PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: June 11, 2017

E. Crabbe
Oracle
I. Minei
Google, Inc.
J. Medved
Cisco Systems, Inc.
R. Varga
Pantheon Technologies SRO
X. Zhang
D. Dhody
Huawei Technologies

December 8, 2016

Optimizations of Label Switched Path State Synchronization Procedures for a Stateful PCE

draft-ietf-pce-stateful-sync-optimizations-07

Abstract

A stateful Path Computation Element (PCE) has access to not only the information disseminated by the network's Interior Gateway Protocol (IGP), but also the set of active paths and their reserved resources for its computation. The additional Label Switched Path (LSP) state information allows the PCE to compute constrained paths while considering individual LSPs and their interactions. This requires a reliable state synchronization mechanism between the PCE and the network, PCE and path computation clients (PCCs), and between cooperating PCEs. The basic mechanism for state synchronization is part of the stateful PCE specification. This draft presents motivations for optimizations to the base state synchronization procedure and specifies the required Path Computation Element Communication Protocol (PCEP) extensions.

Requirements Language

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

Copyright Notice

Table of Contents

1. Introduction 3
2. Terminology 4
3. State Synchronization Avoidance 4
  3.1. Motivation 4
  3.2. State Synchronization Avoidance Procedure 4
  3.3. PCEP Extensions 9
    3.3.1. LSP State Database Version Number TLV 9
    3.3.2. Speaker Entity Identifier TLV 10
4. Incremental State Synchronization 11
  4.1. Motivation 12
  4.2. Incremental Synchronization Procedure 13
5. PCE-triggered Initial Synchronization 15
  5.1. Motivation 15
  5.2. PCE-triggered Initial State Synchronization Procedure . . 15
6. PCE-triggered Re-synchronization 16
  6.1. Motivation 16
  6.2. PCE-triggered State Re-synchronization Procedure 16
7. Advertising Support of Synchronization Optimizations 17
8. IANA Considerations 18
  8.1. PCEP-Error Object 18
  8.2. PCEP TLV Type Indicators 19

8.3. STATEFUL-PCE-CAPABILITY TLV 19
9. Manageability Considerations 20
9.1. Control of Function and Policy 20
9.2. Information and Data Models 20
9.3. Liveness Detection and Monitoring 20
9.4. Verify Correct Operations 20
9.5. Requirements On Other Protocols 20
9.6. Impact On Network Operations 21
10. Security Considerations 21
11. Acknowledgements 21
12. Contributors 21
13. References 21
13.1. Normative References 21
13.2. Informative References 22
Authors' Addresses 22

1. Introduction

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

[I-D.ietf-pce-stateful-pce] describes a set of extensions to PCEP to provide stateful control. A stateful PCE has access to not only the information carried by the network's Interior Gateway Protocol (IGP), but also the set of active paths and their reserved resources for its computations. The additional state allows the PCE to compute constrained paths while considering individual LSPs and their interactions. This requires a reliable state synchronization mechanism between the PCE and the network, PCE and PCC, and between cooperating PCEs. [I-D.ietf-pce-stateful-pce] describes the basic mechanism for state synchronization. This draft specifies following optimizations for state synchronization and the corresponding PCEP procedures and extensions:

- o State Synchronization Avoidance: To skip state synchronization if the state has survived and not changed during session restart. (See Section 3.)

- o Incremental State Synchronization: To do incremental (delta) state synchronization when possible. (See Section 4.)

- o PCE-triggered Initial Synchronization: To let PCE control the timing of the initial state synchronization. (See Section 5.)

- o PCE-triggered Re-synchronization: To let PCE re-synchronize the state for sanity check. (See Section 6.)

2. Terminology

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

This document uses the following terms defined in [I-D.ietf-pce-stateful-pce]: Delegation, Redelegation Timeout Interval, LSP State Report, LSP Update Request, LSP State Database.

Within this document, when describing PCE-PCE communications, the requesting PCE fills the role of a PCC. This provides a saving in documentation without loss of function.

3. State Synchronization Avoidance

3.1. Motivation

The purpose of state synchronization is to provide a checkpoint-in-time state replica of a PCC's LSP state in a stateful PCE. State synchronization is performed immediately after the initialization phase ([RFC5440]). [I-D.ietf-pce-stateful-pce] describes the basic mechanism for state synchronization.

State synchronization is not always necessary following a PCEP session restart. If the state of both PCEP peers did not change, the synchronization phase may be skipped. This can result in significant savings in both control-plane data exchanges and the time it takes for the stateful PCE to become fully operational.

3.2. State Synchronization Avoidance Procedure

State synchronization MAY be skipped following a PCEP session restart if the state of both PCEP peers did not change during the period prior to session re-initialization. To be able to make this determination, state must be exchanged and maintained by both PCE and PCC during normal operation. This is accomplished by keeping track of the changes to the LSP state database, using a version tracking field called the LSP State Database Version Number.

The LSP State Database Version Number, carried in LSP-DB-VERSION TLV (see Section 3.3.1), is owned by a PCC and it MUST be incremented by 1 for each successive change in the PCC's LSP state database. The LSP State Database Version Number MUST start at 1 and may wrap around. Values 0 and 0xFFFFFFFFFFFFFFFF are reserved. If either of the two values are used during LSP state (re)-synchronization, the PCE speaker receiving this node should send back a PCErr with Error-type 20 Error-value TBD (suggested value - 6) 'Received an invalid LSP DB Version Number', and close the PCEP session. Operations that

trigger a change to the local LSP state database include a change in the LSP operational state, delegation of an LSP, removal or setup of an LSP or change in any of the LSP attributes that would trigger a report to the PCE.

If state synchronization avoidance is enabled, a PCC MUST increment its LSP State Database Version Number when the 'Redelegation Timeout Interval' timer expires (see [I-D.ietf-pce-stateful-pce]) for the use of the Redelegation Timeout Interval).

State synchronization avoidance is advertised on a PCEP session during session startup using the INCLUDE-DB-VERSION (S) bit in the capabilities TLV (see Section 7). The peer may move in the network, either physically or logically, which may cause its connectivity details and transport-level identity (such as IP address) to change. To ensure that a PCEP peer can recognize a previously connected peer even in face of such mobility, each PCEP peer includes the SPEAKER-ENTITY-ID TLV described in Section 3.3.2 in the OPEN message.

If both PCEP speakers set the S flag in the OPEN object's STATEFUL-PCE-CAPABILITY TLV to 1, the PCC MUST include the LSP-DB-VERSION TLV in each LSP object of the PCRpt message. If the LSP-DB-VERSION TLV is missing in a PCRpt message, the PCE will generate an error with Error-Type 6 (mandatory object missing) and Error-Value TBD (suggested value - 12) 'LSP-DB-VERSION TLV missing' and close the session. If state synchronization avoidance has not been enabled on a PCEP session, the PCC SHOULD NOT include the LSP-DB-VERSION TLV in the LSP Object and the PCE SHOULD ignore it were it to receive one.

If a PCE's LSP state database survived the restart of a PCEP session, the PCE will include the LSP-DB-VERSION TLV in its OPEN object, and the TLV will contain the last LSP State Database Version Number received on an LSP State Report from the PCC in the previous PCEP session. If a PCC's LSP State Database survived the restart of a PCEP session, the PCC will include the LSP-DB-VERSION TLV in its OPEN object and the TLV will contain the latest LSP State Database Version Number. If a PCEP speaker's LSP state database did not survive the restart of a PCEP session, the PCEP speaker MUST NOT include the LSP-DB-VERSION TLV in the OPEN object.

If both PCEP speakers include the LSP-DB-VERSION TLV in the OPEN Object and the TLV values match, the PCC MAY skip state synchronization. Otherwise, the PCC MUST perform full state synchronization (see [I-D.ietf-pce-stateful-pce]) or incremental state synchronization (see Section 4) to the stateful PCE. If the PCC attempts to skip state synchronization, by setting the SYNC Flag to 0 and PLSP-ID to a non-zero value on the first LSP State Report from the PCC as per [I-D.ietf-pce-stateful-pce], the PCE MUST send

back a PCErr with Error-Type 20 Error-Value TBD (suggested value - 2) 'LSP Database version mismatch', and close the PCEP session.

If state synchronization is required, then prior to completing the initialization phase, the PCE MUST mark any LSPs in the LSP database that were previously reported by the PCC as stale. When the PCC reports an LSP during state synchronization, if the LSP already exists in the LSP database, the PCE MUST update the LSP database and clear the stale marker from the LSP. When it has finished state synchronization, the PCC MUST immediately send an end of synchronization marker. The end of synchronization marker is a Path Computation State Report (PCRpt) message with an LSP object containing a PLSP-ID of 0 and with the SYNC flag set to 0 ([I-D.ietf-pce-stateful-pce]). The LSP-DB-VERSION TLV MUST be included in this PCRpt message. On receiving this state report, the PCE MUST purge any LSPs from the LSP database that are still marked as stale.

Note that a PCE/PCC MAY force state synchronization by not including the LSP-DB-VERSION TLV in its OPEN object.

Since a PCE does not make changes to the LSP State Database Version Number, a PCC should never encounter this TLV in a message from the PCE (other than the OPEN message). A PCC SHOULD ignore the LSP-DB-VERSION TLV, were it to receive one from a PCE.

Figure 1 shows an example sequence where the state synchronization is skipped.

```
                     +-+-+                    +-+-+
                     |PCC|                    |PCE|
                     +-+-+                    +-+-+
                       |                        |
                       |--Open--,               |
                       |  DBv=42 \    ,---Open--|
                       |    S=1   \  /   DBv=42 |
                       |           \/      S=1  |
                       |           /\           |
                       |          /   `-------->| (OK to skip sync)
            (Skip sync) |<--------`              |
                       |            .           |
                       |            .           |
                       |            .           |
                       |                        |
                       |--PCRpt,DBv=43,SYNC=0-->| (Regular
                       |                        |  LSP State Report)
                       |--PCRpt,DBv=44,SYNC=0-->| (Regular
                       |                        |  LSP State Report)
                       |--PCRpt,DBv=45,SYNC=0-->|
                       |                        |
```

Figure 1: State Synchronization Skipped

Figure 2 shows an example sequence where the state synchronization is performed due to LSP state database version mismatch during the PCEP session setup. Note that the same state synchronization sequence would happen if either the PCC or the PCE would not include the LSP-DB-VERSION TLV in their respective Open messages.

```
                     +-+-+                    +-+-+
                     |PCC|                    |PCE|
                     +-+-+                    +-+-+
                       |                        |
                       |--Open--,               |
                       |  DBv=46 \    ,---Open--|
                       |    S=1   \  /   DBv=42 |
                       |           \/      S=1  |
                       |           /\           |
                       |          /   `-------->| (Expect sync)
            (Do sync) |<--------`               |
                       |                        |
                       |--PCRpt,DBv=46,SYNC=1-->| (Sync start)
                       |            .           |
                       |            .           |
                       |            .           |
                       |--PCRpt,DBv=46,SYNC=0-->| (Sync done)
                       |            .           |(Purge LSP State
                       |            .           | if applicable)
                       |            .           |
                       |--PCRpt,DBv=47,SYNC=0-->| (Regular
                       |                        |  LSP State Report)
                       |--PCRpt,DBv=48,SYNC=0-->| (Regular
                       |                        |  LSP State Report)
                       |--PCRpt,DBv=49,SYNC=0-->|
                       |                        |
```

Figure 2: State Synchronization Performed

Figure 3 shows an example sequence where the state synchronization is skipped, but because one or both PCEP speakers set the S Flag to 0, the PCC does not send LSP-DB-VERSION TLVs in subsequent PCRpt messages to the PCE. If the current PCEP session restarts, the PCEP speakers will have to perform state synchronization, since the PCE does not know the PCC's latest LSP State Database Version Number information.

```
                     +-+-+                    +-+-+
                     |PCC|                    |PCE|
                     +-+-+                    +-+-+
                       |                        |
                       |--Open--,               |
                       |  DBv=42 \    ,---Open--|
                       |    S=0   \  /   DBv=42 |
                       |           \/      S=0  |
                       |           /\           |
                       |          /   `-------->| (OK to skip sync)
           (Skip sync) |<--------`              |
                       |            .           |
                       |            .           |
                       |            .           |
                       |------PCRpt,SYNC=0----->| (Regular
                       |                        |  LSP State Report)
                       |------PCRpt,SYNC=0----->| (Regular
                       |                        |  LSP State Report)
                       |------PCRpt,SYNC=0----->|
                       |                        |
```

Figure 3: State Synchronization Skipped, no LSP-DB-VERSION TLVs sent from PCC

3.3. PCEP Extensions

A new INCLUDE-DB-VERSION (S) bit is added in the stateful capabilities TLV (see Section 7 for details).

3.3.1. LSP State Database Version Number TLV

The LSP State Database Version Number (LSP-DB-VERSION) TLV is an optional TLV that MAY be included in the OPEN object and the LSP object.

The format of the LSP-DB-VERSION TLV is shown in the following figure:

```
      0                   1                   2                   3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |           Type=TBD            |            Length=8           |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                 LSP State DB Version Number                   |
     |                                                               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: LSP-DB-VERSION TLV format

The type of the TLV is TBD and it has a fixed length of 8 octets. The value contains a 64-bit unsigned integer, representing the LSP State DB Version Number.

3.3.2. Speaker Entity Identifier TLV

The Speaker Entity Identifier TLV (SPEAKER-ENTITY-ID) is an optional TLV that MAY be included in the OPEN Object when a PCEP speaker wishes to determine if state synchronization can be skipped when a PCEP session is restarted. It contains a unique identifier for the node that does not change during the lifetime of the PCEP speaker. It identifies the PCEP speaker to its peers even if the speaker's IP address is changed.

In case of a remote peer IP address change, a PCEP speaker would learn the speaker entity identifier on receiving the open message but it MAY have already sent its open message without realizing that it is a known PCEP peer. In such a case, either a full synchronization is done or PCEP session is terminated. This may be a local policy decision. The new IP address is associated with the speaker entity identifier for future either way. In the latter case when PCEP session is re-established, it would be correctly associated with speaker entity identifier and not be considered as an unknown peer.

The format of the SPEAKER-ENTITY-ID TLV is shown in the following figure:

```
      0                   1                   2                   3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |           Type=TBD            |       Length (variable)       |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                                                               |
     //                 Speaker Entity Identifier                   //
     |                                                               |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: SPEAKER-ENTITY-ID TLV format

The type of the TLV is TBD and it has a variable length, which MUST be greater than 0. The Value is padded to 4-octet alignment. The padding is not included in the Length field. The value contains the entity identifier of the speaker transmitting this TLV. This identifier is required to be unique within its scope of visibility, which is usually limited to a single domain. It MAY be configured by the operator. Alternatively, it can be derived automatically from a suitably-stable unique identifier, such as a MAC address, serial number, Traffic Engineering Router ID, or similar. In the case of inter-domain connections, the speaker SHOULD prefix its usual identifier with the domain identifier of its residence, such as Autonomous System number, IGP area identifier, or similar.

The relationship between this identifier and entities in the Traffic Engineering database is intentionally left undefined.

From a manageability point of view, a PCE or PCC implementation SHOULD allow the operator to configure this Speaker Entity Identifier.

4. Incremental State Synchronization

[I-D.ietf-pce-stateful-pce] describes the LSP state synchronization mechanism between PCCs and stateful PCEs. During the state synchronization, a PCC sends the information of all its LSPs (i.e., the full LSP-DB) to the stateful PCE. In order to reduce the state synchronization overhead when there is a small number of LSP state change in the network between PCEP session restart, this section defines a mechanism for incremental (Delta) LSP Database (LSP-DB) synchronization.

4.1. Motivation

According to [I-D.ietf-pce-stateful-pce], if a PCE restarts and its LSP-DB survived, PCCs with mismatched LSP State Database Version Number will send all their LSPs information (full LSP-DB) to the stateful PCE, even if only a small number of LSPs underwent state change. It can take a long time and consume large communication channel bandwidth.

Figure 6 shows an example of LSP state synchronization.

```
                  +-----+
                  | PCE |
                  +-----+
                 /
                /
               /
              /
       +-------+            +-------+
       | PCC1 |------------| PCC2 |
       +-------+            +-------+
          |                    |
          |                    |
       +-------+            +-------+
       | PCC3 |------------| PCC4 |
       +-------+            +-------+
```

Figure 6: Topology Example

Assuming there are 320 LSPs in the network, with each PCC having 80 LSPs. During the time when the PCEP session is down, 20 LSPs of each PCC (i.e., 80 LSPs in total), are changed. Hence when PCEP session restarts, the stateful PCE needs to synchronize 320 LSPs with all PCCs. But actually, 240 LSPs stay the same. If performing full LSP state synchronization, it can take a long time to carry out the synchronization of all LSPs. It is especially true when only a low bandwidth communication channel is available (e.g., in-band control channel for optical transport networks) and there is a substantial number of LSPs in the network. Another disadvantage of full LSP synchronization is that it is a waste of communication bandwidth to perform full LSP synchronization given the fact that the number of LSP changes can be small during the time when PCEP session is down.

An incremental (Delta) LSP Database (LSP-DB) state synchronization is described in this section, where only the LSPs underwent state change are synchronized between the session restart. This may include new/modified/deleted LSPs.

4.2. Incremental Synchronization Procedure

[I-D.ietf-pce-stateful-pce] describes state synchronization and Section 3 describes state synchronization avoidance by using LSP-DB-VERSION TLV in its OPEN object. This section extends this idea to only synchronize the delta (changes) in case of version mismatch.

If both PCEP speakers include the LSP-DB-VERSION TLV in the OPEN object and the LSP-DB-VERSION TLV values match, the PCC MAY skip state synchronization. Otherwise, the PCC MUST perform state synchronization. Incremental State synchronization capability is advertised on a PCEP session during session startup using the DELTA-LSP-SYNC-CAPABILITY (D) bit in the capabilities TLV (see Section 7). Instead of dumping full LSP-DB to the stateful PCE again, the PCC synchronizes the delta (changes) as described in Figure 7 when D flag and S flag is set to 1 by both PCC and PCE. Other combinations of D and S flags setting by PCC and PCE result in full LSP-DB synchronization procedure as described in [I-D.ietf-pce-stateful-pce]. The PCC MAY force a full LSP DB synchronization by setting the D flag to zero in the OPEN message.

```
                       +-+-+                   +-+-+
                       |PCC|                   |PCE|
                       +-+-+                   +-+-+
                         |                       |
                         |--Open--,              |
                         |  DBv=46 \    ,---Open--|
                         |    S=1   \  /   DBv=42 |
                         |    D=1    \/      S=1  |
                         |           /\      D=1  |
                         |          /  \          |
                         |         /    `-------->| (Expect Delta sync)
                (Do sync)|<--------`              | (DONOT Purge LSP
                (Delta)  |                        | State)
                         |                        |
     (Delta Sync starts) |--PCRpt,DBv=46,SYNC=1-->|
                         |            .           |
                         |            .           |
                         |            .           |
                         |            .           |
                         |--PCRpt,DBv=46,SYNC=0-->| (Sync done,
                         |                        | PLSP-ID=0)
                         |                        |
                         |--PCRpt,DBv=47,SYNC=0-->| (Regular
                         |                        |  LSP State Report)
                         |--PCRpt,DBv=48,SYNC=0-->| (Regular
                         |                        |  LSP State Report)
                         |--PCRpt,DBv=49,SYNC=0-->|
                         |                        |
```

Figure 7: Incremental Synchronization Procedure

As per Section 3, the LSP State Database Version Number is incremented each time a change is made to the PCC's local LSP State Database. Each LSP is associated with the DB version at the time of its state change. This is needed to determine which LSP and what information needs to be synchronized in incremental state synchronization.

It is not necessary for a PCC to store a complete history of LSP Database change, but rather remember the LSP state changes (including LSP modification, setup and deletion) that happened between the PCEP session(s) restart in order to carry out incremental state synchronization. After the synchronization procedure finishes, the PCC can dump this history information. In the example shown in

Figure 7, the PCC needs to store the LSP state changes that happened between DB Version 43 to 46 and synchronizes these changes only when performing incremental LSP state update. So a PCC needs to remember at least the LSP state changes that happened after an existing PCEP session with a stateful PCE goes down to have any chance of doing incremental synchronisation when the session is re-established.

If a PCC finds out it does not have sufficient information to complete incremental synchronisation after advertising incremental LSP state synchronization capability, it MUST send a PCErr with Error-Type 20 and Error-Value 5 'A PCC indicates to a PCE that it can not complete the state synchronization' (defined in [I-D.ietf-pce-stateful-pce]) and terminate the session. The PCC SHOULD re-establish the session with the D bit set to 0 in the OPEN message.

The other procedures and error checks remain unchanged from the full state synchronization ([I-D.ietf-pce-stateful-pce]).

5. PCE-triggered Initial Synchronization

5.1. Motivation

In networks such as optical transport networks, the control channel between network nodes can be realized through in-band overhead thus has limited bandwidth. With a stateful PCE connected to the network via one network node, it is desirable to control the timing of PCC state synchronization so as not to overload the low communication channel available in the network during the initial synchronization (be it incremental or full) when the session restarts , when there is comparatively large amount of control information needing to be synchronized between the stateful PCE and the network. The method proposed, i.e., allowing PCE to trigger the state synchronization, is similar to the function proposed in Section 6 but is used in different scenarios and for different purposes.

5.2. PCE-triggered Initial State Synchronization Procedure

Support of PCE-triggered initial state synchronization is advertised during session startup using the TRIGGERED-INITIAL-SYNC (F) bit in the STATEFUL-PCE-CAPABILITY TLV (see Section 7).

In order to allow a stateful PCE to control the LSP-DB synchronization after establishing a PCEP session, both PCEP speakers MUST set F bit to 1 in the OPEN message. If the TRIGGERED-INITIAL-SYNC capability is not advertised by a PCE and the PCC receives a PCUpd with the SYNC flag set to 1, it MUST send a PCErr with the SRP-ID-number of the PCUpd, Error-Type 20 and Error-Value TBD (suggested

value - 4) 'Attempt to trigger synchronization when the TRIGGERED-SYNC capability has not been advertised' (see Section 8.1). If the LSP-DB Version is mis-matched, it can send a PCUpd message with PLSP-ID = 0 and SYNC = 1 in order to trigger the LSP-DB synchronization process. In this way, the PCE can control the sequence of LSP synchronization among all the PCCs that are re-establishing PCEP sessions with it. When the capability of PCE control is enabled, only after a PCC receives this message, it will start sending information to the PCE. The PCC SHOULD NOT send PCRpt messages to the stateful PCE before it triggers the State Synchronization. This PCE-triggering capability can be applied to both full and incremental state synchronization. If applied to the later, the PCCs only send information that PCE does not possess, which is inferred from the LSP-DB version information exchanged in the OPEN message (see Section 4.2 for detailed procedure).

Once the initial state synchronization is triggered by the PCE, the procedures and error checks remain unchanged from the full state synchronization ([I-D.ietf-pce-stateful-pce]).

6. PCE-triggered Re-synchronization

6.1. Motivation

The accuracy of the computations performed by the PCE is tied to the accuracy of the view the PCE has on the state of the LSPs. Therefore, it can be beneficial to be able to re-synchronize this state even after the session has been established. The PCE may use this approach to continuously sanity check its state against the network, or to recover from error conditions without having to tear down sessions.

6.2. PCE-triggered State Re-synchronization Procedure

Support of PCE-triggered state synchronization is advertised by both PCEP speakers during session startup using the TRIGGERED-RESYNC (T) bit in the STATEFUL-PCE-CAPABILITY TLV (see Section 7). The PCE can choose to re-synchronize its entire LSP database or a single LSP.

To trigger re-synchronization for an LSP, the PCE MUST first mark the LSP as stale and then send a Path Computation State Update (PCUpd) for it, with the SYNC flag in the LSP object set to 1. The PCE SHOULD NOT include any parameter updates for the LSP, and the PCC SHOULD ignore such updates if the SYNC flag is set. The PCC MUST respond with a PCRpt message with the LSP state, SYNC Flag set to 0 and MUST include the SRP-ID-number of the PCUpd message that triggered the resynchronization.

The PCE can also trigger re-synchronization of the entire LSP database. The PCE MUST first mark all LSPs in the LSP database that were previously reported by the PCC as stale and then send a PCUpd with an LSP object containing a PLSP-ID of 0 and with the SYNC flag set to 1. This PCUpd message is the trigger for the PCC to enter the synchronization phase as described in [I-D.ietf-pce-stateful-pce] and start sending PCRpt messages. After the receipt of the end-of-synchronization marker, the PCE will purge LSPs which were not refreshed. The SRP-ID-number of the PCUpd that triggered the re-synchronization SHOULD be included in each of the PCRpt messages.

If the TRIGGERED-RESYNC capability is not advertised by a PCE and the PCC receives a PCUpd with the SYNC flag set to 1, it MUST send a PCErr with the SRP-ID-number of the PCUpd, Error-Type 20 and Error-Value TBD (suggested value - 4) 'Attempt to trigger synchronization when the TRIGGERED-SYNC capability has not been advertised' (see Section 8.1).

Once the state re-synchronization is triggered by the PCE, the procedures and error checks remain unchanged from the full state synchronization ([I-D.ietf-pce-stateful-pce]). This would also include PCE triggering multiple state re-synchronization requests while synchronization is in progress.

7. Advertising Support of Synchronization Optimizations

Support for each of the optimizations described in this document requires advertising the corresponding capabilities during session establishment time.

New flags are defined for the STATEFUL-PCE-CAPABILITY TLV defined in [I-D.ietf-pce-stateful-pce]. Its format is shown in the following figure:

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |               Type            |            Length=4           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                             Flags                 |F|D|T|I|S|U|
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8: STATEFUL-PCE-CAPABILITY TLV Format

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

U (LSP-UPDATE-CAPABILITY - 1 bit): defined in [I-D.ietf-pce-stateful-pce].

S (INCLUDE-DB-VERSION - 1 bit): if set to 1 by both PCEP Speakers, the PCC will include the LSP-DB-VERSION TLV in each LSP Object. See Section 3.2 for details.

I (LSP-INSTANTIATION-CAPABILITY - 1 bit): defined in [I-D.ietf-pce-pce-initiated-lsp].

T (TRIGGERED-RESYNC - 1 bit): if set to 1 by both PCEP Speakers, the PCE can trigger re-synchronization of LSPs at any point in the life of the session. See Section 6.2 for details.

D (DELTA-LSP-SYNC-CAPABILITY - 1 bit): if set to 1 by a PCEP speaker, it indicates that the PCEP speaker allows incremental (delta) state synchronization. See Section 4.2 for details.

F (TRIGGERED-INITIAL-SYNC - 1 bit): if set to 1 by both PCEP Speakers, the PCE SHOULD trigger initial (first) state synchronization. See Section 5.2 for details.

8. IANA Considerations

This document requests IANA actions to allocate code points for the protocol elements defined in this document.

8.1. PCEP-Error Object

IANA is requested to make the following allocation in the "PCEP-ERROR Object Error Types and Values" registry.

Error-Type	Meaning	Reference
6	Mandatory Object missing	[RFC5440]
	Error-Value= TBD(suggested value 12): LSP-DB-VERSION TLV missing	This document
20	LSP State synchronization error	[I-D.ietf-pce-stateful-pce]
	Error-Value= TBD(suggested value 2): LSP Database version mismatch.	This document
	Error-Value=TBD(suggested value 3): The LSP-DB-VERSION TLV Missing when state synchronization avoidance is enabled.	This document
	Error-Value=TBD(suggested value 4): Attempt to trigger a synchronization when the PCE triggered synchronization capability has not been advertised.	This document
	Error-Value=TBD(suggested value 6): No sufficient LSP change information for incremental LSP state synchronization.	This document
	Error-Value=TBD(suggested value 7): Received an invalid LSP DB Version Number	This document

8.2. PCEP TLV Type Indicators

IANA is requested to make the following allocation in the "PCEP TLV Type Indicators" registry.

Value	Meaning	Reference
TBD(suggested value 23)	LSP-DB-VERSION	This document
TBD(suggested value 24)	SPEAKER-ENTITY-ID	This document

8.3. STATEFUL-PCE-CAPABILITY TLV

The STATEFUL-PCE-CAPABILITY TLV is defined in [I-D.ietf-pce-stateful-pce] and a registry is requested to be created to manage the flags in the TLV. IANA is requested to make the following allocation in the aforementioned registry.

Bit	Description	Reference
TBD(suggested value 26)	TRIGGERED-INITIAL-SYNC	This document
TBD(suggested value 27)	DELTA-LSP-SYNC-CAPABILITY	This document
TBD(suggested value 28)	TRIGGERED-RESYNC	This document
TBD(suggested value 30)	INCLUDE-DB-VERSION	This document

9. Manageability Considerations

All manageability requirements and considerations listed in [RFC5440] and [I-D.ietf-pce-stateful-pce] apply to PCEP protocol extensions defined in this document. In addition, requirements and considerations listed in this section apply.

9.1. Control of Function and Policy

A PCE or PCC implementation MUST allow configuring the state synchronization optimization capabilities as described in this document. The implementation SHOULD also allow the operator to configure the Speaker Entity Identifier (Section 3.3.2).

9.2. Information and Data Models

An implementation SHOULD allow the operator to view the stateful capabilities advertised by each peer, and the current synchronization status with each peer. To serve this purpose, the PCEP MIB module can be extended to include advertised stateful capabilities, and synchronization status.

9.3. Liveness Detection and Monitoring

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

9.4. Verify Correct Operations

Mechanisms defined in this document do not imply any new operation verification requirements in addition to those already listed in [RFC5440] and [I-D.ietf-pce-stateful-pce].

9.5. Requirements On Other Protocols

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

9.6. Impact On Network Operations

Mechanisms defined in this document do not have any impact on network operations in addition to those already listed in [RFC5440] and [I-D.ietf-pce-stateful-pce].

10. Security Considerations

The security considerations listed in [I-D.ietf-pce-stateful-pce] apply to this document as well. However, because the protocol modifications outlined in this document allow the PCE to control state (re)-synchronization timing and sequence, it also introduces a new attack vector: an attacker may flood the PCC with triggered re-synchronization request at a rate which exceeds the PCC's ability to process them, either by spoofing messages or by compromising the PCE itself. The PCC is free to drop any trigger re-synchronization request without additional processing.

11. Acknowledgements

We would like to thank Young Lee, Jonathan Hardwick, Sergio Belotti and Cyril Margaria for their comments and discussions.

12. Contributors

Gang Xie
Huawei Technologies
F3-5-B R&D Center, Huawei Industrial Base, Bantian, Longgang District
Shenzhen, Guangdong, 518129
P.R. China
Email: xiegang09@huawei.com

13. References

13.1. Normative References

[I-D.ietf-pce-stateful-pce]
Crabbe, E., Minei, I., Medved, J., and R. Varga, "PCEP Extensions for Stateful PCE", draft-ietf-pce-stateful-pce-18 (work in progress), December 2016.

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

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <http://www.rfc-editor.org/info/rfc5440>.

13.2. Informative References

[I-D.ietf-pce-pce-initiated-lsp]
Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", draft-ietf-pce-pce-initiated-lsp-07 (work in progress), July 2016.

Authors' Addresses

Edward Crabbe
Oracle

EMail: edward.crabbe@gmail.com

Ina Minei
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043
US

EMail: inaminei@google.com

Jan Medved
Cisco Systems, Inc.
170 West Tasman Dr.
San Jose, CA 95134
US

EMail: jmedved@cisco.com

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

EMail: robert.varga@pantheon.sk

Xian Zhang
Huawei Technologies
F3-5-B R&D Center, Huawei Industrial Base, Bantian, Longgang District
Shenzhen, Guangdong 518129
P.R.China

EMail: zhang.xian@huawei.com

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

EMail: dhruv.ietf@gmail.com