PCE Working Group
Internet-Draft
Intended status: Experimental
Expires: September 17, 2016

Q. Zhao
Z. Li
D. Dhody
Huawei Technologies
C. Zhou
Cisco Systems

March 16, 2016

PCEP Procedures and Protocol Extensions for Using PCE as a Central Controller (PCECC) of LSPs
draft-zhao-pce-pcep-extension-for-pce-controller-03

Abstract

In certain networks deployment scenarios, service providers would like to keep all the existing MPLS functionalities in both MPLS and GMPLS while removing the complexity of existing signalling protocols such as LDP and RSVP-TE. In [I-D.zhao-pce-central-controller-user-cases], we propose to use the PCE [RFC5440] as a central controller (PCECC) so that LSP can be calculated/ signalled/initiated and label forwarding entries are downloaded through a centralized PCE server to each network devices along the LSP 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 and user cases where LSPs are calculated/setup/initiated and label forwarding entries are downloaded through extending the existing PCE architectures and PCEP.

This document also discuss the role of PCECC in Segment Routing (SR).

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 http://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 September 17, 2016.

Copyright Notice

Copyright (c) 2016 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 5
2. Terminology 5
3. PCECC Modes 5
4. PCEP Requirements 6
5. Procedures for Using the PCE as the Central Controller (PCECC) 7
5.1. Stateful PCE Model 7
5.2. New LSP Functions 7
5.3. PCECC Capability Advertisement 8
5.4. PCEP session IP address and TEDB Router ID 9
5.5. LSP Operations 9
5.5.1. Basic PCECC Mode 9
5.5.1.1. PCECC LSP Setup 9
5.5.1.2. Label Operations 11
5.5.2. PCE Initiated PCECC LSP 14
5.5.3. PCECC LSP Update 16
5.5.4. PCECC LSP State Report 19
5.5.5. PCECC Segment Routing (SR) 19
5.5.5.1. PCECC SR-BE 19
5.5.5.2. PCECC SR-TE 20
6. PCEP messages 21
6.1. Label Operations 22
6.1.1. The PCLabelUpd message 22
6.1.2. The PCInitiate message 23
7. PCEP Objects 23
7.1. OPEN Object 23
7.1.1. PCECC Capability TLV 23
7.2. PATH-SETUP-TYPE TLV 24

7.3. Label Object 24
7.3.1. Address TLVs 25
7.4. FEC Object 27
8. Security Considerations 28
9. Manageability Considerations 29
9.1. Control of Function and Policy 29
9.2. Information and Data Models 29
9.3. Liveness Detection and Monitoring 29
9.4. Verify Correct Operations 29
9.5. Requirements On Other Protocols 29
9.6. Impact On Network Operations 29
10. IANA Considerations 29
11. Acknowledgments 29
12. References 29
12.1. Normative References 29
12.2. Informative References 30
Appendix A. Contributor Addresses 32
Appendix B. Label Range Handling 32
B.1. Label Range Reservation 33
B.2. The PCLRResv message 34
B.3. LABEL-RANGE Object 35
Authors' Addresses 36

1. Introduction

In certain network deployment scenarios, service providers would like to have the ability to dynamically adapt to a wide range of customer's requests for the sake of flexible network service delivery, Software Defined Networks(SDN) has provides additional flexibility in how the network is operated compared to the traditional network.

The existing networking ecosystem has become awfully complex and highly demanding in terms of robustness, performance, scalability, flexibility, agility, etc. By migrating to the SDN enabled network from the existing network, service providers and network operators must have a solution which they can evolve easily from the existing network into the SDN enabled network while keeping the network services remain scalable, guarantee robustness and availability etc.

Taking the smooth transition between traditional network and the new SDN enabled network into account, especially from a cost impact assessment perspective, using the existing PCE components from the current network to function as the central controller of the SDN network is one choice, which not only achieves the goal of having a centralized controller, but also leverage the existing PCE network components.

The Path Computation Element communication Protocol (PCEP) provides mechanisms for Path Computation Elements (PCEs) to perform route computations in response to Path Computation Clients (PCCs) requests. PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model [I-D.ietf-pce-stateful-pce] describes a set of extensions to PCEP to enable active control of MPLS-TE and GMPLS tunnels.

[I-D.ietf-pce-pce-initiated-lsp] describes the setup and tear down of PCE-initiated LSPs under the active stateful PCE model, without the need for local configuration on the PCC, thus allowing for a dynamic MPLS network that is centrally controlled and deployed.

[I-D.ietf-pce-remote-initiated-gmpls-lsp] complements [I-D.ietf-pce-pce-initiated-lsp] by addressing the requirements for remote-initiated GMPLS LSPs.

Segment Routing (SR) technology leverage the source routing and tunnelling paradigms. A source node can choose a path without relying on hop-by-hop signalling 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 technology. 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.

A Segment Routed path (SR path) can be derived from an IGP Shortest Path Tree (SPT). Segment Routed Traffic Engineering paths (SR-TE paths) may not follow IGP SPT. Such paths may be chosen by a suitable network planning tool and provisioned on the source node of the SR-TE path.

It is possible to use a stateful PCE for computing one or more SR-TE paths taking into account various constraints and objective functions. Once a path is chosen, the stateful PCE can instantiate an SR-TE path on a PCC using PCEP extensions specified in [I-D.ietf-pce-pce-initiated-lsp] using the SR specific PCEP extensions described in [I-D.ietf-pce-segment-routing].

PCECC may further use PCEP protocol for SR label distribution instead of IGP extensions with some benefits.

Current MPLS label has local meaning. That is, MPLS label is always allocated by the downstream node to the upstream node. Then the MPLS label is only identified by the neighbouring upstream node and downstream node. The label allocation is done locally and signalled through the LDP/RSVP-TE/BGP protocol. To ease the label allocation and signalling mechanism, PCE can be conveniently used as a central

controller with Label download capability. Further PCE can also be used to manage the label range and Segment Routing Global Block (SRGB) etc.

The PCECC solution introduced in [I-D.zhao-pce-central-controller-user-cases] allow for a dynamic MPLS network that is eventually controlled and deployed without the deployment of RSVP-TE protocol or extended IGP protocol with node/ adjacency segment identifiers signalling capability while providing all the key MPLS functionalities needed by the service providers.

This draft specify the procedures and PCEP protocol extensions for using the PCE as the central controller and user cases where LSPs are calculated/setup/initiated/downloaded through extending the existing PCE architectures and PCEP.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

2. Terminology

The following terminology is used in this document.

IGP: Interior Gateway Protocol. Either of the two routing protocols, Open Shortest Path First (OSPF) or Intermediate System to Intermediate System (IS-IS).

PCC: Path Computation Client: any client application requesting a path computation to be performed by a Path Computation Element.

PCE: Path Computation Element. An entity (component, application, or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints.

TE: Traffic Engineering.

3. PCECC Modes

The following PCECC modes are supported -

- Basic PCECC.

- PCECC SR.

* PCECC SR-BE (Best Effort).

* PCECC SR-TE (Traffic Engineered).

In basic PCECC mode, the forwarding is similar to RSVP-TE signalled LSP without the RSVP-TE signalling. The PCECC allocates and download the label entries along the LSP. The rest of processing is similar to the existing stateful PCE mechanism.

In case of SR, there are two modes for SR-BE and SR-TE. For SR-BE, the forwarding is similar to LDP LSP without LDP signalling or IGP-SR extension. The SR Node label are allocated and distributed in the domain centrally by the PCE via PCEP. Each node (PCC) relies on local IGP for the nexthop calculation. For SR-TE, the forwarding uses label stack similar to IGP based SR-TE without IGP-SR extension. The SR node and adjacency labels are allocated and distributed in the domain centrally by the PCE via PCEP by PCECC. Rest of the processing is similar to existing stateful PCE with SR 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 message with a reason code that indicates no support for PCECC.

3. PCEP SHOULD provide a means to identify PCECC based LSP in the PCEP messages.

4. PCEP SHOULD provide a means to update (or cleanup) the label-download or label-map entry to the PCC.

5. Path Computation Client (PCC) supporting this draft MAY support a capability to communicate local label range or global label range or both to PCE. (see xref target="sec_lr"/>)

6. Path Computation Element (PCE) supporting this draft MAY have the capability to negotiate a global label range for a group of clients and communicate the final global label range to PCC. (see xref target="sec_lr"/>)

5. Procedures for Using the PCE as the Central Controller (PCECC)

5.1. Stateful PCE Model

Active stateful PCE is described in [I-D.ietf-pce-stateful-pce]. 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 [I-D.ietf-pce-stateful-pce]. PCRpt message MAYBE used to send PCECC LSP Reports.

(PCInitiate): a PCEP message described in [I-D.ietf-pce-pce-initiated-lsp]. PCInitiate message is used to setup PCE-Initiated LSP based on PCECC mechanism.

(PCUpd): a PCEP message described in [I-D.ietf-pce-stateful-pce]. PCUpd message is used to send PCECC LSP Update.

Per Node Label Operations: a PCEP mechanism to instruct label operations on each node. An approach to this could be -

(PCLabelUpd): a new PCEP message sent by a PCE to a PCC to download or cleanup the Label entry. The PCLabelUpd message described in Section 6.1.1.

(PCInitiate): reuse an existing PCEP message described in [I-D.ietf-pce-pce-initiated-lsp]. PCInitiate message.

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 Update	PCLabelUpd/PCInitiate
Label entry Cleanup	PCLabelUpd/PCInitiate
PCECC Initiated LSP	PCInitiate
PCECC LSP Update	PCUpd
PCECC LSP State Report	PCRpt
PCECC LSP Delegation	PCRpt

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" TLV, described in Section 7.1.1 of this document, in the OPEN Object to advertise its support for PCECC extensions.

The presence of the PCECC Capability TLV in PCC's OPEN Object indicates that the PCC is willing to function as a PCECC client.

The presence of the PCECC Capability 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 PCECC Capability 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 LSP setup/download/label-range reservation if 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 TLV and STATEFUL-PCE-CAPABILITY TLV ([I-D.ietf-pce-stateful-pce]) in OPEN Object to support the extensions defined in this document. If PCECC-CAPABILITY 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.

PCECC-CAPABILITY TLV includes a S-bit to indicate support for PCECC-SR. A PCEP speaker MUST set S-bit in PCECC-CAPABILITY TLV and include SR-PCE-CAPABILITY TLV ([I-D.ietf-pce-segment-routing]) in OPEN Object to support the PCECC SR-TE extensions defined in this

document. If S-bit is set in PCECC-CAPABILITY TLV and SR-PCE-CAPABILITY TLV is not advertised in OPEN Object, PCEP speaker 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 [I-D.ietf-idr-ls-distribution] 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. [I-D.ietf-idr-ls-distribution] 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 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.

5.5.1. Basic PCECC Mode

5.5.1.1. 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 Path Setup Type set for basic PCECC (see Section 7.2) and D (Delegate) flag (see [I-D.ietf-pce-stateful-pce]) set in the LSP object.

LSP-IDENTIFIER TLV MAY be included for PCECC LSP, the LSP-ID SHOULD be generated by the PCE for PCECC LSP. In the first PCRpt message of PCECC LSP, LSP ID of LSP-IDENTIFIER TLV is set to zero.

When a PCE receives PCRpt message with P and D flags set, it MAY generates LSP ID; calculates the path and assigns labels along the path; and setups the path by sending PCLabelUpd or PCInitiate message to each node along the path of the LSP.

In response to the delegate PCRpt message, the PCE SHOULD send the PCUpd message with the same PLSP-ID to the Ingress PCC.

The PCECC LSPs MUST be delegated to a PCE at all times.

LSP deletion operation for PCECC LSP is same as defined in [I-D.ietf-pce-stateful-pce]. If the PCE receives PCRpt message for LSP deletion then it does Label cleanup operation as described in Section 5.5.1.2.1.2 for the corresponding LSP.

The Basic PCECC LSP setup sequence is as shown below using a new message PCLabelUpd.

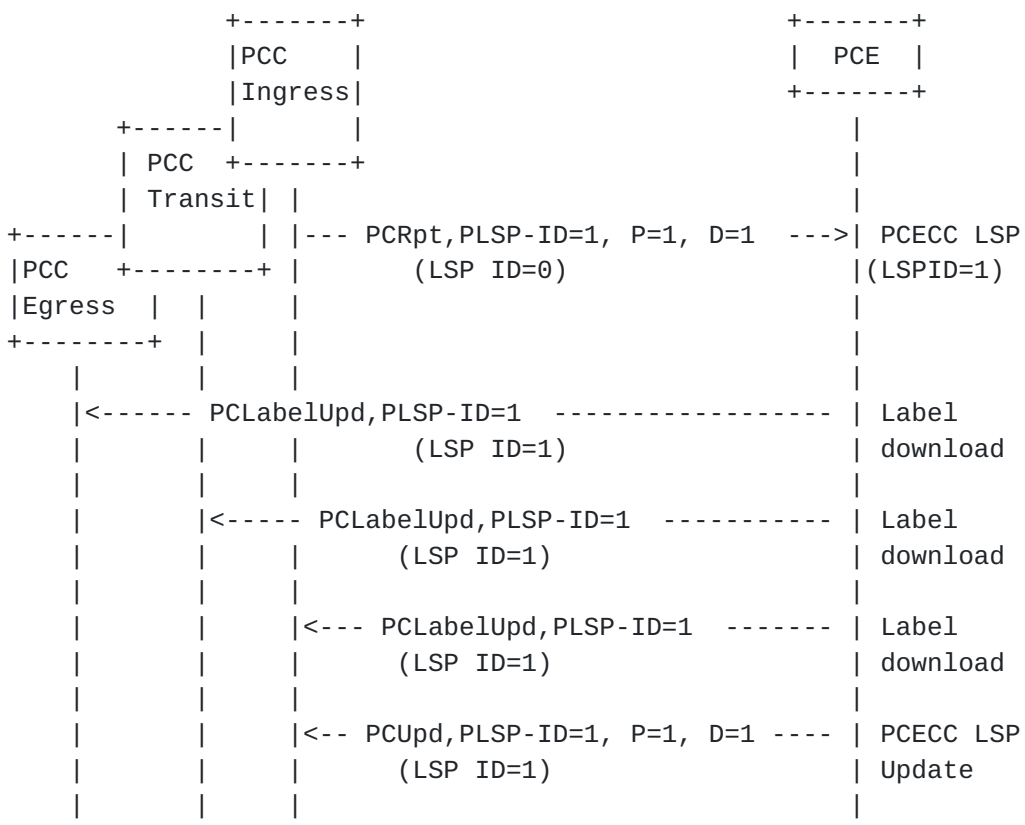
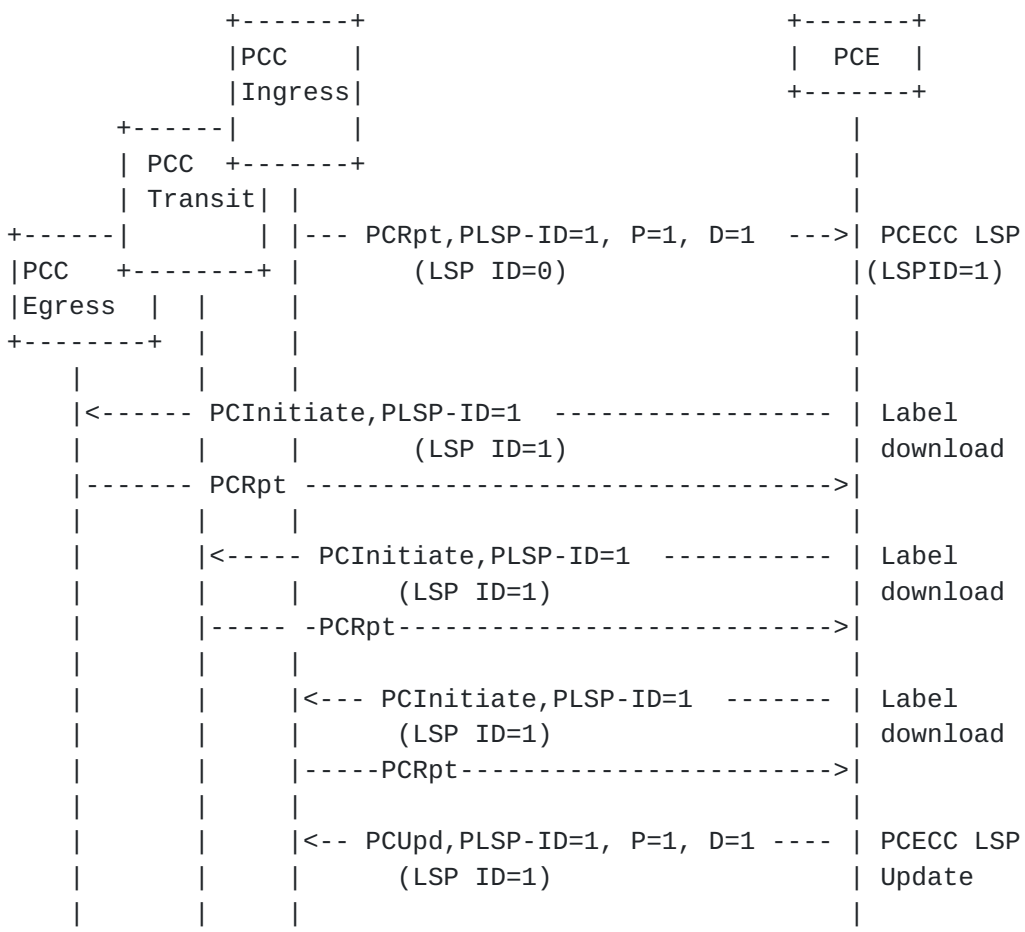


Figure 2: Using PCLabelUpd For Label Operations



Some open issue -

* PCC send PCRpt messages for each PCInitiate message?
* Should one use PCInitiate at Ingress too? How to link this with existing tunnel at the PCC?

Figure 3: Using PCInitiate For Label Operations

The 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.

5.5.1.2. Label Operations

5.5.1.2.1. Via a new PCLabelUpd Message

The new label operations in PCEP can be done via a new message and by defining new PCEP Objects for Label operations and FEC.

5.5.1.2.1.1. Label Download

In order to setup an LSP based on PCECC, the PCE sends a PCLabelUpd message to each node of the LSP to download the Label entry as described in Section 5.5.1.1.

The LSP object in PCLabelUpd MUST include the LSP-IDENTIFIER TLV.

If a node (PCC) receives a PCLabelUpd message but failed to download the Label entry, it MUST send a PCErr message with Error-type=[TBD] (label download failed) and Error-value=[TBD].

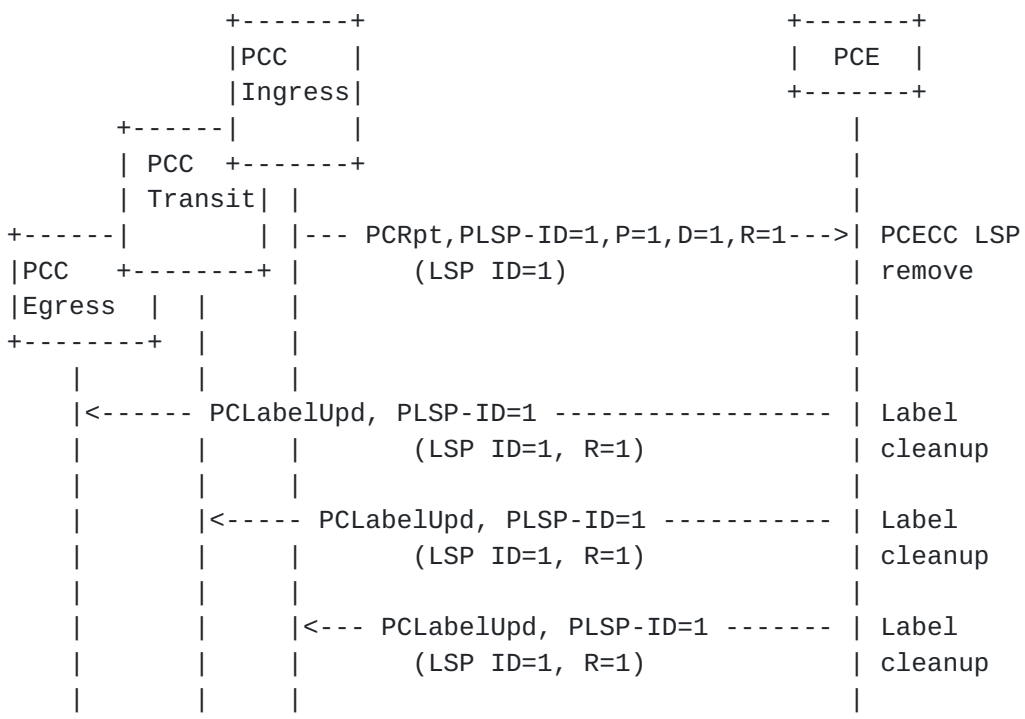
Two new PCEP objects for LABEL and FEC are defined in Section 7.3 and Section 7.4 to encode the Label operations and its associated FEC.

5.5.1.2.1.2. Label Cleanup

In order to delete an LSP based on PCECC, the PCE sends a PCLabelUpd message to each node along the path of the LSP to cleanup the Label entry.

If the PCC receives a PCLabelUpd message but does not recognize the label, the PCC MUST generate a PCErr message with Error-Type 19(Invalid operation) and Error-Value=3, "Unknown Label".

The R flag in SRP object defined in [I-D.ietf-pce-pce-initiated-lsp] specifies the deletion of Label Entry in the new PCLabelUpd message.



5.5.1.2.2. Via an existing PCInitiate Message

The new label operations in PCEP can be done via existing PCInitiate message and by reusing PCEP Objects (and subobjects) to encode Label operations and FEC. A TE LSP can be envisioned as a series of "cross-connects" and ERO in the PCInitiate message can be used to indicate each such cross-connect along the path. The use of PCInitiate message would also require each node to send the PCRpt message. The handling of PLSP-ID, LSP Identifiers, Symbolic path name would need to be taken into consideration.

The PATH-SETUP-TYPE TLV [I-D.ietf-pce-lsp-setup-type] in the SRP object can identify the PCECC LSP.

5.5.1.2.2.1. Label Download

In order to setup an LSP based on PCECC, the PCE sends a PCInitiate message to each node of the LSP to download the Label entry as described in Section 5.5.1.1 represented as a "cross-connect".

If a node (PCC) receives a PCInitiate message with setup type indicated as PCECC, it should not try to initiate a complete tunnel. The error-handling and other processing would remain unchanged.

The ERO in the PCInitiate message would indicate the cross-connect as {input interface, input label, output interface, output label}. For

Egress only input interface/label would be present and for Ingress only output interface/label. For Ingress either PCInitiate is used or the same information can be put in PCUpd message as well.

5.5.1.2.2.2. Label Cleanup

In order to delete an LSP based on PCECC, one can use existing PCInitiate message (with R flag) to each node along the path of the LSP to cleanup the Label entry. The R flag in SRP object defined in [I-D.ietf-pce-pce-initiated-lsp] is used during cleanup.

The error-handling and other processing would remain unchanged.

5.5.2. PCE Initiated PCECC LSP

The LSP Instantiation operation is same as defined in [I-D.ietf-pce-pce-initiated-lsp].

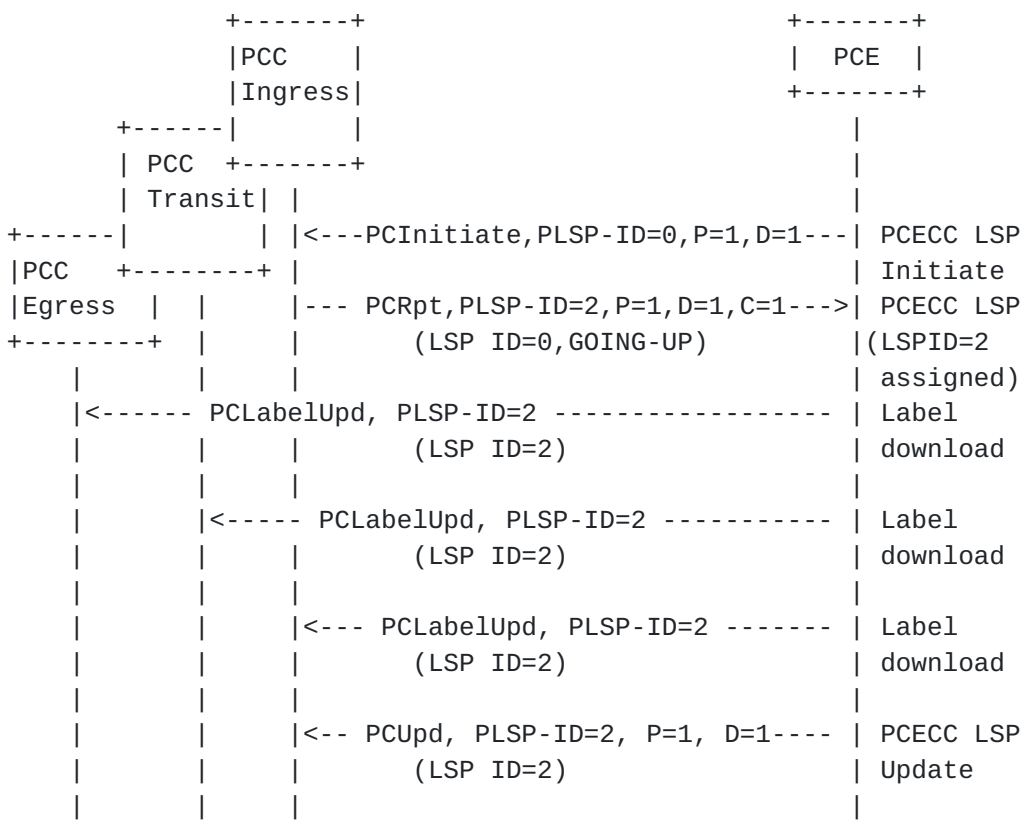
In order to setup a PCE Initiated LSP based on PCECC mechanism, a PCE sends PCInitiate message with Path Setup Type set for basic PCECC (see Section 7.2) to the Ingress PCC.

The Ingress PCC MUST also set D (Delegate) flag (see [I-D.ietf-pce-stateful-pce]) and C (Create) flag (see [I-D.ietf-pce-pce-initiated-lsp]) in LSP object of PCRpt message. The PCC responds with first PCRpt message with the status as "GOING-UP" and assigned PLSP-ID.

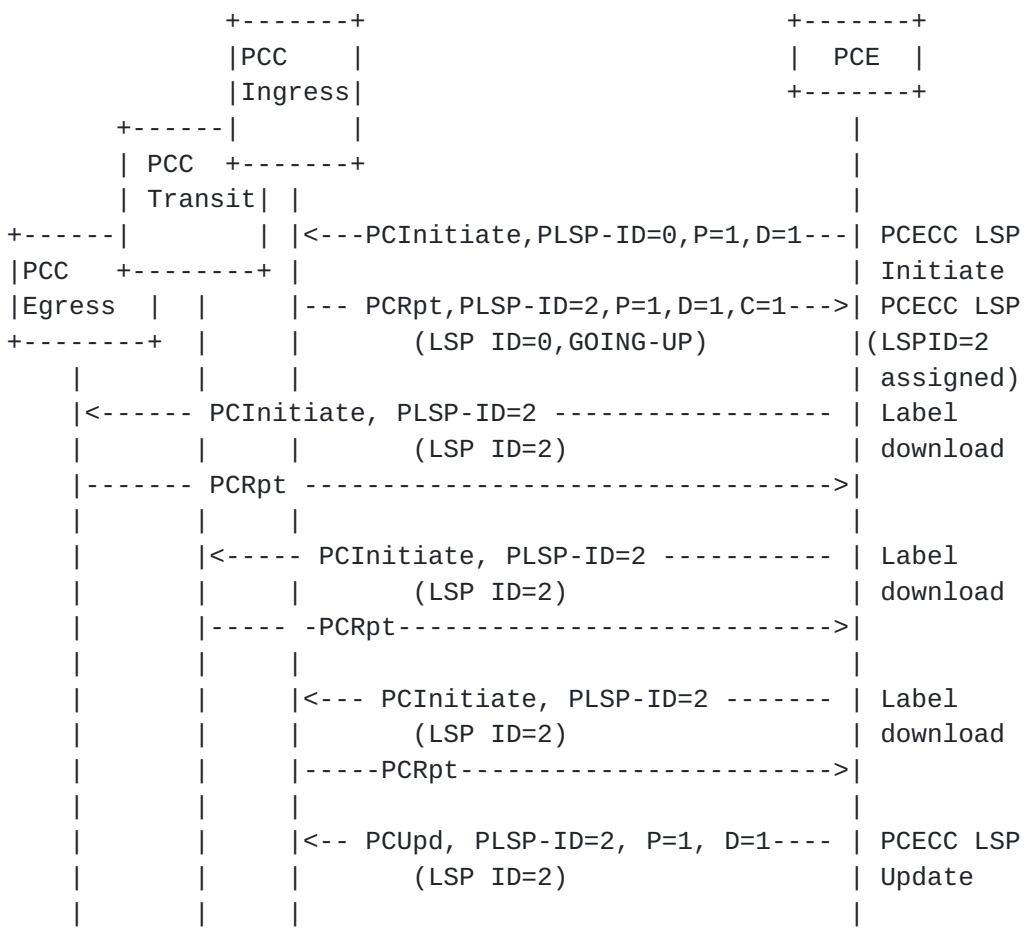
The rest of the PCECC LSP setup operations are same as those described in Section 5.5.1.1.

The LSP deletion operation for PCE Initiated PCECC LSP is same as defined in [I-D.ietf-pce-pce-initiated-lsp]. The PCE should further perform Label entry cleanup operation as described in Section 5.5.1.2.1.2 for the corresponding LSP.

The PCE Initiated PCECC LSP setup sequence is shown below using a new PCLabelUpd message.



The PCE Initiated PCECC LSP setup sequence is shown below using existing PCInitiate message.



Some open issue -

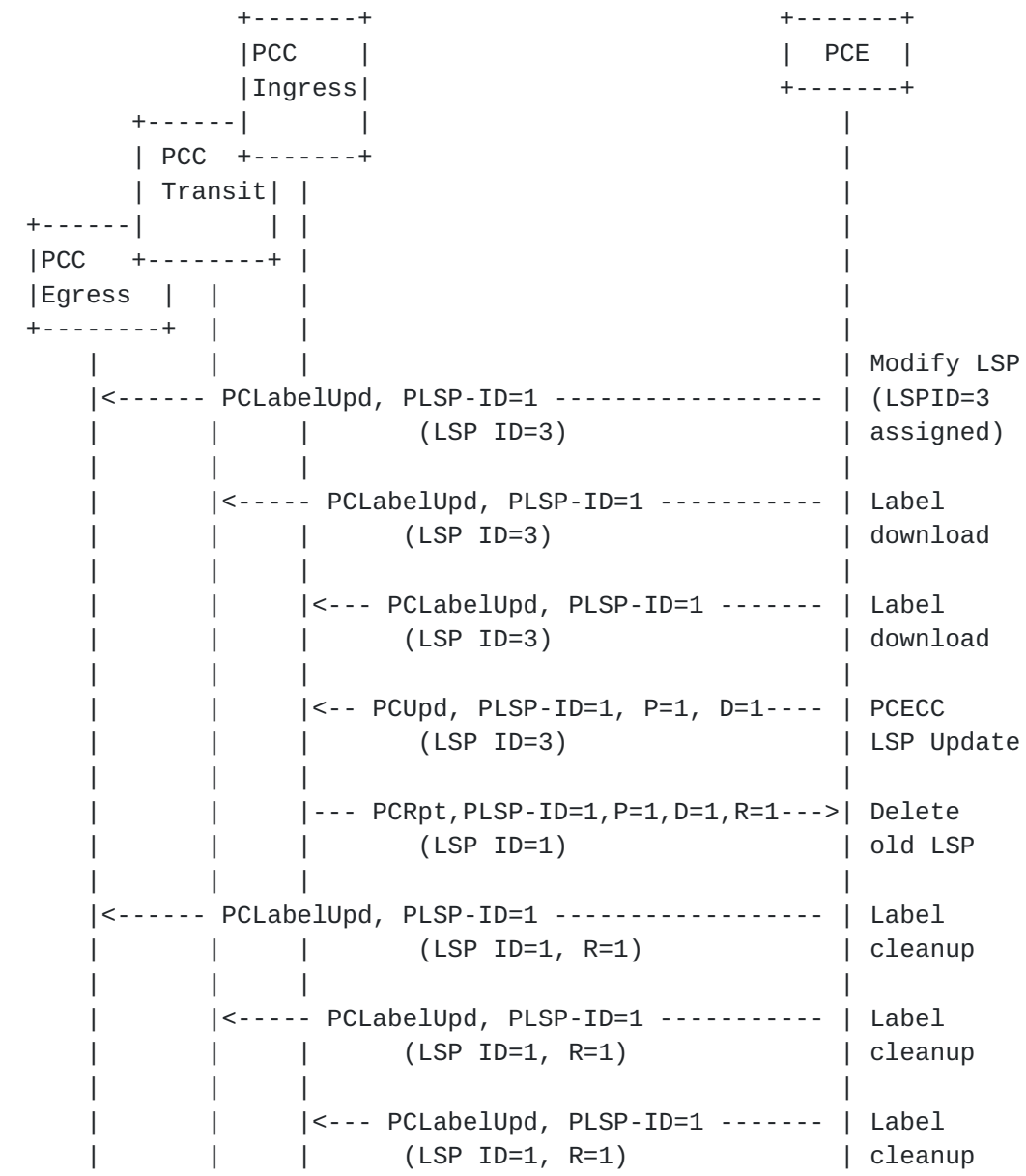
* PCC send PCRpt messages for each PCInitiate message?
* How to link 2 PCInitiate message at Ingress one to initiate the full tunnel, and one to download the cross-connect (label)?

5.5.3. PCECC LSP Update

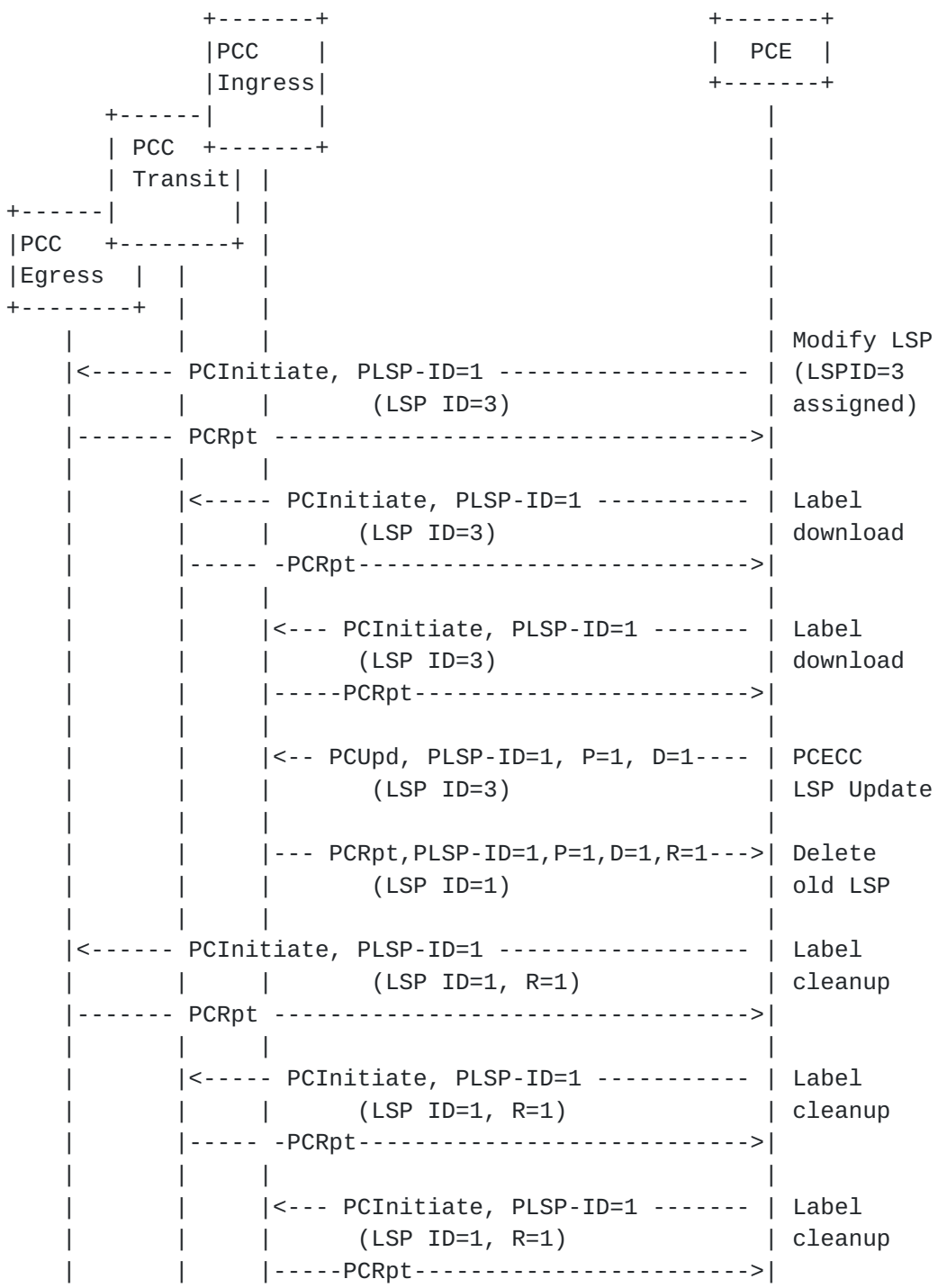
Incase of a modification of PCECC LSP with a new path, a PCE sends a PCUpd message to the Ingress PCC.

When a PCC receives a PCUpd message for an existing LSP, a PCC MAY follow the make-before-break procedure i.e. first download labels for the updated LSP and then switch traffic, before cleaning up the old labels. On successful traffic switch over to the new LSP, PCC sends a PCRpt message to the PCE for the deletion of old LSP. Further the PCE does cleanup operation for the old LSP as described in Section 5.5.1.2.1.2.

The PCECC LSP Update and make-before-break sequence is shown below using a new PCLabelUpd message.



The PCECC LSP Update and make-before-break sequence is shown below using existing PCInitiate message.



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.

5.5.4. 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.

5.5.5. 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) 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.

PCECC SR capability (S-bit) MUST be advertised in Open message.

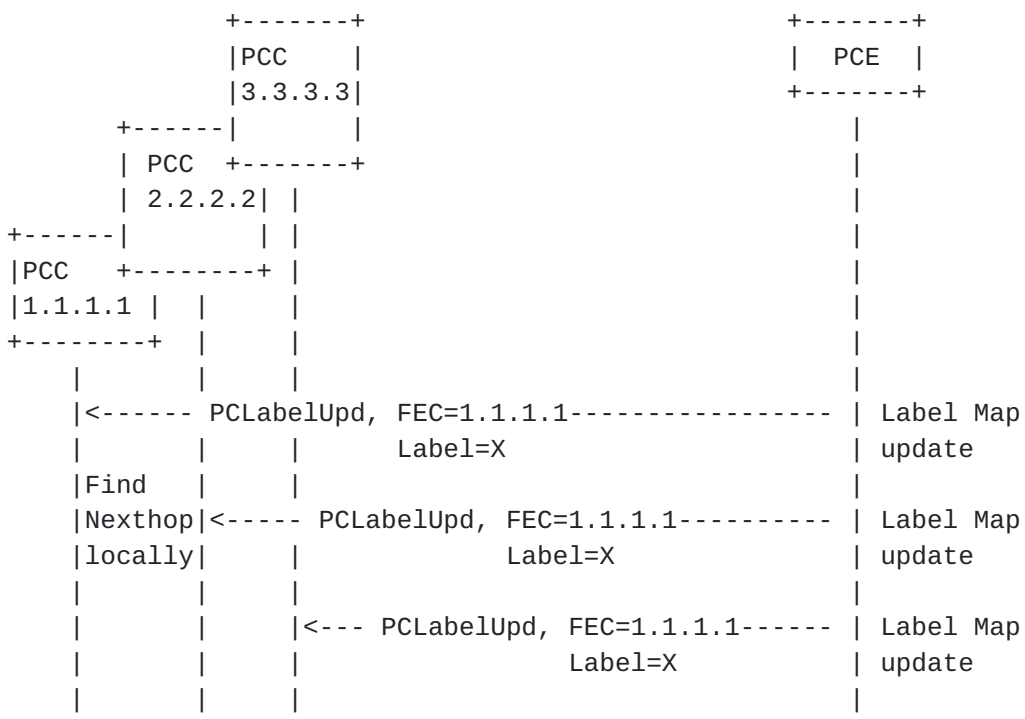
[EDITOR NOTE: Currently only the use of a new message PCLabelUpd is discussed in this section, the editors feel that use of existing PCInitiate message for SR node and adjacency label would be problematic.]

5.5.5.1. PCECC SR-BE

Each node (PCC) is allocated a node-SID (label) by the PCECC. The PCECC sends PCLabelUpd 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-BE node label maps (Section 5.4).

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 PCLabelUpd message in this case MUST NOT have LSP object but use the new FEC object.



The forwarding behaviour 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.

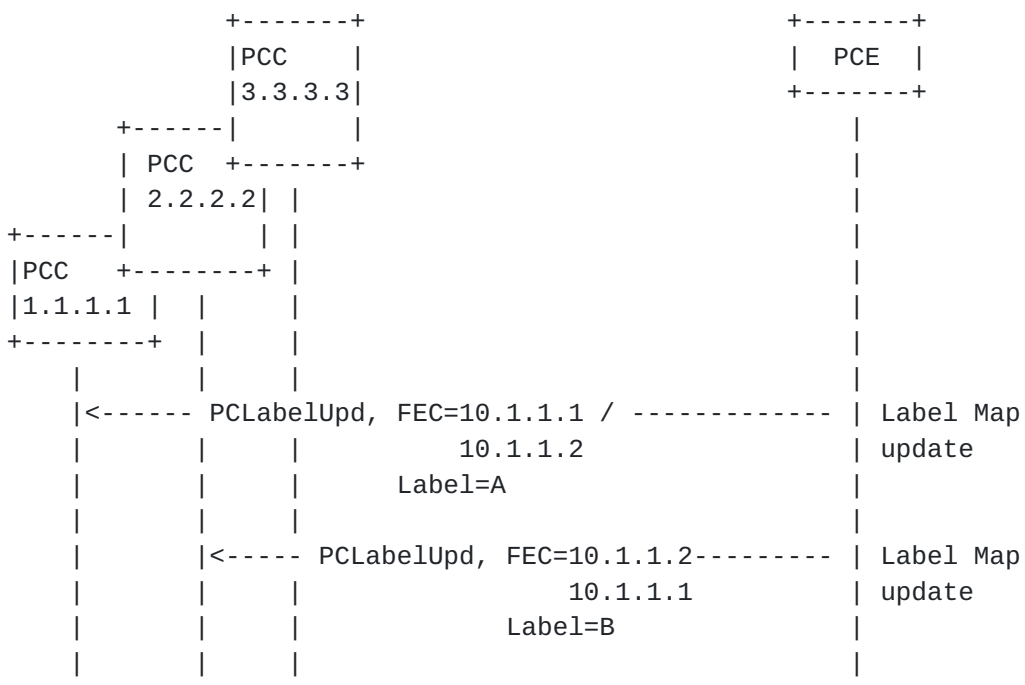
PCE relies on the Node label cleanup using the same PCLabelUpd message.

5.5.5.2. PCECC SR-TE

A Segment Routed Best Effort path (SR-BE path) can be derived from an IGP Shortest Path Tree (SPT) as explained above. On the other hand, SR-TE paths may not follow IGP SPT. Such paths may be chosen by a PCE and provisioned on the source node of the SR-TE path.

[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 SR-TE, apart from node-SID, Adj-SID is used where each adjacency is allocated an Adj-SID (label) by the PCECC. The PCECC sends PCLabelUpd 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-BE, the PCLabelUpd message in this case MUST NOT have LSP object but use the new FEC object.



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-TE (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 PCLabelUpd message.

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 MAY be included in the LSP object for PCECC LSP.

6.1. Label Operations

6.1.1. The PCLabelUpd message

A new Label Update Message (also referred to as PCLabelUpd) is a PCEP message sent by a PCE to a PCC to download label or update the label map. The same message is also used to cleanup the Label entry. The Message-Type field of the PCEP common header for the PCLabelUpd message is set to [TBD].

The format of the PCLabelUpd message is as follows:

```
<PCLabelUpd Message> ::= <Common Header>
                         <pce-label-update-list>
Where:
<pce-label-update-list> ::= <pce-label-update>
                   [<pce-label-update-list>]

<pce-label-update> ::= (<pce-label-download>|<pce-label-map>)

Where:
<pce-label-download> ::= <SRP>
                         <LSP>
                         <label-list>

<pce-label-map> ::= <SRP>
                    <LABEL>
                    <FEC>

<label-list > ::=  <LABEL>
                  [<label-list>]
```

The PCLabelUpd message is used to download label along the path of the LSP for the basic PCECC mode, as well as to update the label map for the Node and Adjacency Label in case of SR.

The SRP object is defined in [I-D.ietf-pce-stateful-pce] and this document extends the use of SRP object in PCLabelUpd message. The SRP object is mandatory and MUST be included in PCLabelUpd message. 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 [I-D.ietf-pce-stateful-pce] and this document extends the use of LSP object in PCLabelUpd message. The LSP is an optional object and used in the basic PCECC mode in

PCLabelUpd message. LSP Identifiers TLV is defined in [I-D.ietf-pce-stateful-pce], it MAY be included in the LSP object in PCLabelUpd message.

The LABEL object is defined in Section 7.3. The LABEL is the mandatory object and MUST be included in PCLabelUpd message. If the LABEL object is missing, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=[TBD] (LABEL object missing). More than one LABEL object MAY be included in the PCLabelUpd message for the transit LSR.

The FEC object is defined in Section 7.4. The FEC is an optional object and used in PCECC SR mode in PCLabelUpd message. The FEC object encodes the Node and Adjacency information of the Label Map.

To cleanup the SRP object must set the R (remove) bit.

6.1.2. The PCInitiate message

Message described in [I-D.ietf-pce-pce-initiated-lsp] continue to apply.

7. PCEP Objects

The PCEP objects defined in this document are compliant with the PCEP object format defined in [RFC5440]. The P flag and the I flag of the PCEP objects defined in this document MUST always be set to 0 on transmission and MUST be ignored on receipt since these flags are exclusively related to path computation requests.

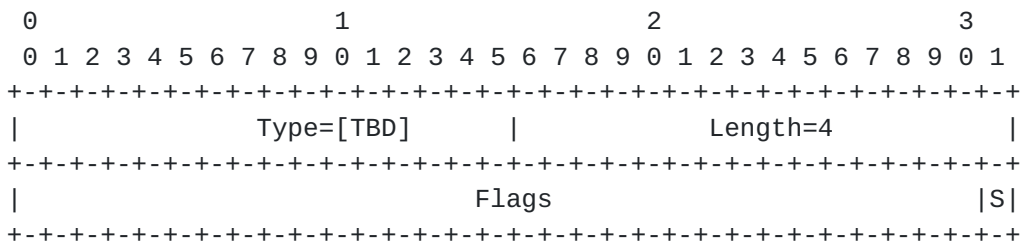
7.1. OPEN Object

This document defines a new optional TLVs for use in the OPEN Object.

7.1.1. PCECC Capability TLV

The PCECC-CAPABILITY TLV is an optional TLV for use in the OPEN Object for PCECC capability advertisement. 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):

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.

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 = 2: Path is setup via Basic PCECC mode.

PST = 1 (defined in [I-D.ietf-pce-segment-routing]) can be reused when Path is setup via PCECC SR-TE mode.

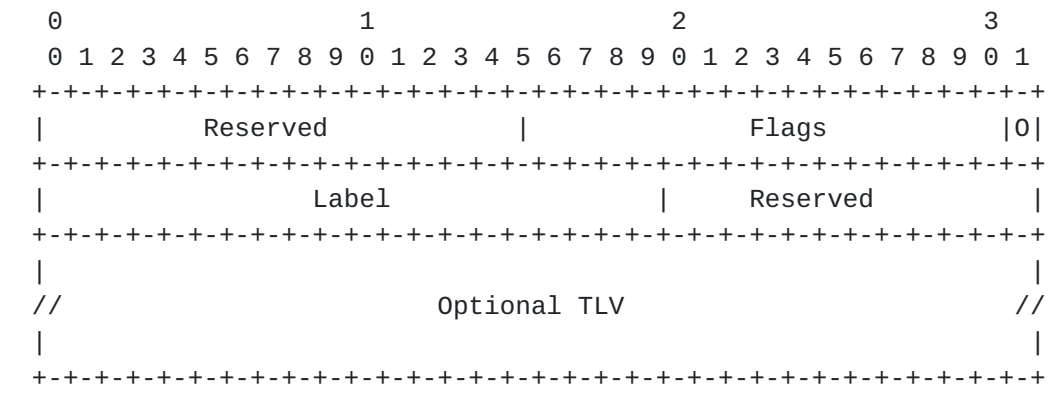
On a PCRpt/PCUpd/PCInitiate message, the PST=2 in PATH-SETUP-TYPE TLV in SRP object indicates that this LSP was setup via a basic PCECC based mechanism; the PST=1 indicates that this LSP was setup via a SR-TE based mechanism where either the labels are allocated by PCE via PCECC mechanism or advertised by IGP.

7.3. Label Object

The LABEL Object is used to specify the Label information and MUST be carried within PCLabelUpd message.

LABEL Object-Class is TBD.

LABEL Object-Type is 1.



The fields in the LABEL object are as follows:

Flags: is used to carry any additional information pertaining to the label. 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 nexthop information (via IPV4-ADDRESS TLV or IPV6-ADDRESS TLV or UNNUMBERED-IPV4-ID-ADDRESS TLV in LABEL 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 LABEL object).

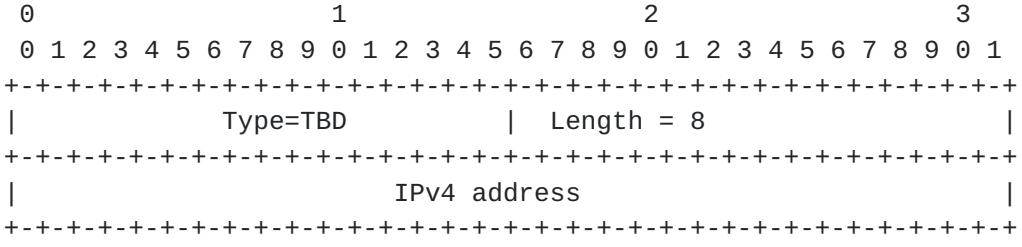
Label (20-bit): The Label information encoded such that the 20 rightmost bits represent a label.

Reserved (12 bit): Set to zero while sending, ignored on receive.

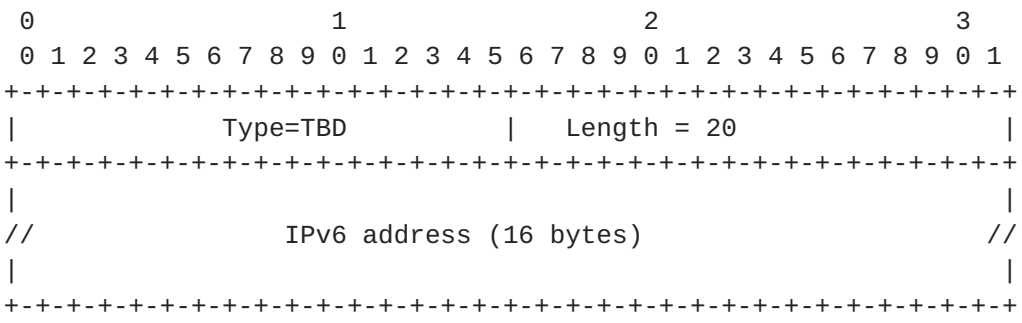
7.3.1. Address TLVs

This document defines the following TLV for the LABEL object to associate the nexthop information incase of an outgoing label and local interface information incase of an incoming label.

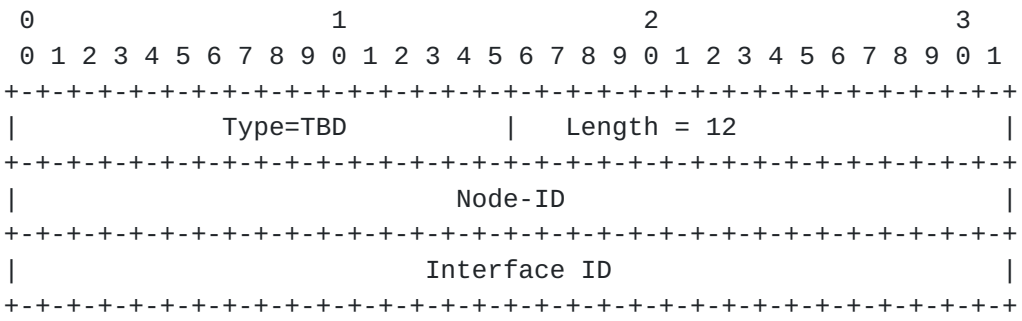
IPV4-ADDRESS TLV:



IPV6-ADDRESS TLV:



UNNUMBERED-IPV4-ID-ADDRESS TLV:



The address TLVs are as follows:

IPV4-ADDRESS TLV: an IPv4 address.

IPV6-ADDRESS TLV: an IPv6 address.

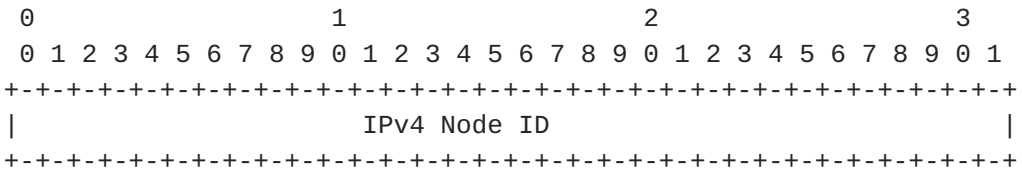
UNNUMBERED-IPV4-ID-ADDRESS TLV: a pair of Node ID / Interface ID tuples.

7.4. FEC Object

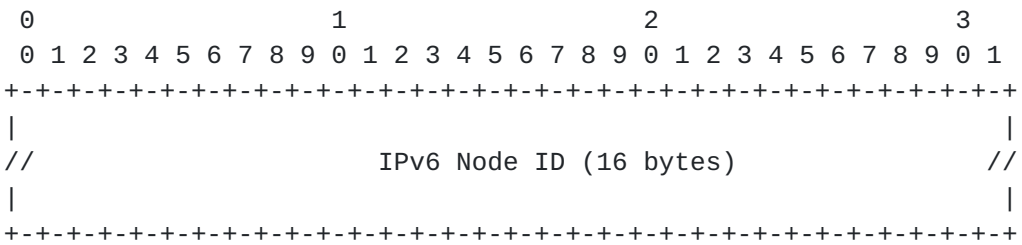
The FEC Object is used to specify the FEC information and MAY be carried within PCLabelUpd message.

FEC Object-Class is TBD.

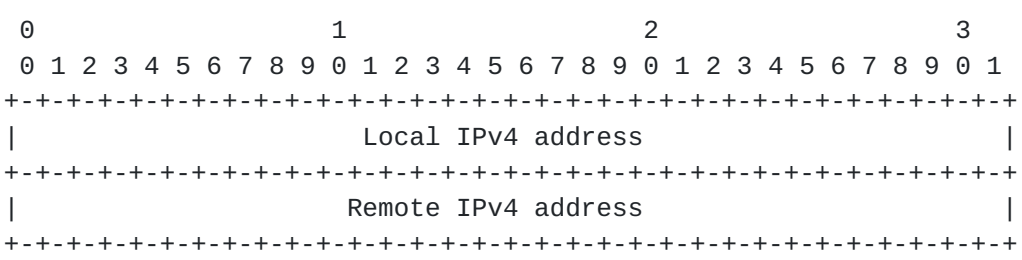
FEC Object-Type is 1 'IPv4 Node ID'.



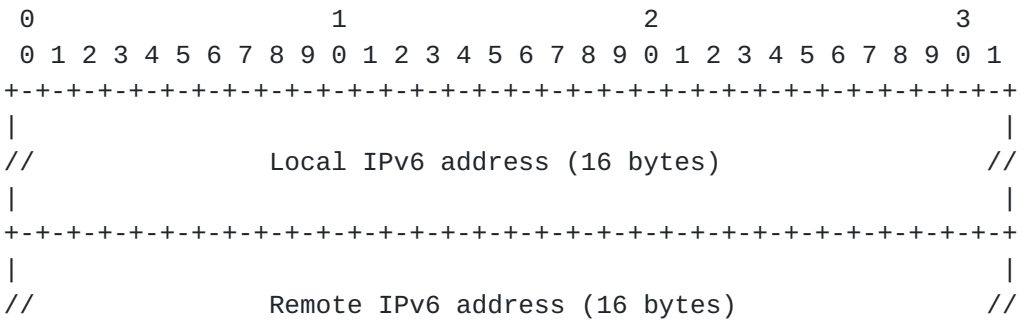
FEC Object-Type is 2 'IPv6 Node ID'.

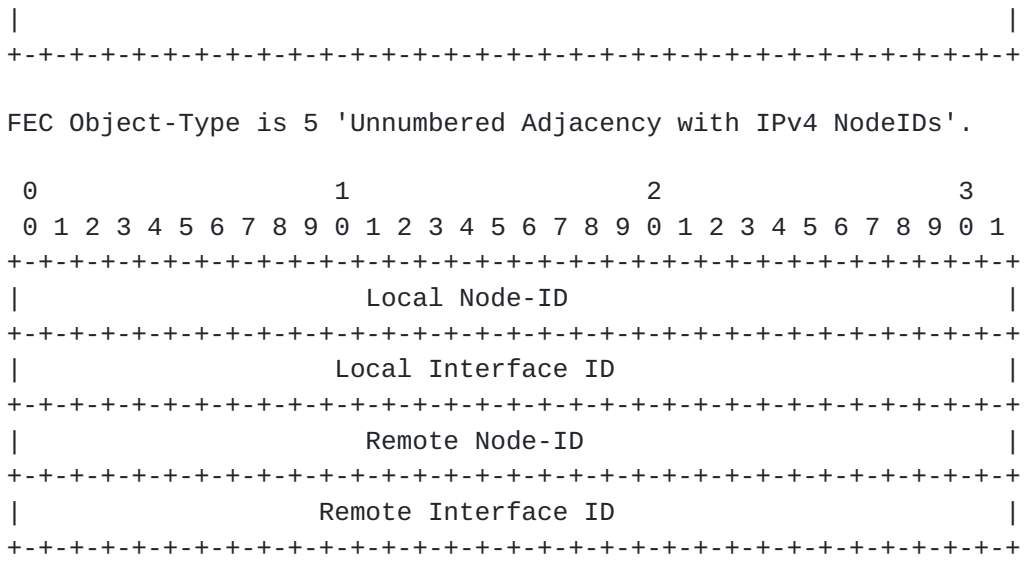


FEC Object-Type is 3 'IPv4 Adjacency'.



FEC Object-Type is 4 'IPv6 Adjacency'.





FEC Object-Type is 5 'Unnumbered Adjacency with IPv4 NodeIDs'.

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.

8. Security Considerations

TBD

9. Manageability Considerations

9.1. Control of Function and Policy

TBD.

9.2. Information and Data Models

TBD.

9.3. Liveness Detection and Monitoring

TBD.

9.4. Verify Correct Operations

TBD.

9.5. Requirements On Other Protocols

TBD.

9.6. Impact On Network Operations

TBD.

10. IANA Considerations

TBD

11. Acknowledgments

We would like to thank Robert Tao, Changjing Yan, Tieying Huang and Avantika for their useful comments and suggestions.

We would like to thank Adrian Farrel for presenting followup of PCECC proposal in IETF 94.

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, <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>.

[I-D.ietf-pce-stateful-pce] Crabbe, E., Minei, I., Medved, J., and R. Varga, "PCEP Extensions for Stateful PCE", draft-ietf-pce-stateful-pce-13 (work in progress), December 2015.

[I-D.ietf-pce-pce-initiated-lsp] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", draft-ietf-pce-pce-initiated-lsp-05 (work in progress), October 2015.

[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-03 (work in progress), March 2016.

[I-D.ietf-idr-ls-distribution] Gredler, H., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and TE Information using BGP", draft-ietf-idr-ls-distribution-13 (work in progress), October 2015.

12.2. 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, <http://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, <http://www.rfc-editor.org/info/rfc4203>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <http://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, <http://www.rfc-editor.org/info/rfc5307>.

[I-D.li-mpls-global-label-framework]
Li, Z., Zhao, Q., Chen, X., Yang, T., and R. Raszuk, "A Framework of MPLS Global Label", draft-li-mpls-global-label-framework-02 (work in progress), July 2014.

[I-D.zhao-pce-central-controller-user-cases]
Zhao, Q., Li, Z., Ke, Z., Fang, L., Zhou, C., and T. Communications, "The Use Cases for Using PCE as the Central Controller(PCECC) of LSPs", draft-zhao-pce-central-controller-user-cases-02 (work in progress), July 2015.

[I-D.ietf-pce-remote-initiated-gmpls-lsp]
Ali, Z., Sivabalan, S., Filsfils, C., Varga, R., Lopez, V., Dios, O., and X. Zhang, "Path Computation Element Communication Protocol (PCEP) Extensions for remote-initiated GMPLS LSP Setup", draft-ietf-pce-remote-initiated-gmpls-lsp-01 (work in progress), July 2015.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-07 (work in progress), December 2015.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., 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-06 (work in progress), December 2015.

[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-06 (work in progress), December 2015.

[I-D.ietf-pce-segment-routing]
Sivabalan, S., Medved, J., Filsfils, C., Crabbe, E., Lopez, V., Tantsura, J., Henderickx, W., and J. Hardwick, "PCEP Extensions for Segment Routing", draft-ietf-pce-segment-routing-06 (work in progress), August 2015.

[I-D.ietf-pce-lsp-setup-type]
Sivabalan, S., Medved, J., Minei, I., Crabbe, E., Varga, R., Tantsura, J., and J. Hardwick, "Conveying path setup type in PCEP messages", draft-ietf-pce-lsp-setup-type-03 (work in progress), June 2015.

Appendix A. Contributor Addresses

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

EMail: udayasree.palle@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

Appendix B. Label Range Handling

This section is created to move the Label range handling text. It would be moved to a separate document in future.

New requirements for PCEP extension regarding Label Range:-

- Path Computation Client (PCC) supporting this draft MAY support a capability to communicate local label range or global label range or both to PCE.

- Path Computation Element (PCE) supporting this draft MAY have the capability to negotiate a global label range for a group of clients and communicate the final global label range to PCC.

Two new flags are added in PCECC-CAPABILITY TLV (Section 7.1.1) to support PCECC:-

L (LOCAL-LABEL-RANGE-CAPABILITY - 1 bit): If set to 1 by a PCEP speaker, it indicates that the PCEP speaker is capable for local label range reservation.

G (GLOBAL-LABEL-RANGE-CAPABILITY - 1 bit): If set to 1 by a PCEP speaker, it indicates that the PCEP speaker capable for global label range reservation.

A new PCEP messages to support PCECC:-

(PCLRResv): a PCEP message sent by a PCC to a PCE to ask for the label range reservation or a PCE to a PCC to send the reserved label range. The PCLRResv message described in Appendix B.2.

B.1. Label Range Reservation

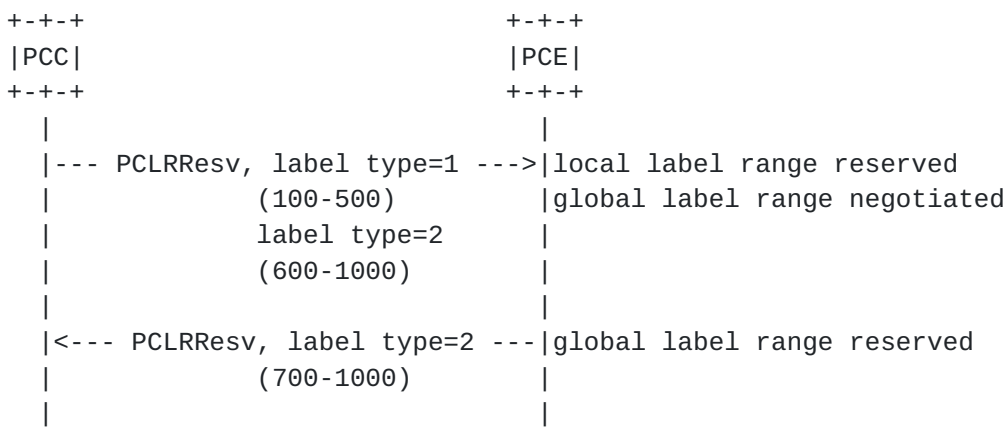
After PCEP initial state synchronization, the label range is reserved.

If L flag is advertised in OPEN Object by PCEP speakers, a PCC reserves a local label range and is communicated using PCLRResv message to a PCE. The PCE maintains the local label range of each node and further during LSP setup, a label is assigned to each node from the corresponding local label range reserved.

If G flag is advertised in OPEN Object by PCEP speakers,a PCC reserves a global label range and is advertised in PCLRResv message to a PCE. The PCE MAY negotiate and reserves the global label range and also sends the negotiated global label range in PCLRResv message to the PCC. Please refer [I-D.li-mpls-global-label-framework] for MPLS global label allocation.

A PCC MUST send PCLRResv message immediately after the initial LSP synchronization completion. A PCE SHOULD NOT send PCLabelUpd message to a PCC before PCLRResv message is received. If the PCC has received PCLabelUpd message and has not initiated label range reservation, it SHOULD send a PCErr message with Error-type=[TBD] (label range not reserved) and Error-value=[TBD].

The label range reservation sequence is shown below.



[Editor's Note: This section of the document would be updated with more details about Label Block Negotiation, Reservation, Adjustment etc in a future revision of the document.]

B.2. The PCLRResv message

A Label Range Reservation message (also referred to as PCLRResv message) is a PCEP message sent by a PCC to a PCE for the reservation of label range or by PCE to PCC to send reserved label range for the network. The Message-Type field of the PCEP common header for the PCLRResv message is set to [TBD].

The format of a PCLRResv message is as follows:

```
PCLRResv Message>::= <Common Header>
                           <label-range>
   Where:

    <label-range> ::= <SRP>
                      <labelrange-list>

   Where
<labelrange-list>::=<LABEL-RANGE>[<labelrange-list>]
```

There are two mandatory objects that MUST be included within each <label-range> in the PCLRResv message: the SRP Object and LABEL-RANGE object.

SRP object is defined in [I-D.ietf-pce-stateful-pce] and this document extends the use of SRP object in PCLRResv message. If the SRP object is missing, the receiving PCE MUST send a PCErr message

with Error-type=6 (Mandatory Object missing) and Error-value=10 (SRP object missing).

PCC generates the value of SRP-ID-number in SRP object of PCLRResv message send to a PCE. The PCE MUST include the same SRP-ID-number in SRP object of PCLRResv message sent to the PCC in response to PCLRResv message.

LABEL-RANGE object is defined in Appendix B.3. If the LABEL-RANGE object is missing, the receiving PCE MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=[TBD] (Label object missing).

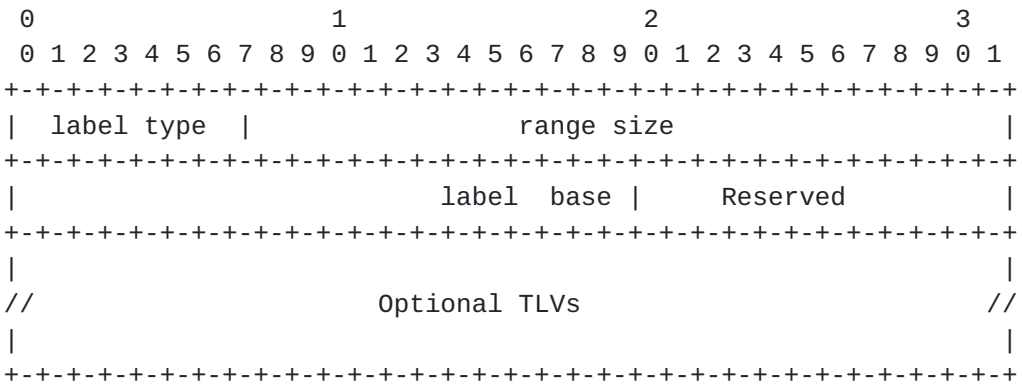
[Editor's Note: This section of the document would be updated with more details about Label Block Negotiation, Reservation, Adjustment etc in a future revision of the document.]

B.3. LABEL-RANGE Object

The LABEL-RANGE object MUST be carried within PCLRResv message. The LABEL-RANGE object is used to carry the label range information based on the label type.

LABEL-RANGE Object-Class is TBD.

LABEL-RANGE Object-Type is 1.



label type (8 bit): The values defined for label type are label type 1 specifies the local label. It means the label range is non negotiable. label type 2 specifies the global label. It means the label range is negotiable. Refer [I-D.li-mpls-global-label-framework] for global label.

Range size (24 bit): It specifies the size of label range.

Label base (20 bit): It specifies the minimum label of label range.

Reserved (12 bit): Set to zero while sending, ignored on receive.

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

Chao Zhou
Cisco Systems

EMail: choa.zhou@cisco.com