Internet Engineering Task Force (IETF)
Request for Comments: 8231
Category: Standards Track
ISSN: 2070-1721

E. Crabbe
Oracle
I. Minei
Google, Inc.
J. Medved
Cisco Systems, Inc.
R. Varga
Pantheon Technologies SRO

September 2017

Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE

Abstract

The Path Computation Element Communication Protocol (PCEP) provides mechanisms for Path Computation Elements (PCEs) to perform path computations in response to Path Computation Client (PCC) requests.

Although PCEP explicitly makes no assumptions regarding the information available to the PCE, it also makes no provisions for PCE control of timing and sequence of path computations within and across PCEP sessions. This document describes a set of extensions to PCEP to enable stateful control of MPLS-TE and GMPLS Label Switched Paths (LSPs) via PCEP.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc8231.

Table of Contents

1. Introduction

[RFC5440] describes the Path Computation Element Communication Protocol (PCEP). PCEP defines the communication between a Path Computation Client (PCC) and a Path Computation Element (PCE), or between PCEs, enabling computation of Multiprotocol Label Switching (MPLS) for Traffic Engineering Label Switched Path (TE LSP) characteristics. Extensions for support of Generalized MPLS (GMPLS) in PCEP are defined in [PCEP-GMPLS].

This document specifies a set of extensions to PCEP to enable stateful control of LSPs within and across PCEP sessions in compliance with [RFC4657]. It includes mechanisms to effect Label Switched Path (LSP) State Synchronization between PCCs and PCEs, delegation of control over LSPs to PCEs, and PCE control of timing and sequence of path computations within and across PCEP sessions.

Extensions to permit the PCE to drive creation of an LSP are defined in [PCE-Init-LSP], which specifies PCE-initiated LSP creation.

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

This document uses the following terms defined in [RFC5440]: PCC, PCE, PCEP Peer, and PCEP speaker.

This document uses the following terms defined in [RFC4655]: Traffic Engineering Database (TED).

This document uses the following terms defined in [RFC3031]: LSP.

This document uses the following terms defined in [RFC8051]: Stateful PCE, Passive Stateful PCE, Active Stateful PCE, Delegation, and LSP State Database.

The following terms are defined in this document:

Revocation: an operation performed by a PCC on a previously delegated LSP. Revocation revokes the rights granted to the PCE in the delegation operation.

Redelegation Timeout Interval: the period of time a PCC waits for, when a PCEP session is terminated, before revoking LSP delegation to a PCE and attempting to redelegate LSPs associated with the terminated PCEP session to an alternate PCE. The Redelegation Timeout Interval is a PCC-local value that can be either operator configured or dynamically computed by the PCC based on local policy.

State Timeout Interval: the period of time a PCC waits for, when a PCEP session is terminated, before flushing LSP state associated with that PCEP session and reverting to operator-defined default parameters or behaviors. The State Timeout Interval is a PCC-local value that can be either operator configured or dynamically computed by the PCC based on local policy.

LSP State Report: an operation to send LSP state (operational/administrative status, LSP attributes configured at the PCC and set by a PCE, etc.) from a PCC to a PCE.

LSP Update Request: an operation where an Active Stateful PCE requests a PCC to update one or more attributes of an LSP and to re-signal the LSP with updated attributes.

SRP-ID-number: a number used to correlate errors and LSP State Reports to LSP Update Requests. It is carried in the Stateful PCE Request Parameter (SRP) object described in Section 7.2.

Within this document, PCEP communications are described through PCC-PCE relationships. The PCE architecture also supports PCE-PCE communication, by having the requesting PCE fill the role of a PCC, as usual.

The message formats in this document are specified using Routing Backus-Naur Format (RBNF) encoding as specified in [RFC5511].

3. Motivation and Objectives for Stateful PCE

3.1. Motivation

[RFC8051] presents several use cases, demonstrating scenarios that benefit from the deployment of a stateful PCE. The scenarios apply equally to MPLS-TE and GMPLS deployments.

3.1.1. Background

Traffic engineering has been a goal of the MPLS architecture since its inception [RFC2702] [RFC3031] [RFC3346]. In the traffic engineering system provided by [RFC3209], [RFC3630], and [RFC5305],

information about network resources utilization is only available as total reserved capacity by the traffic class on a per-interface basis; individual LSP state is available only locally on each Label Edge Router (LER) for its own LSPs. In most cases, this makes good sense, as distribution and retention of total LSP state for all LERs within in the network would be prohibitively costly.

Unfortunately, this visibility in terms of global LSP state may result in a number of issues for some demand patterns, particularly within a common setup and hold priority. This issue affects online traffic engineering systems.

A sufficiently over-provisioned system will by definition have no issues routing its demand on the shortest path. However, lowering the degree to which network over-provisioning is required in order to run a healthy, functioning network is a clear and explicit promise of MPLS architecture. In particular, it has been a goal of MPLS to provide mechanisms to alleviate congestion scenarios in which "traffic streams are inefficiently mapped onto available resources; causing subsets of network resources to become over-utilized while others remain underutilized" [RFC2702].

3.1.2. Why a Stateful PCE?

[RFC4655] defines a stateful PCE to be one in which the PCE maintains "strict synchronization between the PCE and not only the network states (in term of topology and resource information), but also the set of computed paths and reserved resources in use in the network." [RFC4655] also expressed a number of concerns with regard to a stateful PCE, specifically:

- Any reliable synchronization mechanism would result in significant control-plane overhead
- Out-of-band TED synchronization would be complex and prone to race conditions
- Path calculations incorporating total network state would be highly complex

In general, stress on the control plane will be directly proportional to the size of the system being controlled and the tightness of the control loop and indirectly proportional to the amount of over-provisioning in terms of both network capacity and reservation overhead.

Despite these concerns in terms of implementation complexity and scalability, several TE algorithms exist today that have been demonstrated to be extremely effective in large TE systems, providing both rapid convergence and significant benefits in terms of optimality of resource usage [MXMN-TE]. All of these systems share at least two common characteristics: the requirement for both global visibility of a flow (or in this case, a TE LSP) state and for ordered control of path reservations across devices within the system being controlled. While some approaches have been suggested in order to remove the requirements for ordered control (see [MPLS-PC]), these approaches are highly dependent on traffic distribution and do not allow for multiple simultaneous LSP priorities representing Diffserv classes.

The use cases described in [RFC8051] demonstrate a need for visibility into global inter-PCC LSP state in PCE path computations and for PCE control of sequence and timing in altering LSP path characteristics within and across PCEP sessions.

3.1.3. Protocol vs. Configuration

Note that existing configuration tools and protocols can be used to set LSP state, such as a Command Line Interface (CLI) tool. However, this solution has several shortcomings:

- Scale & Performance: configuration operations often have transactional semantics that are typically heavyweight and often require processing of additional configuration portions beyond the state being directly acted upon, with corresponding cost in CPU cycles, negatively impacting both PCC stability LSP Update rate capacity.

- Security: when a PCC opens a configuration channel allowing a PCE to send configuration, a malicious PCE may take advantage of this ability to take over the PCC. In contrast, the PCEP extensions described in this document only allow a PCE control over a very limited set of LSP attributes.

- Interoperability: each vendor has a proprietary information model for configuring LSP state, which limits interoperability of a stateful PCE with PCCs from different vendors. The PCEP extensions described in this document allow for a common information model for LSP state for all vendors.

- Efficient State Synchronization: configuration channels may be heavyweight and unidirectional; therefore, efficient State Synchronization between a PCC and a PCE may be a problem.

3.2. Objectives

The objectives for the protocol extensions to support stateful PCE described in this document are as follows:

- Allow a single PCC to interact with a mix of stateless and stateful PCEs simultaneously using the same protocol, i.e., PCEP.
- Support efficient LSP State Synchronization between the PCC and one or more active or passive stateful PCEs.
- Allow a PCC to delegate control of its LSPs to an active stateful PCE such that a given LSP is under the control of a single PCE at any given time.
  - A PCC may revoke this delegation at any time during the lifetime of the LSP. If LSP delegation is revoked while the PCEP session is up, the PCC MUST notify the PCE about the revocation.
  - A PCE may return an LSP delegation at any point during the lifetime of the PCEP session. If LSP delegation is returned by the PCE while the PCEP session is up, the PCE MUST notify the PCC about the returned delegation.
- Allow a PCE to control computation timing and update timing across all LSPs that have been delegated to it.
- Enable uninterrupted operation of a PCC's LSPs in the event of a PCE failure or while control of LSPs is being transferred between PCEs.

4. New Functions to Support Stateful PCEs

Several new functions are required in PCEP to support stateful PCEs. A function can be initiated either from a PCC towards a PCE (C-E) or from a PCE towards a PCC (E-C). The new functions are:

Capability advertisement (E-C,C-E): both the PCC and the PCE must announce during PCEP session establishment that they support PCEP Stateful PCE extensions defined in this document.

LSP State Synchronization (C-E): after the session between the PCC and a stateful PCE is initialized, the PCE must learn the state of a PCC's LSPs before it can perform path computations or update LSP attributes in a PCC.

LSP Update Request (E-C): a PCE requests modification of attributes on a PCC's LSP.

LSP State Report (C-E): a PCC sends an LSP State Report to a PCE whenever the state of an LSP changes.

LSP control delegation (C-E,E-C): a PCC grants to a PCE the right to update LSP attributes on one or more LSPs; the PCE becomes the authoritative source of the LSP's attributes as long as the delegation is in effect (see Section 5.7); the PCC may withdraw the delegation or the PCE may give up the delegation at any time.

Similarly to [RFC5440], no assumption is made about the discovery method used by a PCC to discover a set of PCEs (e.g., via static configuration or dynamic discovery) and on the algorithm used to select a PCE.

5. Overview of Protocol Extensions

5.1. LSP State Ownership

In PCEP (defined in [RFC5440]), LSP state and operation are under the control of a PCC (a PCC may be a Label Switching Router (LSR) or a management station). Attributes received from a PCE are subject to PCC's local policy. The PCEP extensions described in this document do not change this behavior.

An active stateful PCE may have control of a PCC's LSPs that were delegated to it, but the LSP state ownership is retained by the PCC. In particular, in addition to specifying values for LSP's attributes, an active stateful PCE also decides when to make LSP modifications.

Retaining LSP state ownership on the PCC allows for:

- a PCC to interact with both stateless and stateful PCEs at the same time
- a stateful PCE to only modify a small subset of LSP parameters, i.e., to set only a small subset of the overall LSP state; other parameters may be set by the operator, for example, through CLI commands
- a PCC to revert delegated LSP to an operator-defined default or to delegate the LSPs to a different PCE, if the PCC gets disconnected from a PCE with currently delegated LSPs

5.2. New Messages

In this document, we define the following new PCEP messages:

Path Computation State Report (PCRpt): a PCEP message sent by a PCC to a PCE to report the status of one or more LSPs. Each LSP State Report in a PCRpt message MAY contain the actual LSP's path, bandwidth, operational and administrative status, etc. An LSP Status Report carried on a PCRpt message is also used in delegation or revocation of control of an LSP to/from a PCE. The PCRpt message is described in Section 6.1.

Path Computation Update Request (PCUpd): a PCEP message sent by a PCE to a PCC to update LSP parameters, on one or more LSPs. Each LSP Update Request on a PCUpd message MUST contain all LSP parameters that a PCE wishes to be set for a given LSP. An LSP Update Request carried on a PCUpd message is also used to return LSP delegations if at any point PCE no longer desires control of an LSP. The PCUpd message is described in Section 6.2.

The new functions defined in Section 4 are mapped onto the new messages as shown in the following table.

Function	Message
Capability Advertisement (E-C,C-E)	Open
State Synchronization (C-E)	PCRpt
LSP State Report (C-E)	PCRpt
LSP Control Delegation (C-E,E-C)	PCRpt, PCUpd
LSP Update Request (E-C)	PCUpd

Table 1: New Function to Message Mapping

5.3. Error Reporting

Error reporting is done using the procedures defined in [RFC5440] and reusing the applicable error types and error values of [RFC5440] wherever appropriate. The current document defines new error values for several error types to cover failures specific to stateful PCE.

5.4. Capability Advertisement

During the PCEP initialization phase, PCEP speakers (PCE or PCC) advertise their support of PCEP Stateful PCE extensions. A PCEP speaker includes the "STATEFUL-PCE-CAPABILITY TLV", described in Section 7.1.1, in the OPEN object to advertise its support for PCEP

Stateful PCE extensions. The STATEFUL-PCE-CAPABILITY TLV includes the 'LSP Update' flag that indicates whether the PCEP speaker supports LSP parameter updates.

The presence of the STATEFUL-PCE-CAPABILITY TLV in PCC's OPEN object indicates that the PCC is willing to send LSP State Reports whenever LSP parameters or operational status changes.

The presence of the STATEFUL-PCE-CAPABILITY TLV in PCE's OPEN message indicates that the PCE is interested in receiving LSP State Reports whenever LSP parameters or operational status changes.

The PCEP extensions for stateful PCEs MUST NOT be used if one or both PCEP speakers have not included the STATEFUL-PCE-CAPABILITY TLV in their respective OPEN message. If the PCEP speaker on the PCC supports the extensions of this specification but did not advertise this capability, then upon receipt of a PCUpd message from the PCE, it MUST generate a PCEP Error (PCErr) with Error-type=19 (Invalid Operation) and error-value 2 (Attempted LSP Update Request if the stateful PCE capability was not advertised)(see Section 8.5), and it SHOULD terminate the PCEP session. If the PCEP Speaker on the PCE supports the extensions of this specification but did not advertise this capability, then upon receipt of a PCRpt message from the PCC, it MUST generate a PCErr with Error-type=19 (Invalid Operation) and error-value 5 (Attempted LSP State Report if stateful PCE capability was not advertised) (see Section 8.5), and it SHOULD terminate the PCEP session.

LSP delegation and LSP Update operations defined in this document may only be used if both PCEP speakers set the LSP-UPDATE-CAPABILITY flag in the STATEFUL-PCE-CAPABILITY TLV to 'Updates Allowed (U flag = 1)'. If this is not the case and LSP delegation or LSP Update operations are attempted, then a PCErr with Error-type=19 (Invalid Operation) and error-value 1 (Attempted LSP Update Request for a non-delegated LSP) (see Section 8.5) MUST be generated. Note that, even if one of the PCEP speakers does not set the LSP-UPDATE-CAPABILITY flag in its STATEFUL-PCE-CAPABILITY TLV, a PCE can still operate as a passive stateful PCE by accepting LSP State Reports from the PCC in order to build and maintain an up-to-date view of the state of the PCC's LSPs.

5.5. IGP Extensions for Stateful PCE Capabilities Advertisement

When PCCs are LSRs participating in the IGP (OSPF or IS-IS), and PCEs are either LSRs or servers also participating in the IGP, an effective mechanism for PCE discovery within an IGP routing domain consists of utilizing IGP advertisements. Extensions for the advertisement of PCE Discovery Information are defined for OSPF and for IS-IS in [RFC5088] and [RFC5089], respectively.

The PCE-CAP-FLAGS sub-TLV, defined in [RFC5089], is an optional sub-TLV used to advertise PCE capabilities. It MAY be present within the PCE Discovery (PCED) sub-TLV carried by OSPF or IS-IS. [RFC5088] and [RFC5089] provide the description and processing rules for this sub-TLV when carried within OSPF and IS-IS, respectively.

The format of the PCE-CAP-FLAGS sub-TLV is included below for easy reference:

Type: 5

Length: Multiple of 4.

Value: This contains an array of units of 32-bit flags with the most significant bit as 0. Each bit represents one PCE capability.

PCE capability bits are defined in [RFC5088]. This document defines new capability bits for the stateful PCE as follows:

Bit	Capability
11	Active stateful PCE capability
12	Passive stateful PCE capability

Note that while active and passive stateful PCE capabilities may be advertised during discovery, PCEP speakers that wish to use stateful PCEP MUST negotiate stateful PCEP capabilities during PCEP session setup, as specified in the current document. A PCC MAY initiate stateful PCEP capability negotiation at PCEP session setup even if it did not receive any IGP PCE capability advertisements.

5.6. State Synchronization

The purpose of State Synchronization is to provide a checkpoint-in-time state replica of a PCC's LSP state in a PCE. State Synchronization is performed immediately after the initialization phase [RFC5440].

During State Synchronization, a PCC first takes a snapshot of the state of its LSPs, then it sends the snapshot to a PCE in a sequence of LSP State Reports. Each LSP State Report sent during State Synchronization has the SYNC flag in the LSP object set to 1. The set of LSPs for which state is synchronized with a PCE is determined by the PCC's local configuration (see more details in Section 9.1) and MAY also be determined by stateful PCEP capabilities defined in other documents, such as [RFC8232].

The end of the synchronization marker is a PCRpt message with the SYNC flag set to 0 for an LSP object with PLSP-ID equal to the reserved value 0 (see Section 7.3). In this case, the LSP object SHOULD NOT include the SYMBOLIC-PATH-NAME TLV and SHOULD include the LSP-IDENTIFIERS TLV with the special value of all zeroes. The PCRpt message MUST include an empty Explicit Route Object (ERO) as its intended path and SHOULD NOT include the optional Record Route Object (RRO) for its actual path. If the PCC has no state to synchronize, it SHOULD only send the end of the synchronization marker.

A PCE SHOULD NOT send PCUpd messages to a PCC before State Synchronization is complete. A PCC SHOULD NOT send PCReq messages to a PCE before State Synchronization is complete. This is to allow the PCE to get the best possible view of the network before it starts computing new paths.

Either the PCE or the PCC MAY terminate the session using the PCEP session termination procedures during the synchronization phase. If the session is terminated, the PCE MUST clean up the state it received from this PCC. The session re-establishment MUST be re-attempted per the procedures defined in [RFC5440], including use of a backoff timer.

If the PCC encounters a problem that prevents it from completing the LSP State Synchronization, it MUST send a PCErr message with error-type 20 (LSP State Synchronization Error) and error-value 5 (indicating an internal PCC error) to the PCE and terminate the session.

The PCE does not send positive acknowledgments for properly received synchronization messages. It MUST respond with a PCErr message with Error-type=20 (LSP State Synchronization Error) and error-value 1 (indicating an error in processing the PCRpt) (see Section 8.5) if it encounters a problem with the LSP State Report it received from the PCC, and it MUST terminate the session.

A PCE implementing a limit on the resources a single PCC can occupy MUST send a PCEP Notify (PCNtf) message with Notification Type 4 (Stateful PCE resource limit exceeded) and Notification Value 1 (Entering resource limit exceeded state) in response to the PCRpt message triggering this condition in the synchronization phase and MUST terminate the session.

The successful State Synchronization sequence is shown in Figure 1.

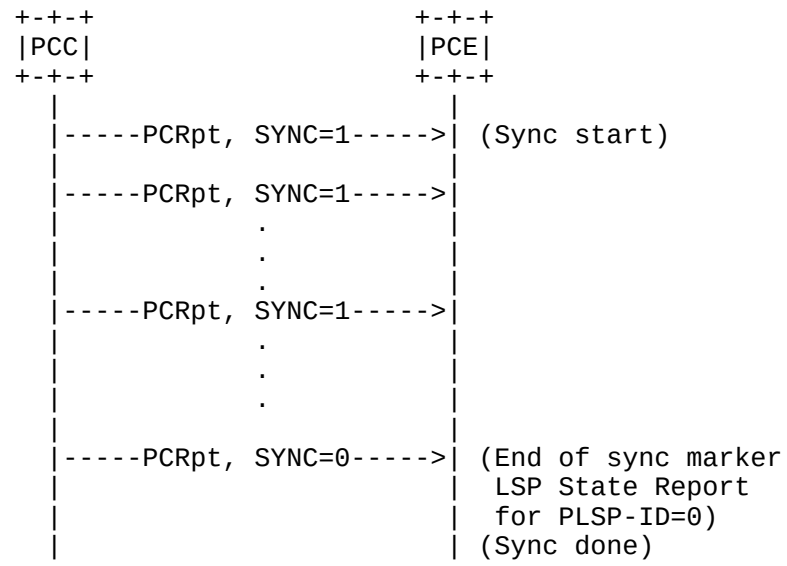

Figure 1: Successful State Synchronization

The sequence where the PCE fails during the State Synchronization phase is shown in Figure 2.

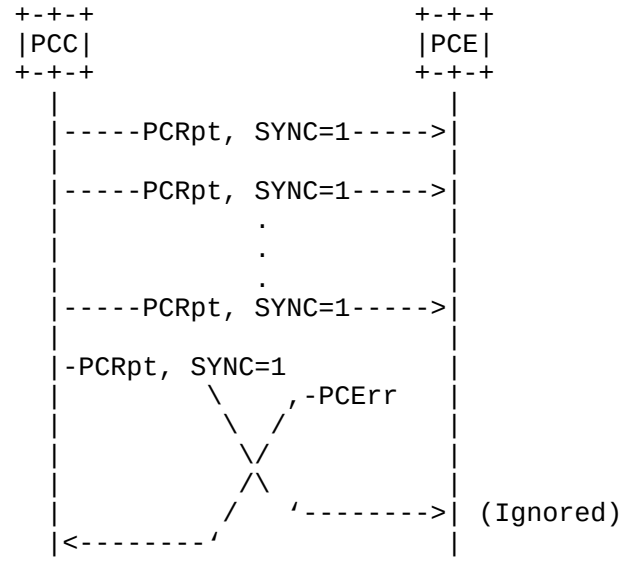

Figure 2: Failed State Synchronization (PCE Failure)

The sequence where the PCC fails during the State Synchronization phase is shown in Figure 3.

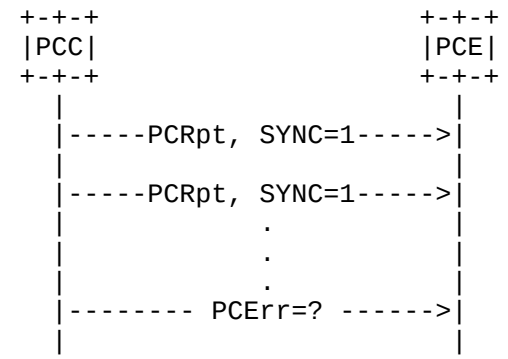

Figure 3: Failed State Synchronization (PCC Failure)

Optimizations to the synchronization procedures and alternate mechanisms of providing the synchronization function are outside the scope of this document and are discussed elsewhere (see [RFC8232]).

5.7. LSP Delegation

If during capability advertisement both the PCE and the PCC have indicated that they support LSP Update, then the PCC may choose to grant the PCE a temporary right to update (a subset of) LSP attributes on one or more LSPs. This is called "LSP delegation", and it MAY be performed at any time after the initialization phase, including during the State Synchronization phase.

A PCE MAY return an LSP delegation at any time if it no longer wishes to update the LSP's state. A PCC MAY revoke an LSP delegation at any time. Delegation, Revocation, and Return are done individually for each LSP.

In the event of a delegation being rejected or returned by a PCE, the PCC SHOULD react based on local policy. It can, for example, either retry delegating to the same PCE using an exponentially increasing timer or delegate to an alternate PCE.

5.7.1. Delegating an LSP

A PCC delegates an LSP to a PCE by setting the Delegate flag in the LSP State Report to 1. If the PCE does not accept the LSP delegation, it MUST immediately respond with an empty LSP Update Request that has the Delegate flag set to 0. If the PCE accepts the LSP delegation, it MUST set the Delegate flag to 1 when it sends an

LSP Update Request for the delegated LSP (note that this may occur at a later time). The PCE MAY also immediately acknowledge a delegation by sending an empty LSP Update Request that has the Delegate flag set to 1.

The delegation sequence is shown in Figure 4.

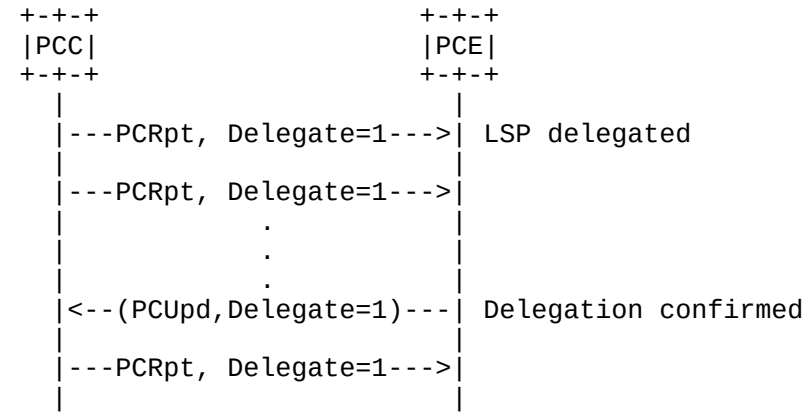

Figure 4: Delegating an LSP

Note that for an LSP to remain delegated to a PCE, the PCC MUST set the Delegate flag to 1 on each LSP State Report sent to the PCE.

5.7.2. Revoking a Delegation

5.7.2.1. Explicit Revocation

When a PCC decides that a PCE is no longer permitted to modify an LSP, it revokes that LSP's delegation to the PCE. A PCC may revoke an LSP delegation at any time during the LSP's lifetime. A PCC revoking an LSP delegation MAY immediately remove the updated parameters provided by the PCE and revert to the operator-defined parameters, but to avoid traffic loss, it SHOULD do so in a make-before-break fashion. If the PCC has received but not yet acted on PCUpd messages from the PCE for the LSP whose delegation is being revoked, then it SHOULD ignore these PCUpd messages when processing the message queue. All effects of all messages for which processing started before the revocation took place MUST be allowed to complete, and the result MUST be given the same treatment as any LSP that had been previously delegated to the PCE (e.g., the state MAY immediately revert to the operator-defined parameters).

If a PCEP session with the PCE to which the LSP is delegated exists in the UP state during the revocation, the PCC MUST notify that PCE by sending an LSP State Report with the Delegate flag set to 0, as shown in Figure 5.

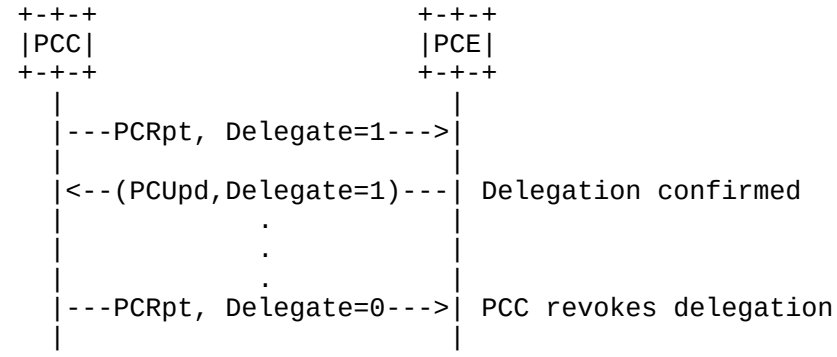

Figure 5: Revoking a Delegation

After an LSP delegation has been revoked, a PCE can no longer update an LSP's parameters; an attempt to update parameters of a non-delegated LSP will result in the PCC sending a PCErr message with Error-type=19 (Invalid Operation) and error-value 1 (Attempted LSP Update Request for a non-delegated LSP) (see Section 8.5).

5.7.2.2. Revocation on Redelegation Timeout

When a PCC's PCEP session with a PCE terminates unexpectedly, the PCC MUST wait the time interval specified in the Redelegation Timeout Interval before revoking LSP delegations to that PCE and attempting to redelegate LSPs to an alternate PCE. If a PCEP session with the original PCE can be re-established before the Redelegation Timeout Interval timer expires, LSP delegations to the PCE remain intact.

Likewise, when a PCC's PCEP session with a PCE terminates unexpectedly, and the PCC does not succeed in redelegating its LSPs, the PCC MUST wait for the State Timeout Interval before flushing any LSP state associated with that PCE. Note that the State Timeout Interval timer may expire before the PCC has redelegated the LSPs to another PCE, for example, if a PCC is not connected to any active stateful PCE or if no connected active stateful PCE accepts the delegation. In this case, the PCC MUST flush any LSP state set by the PCE upon expiration of the State Timeout Interval and revert to operator-defined default parameters or behaviors. This operation SHOULD be done in a make-before-break fashion.

The State Timeout Interval MUST be greater than or equal to the Redelegation Timeout Interval and MAY be set to infinity (meaning that until the PCC specifically takes action to change the parameters set by the PCE, they will remain intact).

5.7.3. Returning a Delegation

In order to keep a delegation, a PCE MUST set the Delegate flag to 1 on each LSP Update Request sent to the PCC. A PCE that no longer wishes to update an LSP's parameters SHOULD return the LSP delegation back to the PCC by sending an empty LSP Update Request that has the Delegate flag set to 0. If a PCC receives an LSP Update Request with the Delegate flag set to 0 (whether the LSP Update Request is empty or not), it MUST treat this as a delegation return.

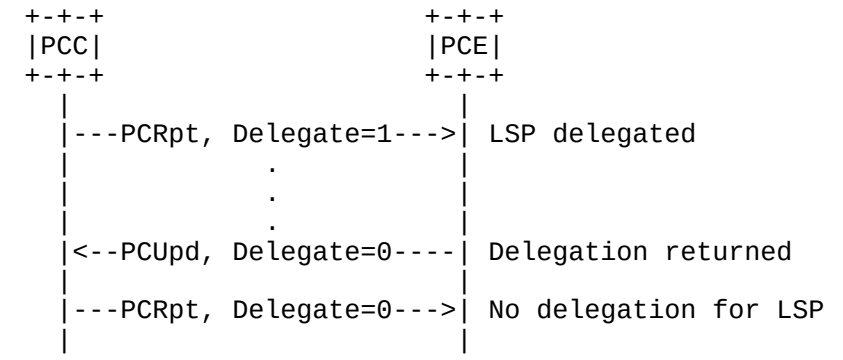

Figure 6: Returning a Delegation

If a PCC cannot delegate an LSP to a PCE (for example, if a PCC is not connected to any active stateful PCE or if no connected active stateful PCE accepts the delegation), the LSP delegation on the PCC will timeout within a configurable Redelegation Timeout Interval, and the PCC MUST flush any LSP state set by a PCE at the expiration of the State Timeout Interval and revert to operator-defined default parameters or behaviors.

5.7.4. Redundant Stateful PCEs

In a redundant configuration where one PCE is backing up another PCE, the backup PCE may have only a subset of the LSPs in the network delegated to it. The backup PCE does not update any LSPs that are not delegated to it. In order to allow the backup to operate in a hot-standby mode and avoid the need for State Synchronization in case the primary fails, the backup receives all LSP State Reports from a PCC. When the primary PCE for a given LSP set fails, after expiry of the Redelegation Timeout Interval, the PCC SHOULD delegate to the

redundant PCE all LSPs that had been previously delegated to the failed PCE. Assuming that the State Timeout Interval had been configured to be greater than the Redelegation Timeout Interval (as MANDATORY), and assuming that the primary and redundant PCEs take similar decisions, this delegation change will not cause any changes to the LSP parameters.

5.7.5. Redelegation on PCE Failure

On failure, the goal is to: 1) avoid any traffic loss on the LSPs that were updated by the PCE that crashed, 2) minimize the churn in the network in terms of ownership of the LSPs, 3) not leave any "orphan" (undelegated) LSPs, and 4) be able to control when the state that was set by the PCE can be changed or purged. The values chosen for the Redelegation Timeout and State Timeout values affect the ability to accomplish these goals.

This section summarizes the behavior with regards to LSP delegation and LSP state on a PCE failure.

If the PCE crashes but recovers within the Redelegation Timeout, both the delegation state and the LSP state are kept intact.

If the PCE crashes but does not recover within the Redelegation Timeout, the delegation state is returned to the PCC. If the PCC can redelegate the LSPs to another PCE, and that PCE accepts the delegations, there will be no change in LSP state. If the PCC cannot redelegate the LSPs to another PCE, then upon expiration of the State Timeout Interval, the state set by the PCE is removed and the LSP reverts to operator-defined parameters, which may cause a change in the LSP state. Note that an operator may choose to use an infinite State Timeout Interval if he wishes to maintain the PCE state indefinitely. Note also that flushing the state should be implemented using make-before-break to avoid traffic loss.

If there is a standby PCE, the Redelegation Timeout may be set to 0 through policy on the PCC, causing the LSPs to be redelegated immediately to the PCC, which can delegate them immediately to the standby PCE. Assuming that the PCC can redelegate the LSP to the standby PCE within the State Timeout Interval, and assuming the standby PCE takes similar decisions as the failed PCE, the LSP state will be kept intact.

5.8. LSP Operations

5.8.1. Passive Stateful PCE Path Computation Request/Response

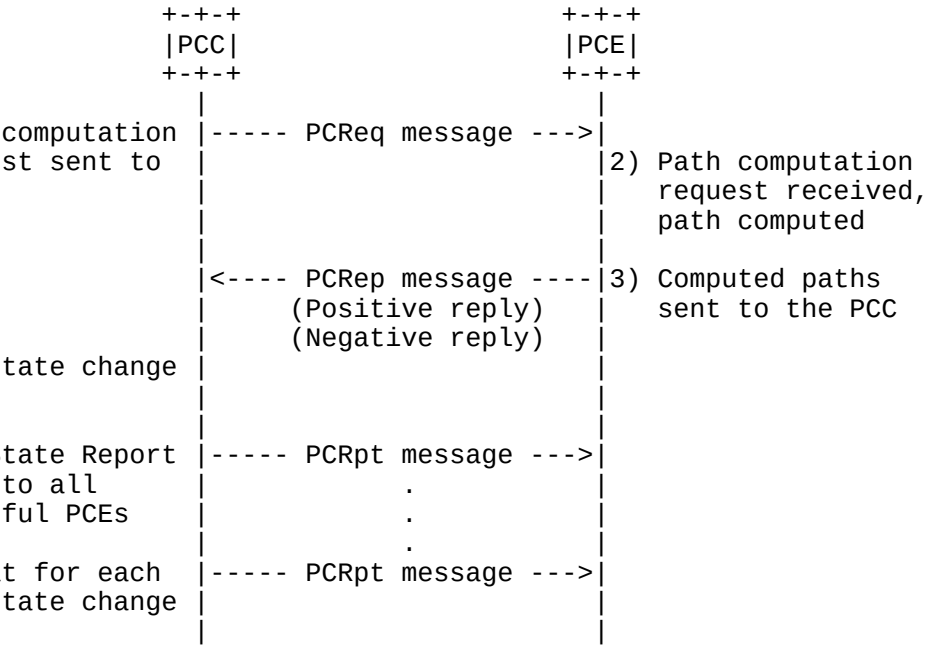

Figure 7: Passive Stateful PCE Path Computation Request/Response

Once a PCC has successfully established a PCEP session with a passive stateful PCE and the PCC's LSP state is synchronized with the PCE (i.e., the PCE knows about all of the PCC's existing LSPs), if an event is triggered that requires the computation of a set of paths, the PCC sends a path computation request to the PCE ([RFC5440], Section 4.2.3). The PCReq message MAY contain the LSP object to identify the LSP for which the path computation is requested.

Upon receiving a path computation request from a PCC, the PCE triggers a path computation and returns either a positive or a negative reply to the PCC ([RFC5440], Section 4.2.4).

Upon receiving a positive path computation reply, the PCC receives a set of computed paths and starts to set up the LSPs. For each LSP, it MAY send an LSP State Report carried on a PCRpt message to the PCE, indicating that the LSP's status is "Going-up".

Once an LSP is up or active, the PCC MUST send an LSP State Report carried on a PCRpt message to the PCE, indicating that the LSP's status is 'Up' or 'Active', respectively. If the LSP could not be set up, the PCC MUST send an LSP State Report indicating that the LSP is 'Down' and stating the cause of the failure. Note that due to timing constraints, the LSP status may change from 'Going-up' to 'Up' (or 'Down') before the PCC has had a chance to send an LSP State Report indicating that the status is 'Going-up'. In such cases, the PCC MAY choose to only send the PCRpt indicating the latest status ('Active', 'Up', or 'Down').

Upon receiving a negative reply from a PCE, a PCC MAY resend a modified request or take any other appropriate action. For each requested LSP, it SHOULD also send an LSP State Report carried on a PCRpt message to the PCE, indicating that the LSP's status is 'Down'.

There is no direct correlation between PCRep and PCRpt messages. For a given LSP, multiple LSP State Reports will follow a single PCRep message, as a PCC notifies a PCE of the LSP's state changes.

A PCC MUST send each LSP State Report to each stateful PCE that is connected to the PCC.

Note that a single PCRpt message MAY contain multiple LSP State Reports.

The passive stateful model for stateful PCEs is described in [RFC4655], Section 6.8.

5.8.2. Switching from Passive Stateful to Active Stateful

This section deals with the scenario of an LSP transitioning from a passive stateful to an active stateful mode of operation. When the LSP has no working path, prior to delegating the LSP, the PCC MUST first use the procedure defined in Section 5.8.1 to request the initial path from the PCE. This is required because the action of delegating the LSP to a PCE using a PCRpt message is not an explicit request to the PCE to compute a path for the LSP. The only explicit way for a PCC to request a path from the PCE is to send a PCReq message. The PCRpt message MUST NOT be used by the PCC to attempt to request a path from the PCE.

When the LSP is delegated after its setup, it may be useful for the PCC to communicate to the PCE the locally configured intended configuration parameters, so that the PCE may reuse them in its computations. Such parameters MAY be acquired through an out-of-band channel, or MAY be communicated in the PCRpt message delegating the LSPs, by including them as part of the intended-attribute-list as

explained in Section 6.1. An implementation MAY allow policies on the PCC to determine the configuration parameters to be sent to the PCE.

5.8.3. Active Stateful PCE LSP Update

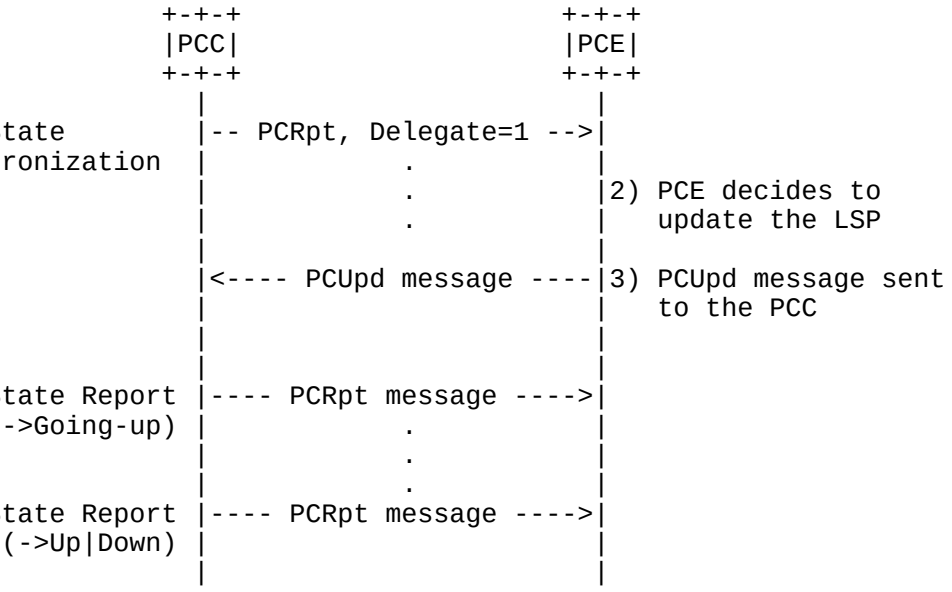

Figure 8: Active Stateful PCE

Once a PCC has successfully established a PCEP session with an active stateful PCE, the PCC's LSP state is synchronized with the PCE (i.e., the PCE knows about all of the PCC's existing LSPs). After LSPs have been delegated to the PCE, the PCE can modify LSP parameters of delegated LSPs.

To update an LSP, a PCE MUST send the PCC an LSP Update Request using a PCUpd message. The LSP Update Request contains a variety of objects that specify the set of constraints and attributes for the LSP's path. Each LSP Update Request MUST have a unique identifier, the SRP-ID-number, carried in the SRP object described in Section 7.2. The SRP-ID-number is used to correlate errors and state reports to LSP Update Requests. A single PCUpd message MAY contain multiple LSP Update Requests.

Upon receiving a PCUpd message, the PCC starts to set up LSPs specified in LSP Update Requests carried in the message. For each LSP, it MAY send an LSP State Report carried on a PCRpt message to the PCE, indicating that the LSP's status is 'Going-up'. If the PCC

decides that the LSP parameters proposed in the PCUpd message are unacceptable, it MUST report this error by including the LSP-ERROR-CODE TLV (Section 7.3.3) with LSP error-value="Unacceptable parameters" in the LSP object in the PCRpt message to the PCE. Based on local policy, it MAY react further to this error by revoking the delegation. If the PCC receives a PCUpd message for an LSP object identified with a PLSP-ID that does not exist on the PCC, it MUST generate a PCErr with Error-type=19 (Invalid Operation), error-value 3, (Attempted LSP Update Request for an LSP identified by an unknown PSP-ID) (see Section 8.5).

Once an LSP is up, the PCC MUST send an LSP State Report (PCRpt message) to the PCE, indicating that the LSP's status is 'Up'. If the LSP could not be set up, the PCC MUST send an LSP State Report indicating that the LSP is 'Down' and stating the cause of the failure. A PCC MAY compress LSP State Reports to only reflect the most up to date state, as discussed in the previous section.

A PCC MUST send each LSP State Report to each stateful PCE that is connected to the PCC.

PCErr and PCRpt messages triggered as a result of a PCUpd message MUST include the SRP-ID-number from the PCUpd. This provides correlation of requests and errors and acknowledgement of state processing. The PCC MAY compress the state when processing PCUpd. In this case, receipt of a higher SRP-ID-number implicitly acknowledges processing all the updates with a lower SRP-ID-number for the specific LSP (as per Section 7.2).

A PCC MUST NOT send to any PCE a path computation request for a delegated LSP. Should the PCC decide it wants to issue a Path Computation Request on a delegated LSP, it MUST perform the Delegation Revocation procedure first.

5.9. LSP Protection

LSP protection and interaction with stateful PCE, as well as the extensions necessary to implement this functionality, will be discussed in a separate document.

5.10. PCEP Sessions

A permanent PCEP session MUST be established between a stateful PCE and the PCC. In the case of session failure, session re-establishment MUST be re-attempted per the procedures defined in [RFC5440].

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. For each PCEP message type, a set of rules is defined that specifies the set of objects that the message can carry.

6.1. The PCRpt Message

A Path Computation LSP State Report message (also referred to as a PCRpt message) is a PCEP message sent by a PCC to a PCE to report the current state of an LSP. A PCRpt message can carry more than one LSP State Reports. A PCC can send an LSP State Report either in response to an LSP Update Request from a PCE or asynchronously when the state of an LSP changes. The Message-Type field of the PCEP common header for the PCRpt message is 10.

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> ::= [<SRP>]
                      <LSP>
                      <path>
 Where:
   <path>::= <intended-path>
             [<actual-attribute-list><actual-path>]
             <intended-attribute-list>

   <actual-attribute-list>::=[<BANDWIDTH>]
                              [<metric-list>]
```

Where:

<intended-path> is represented by the ERO object defined in Section 7.9 of [RFC5440].

<actual-attribute-list> consists of the actual computed and signaled values of the <BANDWIDTH> and <metric-lists> objects defined in [RFC5440].

<actual-path> is represented by the RRO object defined in Section 7.10 of [RFC5440].

<intended-attribute-list> is the attribute-list defined in Section 6.5 of [RFC5440] and extended by PCEP extensions.

The SRP object (see Section 7.2) is OPTIONAL. If the PCRpt message is not in response to a PCupd message, the SRP object MAY be omitted.

When the PCC does not include the SRP object, the PCE MUST treat this as an SRP object with an SRP-ID-number equal to the reserved value 0x00000000. The reserved value 0x00000000 indicates that the state reported is not a result of processing a PCUpd message.

If the PCRpt message is in response to a PCUpd message, the SRP object MUST be included and the value of the SRP-ID-number in the SRP object MUST be the same as that sent in the PCUpd message that triggered the state that is reported. If the PCC compressed several PCUpd messages for the same LSP by only processing the one with the highest number, then it should use the SRP-ID-number of that request. No state compression is allowed for state reporting, e.g., PCRpt messages MUST NOT be pruned from the PCC's egress queue even if subsequent operations on the same LSP have been completed before the PCRpt message has been sent to the TCP stack. The PCC MUST explicitly report state changes (including removal) for paths it manages.

The LSP object (see Section 7.3) is REQUIRED, and it MUST be included in each LSP State Report on the PCRpt message. If the LSP object is missing, the receiving PCE MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value 8 (LSP object missing).

If the LSP transitioned to non-operational state, the PCC SHOULD include the LSP-ERROR-TLV (Section 7.3.3) with the relevant LSP Error Code to report the error to the PCE.

The intended path, represented by the ERO object, is REQUIRED. If the ERO object is missing, the receiving PCE MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value 9 (ERO object missing). The ERO may be empty if the PCE does not have a path for a delegated LSP.

The actual path, represented by the RRO object, SHOULD be included in a PCRpt by the PCC when the path is up or active, but it MAY be omitted if the path is down due to a signaling error or another failure.

The intended-attribute-list maps to the attribute-list in Section 6.5 of [RFC5440] and is used to convey the requested parameters of the LSP path. This is needed in order to support the switch from passive

to active stateful PCE as described in Section 5.8.2. When included as part of the intended-attribute-list, the meaning of the BANDWIDTH object is the requested bandwidth as intended by the operator. In this case, the BANDWIDTH Object-Type of 1 SHOULD be used. Similarly, to indicate a limiting constraint, the METRIC object SHOULD be included as part of the intended-attribute-list with the B flag set and with a specific metric value. To indicate the optimization metric, the METRIC object SHOULD be included as part of the intended-attribute-list with the B flag unset and the metric value set to zero. Note that the intended-attribute-list is optional and thus may be omitted. In this case, the PCE MAY use the values in the actual-attribute-list as the requested parameters for the path.

The actual-attribute-list consists of the actual computed and signaled values of the BANDWIDTH and METRIC objects defined in [RFC5440]. When included as part of the actual-attribute-list, Object-Type 2 [RFC5440] SHOULD be used for the BANDWIDTH object, and the C flag SHOULD be set in the METRIC object [RFC5440].

Note that the ordering of intended-path, actual-attribute-list, actual-path, and intended-attribute-list is chosen to retain compatibility with implementations of an earlier version of this standard.

A PCE may choose to implement a limit on the resources a single PCC can occupy. If a PCRpt is received that causes the PCE to exceed this limit, the PCE MUST notify the PCC using a PCNtf message with Notification Type 4 (Stateful PCE resource limit exceeded) and Notification Value 1 (Entering resource limit exceeded state), and it MUST terminate the session.

6.2. The PCUpd Message

A Path Computation LSP Update Request message (also referred to as PCUpd message) is a PCEP message sent by a PCE to a PCC to update attributes of an LSP. A PCUpd message can carry more than one LSP Update Request. The Message-Type field of the PCEP common header for the PCUpd message is 11.

The format of a PCUpd message is as follows:

```
<PCUpd Message> ::= <Common Header>
                    <update-request-list>
```

Where:

```
<update-request-list> ::= <update-request>[<update-request-list>]

<update-request> ::= <SRP>
                     <LSP>
                     <path>
```

Where:

```
<path>::= <intended-path><intended-attribute-list>
```

Where:

<intended-path> is represented by the ERO object defined in Section 7.9 of [RFC5440].

<intended-attribute-list> is the attribute-list defined in [RFC5440] and extended by PCEP extensions.

There are three mandatory objects that MUST be included within each LSP Update Request in the PCUpd message: the SRP object (see Section 7.2), the LSP object (see Section 7.3) and the ERO object (as defined in [RFC5440], which represents the intended path. 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). If the LSP object is missing, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=8 (LSP object missing). If the ERO object is missing, the receiving PCC MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=9 (ERO object missing).

The ERO in the PCUpd may be empty if the PCE cannot find a valid path for a delegated LSP. One typical situation resulting in this empty ERO carried in the PCUpd message is that a PCE can no longer find a strict SRLG-disjoint path for a delegated LSP after a link failure. The PCC SHOULD implement a local policy to decide the appropriate action to be taken: either tear down the LSP or revoke the delegation and use a locally computed path, or keep the existing LSP.

A PCC only acts on an LSP Update Request if permitted by the local policy configured by the network manager. Each LSP Update Request that the PCC acts on results in an LSP setup operation. An LSP Update Request MUST contain all LSP parameters that a PCE wishes to

be set for the LSP. A PCC MAY set missing parameters from locally configured defaults. If the LSP specified in the Update Request is already up, it will be re-signaled.

The PCC SHOULD minimize the traffic interruption and MAY use the make-before-break procedures described in [RFC3209] in order to achieve this goal. If the make-before-break procedures are used, two paths will briefly coexist. The PCC MUST send separate PCRpt messages for each, identified by the LSP-IDENTIFIERS TLV. When the old path is torn down after the head end switches over the traffic, this event MUST be reported by sending a PCRpt message with the LSP-IDENTIFIERS-TLV of the old path and the R bit set. The SRP-ID-number that the PCC associates with this PCRpt MUST be 0x00000000. Thus, a make-before-break operation will typically result in at least two PCRpt messages, one for the new path and one for the removal of the old path (more messages may be possible if intermediate states are reported).

If the path setup fails due to an RSVP signaling error, the error is reported to the PCE. The PCC will not attempt to re-signal the path until it is prompted again by the PCE with a subsequent PCUpd message.

A PCC MUST respond with an LSP State Report to each LSP Update Request it processed to indicate the resulting state of the LSP in the network (even if this processing did not result in changing the state of the LSP). The SRP-ID-number included in the PCRpt MUST match that in the PCUpd. A PCC MAY respond with multiple LSP State Reports to report LSP setup progress of a single LSP. In that case, the SRP-ID-number MUST be included for the first message; for subsequent messages, the reserved value 0x00000000 SHOULD be used.

Note that a PCC MUST process all LSP Update Requests -- for example, an LSP Update Request is sent when a PCE returns delegation or puts an LSP into non-operational state. The protocol relies on TCP for message-level flow control.

If the rate of PCUpd messages sent to a PCC for the same target LSP exceeds the rate at which the PCC can signal LSPs into the network, the PCC MAY perform state compression on its ingress queue. The compression algorithm is based on the fact that each PCUpd request contains the complete LSP state the PCE wishes to be set and works as follows: when the PCC starts processing a PCUpd message at the head of its ingress queue, it may search the queue forward for more recent PCUpd messages pertaining to that particular LSP, prune all but the latest one from the queue, and process only the last one as that request contains the most up-to-date desired state for the LSP. The PCC MUST NOT send PCRpt nor PCErr messages for requests that were

pruned from the queue in this way. This compression step may be performed only while the LSP is not being signaled, e.g., if two PCUpd arrive for the same LSP in quick succession and the PCC started the signaling of the changes relevant to the first PCUpd, then it MUST wait until the signaling finishes (and report the new state via a PCRpt) before attempting to apply the changes indicated in the second PCUpd.

Note also that it is up to the PCE to handle inter-LSP dependencies; for example, if ordering of LSP setups is required, the PCE has to wait for an LSP State Report for a previous LSP before starting the update of the next LSP.

If the PCUpd cannot be satisfied (for example, due to an unsupported object or a TLV), the PCC MUST respond with a PCErr message indicating the failure (see Section 7.3.3).

6.3. The PCErr Message

If the stateful PCE capability has been advertised on the PCEP session, the PCErr message MAY include the SRP object. If the error reported is the result of an LSP Update Request, then the SRP-ID-number MUST be the one from the PCUpd that triggered the error. If the error is unsolicited, the SRP object MAY be omitted. This is equivalent to including an SRP object with the SRP-ID-number equal to the reserved value 0x00000000.

The format of a PCErr message from [RFC5440] is extended as follows:

```
<PCErr Message> ::= <Common Header>
                    ( <error-obj-list> [<Open>] ) | <error>
                    [<error-list>]

<error-obj-list>::=<PCEP-ERROR>[<error-obj-list>]

<error>::=[<request-id-list> | <stateful-request-id-list>]
           <error-obj-list>

<request-id-list>::=<RP>[<request-id-list>]

<stateful-request-id-list>::=<SRP>[<stateful-request-id-list>]

<error-list>::=<error>[<error-list>]
```

6.4. The PCReq Message

A PCC MAY include the LSP object in the PCReq message (see Section 7.3) if the stateful PCE capability has been negotiated on a PCEP session between the PCC and a PCE.

The definition of the PCReq message from [RFC5440] is extended to optionally include the LSP object after the END-POINTS object. The encoding from [RFC5440] will become:

```
<PCReq Message>::= <Common Header>
                   [<svec-list>]
                   <request-list>
```

Where:

```
<svec-list>::=<SVEC>[<svec-list>]
<request-list>::=<request>[<request-list>]

<request>::= <RP>
             <END-POINTS>
             [<LSP>]
             [<LSPA>]
             [<BANDWIDTH>]
             [<metric-list>]
             [<RRO>[<BANDWIDTH>]]
             [<IRO>]
             [<LOAD-BALANCING>]
```

6.5. The PCRep Message

A PCE MAY include the LSP object in the PCRep message (see Section 7.3) if the stateful PCE capability has been negotiated on a PCEP session between the PCC, and the PCE and the LSP object were included in the corresponding PCReq message from the PCC.

The definition of the PCRep message from [RFC5440] is extended to optionally include the LSP object after the Request Parameter (RP) object. The encoding from [RFC5440] will become:

```
<PCRep Message> ::= <Common Header>
                    <response-list>
```

Where:

<response-list>::=<response>[<response-list>]

<response>::=<RP>
[<LSP>]
[<NO-PATH>]
[<attribute-list>]
[<path-list>]

7. Object Formats

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

7.1. OPEN Object

This document defines one new optional TLV for use in the OPEN object.

7.1.1. STATEFUL-PCE-CAPABILITY TLV

The STATEFUL-PCE-CAPABILITY TLV is an optional TLV for use in the OPEN object for stateful PCE capability advertisement. Its format is shown in the following figure:

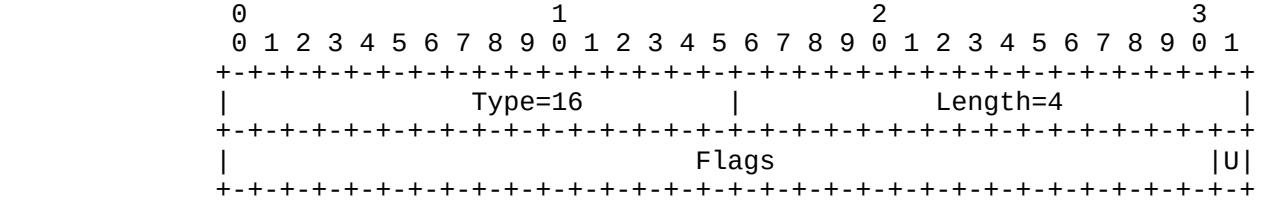

Figure 9: STATEFUL-PCE-CAPABILITY TLV Format

The type (16 bits) of the TLV is 16. The length field is 16 bits long and has a fixed value of 4.

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

U (LSP-UPDATE-CAPABILITY - 1 bit): if set to 1 by a PCC, the U flag indicates that the PCC allows modification of LSP parameters; if set to 1 by a PCE, the U flag indicates that the PCE is capable of

updating LSP parameters. The LSP-UPDATE-CAPABILITY flag must be advertised by both a PCC and a PCE for PCUpd messages to be allowed on a PCEP session.

Unassigned bits are considered reserved. They MUST be set to 0 on transmission and MUST be ignored on receipt.

A PCEP speaker operating in passive stateful PCE mode advertises the stateful PCE capability with the U flag set to 0. A PCEP speaker operating in active stateful PCE mode advertises the stateful PCE capability with the U flag set to 1.

Advertisement of the stateful PCE capability implies support of LSPs that are signaled via RSVP, as well as the objects, TLVs, and procedures defined in this document.

7.2. SRP Object

The SRP (Stateful PCE Request Parameters) object MUST be carried within PCUpd messages and MAY be carried within PCRpt and PCErr messages. The SRP object is used to correlate between update requests sent by the PCE and the error reports and state reports sent by the PCC.

SRP Object-Class is 33.

SRP Object-Type is 1.

The format of the SRP object body is shown in Figure 10:

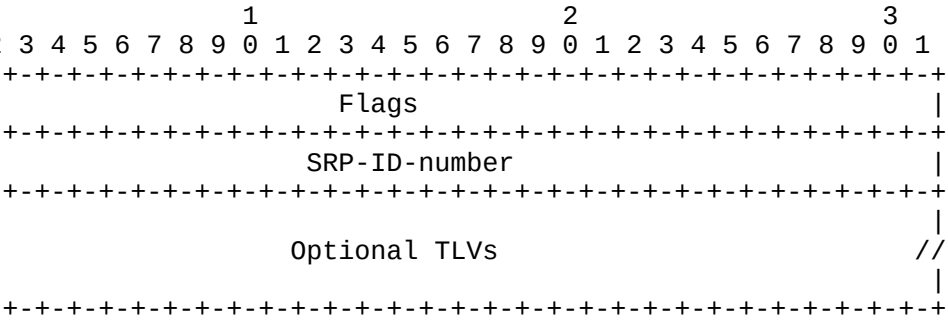

Figure 10: The SRP Object Format

The SRP object body has a variable length and may contain additional TLVs.

Flags (32 bits): None defined yet.

SRP-ID-number (32 bits): The SRP-ID-number value in the scope of the current PCEP session uniquely identifies the operation that the PCE has requested the PCC to perform on a given LSP. The SRP-ID-number is incremented each time a new request is sent to the PCC, and it may wrap around.

The values 0x00000000 and 0xFFFFFFFF are reserved.

Optional TLVs MAY be included within the SRP object body. The specification of such TLVs is outside the scope of this document.

Every request to update an LSP receives a new SRP-ID-number. This number is unique per PCEP session and is incremented each time an operation is requested from the PCE. Thus, for a given LSP, there may be more than one SRP-ID-number unacknowledged at a given time. The value of the SRP-ID-number is echoed back by the PCC in PCErr and PCRpt messages to allow for correlation between requests made by the PCE and errors or state reports generated by the PCC. If the error or report was not a result of a PCE operation (for example, in the case of a link down event), the reserved value of 0x00000000 is used for the SRP-ID-number. The absence of the SRP object is equivalent to an SRP object with the reserved value of 0x00000000. An SRP-ID-number is considered unacknowledged and cannot be reused until a PCErr or PCRpt arrives with an SRP-ID-number equal or higher for the same LSP. In case of SRP-ID-number wrapping, the last SRP-ID-number before the wrapping MUST be explicitly acknowledged, to avoid a situation where SRP-ID-numbers remain unacknowledged after the wrap. This means that the PCC may need to issue two PCUpd messages on detecting a wrap.

7.3. LSP Object

The LSP object MUST be present within PCRpt and PCUpd messages. The LSP object MAY be carried within PCReq and PCRep messages if the stateful PCE capability has been negotiated on the session. The LSP object contains a set of fields used to specify the target LSP, the operation to be performed on the LSP, and LSP delegation. It also contains a flag indicating to a PCE that the LSP State Synchronization is in progress. This document focuses on LSPs that are signaled with RSVP; many of the TLVs used with the LSP object mirror RSVP state.

LSP Object-Class is 32.

LSP Object-Type is 1.

The format of the LSP object body is shown in Figure 11:

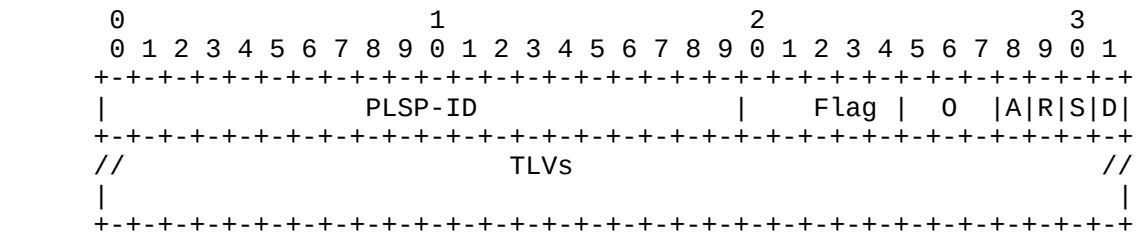

Figure 11: The LSP Object Format

PLSP-ID (20 bits): A PCEP-specific identifier for the LSP. A PCC creates a unique PLSP-ID for each LSP that is constant for the lifetime of a PCEP session. The PCC will advertise the same PLSP-ID on all PCEP sessions it maintains at a given time. The mapping of the symbolic path name to PLSP-ID is communicated to the PCE by sending a PCRpt message containing the SYMBOLIC-PATH-NAME TLV. All subsequent PCEP messages then address the LSP by the PLSP-ID. The values of 0 and 0xFFFFF are reserved. Note that the PLSP-ID is a value that is constant for the lifetime of the PCEP session, during which time for an RSVP-signaled LSP there might be different RSVP identifiers (LSP-id, tunnel-id) allocated to it.

Flags (12 bits), starting from the least significant bit:

D (Delegate - 1 bit): On a PCRpt message, the D flag set to 1 indicates that the PCC is delegating the LSP to the PCE. On a PCUpd message, the D flag set to 1 indicates that the PCE is confirming the LSP delegation. To keep an LSP delegated to the PCE, the PCC must set the D flag to 1 on each PCRpt message for the duration of the delegation -- the first PCRpt with the D flag set to 0 revokes the delegation. To keep the delegation, the PCE must set the D flag to 1 on each PCUpd message for the duration of the delegation -- the first PCUpd with the D flag set to 0 returns the delegation.

S (SYNC - 1 bit): The S flag MUST be set to 1 on each PCRpt sent from a PCC during State Synchronization. The S flag MUST be set to 0 in other messages sent from the PCC. When sending a PCUpd message, the PCE MUST set the S flag to 0.

R (Remove - 1 bit): On PCRpt messages, the R flag indicates that the LSP has been removed from the PCC and the PCE SHOULD remove all state from its database. Upon receiving an LSP State Report with the R flag set to 1 for an RSVP-signaled LSP, the PCE SHOULD remove all state for the path identified by the LSP-IDENTIFIERS

TLV from its database. When the all-zeros LSP-IDENTIFIERS TLV is used, the PCE SHOULD remove all state for the PLSP-ID from its database. When sending a PCUpd message, the PCE MUST set the R flag to 0.

A (Administrative - 1 bit): On PCRpt messages, the A flag indicates the PCC's target operational status for this LSP. On PCUpd messages, the A flag indicates the LSP status that the PCE desires for this LSP. In both cases, a value of '1' means that the desired operational state is active, and a value of '0' means that the desired operational state is inactive. A PCC ignores the A flag on a PCUpd message unless the operator's policy allows the PCE to control the corresponding LSP's administrative state.

O (Operational - 3 bits): On PCRpt messages, the O field represents the operational status of the LSP.

The following values are defined:

0 - DOWN: not active.

1 - UP: signaled.

2 - ACTIVE: up and carrying traffic.

3 - GOING-DOWN: LSP is being torn down, and resources are being released.

4 - GOING-UP: LSP is being signaled.

5-7 - Reserved: these values are reserved for future use.

Unassigned bits are reserved for future uses. They MUST be set to 0 on transmission and MUST be ignored on receipt. When sending a PCUpd message, the PCE MUST set the O field to 0.

TLVs that may be included in the LSP object are described in the following sections. Other optional TLVs, that are not defined in this document, MAY also be included within the LSP object body.

7.3.1. LSP-IDENTIFIERS TLVs

The LSP-IDENTIFIERS TLV MUST be included in the LSP object in PCRpt messages for RSVP-signaled LSPs. If the TLV is missing, the PCE will generate an error with Error-type=6 (Mandatory Object missing) and error-value 11 (LSP-IDENTIFIERS TLV missing) and close the session. The LSP-IDENTIFIERS TLV MAY be included in the LSP object in PCUpd messages for RSVP-signaled LSPs. The special value of all zeros for

this TLV is used to refer to all paths pertaining to a particular PLSP-ID. There are two LSP-IDENTIFIERS TLVs, one for IPv4 and one for IPv6.

It is the responsibility of the PCC to send to the PCE the identifiers for each RSVP incarnation of the tunnel. For example, in a make-before-break scenario, the PCC MUST send a separate PCRpt for the old and reoptimized paths and explicitly report removal of any of these paths using the R bit in the LSP object.

The format of the IPV4-LSP-IDENTIFIERS TLV is shown in the following figure:

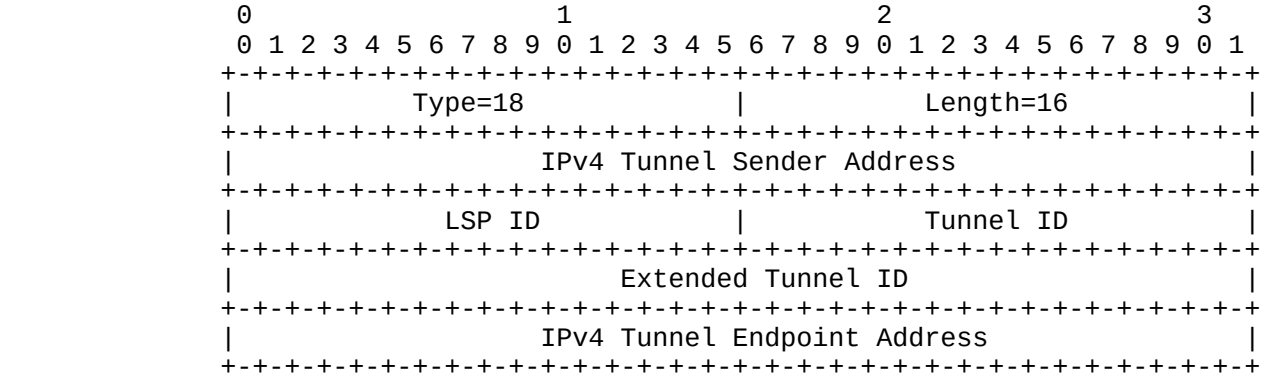

Figure 12: IPV4-LSP-IDENTIFIERS TLV Format

The type (16 bits) of the TLV is 18. The length field is 16 bits long and has a fixed value of 16. The value contains the following fields:

IPv4 Tunnel Sender Address: contains the sender node's IPv4 address, as defined in [RFC3209], Section 4.6.2.1, for the LSP_TUNNEL_IPv4 Sender Template Object.

LSP ID: contains the 16-bit 'LSP ID' identifier defined in [RFC3209], Section 4.6.2.1 for the LSP_TUNNEL_IPv4 Sender Template Object. A value of 0 MUST be used if the LSP is not yet signaled.

Tunnel ID: contains the 16-bit 'Tunnel ID' identifier defined in [RFC3209], Section 4.6.1.1 for the LSP_TUNNEL_IPv4 Session Object.

Extended Tunnel ID: contains the 32-bit 'Extended Tunnel ID' identifier defined in [RFC3209], Section 4.6.1.1 for the LSP_TUNNEL_IPv4 Session Object.

IPv4 Tunnel Endpoint Address: contains the egress node's IPv4 address, as defined in [RFC3209], Section 4.6.1.1, for the LSP_TUNNEL_IPv4 Sender Template Object.

The format of the IPV6-LSP-IDENTIFIERS TLV is shown in the following figure:

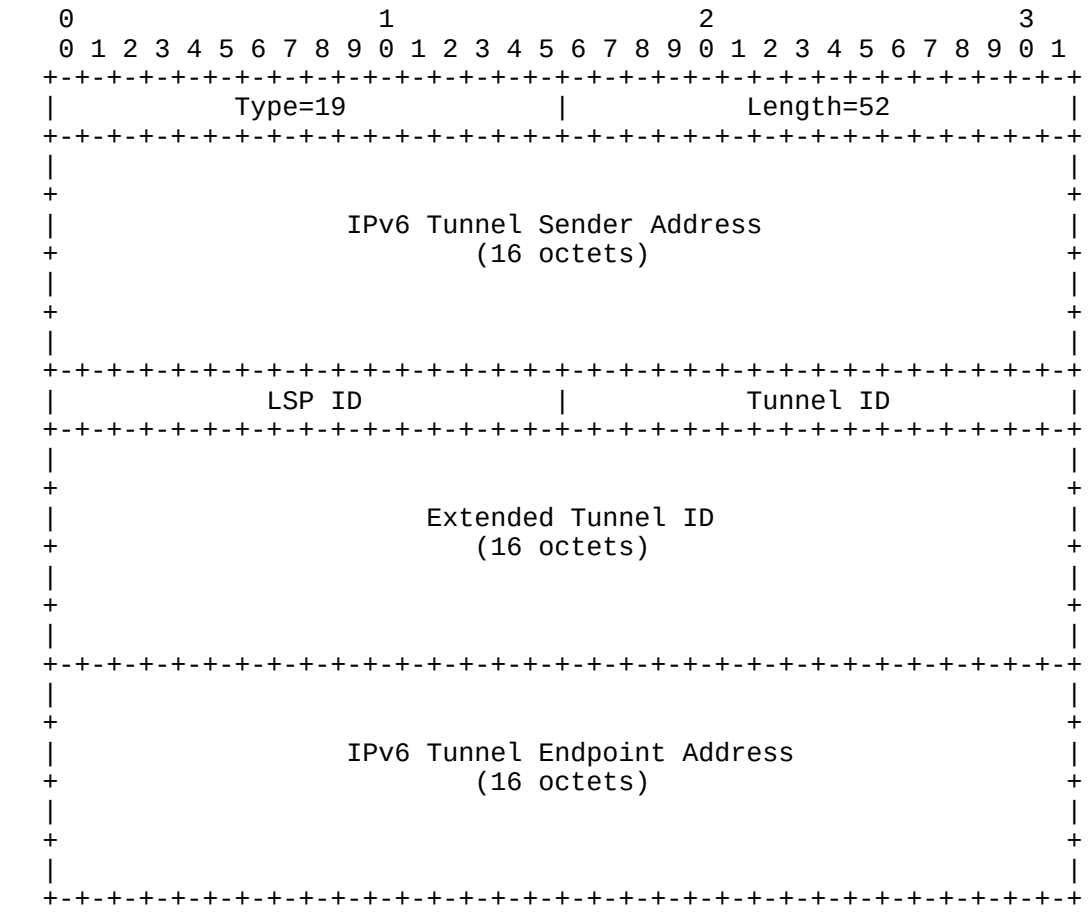

Figure 13: IPV6-LSP-IDENTIFIERS TLV Format

The type (16 bits) of the TLV is 19. The length field is 16 bits long and has a fixed value of 52. The value contains the following fields:

IPv6 Tunnel Sender Address: contains the sender node's IPv6 address, as defined in [RFC3209], Section 4.6.2.2, for the LSP_TUNNEL_IPv6 Sender Template Object.

LSP ID: contains the 16-bit 'LSP ID' identifier defined in [RFC3209], Section 4.6.2.2 for the LSP_TUNNEL_IPv6 Sender Template Object. A value of 0 MUST be used if the LSP is not yet signaled.

Tunnel ID: contains the 16-bit 'Tunnel ID' identifier defined in [RFC3209], Section 4.6.1.2 for the LSP_TUNNEL_IPv6 Session Object.

Extended Tunnel ID: contains the 128-bit 'Extended Tunnel ID' identifier defined in [RFC3209], Section 4.6.1.2 for the LSP_TUNNEL_IPv6 Session Object.

IPv6 Tunnel Endpoint Address: contains the egress node's IPv6 address, as defined in [RFC3209], Section 4.6.1.2, for the LSP_TUNNEL_IPv6 Session Object.

The Tunnel ID remains constant over the lifetime of a tunnel.

7.3.2. Symbolic Path Name TLV

Each LSP MUST have a symbolic path name that is unique in the PCC. The symbolic path name is a human-readable string that identifies an LSP in the network. The symbolic path name MUST remain constant throughout an LSP's lifetime, which may span across multiple consecutive PCEP sessions and/or PCC restarts. The symbolic path name MAY be specified by an operator in a PCC's configuration. If the operator does not specify a unique symbolic name for an LSP, then the PCC MUST auto-generate one.

The PCE uses the symbolic path name as a stable identifier for the LSP. If the PCEP session restarts, or the PCC restarts, or the PCC re-delegates the LSP to a different PCE, the symbolic path name for the LSP remains constant and can be used to correlate across the PCEP session instances.

The other protocol identifiers for the LSP cannot reliably be used to identify the LSP across multiple PCEP sessions, for the following reasons.

- The PLSP-ID is unique only within the scope of a single PCEP session.
- The LSP-IDENTIFIERS TLV is only guaranteed to be present for LSPs that are signaled with RSVP-TE, and it may change during the lifetime of the LSP.

The SYMBOLIC-PATH-NAME TLV MUST be included in the LSP object in the LSP State Report (PCRpt) message when during a given PCEP session an LSP is first reported to a PCE. A PCC sends to a PCE the first LSP

State Report either during State Synchronization or when a new LSP is configured at the PCC.

The initial PCRpt creates a binding between the symbolic path name and the PLSP-ID for the LSP that lasts for the duration of the PCEP session. The PCC MAY omit the symbolic path name from subsequent LSP State Reports for that LSP on that PCEP session, and just use the PLSP-ID.

The format of the SYMBOLIC-PATH-NAME TLV is shown in the following figure:

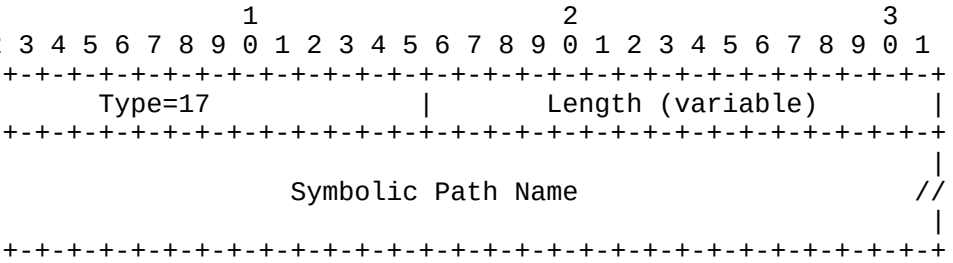

Figure 14: SYMBOLIC-PATH-NAME TLV Format

Type (16 bits): the type is 17.

Length (16 bits): indicates the total length of the TLV in octets and MUST be greater than 0. The TLV MUST be zero-padded so that the TLV is 4-octet aligned.

Symbolic Path Name (variable): symbolic name for the LSP, unique in the PCC. It SHOULD be a string of printable ASCII characters, without a NULL terminator.

7.3.3. LSP Error Code TLV

The LSP Error Code TLV is an optional TLV for use in the LSP object to convey error information. When an LSP Update Request fails, an LSP State Report MUST be sent to report the current state of the LSP, and it SHOULD contain the LSP-ERROR-CODE TLV indicating the reason for the failure. Similarly, when a PCRpt is sent as a result of an LSP transitioning to non-operational state, the LSP-ERROR-CODE TLV SHOULD be included to indicate the reason for the transition.

The format of the LSP-ERROR-CODE TLV is shown in the following figure:

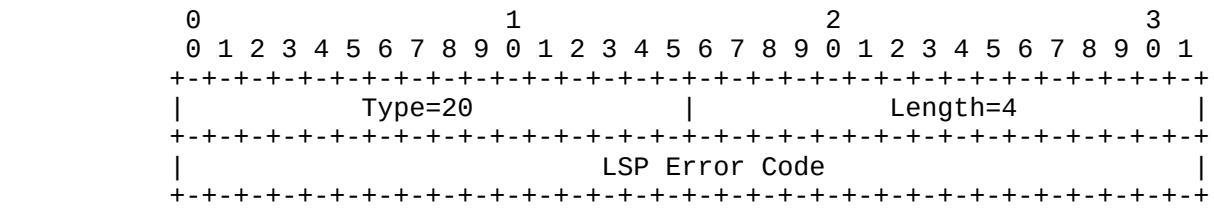

Figure 15: LSP-ERROR-CODE TLV Format

The type (16 bits) of the TLV is 20. The length field is 16 bits long and has a fixed value of 4. The value contains an error code that indicates the cause of the failure.

The following LSP Error Codes are currently defined:

Value	Description
1	Unknown reason
2	Limit reached for PCE-controlled LSPs
3	Too many pending LSP Update Requests
4	Unacceptable parameters
5	Internal error
6	LSP administratively brought down
7	LSP preempted
8	RSVP signaling error

7.3.4. RSVP Error Spec TLV

The RSVP-ERROR-SPEC TLV is an optional TLV for use in the LSP object to carry RSVP error information. It includes the RSVP ERROR_SPEC or USER_ERROR_SPEC object ([RFC2205] and [RFC5284]), which were returned to the PCC from a downstream node. If the setup of an LSP fails at a downstream node that returned an ERROR_SPEC to the PCC, the PCC SHOULD include in the PCRpt for this LSP the LSP-ERROR-CODE TLV with LSP Error Code = "RSVP signaling error" and the RSVP-ERROR-SPEC TLV with the relevant RSVP ERROR-SPEC or USER_ERROR_SPEC object.

The format of the RSVP-ERROR-SPEC TLV is shown in the following figure:

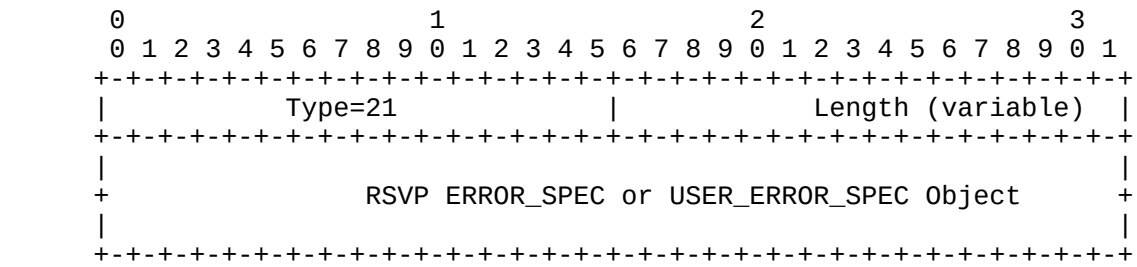

Figure 16: RSVP-ERROR-SPEC TLV Format

Type (16 bits): the type is 21.

Length (16 bits): indicates the total length of the TLV in octets. The TLV MUST be zero-padded so that the TLV is 4-octet aligned.

Value (variable): contains the RSVP ERROR_SPEC or USER_ERROR_SPEC object, as specified in [RFC2205] and [RFC5284], including the object header.

8. IANA Considerations

The code points described below have been allocated for the protocol elements defined in this document.

8.1. PCE Capabilities in IGP Advertisements

The following bits have been registered in the "Path Computation Element (PCE) Capability Flags" subregistry of the "Open Shortest Path First (OSPF) Parameters" registry:

Bit	Description	Reference
11	Active stateful PCE capability	This document
12	Passive stateful PCE capability	This document

8.2. PCEP Messages

The following message types have been allocated within the "PCEP Messages" subregistry of the "Path Computation Element Protocol (PCEP) Numbers" registry:

Value	Description	Reference
10	Report	This document
11	Update	This document

8.3. PCEP Objects

The following object-class values and object types have been allocated within the "PCEP Objects" subregistry of the "Path Computation Element Protocol (PCEP) Numbers" registry:

Object-Class Value	Name	Reference
32	LSP Object-Type 0: Reserved 1: LSP	This document
33	SRP Object-Type 0: Reserved 1: SRP	This document

8.4. LSP Object

A new subregistry, named "LSP Object Flag Field", has been created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the Flag field of the LSP object. New values are assigned by Standards Action [RFC8126]. Each bit should be tracked with the following qualities:

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

The following values are defined in this document:

Bit	Description	Reference
0-4	Unassigned	This document
5-7	Operational (3 bits)	This document
8	Administrative	This document
9	Remove	This document
10	SYNC	This document
11	Delegate	This document

8.5. PCEP-Error Object

The following error types and error values have been registered within the "PCEP-ERROR Object Error Types and Values" subregistry of the "Path Computation Element Protocol (PCEP) Numbers" registry:

Error-Type	Meaning
6	Mandatory Object missing
	Error-value 8: LSP object missing 9: ERO object missing 10: SRP object missing 11: LSP-IDENTIFIERS TLV missing
19	Invalid Operation
	Error-value 1: Attempted LSP Update Request for a non-delegated LSP. The PCEP-ERROR object is followed by the LSP object that identifies the LSP. 2: Attempted LSP Update Request if the stateful PCE capability was not advertised. 3: Attempted LSP Update Request for an LSP identified by an unknown PLSP-ID. 5: Attempted LSP State Report if stateful PCE capability was not advertised.
20	LSP State Synchronization Error
	Error-value 1: A PCE indicates to a PCC that it cannot process (an otherwise valid) LSP State Report. The PCEP-ERROR object is followed by the LSP object that identifies the LSP. 5: A PCC indicates to a PCE that it cannot complete the State Synchronization.

8.6. Notification Object

The following Notification Types and Notification Values have been allocated within the "Notification Object" subregistry of the "Path Computation Element Protocol (PCEP) Numbers" registry:

```
Notification-Type  Name
4     Stateful PCE resource limit exceeded

          Notification-value
          1:   Entering resource limit exceeded state
          2:   Deprecated
```

Note that the early allocation included an additional Notification Value 2 for "Exiting resource limit exceeded state". This Notification Value is no longer required and has been marked as "Deprecated".

8.7. PCEP TLV Type Indicators

The following TLV Type Indicator values have been registered within the "PCEP TLV Type Indicators" subregistry of the "Path Computation Element Protocol (PCEP) Numbers" registry:

Value	Description	Reference
16	STATEFUL-PCE-CAPABILITY	This document
17	SYMBOLIC-PATH-NAME	This document
18	IPV4-LSP-IDENTIFIERS	This document
19	IPV6-LSP-IDENTIFIERS	This document
20	LSP-ERROR-CODE	This document
21	RSVP-ERROR-SPEC	This document

8.8. STATEFUL-PCE-CAPABILITY TLV

A new subregistry, named "STATEFUL-PCE-CAPABILITY TLV Flag Field", has been created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the Flag field in the STATEFUL-PCE-CAPABILITY TLV of the PCEP OPEN object (class = 1). New values are assigned by Standards Action [RFC8126]. Each bit should be tracked with the following qualities:

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

The following values are defined in this document:

Value	Description	Reference
31	LSP-UPDATE-CAPABILITY	This document

8.9. LSP-ERROR-CODE TLV

A new subregistry, named "LSP-ERROR-CODE TLV Error Code Field", has been created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the LSP Error Code field of the LSP-ERROR-CODE TLV. This field specifies the reason for failure to update the LSP.

New values are assigned by Standards Action [RFC8126]. Each value should be tracked with the following qualities: value, meaning, and defining RFC. The following values are defined in this document:

Value	Meaning
0	Reserved
1	Unknown reason
2	Limit reached for PCE-controlled LSPs
3	Too many pending LSP Update Requests
4	Unacceptable parameters
5	Internal error
6	LSP administratively brought down
7	LSP preempted
8	RSVP signaling error

9. Manageability Considerations

All manageability requirements and considerations listed in [RFC5440] apply to the PCEP extensions defined in this document. In addition, requirements and considerations listed in this section apply.

9.1. Control Function and Policy

In addition to configuring specific PCEP session parameters, as specified in [RFC5440], Section 8.1, a PCE or PCC implementation MUST allow configuring the stateful PCEP capability and the LSP Update capability. A PCC implementation SHOULD allow the operator to specify multiple candidate PCEs for and a delegation preference for each candidate PCE. A PCC SHOULD allow the operator to specify an LSP delegation policy where LSPs are delegated to the most-preferred online PCE. A PCC MAY allow the operator to specify different LSP delegation policies.

A PCC implementation that allows concurrent connections to multiple PCEs SHOULD allow the operator to group the PCEs by administrative domains, and it MUST NOT advertise LSP existence and state to a PCE if the LSP is delegated to a PCE in a different group.

A PCC implementation SHOULD allow the operator to specify whether the PCC will advertise LSP existence and state for LSPs that are not controlled by any PCE (for example, LSPs that are statically configured at the PCC).

A PCC implementation SHOULD allow the operator to specify both the Redelegation Timeout Interval and the State Timeout Interval. The default value of the Redelegation Timeout Interval SHOULD be set to 30 seconds. An operator MAY also configure a policy that will dynamically adjust the Redelegation Timeout Interval, for example setting it to zero when the PCC has an established session to a backup PCE. The default value for the State Timeout Interval SHOULD be set to 60 seconds.

After the expiration of the State Timeout Interval, the LSP reverts to operator-defined default parameters. A PCC implementation MUST allow the operator to specify the default LSP parameters. To achieve a behavior where the LSP retains the parameters set by the PCE until such time that the PCC makes a change to them, a State Timeout Interval of infinity SHOULD be used. Any changes to LSP parameters SHOULD be done in a make-before-break fashion.

LSP delegation is controlled by operator-defined policies on a PCC. LSPs are delegated individually -- different LSPs may be delegated to different PCEs. An LSP is delegated to at most one PCE at any given

point in time. A PCC implementation SHOULD support the delegation policy, when all PCC's LSPs are delegated to a single PCE at any given time. Conversely, the policy revoking the delegation for all PCC's LSPs SHOULD also be supported.

A PCC implementation SHOULD allow the operator to specify delegation priority for PCEs. This effectively defines the primary PCE and one or more backup PCEs to which a primary PCE's LSPs can be delegated when the primary PCE fails.

Policies defined for stateful PCEs and PCCs should eventually fit in the policy-enabled path computation framework defined in [RFC5394], and the framework should be extended to support stateful PCEs.

9.2. Information and Data Models

The PCEP YANG module [PCEP-YANG] should include:

- advertised stateful capabilities and synchronization status per PCEP session.
- the delegation status of each configured LSP.

The PCEP MIB [RFC7420] could also be updated to include this information.

9.3. Liveness Detection and Monitoring

PCEP extensions defined in this document do not require any new mechanisms beyond those already defined in [RFC5440], Section 8.3.

9.4. Verifying Correct Operation

Mechanisms defined in [RFC5440], Section 8.4 also apply to PCEP extensions defined in this document. In addition to monitoring parameters defined in [RFC5440], a stateful PCC-side PCEP implementation SHOULD provide the following parameters:

- Total number of LSP Updates
- Number of successful LSP Updates
- Number of dropped LSP Updates
- Number of LSP Updates where LSP setup failed

A PCC implementation SHOULD provide a command to show for each LSP whether it is delegated, and if so, to which PCE.

A PCC implementation SHOULD allow the operator to manually revoke LSP delegation.

9.5. Requirements on Other Protocols and Functional Components

PCEP extensions defined in this document do not put new requirements on other protocols.

9.6. Impact on Network Operation

Mechanisms defined in [RFC5440], Section 8.6 also apply to PCEP extensions defined in this document.

Additionally, a PCEP implementation SHOULD allow a limit to be placed on the number of LSPs delegated to the PCE and on the rate of PCUpd and PCRpt messages sent by a PCEP speaker and processed from a peer. It SHOULD also allow sending a notification when a rate threshold is reached.

A PCC implementation SHOULD allow a limit to be placed on the rate of LSP Updates to the same LSP to avoid signaling overload discussed in Section 10.3.

10. Security Considerations

10.1. Vulnerability

This document defines extensions to PCEP to enable stateful PCEs. The nature of these extensions and the delegation of path control to PCEs results in more information being available for a hypothetical adversary and a number of additional attack surfaces that must be protected.

The security provisions described in [RFC5440] remain applicable to these extensions. However, because the protocol modifications outlined in this document allow the PCE to control path computation timing and sequence, the PCE defense mechanisms described in [RFC5440], Section 7.2 are also now applicable to PCC security.

As a general precaution, it is RECOMMENDED that these PCEP extensions only be activated on authenticated and encrypted sessions across PCEs and PCCs belonging to the same administrative authority, using Transport Layer Security (TLS) [PCEPS], as per the recommendations and best current practices in [RFC7525].

The following sections identify specific security concerns that may result from the PCEP extensions outlined in this document along with recommended mechanisms to protect PCEP infrastructure against related attacks.

10.2. LSP State Snooping

The stateful nature of this extension explicitly requires LSP status updates to be sent from PCC to PCE. While this gives the PCE the ability to provide more optimal computations to the PCC, it also provides an adversary with the opportunity to eavesdrop on decisions made by network systems external to PCE. This is especially true if the PCC delegates LSPs to multiple PCEs simultaneously.

Adversaries may gain access to this information by eavesdropping on unsecured PCEP sessions and might then use this information in various ways to target or optimize attacks on network infrastructure, for example, by flexibly countering anti-DDoS measures being taken to protect the network or by determining choke points in the network where the greatest harm might be caused.

PCC implementations that allow concurrent connections to multiple PCEs SHOULD allow the operator to group the PCEs by administrative domains, and they MUST NOT advertise LSP existence and state to a PCE if the LSP is delegated to a PCE in a different group.

10.3. Malicious PCE

The LSP delegation mechanism described in this document allows a PCC to grant effective control of an LSP to the PCE for the duration of a PCEP session. While this enables PCE control of the timing and sequence of path computations within and across PCEP sessions, it also introduces a new attack vector: an attacker may flood the PCC with PCUpd messages at a rate that exceeds either the PCC's ability to process them or the network's ability to signal the changes, by either spoofing messages or compromising the PCE itself.

A PCC is free to revoke an LSP delegation at any time without needing any justification. A defending PCC can do this by enqueueing the appropriate PCRpt message. As soon as that message is enqueued in the session, the PCC is free to drop any incoming PCUpd messages without additional processing.

10.4. Malicious PCC

A stateful session also results in an increased attack surface by placing a requirement for the PCE to keep an LSP state replica for each PCC. It is RECOMMENDED that PCE implementations provide a limit on resources a single PCC can occupy. A PCE implementing such a limit MUST send a PCNtf message with notification-type 4 (Stateful PCE resource limit exceeded) and notification-value 1 (Entering resource limit exceeded state) upon receiving an LSP State Report causing it to exceed this threshold.

Delegation of LSPs can create further strain on PCE resources and a PCE implementation MAY preemptively give back delegations if it finds itself lacking the resources needed to effectively manage the delegation. Since the delegation state is ultimately controlled by the PCC, PCE implementations SHOULD provide throttling mechanisms to prevent strain created by flaps of either a PCEP session or an LSP delegation.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, DOI 10.17487/RFC2205, September 1997, <https://www.rfc-editor.org/info/rfc2205>.

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

[RFC5088] Le Roux, JL., Ed., Vasseur, JP., Ed., Ikejiri, Y., and R. Zhang, "OSPF Protocol Extensions for Path Computation Element (PCE) Discovery", RFC 5088, DOI 10.17487/RFC5088, January 2008, <https://www.rfc-editor.org/info/rfc5088>.

[RFC5089] Le Roux, JL., Ed., Vasseur, JP., Ed., Ikejiri, Y., and R. Zhang, "IS-IS Protocol Extensions for Path Computation Element (PCE) Discovery", RFC 5089, DOI 10.17487/RFC5089, January 2008, <https://www.rfc-editor.org/info/rfc5089>.

[RFC5284] Swallow, G. and A. Farrel, "User-Defined Errors for RSVP", RFC 5284, DOI 10.17487/RFC5284, August 2008, <https://www.rfc-editor.org/info/rfc5284>.

[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>.

[RFC5511] Farrel, A., "Routing Backus-Naur Form (RBNF): A Syntax Used to Form Encoding Rules in Various Routing Protocol Specifications", RFC 5511, DOI 10.17487/RFC5511, April 2009, <https://www.rfc-editor.org/info/rfc5511>.

[RFC8051] Zhang, X., Ed. and I. Minei, Ed., "Applicability of a Stateful Path Computation Element (PCE)", RFC 8051, DOI 10.17487/RFC8051, January 2017, <https://www.rfc-editor.org/info/rfc8051>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

11.2. Informative References

[MPLS-PC] Chaieb, I., Le Roux, JL., and B. Cousin, "Improved MPLS-TE LSP Path Computation using Preemption", Global Information Infrastructure Symposium, DOI 10.1109/GIIS.2007.4404195, July 2007.

[MXMN-TE] Danna, E., Mandal, S., and A. Singh, "A practical algorithm for balancing the max-min fairness and throughput objectives in traffic engineering", INFOCOM, 2012 Proceedings IEEE, pp. 846-854, DOI 10.1109/INFCOM.2012.6195833, March 2012.

[PCE-Init-LSP] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", Work in Progress, draft-ietf-pce-pce-initiated-lsp-10, June 2017.

[PCEP-GMPLS] Margaria, C., de Dios, O., and F. Zhang, "PCEP extensions for GMPLS", Work in Progress, draft-ietf-pce-gmpls-pcep-extensions-11, October 2015.

[PCEP-YANG] Dhody, D., Hardwick, J., Beeram, V., and j. jefftant@gmail.com, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", Work in Progress, draft-ietf-pce-pcep-yang-05, June 2017.

[PCEPS] Lopez, D., de Dios, O., Wu, Q., and D. Dhody, "Secure Transport for PCEP", Work in Progress, draft-ietf-pce-pceps-18, September 2017.

[RFC2702] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and J. McManus, "Requirements for Traffic Engineering Over MPLS", RFC 2702, DOI 10.17487/RFC2702, September 1999, <https://www.rfc-editor.org/info/rfc2702>.

[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>.

[RFC3346] Boyle, J., Gill, V., Hannan, A., Cooper, D., Awduche, D., Christian, B., and W. Lai, "Applicability Statement for Traffic Engineering with MPLS", RFC 3346, DOI 10.17487/RFC3346, August 2002, <https://www.rfc-editor.org/info/rfc3346>.

[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>.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <https://www.rfc-editor.org/info/rfc4655>.

[RFC4657] Ash, J., Ed. and J. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, DOI 10.17487/RFC4657, September 2006, <https://www.rfc-editor.org/info/rfc4657>.

[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>.

[RFC5394] Bryskin, I., Papadimitriou, D., Berger, L., and J. Ash, "Policy-Enabled Path Computation Framework", RFC 5394, DOI 10.17487/RFC5394, December 2008, <https://www.rfc-editor.org/info/rfc5394>.

[RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <https://www.rfc-editor.org/info/rfc7420>.

[RFC7525] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May 2015, <https://www.rfc-editor.org/info/rfc7525>.

[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, <http://www.rfc-editor.org/info/rfc8232>.

Acknowledgements

We would like to thank Adrian Farrel, Cyril Margaria, and Ramon Casellas for their contributions to this document.

We would like to thank Shane Amante, Julien Meuric, Kohei Shiomoto, Paul Schultz, and Raveendra Torvi for their comments and suggestions. Thanks also to Jon Hardwick, Oscar Gonzales de Dios, Tomas Janciga, Stefan Kobza, Kexin Tang, Matej Spanik, Jon Parker, Marek Zavodsky, Ambrose Kwong, Ashwin Sampath, Calvin Ying, Mustapha Aissaoui, Stephane Litkowski, and Olivier Dugeon for helpful comments and discussions.

Contributors

The following people contributed substantially to the content of this document and should be considered coauthors:

Xian Zhang
Huawei Technology
F3-5-B R&D Center
Huawei Industrial Base, Bantian, Longgang District
Shenzhen, Guangdong 518129
China
Email: zhang.xian@huawei.com

Dhruv Dhody
Huawei Technology
Leela Palace
Bangalore, Karnataka 560008
INDIA
Email: dhruv.dhody@huawei.com

Siva Sivabalan
Cisco Systems, Inc.
2000 Innovation Drive
Kanata, Ontario K2K 3E8
Canada
Email: msiva@cisco.com

Authors' Addresses

Edward Crabbe
Oracle
1501 4th Ave, suite 1800
Seattle, WA 98101
United States of America

Email: edward.crabbe@oracle.com

Ina Minei
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043
United States of America

Email: inaminei@google.com

Jan Medved
Cisco Systems, Inc.
170 West Tasman Dr.
San Jose, CA 95134
United States of America

Email: jmedved@cisco.com

Robert Varga
Pantheon Technologies SRO
Mlynske Nivy 56
Bratislava 821 05
Slovakia

Email: robert.varga@pantheon.tech