

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.

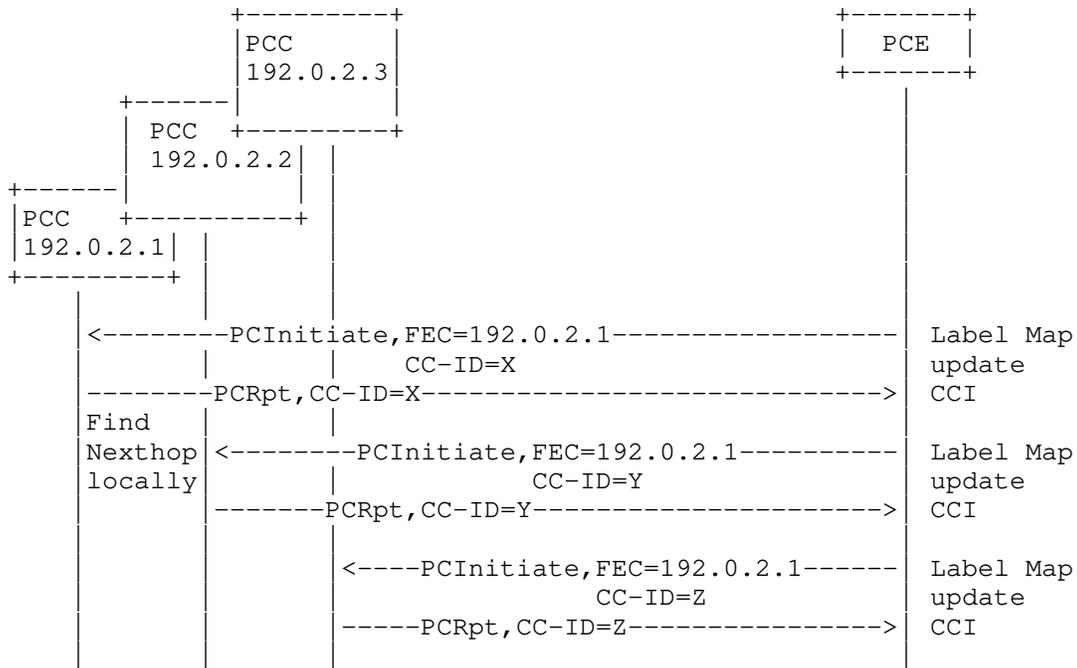
5.5.1.1. PCECC SR Node/Prefix SID allocation

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.

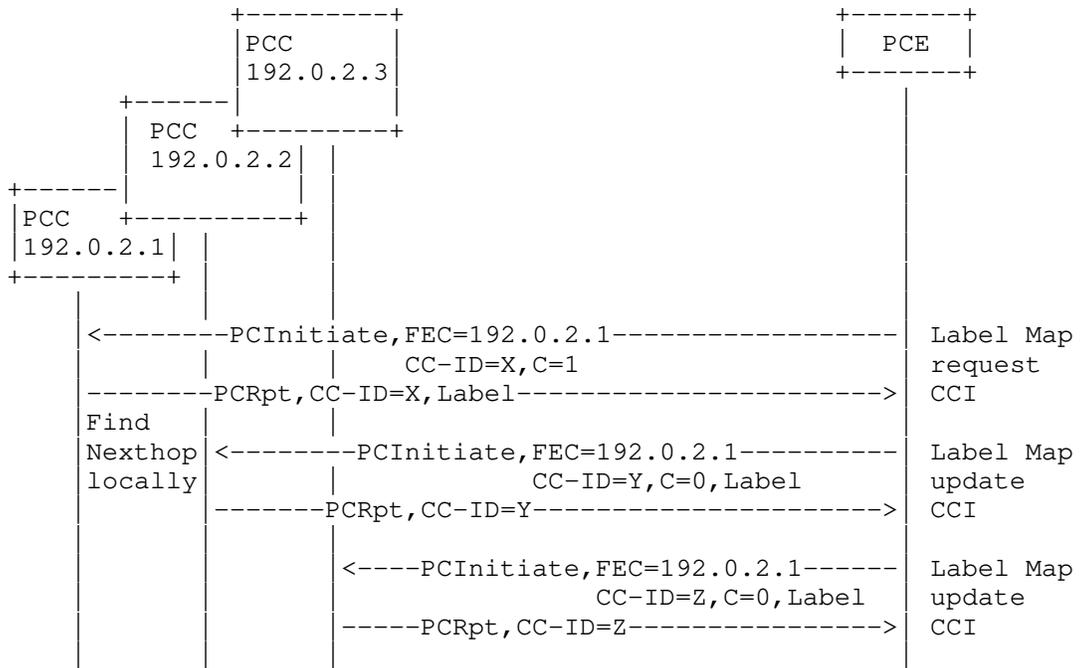


The forwarding behavior and the end result is similar to IGP based "Node-SID" in SR. Thus, from anywhere in the domain, it enforces the ECMP-aware shortest-path forwarding of the packet towards the related node as per [RFC8402].

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.

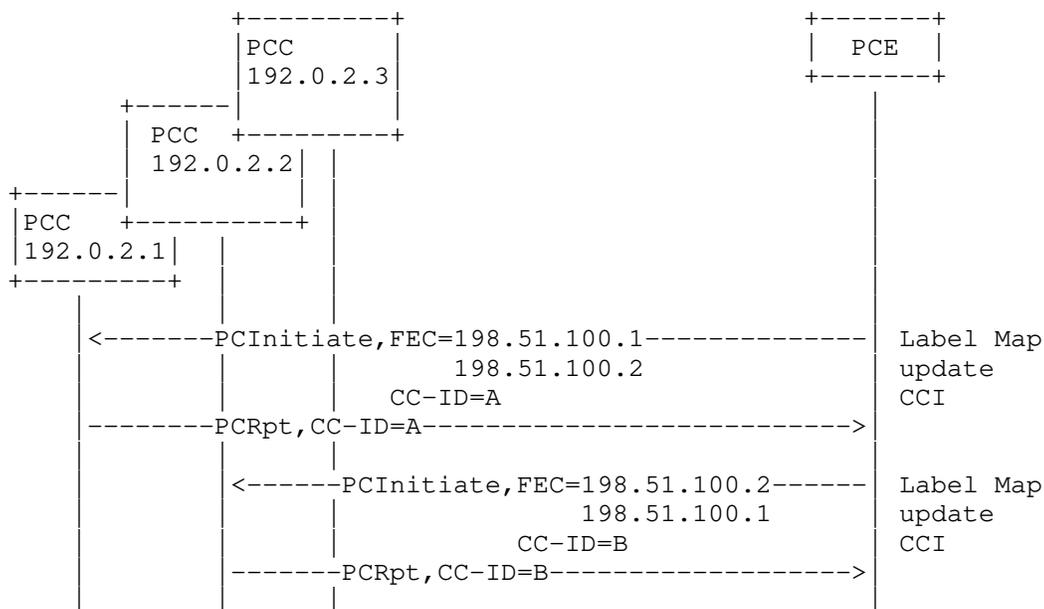
In the case where the label/SID allocation is made by the PCC itself (see Section 5.5.1.6), the PCE could request an allocation to be made by the PCC, and where the PCC would send a PCRpt with the allocated label/SID encoded in the CC-ID object as shown in Figure 2.



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.



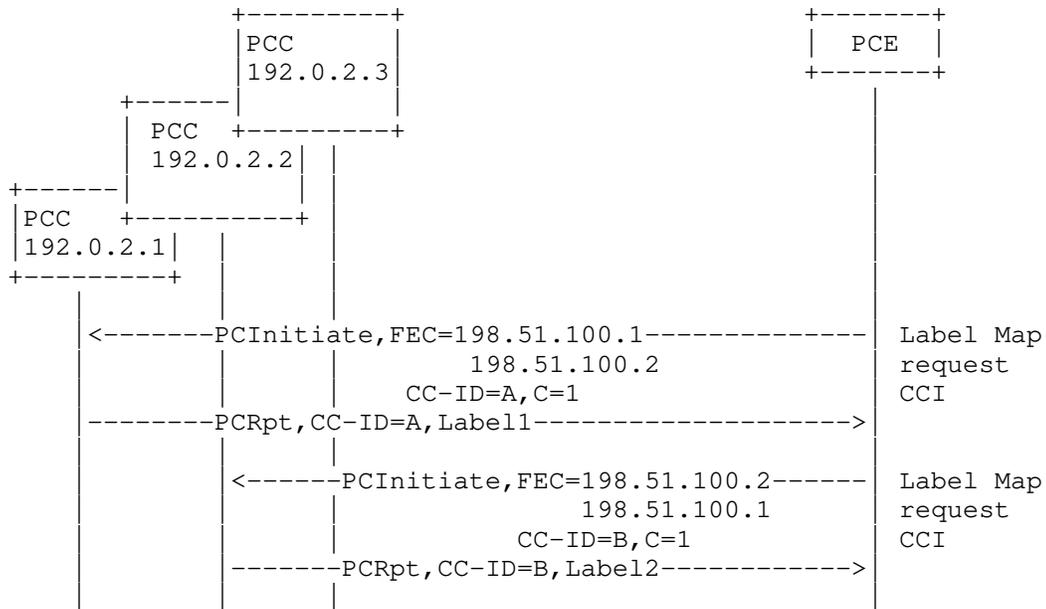
The forwarding behavior and the end result is similar to IGP based "Adj-SID" in SR.

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).

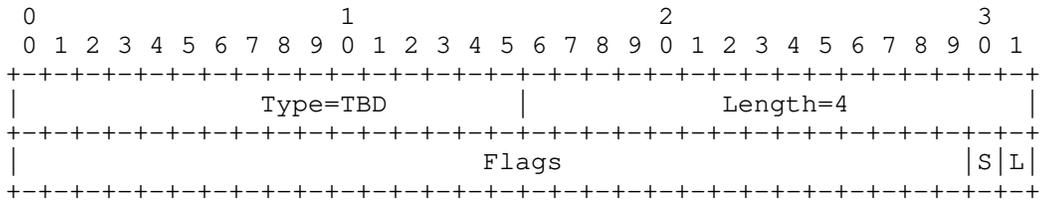
7. PCEP Objects

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.

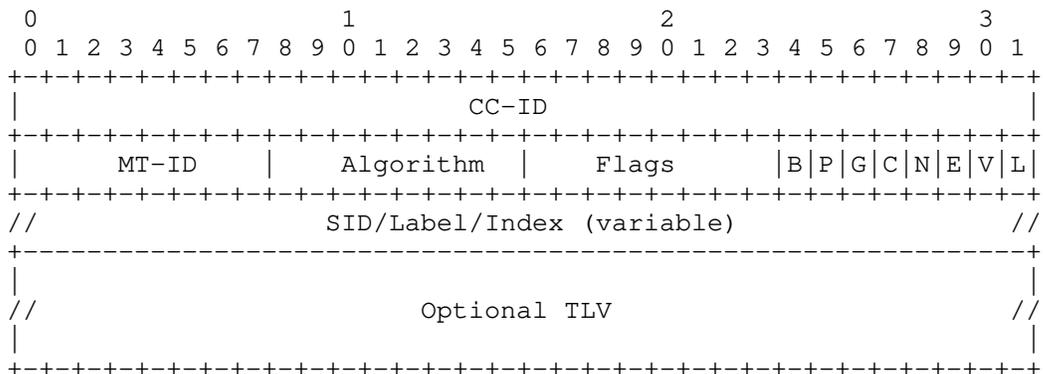
7.2. SR-TE Path Setup

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.

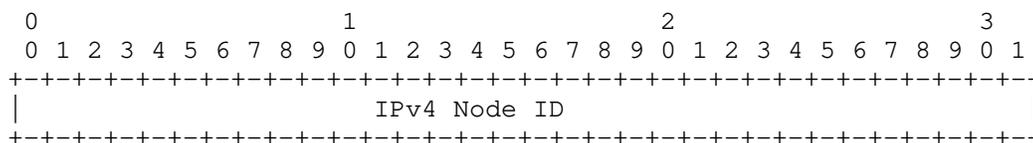
A 24-bit label where the 20 rightmost bits are used for encoding the label value.

7.4. FEC Object

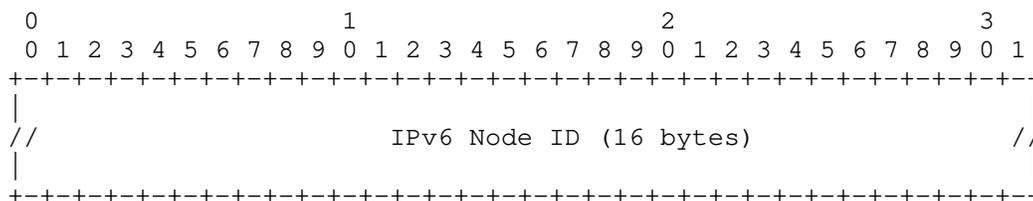
The FEC Object is used to specify the FEC information and MAY be carried within PCInitiate or PCRpt message.

FEC Object-Class is TBD3.

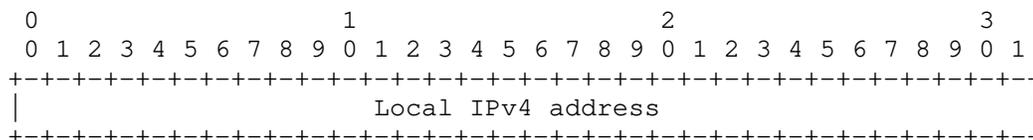
FEC Object-Type is 1 'IPv4 Node ID'.



FEC Object-Type is 2 'IPv6 Node ID'.



FEC Object-Type is 3 'IPv4 Adjacency'.



```

| Remote IPv4 address |
+-----+

```

FEC Object-Type is 4 'IPv6 Adjacency'.

```

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 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 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	Invalid operation.	Error-value = TBD5 : FEC object missing
		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] 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>>.
- [RFC8660] Bashandy, A., Ed., Filsfils, 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., Filsfils, 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., Filsfils, 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

E-Mail: atokar@cisco.com

Authors' Addresses

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

E-Mail: lizhenbin@huawei.com

Shuping Peng
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

E-Mail: pengshuping@huawei.com

Mahendra Singh Negi
RtBrick Inc
N-17L, 18th Cross Rd, HSR Layout
Bangalore, Karnataka 560102
India

E-Mail: mahend.ietf@gmail.com

Quintin Zhao
Etheric Networks
1009 S CLAREMONT ST
SAN MATEO, CA 94402
USA

E-Mail: qzhao@ethericnetworks.com

Chao Zhou
HPE

EMail: chaozhou_us@yahoo.com