-routing-policy-21, 19 March 2022, <<https://www.ietf.org/archive/id/draft-ietf-spring-segment-routing-policy-21.txt>>.
- [I-D.dhody-pce-pcep-extension-pce-controller-srv6] Li, Z., Peng, S., Geng, X., and M. Negi, "Path Computation Element Communication Protocol (PCEP) Procedures and Extensions for Using the PCE as a Central Controller (PCECC) for SRv6 SID Allocation and Distribution.", Work in Progress, Internet-Draft, draft-dhody-pce-pcep-extension-pce-controller-srv6-08, 6 March 2022, <<https://www.ietf.org/internet-drafts/draft-dhody-pce-pcep-extension-pce-controller-srv6-08.txt>>.

[RFC8986] Filtsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <<https://www.rfc-editor.org/info/rfc8986>>.

[I-D.ietf-pce-binding-label-sid] Sivabalan, S., Filtsfils, C., Tantsura, J., Previdi, S., and C. L. (editor), "Carrying Binding Label/Segment Identifier (SID) in PCE-based Networks.", Work in Progress, Internet-Draft, draft-ietf-pce-binding-label-sid-14, 3 March 2022, <<https://www.ietf.org/archive/id/draft-ietf-pce-binding-label-sid-14.txt>>.

#### Appendix A. Contributors

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

EEmail: [dhruv.ietf@gmail.com](mailto:dhruv.ietf@gmail.com)

Mach Chen  
Huawei Technologies  
China

EEmail: [Mach.chen@huawei.com](mailto:Mach.chen@huawei.com)

Zhenbin Li  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

EEmail: [lizhenbin@huawei.com](mailto:lizhenbin@huawei.com)

Jie Dong  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

EEmail: [jie.dong@huawei.com](mailto:jie.dong@huawei.com)

## Authors' Addresses

Cheng Li  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing  
100095  
China  
Email: c.l@huawei.com

Hang Shi  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing  
100095  
China  
Email: shihang9@huawei.com

Aijun Wang  
China Telecom  
Beiqijia Town,  
Beijing  
Changping District, 102209  
China  
Email: wangaj3@chinatelecom.cn

Weiqiang Cheng  
China Mobile  
Email: chengweiqiang@chinamobile.com

Chao Zhou  
HPE  
Email: chaozhou\_us@yahoo.com

PCE Working Group  
Internet-Draft  
Updates: 8281 (if approved)  
Intended status: Standards Track  
Expires: December 22, 2018

C. Li  
M. Chen  
D. Dhody  
Z. Li  
J. Dong  
Huawei Technologies  
June 20, 2018

PCEP Extension for Segment Routing (SR) Bi-directional Associated Paths  
draft-li-pce-sr-bidir-path-00

Abstract

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. The Stateful PCE extensions allow stateful control of Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Label Switched Paths (LSPs) using PCEP. Furthermore, PCEP can be used for computing paths in SR networks.

This document defines PCEP extensions for grouping two reverse unidirectional SR Paths into an Associated Bidirectional SR path when using a Stateful PCE for both PCE-Initiated and PCC-Initiated LSPs as well as when using a Stateless PCE.

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.

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 22, 2018.

#### Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 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

|                                                                               |    |
|-------------------------------------------------------------------------------|----|
| 1. Introduction . . . . .                                                     | 2  |
| 2. Terminology . . . . .                                                      | 4  |
| 3. PCEP Extension for Bi-directional SR Path . . . . .                        | 4  |
| 3.1. Double-sided Bidirectional SR Path Association Group<br>Object . . . . . | 4  |
| 4. Bi-directional Flag . . . . .                                              | 5  |
| 5. Procedures of Bi-directional Path Computation . . . . .                    | 5  |
| 5.1. PCE Initiated SR Paths . . . . .                                         | 6  |
| 5.2. PCC Initiated SR Paths . . . . .                                         | 7  |
| 5.3. Stateless PCE . . . . .                                                  | 8  |
| 5.4. Error Handling . . . . .                                                 | 8  |
| 6. IANA Considerations . . . . .                                              | 9  |
| 7. Security Considerations . . . . .                                          | 9  |
| 8. Acknowledgments . . . . .                                                  | 9  |
| 9. References . . . . .                                                       | 9  |
| 9.1. Normative References . . . . .                                           | 9  |
| 9.2. Informative References . . . . .                                         | 10 |
| Authors' Addresses . . . . .                                                  | 11 |

#### 1. Introduction

Segment routing (SR) [I-D.ietf-spring-segment-routing] leverages the source routing and tunneling paradigms. SR supports to steer packets into an explicit forwarding path according to the Segment Routing Policy (SR Policy) [I-D.ietf-spring-segment-routing-policy] at the ingress node.

However, the SR Policies defined in [I-D.ietf-spring-segment-routing-policy] only supports uni-directional SR paths. For supporting bi-directional paths [I-D.cheng-spring-mpls-path-segment], new SR policies carrying Path ID and bi-directional path information are defined in [I-D.li-idr-sr-policy-path-segment-distribution].

[RFC5440] describes the Path Computation Element (PCE) Communication Protocol (PCEP). PCEP enables the communication between a Path Computation Client (PCC) and a PCE, or between PCE and PCE, for the purpose of computation of Multiprotocol Label Switching (MPLS) as well as Generalized MPLS (GMPLS) Traffic Engineering Label Switched Path (TE LSP) characteristics.

[RFC8231] specifies a set of extensions to PCEP to enable stateful control of TE LSPs within and across PCEP sessions in compliance with [RFC4657]. It includes mechanisms to effect LSP State Synchronization between PCCs and PCEs, delegation of control over LSPs to PCEs, and PCE control of timing and sequence of path computations within and across PCEP sessions. The model of operation where LSPs are initiated from the PCE is described in [RFC8281].

[I-D.ietf-pce-segment-routing] specifies extensions to the Path Computation Element Protocol (PCEP) [RFC5440] for SR networks, that allow a stateful PCE to compute and initiate SR-TE paths, as well as a PCC to request, report or delegate SR paths.

[I-D.negi-pce-segment-routing-ipv6] extend PCEP to support SR for IPv6 data plane.

[I-D.ietf-pce-association-group] introduces a generic mechanism to create a grouping of LSPs which can then be used to define associations between a set of LSPs and/or a set of attributes, for example primary and secondary LSP associations, and is equally applicable to the active and passive modes of a Stateful PCE [RFC8231] or a stateless PCE [RFC5440].

[I-D.ietf-pce-association-bidir] defines PCEP extensions for grouping two reverse unidirectional MPLS TE LSPs into an Associated Bidirectional LSP when using a Stateful PCE for both PCE-Initiated and PCC-Initiated LSPs as well as when using a Stateless PCE.

This document extends the bidirectional association to segment routing by specifying PCEP extensions for grouping two reverse unidirectional SR paths into a bi-directional SR path.

[I-D.li-pce-sr-path-segment] defines a procedure for path ID in PCEP for SR by defining the PATH-ID TLV. The path ID can be a path segment in SR-MPLS [I-D.cheng-spring-mpls-path-segment], or a path ID

in SRv6 [I-D.li-spring-passive-pm-for-srv6-npl], or other IDs that can identify an SR path. The PATH-ID MUST be included for associated bidirectional SR paths.

## 2. Terminology

This memo makes use of the terms defined in [I-D.ietf-pce-segment-routing]. The reader is assumed to be familiar with the terminology defined in [RFC5440], [RFC8231], [RFC8281], [I-D.ietf-pce-association-group] and [I-D.ietf-pce-association-bidir].

## 3. PCEP Extension for Bi-directional SR Path

As per [I-D.ietf-pce-association-group], LSPs are associated by adding them to a common association group. [I-D.ietf-pce-association-bidir] specifies PCEP extensions for grouping two reverse unidirectional MPLS-TE LSPs into an Associated Bidirectional LSP for both single-sided and double-sided initiation cases by defining two new Bidirectional LSP Association Groups.

This document extends the procedure for SR bidirectional associated paths by defining a new bidirectional association type (i.e. Double-sided Bi-directional SR Path Association Group). The document further describe the mechanism of associating two unidirectional SR path into a bidirectional SR path. [I-D.li-pce-sr-path-segment] defines a procedure for path ID in PCEP for SR by defining the PATH-ID TLV. The bidirectional SR path MUST also use the PATH-ID TLV.

### 3.1. Double-sided Bidirectional SR Path Association Group Object

As defined in [I-D.ietf-pce-association-bidir], two LSPs are associated as a bi-directional MPLS-TE LSP by a common bi-directional LSP association group. For associating two SR paths, this document defines a new association group called 'Double-sided Bidirectional SR Path Association Group' as follows:

- o Association Type (TBD) = Double-sided Bidirectional SR Path Association Group

Similar to other bidirectional associations, this Association Type is operator-configured in nature and statically created by the operator on the PCEP peers. The paths belonging to this association is conveyed via PCEP messages to the PCEP peer. Operator-configured Association Range TLV [I-D.ietf-pce-association-group] MUST NOT be sent for these Association Types, and MUST be ignored, so that the entire range of association ID can be used for them. The handling of the Association ID, Association Source, optional Global Association

Source and optional Extended Association ID in this association are set in the same way as [I-D.ietf-pce-association-bidir].

A member of the Double-sided Bi-directional SR Path Association Group can take the role of a forward or reverse SR path and follows the rules similar to the rules defined in [I-D.ietf-pce-association-bidir] for LSPs.

- o An SR path (forward or reverse) can not be part of more than one Double-sided Bi-directional SR Path Association Group.
- o The endpoints of the SR paths in this associations cannot be different.

For describing the SR paths in this association group, such as direction and co-routed information, this association group reuses the Bi-directional LSP Association Group TLV defined in [I-D.ietf-pce-association-bidir]. All fields and processing rules are as per [I-D.ietf-pce-association-bidir].

#### 4. Bi-directional Flag

As defined in [RFC5440], the B-flag in RP object MUST be set when the PCC specifies that the path computation request relates to a bi-directional TE LSP. In this document, the B-flag also MUST be set when the PCC specifies that the path computation request relates to a bi-directional SR path. Likely, when a stateful PCE initiates or updates a bi-directional SR paths including LSPs and SR paths, the B-flag in SRP object [I-D.ietf-pce-pcep-stateful-pce-gmpls] MUST be set as well.

#### 5. Procedures of Bi-directional Path Computation

Two uni-directional SR paths can be associated by the association group object as specified in [I-D.ietf-pce-association-group]. A bidirectional LSP association group object is defined in [I-D.ietf-pce-association-bidir] (for MPLS-TE). This documents extends the mechanism for bidirectional SR paths. Two SR paths can be associated together by including the Bi-directional SR Path Association Group in the PCEP messages. The PATH-ID TLV [I-D.li-pce-sr-path-segment] MUST also be included in the LSP object for these SR paths.

There is also a need to include the reverse direction path in the PCEP messages, to do this the PCE SHOULD inform the reverse SR path to the ingress PCC and vice versa. To achieve this a PCInitiate message for the reverse SR path is sent to the ingress PCC and a PCInitiate message for the forward SR path is sent to the egress PCC

(with the same association group). These PCInitiate message MUST not trigger initiation of SR paths. The information of reverse direction path can be used for several scenarios, such as directed BFD [I-D.ietf-mppls-bfd-directed].

## 5.1. PCE Initiated SR Paths

As specified in [I-D.ietf-pce-association-group] Bidirectional SR Association Group can be created by a Stateful PCE.

- o Stateful PCE can create and update the forward and reverse SR path independently for Double-sided Bi-directional SR Path Association Groups.
- o Stateful PCE can establish and remove the association relationship on a per SR path basis.
- o Stateful PCE can create and update the SR path and the association on a PCC via PCInitiate and PCUpd messages, respectively, using the procedures described in [I-D.ietf-pce-association-group].
- o The Path-ID TLV MUST be included for each SR path in the LSP object.
- o The opposite direction SR path (LSP2(R) at S, LSP1(F) at D ) SHOULD be informed via PCInitiate message with the matching association group.

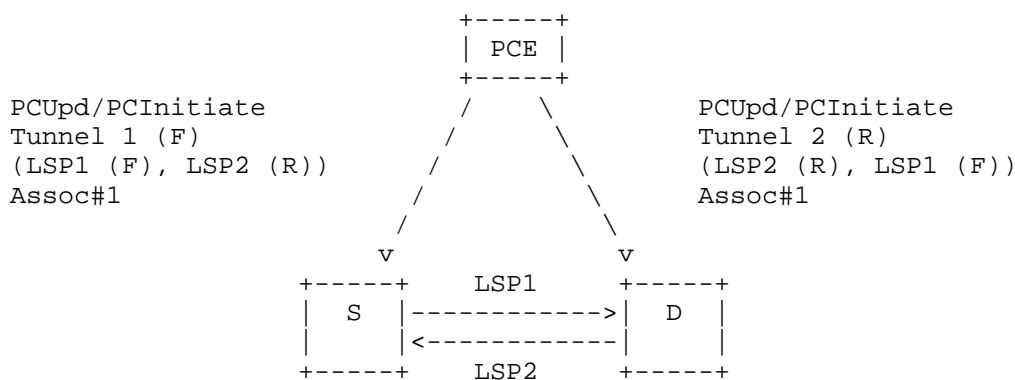


Figure 1: PCE-Initiated Double-sided Bidirectional SR Path

## 5.2. PCC Initiated SR Paths

As specified in [I-D.ietf-pce-association-group] Bidirectional SR Association Group can also be created by a PCC.

- o PCC can create and update the forward and reverse SR paths independently for Double-sided Bi-directional SR Path Association Groups.
- o PCC can establish and remove the association relationship on a per SR path basis.
- o PCC MUST report the change in the association group of an SR path to PCE(s) via PCRpt message.
- o PCC can report the forward and reverse SR paths independently to PCE(s) via PCRpt message.
- o PCC can delegate the forward and reverse SR paths independently to a Stateful PCE, where PCE would control the SR paths.
- o Stateful PCE can update the SR paths in the Double-sided Bi-directional SR Path Association Group via PCUpd message, using the procedures described in [I-D.ietf-pce-association-group].
- o The Path-ID TLV MUST be handled as defined in [I-D.li-pce-sr-path-segment].
- o The opposite direction SR path (LSP2(R) at S, LSP1(F) at D ) SHOULD be informed via PCInitiate message with the matching association group.

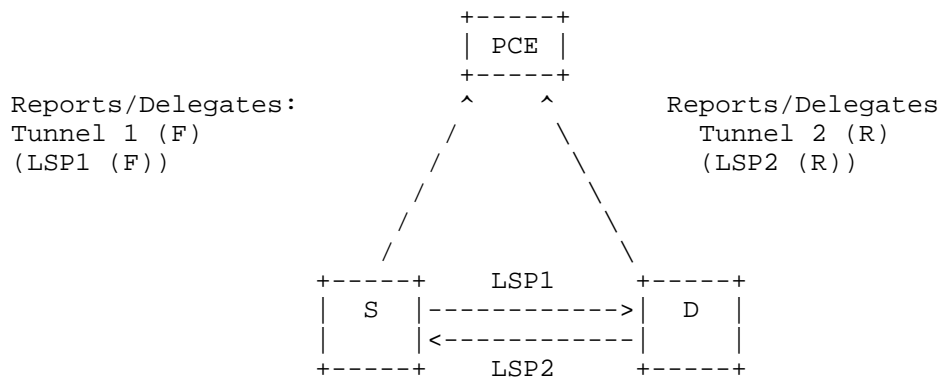
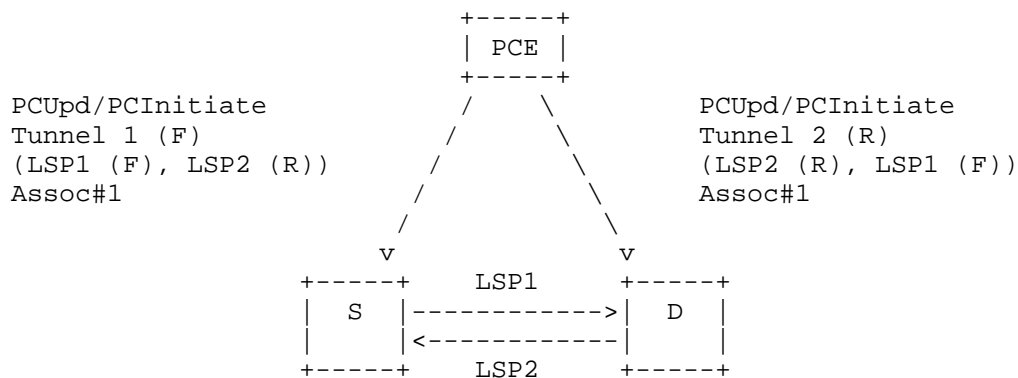


Figure 2a: PCC-Initiated Double-sided Bidirectional SR Path

Figure 2b: PCC-Initiated Double-sided Bidirectional SR Path  
along with opposite direction SR path

### 5.3. Stateless PCE

As defined in [I-D.ietf-pce-association-bidir], for a stateless PCE, it might be useful to associate a path computation request to an association group, thus enabling it to associate a common set of configuration parameters or behaviors with the request. A PCC can request co-routed or non co-routed forward and reverse direction paths from a stateless PCE for a bidirectional LSP association group.

### 5.4. Error Handling

The error handling as described in [I-D.ietf-pce-association-bidir] continue to apply.

## 6. IANA Considerations

TBA

## 7. Security Considerations

TBA

## 8. Acknowledgments

## 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>>.
- [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>>.
- [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>>.
- [I-D.ietf-pce-association-group]  
Minei, I., Crabbe, E., Sivabalan, S., Ananthakrishnan, H., Dhody, D., and Y. Tanaka, "PCEP Extensions for Establishing Relationships Between Sets of LSPs", draft-ietf-pce-association-group-06 (work in progress), June 2018.

- [I-D.ietf-pce-association-bidir]  
Barth, C., Gandhi, R., and B. Wen, "PCEP Extensions for Associated Bidirectional Label Switched Paths (LSPs)", draft-ietf-pce-association-bidir-01 (work in progress), May 2018.
- [I-D.ietf-pce-pcep-stateful-pce-gmpls]  
Zhang, X., Lee, Y., Zhang, F., Casellas, R., Dios, O., and Z. Ali, "Path Computation Element (PCE) Protocol Extensions for Stateful PCE Usage in GMPLS-controlled Networks", draft-ietf-pce-pcep-stateful-pce-gmpls-08 (work in progress), February 2018.
- [I-D.negi-pce-segment-routing-ipv6]  
Negi, M., Kaladharan, P., Dhody, D., and S. Sivabalan, "PCEP Extensions for Segment Routing leveraging the IPv6 data plane", draft-negi-pce-segment-routing-ipv6-01 (work in progress), March 2018.
- [I-D.li-pce-sr-path-segment]  
Li, C., Chen, M., Dong, J., Li, Z., and D. Dhody, "Path Computation Element Communication Protocol (PCEP) Extension for Path Identification in Segment Routing (SR)", draft-li-pce-sr-path-segment-00 (work in progress), June 2018.

## 9.2. Informative References

- [RFC4657] Ash, J., Ed. and J. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, DOI 10.17487/RFC4657, September 2006, <<https://www.rfc-editor.org/info/rfc4657>>.
- [I-D.ietf-pce-segment-routing]  
Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "PCEP Extensions for Segment Routing", draft-ietf-pce-segment-routing-11 (work in progress), November 2017.
- [I-D.ietf-mpls-bfd-directed]  
Mirsky, G., Tantsura, J., Varlashkin, I., and M. Chen, "Bidirectional Forwarding Detection (BFD) Directed Return Path", draft-ietf-mpls-bfd-directed-08 (work in progress), December 2017.

[I-D.ietf-spring-segment-routing-policy]

Filsfils, C., Sivabalan, S., daniel.voyer@bell.ca, d., bogdanov@google.com, b., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-01 (work in progress), June 2018.

[I-D.ietf-spring-segment-routing]

Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-15 (work in progress), January 2018.

[I-D.cheng-spring-mpls-path-segment]

Cheng, W., Wang, L., Li, H., Chen, M., Zigler, R., and S. Zhan, "Path Segment in MPLS Based Segment Routing Network", draft-cheng-spring-mpls-path-segment-01 (work in progress), March 2018.

[I-D.li-spring-passive-pm-for-srv6-np]

Li, C. and M. Chen, "Passive Performance Measurement for SRv6 Network Programming", draft-li-spring-passive-pm-for-srv6-np-00 (work in progress), March 2018.

[I-D.li-idr-sr-policy-path-segment-distribution]

Li, C., Chen, M., Dong, J., and Z. Li, "Segment Routing Policies for Path Segment and Bi-directional Path", draft-li-idr-sr-policy-path-segment-distribution-00 (work in progress), April 2018.

Authors' Addresses

Cheng Li  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: chengli13@huawei.com

Mach(Guoyi) Chen  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: Mach.chen@huawei.com

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

Email: dhruv.ietf@gmail.com

Zhenbin Li  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: lizhenbin@huawei.com

Jie Dong  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: jie.dong@huawei.com

PCE Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: December 22, 2018

C. Li  
M. Chen  
J. Dong  
Z. Li  
D. Dhody  
Huawei Technologies  
June 20, 2018

Path Computation Element Communication Protocol (PCEP) Extension for  
Path Identification in Segment Routing (SR)  
draft-li-pce-sr-path-segment-00

## Abstract

The Path Computation Element (PCE) provides path computation functions in support of traffic engineering in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks.

The Source Packet Routing in Networking (SPRING) architecture describes how Segment Routing (SR) can be used to steer packets through an IPv6 or MPLS network using the source routing paradigm. A Segment Routed Path can be derived from a variety of mechanisms, including an IGP Shortest Path Tree (SPT), explicit configuration, or a Path Computation Element (PCE).

Path identification is needed for several use cases such as performance measurement in Segment Routing (SR) network. This document specifies extensions to the Path Computation Element Protocol (PCEP) to support requesting, replying, reporting and updating the path identifier between PCEP speakers.

## 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.

## 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 22, 2018.

#### Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 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

|                                                         |    |
|---------------------------------------------------------|----|
| 1. Introduction . . . . .                               | 3  |
| 2. Terminology . . . . .                                | 4  |
| 3. Overview of Path ID Extensions in PCEP . . . . .     | 4  |
| 4. Objects and TLVs . . . . .                           | 5  |
| 4.1. The OPEN Object . . . . .                          | 5  |
| 4.1.1. The SR PCE Capability sub-TLV . . . . .          | 5  |
| 4.1.2. The SRv6 PCE Capability sub-TLV . . . . .        | 5  |
| 4.1.3. PCECC-CAPABILITY sub-TLV . . . . .               | 5  |
| 4.2. LSP Object . . . . .                               | 6  |
| 4.2.1. Path ID TLV . . . . .                            | 6  |
| 4.3. FEC Object . . . . .                               | 8  |
| 4.4. CCI Object . . . . .                               | 9  |
| 5. Operations . . . . .                                 | 9  |
| 5.1. PCC Allocated Path ID . . . . .                    | 9  |
| 5.1.1. Ingress PCC Allocated Path ID . . . . .          | 9  |
| 5.2. PCE Allocated Path ID . . . . .                    | 10 |
| 5.2.1. PCE Controlled ID Spaces Advertisement . . . . . | 10 |
| 5.2.2. Ingress PCC request Path ID to PCE . . . . .     | 10 |
| 5.2.3. PCE allocated Path ID on its own . . . . .       | 11 |
| 5.3. Two Label Solution . . . . .                       | 12 |
| 6. IANA Considerations . . . . .                        | 12 |

|                                       |    |
|---------------------------------------|----|
| 7. Security Considerations . . . . .  | 12 |
| 8. Acknowledgments . . . . .          | 13 |
| 9. References . . . . .               | 13 |
| 9.1. Normative References . . . . .   | 13 |
| 9.2. Informative References . . . . . | 14 |
| Authors' Addresses . . . . .          | 15 |

## 1. Introduction

[RFC5440] describes the Path Computation Element (PCE) Communication Protocol (PCEP). PCEP enables the communication between a Path Computation Client (PCC) and a PCE, or between PCE and PCE, for the purpose of computation of Multiprotocol Label Switching (MPLS) as well as Generalized MPLS (GMPLS) Traffic Engineering Label Switched Path (TE LSP) characteristics.

[RFC8231] specifies a set of extensions to PCEP to enable stateful control of TE LSPs within and across PCEP sessions in compliance with [RFC4657]. It includes mechanisms to effect LSP State Synchronization between PCCs and PCEs, delegation of control over LSPs to PCEs, and PCE control of timing and sequence of path computations within and across PCEP sessions. The model of operation where LSPs are initiated from the PCE is described in [RFC8281].

[I-D.zhao-pce-pcep-extension-for-pce-controller] specify the procedures and PCEP protocol extensions for using the PCE as the central controller for static LSPs, where LSPs can be provisioned as explicit label instructions at each hop on the end-to-end path.

Segment routing (SR) [I-D.ietf-spring-segment-routing] leverages the source routing and tunneling paradigms. SR supports steering packets into an explicit forwarding path according to the Segment Routing Policy (SR Policy) [I-D.ietf-spring-segment-routing-policy] at the ingress node.

An SR path needs to be identified in some use cases such as performance measurement. For identifying an SR path, [I-D.cheng-spring-mpls-path-segment] introduces a new segment that is referred to as Path Segment.

[I-D.li-idr-sr-policy-path-segment-distribution] defines extensions to BGP to distribute SR policies carrying Path segment identifier.

[I-D.ietf-pce-segment-routing] specifies extensions to the Path Computation Element Protocol (PCEP) [RFC5440] for SR networks, that allow a stateful PCE to compute and initiate SR-TE paths, as well as a PCC to request, report or delegate SR paths.

[I-D.negi-pce-segment-routing-ipv6] extend PCEP to support SR for IPv6 data plane.

[I-D.zhao-pce-pcep-extension-pce-controller-sr] specifies the procedures and PCEP protocol extensions when a PCE-based controller is also responsible for configuring the forwarding actions on the routers (SR SID distribution in this case), in addition to computing the paths for packet flows in a segment routing network and telling the edge routers what instructions to attach to packets as they enter the network.

This document specifies a mechanism to carry the SR path identification information in PCEP messages [RFC5440] [RFC8231] [RFC8281]. The path ID can be a path segment in SR-MPLS [I-D.cheng-spring-mpls-path-segment], or a path ID in SRv6 [I-D.li-spring-passive-pm-for-srv6-np], or other IDs that can identify the SR path. This document also extends the PCECC-SR mechanism to inform the path ID to the egress PCC.

## 2. Terminology

This memo makes use of the terms defined in [RFC4655], [I-D.ietf-pce-segment-routing], and [I-D.ietf-spring-segment-routing].

## 3. Overview of Path ID Extensions in PCEP

This document specifies a mechanism of encoding (and allocating) path ID (path segment in case of MPLS, path ID in case of IPv6, etc) in PCEP extensions. For supporting path ID in PCEP, several TLVs and flags are defined. The formats of the objects and TLVs are described in Section 4. The procedures of path ID allocation are described in Section 5.

There are various modes of operations, such as -

- o The path ID can be allocated by Ingress PCC itself and informed to the PCE. The PCE can then inform the egress PCC.
- o The PCC can also request PCE to allocate the path ID, in this case, the PCE would allocate and inform the assigned path ID to the ingress/egress PCC using PCEP messages.
- o For PCE can allocate a path ID on its own accord and inform the ingress/egress PCC , useful for PCE-initiated LSPs.

The path ID information to the ingress PCC and PCE is exchanged via an extension to [I-D.ietf-pce-segment-routing] and [I-D.negi-pce-segment-routing-ipv6]. The path ID information to the egress PCC is informed via an extension to the PCECC-SR procedures [I-D.zhao-pce-pcep-extension-pce-controller-sr].

For the PCE to allocate a path ID, the PCE MUST be aware of the path ID space from the PCCs. This is done via mechanism as described in [I-D.li-pce-controlled-id-space].

[Editor's note - There is currently no mechanism for the PCE to ask PCC to allocate path ID. Further discussion is needed to check if that would be useful in any way.]

#### 4. Objects and TLVs

##### 4.1. The OPEN Object

###### 4.1.1. The SR PCE Capability sub-TLV

[I-D.ietf-pce-segment-routing] defined a new Path Setup Type (PST) and SR-PCE-CAPABILITY sub-TLV for SR. PCEP speakers use this sub-TLV to exchange information about their SR capability. The TLV includes a Flags field and one bit (L-flag) was allocated in [I-D.ietf-pce-segment-routing].

This document adds an additional flag for path ID allocation, as follows -

P (Path Identification bit): A PCEP speaker sets this flag to 1 to indicate that it has the capability to encode SR path identification (path segment, as per [I-D.cheng-spring-mpls-path-segment]).

###### 4.1.2. The SRv6 PCE Capability sub-TLV

[I-D.negi-pce-segment-routing-ipv6] defined a new Path Setup Type (PST) and SRv6-PCE-CAPABILITY sub-TLV for SRv6. PCEP speakers use this sub-TLV to exchange information about their SRv6 capability. The TLV includes a Flags field and one bit (L-flag) was allocated in [I-D.negi-pce-segment-routing-ipv6].

This document adds an additional flag for path ID allocation, as follows -

P (Path Identification bit): A PCEP speaker sets this flag to 1 to indicate that it has the capability to encode SRv6 path identification.

###### 4.1.3. PCECC-CAPABILITY sub-TLV

The PCECC Capability as per [I-D.zhao-pce-pcep-extension-pce-controller-sr] MUST also be advertised on the egress PCEP session, along with the SR sub-TLVs.

This is needed to ensure that the PCE can inform the egress PCC of the path ID via PCECC mechanism as described in this document.

#### 4.2. LSP Object

The LSP Object is defined in Section 7.3 of [RFC8231]. This document adds the following flags to the LSP Object:

P (Path Identification Allocation bit): If the bit is set to 1, it indicates that the path identifier needs to be allocated by the PCE for this LSP. A PCC would set this bit to 1 to request for allocation of path identifier by the PCE in the PCReq or PCRpt message. A PCE would also set this bit to 1 to indicate that the path identifier is allocated by PCE and encoded in the PCRep, PCUpd or PCInitiate message (the PATH-ID TLV MUST be present in LSP object).

##### 4.2.1. Path ID TLV

The PATH-ID TLV is an optional TLV for use in the LSP Object for path ID allocation. The type of this TLV is to be allocated by IANA. The format is shown below.

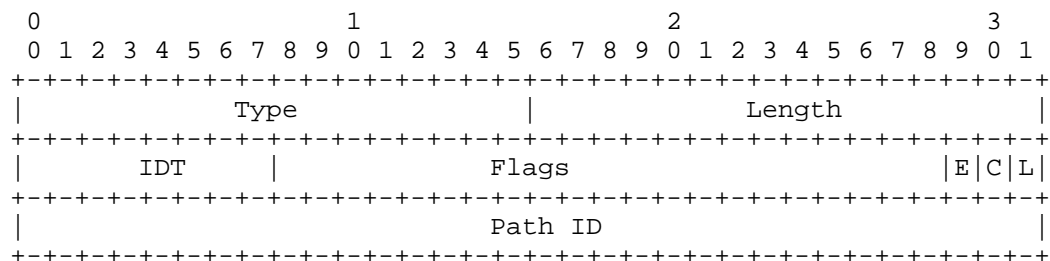


Figure 1: The PATH-ID TLV Format

The type (16-bit) of the TLV is TBA (to be allocated by IANA). The length (16-bit) has a fixed value of 8 octets. The value contains the following fields:

IDT (The ID type - 8 bits): The IDT field specifies the type of the Path ID field, which carries a Path ID corresponding to the SR path.

- \* 0: MPLS Path segment, which is an MPLS label as defined in [I-D.cheng-spring-mpls-path-segment]. The PST type MUST be set to SR (MPLS).

- \* 1: SRv6 Path ID, which is a 4-octet integer as defined in [I-D.li-spring-passive-pm-for-srv6-npl]. The PST type MUST be set to SRv6.

Flags (24 bits): Three flag is currently defined:

- \* L-Bit (Local/Global - 1 bit): If set, then the Path ID carried by the PATH-ID TLV has local significance. If not set, then the Path ID carried by this TLV has global significance (i.e. Path ID is global within an SR domain).
- \* C-bit (PCC/PCE - 1 bit): If set, then the Path ID carried by the PATH-ID TLV has been allocated by the PCC. If not set, then the Path ID carried by this TLV has been allocated by the PCE.
- \* E-bit (Egress/Ingress - 1 bit): If set, then the Path ID carried by the PATH-ID TLV has been allocated from the Egress Path ID space. If not set, then the Path ID carried by this TLV has been allocated from the Ingress Path ID space.
- \* The unassigned bits MUST be set to 0 and MUST be ignored at receipt.

Path ID: The path ID of an SR path. The path ID type is indicated by the ID Type field. It can be a path segment [I-D.cheng-spring-mpls-path-segment] in MPLS label format. Or it can be a 4 octets integer ID as defined in [I-D.li-spring-passive-pm-for-srv6-npl] or other IDs that can identify a path.

Only one instance of each TLV is processed, if more than one TLV of each type is included, the first one is processed and others MUST be ignored.

When the path ID allocation is enable, a PATH-ID TLV SHOULD be included in the LSP object.

If the path ID is allocated by the ingress node, a PATH-ID TLV MUST be included in a LSP object (with C-bit set and E-bit is unset) in the PCEP message from PCC. In this case the P flag in LSP object is set to 0.

If the PCC request the path ID to be allocated by the PCE, P flag in LSP object is set to 1 and Path ID TLV MAY be skipped. After the PCE has allocated a path ID, it MUST include the PATH-ID TLV in a LSP object (with C-bit unset), the E-bit is set by PCE based on the path ID space from which the allocation is made.

If the PCE allocated the path ID on its own accord, a PATH-ID TLV MUST be included in a LSP object (with C-bit unset), the E-bit is set by PCE based on the path ID space from which the allocation is made.

#### 4.3. FEC Object

The FEC Object [I-D.zhao-pce-pcep-extension-pce-controller-sr] is used to specify the FEC information and MAY be carried within PCInitiate or PCRpt message for the PCECC-SR operations. The PCE MUST inform the Path Identification information to the Egress PCC. To do this, this document extends the procedures of [I-D.zhao-pce-pcep-extension-pce-controller-sr] by defining a new FEC object type for Path.

FEC Object-Type is TBD 'Path'.

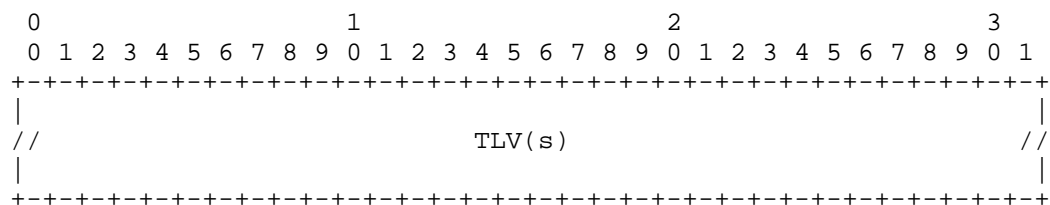


Figure 2: The path FEC object Format

One or more following TLV(s) are allowed in the 'path' FEC object -

- o SYMBOLIC-PATH-NAME TLV: As defined in [RFC8231], it is a human-readable string that identifies an LSP in the network.
- o LSP-IDENTIFIERS TLVs: As defined in [RFC8231], it is optional for SR, but could be used to encode the source, destination and other identification information for the path.
- o SPEAKER-ENTITY-ID TLV: As defined in [RFC8232], a unique identifier for the PCEP speaker, it is used to identify the Ingress PCC.

Either SYMBOLIC-PATH-NAME TLV or LSP-IDENTIFIERS TLV MUST be included. SPEAKER-ENTITY-ID TLV is optional. Only one instance of each TLV is processed, if more than one TLV of each type is included, the first one is processed and others MUST be ignored.

#### 4.4. CCI Object

The Central Control Instructions (CCI) Object is used by the PCE to specify the forwarding instructions is defined in [I-D.zhao-pce-pcep-extension-for-pce-controller]. Further [I-D.zhao-pce-pcep-extension-pce-controller-sr] defined a CCI object type for SR.

The Path ID information is encoded directly in the CCI SR object. The Path ID TLV as described in the Section 4.2.1, MAY also be included in the CCI SR object.

#### 5. Operations

The path ID allocation and encoding is as per the stateful PCE operations for segment routing. The procedures are as per the corresponding extensions defined in [I-D.ietf-pce-segment-routing] and [I-D.negi-pce-segment-routing-ipv6] (which are further based on [RFC8231] and [RFC8281]). The additional operations for path identification are defined in this section.

To notify the path ID to the Egress PCC, the procedures are as per the PCECC-SR [I-D.zhao-pce-pcep-extension-pce-controller-sr] (which is based on [I-D.zhao-pce-pcep-extension-for-pce-controller]). The additional operations are defined in this section.

##### 5.1. PCC Allocated Path ID

###### 5.1.1. Ingress PCC Allocated Path ID

The Ingress PCC could allocate the Path ID and inform the PCE via the PCRpt message as per [RFC8231]. The PATH-ID TLV MUST be included in a LSP object in the PCEP message from PCC. The P flag in LSP object is set to 0. On receiving this report, the PCE updates the information in its database. The active PCE (where the LSP is delegated) further informs the egress about the path ID allocated by the PCC using the PCInitiate message as described in [I-D.zhao-pce-pcep-extension-pce-controller-sr].

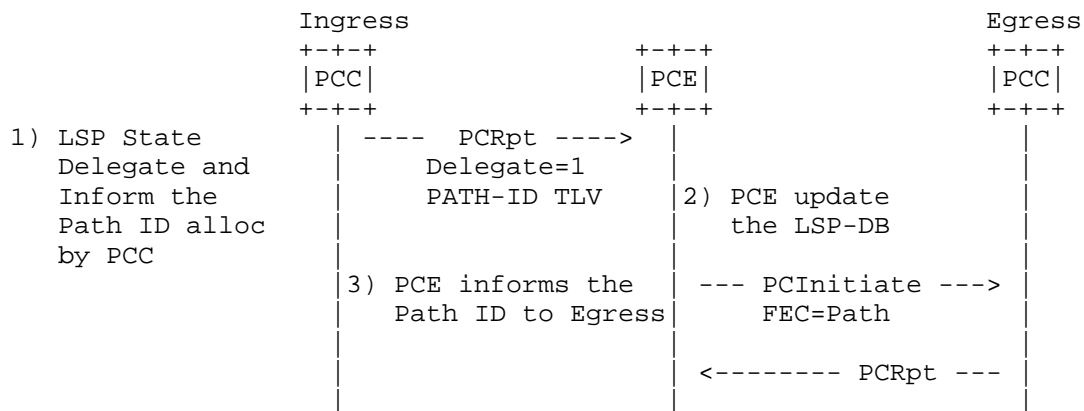


Figure 3: Ingress PCC Allocated Path ID

## 5.2. PCE Allocated Path ID

### 5.2.1. PCE Controlled ID Spaces Advertisement

For allocating the path IDs to SR paths by the PCEs, the PCE controlled ID Spaces MUST be known at PCEs via configurations or any other mechanism. The PCE controlled ID spaces MAY be advertised as described in [I-D.li-pce-controlled-id-space].

### 5.2.2. Ingress PCC request Path ID to PCE

The ingress PCC could request the path ID to be allocated by the PCE via PCRpt message as per [RFC8231]. The delegate flag (D-flag) MUST also be set for this LSP. The PATH-ID TLV MAY be included with Path ID set to 0x0000. The active PCE would allocated the path ID as per the PATH-ID flags and in case PATH-ID is not included, the PCE MUST act based on the local policy. The PCE would further respond to Ingress PCC with PCUpd message as per [RFC8231] and MUST include the PATH-ID TLV in a LSP object. The PCE would further inform the egress PCC about the path ID allocated by the PCE using the PCInitiate message as described in [I-D.zhao-pce-pcep-extension-pce-controller-sr].

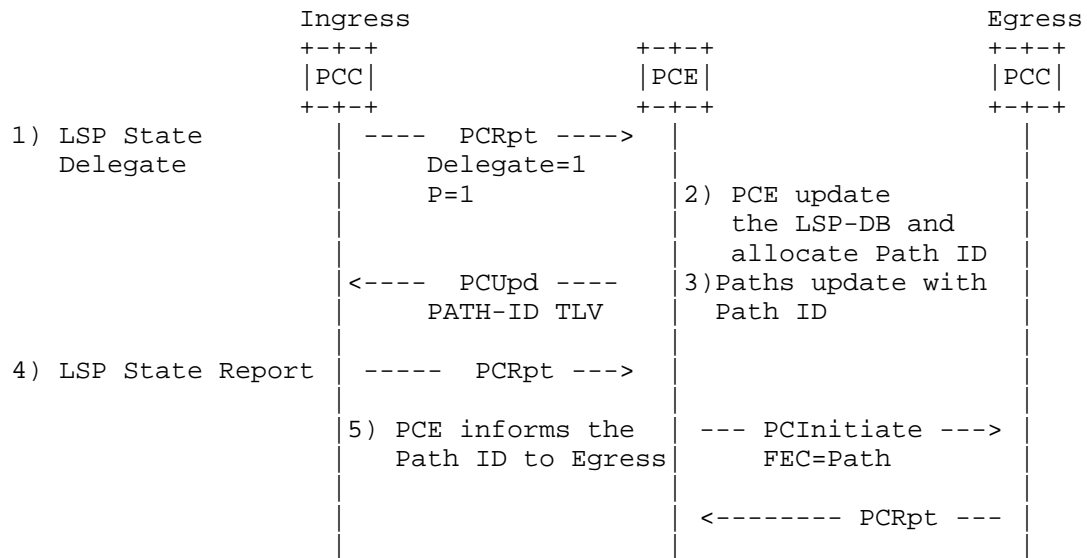


Figure 4: Ingress PCC request Path ID to PCE

## 5.2.3. PCE allocated Path ID on its own

The PCE could allocate the path ID on its own accord for a PCE-Initiated (or delegated LSP). The allocated path ID needs to be informed to the Ingress and Egress PCC. The PCE would use the PCInitiate message [RFC8281] or PCUpd message [RFC8231] towards the Ingress PCC and MUST include the PATH-ID TLV in the LSP object. The PCE would further inform the egress PCC about the path ID allocated by the PCE using the PCInitiate message as described in [I-D.zhao-pce-pcep-extension-pce-controller-sr].

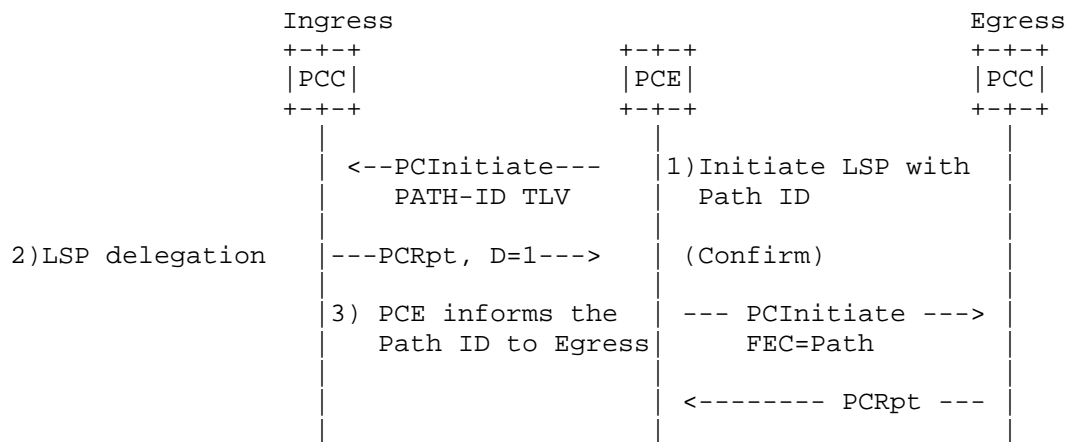


Figure 5: PCE allocated Path ID on its own

### 5.3. Two Label Solution

[I-D.cheng-spring-mpls-path-segment] describe a Path Segment to uniquely identify an SR path in a specific context. (e.g., in the context of the egress node or ingress node of an SR path, or within an SR domain). It further describes two solution based on 'one label' or 'two labels' solution. For the latter, two segments (Source segment and Path segment) are used to identify an SR path where the source segment is a global node segment which can uniquely identify a node within the SR domain (it is NOT used for forwarding and indicates that a Path segment immediately follows). The combination of Source segment and Path segment uniquely identify an SR Path with an SR domain.

The procedure described in this document allocates and encode the Path Segment only. It is expected that the Egress PCC is aware of the Source segment by some other procedures. These procedures could be IGP, PCECC-SR, or some other mechanisms.

### 6. IANA Considerations

TBA

### 7. Security Considerations

TBA

## 8. Acknowledgments

## 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>>.
- [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>>.
- [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>>.
- [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>>.
- [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>>.
- [I-D.ietf-pce-segment-routing] Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "PCEP Extensions for Segment Routing", draft-ietf-pce-segment-routing-11 (work in progress), November 2017.

- [I-D.negi-pce-segment-routing-ipv6]  
Negi, M., Kaladharan, P., Dhody, D., and S. Sivabalan,  
"PCEP Extensions for Segment Routing leveraging the IPv6  
data plane", draft-negi-pce-segment-routing-ipv6-01 (work  
in progress), March 2018.
- [I-D.li-pce-controlled-id-space]  
Li, C., Chen, M., Dong, J., Li, Z., and D. Dhody, "PCE  
Controlled ID Space", draft-li-pce-controlled-id-space-00  
(work in progress), June 2018.
- [I-D.zhao-pce-pcep-extension-pce-controller-sr]  
Zhao, Q., Li, Z., Dhody, D., Karunanithi, S., Farrel, A.,  
and C. Zhou, "PCEP Procedures and Protocol Extensions for  
Using PCE as a Central Controller (PCECC) of SR-LSPs",  
draft-zhao-pce-pcep-extension-pce-controller-sr-03 (work  
in progress), June 2018.
- [I-D.zhao-pce-pcep-extension-for-pce-controller]  
Zhao, Q., Li, Z., Dhody, D., Karunanithi, S., Farrel, A.,  
and C. Zhou, "PCEP Procedures and Protocol Extensions for  
Using PCE as a Central Controller (PCECC) of LSPs", draft-  
zhao-pce-pcep-extension-for-pce-controller-08 (work in  
progress), June 2018.

## 9.2. 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,  
<<https://www.rfc-editor.org/info/rfc4655>>.
- [RFC4657] Ash, J., Ed. and J. Le Roux, Ed., "Path Computation  
Element (PCE) Communication Protocol Generic  
Requirements", RFC 4657, DOI 10.17487/RFC4657, September  
2006, <<https://www.rfc-editor.org/info/rfc4657>>.
- [I-D.ietf-spring-segment-routing]  
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B.,  
Litkowski, S., and R. Shakir, "Segment Routing  
Architecture", draft-ietf-spring-segment-routing-15 (work  
in progress), January 2018.
- [I-D.ietf-spring-segment-routing-policy]  
Filsfils, C., Sivabalan, S., daniel.voyer@bell.ca, d.,  
bogdanov@google.com, b., and P. Mattes, "Segment Routing  
Policy Architecture", draft-ietf-spring-segment-routing-  
policy-01 (work in progress), June 2018.

[I-D.cheng-spring-mpls-path-segment]

Cheng, W., Wang, L., Li, H., Chen, M., Zigler, R., and S. Zhan, "Path Segment in MPLS Based Segment Routing Network", draft-cheng-spring-mpls-path-segment-01 (work in progress), March 2018.

[I-D.li-spring-passive-pm-for-srv6-np]

Li, C. and M. Chen, "Passive Performance Measurement for SRv6 Network Programming", draft-li-spring-passive-pm-for-srv6-np-00 (work in progress), March 2018.

[I-D.li-idr-sr-policy-path-segment-distribution]

Li, C., Chen, M., Dong, J., and Z. Li, "Segment Routing Policies for Path Segment and Bi-directional Path", draft-li-idr-sr-policy-path-segment-distribution-00 (work in progress), April 2018.

Authors' Addresses

Cheng Li  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: chenglil3@huawei.com

Mach(Guoyi) Chen  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: Mach.chen@huawei.com

Jie Dong  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: jie.dong@huawei.com

Zhenbin Li  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: lizhenbin@huawei.com

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

Email: dhruv.ietf@gmail.com

PCE Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: February 20, 2020

C. Li  
M. Chen  
Huawei Technologies  
W. Cheng  
China Mobile  
J. Dong  
Z. Li  
Huawei Technologies  
R. Gandhi  
Cisco Systems, Inc.  
Q. Xiong  
ZTE Corporation  
August 19, 2019

Path Computation Element Communication Protocol (PCEP) Extension for  
Path Segment in Segment Routing (SR)  
draft-li-pce-sr-path-segment-08

Abstract

The Path Computation Element (PCE) provides path computation functions in support of traffic engineering in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks.

The Source Packet Routing in Networking (SPRING) architecture describes how Segment Routing (SR) can be used to steer packets through an IPv6 or MPLS network using the source routing paradigm. A Segment Routed Path can be derived from a variety of mechanisms, including an IGP Shortest Path Tree (SPT), explicit configuration, or a Path Computation Element (PCE).

Path identification is needed for several use cases such as performance measurement in Segment Routing (SR) network. This document specifies extensions to the Path Computation Element Communication Protocol (PCEP) to support requesting, replying, reporting and updating the Path Segment ID (Path SID) between PCEP speakers.

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 February 20, 2020.

#### 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 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

|                                                                |    |
|----------------------------------------------------------------|----|
| 1. Introduction . . . . .                                      | 3  |
| 2. Terminology . . . . .                                       | 4  |
| 2.1. Requirements Language . . . . .                           | 4  |
| 3. Overview of Path Segment Extensions in PCEP . . . . .       | 4  |
| 4. Objects and TLVs . . . . .                                  | 5  |
| 4.1. The OPEN Object . . . . .                                 | 5  |
| 4.1.1. The SR PCE Capability sub-TLV . . . . .                 | 5  |
| 4.1.2. PCECC-CAPABILITY sub-TLV . . . . .                      | 6  |
| 4.2. LSP Object . . . . .                                      | 6  |
| 4.2.1. Path Segment TLV . . . . .                              | 7  |
| 4.3. FEC Object . . . . .                                      | 8  |
| 4.4. CCI Object . . . . .                                      | 9  |
| 5. Operations . . . . .                                        | 10 |
| 5.1. Stateful PCE Operation . . . . .                          | 10 |
| 5.1.1. Ingress PCC-Initiated Path Segment Allocation . . . . . | 10 |
| 5.1.2. PCE Initiated Path Segment Allocation . . . . .         | 12 |
| 5.2. PCECC Based Operation . . . . .                           | 13 |
| 5.2.1. PCE Controlled Label Spaces Advertisement . . . . .     | 13 |
| 5.2.2. PCECC based Path Segment Allocation . . . . .           | 13 |
| 6. Dataplane Considerations . . . . .                          | 15 |
| 7. Implementation Status . . . . .                             | 16 |
| 7.1. Huawei's Commercial Delivery . . . . .                    | 16 |
| 7.2. ZTE's Commercial Delivery . . . . .                       | 17 |
| 8. IANA Considerations . . . . .                               | 17 |

|             |                                             |    |
|-------------|---------------------------------------------|----|
| 8.1.        | SR PCE Capability Flags . . . . .           | 17 |
| 8.2.        | New LSP Flag Registry . . . . .             | 17 |
| 8.3.        | New PCEP TLV . . . . .                      | 17 |
| 8.3.1.      | Path Segment TLV . . . . .                  | 18 |
| 8.4.        | New FEC Type Registry . . . . .             | 18 |
| 8.5.        | PCEP Error Type and Value . . . . .         | 19 |
| 9.          | Security Considerations . . . . .           | 19 |
| 10.         | Manageability Considerations . . . . .      | 19 |
| 10.1.       | Control of Function and Policy . . . . .    | 19 |
| 10.2.       | Information and Data Models . . . . .       | 20 |
| 10.3.       | Liveness Detection and Monitoring . . . . . | 20 |
| 10.4.       | Verify Correct Operations . . . . .         | 20 |
| 10.5.       | Requirements On Other Protocols . . . . .   | 20 |
| 10.6.       | Impact On Network Operations . . . . .      | 20 |
| 11.         | Acknowledgments . . . . .                   | 20 |
| 12.         | References . . . . .                        | 20 |
| 12.1.       | Normative References . . . . .              | 20 |
| 12.2.       | Informative References . . . . .            | 22 |
| Appendix A. | Contributors . . . . .                      | 23 |
| Appendix B. | SRv6 extensions . . . . .                   | 23 |
| B.1.        | The SRv6 PCE Capability sub-TLV . . . . .   | 24 |
| B.2.        | SRv6 PCE Capability Flags . . . . .         | 24 |
| B.3.        | Path Segment TLV . . . . .                  | 25 |
| Authors'    | Addresses . . . . .                         | 25 |

## 1. Introduction

[RFC5440] describes the Path Computation Element (PCE) Communication Protocol (PCEP). PCEP enables the communication between a Path Computation Client (PCC) and a PCE, or between PCE and PCE, for the purpose of computation of Multiprotocol Label Switching (MPLS) as well as Generalized MPLS (GMPLS) Traffic Engineering Label Switched Path (TE LSP) characteristics.

[RFC8231] specifies a set of extensions to PCEP to enable stateful control of TE LSPs within and across PCEP sessions in compliance with [RFC4657]. It includes mechanisms to effect LSP State Synchronization between PCCs and PCEs, delegation of control over LSPs to PCEs, and PCE control of timing and sequence of path computations within and across PCEP sessions. The model of operation where LSPs are initiated from the PCE is described in [RFC8281].

[I-D.ietf-pce-pcep-extension-for-pce-controller] specify the procedures and PCEP protocol extensions for using the PCE as the central controller for static LSPs, where LSPs can be provisioned as explicit label instructions at each hop on the end-to-end path.

Segment routing (SR) [RFC8402] leverages the source routing and tunneling paradigms and supports steering packets into an explicit forwarding path at the ingress node.

An SR path needs to be identified in some use cases such as performance measurement. For identifying an SR path, [I-D.ietf-spring-mpls-path-segment] introduces a new segment that is referred to as Path Segment.

[I-D.ietf-pce-segment-routing] specifies extensions to the Path Computation Element Protocol (PCEP) [RFC5440] for SR networks, that allow a stateful PCE to compute and initiate SR-TE paths, as well as a PCC to request, report or delegate SR paths.

[I-D.zhao-pce-pcep-extension-pce-controller-sr] specifies the procedures and PCEP protocol extensions when a PCE-based controller is also responsible for configuring the forwarding actions on the routers (SR SID distribution in this case), in addition to computing the paths for packet flows in a segment routing network and telling the edge routers what instructions to attach to packets as they enter the network.

This document specifies a mechanism to carry the SR path identification information in PCEP messages [RFC5440] [RFC8231] [RFC8281]. The SR path identifier can be a Path Segment in SR-MPLS [I-D.ietf-spring-mpls-path-segment], or other IDs that can identify an SR path. This document also extends the PCECC-SR mechanism to inform the Path Segment to the egress PCC.

## 2. Terminology

This memo makes use of the terms defined in [RFC4655], [I-D.ietf-pce-segment-routing], and [RFC8402].

### 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. Overview of Path Segment Extensions in PCEP

This document specifies a mechanism of allocating Path Segment and extends PCEP to encode it in PCEP messages. For supporting Path Segment in PCEP, several TLVs and flags are defined. The formats of

the objects and TLVs are described in Section 4. The procedures of Path Segment allocation are described in Section 5.

There are various modes of operations, such as -

- o The Path Segment can be allocated by Egress PCC. The PCE should request the Path Segment from Egress PCC.
- o The PCE can allocate a Path Segment on its own accord and inform the ingress/egress PCC, useful for PCE-initiated LSPs.
- o Ingress PCC can also request PCE to allocate the Path Segment, in this case, the PCE would either allocate and inform the assigned Path Segment to the ingress/egress PCC using PCEP messages, or first request egress PCC for Path Segment and then inform it to the ingress PCC.

The path information to the ingress PCC and PCE is exchanged via an extension to [I-D.ietf-pce-segment-routing] and [I-D.ietf-pce-segment-routing-ipv6]. The Path Segment information (for SR-MPLS) to the egress PCC can be informed via an extension to the PCECC-SR procedures [I-D.zhao-pce-pcep-extension-pce-controller-sr].

For the PCE to allocate a Path Segment on its own, the PCE needs to be aware of the MPLS label space from the PCCs. This is done via mechanism as described in [I-D.li-pce-controlled-id-space]. Otherwise, the PCE should request the egress PCC for Path Segment allocation.

#### 4. Objects and TLVs

##### 4.1. The OPEN Object

###### 4.1.1. The SR PCE Capability sub-TLV

[I-D.ietf-pce-segment-routing] defined a new Path Setup Type (PST) and SR-PCE-CAPABILITY sub-TLV for SR-MPLS. PCEP speakers use this sub-TLV to exchange information about their SR capability. The TLV defines a Flags field [I-D.ietf-pce-segment-routing].

This document adds an additional flag for Path Segment allocation, as follows -

- o P (Path Segment Identification bit): A PCEP speaker sets this flag to 1 to indicate that it has the capability to encode SR path identification (Path Segment, as per [I-D.ietf-spring-mpls-path-segment]).

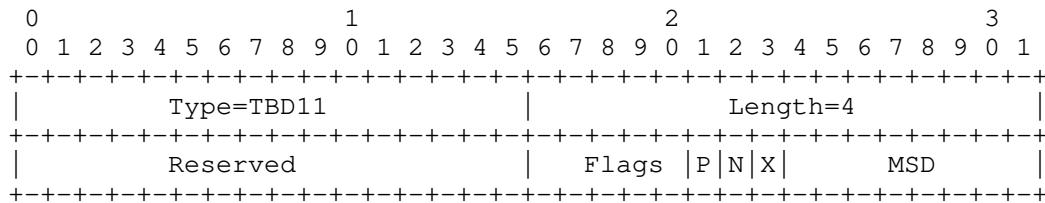


Figure 1: P-flag in SR-PCE-CAPABILITY TLV

The figure is included for the ease of the reader and will be removed at the time of publication.

#### 4.1.2. PCECC-CAPABILITY sub-TLV

Along with the SR sub-TLVs, the PCECC Capability as per [I-D.zhao-pce-pcep-extension-pce-controller-sr] should be advertised if the PCE allocates the Path Segment and acts as a Central Controller that manages the Label space.

The PCECC Capability should be advertised on the egress PCEP session, along with the SR sub-TLVs. This is needed to ensure that the PCE can use the PCECC objects/mechanism to request/inform the egress PCC of the Path Segment as described in Section 5.2.

#### 4.2. LSP Object

The LSP Object is defined in Section 7.3 of [RFC8231]. This document adds a flag in the LSP Object:

- o P (PCE Allocation bit): If the bit is set to 1, it indicates that the PCC requests PCE to make allocations for this LSP. The TLV in LSP object identifies what should be allocated, such as Path Segment or Binding Segment. A PCC would set this bit to 1 and include a PATH-SEGMENT TLV in the LSP object to request for allocation of Path Segment by the PCE in the PCEP message. A PCE would also set this bit to 1 and include a PATH-SEGMENT TLV to indicate that the Path Segment is allocated by PCE and encoded in the PCEP message towards PCC. Further, a PCE would set this bit to 0 and include a PATH-SEGMENT TLV in the LSP object to indicate that the Path Segment should be allocated by the PCC as described in Section 5.1.1.

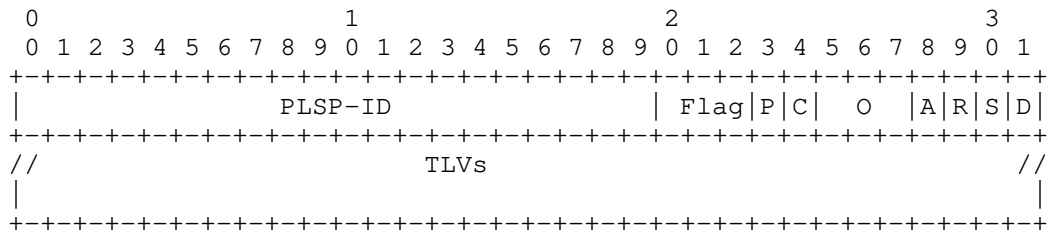


Figure 2: P-flag in LSP Object

The figure is included for the ease of the reader and will be removed at the time of publication.

#### 4.2.1. Path Segment TLV

The PATH-SEGMENT TLV is an optional TLV for use in the LSP Object for Path Segment allocation. The type of this TLV is to be allocated by IANA (TBA4). The format is as shown below.

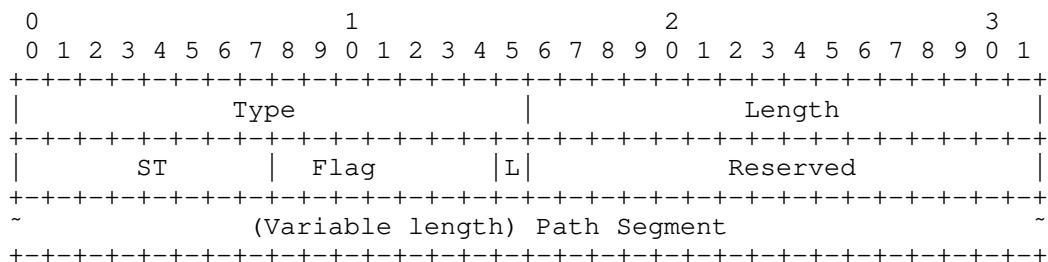


Figure 3: The PATH-SEGMENT TLV Format

The type (16-bit) of the TLV is TBA4 (to be allocated by IANA). The length (16-bit) has a variable length. The value contains the following fields:

- o ST (The Segment type - 8 bits): The ST field specifies the type of the Path Segment field, which carries a Path Segment corresponding to the SR path.
  - \* 0: MPLS Path Segment, which is an MPLS label as defined in [I-D.ietf-spring-mpls-path-segment]. The PST type MUST be set to SR (MPLS).
  - \* 1-255: Reserved for future use.
- o Flags (8 bits): One flag is currently defined:

- \* L-Bit (Local/Global - 1 bit): If set, then the Path Segment carried by the PATH-SEGMENT TLV has local significance. If not set, then the Path Segment carried by this TLV has global significance (i.e. Path Segment is global within an SR domain).
- \* The unassigned bits MUST be set to 0 and MUST be ignored at receipt.
- o Reserved (16 bits): MUST be set to 0 and MUST be ignored at receipt.
- o Path Segment: The Path Segment of an SR path. The Path Segment type is indicated by the ST field. When the ST is 0, it is a MPLS Path Segment [I-D.ietf-spring-mpls-path-segment] in the MPLS label format.

In general, only one instance of PATH-SEGMENT TLV will be included in LSP object. If more than one PATH-SEGMENT TLV is included, the first one is processed and others MUST be ignored. Multiple Path Segment allocation for use cases like alternate-making will be considered in future version of this draft.

When the Path Segment allocation is enabled, a PATH-SEGMENT TLV MUST be included in the LSP object.

If the label space is maintained by PCC itself, and the Path Segment is allocated by Egress PCC, then the PCE should request the Path Segment from Egress PCC as described in Section 5.1.1. In this case, the PCE should send a PCUpdate or PCInitiate message to the egress PCC to request the Path Segment. The P-flag in LSP should be unset in this case.

If a PCEP node does not recognize the PATH-SEGMENT TLV, it would behave in accordance with [RFC5440] and ignore the TLV. If a PCEP node recognizes the TLV but does not support the TLV, it MUST send PCErr with Error-Type = 2 (Capability not supported).

#### 4.3. FEC Object

The FEC Object [I-D.zhao-pce-pcep-extension-pce-controller-sr] is used to specify the FEC information and carried within PCInitiate or PCRpt message for the PCECC-SR operations. The PCE MUST inform the Path Identification information to the Egress PCC. To do this, this document extends the procedures of [I-D.zhao-pce-pcep-extension-pce-controller-sr] by defining a new FEC object type for Path.

FEC Object-Type is TBA6 'Path'.

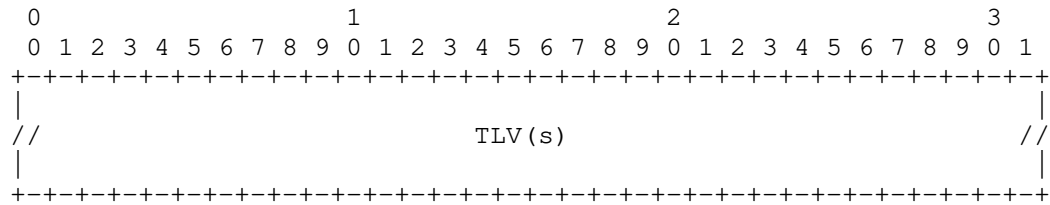


Figure 4: The path FEC object Format

One or more following TLV(s) are allowed in the 'path' FEC object -

- o SYMBOLIC-PATH-NAME TLV: As defined in [RFC8231], it is a human-readable string that identifies an LSP in the network.
- o LSP-IDENTIFIERS TLVs: As defined in [RFC8231], it is optional for SR, but could be used to encode the source, destination and other identification information for the path.
- o SPEAKER-ENTITY-ID TLV: As defined in [RFC8232], a unique identifier for the PCEP speaker, it is used to identify the Ingress PCC.

Either SYMBOLIC-PATH-NAME TLV or LSP-IDENTIFIERS TLV MUST be included. SPEAKER-ENTITY-ID TLV is optional. Only one instance of each TLV is processed, if more than one TLV of each type is included, the first one is processed and others MUST be ignored.

#### 4.4. CCI Object

The Central Control Instructions (CCI) Object is used by the PCE to specify the forwarding instructions is defined in [I-D.ietf-pce-pcep-extension-for-pce-controller]. Further [I-D.zhao-pce-pcep-extension-pce-controller-sr] defined a CCI object type for SR.

The Path Segment information is encoded directly in the CCI SR object. The Path Segment TLV as described in the Section 4.2.1, MUST also be included in the CCI SR object as the TLV (as it includes additional information regarding the Path Segment identifier). The C flag in CCI object is used to indicate if the allocation needs to be done by the PCC.

## 5. Operations

The Path Segment allocation and encoding is as per the Stateful PCE operations for segment routing. The procedures are as per the corresponding extensions defined in [I-D.ietf-pce-segment-routing] and [I-D.ietf-pce-segment-routing-ipv6] (which are further based on [RFC8231] and [RFC8281]). The additional operations for Path Segment are defined in this section.

To notify (or request) the Path Segment to the Egress PCC, the procedures are as per the PCECC-SR [I-D.zhao-pce-pcep-extension-pce-controller-sr] (which is based on [I-D.ietf-pce-pcep-extension-for-pce-controller]). The additional operations are defined in this section.

### 5.1. Stateful PCE Operation

As defined in [I-D.ietf-spring-mpls-path-segment], a Path Segment can be allocated by the egress PCC. In this case, the label space is maintained on the PCC itself.

This section describes the mechanism of Path Segment allocation by using PCInitiate and PCUpd message in Stateful PCE model.

#### 5.1.1. Ingress PCC-Initiated Path Segment Allocation

The ingress PCC could request the Path Segment to be allocated by the PCE via PCRpt message. The delegate flag (D-flag) MUST also be set for this LSP. Also, the P-flag in the LSP object MUST be set.

On receiving a delegation request with Path Segment allocation request from an ingress PCC, a stateful PCE requests the egress PCC to allocate a Path Segment.

The PATH-SEGMENT TLV MUST be included in an LSP object in the PCInitiate message sent from the PCE to the egress to request Path Segment allocation by the egress PCC. The P flag in LSP object MUST be set to 0. This PCInitiate message to egress PCC would be the similar to the one sent to ingress PCC as per [I-D.ietf-pce-segment-routing], but the egress PCC would only allocate the Path Segment and would not trigger the LSP initiation operation (as it would be the egress for this LSP).

If the value of Path Segment is 0x0, it indicates that the PCE is requesting a Path Segment for this LSP. If the Path Segment is set to a value 'n' and the P flag is unset in the LSP object, it indicates that the PCE requests a specific value 'n' of Path Segment. If the Path Segment is allocated successfully, the egress PCC reports

the Path Segment via PCRpt message with PATH-SEGMENT TLV in LSP object. Else, it MUST send a PCErr message with Error-Type = TBA7 ("Path SID failure") and Error Value = 1 ("Invalid SID"). If the value of Path Segment is valid, but the PCC is unable to allocate the Path Segment, it MUST send a PCErr message with Error-Type = TBA7 ("Path SID failure") and Error Value = 2 ("Unable to allocate the specified label/SID").

Once the PCE receives the PCRpt message, it can obtain the Path Segment information from the egress PCC and then update the path with Path Segment by sending PCUpd message to the ingress PCC.

If the Path Segment is updated successfully, the ingress PCC will acknowledge with a PCRpt message to the PCE. In case of error, an PCErr message with Error-Type = TBA7 ("Path SID failure") and Error Value = 1 ("Invalid SID") will be sent back to the PCE. The PCE MUST roll back the Path Segment value to the previous value (if any) by sending a PCUpd message to synchronize with the egress PCC.

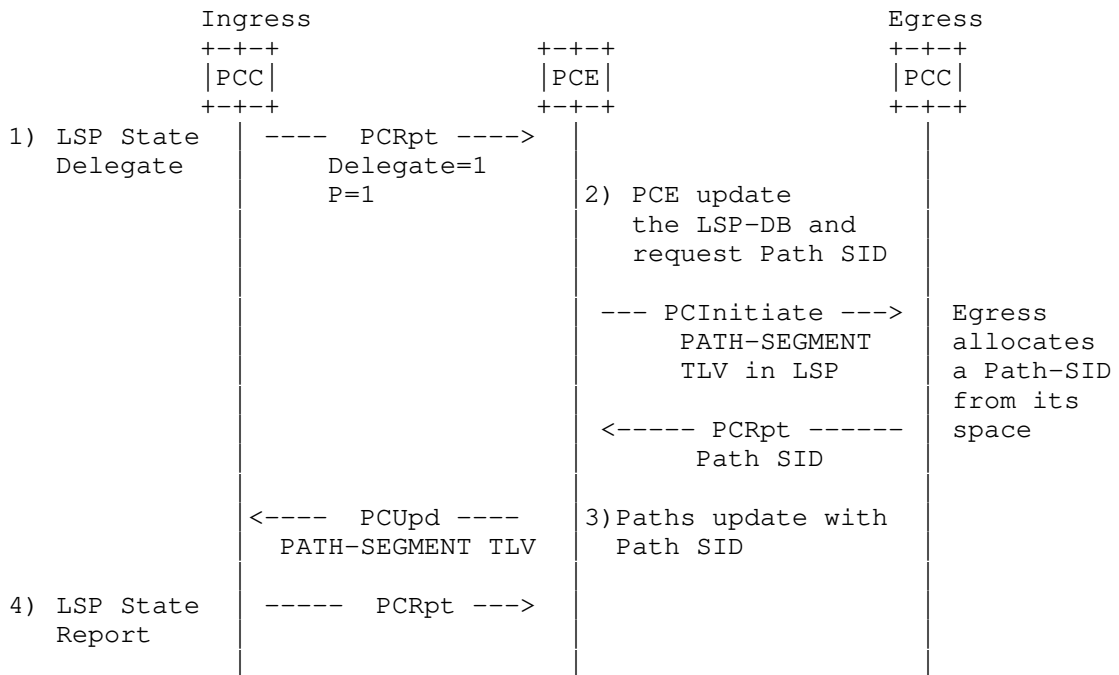


Figure 5: Ingress PCC-Initiated Path Segment Allocation

If the ingress PCC wishes to withdraw or modify a previously reported Path Segment value, it MUST send a PCRpt message without any PATH-

SEGMENT TLV or with the PATH-SEGMENT TLV containing the new Path Segment respectively. In this case, the PCE should synchronize with egress PCC via PCUpd message.

The Path Segment MUST be withdrawn when the corresponding LSP is removed. When the LSP is deleted, the PCE MUST request the egress PCC to withdraw the LSP and associated Path Segment via PCInitiate message with the R flag is set in the SRP object.

If an egress PCC receives a valid Path Segment value from a PCE which is different than the current Path Segment, it MUST try to allocate the new value. If the new Path Segment is successfully allocated, the egress PCC MUST report the new value to the PCE. Otherwise, it MUST send a PCErr message with Error-Type = TBA7 ("Path label/SID failure") and Error Value = 2 ("Unable to allocate the specified label/SID").

#### 5.1.1.2. PCE Initiated Path Segment Allocation

A stateful PCE also can initiate or update an LSP with Path Segment actively via requesting the egress PCC to allocate a Path Segment.

If a PCE wishes to modify a previously requested Path Segment value or allocate a Path Segment for an PCE-Initiated LSP, it MUST request the egress PCC to allocate a new value by sending a PCUpd message to the egress PCC with PATH-SEGMENT TLV containing the new Path Segment value. Also, the P flag in LSP object is unset. Absence of the PATH-SEGMENT TLV in PCUpd message means that the PCE wishes to withdraw the Path Segment.

The mechanism of requesting Path Segment is as per Section 5.1.1.

Once the PCE receives the PCRpt message, it can obtain the Path Segment information from the egress PCC and then update or initiate an LSP with Path Segment.

If the SR-Path is setup, the ingress PCC will acknowledge with a PCRpt message to the PCE. In case of error, as described in [I-D.ietf-pce-segment-routing], an PCErr message will be sent back to the PCE. The PCE MUST request the egress PCC to withdraw the LSP and associated Path Segment via PCInitiate message with the R flag is set in the SRP object.

If the Path Segment is updated successfully, the ingress PCC will acknowledge with a PCRpt message to the PCE. In case of error, an PCErr message with Error-Type = TBA7 ("Path SID failure") and Error Value = 1 ("Invalid SID") will be sent back to the PCE. The PCE MUST

roll back the Path Segment value to the previous value (if any) by sending a PCUpd message to synchronize with the egress PCC.

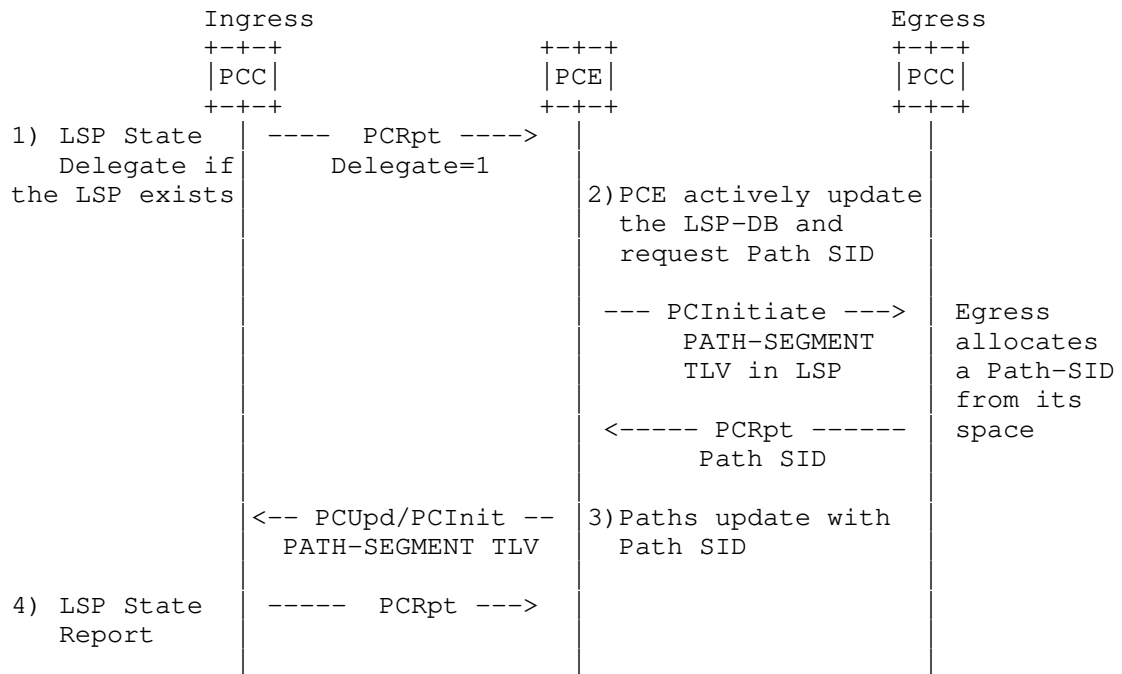


Figure 6: Stateful PCE-Initiated Path Segment Allocation

## 5.2. PCECC Based Operation

### 5.2.1. PCE Controlled Label Spaces Advertisement

For allocating the Path Segments to SR paths by the PCEs, the PCE controlled label space MUST be known at PCEs via configurations or any other mechanisms. The PCE controlled label spaces MAY be advertised as described in [I-D.li-pce-controlled-id-space].

### 5.2.2. PCECC based Path Segment Allocation

#### 5.2.2.1. PCECC-Initiated

The PCE could allocate the Path Segment on its own for a PCE-Initiated (or delegated LSP). The allocated Path Segment needs to be informed to the Ingress and Egress PCC. The PCE would use the PCInitiate message [RFC8281] or PCUpd message [RFC8231] towards the Ingress PCC and MUST include the PATH-SEGMENT TLV in the LSP object.

The PCE would further inform the egress PCC about the Path Segment allocated by the PCE using the PCInitiate message as described in [I-D.zhao-pce-pcep-extension-pce-controller-sr].

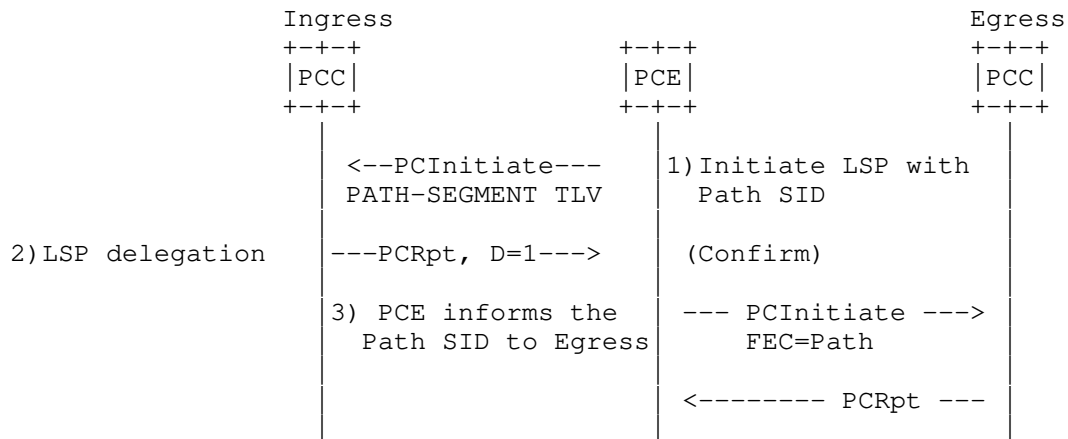


Figure 7: PCE allocated Path Segment on its own

#### 5.2.2.2. Ingress PCC-Initiated PCECC

The ingress PCC could request the Path Segment to be allocated by the PCE via PCRpt message as per [RFC8231]. The delegate flag (D-flag) MUST also be set for this LSP. Also, the P-flag in the LSP object MUST be set.

A PATH-SEGMENT TLV MUST be included in the LSP object. If the value of Path Segment is 0x0, it indicates that the Ingress PCC is requesting a Path Segment for this LSP. If the Path Segment is set to a value 'n', it indicates that the ingress PCC requests a specific value 'n' of Path Segment.

If the Path Segment is allocated successfully, the PCE would further respond to Ingress PCC with PCUpd message as per [RFC8231] and MUST include the PATH-SEGMENT TLV in a LSP object. Else, it MUST send a PCErr message with Error-Type = TBA7 ("Path SID failure") and Error Value = 1 ("Invalid SID"). If the value of Path Segment is valid, but the PCC is unable to allocate the Path Segment, it MUST send a PCErr message with Error-Type = TBA7 ("Path SID failure") and Error Value = 2 ("Unable to allocate the specified label/SID").

The active PCE would allocate the Path Segment as per the PATH-SEGMENT flags and in case PATH-SEGMENT is not included, the PCE MUST act based on the local policy.

The PCE would further inform the egress PCC about the Path Segment allocated by the PCE using the PCInitiate message as described in [I-D.zhao-pce-pcep-extension-pce-controller-sr].

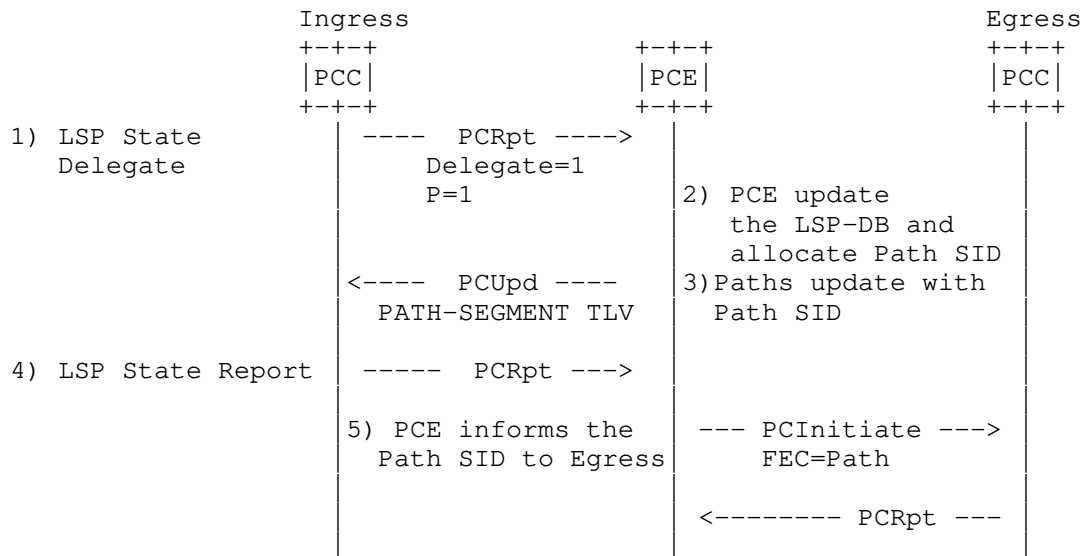


Figure 8: Ingress PCC request Path Segment to PCE

## 6. Dataplane Considerations

As described in [I-D.ietf-spring-mpls-path-segment], in an SR-MPLS network, when a packet is transmitted along an SR path, the labels in the MPLS label stack will be swapped or popped. So that no label or only the last label may be left in the MPLS label stack when the packet reaches the egress node. Thus, the egress node cannot determine from which SR path the packet comes. For this reason, it introduces the Path Segment.

Apart from allocation and encoding of the Path Segment (described in this document) for the LSP, it would also be included in the SID/Label stack of the LSP (usually for processing by the egress). To support this, the Path Segment MAY also be a part of SR-ERO as prepared by the PCE as per [I-D.ietf-pce-segment-routing]. The PCC MAY also include the Path Segment while preparing the label stack based on the local policy and use-case.

It is important that the PCE learns the Maximum SID Depth (MSD) that can be imposed at each node/link of a given SR path to ensure that the SID stack depth does not exceed the number of SIDs the node is

capable of imposing. As a new type of segment, Path Segment will be inserted in the SID list just like other SIDs. Thus, the PCE needs to consider the affect of Path Segment when computing a LSP with Path Segment allocation.

## 7. Implementation Status

[Note to the RFC Editor - remove this section before publication, as well as remove the reference to [RFC7942].

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in [RFC7942]. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to [RFC7942], "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature. It is up to the individual working groups to use this information as they see fit".

### 7.1. Huawei's Commercial Delivery

The feature is developing based on Huawei VRP8.

- o Organization: Huawei
- o Implementation: Huawei's Commercial Delivery implementation based on VRP8.
- o Description: The implementation is under development and follows the mechanism as defined in section-5.1.1.
- o Maturity Level: Product
- o Contact: tanren@huawei.com

## 7.2. ZTE's Commercial Delivery

- o Organization: ZTE
- o Implementation: ZTE's Commercial Delivery implementation based on Rosng v8.
- o Description: The implementation is under development and follows the mechanism as defined in section-5.1.1.
- o Maturity Level: Product
- o Contact: zhan.shuangping@zte.com.cn

## 8. IANA Considerations

### 8.1. SR PCE Capability Flags

SR PCE Capability TLV is defined in [I-D.ietf-pce-segment-routing], and the registry to manage the Flag field of the SR PCE Capability TLV is requested in [I-D.ietf-pce-segment-routing]. IANA is requested to make the following allocation in the "SR Capability Flag Field" sub-registry.

| Bit  | Description                             | Reference     |
|------|-----------------------------------------|---------------|
| TBA1 | Path Segment Allocation is supported(P) | This document |

### 8.2. New LSP Flag Registry

[RFC8231] defines the LSP object; per that RFC, IANA created a registry to manage the value of the LSP object's Flag field. IANA has allocated a new bit in the "LSP Object Flag Field" sub-registry, as follows:

| Bit  | Description                            | Reference     |
|------|----------------------------------------|---------------|
| TBA3 | Request for Path Segment Allocation(P) | This document |

### 8.3. New PCEP TLV

IANA is requested to add the assignment of a new allocation in the existing "PCEP TLV Type Indicators" sub-registry as follows:

| Value | Description      | Reference     |
|-------|------------------|---------------|
| TBA4  | PATH-SEGMENT TLV | This document |

#### 8.3.1. Path Segment TLV

This document requests that a new sub-registry named "PATH-SEGMENT TLV Segment Type (ST) Field" to be created to manage the value of the ST field in the PATH-SEGMENT TLV.

| Value | Description                    | Reference     |
|-------|--------------------------------|---------------|
| 0     | MPLS Path Segment (MPLS label) | This document |
| 1-255 | Reserved for future use        | This document |

Further, this document also requests that a new sub-registry named "PATH-SEGMENT TLV Flag Field" to be created to manage the Flag field in the PATH-SEGMENT TLV. New values are assigned by Standards Action [RFC8126]. Each bit should be tracked with the following qualities:

- o Bit number (counting from bit 0 as the most significant bit)
- o Capability description
- o Defining RFC

| Bit | Description            | Reference     |
|-----|------------------------|---------------|
| 7   | Local Signification(L) | This document |

#### 8.4. New FEC Type Registry

A new PCEP object called FEC is defined in [I-D.zhao-pce-pcep-extension-pce-controller-sr]. IANA is requested to allocate a new Object-Type for FEC object in the "PCEP Objects" sub-registry.

| Value | Description | Reference     |
|-------|-------------|---------------|
| TBA6  | Path        | This document |

### 8.5. PCEP Error Type and Value

IANA is requested to allocate code-points in the "PCEP-ERROR Object Error Types and Values" sub-registry for the following new error-types and error-values:

| Error-Type | Meaning                                                                                                      | Reference     |
|------------|--------------------------------------------------------------------------------------------------------------|---------------|
| TBA7       | Path SID failure:<br>Error-value = 1<br>Invalid SID<br><br>Error-value = 2<br>Unable to allocate<br>Path SID | This document |

### 9. Security Considerations

The security considerations described in [RFC5440], [RFC8231], [RFC8281] and [I-D.ietf-pce-segment-routing] are applicable to this specification. No additional security measure is required.

As described [I-D.ietf-pce-segment-routing] and [I-D.ietf-pce-pcep-extension-for-pce-controller], SR allows a network controller to instantiate and control paths in the network. A rogue PCE can manipulate Path SID allocations to have impact based on the usage of Path SID such as accounting, bi-directional etc.

Thus, as per [RFC8231], it is RECOMMENDED that these PCEP extensions only be activated on authenticated and encrypted sessions across PCEs and PCCs belonging to the same administrative authority, using Transport Layer Security (TLS) [RFC8253], as per the recommendations and best current practices in [RFC7525] (unless explicitly set aside in [RFC8253]).

### 10. Manageability Considerations

All manageability requirements and considerations listed in [RFC5440], [RFC8231], and [I-D.ietf-pce-segment-routing] apply to PCEP protocol extensions defined in this document. In addition, requirements and considerations listed in this section apply.

#### 10.1. Control of Function and Policy

A PCEP implementation SHOULD allow the operator to configure the policy based on which it allocates the Path SID. This includes the Path SID scope.

## 10.2. Information and Data Models

The PCEP YANG module is defined in [I-D.ietf-pce-pcep-yang]. In future, this YANG module should be extended or augmented to provide the following additional information relating to Path SID.

An implementation SHOULD allow the operator to view the Path SID allocated to the LSP as well as Path SID as part of the computed SID list for the SR path.

## 10.3. 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].

## 10.4. Verify Correct Operations

Mechanisms defined in this document do not imply any new operation verification requirements in addition to those already listed in [RFC5440], [RFC8231], and [I-D.ietf-pce-segment-routing] .

## 10.5. Requirements On Other Protocols

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

## 10.6. Impact On Network Operations

Mechanisms defined in [RFC5440], [RFC8231], and [I-D.ietf-pce-segment-routing] also apply to PCEP extensions defined in this document. Further, the mechanism described in this document can help the operator to request control of the LSPs at a particular PCE.

## 11. Acknowledgments

TBA

## 12. References

### 12.1. 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>>.

- [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>>.
- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <<https://www.rfc-editor.org/info/rfc8126>>.
- [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>>.
- [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>>.
- [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>>.
- [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>>.

- [I-D.ietf-pce-segment-routing]  
Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W.,  
and J. Hardwick, "PCEP Extensions for Segment Routing",  
draft-ietf-pce-segment-routing-16 (work in progress),  
March 2019.
- [I-D.ietf-pce-segment-routing-ipv6]  
Negi, M., Li, C., Sivabalan, S., Kaladharan, P., and Y.  
Zhu, "PCEP Extensions for Segment Routing leveraging the  
IPv6 data plane", draft-ietf-pce-segment-routing-ipv6-02  
(work in progress), April 2019.
- [I-D.zhao-pce-pcep-extension-pce-controller-sr]  
Zhao, Q., Li, Z., Negi, M., and C. Zhou, "PCEP Procedures  
and Protocol Extensions for Using PCE as a Central  
Controller (PCECC) of SR-LSPs", draft-zhao-pce-pcep-  
extension-pce-controller-sr-05 (work in progress), July  
2019.
- [I-D.ietf-pce-pcep-extension-for-pce-controller]  
Zhao, Q., Li, Z., Negi, M., and C. Zhou, "PCEP Procedures  
and Protocol Extensions for Using PCE as a Central  
Controller (PCECC) of LSPs", draft-ietf-pce-pcep-  
extension-for-pce-controller-02 (work in progress), July  
2019.
- [I-D.li-spring-srv6-path-segment]  
Li, C., Cheng, W., Chen, M., Dhody, D., Li, Z., Dong, J.,  
and R. Gandhi, "Path Segment for SRv6 (Segment Routing in  
IPv6)", draft-li-spring-srv6-path-segment-03 (work in  
progress), August 2019.

## 12.2. 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,  
<<https://www.rfc-editor.org/info/rfc4655>>.
- [RFC4657] Ash, J., Ed. and J. Le Roux, Ed., "Path Computation  
Element (PCE) Communication Protocol Generic  
Requirements", RFC 4657, DOI 10.17487/RFC4657, September  
2006, <<https://www.rfc-editor.org/info/rfc4657>>.
- [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L.,  
Decraene, B., Litkowski, S., and R. Shakir, "Segment  
Routing Architecture", RFC 8402, DOI 10.17487/RFC8402,  
July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.

[RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", BCP 205, RFC 7942, DOI 10.17487/RFC7942, July 2016, <<https://www.rfc-editor.org/info/rfc7942>>.

[I-D.li-pce-controlled-id-space]

Li, C., Chen, M., Dong, J., Li, Z., Wang, A., Cheng, W., and C. Zhou, "PCE Controlled ID Space", draft-li-pce-controlled-id-space-03 (work in progress), June 2019.

[I-D.ietf-spring-mpls-path-segment]

Cheng, W., Li, H., Chen, M., Gandhi, R., and R. Zigler, "Path Segment in MPLS Based Segment Routing Network", draft-ietf-spring-mpls-path-segment-00 (work in progress), March 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.

#### Appendix A. Contributors

The following people have substantially contributed to this document:

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

Email: [dhruv.ietf@gmail.com](mailto:dhruv.ietf@gmail.com)

Zafar Ali  
Cisco Systems, Inc.

Email: [zali@cisco.com](mailto:zali@cisco.com)

#### Appendix B. SRv6 extensions

This section would be rolled into the document once the SPRING WG adopts SRv6 path segment.

### B.1. The SRv6 PCE Capability sub-TLV

[I-D.ietf-pce-segment-routing-ipv6] defined a new Path Setup Type (PST) and SRv6-PCE-CAPABILITY sub-TLV for SRv6. PCEP speakers use this sub-TLV to exchange information about their SRv6 capability. The TLV includes a Flags field and one bit (L-flag) was allocated in [I-D.ietf-pce-segment-routing-ipv6].

This document adds an additional flag for Path Segment allocation, as follows -

- o P (Path Segment Identification bit): A PCEP speaker sets this flag to 1 to indicate that it has the capability to encode SRv6 path identification. (Path Segment, as per [I-D.li-spring-srv6-path-segment]).

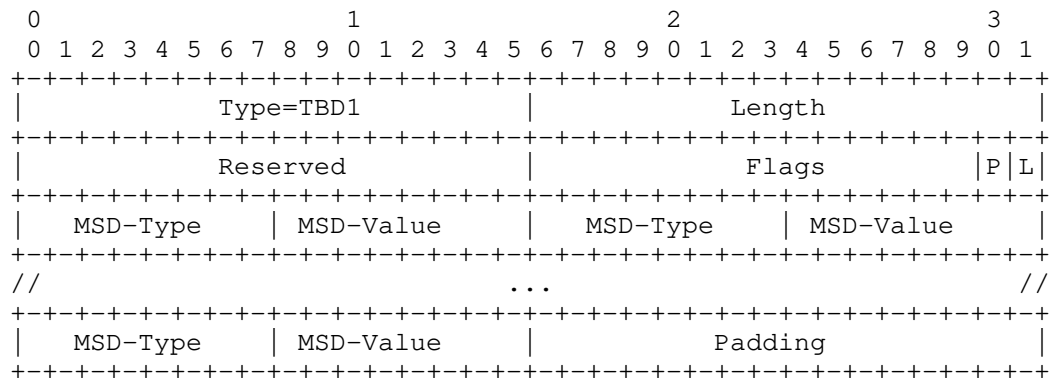


Figure 9: P-flag in SRv6-PCE-CAPABILITY TLV

The figure is included for the ease of the reader and can be removed at the time of publication.

### B.2. SRv6 PCE Capability Flags

SRv6 PCE Capability TLV is defined in [I-D.ietf-pce-segment-routing-ipv6], and the registry to manage the Flag field of the SRv6 PCE Capability Flags is requested in [I-D.ietf-pce-segment-routing-ipv6]. IANA is requested to make the following allocation in the aforementioned registry.

| Bit  | Description                             | Reference     |
|------|-----------------------------------------|---------------|
| TBA2 | Path Segment Allocation is supported(P) | This document |

## B.3. Path Segment TLV

A new assignment should be done to the "PATH-SEGMENT TLV Segment Type (ST) Field" sub-registry for SRv6.

| Value | Description                   | Reference     |
|-------|-------------------------------|---------------|
| 1     | SRv6 Path Segment (IPv6 addr) | This document |
| 2-255 | Reserved for future use       | This document |

## Authors' Addresses

Cheng Li  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: chengli13@huawei.com

Mach(Guoyi) Chen  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: Mach.chen@huawei.com

Weiqiang Cheng  
China Mobile  
China

Email: chengweiqiang@chinamobile.com

Jie Dong  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: jie.dong@huawei.com

Zhenbin Li  
Huawei Technologies  
Huawei Campus, No. 156 Beiqing Rd.  
Beijing 100095  
China

Email: lizhenbin@huawei.com

Rakesh Gandhi  
Cisco Systems, Inc.  
Canada

Email: rgandhi@cisco.com

Quan Xiong  
ZTE Corporation  
China

Email: xiong.quan@zte.com.cn

PCE Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: December 20, 2018

Q. Zhao  
Z. Li  
D. Dhody  
S. Karunanithi  
Huawei Technologies  
A. Farrel  
Juniper Networks, Inc  
C. Zhou  
Cisco Systems  
June 18, 2018

PCEP Procedures and Protocol Extensions for Using PCE as a Central  
Controller (PCECC) of LSPs  
draft-zhao-pce-pcep-extension-for-pce-controller-08

Abstract

The Path Computation Element (PCE) is a core component of Software-Defined Networking (SDN) systems. It can compute optimal paths for traffic across a network and can also update the paths to reflect changes in the network or traffic demands.

PCE was developed to derive paths for MPLS Label Switched Paths (LSPs), which are supplied to the head end of the LSP using the Path Computation Element Communication Protocol (PCEP). But SDN has a broader applicability than signaled (G)MPLS traffic-engineered (TE) networks, and the PCE may be used to determine paths in a range of use cases. PCEP has been proposed as a control protocol for use in these environments to allow the PCE to be fully enabled as a central controller.

A PCE-based central controller (PCECC) can simplify the processing of a distributed control plane by blending it with elements of SDN and without necessarily completely replacing it. Thus, the LSP can be calculated/setup/initiated and the label forwarding entries can also be downloaded through a centralized PCE server to each network devices along the path while leveraging the existing PCE technologies as much as possible.

This document specifies the procedures and PCEP protocol extensions for using the PCE as the central controller.

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 20, 2018.

#### Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 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

|                                                                                |    |
|--------------------------------------------------------------------------------|----|
| 1. Introduction . . . . .                                                      | 3  |
| 1.1. Requirements Language . . . . .                                           | 5  |
| 2. Terminology . . . . .                                                       | 5  |
| 3. Basic PCECC Mode . . . . .                                                  | 5  |
| 4. PCEP Requirements . . . . .                                                 | 5  |
| 5. Procedures for Using the PCE as the Central Controller<br>(PCECC) . . . . . | 6  |
| 5.1. Stateful PCE Model . . . . .                                              | 6  |
| 5.2. New LSP Functions . . . . .                                               | 6  |
| 5.3. PCECC Capability Advertisement . . . . .                                  | 7  |
| 5.4. LSP Operations . . . . .                                                  | 8  |
| 5.4.1. Basic PCECC LSP Setup . . . . .                                         | 8  |
| 5.4.2. Central Control Instructions . . . . .                                  | 10 |
| 5.4.2.1. Label Download . . . . .                                              | 10 |
| 5.4.2.2. Label Cleanup . . . . .                                               | 11 |
| 5.4.3. PCE Initiated PCECC LSP . . . . .                                       | 12 |
| 5.4.4. PCECC LSP Update . . . . .                                              | 14 |
| 5.4.5. Re Delegation and Cleanup . . . . .                                     | 16 |
| 5.4.6. Synchronization of Central Controllers Instructions .                   | 16 |

|                                                  |    |
|--------------------------------------------------|----|
| 5.4.7. PCECC LSP State Report . . . . .          | 16 |
| 6. PCEP messages . . . . .                       | 16 |
| 6.1. The PCInitiate message . . . . .            | 17 |
| 6.2. The PCRpt message . . . . .                 | 18 |
| 7. PCEP Objects . . . . .                        | 19 |
| 7.1. OPEN Object . . . . .                       | 19 |
| 7.1.1. PCECC Capability sub-TLV . . . . .        | 19 |
| 7.2. PATH-SETUP-TYPE TLV . . . . .               | 20 |
| 7.3. CCI Object . . . . .                        | 20 |
| 7.3.1. Address TLVs . . . . .                    | 21 |
| 8. Security Considerations . . . . .             | 23 |
| 8.1. Malicious PCE . . . . .                     | 23 |
| 9. Manageability Considerations . . . . .        | 23 |
| 9.1. Control of Function and Policy . . . . .    | 23 |
| 9.2. Information and Data Models . . . . .       | 23 |
| 9.3. Liveness Detection and Monitoring . . . . . | 23 |
| 9.4. Verify Correct Operations . . . . .         | 23 |
| 9.5. Requirements On Other Protocols . . . . .   | 23 |
| 9.6. Impact On Network Operations . . . . .      | 24 |
| 10. IANA Considerations . . . . .                | 24 |
| 10.1. PCEP TLV Type Indicators . . . . .         | 24 |
| 10.2. New Path Setup Type Registry . . . . .     | 24 |
| 10.3. PCEP Object . . . . .                      | 24 |
| 10.4. CCI Object Flag Field . . . . .            | 24 |
| 10.5. PCEP-Error Object . . . . .                | 25 |
| 11. Acknowledgments . . . . .                    | 25 |
| 12. References . . . . .                         | 25 |
| 12.1. Normative References . . . . .             | 25 |
| 12.2. Informative References . . . . .           | 26 |
| Appendix A. Contributor Addresses . . . . .      | 29 |
| Authors' Addresses . . . . .                     | 29 |

## 1. Introduction

The Path Computation Element (PCE) [RFC4655] was developed to offload path computation function from routers in an MPLS traffic-engineered network. Since then, the role and function of the PCE has grown to cover a number of other uses (such as GMPLS [RFC7025]) and to allow delegated control [RFC8231] and PCE-initiated use of network resources [RFC8281].

According to [RFC7399], Software-Defined Networking (SDN) refers to a separation between the control elements and the forwarding components so that software running in a centralized system, called a controller, can act to program the devices in the network to behave in specific ways. A required element in an SDN architecture is a component that plans how the network resources will be used and how the devices will be programmed. It is possible to view this

component as performing specific computations to place traffic flows within the network given knowledge of the availability of network resources, how other forwarding devices are programmed, and the way that other flows are routed. This is the function and purpose of a PCE, and the way that a PCE integrates into a wider network control system (including an SDN system) is presented in [RFC7491].

In early PCE implementations, where the PCE was used to derive paths for MPLS Label Switched Paths (LSPs), paths were requested by network elements (known as Path Computation Clients (PCCs)), and the results of the path computations were supplied to network elements using the Path Computation Element Communication Protocol (PCEP) [RFC5440]. This protocol was later extended to allow a PCE to send unsolicited requests to the network for LSP establishment [RFC8281].

[RFC8283] introduces the architecture for PCE as a central controller as an extension of the architecture described in [RFC4655] and assumes the continued use of PCEP as the protocol used between PCE and PCC. [RFC8283] further examines the motivations and applicability for PCEP as a Southbound Interface (SBI), and introduces the implications for the protocol. [I-D.ietf-teas-pcecc-use-cases] describes the use cases for the PCECC architecture.

A PCE-based central controller (PCECC) can simplify the processing of a distributed control plane by blending it with elements of SDN and without necessarily completely replacing it. Thus, the LSP can be calculated/setup/initiated and the label forwarding entries can also be downloaded through a centralized PCE server to each network devices along the path while leveraging the existing PCE technologies as much as possible.

This draft specify the procedures and PCEP protocol extensions for using the PCE as the central controller for static LSPs, where LSPs can be provisioned as explicit label instructions at each hop on the end-to-end path. Each router along the path must be told what label-forwarding instructions to program and what resources to reserve. The PCE-based controller keeps a view of the network and determines the paths of the end-to-end LSPs, and the controller uses PCEP to communicate with each router along the path of the end-to-end LSP.

The extension for PCECC in Segment Routing (SR) is specified in a separate draft [I-D.zhao-pce-pcep-extension-pce-controller-sr].

### 1.1. 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.

## 2. Terminology

Terminologies used in this document is same as described in the draft [RFC8283] and [I-D.ietf-teas-pcecc-use-cases].

## 3. Basic PCECC Mode

In this mode LSPs are provisioned as explicit label instructions at each hop on the end-to-end path. Each router along the path must be told what label forwarding instructions to program and what resources to reserve. The controller uses PCEP to communicate with each router along the path of the end-to-end LSP.

Note that the PCE-based controller will take responsibility for managing some part of the MPLS label space for each of the routers that it controls, and may take wider responsibility for partitioning the label space for each router and allocating different parts for different uses. This is also described in section 3.1.2. of [RFC8283]. For the purpose of this document, it is assumed that label range to be used by a PCE is known and set on both PCEP peers. A future extension could add this capability to advertise the range via possible PCEP extensions as well. The rest of processing is similar to the existing stateful PCE mechanism.

## 4. PCEP Requirements

Following key requirements associated PCECC should be considered when designing the PCECC based solution:

1. PCEP speaker supporting this draft MUST have the capability to advertise its PCECC capability to its peers.
2. PCEP speaker not supporting this draft MUST be able to reject PCECC related extensions with a error reason code that indicates that this feature is not supported.
3. PCEP speaker MUST provide a means to identify PCECC based LSP in the PCEP messages.

4. PCEP procedures SHOULD provide a means to update (or cleanup) the label- download entry to the PCC.
  5. PCEP procedures SHOULD provide a means to synchronize the labels between PCE to PCC in PCEP messages.
5. Procedures for Using the PCE as the Central Controller (PCECC)
- 5.1. Stateful PCE Model

Active stateful PCE is described in [RFC8231]. PCE as a central controller (PCECC) reuses existing Active stateful PCE mechanism as much as possible to control the LSP.

#### 5.2. New LSP Functions

This document defines the following new PCEP messages and extends the existing messages to support PCECC:

(PCRpt): a PCEP message described in [RFC8231]. PCRpt message is used to send PCECC LSP Reports. It is also extended to report the set of Central Controller's Instructions (CCI) (label forwarding instructions in the context of this document) received from the PCE. See Section 5.4.6 for more details.

(PCInitiate): a PCEP message described in [RFC8281]. PCInitiate message is used to setup PCE-Initiated LSP based on PCECC mechanism. It is also extended for Central Controller's Instructions (CCI) (download or cleanup the Label forwarding instructions in the context of this document) on all nodes along the path.

(PCUpd): a PCEP message described in [RFC8231]. PCUpd message is used to send PCECC LSP Update.

The new LSP functions defined in this document are mapped onto the messages as shown in the following table.

| Function                       | Message    |
|--------------------------------|------------|
| PCECC Capability advertisement | Open       |
| Label entry Add                | PCInitiate |
| Label entry Cleanup            | PCInitiate |
| PCECC Initiated LSP            | PCInitiate |
| PCECC LSP Update               | PCUpd      |
| PCECC LSP State Report         | PCRpt      |
| PCECC LSP Delegation           | PCRpt      |
| PCECC Label Report             | PCRpt      |

This document specifies a new object CCI (see Section 7.3) for the encoding of central controller's instructions. In the scope of this document this is limited to Label forwarding instructions. The CC-ID is the unique identifier for the central controller's instructions in PCEP. The PCEP messages are extended in this document to handle the PCECC operations.

### 5.3. PCECC Capability Advertisement

During PCEP Initialization Phase, PCEP Speakers (PCE or PCC) advertise their support of PCECC extensions.

This document defines a new Path Setup Type (PST) [I-D.ietf-pce-lsp-setup-type] for PCECC, as follows:

- o PST = TBD: Path is setup via PCECC mode.

A PCEP speaker MUST indicate its support of the function described in this document by sending a PATH-SETUP-TYPE-CAPABILITY TLV in the OPEN object with this new PST included in the PST list.

This document also defines the PCECC Capability sub-TLV Section 7.1.1. PCEP speakers use this sub-TLV to exchange information about their PCECC capability. If a PCEP speaker includes PST=TBD in the PST List of the PATH-SETUP-TYPE-CAPABILITY TLV then it MUST also include the PCECC Capability sub-TLV inside the PATH-SETUP-TYPE-CAPABILITY TLV.

The presence of the PST and PCECC Capability sub-TLV in PCC's OPEN Object indicates that the PCC is willing to function as a PCECC client.

The presence of the PST and PCECC Capability sub-TLV in PCE's OPEN message indicates that the PCE is interested in function as a PCECC server.

The PCEP protocol extensions for PCECC MUST NOT be used if one or both PCEP Speakers have not included the PST or the PCECC Capability sub-TLV in their respective OPEN message. If the PCEP Speakers support the extensions of this draft but did not advertise this capability then a PCerr message with Error-Type=19(Invalid Operation) and Error-Value=TBD (Attempted PCECC operations when PCECC capability was not advertised) will be generated and the PCEP session will be terminated.

A PCC or a PCE MUST include both PCECC-CAPABILITY sub-TLV and STATEFUL-PCE-CAPABILITY TLV ([RFC8231]) (with I flag set [RFC8281]) in OPEN Object to support the extensions defined in this document. If PCECC-CAPABILITY sub-TLV is advertised and STATEFUL-PCE-CAPABILITY TLV is not advertised in OPEN Object, it SHOULD send a PCerr message with Error-Type=19 (Invalid Operation) and Error-value=TBD (stateful PCE capability was not advertised) and terminate the session.

#### 5.4. LSP Operations

The PCEP messages pertaining to PCECC MUST include PATH-SETUP-TYPE TLV [I-D.ietf-pce-lsp-setup-type] in the SRP object to clearly identify the PCECC LSP is intended.

##### 5.4.1. Basic PCECC LSP Setup

In order to setup a LSP based on PCECC mechanism, a PCC MUST delegate the LSP by sending a PCRpt message with PST set for PCECC (see Section 7.2) and D (Delegate) flag (see [RFC8231]) set in the LSP object.

LSP-IDENTIFIER TLV MUST be included for PCECC LSP, the tuple uniquely identifies the LSP in the network. The LSP object is included in central controller's instructions (label download) to identify the PCECC LSP for this instruction. The PLSP-ID is the original identifier used by the ingress PCC, so the transit LSR could have multiple central controller instructions that have the same PLSP-ID. The PLSP-ID in combination with the source (in LSP-IDENTIFIER TLV) MUST be unique. The PLSP-ID is included for maintainability reasons. As per [RFC8281], the LSP object could include SPEAKER-ENTITY-ID TLV to identify the PCE that initiated these instructions. Also the CC-ID is unique on the PCEP session as described in Section 7.3.

When a PCE receives PCRpt message with D flags and PST Type set, it calculates the path and assigns labels along the path; and set up the

path by sending PCInitiate message to each node along the path of the LSP. The PCC generates a Path Computation State Report (PCRpt) and include the central controller's instruction (CCI) and the identified LSP. The CC-ID is uniquely identify the central controller's instruction within PCEP. The PCC further responds with the PCRpt messages including the CCI and LSP objects.

Once the central controller's instructions (label operations) are completed, the PCE SHOULD send the PCUpd message to the Ingress PCC. The PCUpd message is as per [RFC8231] SHOULD include the path information as calculated by the PCE.

Note that the PCECC LSPs MUST be delegated to a PCE at all times.

LSP deletion operation for PCECC LSP is same as defined in [RFC8231]. If the PCE receives PCRpt message for LSP deletion then it does Label cleanup operation as described in Section 5.4.2.2 for the corresponding LSP.

The Basic PCECC LSP setup sequence is as shown below.

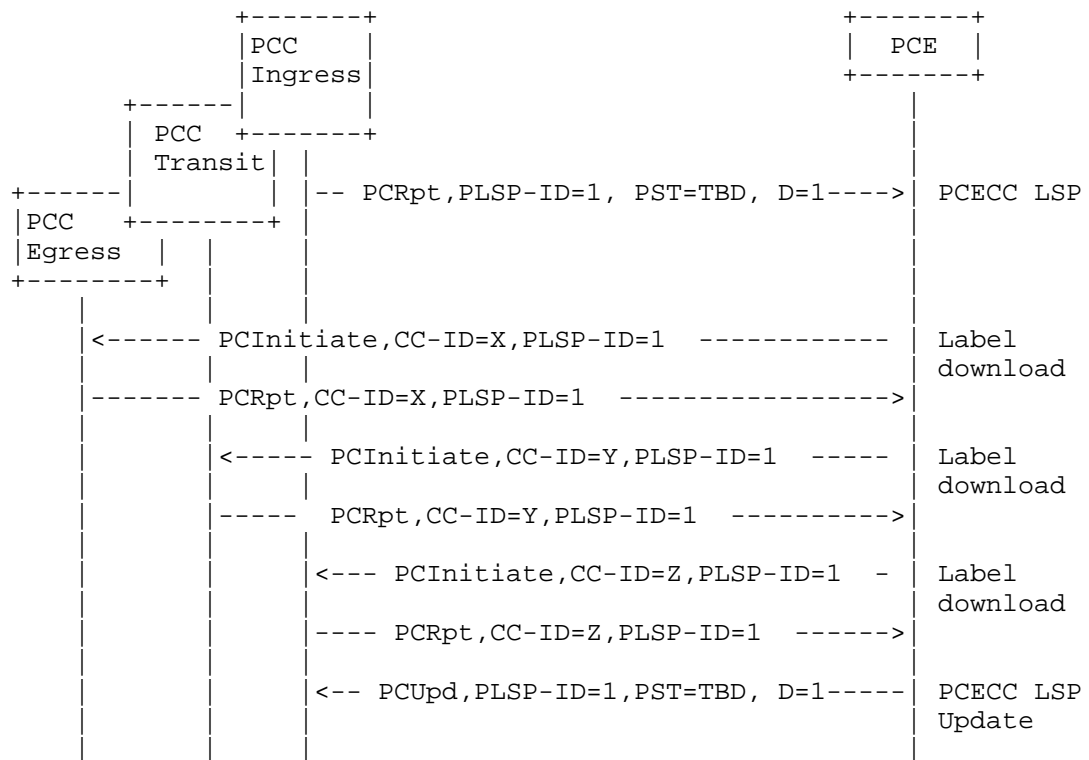


Figure 2: Basic PCECC LSP setup

The PCECC LSP are considered to be 'up' by default (on receipt of PCUpd message from PCE). The Ingress MAY further choose to deploy a data plane check mechanism and report the status back to the PCE via PCRpt message.

#### 5.4.2. Central Control Instructions

The new central controller's instructions (CCI) for the label operations in PCEP is done via the PCInitiate message, by defining a new PCEP Objects for CCI operations. Local label range of each PCC is assumed to be known at both the PCC and the PCE.

##### 5.4.2.1. Label Download

In order to setup an LSP based on PCECC, the PCE sends a PCInitiate message to each node along the path to download the Label instruction as described in Section 5.4.1.

The CCI object MUST be included, along with the LSP object in the PCInitiate message. The LSP-IDENTIFIER TLV MUST be included in LSP object. The SPEAKER-ENTITY-ID TLV SHOULD be included in LSP object.

If a node (PCC) receives a PCInitiate message which includes a Label to download as part of CCI, that is out of the range set aside for the PCE, it MUST send a PCErr message with Error-type=TBD (PCECC failure) and Error-value=TBD (Label out of range) and MUST include the SRP object to specify the error is for the corresponding label update via PCInitiate message. If a PCC receives a PCInitiate message but failed to download the Label entry, it MUST send a PCErr message with Error-type=TBD (PCECC failure) and Error-value=TBD (instruction failed) and MUST include the SRP object to specify the error is for the corresponding label update via PCInitiate message.

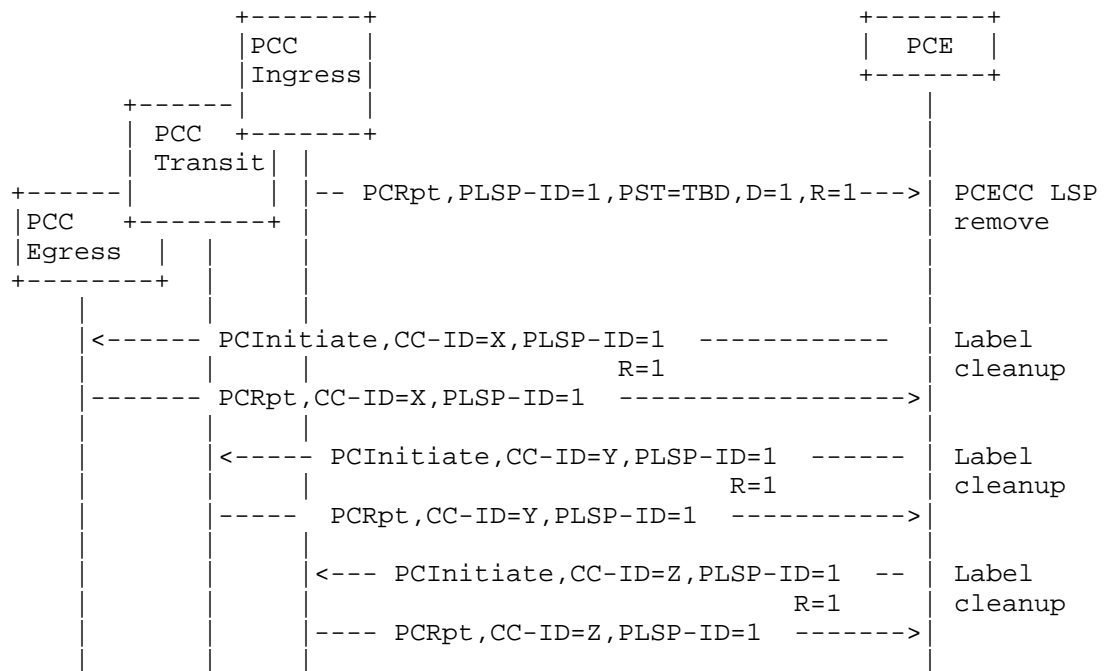
New PCEP object for central control instructions (CCI) is defined in Section 7.3.

#### 5.4.2.2. Label Cleanup

In order to delete an LSP based on PCECC, the PCE sends a central controller instructions via a PCInitiate message to each node along the path of the LSP to cleanup the Label forwarding instruction.

If the PCC receives a PCInitiate message but does not recognize the label in the CCI, the PCC MUST generate a PCErr message with Error-Type 19(Invalid operation) and Error-Value=TBD, "Unknown Label" and MUST include the SRP object to specify the error is for the corresponding label cleanup (via PCInitiate message).

The R flag in the SRP object defined in [RFC8281] specifies the deletion of Label Entry in the PCInitiate message.



As per [RFC8281], following the removal of the Label forwarding instruction, the PCC MUST send a PCRpt message. The SRP object in the PCRpt MUST include the SRP-ID-number from the PCInitiate message that triggered the removal. The R flag in the SRP object MUST be set.

#### 5.4.3. PCE Initiated PCECC LSP

The LSP Instantiation operation is same as defined in [RFC8281].

In order to setup a PCE Initiated LSP based on the PCECC mechanism, a PCE sends PCInitiate message with Path Setup Type set for PCECC (see Section 7.2) to the Ingress PCC.

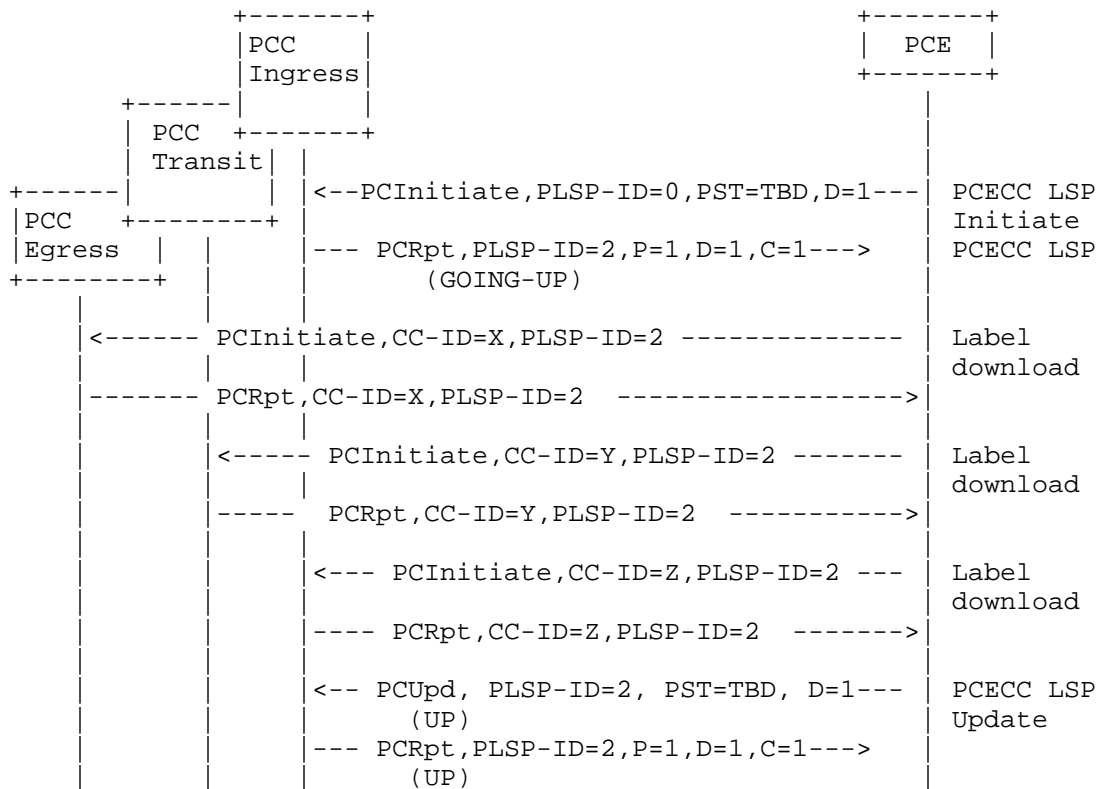
The Ingress PCC MUST also set D (Delegate) flag (see [RFC8231]) and C (Create) flag (see [RFC8281]) in LSP object of PCRpt message. The PCC responds with first PCRpt message with the status as "GOING-UP" and assigned PLSP-ID.

Note that the label forwarding instructions from PCECC are send after the initial PCInitiate and PCRpt exchange. This is done so that the PLSP-ID and other LSP identifiers can be obtained from the ingress and can be included in the label forwarding instruction in the next

PCInitiate message. The rest of the PCECC LSP setup operations are same as those described in Section 5.4.1.

The LSP deletion operation for PCE Initiated PCECC LSP is same as defined in [RFC8281]. The PCE should further perform Label entry cleanup operation as described in Section 5.4.2.2 for the corresponding LSP.

The PCE Initiated PCECC LSP setup sequence is shown below -

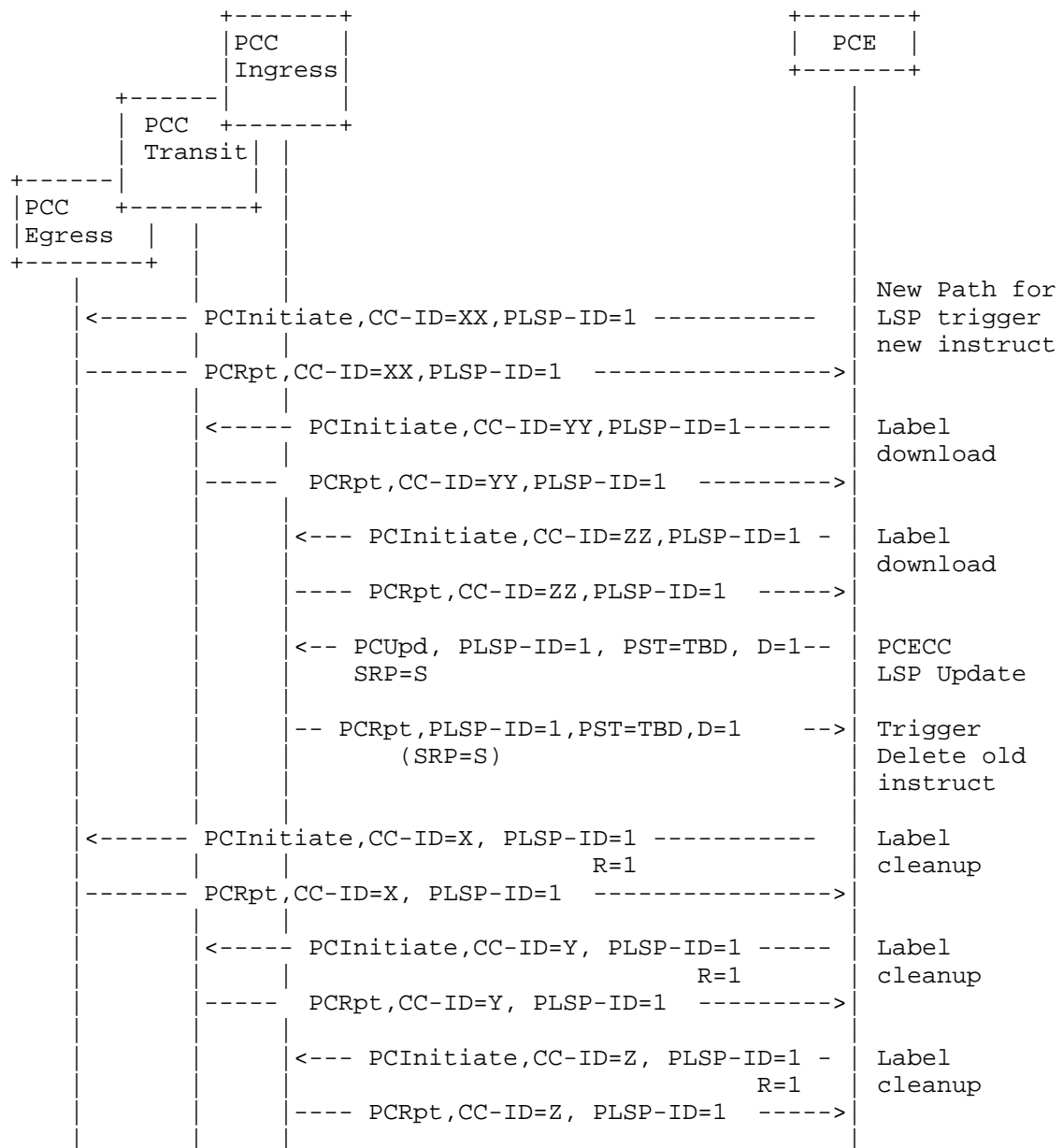


Once the label operations are completed, the PCE SHOULD send the PCUpd message to the Ingress PCC. The PCUpd message is as per [RFC8231].

#### 5.4.4. PCECC LSP Update

In case of a modification of PCECC LSP with a new path, a PCE sends a PCUpd message to the Ingress PCC. But to follow the make-before-break procedures, the PCECC first update new instructions based on the updated LSP and then update to ingress to switch traffic, before cleaning up the old instructions. A new CC-ID is used to identify the updated instruction, the existing identifiers in the LSP object identify the existing LSP. Once new instructions are downloaded, the PCE further updates the new path at the ingress which triggers the traffic switch on the updated path. The Ingress PCC acknowledges with a PCRpt message, on receipt of PCRpt message, the PCE does cleanup operation for the old LSP as described in Section 5.4.2.2.

The PCECC LSP Update sequence is shown below -



The modified PCECC LSP are considered to be 'up' by default. The Ingress MAY further choose to deploy a data plane check mechanism and report the status back to the PCE via PCRpt message.

#### 5.4.5. Re Delegation and Cleanup

As described in [RFC8281], a new PCE can gain control over the orphaned LSP. In case of PCECC LSP, the new PCE MUST also gain control over the central controllers instructions in the same way by sending a PCInitiate message that includes the SRP, LSP and CCI objects and carries the CC-ID and PLSP-ID identifying the instruction, it wants to take control of.

Further, as described in [RFC8281], the State Timeout Interval timer ensures that a PCE crash does not result in automatic and immediate disruption for the services using PCE-initiated LSPs. Similarly the central controller instructions are not removed immediately upon PCE failure. Instead, they are cleaned up on the expiration of this timer. This allows for network cleanup without manual intervention. The PCC MUST support removal of CCI as one of the behaviors applied on expiration of the State Timeout Interval timer.

#### 5.4.6. Synchronization of Central Controllers Instructions

The purpose of Central Controllers Instructions synchronization (labels in the context of this document) is to make sure that the PCE's view of CCI (Labels) matches with the PCC's Label allocation. This synchronization is performed as part of the LSP state synchronization as described in [RFC8231] and [RFC8233].

As per LSP State Synchronization [RFC8231], a PCC reports the state of its LSPs to the PCE using PCRpt messages and as per [RFC8281], PCE would initiate any missing LSPs and/or remove any LSPs that are not wanted. The same PCEP messages and procedure is also used for the Central Controllers Instructions synchronization. The PCRpt message includes the CCI and the LSP object to report the label forwarding instructions. The PCE would further remove any unwanted instructions or initiate any missing instructions.

#### 5.4.7. PCECC LSP State Report

As mentioned before, an Ingress PCC MAY choose to apply any OAM mechanism to check the status of LSP in the Data plane and MAY further send its status in PCRpt message to the PCE.

### 6. PCEP messages

As defined in [RFC5440], a PCEP message consists of a common header followed by a variable-length body made of a set of objects that can be either mandatory or optional. An object is said to be mandatory in a PCEP message when the object must be included for the message to be considered valid. For each PCEP message type, a set of rules is

defined that specify the set of objects that the message can carry. An implementation MUST form the PCEP messages using the object ordering specified in this document.

LSP-IDENTIFIERS TLV MUST be included in the LSP object for PCECC LSP.

### 6.1. The PCInitiate message

The PCInitiate message [RFC8281] can be used to download or remove the labels, the message has been extended as shown below -

```
<PCInitiate Message> ::= <Common Header>
 <PCE-initiated-lsp-list>
```

Where:

```
<Common Header> is defined in [RFC5440]
```

```
<PCE-initiated-lsp-list> ::= <PCE-initiated-lsp-request>
 [<PCE-initiated-lsp-list>]
```

```
<PCE-initiated-lsp-request> ::=
 (<PCE-initiated-lsp-instantiation>|
 <PCE-initiated-lsp-deletion>|
 <PCE-initiated-lsp-central-control>)
```

```
<PCE-initiated-lsp-central-control> ::= <SRP>
 <LSP>
 <cci-list>
```

```
<cci-list> ::= <CCI>
 [<cci-list>]
```

Where:

```
<PCE-initiated-lsp-instantiation> and
<PCE-initiated-lsp-deletion> are as per
[RFC8281].
```

The LSP and SRP object is defined in [RFC8231].

When PCInitiate message is used for central controller's instructions (labels), the SRP, LSP and CCI objects MUST be present. The SRP object is defined in [RFC8231] and if the SRP object is missing, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=10 (SRP object missing). The LSP object is defined in [RFC8231] and if the LSP object is missing, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=8 (LSP object missing). The CCI

object is defined in Section 7.3 and if the CCI object is missing, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=TBD (CCI object missing). More than one CCI object MAY be included in the PCInitiate message for the transit LSR.

To cleanup the SRP object must set the R (remove) bit.

At max two instances of CCI object would be included in case of transit LSR to encode both in-coming and out-going label forwarding instructions. Other instances MUST be ignored.

## 6.2. The PCRpt message

The PCRpt message can be used to report the labels that were allocated by the PCE, to be used during the state synchronization phase.

```
<PCRpt Message> ::= <Common Header>
 <state-report-list>
```

Where:

```
<state-report-list> ::= <state-report>[<state-report-list>]
```

```
<state-report> ::= (<lsp-state-report>|
 <central-control-report>)
```

```
<lsp-state-report> ::= [<SRP>]
 <LSP>
 <path>
```

```
<central-control-report> ::= [<SRP>]
 <LSP>
 <cci-list>
```

```
<cci-list> ::= <CCI>
 [<cci-list>]
```

Where:

<path> is as per [RFC8231] and the LSP and SRP object are also defined in [RFC8231].

When PCRpt message is used to report the central controller's instructions (labels), the LSP and CCI objects MUST be present. The LSP object is defined in [RFC8231] and if the LSP object is missing, the receiving PCE MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=8 (LSP object missing).

The CCI object is defined in Section 7.3 and if the CCI object is missing, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=TBD (CCI object missing). Two CCI object can be included in the PCRpt message for the transit LSR.

## 7. PCEP Objects

The PCEP objects defined in this document are compliant with the PCEP object format defined in [RFC5440].

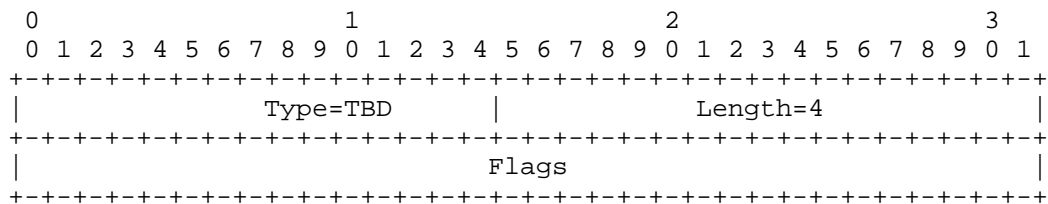
### 7.1. OPEN Object

This document defines a new optional TLVs for use in the OPEN Object.

#### 7.1.1. PCECC Capability sub-TLV

The PCECC-CAPABILITY sub-TLV is an optional TLV for use in the OPEN Object for PCECC capability advertisement in PATH-SETUP-TYPE-CAPABILITY TLV. Advertisement of the PCECC capability implies support of LSPs that are setup through PCECC as per PCEP extensions defined in this document.

Its format is shown in the following figure:



The type of the TLV is TBD and it has a fixed length of 4 octets.

The value comprises a single field - Flags (32 bits).

No flags are assigned right now.

Unassigned bits are considered reserved. They MUST be set to 0 on transmission and MUST be ignored on receipt.

## 7.2. PATH-SETUP-TYPE TLV

The PATH-SETUP-TYPE TLV is defined in [I-D.ietf-pce-lsp-setup-type]; this document defines a new PST value:

- o PST = TBD: Path is setup via PCECC mode.

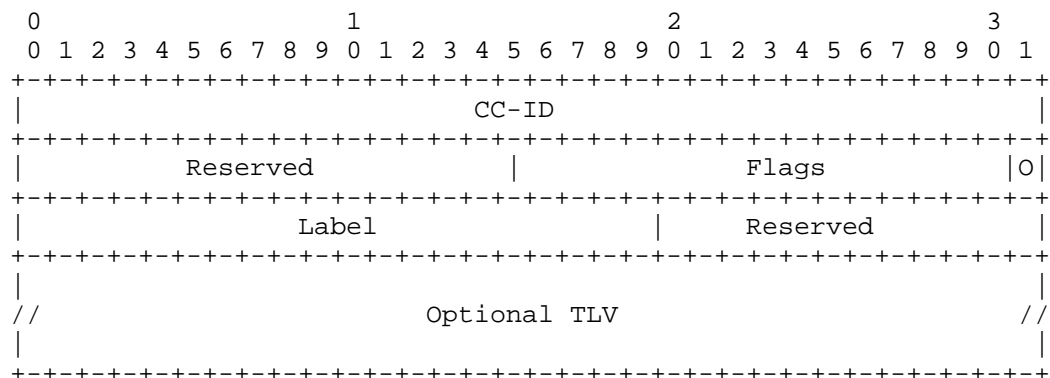
On a PCRpt/PCUpd/PCInitiate message, the PST=TBD in PATH-SETUP-TYPE TLV in SRP object indicates that this LSP was setup via a PCECC based mechanism.

## 7.3. CCI Object

The Central Control Instructions (CCI) Object is used by the PCE to specify the forwarding instructions (Label information in the context of this document) to the PCC, and MAY be carried within PCInitiate or PCRpt message for label download.

CCI Object-Class is TBD.

CCI Object-Type is 1 for the MPLS Label.



The fields in the CCI object are as follows:

**CC-ID:** A PCEP-specific identifier for the CCI information. A PCE creates an CC-ID for each instruction, the value is unique within the scope of the PCE and is constant for the lifetime of a PCEP session. The values 0 and 0xFFFFFFFF are reserved and MUST NOT be used.

**Flags:** is used to carry any additional information pertaining to the CCI. Currently, the following flag bit is defined:

- \* O bit(Out-label) : If the bit is set, it specifies the label is the OUT label and it is mandatory to encode the next-hop information (via IPV4-ADDRESS TLV or IPV6-ADDRESS TLV or UNNUMBERED-IPV4-ID-ADDRESS TLV in the CCI object). If the bit is not set, it specifies the label is the IN label and it is optional to encode the local interface information (via IPV4-ADDRESS TLV or IPV6-ADDRESS TLV or UNNUMBERED-IPV4-ID-ADDRESS TLV in the CCI object).

Label (20-bit): The Label information.

Reserved (12 bit): Set to zero while sending, ignored on receive.

#### 7.3.1. Address TLVs

This document defines the following TLVs for the CCI object to associate the next-hop information in case of an outgoing label and local interface information in case of an incoming label.

## IPv4-ADDRESS TLV:

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| Type=TBD |
+-----+-----+-----+-----+-----+-----+-----+-----+
| IPv4 address |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

## IPv6-ADDRESS TLV:

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| Type=TBD |
+-----+-----+-----+-----+-----+-----+-----+-----+
| // //
| IPv6 address (16 bytes) |
| // //
+-----+-----+-----+-----+-----+-----+-----+-----+

```

## UNNUMBERED-IPv4-ID-ADDRESS TLV:

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| Type=TBD |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Node-ID |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Interface ID |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

The address TLVs are as follows:

IPv4-ADDRESS TLV: an IPv4 address.

IPv6-ADDRESS TLV: an IPv6 address.

UNNUMBERED-IPv4-ID-ADDRESS TLV: a pair of Node ID / Interface ID tuples.

## 8. Security Considerations

The security considerations described in [RFC8231] and [RFC8281] apply to the extensions described in this document. Additional considerations related to a malicious PCE are introduced.

### 8.1. Malicious PCE

PCE has complete control over PCC to update the labels and can cause the LSP's to behave inappropriate and cause cause major impact to the network. As a general precaution, it is RECOMMENDED that these PCEP extensions only be activated on authenticated and encrypted sessions across PCEs and PCCs belonging to the same administrative authority, using Transport Layer Security (TLS) [RFC8253], as per the recommendations and best current practices in [RFC7525].

## 9. Manageability Considerations

### 9.1. Control of Function and Policy

A PCE or PCC implementation SHOULD allow to configure to enable/disable PCECC capability as a global configuration.

### 9.2. Information and Data Models

[RFC7420] describes the PCEP MIB, this MIB can be extended to get the PCECC capability status.

The PCEP YANG module [I-D.ietf-pce-pcep-yang] could be extended to enable/disable PCECC capability.

### 9.3. 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].

### 9.4. 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].

### 9.5. Requirements On Other Protocols

PCEP extensions defined in this document do not put new requirements on other protocols.

## 9.6. Impact On Network Operations

PCEP extensions defined in this document do not put new requirements on network operations.

## 10. IANA Considerations

### 10.1. PCEP TLV Type Indicators

IANA is requested to confirm the early allocation of the following TLV Type Indicator values within the "PCEP TLV Type Indicators" sub-registry of the PCEP Numbers registry, and to update the reference in the registry to point to this document, when it is an RFC:

| Value | Meaning                        | Reference     |
|-------|--------------------------------|---------------|
| TBD   | PCECC-CAPABILITY               | This document |
| TBD   | IPV4-ADDRESS TLV               | This document |
| TBD   | IPV6-ADDRESS TLV               | This document |
| TBD   | UNNUMBERED-IPV4-ID-ADDRESS TLV | This document |

### 10.2. New Path Setup Type Registry

IANA is requested to allocate new PST Field in PATH- SETUP-TYPE TLV. The allocation policy for this new registry should be by IETF Consensus. The new registry should contain the following value:

| Value | Description                                        | Reference     |
|-------|----------------------------------------------------|---------------|
| TBD   | Traffic engineering path is setup using PCECC mode | This document |

### 10.3. PCEP Object

IANA is requested to allocate new registry for CCI PCEP object.

| Object-Class Value | Name            | Reference     |
|--------------------|-----------------|---------------|
| TBD                | CCI Object-Type | This document |
|                    | 1               | MPLS Label    |

### 10.4. CCI Object Flag Field

IANA is requested to create a registry to manage the Flag field of the CCI object.

One bit to be defined for the CCI Object flag field in this document:

Codespace of the Flag field (CCI Object)

| Bit | Description                     | Reference     |
|-----|---------------------------------|---------------|
| 7   | Specifies label<br>is out label | This document |

### 10.5. PCEP-Error Object

IANA is requested to allocate new error types and error values within the "PCEP-ERROR Object Error Types and Values" sub-registry of the PCEP Numbers registry for the following errors:

| Error-Type | Meaning                   |                                                                     |
|------------|---------------------------|---------------------------------------------------------------------|
| -----      | -----                     |                                                                     |
| 19         | Invalid operation.        |                                                                     |
|            | Error-value = TBD :       | Attempted PCECC operations when PCECC capability was not advertised |
|            | Error-value = TBD :       | Stateful PCE capability was not advertised                          |
|            | Error-value = TBD :       | Unknown Label                                                       |
| 6          | Mandatory Object missing. |                                                                     |
|            | Error-value = TBD :       | CCI object missing                                                  |
| TBD        | PCECC failure.            |                                                                     |
|            | Error-value = TBD :       | Label out of range.                                                 |
|            | Error-value = TBD :       | Instruction failed.                                                 |

### 11. Acknowledgments

We would like to thank Robert Tao, Changjing Yan, Tieying Huang and Avantika for their useful comments and suggestions.

### 12. References

#### 12.1. 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>>.

- [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>>.
- [RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <<https://www.rfc-editor.org/info/rfc7420>>.
- [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>>.
- [RFC8233] Dhody, D., Wu, Q., Manral, V., Ali, Z., and K. Kumaki, "Extensions to the Path Computation Element Communication Protocol (PCEP) to Compute Service-Aware Label Switched Paths (LSPs)", RFC 8233, DOI 10.17487/RFC8233, September 2017, <<https://www.rfc-editor.org/info/rfc8233>>.
- [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>>.

## 12.2. 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, <<https://www.rfc-editor.org/info/rfc4655>>.

- [RFC7025] Otani, T., Ogaki, K., Caviglia, D., Zhang, F., and C. Margaria, "Requirements for GMPLS Applications of PCE", RFC 7025, DOI 10.17487/RFC7025, September 2013, <<https://www.rfc-editor.org/info/rfc7025>>.
- [RFC7399] Farrel, A. and D. King, "Unanswered Questions in the Path Computation Element Architecture", RFC 7399, DOI 10.17487/RFC7399, October 2014, <<https://www.rfc-editor.org/info/rfc7399>>.
- [RFC7491] King, D. and A. Farrel, "A PCE-Based Architecture for Application-Based Network Operations", RFC 7491, DOI 10.17487/RFC7491, March 2015, <<https://www.rfc-editor.org/info/rfc7491>>.
- [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>>.
- [RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <<https://www.rfc-editor.org/info/rfc8283>>.
- [I-D.ietf-teas-pcecc-use-cases]  
Zhao, Q., Li, Z., Khasanov, B., Ke, Z., Fang, L., Zhou, C., Communications, T., and A. Rachitskiy, "The Use Cases for Using PCE as the Central Controller(PCECC) of LSPs", draft-ietf-teas-pcecc-use-cases-01 (work in progress), May 2017.
- [I-D.ietf-pce-lsp-setup-type]  
Sivabalan, S., Tantsura, J., Minei, I., Varga, R., and J. Hardwick, "Conveying path setup type in PCEP messages", draft-ietf-pce-lsp-setup-type-10 (work in progress), May 2018.
- [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-07 (work in progress), March 2018.

[I-D.zhao-pce-pcep-extension-pce-controller-sr]

Zhao, Q., Li, Z., Dhody, D., Karunanithi, S., Farrel, A.,  
and C. Zhou, "PCEP Procedures and Protocol Extensions for  
Using PCE as a Central Controller (PCECC) of SR-LSPs",  
draft-zhao-pce-pcep-extension-pce-controller-sr-02 (work  
in progress), March 2018.

## Appendix A. Contributor Addresses

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

EMail: udayasreereddy@gmail.com

Mahendra Singh Negi  
Huawei Technologies  
Divyashree Techno Park, Whitefield  
Bangalore, Karnataka 560066  
India

EMail: mahendrasingh@huawei.com

Katherine Zhao  
Huawei Technologies  
2330 Central Expressway  
Santa Clara, CA 95050  
USA

EMail: katherine.zhao@huawei.com

Boris Zhang  
Telus Ltd.  
Toronto  
Canada

EMail: boris.zhang@telus.com

## Authors' Addresses

Quintin Zhao  
Huawei Technologies  
125 Nagog Technology Park  
Acton, MA 01719  
USA

EMail: quintin.zhao@huawei.com

Zhenbin Li  
Huawei Technologies  
Huawei Bld., No.156 Beiqing Rd.  
Beijing 100095  
China

EMail: lizhenbin@huawei.com

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

EMail: dhruv.ietf@gmail.com

Satish Karunanithi  
Huawei Technologies  
Divyashree Techno Park, Whitefield  
Bangalore, Karnataka 560066  
India

EMail: satishk@huawei.com

Adrian Farrel  
Juniper Networks, Inc  
UK

EMail: adrian@olddog.co.uk

Chao Zhou  
Cisco Systems

EMail: choa.zhou@cisco.com

PCE Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: December 20, 2018

Q. Zhao  
Z. Li  
D. Dhody  
S. Karunanithi  
Huawei Technologies  
A. Farrel  
Juniper Networks, Inc  
C. Zhou  
Cisco Systems  
June 18, 2018

PCEP Procedures and Protocol Extensions for Using PCE as a Central  
Controller (PCECC) of SR-LSPs  
draft-zhao-pce-pcep-extension-pce-controller-sr-03

Abstract

The Path Computation Element (PCE) is a core component of Software-Defined Networking (SDN) systems. It can compute optimal paths for traffic across a network and can also update the paths to reflect changes in the network or traffic demands.

PCE was developed to derive paths for MPLS Label Switched Paths (LSPs), which are supplied to the head end of the LSP using the Path Computation Element Communication Protocol (PCEP). But SDN has a broader applicability than signaled (G)MPLS traffic-engineered (TE) networks, and the PCE may be used to determine paths in a range of use cases. PCEP has been proposed as a control protocol for use in these environments to allow the PCE to be fully enabled as a central controller.

A PCE-based central controller (PCECC) can simplify the processing of a distributed control plane by blending it with elements of SDN and without necessarily completely replacing it. Thus, the LSP can be calculated/setup/initiated and the label forwarding entries can also be downloaded through a centralized PCE server to each network devices along the path while leveraging the existing PCE technologies as much as possible.

This document specifies the procedures and PCEP protocol extensions when a PCE-based controller is also responsible for configuring the forwarding actions on the routers, in addition to computing the paths for packet flows in a segment routing network and telling the edge routers what instructions to attach to packets as they enter the network.

## 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 20, 2018.

## Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 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

|                                                                                                   |   |
|---------------------------------------------------------------------------------------------------|---|
| 1. Introduction . . . . .                                                                         | 3 |
| 1.1. Requirements Language . . . . .                                                              | 5 |
| 2. Terminology . . . . .                                                                          | 5 |
| 3. PCECC SR . . . . .                                                                             | 5 |
| 4. PCEP Requirements . . . . .                                                                    | 6 |
| 5. Procedures for Using the PCE as the Central Controller<br>(PCECC) in Segment Routing . . . . . | 6 |
| 5.1. Stateful PCE Model . . . . .                                                                 | 6 |
| 5.2. New LSP Functions . . . . .                                                                  | 6 |
| 5.3. PCECC Capability Advertisement . . . . .                                                     | 7 |
| 5.4. PCEP session IP address and TEDB Router ID . . . . .                                         | 7 |
| 5.5. LSP Operations . . . . .                                                                     | 7 |
| 5.5.1. PCECC Segment Routing (SR) . . . . .                                                       | 8 |
| 5.5.1.1. PCECC SR Node/Prefix SID allocation . . . . .                                            | 8 |

|                    |                                                |    |
|--------------------|------------------------------------------------|----|
| 5.5.1.2.           | PCECC SR Adjacency Label allocation . . . . .  | 9  |
| 5.5.1.3.           | Redundant PCEs . . . . .                       | 11 |
| 5.5.1.4.           | Re Delegation and Cleanup . . . . .            | 11 |
| 5.5.1.5.           | Synchronization of Label Allocations . . . . . | 11 |
| 6.                 | PCEP messages . . . . .                        | 11 |
| 6.1.               | Central Control Instructions . . . . .         | 11 |
| 6.1.1.             | The PCInitiate message . . . . .               | 11 |
| 6.1.2.             | The PCRpt message . . . . .                    | 12 |
| 7.                 | PCEP Objects . . . . .                         | 13 |
| 7.1.               | OPEN Object . . . . .                          | 13 |
| 7.1.1.             | PCECC Capability sub-TLV . . . . .             | 13 |
| 7.2.               | PATH-SETUP-TYPE TLV . . . . .                  | 14 |
| 7.3.               | CCI Object . . . . .                           | 14 |
| 7.4.               | FEC Object . . . . .                           | 15 |
| 8.                 | Security Considerations . . . . .              | 17 |
| 9.                 | Manageability Considerations . . . . .         | 17 |
| 9.1.               | Control of Function and Policy . . . . .       | 17 |
| 9.2.               | Information and Data Models . . . . .          | 18 |
| 9.3.               | Liveness Detection and Monitoring . . . . .    | 18 |
| 9.4.               | Verify Correct Operations . . . . .            | 18 |
| 9.5.               | Requirements On Other Protocols . . . . .      | 18 |
| 9.6.               | Impact On Network Operations . . . . .         | 18 |
| 10.                | IANA Considerations . . . . .                  | 18 |
| 10.1.              | PCECC-CAPABILITY TLV . . . . .                 | 18 |
| 10.2.              | New Path Setup Type Registry . . . . .         | 19 |
| 10.3.              | PCEP Object . . . . .                          | 19 |
| 10.4.              | PCEP-Error Object . . . . .                    | 19 |
| 11.                | Acknowledgments . . . . .                      | 19 |
| 12.                | References . . . . .                           | 20 |
| 12.1.              | Normative References . . . . .                 | 20 |
| 12.2.              | Informative References . . . . .               | 21 |
| Appendix A.        | Contributor Addresses . . . . .                | 24 |
| Authors' Addresses | . . . . .                                      | 24 |

## 1. Introduction

The Path Computation Element (PCE) [RFC4655] was developed to offload path computation function from routers in an MPLS traffic-engineered network. Since then, the role and function of the PCE has grown to cover a number of other uses (such as GMPLS [RFC7025]) and to allow delegated control [RFC8231] and PCE-initiated use of network resources [RFC8281].

According to [RFC7399], Software-Defined Networking (SDN) refers to a separation between the control elements and the forwarding components so that software running in a centralized system, called a controller, can act to program the devices in the network to behave in specific ways. A required element in an SDN architecture is a

component that plans how the network resources will be used and how the devices will be programmed. It is possible to view this component as performing specific computations to place traffic flows within the network given knowledge of the availability of network resources, how other forwarding devices are programmed, and the way that other flows are routed. This is the function and purpose of a PCE, and the way that a PCE integrates into a wider network control system (including an SDN system) is presented in [RFC7491].

In early PCE implementations, where the PCE was used to derive paths for MPLS Label Switched Paths (LSPs), paths were requested by network elements (known as Path Computation Clients (PCCs)), and the results of the path computations were supplied to network elements using the Path Computation Element Communication Protocol (PCEP) [RFC5440]. This protocol was later extended to allow a PCE to send unsolicited requests to the network for LSP establishment [RFC8281].

[RFC8283] introduces the architecture for PCE as a central controller as an extension of the architecture described in [RFC4655] and assumes the continued use of PCEP as the protocol used between PCE and PCC. [RFC8283] further examines the motivations and applicability for PCEP as a Southbound Interface (SBI), and introduces the implications for the protocol. [I-D.ietf-teas-pcecc-use-cases] describes the use cases for the PCECC architecture.

[I-D.zhao-pce-pcep-extension-for-pce-controller] specify the procedures and PCEP protocol extensions for using the PCE as the central controller for static LSPs, where LSPs can be provisioned as explicit label instructions at each hop on the end-to-end path.

Segment Routing (SR) technology leverages the source routing and tunneling paradigms. A source node can choose a path without relying on hop-by-hop signaling protocols such as LDP or RSVP-TE. Each path is specified as a set of "segments" advertised by link-state routing protocols (IS-IS or OSPF). [I-D.ietf-spring-segment-routing] provides an introduction to SR architecture. The corresponding IS-IS and OSPF extensions are specified in [I-D.ietf-isis-segment-routing-extensions] and [I-D.ietf-ospf-segment-routing-extensions], respectively. It relies on a series of forwarding instructions being placed in the header of a packet. The segment routing architecture supports operations that can be used to steer packet flows in a network, thus providing a form of traffic engineering. [I-D.ietf-pce-segment-routing] specify the SR specific PCEP extensions.

PCECC may further use PCEP protocol for SR SID (Segment Identifier) distribution on the SR nodes with some benefits.

This document specifies the procedures and PCEP protocol extensions when a PCE-based controller is also responsible for configuring the forwarding actions on the routers (SR SID distribution in this case), in addition to computing the paths for packet flows in a segment routing network and telling the edge routers what instructions to attach to packets as they enter the network.

### 1.1. 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.

## 2. Terminology

Terminologies used in this document is same as described in the draft [RFC8283] and [I-D.ietf-teas-pcecc-use-cases].

## 3. PCECC SR

[I-D.ietf-pce-segment-routing] specifies extensions to PCEP that allow a stateful PCE to compute, update or initiate SR-TE paths. An ingress node of an SR-TE path appends all outgoing packets with a list of MPLS labels (SIDs). This is encoded in SR-ERO subobject, capable of carrying a label (SID) as well as the identity of the node/adjacency label (SID).

The notion of segment and SID is defined in [I-D.ietf-spring-segment-routing], which fits the MPLS architecture [RFC3031] as the label which is managed by a local allocation process of LSR (similarly to other MPLS signaling protocols) [I-D.ietf-spring-segment-routing-mpls]. The SR information such as node/adjacency label (SID) is flooded via IGP as specified in [I-D.ietf-isis-segment-routing-extensions] and [I-D.ietf-ospf-segment-routing-extensions].

As per [RFC8283], PCE as a central controller can allocate and provision the node/prefix/adjacency label (SID) via PCEP.

Rest of the processing is similar to existing stateful PCE with SR mechanism.

For the purpose of this document, it is assumed that label range to be used by a PCE is set on both PCEP peers. Further, a global label range is assumed to be set on all PCEP peers in the SR domain.

#### 4. PCEP Requirements

Following key requirements for PCECC-SR should be considered when designing the PCECC based solution:

- o PCEP speaker supporting this draft MUST have the capability to advertise its PCECC-SR capability to its peers.
- o PCEP speaker not supporting this draft MUST be able to reject PCECC-SR related message with a reason code that indicates no support for PCECC.
- o PCEP procedures MUST provide a means to update (or cleanup) the label- map entry to the PCC.
- o PCEP procedures SHOULD provide a means to synchronize the SR labels allocations between PCE to PCC in the PCEP messages.

#### 5. Procedures for Using the PCE as the Central Controller (PCECC) in Segment Routing

##### 5.1. Stateful PCE Model

Active stateful PCE is described in [RFC8231]. PCE as a central controller (PCECC) reuses existing Active stateful PCE mechanism as much as possible to control the LSP.

##### 5.2. New LSP Functions

This document uses the same PCEP messages and its extensions which are described in [I-D.zhao-pce-pcep-extension-for-pce-controller] for PCECC-SR as well.

PCEP messages PCRpt, PCInitiate, PCUpd are also used to send LSP Reports, LSP setup and LSP update respectively. The extended PCInitiate message in [I-D.zhao-pce-pcep-extension-for-pce-controller] is used to download or cleanup central controller's instructions (CCIs) (SR SID in scope of this document). The extended PCRpt message described in [I-D.zhao-pce-pcep-extension-for-pce-controller] is also used to report the CCIs (SR SIDs) from PCC to PCE.

[I-D.zhao-pce-pcep-extension-for-pce-controller] specify an object called CCI for the encoding of central controller's instructions. This document extends the CCI by defining a new object-type for segment routing. The PCEP messages are extended in this document to handle the PCECC operations for SR.

### 5.3. PCECC Capability Advertisement

During PCEP Initialization Phase, PCEP Speakers (PCE or PCC) advertise their support of PCECC extensions. A PCEP Speaker includes the "PCECC Capability" sub-TLV, described in [I-D.zhao-pce-pcep-extension-for-pce-controller].

A new S-bit is added in PCECC-CAPABILITY sub-TLV to indicate support for PCECC-SR. A PCC MUST set S-bit in PCECC-CAPABILITY sub-TLV and include SR-PCE-CAPABILITY sub-TLV ([I-D.ietf-pce-segment-routing]) in OPEN Object (inside the the PATH-SETUP-TYPE-CAPABILITY TLV) to support the PCECC SR extensions defined in this document. If S-bit is set in PCECC-CAPABILITY sub-TLV and SR-PCE-CAPABILITY sub-TLV is not advertised in OPEN Object, PCE SHOULD send a PCERR message with Error-Type=19 (Invalid Operation) and Error-value=TBD(SR capability was not advertised) and terminate the session.

### 5.4. PCEP session IP address and TEDB Router ID

PCE may construct its TEDB by participating in the IGP ([RFC3630] and [RFC5305] for MPLS-TE; [RFC4203] and [RFC5307] for GMPLS). An alternative is offered by BGP-LS [RFC7752] and [I-D.dhodylee-pce-pcep-ls].

PCEP [RFC5440] speaker MAY use any IP address while creating a TCP session. It is important to link the session IP address with the Router ID in TEDB for successful PCECC operations.

During PCEP Initialization Phase, PCC SHOULD advertise the TE mapping information. Thus a PCC includes the "Node Attributes TLV" [I-D.dhodylee-pce-pcep-ls] with "IPv4/IPv6 Router-ID of Local Node", in the OPEN Object for this purpose. [RFC7752] describes the usage as auxiliary Router-IDs that the IGP might be using, e.g., for TE purposes. If there are more than one auxiliary Router-ID of a given type, then multiple TLVs are used to encode them.

If "IPv4/IPv6 Router-ID" TLV is not present, the TCP session IP address is directly used for the mapping purpose.

### 5.5. LSP Operations

The PCEP messages pertaining to PCECC-SR MUST include PATH-SETUP-TYPE TLV [I-D.ietf-pce-lsp-setup-type] with PST=TBD in the SRP object to clearly identify the PCECC-SR LSP is intended.

#### 5.5.1. PCECC Segment Routing (SR)

Segment Routing (SR) as described in [I-D.ietf-spring-segment-routing] depends on "segments" that are advertised by Interior Gateway Protocols (IGPs). The SR-node allocates and advertises the SID (node, adj etc) and flood via the IGP. This document proposes a new mechanism where PCE allocates the SID (label/index/SID) centrally and uses PCEP to advertise the SID. In some deployments PCE (and PCEP) are better suited than IGP because of centralized nature of PCE and direct TCP based PCEP session to the node.

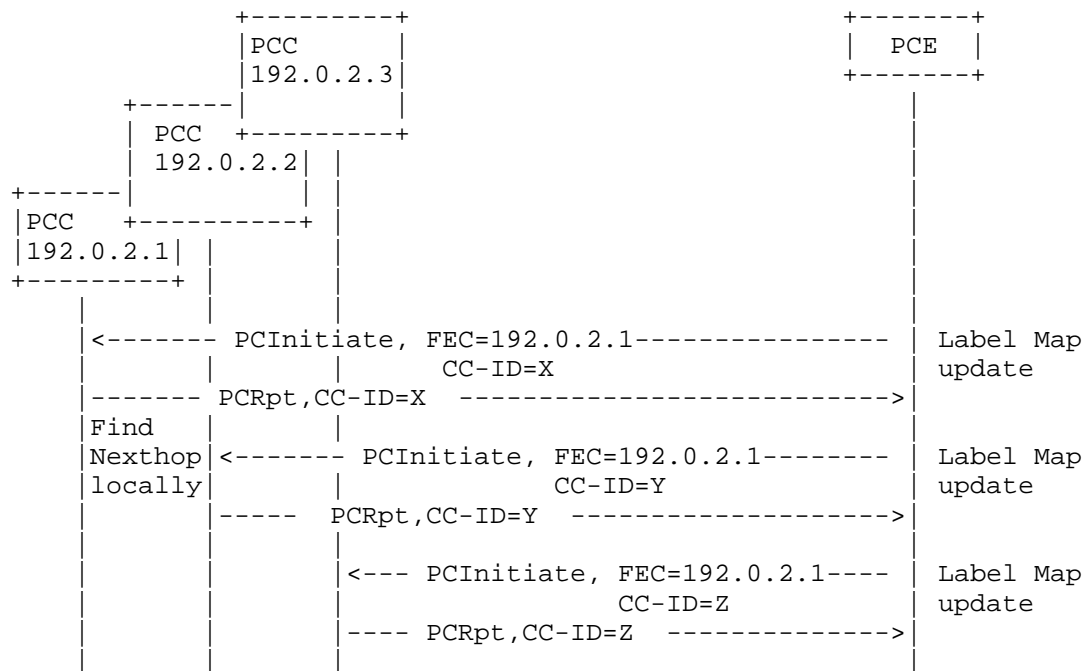
##### 5.5.1.1. PCECC SR Node/Prefix SID allocation

Each node (PCC) is allocated a node-SID by the PCECC. The PCECC sends PCInitiate message to update the label map of each node to all the nodes in the domain. The TE router ID is determined from the TEDB or from "IPv4/IPv6 Router-ID" Sub-TLV [I-D.dhodylee-pce-pcep-ls], in the OPEN Object Section 5.4.

It is RECOMMENDED that PCEP session with PCECC SR capability to use a different session IP address during TCP session establishment than the node Router ID in TEDB, to make sure that the PCEP session does not get impacted by the SR Node/Prefix Label maps (Section 5.4).

If a node (PCC) receives a PCInitiate message with a CCI encoding a SID, out of the range set aside for the SRGB, it MUST send a PCErr message with Error-type=TBD (PCECC failure) and Error-value=TBD (SID out of range) and MUST include the SRP object to specify the error is for the corresponding label update via PCInitiate message.

On receiving the label map, each node (PCC) uses the local information to determine the next-hop and download the label forwarding instructions accordingly. The PCInitiate message in this case MUST NOT have LSP object but uses the new FEC object defined in this document.



The forwarding behavior and the end result is similar to IGP based "Node-SID" in SR. Thus, from anywhere in the domain, it enforces the ECMP-aware shortest-path forwarding of the packet towards the related node.

PCE relies on the Node/Prefix Label cleanup using the same PCInitiate message.

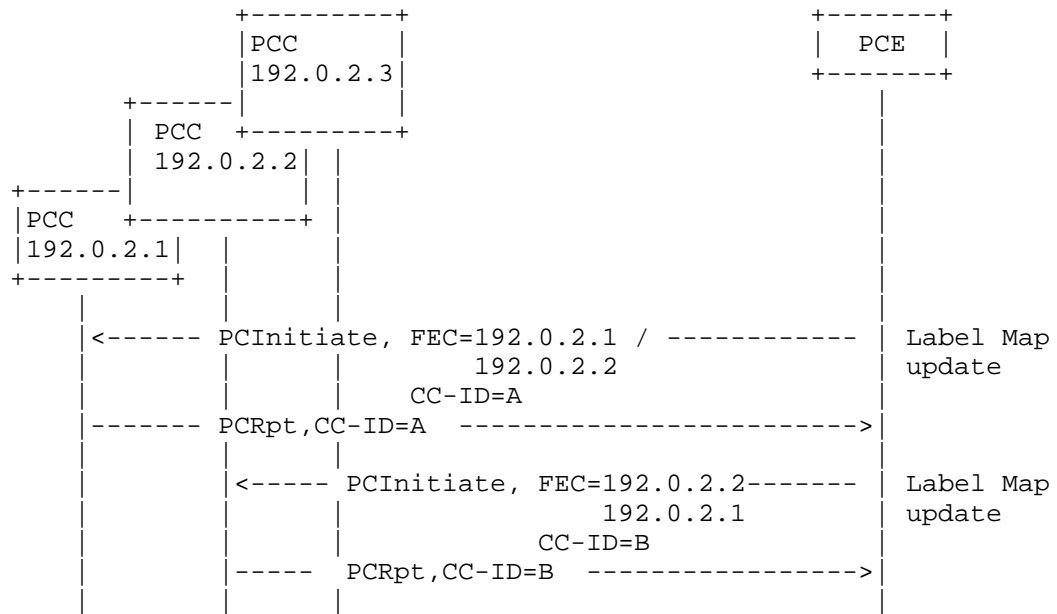
The above example Figure 1 depict FEC and PCEP speakers that uses IPv4 address. Similarly IPv6 address (such as 2001:DB8::1) can be used during PCEP session establishment as well in FEC object as described in this specification.

#### 5.5.1.2. PCECC SR Adjacency Label allocation

[I-D.ietf-pce-segment-routing] extends PCEP to allow a stateful PCE to compute and initiate SR-TE paths, as well as a PCC to request a path subject to certain constraint(s) and optimization criteria in SR networks.

For PCECC SR, apart from node-SID, Adj-SID is used where each adjacency is allocated an Adj-SID by the PCECC. The PCECC sends PCInitiate message to update the label map of each Adj to the

corresponding nodes in the domain. Each node (PCC) download the label forwarding instructions accordingly. Similar to SR Node/Prefix Label allocation, the PCInitiate message in this case MUST NOT have LSP object but uses the new FEC object defined in this document.



The forwarding behavior and the end result is similar to IGP based "Adj-SID" in SR.

The Path Setup Type for segment routing MUST be set for PCECC SR = TBD (see Section 7.2). All PCEP procedures and mechanism are similar to [I-D.ietf-pce-segment-routing].

PCE relies on the Adj label cleanup using the same PCInitiate message.

The above example Figure 2 depict FEC and PCEP speakers that uses IPv4 address. Similarly IPv6 address (such as 2001:DB8::1, 2001:DB8::2) can be used during PCEP session establishment as well in FEC object as described in this specification.

#### 5.5.1.3. Redundant PCEs

[I-D.litkowski-pce-state-sync] describes synchronization mechanism between the stateful PCEs. The SR SIDs allocated by a PCE MUST also be synchronized among PCEs for PCECC SR state synchronization. Note that the SR SIDs are independent to the PCECC-SR LSP, and remains intact till any topology change. The redundant PCEs MUST have a common view of all SR SIDs allocated in the domain.

#### 5.5.1.4. Re Delegation and Cleanup

[I-D.zhao-pce-pcep-extension-for-pce-controller] describes the action needed for CCIs for the Basic PCECC LSP on this terminated session. Similarly actions should be applied for the SR SID as well.

#### 5.5.1.5. Synchronization of Label Allocations

[I-D.zhao-pce-pcep-extension-for-pce-controller] describes the synchronization of Central Controller's Instructions (CCI) via LSP state synchronization as described in [RFC8231] and [RFC8232]. Same procedures should be applied for SR SIDs as well.

### 6. PCEP messages

As defined in [RFC5440], a PCEP message consists of a common header followed by a variable-length body made of a set of objects that can be either mandatory or optional. An object is said to be mandatory in a PCEP message when the object must be included for the message to be considered valid. For each PCEP message type, a set of rules is defined that specify the set of objects that the message can carry. An implementation MUST form the PCEP messages using the object ordering specified in this document.

#### 6.1. Central Control Instructions

##### 6.1.1. The PCInitiate message

The PCInitiate Message defined in [RFC8281] and extended in [I-D.zhao-pce-pcep-extension-for-pce-controller] is further extended to support SR based central control instructions.

The format of the extended PCInitiate message is as follows:

```

<PCInitiate Message> ::= <Common Header>
 <PCE-initiated-lsp-list>

```

Where:

<Common Header> is defined in [RFC5440]

```

<PCE-initiated-lsp-list> ::= <PCE-initiated-lsp-request>
 [<PCE-initiated-lsp-list>]

```

```

<PCE-initiated-lsp-request> ::=
 (<PCE-initiated-lsp-instantiation>|
 <PCE-initiated-lsp-deletion>|
 <PCE-initiated-lsp-central-control>)

```

```

<PCE-initiated-lsp-central-control> ::= <SRP>
 (<LSP>
 <cci-list>)|
 (<FEC>
 <CCI>)

```

```

<cci-list> ::= <CCI>
 [<cci-list>]

```

Where:

<PCE-initiated-lsp-instantiation> and  
 <PCE-initiated-lsp-deletion> are as per  
 [RFC8281].

The LSP and SRP object is defined in [RFC8231].

When PCInitiate message is used to distribute SR SIDs, the SRP, FEC and CCI objects MUST be present. The error handling for missing SRP or CCI object is as per [I-D.zhao-pce-pcep-extension-for-pce-controller]. If the FEC object is missing, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=TBD (FEC object missing).

To cleanup the SRP object must set the R (remove) bit.

#### 6.1.2. The PCRpt message

The PCRpt message can be used to report the SR instructions received from the central controller (PCE) during the state synchronization phase.

The format of the PCRpt message is as follows:

```
<PCRpt Message> ::= <Common Header>
 <state-report-list>
```

Where:

```
<state-report-list> ::= <state-report>[<state-report-list>]
```

```
<state-report> ::= (<lsp-state-report>|
 <central-control-report>)
```

```
<lsp-state-report> ::= [<SRP>]
 <LSP>
 <path>
```

```
<central-control-report> ::= [<SRP>]
 (<LSP>
 <cci-list>)|
 (<FEC>
 <CCI>)
```

```
<cci-list> ::= <CCI>
 [<cci-list>]
```

Where:

<path> is as per [RFC8231] and the LSP and SRP object are also defined in [RFC8231].

When PCRpt message is used to report the label map allocations, the FEC and CCI objects MUST be present. The error handling for CCI object is as per [I-D.zhao-pce-pcep-extension-for-pce-controller]. If the FEC object is missing, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=TBD (FEC object missing).

## 7. PCEP Objects

### 7.1. OPEN Object

#### 7.1.1. PCECC Capability sub-TLV

[I-D.zhao-pce-pcep-extension-for-pce-controller] defined the PCECC-CAPABILITY TLV.

A new S-bit is defined in PCECC-CAPABILITY sub-TLV for PCECC-SR:



S (PCECC-SR-CAPABILITY - 1 bit): If set to 1 by a PCEP speaker, it indicates that the PCEP speaker is capable for PCECC-SR capability and PCE would allocate node and Adj label on this session.

## 7.2. PATH-SETUP-TYPE TLV

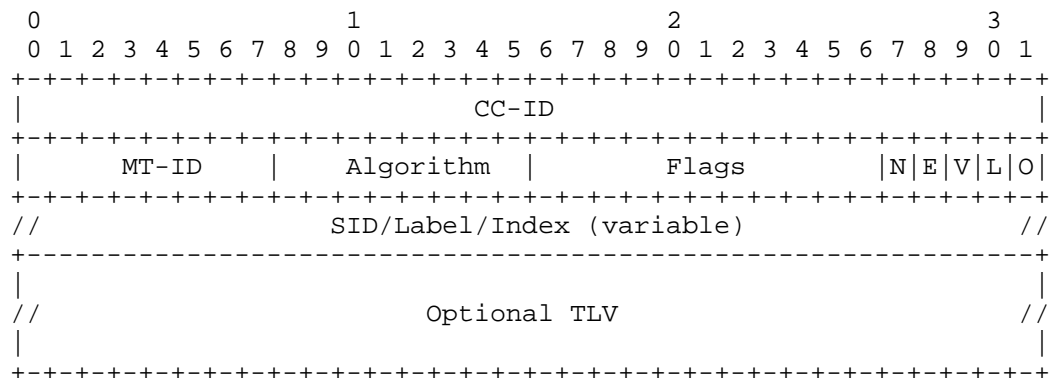
The PATH-SETUP-TYPE TLV is defined in [I-D.ietf-pce-lsp-setup-type]. PST = TBD is used when Path is setup via PCECC SR mode.

On a PCRpt/PCUpd/PCInitiate message, the PST=TBD indicates that this LSP was setup via a PCECC-SR based mechanism where either the SIDs were allocated/instructed by PCE via PCECC mechanism.

## 7.3. CCI Object

The Central Control Instructions (CCI) Object is used by the PCE to specify the forwarding instructions is defined in [I-D.zhao-pce-pcep-extension-for-pce-controller]. This document defines another object-type for SR purpose.

CCI Object-Type is TBD for SR as below -



The field CC-ID is as described in [I-D.zhao-pce-pcep-extension-for-pce-controller]. Following new fields are defined for CCI Object-Type TBD -

MT-ID: Multi-Topology ID (as defined in [RFC4915]).

Algorithm: Single octet identifying the algorithm the SID is associated with. See [I-D.ietf-ospf-segment-routing-extensions].

Flags: is used to carry any additional information pertaining to the CCI. The O bit was defined in [I-D.zhao-pce-pcep-extension-for-pce-controller], this document further defines following bits-

- \* L-Bit (Local/Global): If set, then the value/index carried by the CCI object has local significance. If not set, then the value/index carried by this object has global significance.
- \* V-Bit (Value/Index): If set, then the CCI carries an absolute value. If not set, then the CCI carries an index.
- \* E-Bit (Explicit-Null): If set, any upstream neighbor of the node that advertised the SID MUST replace the SID with the Explicit-NULL label (0 for IPv4) before forwarding the packet.
- \* N-Bit (No-PHP): If set, then the penultimate hop MUST NOT pop the SID before delivering packets to the node that advertised the SID.

SID/Label/Index: According to the V and L flags, it contains either:

A 32-bit index defining the offset in the SID/Label space advertised by this router.

A 24-bit label where the 20 rightmost bits are used for encoding the label value.

#### 7.4. FEC Object

The FEC Object is used to specify the FEC information and MAY be carried within PCInitiate or PCRpt message.

FEC Object-Class is TBD.

FEC Object-Type is 1 'IPv4 Node ID'.

|   |   |   |   |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
|---|---|---|---|

```

 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+
| IPv4 Node ID |
+-----+

```

FEC Object-Type is 2 'IPv6 Node ID'.

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+
| IPv6 Node ID (16 bytes) |
| |
+-----+-----+-----+-----+

```

FEC Object-Type is 3 'IPv4 Adjacency'.

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+
| Local IPv4 address |
+-----+-----+-----+-----+
| Remote IPv4 address |
+-----+-----+-----+-----+

```

FEC Object-Type is 4 'IPv6 Adjacency'.

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+
| Local IPv6 address (16 bytes) |
| |
+-----+-----+-----+-----+
| Remote IPv6 address (16 bytes) |
| |
+-----+-----+-----+-----+

```

FEC Object-Type is 5 'Unnumbered Adjacency with IPv4 NodeIDs'.

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+
| Local Node-ID |
+-----+-----+-----+-----+
| Local Interface ID |
+-----+-----+-----+-----+

```

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Remote Node-ID |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| Remote Interface ID |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

The FEC objects are as follows:

IPv4 Node ID: where IPv4 Node ID is specified as an IPv4 address of the Node. FEC Object-type is 1, and the Object-Length is 4 in this case.

IPv6 Node ID: where IPv6 Node ID is specified as an IPv6 address of the Node. FEC Object-type is 2, and the Object-Length is 16 in this case.

IPv4 Adjacency: where Local and Remote IPv4 address is specified as pair of IPv4 address of the adjacency. FEC Object-type is 3, and the Object-Length is 8 in this case.

IPv6 Adjacency: where Local and Remote IPv6 address is specified as pair of IPv6 address of the adjacency. FEC Object-type is 4, and the Object-Length is 32 in this case.

Unnumbered Adjacency with IPv4 NodeID: where a pair of Node ID / Interface ID tuples is used. FEC Object-type is 5, and the Object-Length is 16 in this case.

Binding ID: TBD

## 8. Security Considerations

The security considerations described in [I-D.zhao-pce-pcep-extension-for-pce-controller] apply to the extensions described in this document.

## 9. Manageability Considerations

### 9.1. Control of Function and Policy

A PCE or PCC implementation SHOULD allow to configure to enable/disable PCECC SR capability as a global configuration.

## 9.2. Information and Data Models

[RFC7420] describes the PCEP MIB, this MIB can be extended to get the PCECC SR capability status.

The PCEP YANG module [I-D.ietf-pce-pcep-yang] could be extended to enable/disable PCECC SR capability.

## 9.3. 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].

## 9.4. 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].

## 9.5. Requirements On Other Protocols

PCEP extensions defined in this document do not put new requirements on other protocols.

## 9.6. Impact On Network Operations

PCEP implementation SHOULD allow a limit to be placed on the rate of PCLabelUpd messages sent by PCE and processed by PCC. It SHOULD also allow sending a notification when a rate threshold is reached.

## 10. IANA Considerations

### 10.1. PCECC-CAPABILITY TLV

[I-D.zhao-pce-pcep-extension-for-pce-controller] defines the PCECC-CAPABILITY TLV and requests that IANA creates a registry to manage the value of the PCECC-CAPABILITY TLV's Flag field. IANA is requested to allocate a new bit in the PCECC-CAPABILITY TLV Flag Field registry, as follows:

| Bit | Description              | Reference     |
|-----|--------------------------|---------------|
| 31  | S((PCECC-SR-CAPABILITY)) | This document |

## 10.2. New Path Setup Type Registry

IANA is requested to allocate new PST Field in PATH- SETUP-TYPE TLV. The allocation policy for this new registry should be by IETF Consensus. The new registry should contain the following value:

| Value | Description                                           | Reference     |
|-------|-------------------------------------------------------|---------------|
| TBD   | Traffic engineering path is setup using PCECC-SR mode | This document |

## 10.3. PCEP Object

IANA is requested to allocate new registry for FEC PCEP object.

| Object-Class | Value | Name            | Reference                             |
|--------------|-------|-----------------|---------------------------------------|
| TBD          |       | FEC             | This document                         |
|              |       | Object-Type : 1 | IPv4 Node ID                          |
|              |       | Object-Type : 2 | IPv6 Node ID                          |
|              |       | Object-Type : 3 | IPv4 Adjacency                        |
|              |       | Object-Type : 4 | IPv6 Adjacency                        |
|              |       | Object-Type : 5 | Unnumbered Adjacency with IPv4 NodeID |

## 10.4. PCEP-Error Object

IANA is requested to allocate new error types and error values within the "PCEP-ERROR Object Error Types and Values" sub-registry of the PCEP Numbers registry for the following errors:

| Error-Type | Meaning                                   |                                  |
|------------|-------------------------------------------|----------------------------------|
| -----      | -----                                     |                                  |
| 6          | Mandatory Object missing.                 |                                  |
| 19         | Error-value = TBD :<br>Invalid operation. | FEC object missing               |
|            | Error-value = TBD :                       | SR capability was not advertised |

## 11. Acknowledgments

We would like to thank Robert Tao, Changjing Yan, Tieying Huang and Avantika for their useful comments and suggestions.

## 12. References

### 12.1. 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>>.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <<https://www.rfc-editor.org/info/rfc3630>>.
- [RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <<https://www.rfc-editor.org/info/rfc4203>>.
- [RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <<https://www.rfc-editor.org/info/rfc4915>>.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <<https://www.rfc-editor.org/info/rfc5305>>.
- [RFC5307] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 5307, DOI 10.17487/RFC5307, October 2008, <<https://www.rfc-editor.org/info/rfc5307>>.
- [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>>.
- [RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <<https://www.rfc-editor.org/info/rfc7420>>.

- [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <<https://www.rfc-editor.org/info/rfc7752>>.
- [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>>.

## 12.2. Informative References

- [RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <<https://www.rfc-editor.org/info/rfc3031>>.
- [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>>.
- [RFC7025] Otani, T., Ogaki, K., Caviglia, D., Zhang, F., and C. Margaria, "Requirements for GMPLS Applications of PCE", RFC 7025, DOI 10.17487/RFC7025, September 2013, <<https://www.rfc-editor.org/info/rfc7025>>.
- [RFC7399] Farrel, A. and D. King, "Unanswered Questions in the Path Computation Element Architecture", RFC 7399, DOI 10.17487/RFC7399, October 2014, <<https://www.rfc-editor.org/info/rfc7399>>.
- [RFC7491] King, D. and A. Farrel, "A PCE-Based Architecture for Application-Based Network Operations", RFC 7491, DOI 10.17487/RFC7491, March 2015, <<https://www.rfc-editor.org/info/rfc7491>>.

- [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>>.
- [RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <<https://www.rfc-editor.org/info/rfc8283>>.
- [I-D.ietf-teas-pcecc-use-cases]  
Zhao, Q., Li, Z., Khasanov, B., Ke, Z., Fang, L., Zhou, C., Communications, T., and A. Rachitskiy, "The Use Cases for Using PCE as the Central Controller(PCECC) of LSPs", draft-ietf-teas-pcecc-use-cases-01 (work in progress), May 2017.
- [I-D.ietf-pce-lsp-setup-type]  
Sivabalan, S., Tantsura, J., Minei, I., Varga, R., and J. Hardwick, "Conveying path setup type in PCEP messages", draft-ietf-pce-lsp-setup-type-10 (work in progress), May 2018.
- [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-07 (work in progress), March 2018.
- [I-D.zhao-pce-pcep-extension-for-pce-controller]  
Zhao, Q., Li, Z., Dhody, D., Karunanithi, S., Farrel, A., and C. Zhou, "PCEP Procedures and Protocol Extensions for Using PCE as a Central Controller (PCECC) of LSPs", draft-zhao-pce-pcep-extension-for-pce-controller-07 (work in progress), March 2018.
- [I-D.ietf-pce-segment-routing]  
Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "PCEP Extensions for Segment Routing", draft-ietf-pce-segment-routing-11 (work in progress), November 2017.

- [I-D.ietf-isis-segment-routing-extensions]  
Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and J. Tantsura, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-17 (work in progress), June 2018.
- [I-D.ietf-ospf-segment-routing-extensions]  
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", draft-ietf-ospf-segment-routing-extensions-25 (work in progress), April 2018.
- [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-03 (work in progress), April 2018.
- [I-D.dhodylee-pce-pcep-ls]  
Dhody, D., Lee, Y., and D. Ceccarelli, "PCEP Extension for Distribution of Link-State and TE Information.", draft-dhodylee-pce-pcep-ls-10 (work in progress), March 2018.
- [I-D.ietf-spring-segment-routing]  
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-15 (work in progress), January 2018.
- [I-D.ietf-spring-segment-routing-mpls]  
Bashandy, A., Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with MPLS data plane", draft-ietf-spring-segment-routing-mpls-14 (work in progress), June 2018.

#### Appendix A. Contributor Addresses

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

Email: udayasreereddy@gmail.com

Mahendra Singh Negi  
Huawei Technologies  
Divyashree Techno Park, Whitefield  
Bangalore, Karnataka 560066  
India

Email: mahendrasingh@huawei.com

Katherine Zhao  
Huawei Technologies  
2330 Central Expressway  
Santa Clara, CA 95050  
USA

Email: katherine.zhao@huawei.com

Boris Zhang  
Telus Ltd.  
Toronto  
Canada

Email: boris.zhang@telus.com

#### Authors' Addresses

Quintin Zhao  
Huawei Technologies  
125 Nagog Technology Park  
Acton, MA 01719  
USA

Email: quintin.zhao@huawei.com

Zhenbin Li  
Huawei Technologies  
Huawei Bld., No.156 Beiqing Rd.  
Beijing 100095  
China

EMail: lizhenbin@huawei.com

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

EMail: dhruv.ietf@gmail.com

Satish Karunanithi  
Huawei Technologies  
Divyashree Techno Park, Whitefield  
Bangalore, Karnataka 560066  
India

EMail: satishk@huawei.com

Adrian Farrel  
Juniper Networks, Inc  
UK

EMail: adrian@olddog.co.uk

Chao Zhou  
Cisco Systems

EMail: choa.zhou@cisco.com

PCE Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: May 29, 2021

Z. Li  
S. Peng  
Huawei Technologies  
M. Negi  
RtBrick Inc  
Q. Zhao  
Etheric Networks  
C. Zhou  
HPE  
November 25, 2020

PCEP Procedures and Protocol Extensions for Using PCE as a Central  
Controller (PCECC) for Segment Routing (SR) MPLS Segment Identifier  
(SID) Allocation and Distribution.  
draft-zhao-pce-pcep-extension-pce-controller-sr-09

## Abstract

The Path Computation Element (PCE) is a core component of Software-Defined Networking (SDN) systems. It can compute optimal paths for traffic across a network and can also update the paths to reflect changes in the network or traffic demands.

PCE was developed to derive paths for MPLS Label Switched Paths (LSPs), which are supplied to the head end of the LSP using the Path Computation Element Communication Protocol (PCEP). But SDN has a broader applicability than signaled (G)MPLS traffic-engineered (TE) networks, and the PCE may be used to determine paths in a range of use cases. PCEP has been proposed as a control protocol for use in these environments to allow the PCE to be fully enabled as a central controller.

A PCE-based Central Controller (PCECC) can simplify the processing of a distributed control plane by blending it with elements of SDN and without necessarily completely replacing it. Thus, the LSP can be calculated/set up/initiated and the label forwarding entries can also be downloaded through a centralized PCE server to each network device along the path while leveraging the existing PCE technologies as much as possible.

This document specifies the procedures and PCEP extensions when a PCE-based controller is also responsible for configuring the forwarding actions on the routers, in addition to computing the paths for packet flows in a segment routing (SR) network and telling the edge routers what instructions to attach to packets as they enter the network. PCECC is further enhanced for SR SID (Segment Identifier) allocation and distribution.

## 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 May 29, 2021.

## Copyright Notice

Copyright (c) 2020 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 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

|                                                                                              |   |
|----------------------------------------------------------------------------------------------|---|
| 1. Introduction . . . . .                                                                    | 3 |
| 1.1. Requirements Language . . . . .                                                         | 5 |
| 2. Terminology . . . . .                                                                     | 5 |
| 3. PCECC SR . . . . .                                                                        | 5 |
| 4. PCEP Requirements . . . . .                                                               | 6 |
| 5. Procedures for Using the PCE as a Central Controller (PCECC) in Segment Routing . . . . . | 6 |
| 5.1. Stateful PCE Model . . . . .                                                            | 6 |
| 5.2. New LSP Functions . . . . .                                                             | 6 |
| 5.3. PCECC Capability Advertisement . . . . .                                                | 7 |
| 5.4. PCEP session IP address and TED Router ID . . . . .                                     | 7 |
| 5.5. LSP Operations . . . . .                                                                | 8 |
| 5.5.1. PCECC Segment Routing (SR) . . . . .                                                  | 8 |
| 5.5.1.1. PCECC SR Node/Prefix SID allocation . . . . .                                       | 8 |

|             |                                                   |    |
|-------------|---------------------------------------------------|----|
| 5.5.1.2.    | PCECC SR Adjacency Label allocation . . . . .     | 10 |
| 5.5.1.3.    | Redundant PCEs . . . . .                          | 12 |
| 5.5.1.4.    | Re Delegation and Clean up . . . . .              | 12 |
| 5.5.1.5.    | Synchronization of Label Allocations . . . . .    | 13 |
| 5.5.1.6.    | PCC-Based Allocations . . . . .                   | 13 |
| 5.5.1.7.    | Binding SID . . . . .                             | 13 |
| 6.          | PCEP Messages . . . . .                           | 14 |
| 6.1.        | Central Control Instructions . . . . .            | 14 |
| 6.1.1.      | The PCInitiate Message . . . . .                  | 14 |
| 6.1.2.      | The PCRpt message . . . . .                       | 15 |
| 7.          | PCEP Objects . . . . .                            | 16 |
| 7.1.        | OPEN Object . . . . .                             | 16 |
| 7.1.1.      | PCECC Capability sub-TLV . . . . .                | 16 |
| 7.2.        | SR-TE Path Setup . . . . .                        | 17 |
| 7.3.        | CCI Object . . . . .                              | 17 |
| 7.4.        | FEC Object . . . . .                              | 19 |
| 8.          | Implementation Status . . . . .                   | 21 |
| 8.1.        | Huawei's Proof of Concept based on ONOS . . . . . | 22 |
| 9.          | Security Considerations . . . . .                 | 22 |
| 10.         | Manageability Considerations . . . . .            | 22 |
| 10.1.       | Control of Function and Policy . . . . .          | 22 |
| 10.2.       | Information and Data Models . . . . .             | 23 |
| 10.3.       | Liveness Detection and Monitoring . . . . .       | 23 |
| 10.4.       | Verify Correct Operations . . . . .               | 23 |
| 10.5.       | Requirements On Other Protocols . . . . .         | 23 |
| 10.6.       | Impact On Network Operations . . . . .            | 23 |
| 11.         | IANA Considerations . . . . .                     | 23 |
| 11.1.       | PCECC-CAPABILITY sub-TLV . . . . .                | 23 |
| 11.2.       | PCEP Object . . . . .                             | 24 |
| 11.3.       | PCEP-Error Object . . . . .                       | 24 |
| 11.4.       | CCI Object Flag Field for SR . . . . .            | 24 |
| 12.         | Acknowledgments . . . . .                         | 25 |
| 13.         | References . . . . .                              | 25 |
| 13.1.       | Normative References . . . . .                    | 25 |
| 13.2.       | Informative References . . . . .                  | 27 |
| Appendix A. | Contributor Addresses . . . . .                   | 30 |
| Authors'    | Addresses . . . . .                               | 31 |

## 1. Introduction

The Path Computation Element (PCE) [RFC4655] was developed to offload the path computation function from routers in an MPLS traffic-engineered network. Since then, the role and function of the PCE has grown to cover a number of other uses (such as GMPLS [RFC7025]) and to allow delegated control [RFC8231] and PCE-initiated use of network resources [RFC8281].

According to [RFC7399], Software-Defined Networking (SDN) refers to a separation between the control elements and the forwarding components so that software running in a centralized system, called a controller, can act to program the devices in the network to behave in specific ways. A required element in an SDN architecture is a component that plans how the network resources will be used and how the devices will be programmed. It is possible to view this component as performing specific computations to place traffic flows within the network given knowledge of the availability of network resources, how other forwarding devices are programmed, and the way that other flows are routed. This is the function and purpose of a PCE, and the way that a PCE integrates into a wider network control system (including an SDN system) is presented in [RFC7491].

In early PCE implementations, where the PCE was used to derive paths for MPLS Label Switched Paths (LSPs), paths were requested by network elements (known as Path Computation Clients (PCCs)), and the results of the path computations were supplied to network elements using the Path Computation Element Communication Protocol (PCEP) [RFC5440]. This protocol was later extended to allow a PCE to send unsolicited requests to the network for LSP establishment [RFC8281].

[RFC8283] introduces the architecture for PCE as a central controller as an extension of the architecture described in [RFC4655] and assumes the continued use of PCEP as the protocol used between PCE and PCC. [RFC8283] further examines the motivations and applicability for PCEP as a Southbound Interface (SBI), and introduces the implications for the protocol. [I-D.ietf-teas-pcecc-use-cases] describes the use cases for the PCE-based Central Controller (PCECC) architecture.

[I-D.ietf-pce-pcep-extension-for-pce-controller] specify the procedures and PCEP extensions for using the PCE as the central controller for static LSPs, where LSPs can be provisioned as explicit label instructions at each hop on the end-to-end path.

Segment Routing (SR) technology leverages the source routing and tunneling paradigms. A source node can choose a path without relying on hop-by-hop signaling protocols such as LDP or RSVP-TE. Each path is specified as a set of "segments" advertised by link-state routing protocols (IS-IS or OSPF). [RFC8402] provides an introduction to SR architecture. The corresponding IS-IS and OSPF extensions are specified in [RFC8667] and [RFC8665], respectively. It relies on a series of forwarding instructions being placed in the header of a packet. The segment routing architecture supports operations that can be used to steer packet flows in a network, thus providing a form of traffic engineering. [RFC8664] specify the SR specific PCEP extensions.

PCECC may further use PCEP for SR SID (Segment Identifier) allocation and distribution on the SR nodes with some benefits.

This document specifies the procedures and PCEP extensions when a PCE-based controller is also responsible for configuring the forwarding actions on the routers (SR SID allocation and distribution in this case), in addition to computing the paths for packet flows in a segment routing network and telling the edge routers what instructions to attach to packets as they enter the network.

Only SR using MPLS dataplane (SR-MPLS) is in the scope of this document. Refer [I-D.dhody-pce-pcep-extension-pce-controller-srv6] for use of PCECC technique for SR in IPv6 (SRv6) dataplane.

### 1.1. 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.

## 2. Terminology

Terminologies used in this document is the same as described in the draft [RFC8283] and [I-D.ietf-teas-pcecc-use-cases].

## 3. PCECC SR

[RFC8664] specifies extensions to PCEP that allow a stateful PCE to compute, update, or initiate SR-TE paths. An ingress node of an SR-TE path appends all outgoing packets with a list of MPLS labels (SIDs). This is encoded in SR-ERO subobject, capable of carrying a label (SID) as well as the identity of the node/adjacency label (SID).

The notion of segment and SID is defined in [RFC8402], which fits the MPLS architecture [RFC3031] as the label which is managed by a local allocation process of LSR (similarly to other MPLS signaling protocols) [RFC8660]. The SR information such as node/adjacency label (SID) is flooded via IGP as specified in [RFC8667] and [RFC8665].

As per [RFC8283], PCE as a central controller can allocate and provision the node/prefix/adjacency label (SID) via PCEP.

The rest of the processing is similar to existing stateful PCE with SR mechanism.

For the purpose of this document, it is assumed that the label range to be used by a PCE is set on both PCEP peers. Further, a global label range is assumed to be set on all PCEP peers in the SR domain. This document also allows a case where the label space is maintained by PCC itself, and the labels are allocated by the PCC, in this case, the PCE should request the allocation from PCC as described in Section 5.5.1.6.

#### 4. PCEP Requirements

Following key requirements for PCECC-SR should be considered when designing the PCECC-based solution:

- o A PCEP speaker supporting this draft needs to have the capability to advertise its PCECC-SR capability to its peers.
- o PCEP procedures need to allow for PCC-based label/SID allocations.
- o PCEP procedures need means to update (or clean up) the label-map entry to the PCC.
- o PCEP procedures need to provide a mean to synchronize the SR labels allocations between the PCE to the PCC via PCEP messages.

#### 5. Procedures for Using the PCE as a Central Controller (PCECC) in Segment Routing

##### 5.1. Stateful PCE Model

Active stateful PCE is described in [RFC8231]. PCE as a Central Controller (PCECC) reuses the existing active stateful PCE mechanism as much as possible to control the LSPs.

##### 5.2. New LSP Functions

Several new functions are required in PCEP to support PCECC as described in [I-D.ietf-pce-pcep-extension-for-pce-controller]. This document reuses the existing messages to support PCECC-SR.

The PCEP messages PCRpt, PCInitiate, PCUpd are used to send LSP Reports, LSP setup, and LSP update respectively. The extended PCInitiate message described in [I-D.ietf-pce-pcep-extension-for-pce-controller] is used to download or clean up central controller's instructions (CCIs) (SR SID in the scope of this document). The extended PCRpt message described in [I-D.ietf-pce-pcep-extension-for-pce-controller] is also used to report the CCIs (SR SIDs) from PCC to PCE.

[I-D.ietf-pce-pcep-extension-for-pce-controller] specify an object called CCI for the encoding of the central controller's instructions. This document extends the CCI by defining a new object-type for segment routing. The PCEP messages are extended in this document to handle the PCECC operations for SR.

### 5.3. PCECC Capability Advertisement

During PCEP Initialization Phase, PCEP Speakers (PCE or PCC) advertise their support of PCECC extensions. A PCEP Speaker includes the "PCECC Capability" sub-TLV, described in [I-D.ietf-pce-pcep-extension-for-pce-controller].

A new S-bit is added in the PCECC-CAPABILITY sub-TLV to indicate support for PCECC-SR. A PCC MUST set the S-bit in the PCECC-CAPABILITY sub-TLV and include the SR-PCE-CAPABILITY sub-TLV ([RFC8664]) in the OPEN Object (inside the PATH-SETUP-TYPE-CAPABILITY TLV) to support the PCECC SR extensions defined in this document. If the S-bit is set in the PCECC-CAPABILITY sub-TLV and the SR-PCE-CAPABILITY sub-TLV is not advertised in the OPEN Object, PCE SHOULD send a PCErr message with Error-Type=19 (Invalid Operation) and Error-value=TBd4 (SR capability was not advertised) and terminate the session.

The rest of the processing is as per [I-D.ietf-pce-pcep-extension-for-pce-controller].

### 5.4. PCEP session IP address and TED Router ID

A PCE may construct its Traffic Engineering Database (TED) by participating in the IGP ([RFC3630] and [RFC5305] for MPLS-TE; [RFC4203] and [RFC5307] for GMPLS). An alternative is offered by BGP-LS [RFC7752] and [I-D.dhodylee-pce-pcep-ls].

A PCEP [RFC5440] speaker could use any local IP address while creating a TCP session. It is important to link the session IP address with the Router ID in TED for successful PCECC operations.

During PCEP Initialization Phase, the PCC SHOULD advertise the TE mapping information by including the "Node Attributes TLV" [I-D.dhodylee-pce-pcep-ls] with "IPv4/IPv6 Router-ID of Local Node", in the OPEN Object for this purpose. [RFC7752] describes the usage as auxiliary Router-IDs that the IGP might be using, e.g., for TE purposes. If there are more than one auxiliary Router-ID of a given type, then multiple TLVs are used to encode them.

If "IPv4/IPv6 Router-ID" TLV is not present, the TCP session IP address is directly used for mapping purpose.

## 5.5. LSP Operations

[RFC8664] specify the PCEP extension to allow a stateful PCE to compute and initiate SR-TE paths, as well as a PCC to request a path subject to certain constraint(s) and optimization criteria in SR networks.

The Path Setup Type for segment routing (PST=1) is used on the PCEP session with the Ingress as per [RFC8664].

### 5.5.1. PCECC Segment Routing (SR)

Segment Routing (SR) as described in [RFC8402] depends on "segments" that are advertised by Interior Gateway Protocols (IGPs). The SR-node allocates and advertises the SID (node, adj, etc) and flood them via the IGP. This document proposes a new mechanism where PCE allocates the SID (label/index/SID) centrally and uses PCEP to advertise them. In some deployments, PCE (and PCEP) are better suited than IGP because of the centralized nature of PCE and direct TCP based PCEP sessions to the node.

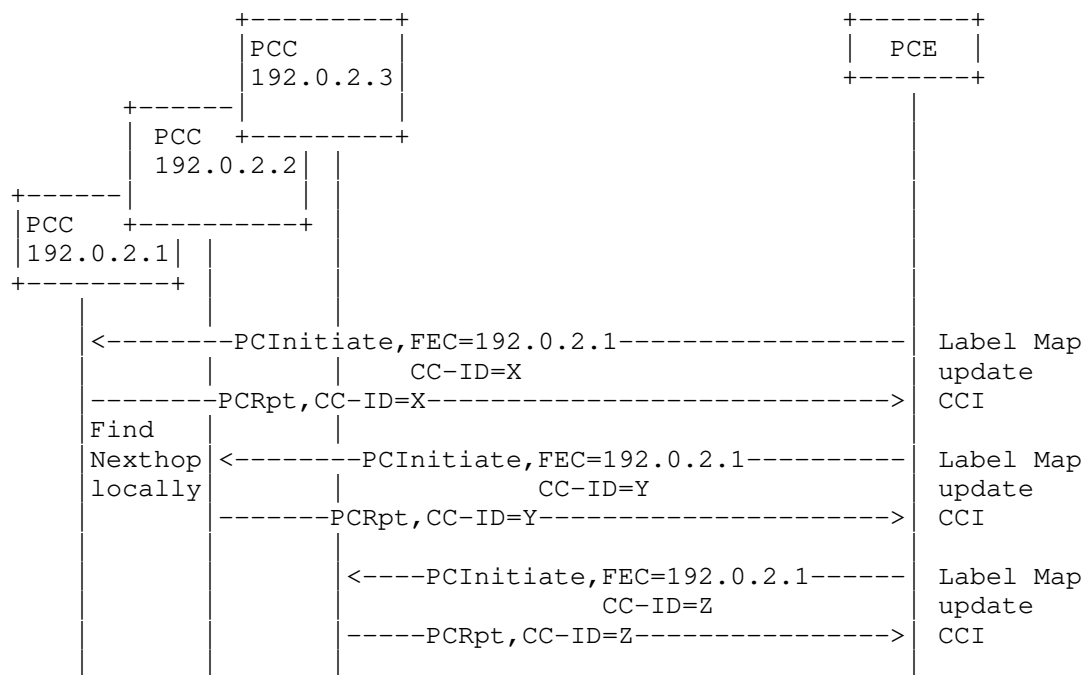
#### 5.5.1.1. PCECC SR Node/Prefix SID allocation

Each node (PCC) is allocated a node-SID by the PCECC. The PCECC sends PCInitiate message to update the label map of each node to all the nodes in the domain. The TE router ID is determined from the TED or from "IPv4/IPv6 Router-ID" Sub-TLV [I-D.dhodylee-pce-pcep-ls], in the OPEN Object Section 5.4.

It is RECOMMENDED that PCEP session with PCECC-SR capability to use a different session IP address during TCP session establishment than the node Router ID in TEDB, to make sure that the PCEP session does not get impacted by the SR Node/Prefix Label maps (Section 5.4).

If a node (PCC) receives a PCInitiate message with a CCI encoding a SID, out of the range set aside for the SR Global Block (SRGB), it MUST send a PCErr message with Error-type=TBD (PCECC failure) and Error-value=TBD (Label out of range) (defined in [I-D.ietf-pce-pcep-extension-for-pce-controller]) and MUST include the SRP object to specify the error is for the corresponding central control instruction via the PCInitiate message.

On receiving the label map, each node (PCC) uses the local routing information to determine the next-hop and download the label forwarding instructions accordingly. The PCInitiate message in this case does not use the LSP object but uses a new FEC object defined in this document.

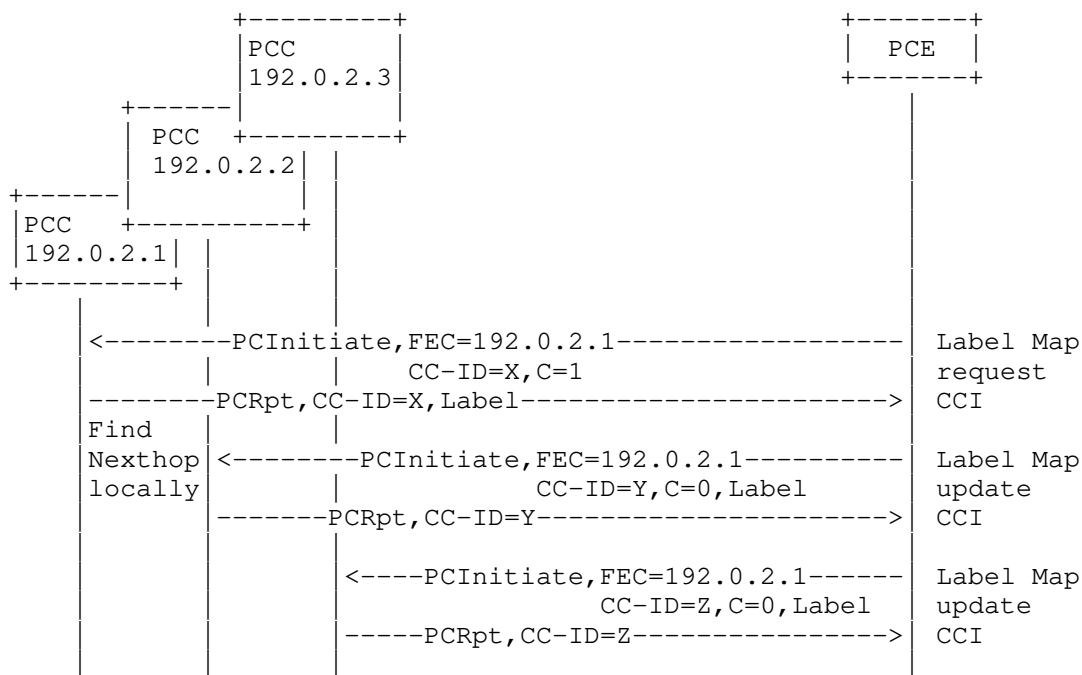


The forwarding behavior and the end result is similar to IGP based "Node-SID" in SR. Thus, from anywhere in the domain, it enforces the ECMP-aware shortest-path forwarding of the packet towards the related node as per [RFC8402].

PCE relies on the Node/Prefix Label clean up using the same PCInitiate message as per [RFC8281].

The above example Figure 1 depicts the FEC and PCEP speakers that uses IPv4 address. Similarly an IPv6 address (such as 2001:DB8::1) can be used during PCEP session establishment in the FEC object as described in this specification.

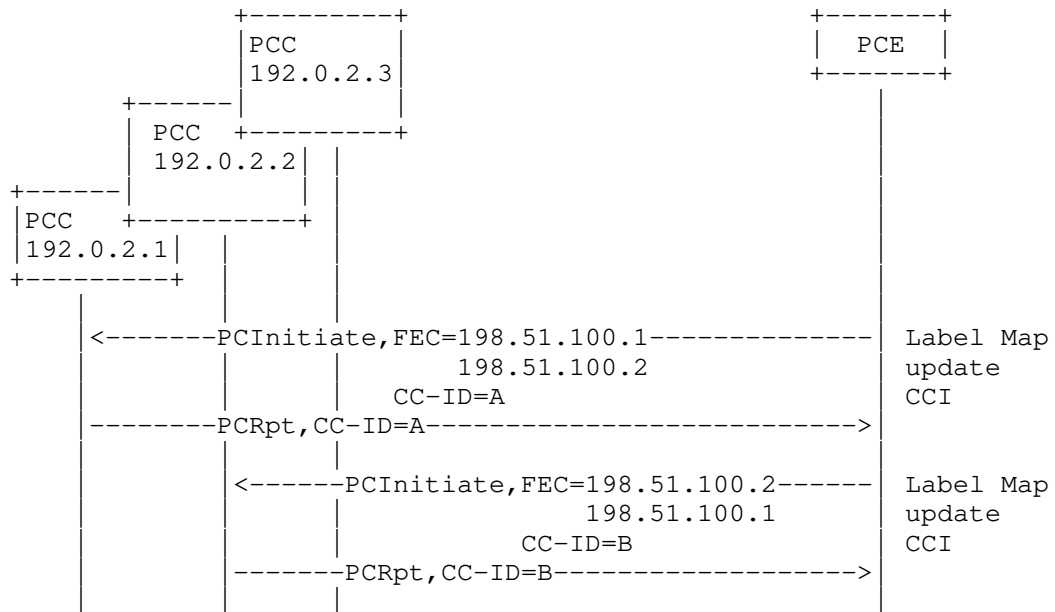
In the case where the label/SID allocation is made by the PCC itself (see Section 5.5.1.6), the PCE could request an allocation to be made by the PCC, and where the PCC would send a PCRpt with the allocated label/SID encoded in the CC-ID object as shown in Figure 2.



It should be noted that in this example, the request is made to the node 192.0.2.1 with C bit set in the CCI object to indicate that the allocation needs to be done by this PCC and it responds with the allocated label/SID to the PCE. The PCE would further inform the other PCCs in the network about the label-map allocation without setting the C bit.

#### 5.5.1.2. PCECC SR Adjacency Label allocation

For PCECC-SR, apart from node-SID, Adj-SID is used where each adjacency is allocated an Adj-SID by the PCECC. The PCECC sends the PCInitiate message to update the label map of each adjacency to the corresponding nodes in the domain. Each node (PCC) download the label forwarding instructions accordingly. Similar to SR Node/Prefix Label allocation, the PCInitiate message in this case does not use the LSP object but uses the new FEC object defined in this document.



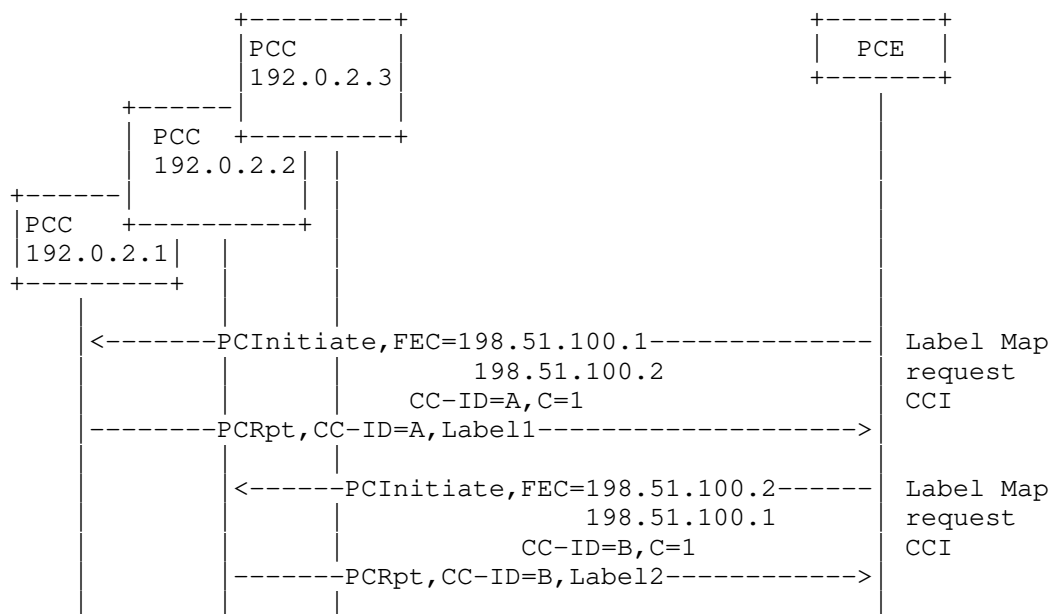
The forwarding behavior and the end result is similar to IGP based "Adj-SID" in SR.

PCE relies on the Adj label clean up using the same PCInitiate message as per [RFC8281].

The above example Figure 3 depicts FEC object and PCEP speakers that uses an IPv4 address. Similarly an IPv6 address (such as 2001:DB8::1, 2001:DB8::2) can be used during the PCEP session establishment in the FEC object as described in this specification.

The handling of adjacencies on the LAN subnetworks is specified in [RFC8402]. PCECC MUST assign Adj-SID for every pair of routers in the LAN. The rest of the protocol mechanism remains the same.

In the case where the label/SID map allocation is made by the PCC itself (see Section 5.5.1.6), the PCE could request an allocation to be made by the PCC, and where the PCC would send a PCRpt with the allocated label/SID encoded in the CC-ID object as shown in Figure 4.



In this example, the request is made to the node 192.0.2.1 with the C bit set in the CCI object to indicate that the allocation needs to be done by this PCC for the adjacency (198.51.100.1 - 198.51.100.2) and it responds with the allocated label/SID to the PCE. Similarly, another request is made to the node 192.0.2.2 with the C bit set in the CCI object to indicate that the allocation needs to be done by this PCC for the adjacency (198.51.100.2 - 198.51.100.1).

#### 5.5.1.3. Redundant PCEs

[I-D.litkowski-pce-state-sync] describes the synchronization mechanism between the stateful PCEs. The SR SIDs allocated by a PCE MUST also be synchronized among PCEs for PCECC SR state synchronization. Note that the SR SIDs are independent of the SR-TE LSPs, and remains intact till any topology change. The redundant PCEs MUST have a common view of all SR SIDs allocated in the domain.

#### 5.5.1.4. Re Delegation and Clean up

[I-D.ietf-pce-pcep-extension-for-pce-controller] describes the action needed for CCIs for the static LSPs on a terminated session. Same holds true for the CCI for SR SID as well.

#### 5.5.1.5. Synchronization of Label Allocations

[I-D.ietf-pce-pcep-extension-for-pce-controller] describes the synchronization of Central Controller's Instructions (CCI) via LSP state synchronization as described in [RFC8231] and [RFC8232]. Same procedures are applied for the CCI for SR SID as well.

#### 5.5.1.6. PCC-Based Allocations

The PCE can request the PCC to allocate the label/SID using the PCInitiate message. The C flag in the CCI object is set to 1 to indicate that the allocation needs to be done by the PCC. The PCC would allocate the SID/Label/Index and would report to the PCE using the PCRpt message.

If the value of the SID/Label/Index is 0 and the C flag is set to 1, it indicates that the PCE is requesting the allocation to be done by the PCC. If the SID/Label/Index is 'n' and the C flag is set to 1 in the CCI object, it indicates that the PCE requests a specific value 'n' for the SID/Label/Index. If the allocation is successful, the PCC should report via PCRpt message with the CCI object. Else, it MUST send a PCErr message with Error-Type = TBD ("PCECC failure") and Error Value = TBD ("Invalid CCI") (defined in [I-D.ietf-pce-pcep-extension-for-pce-controller]). If the value of the SID/Label/Index in the CCI object is valid, but the PCC is unable to allocate it, it MUST send a PCErr message with Error-Type = TBD ("PCECC failure") and Error Value = TBD ("Unable to allocate the specified CCI") (defined in [I-D.ietf-pce-pcep-extension-for-pce-controller]).

If the PCC wishes to withdraw or modify the previously assigned label/SID, it MUST send a PCRpt message without any SID/Label/Index or with the SID/Label/Index containing the new value respectively in the CCI object. The PCE would further trigger the removal of the central controller instruction as per this document.

#### 5.5.1.7. Binding SID

A PCECC can allocate and provision the node/prefix/adjacency label (SID) via PCEP. Another SID called binding SID is described in [I-D.ietf-pce-binding-label-sid], the PCECC mechanism can also be used to allocate the binding SID.

A procedure for binding label/SID allocation is described in [I-D.ietf-pce-pcep-extension-for-pce-controller] and is applicable for all path setup types (including SR paths).

## 6. PCEP Messages

As defined in [RFC5440], a PCEP message consists of a common header followed by a variable-length body made of a set of objects that can be either mandatory or optional. An object is said to be mandatory in a PCEP message when the object must be included for the message to be considered valid. For each PCEP message type, a set of rules is defined that specify the set of objects that the message can carry. An implementation MUST form the PCEP messages using the object ordering specified in this document.

### 6.1. Central Control Instructions

#### 6.1.1. The PCInitiate Message

The PCInitiate message defined in [RFC8281] and extended in [I-D.ietf-pce-pcep-extension-for-pce-controller] is further extended to support SR based central control instructions.

The format of the extended PCInitiate message is as follows:

```
<PCInitiate Message> ::= <Common Header>
 <PCE-initiated-lsp-list>
```

Where:

<Common Header> is defined in [RFC5440]

```
<PCE-initiated-lsp-list> ::= <PCE-initiated-lsp-request>
 [<PCE-initiated-lsp-list>]
```

```
<PCE-initiated-lsp-request> ::=
 (<PCE-initiated-lsp-instantiation> |
 <PCE-initiated-lsp-deletion> |
 <PCE-initiated-lsp-central-control>)
```

```
<PCE-initiated-lsp-central-control> ::= <SRP>
 (<LSP>
 <cci-list>) |
 (<FEC>
 <CCI>)
```

```
<cci-list> ::= <CCI>
 [<cci-list>]
```

Where:

<PCE-initiated-lsp-instantiation> and  
<PCE-initiated-lsp-deletion> are as per  
[RFC8281].

The LSP and SRP object is defined in [RFC8231].

When the PCInitiate message is used to distribute SR SIDs, the SRP, the FEC and the CCI objects MUST be present. The error handling for missing SRP or CCI object is as per [I-D.ietf-pce-pcep-extension-for-pce-controller]. If the FEC object is missing, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=TBD5 (FEC object missing).

To clean up, the R (remove) bit in the SRP object and the corresponding FEC and the CCI object are included.

#### 6.1.2. The PCRpt message

The PCRpt message can be used to report the SR central controller instructions received from the PCECC during the state synchronization phase or as an acknowledgment to the PCInitiate message.

The format of the PCRpt message is as follows:

```

<PCRpt Message> ::= <Common Header>
 <state-report-list>

```

Where:

```

<state-report-list> ::= <state-report> [<state-report-list>]

```

```

<state-report> ::= (<lsp-state-report> |
 <central-control-report>)

```

```

<lsp-state-report> ::= [<SRP>]
 <LSP>
 <path>

```

```

<central-control-report> ::= [<SRP>]
 (<LSP>
 <cci-list>) |
 (<FEC>
 <CCI>)

```

```

<cci-list> ::= <CCI>
 [<cci-list>]

```

Where:

<path> is as per [RFC8231] and the LSP and SRP object are also defined in [RFC8231].

When PCRpt message is used to report the label map allocations, the FEC and CCI objects MUST be present. The error handling for the missing CCI object is as per [I-D.ietf-pce-pcep-extension-for-pce-controller]. If the FEC object is missing, the receiving PCE MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=TBD5 (FEC object missing).

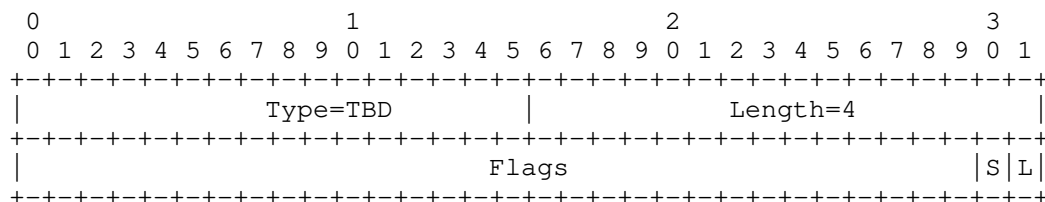
## 7. PCEP Objects

### 7.1. OPEN Object

#### 7.1.1. PCECC Capability sub-TLV

[I-D.ietf-pce-pcep-extension-for-pce-controller] defined the PCECC-CAPABILITY sub-TLV.

A new S-bit is defined in PCECC-CAPABILITY sub-TLV for PCECC-SR:



[Editor's Note - The above figure is included for ease of the reader but should be removed before publication.]

S (PCECC-SR-CAPABILITY - 1 bit - TBD1): If set to 1 by a PCEP speaker, it indicates that the PCEP speaker is capable of PCECC-SR capability and the PCE allocates the Node and Adj label/SID on this session.

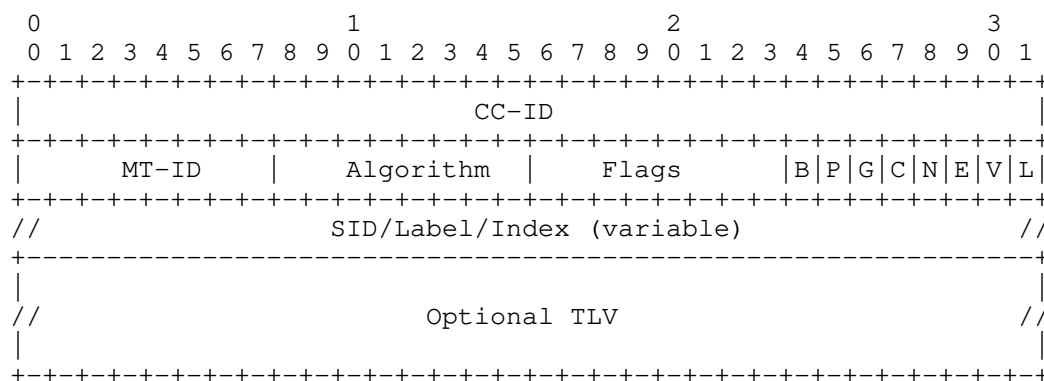
## 7.2. SR-TE Path Setup

The PATH-SETUP-TYPE TLV is defined in [RFC8408]. A PST value of 1 is used when Path is setup via SR mode as per [RFC8664]. The procedure for SR-TE path setup as specified in [RFC8664] remains unchanged.

## 7.3. CCI Object

The Central Control Instructions (CCI) Object used by the PCE to specify the controller instructions is defined in [I-D.ietf-pce-pcep-extension-for-pce-controller]. This document defines another object-type for SR-MPLS purpose.

CCI Object-Type is TBD6 for SR-MPLS as below -



The field CC-ID is as described in [I-D.ietf-pce-pcep-extension-for-pce-controller]. Following new fields are defined for CCI Object-Type TBD6 -

MT-ID: Multi-Topology ID (as defined in [RFC4915]).

Algorithm: Single octet identifying the algorithm the SID is associated with. See [RFC8665].

Flags: is used to carry any additional information pertaining to the CCI. The following bits are defined -

- \* L-Bit (Local/Global): If set, then the value/index carried by the CCI object has local significance. If not set, then the value/index carried by this object has global significance.
- \* V-Bit (Value/Index): If set, then the CCI carries an absolute value. If not set, then the CCI carries an index.
- \* E-Bit (Explicit-Null): If set, any upstream neighbor of the node that advertised the SID MUST replace the SID with the Explicit-NULL label (0 for IPv4) before forwarding the packet.
- \* N-Bit (No-PHP): If set, then the penultimate hop MUST NOT pop the SID before delivering packets to the node that advertised the SID.
- \* C-Bit (PCC Allocation): If the bit is set to 1, it indicates that the allocation needs to be done by the PCC for this central controller instruction. A PCE set this bit to request the PCC to make an allocation from its SR label/ID space. A PCC would set this bit to indicate that it has allocated the CC-ID and report it to the PCE.
- \* Following bits are applicable when the SID represents an Adj-SID only, it MUST be ignored for others -
  - + G-Bit (Group): When set, the G-Flag indicates that the Adj-SID refers to a group of adjacencies (and therefore MAY be assigned to other adjacencies as well).
  - + P-Bit (Persistent): When set, the P-Flag indicates that the Adj-SID is persistently allocated, i.e., the Adj-SID value remains consistent across router restart and/or interface flap.
  - + B-Bit (Backup): If set, the Adj-SID refers to an adjacency that is eligible for protection (e.g., using IP Fast Reroute

or MPLS-FRR (MPLS-Fast Reroute) as described in Section 2.1 of [RFC8402].

- + All unassigned bits MUST be set to zero at transmission and ignored at receipt.

SID/Label/Index: According to the V and L flags, it contains either:

A 32-bit index defining the offset in the SID/Label space advertised by this router.

A 24-bit label where the 20 rightmost bits are used for encoding the label value.

#### 7.4. FEC Object

The FEC Object is used to specify the FEC information and MAY be carried within PCInitiate or PCRpt message.

FEC Object-Class is TBD3.

FEC Object-Type is 1 'IPv4 Node ID'.

```

0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| IPv4 Node ID |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

FEC Object-Type is 2 'IPv6 Node ID'.

```

0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| IPv6 Node ID (16 bytes) |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

FEC Object-Type is 3 'IPv4 Adjacency'.

```

0 1 2 3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| Local IPv4 address |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

```

|----- Remote IPv4 address -----|
+-----+

```

FEC Object-Type is 4 'IPv6 Adjacency'.

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+
|
// Local IPv6 address (16 bytes) //
|
+-----+
|
// Remote IPv6 address (16 bytes) //
|
+-----+

```

FEC Object-Type is 5 'Unnumbered Adjacency with IPv4 NodeIDs'.

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+
| Local Node-ID |
+-----+
| Local Interface ID |
+-----+
| Remote Node-ID |
+-----+
| Remote Interface ID |
+-----+

```

FEC Object-Type is 6 'Linklocal IPv6 Adjacency'.

```

 0 1 2 3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+
// Local IPv6 address (16 octets) //
+-----+
| Local Interface ID |
+-----+
// Remote IPv6 address (16 octets) //
+-----+
| Remote Interface ID |
+-----+

```

The FEC objects are as follows:

IPv4 Node ID: where IPv4 Node ID is specified as an IPv4 address of the Node. FEC Object-type is 1, and the Object-Length is 4 in this case.

IPv6 Node ID: where IPv6 Node ID is specified as an IPv6 address of the Node. FEC Object-type is 2, and the Object-Length is 16 in this case.

IPv4 Adjacency: where Local and Remote IPv4 address is specified as pair of IPv4 addresses of the adjacency. FEC Object-type is 3, and the Object-Length is 8 in this case.

IPv6 Adjacency: where Local and Remote IPv6 address is specified as pair of IPv6 addresses of the adjacency. FEC Object-type is 4, and the Object-Length is 32 in this case.

Unnumbered Adjacency with IPv4 NodeID: where a pair of Node ID / Interface ID tuple is used. FEC Object-type is 5, and the Object-Length is 16 in this case.

Linklocal IPv6 Adjacency: where a pair of (global IPv6 address, interface ID) tuple is used. FEC object-type is 6, and the Object-Length is 40 in this case.

## 8. Implementation Status

[Note to the RFC Editor - remove this section before publication, as well as remove the reference to RFC 7942.]

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in [RFC7942]. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to [RFC7942], "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature.

It is up to the individual working groups to use this information as they see fit".

#### 8.1. Huawei's Proof of Concept based on ONOS

The PCE function was developed in the ONOS open source platform. This extension was implemented on a private version as a proof of concept for PCECC.

- o Organization: Huawei
- o Implementation: Huawei's PoC based on ONOS
- o Description: PCEP as a southbound plugin was added to ONOS. To support PCECC-SR, an earlier version of this I-D was implemented. Refer <https://wiki.onosproject.org/display/ONOS/PCEP+Protocol>
- o Maturity Level: Prototype
- o Coverage: Partial
- o Contact: satishk@huawei.com

#### 9. Security Considerations

The security considerations described in [I-D.ietf-pce-pcep-extension-for-pce-controller] apply to the extensions described in this document.

As per [RFC8231], it is RECOMMENDED that these PCEP extensions only be activated on authenticated and encrypted sessions across PCEs and PCCs belonging to the same administrative authority, using Transport Layer Security (TLS) [RFC8253] as per the recommendations and best current practices in [RFC7525] (unless explicitly set aside in [RFC8253]).

#### 10. Manageability Considerations

##### 10.1. Control of Function and Policy

A PCE or PCC implementation SHOULD allow to configure to enable/disable PCECC SR capability as a global configuration. The implementation SHOULD also allow setting the local IP address used by the PCEP session.

## 10.2. Information and Data Models

[RFC7420] describes the PCEP MIB, this MIB can be extended to get the PCECC SR capability status.

The PCEP YANG module [I-D.ietf-pce-pcep-yang] could be extended to enable/disable PCECC SR capability.

## 10.3. 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].

## 10.4. 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].

## 10.5. Requirements On Other Protocols

PCEP extensions defined in this document do not put new requirements on other protocols.

## 10.6. Impact On Network Operations

PCEP implementation SHOULD allow a limit to be placed on the rate of PCLabelUpd messages sent by PCE and processed by PCC. It SHOULD also allow sending a notification when a rate threshold is reached.

## 11. IANA Considerations

### 11.1. PCECC-CAPABILITY sub-TLV

[I-D.ietf-pce-pcep-extension-for-pce-controller] defines the PCECC-CAPABILITY sub-TLV and requests that IANA to create a new sub-registry to manage the value of the PCECC-CAPABILITY sub-TLV's Flag field.

IANA is requested to allocate a new bit in the PCECC-CAPABILITY sub-TLV Flag Field sub-registry, as follows:

| Bit  | Description | Reference     |
|------|-------------|---------------|
| TBD1 | SR          | This document |

## 11.2. PCEP Object

IANA is requested to allocate new code-points for the new FEC object and a new Object-Type for CCI object in "PCEP Objects" sub-registry as follows:

| Object-Class Value | Name | Object-Type                              | Reference     |
|--------------------|------|------------------------------------------|---------------|
| TBD3               | FEC  | 1: IPv4 Node ID                          | This document |
|                    |      | 2: IPv6 Node ID                          | This document |
|                    |      | 3: IPv4 Adjacency                        | This document |
|                    |      | 4: IPv6 Adjacency                        | This document |
|                    |      | 5: Unnumbered Adjacency with IPv4 NodeID | This document |
|                    |      | 6: Linklocal IPv6 Adjacency              | This document |
| TBD                | CCI  | TBD6: SR-MPLS                            | This document |

## 11.3. PCEP-Error Object

IANA is requested to allocate a new error-value within the "PCEP-ERROR Object Error Types and Values" sub-registry of the PCEP Numbers registry for the following errors:

| Error-Type | Meaning                   |                                  |
|------------|---------------------------|----------------------------------|
| 6          | Mandatory Object missing. |                                  |
| 19         | Error-value = TBD5 :      | FEC object missing               |
|            | Invalid operation.        |                                  |
|            | Error-value = TBD4 :      | SR capability was not advertised |

## 11.4. CCI Object Flag Field for SR

IANA is requested to create a new sub-registry to manage the Flag field of the CCI Object-Type=TBD6 for SR called "CCI Object Flag Field for SR". New values are to be assigned by Standards Action [RFC8126]. Each bit should be tracked with the following qualities:

- o Bit number (counting from bit 0 as the most significant bit)
- o Capability description
- o Defining RFC

Following bits are defined for the CCI Object flag field for SR in this document as follows:

| Bit | Description            | Reference     |
|-----|------------------------|---------------|
| 0-7 | Unassigned             | This document |
| 8   | B-Bit - Backup         | This document |
| 9   | P-Bit - Persistent     | This document |
| 10  | G-Bit - Group          | This document |
| 11  | C-Bit - PCC Allocation | This document |
| 12  | N-Bit - No-PHP         | This document |
| 13  | E-Bit - Explicit-Null  | This document |
| 14  | V-Bit - Value/Index    | This document |
| 15  | L-Bit - Local/Global   | This document |

## 12. Acknowledgments

We would like to thank Robert Tao, Changjing Yan, Tieying Huang, Avantika, and Aijun Wang for their useful comments and suggestions.

## 13. References

### 13.1. 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>>.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <<https://www.rfc-editor.org/info/rfc3630>>.
- [RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <<https://www.rfc-editor.org/info/rfc4203>>.
- [RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <<https://www.rfc-editor.org/info/rfc4915>>.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <<https://www.rfc-editor.org/info/rfc5305>>.
- [RFC5307] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 5307, DOI 10.17487/RFC5307, October 2008, <<https://www.rfc-editor.org/info/rfc5307>>.

- [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>>.
- [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <<https://www.rfc-editor.org/info/rfc7752>>.
- [RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", BCP 205, RFC 7942, DOI 10.17487/RFC7942, July 2016, <<https://www.rfc-editor.org/info/rfc7942>>.
- [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>>.
- [RFC8408] Sivabalan, S., Tantsura, J., Minei, I., Varga, R., and J. Hardwick, "Conveying Path Setup Type in PCE Communication Protocol (PCEP) Messages", RFC 8408, DOI 10.17487/RFC8408, July 2018, <<https://www.rfc-editor.org/info/rfc8408>>.
- [RFC8664] Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing", RFC 8664, DOI 10.17487/RFC8664, December 2019, <<https://www.rfc-editor.org/info/rfc8664>>.

[I-D.ietf-pce-pcep-extension-for-pce-controller]  
Li, Z., Peng, S., Negi, M., Zhao, Q., and C. Zhou, "PCEP Procedures and Protocol Extensions for Using PCE as a Central Controller (PCECC) of LSPs", draft-ietf-pce-pcep-extension-for-pce-controller-08 (work in progress), November 2020.

### 13.2. Informative References

- [RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <<https://www.rfc-editor.org/info/rfc3031>>.
- [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>>.
- [RFC7025] Otani, T., Ogaki, K., Caviglia, D., Zhang, F., and C. Margaria, "Requirements for GMPLS Applications of PCE", RFC 7025, DOI 10.17487/RFC7025, September 2013, <<https://www.rfc-editor.org/info/rfc7025>>.
- [RFC7399] Farrel, A. and D. King, "Unanswered Questions in the Path Computation Element Architecture", RFC 7399, DOI 10.17487/RFC7399, October 2014, <<https://www.rfc-editor.org/info/rfc7399>>.
- [RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <<https://www.rfc-editor.org/info/rfc7420>>.
- [RFC7491] King, D. and A. Farrel, "A PCE-Based Architecture for Application-Based Network Operations", RFC 7491, DOI 10.17487/RFC7491, March 2015, <<https://www.rfc-editor.org/info/rfc7491>>.
- [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>>.

- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <<https://www.rfc-editor.org/info/rfc8126>>.
- [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>>.
- [RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <<https://www.rfc-editor.org/info/rfc8283>>.
- [RFC8402] Filssils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.
- [RFC8660] Bashandy, A., Ed., Filssils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <<https://www.rfc-editor.org/info/rfc8660>>.
- [RFC8665] Psenak, P., Ed., Previdi, S., Ed., Filssils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", RFC 8665, DOI 10.17487/RFC8665, December 2019, <<https://www.rfc-editor.org/info/rfc8665>>.
- [RFC8667] Previdi, S., Ed., Ginsberg, L., Ed., Filssils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", RFC 8667, DOI 10.17487/RFC8667, December 2019, <<https://www.rfc-editor.org/info/rfc8667>>.
- [I-D.ietf-teas-pcecc-use-cases]  
Li, Z., Khasanov, B., Dhody, D., Zhao, Q., Ke, Z., Fang, L., Zhou, C., Communications, T., Rachitskiy, A., and A. Gulida, "The Use Cases for Path Computation Element (PCE) as a Central Controller (PCECC).", draft-ietf-teas-pcecc-use-cases-06 (work in progress), September 2020.

[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-15 (work in progress), October 2020.

[I-D.ietf-pce-binding-label-sid]

Sivabalan, S., Filsfils, C., Tantsura, J., Hardwick, J., Previdi, S., and C. Li, "Carrying Binding Label/Segment-ID in PCE-based Networks.", draft-ietf-pce-binding-label-sid-05 (work in progress), October 2020.

[I-D.litkowski-pce-state-sync]

Litkowski, S., Sivabalan, S., Li, C., and H. Zheng, "Inter Stateful Path Computation Element (PCE) Communication Procedures.", draft-litkowski-pce-state-sync-09 (work in progress), November 2020.

[I-D.dhodylee-pce-pcep-ls]

Dhody, D., Peng, S., Lee, Y., Ceccarelli, D., Wang, A., and G. Mishra, "PCEP extensions for Distribution of Link-State and TE Information", draft-dhodylee-pce-pcep-ls-19 (work in progress), November 2020.

[I-D.dhody-pce-pcep-extension-pce-controller-srv6]

Li, Z., Peng, S., Geng, X., and M. Negi, "PCEP Procedures and Protocol Extensions for Using PCE as a Central Controller (PCECC) for SRv6", draft-dhody-pce-pcep-extension-pce-controller-srv6-05 (work in progress), November 2020.

## Appendix A. Contributor Addresses

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

EMail: dhruv.ietf@gmail.com

Satish Karunanithi  
Huawei Technologies  
Divyashree Techno Park, Whitefield  
Bangalore, Karnataka 560066  
India

EMail: satishk@huawei.com

Adrian Farrel  
Juniper Networks, Inc  
UK

EMail: adrian@olddog.co.uk

Xuesong Geng  
Huawei Technologies  
China

Email: gengxuesong@huawei.com

Udayasree Palle

EMail: udayasreereddy@gmail.com

Katherine Zhao  
Huawei Technologies  
2330 Central Expressway  
Santa Clara, CA 95050  
USA

EMail: katherine.zhao@huawei.com

Boris Zhang  
Telus Ltd.  
Toronto  
Canada

EMail: boris.zhang@telus.com

Alex Tokar  
Cisco Systems  
Slovak Republic

EMail: atokar@cisco.com

#### Authors' Addresses

Zhenbin Li  
Huawei Technologies  
Huawei Bld., No.156 Beiqing Rd.  
Beijing 100095  
China

EMail: lizhenbin@huawei.com

Shuping Peng  
Huawei Technologies  
Huawei Bld., No.156 Beiqing Rd.  
Beijing 100095  
China

EMail: pengshuping@huawei.com

Mahendra Singh Negi  
RtBrick Inc  
N-17L, 18th Cross Rd, HSR Layout  
Bangalore, Karnataka 560102  
India

EMail: mahend.ietf@gmail.com

Quintin Zhao  
Etheric Networks  
1009 S CLAREMONT ST  
SAN MATEO, CA 94402  
USA

EMail: qzhao@ethericnetworks.com

Chao Zhou  
HPE

EMail: [chaozhou\\_us@yahoo.com](mailto:chaozhou_us@yahoo.com)