PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 29, 2017

D. Dhody
Y. Lee
Huawei Technologies
D. Ceccarelli
Ericsson

June 27, 2017

PCEP Extension for Distribution of Link-State and TE Information. draft-dhodylee-pce-pcep-ls-08

Abstract

In order to compute and provide optimal paths, Path Computation Elements (PCEs) require an accurate and timely Traffic Engineering Database (TED). Traditionally this TED has been obtained from a link state (LS) routing protocol supporting traffic engineering extensions.

This document extends the Path Computation Element Communication Protocol (PCEP) with Link-State and TE Information.

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on December 29, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 3
1.1. Requirements Language 5
2. Terminology 5
3. Applicability 5
4. Requirements for PCEP extension 6
5. New Functions to distribute link-state (and TE) via PCEP . . 7
6. Overview of Extension to PCEP 7
6.1. New Messages 7
6.2. Capability Advertisement 7
6.3. Initial Link-State (and TE) Synchronization 8
6.3.1. Optimizations for LS Synchronization 10
6.4. LS Report 11
7. Transport 11
8. PCEP Messages 11
8.1. LS Report Message 11
8.2. The PCErr Message 12
9. Objects and TLV 12
9.1. TLV Format 13
9.2. Open Object 13
9.2.1. LS Capability TLV 13
9.3. LS Object 14
9.3.1. Routing Universe TLV 15
9.3.2. Route Distinguisher TLV 16
9.3.3. Virtual Network TLV 17
9.3.4. Local Node Descriptors TLV 17
9.3.5. Remote Node Descriptors TLV 18
9.3.6. Node Descriptors Sub-TLVs 19
9.3.7. Link Descriptors TLV 19
9.3.8. Prefix Descriptors TLV 20
9.3.9. PCEP-LS Attributes 21
9.3.9.1. Node Attributes TLV 21
9.3.9.2. Link Attributes TLV 22
9.3.9.3. Prefix Attributes TLV 24
9.3.10. Removal of an Attribute 25
10. Other Considerations 25
10.1. Inter-AS Links 25
11. Security Considerations 25
12. Manageability Considerations 25
12.1. Control of Function and Policy 26
12.2. Information and Data Models 26

12.3. Liveness Detection and Monitoring 26
12.4. Verify Correct Operations 27
12.5. Requirements On Other Protocols 27
12.6. Impact On Network Operations 27
13. IANA Considerations 27
13.1. PCEP Messages 27
13.2. PCEP Objects 27
13.3. LS Object 28
13.4. PCEP-Error Object 28
13.5. PCEP TLV Type Indicators 29
13.6. PCEP-LS Sub-TLV Type Indicators 29
14. TLV/Sub-TLV Code Points Summary 32
15. Implementation Status 32
15.1. Hierarchical Transport PCE controllers 32
15.2. ONOS-based Controller (MDSC and PNC) 33
16. Acknowledgments 33
17. References 33
17.1. Normative References 33
17.2. Informative References 34
Appendix A. Relevant OSPF TLV and sub-TLV 38
Appendix B. Examples 39
B.1. All Nodes 39
B.2. Designated Node 40
B.3. Between PCEs 40
Appendix C. Contributor Addresses 42
Authors' Addresses 42

1. Introduction

In Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS), a Traffic Engineering Database (TED) is used in computing paths for connection oriented packet services and for circuits. The TED contains all relevant information that a Path Computation Element (PCE) needs to perform its computations. It is important that the TED be complete and accurate each time, the PCE performs a path computation.

In MPLS and GMPLS, interior gateway routing protocols (IGPs) have been used to create and maintain a copy of the TED at each node running the IGP. One of the benefits of the PCE architecture [RFC4655] is the use of computationally more sophisticated path computation algorithms and the realization that these may need enhanced processing power not necessarily available at each node participating in an IGP.

Section 4.3 of [RFC4655] describes the potential load of the TED on a network node and proposes an architecture where the TED is maintained by the PCE rather than the network nodes. However, it does not

describe how a PCE would obtain the information needed to populate its TED. PCE may construct its TED by participating in the IGP ([RFC3630] and [RFC5305] for MPLS-TE; [RFC4203] and [RFC5307] for GMPLS). An alternative is offered by BGP-LS [RFC7752] .

[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. PCC can delegate the rights to modify the LSP parameters to an Active Stateful PCE. This requires PCE to quickly be updated on any changes in the Topology and TEDB, so that PCE can meet the need for updating LSPs effectively and in a timely manner. The fastest way for a PCE to be updated on TED changes is via a direct interface with each network node and with incremental update from each network node with only the attribute that is modified.

[I-D.ietf-pce-pce-initiated-lsp] describes the setup, maintenance and teardown of PCE-initiated LSPs under the stateful PCE model, without the need for local configuration on the PCC, thus allowing for a dynamic network that is centrally controlled and deployed. This model requires timely topology and TED update at the PCE.

[RFC5440] describes the specifications for the Path Computation Element Communication Protocol (PCEP). PCEP specifies the communication between a Path Computation Client (PCC) and a Path Computation Element (PCE), or between two PCEs based on the PCE architecture [RFC4655].

This document describes a mechanism by which Link State and TE information can be collected from networks and shared with PCE using the PCEP itself. This is achieved using a new PCEP message format. The mechanism is applicable to physical and virtual links as well as further subjected to various policies.

A network node maintains one or more databases for storing link-state and TE information about nodes and links in any given area. Link attributes stored in these databases include: local/remote IP addresses, local/ remote interface identifiers, link metric and TE metric, link bandwidth, reservable bandwidth, per CoS class reservation state, preemption and Shared Risk Link Groups (SRLG). The node's PCEP process can retrieve topology from these databases and distribute it to a PCE, either directly or via another PCEP Speaker, using the encoding specified in this document.

Further [RFC6805] describes Hierarchical-PCE architecture, where a parent PCE maintains a domain topology map. To build this domain topology map, the child PCE can carry the border nodes and inter-

domain link information to the parent PCE using the mechanism described in this document. Further as described in [I-D.ietf-pce-applicability-actn], the child PCE can also transport abstract Link-State and TE information from child PCE to a Parent PCE using the mechanism described in this document to build an abstract topology at the parent PCE.

[I-D.ietf-pce-stateful-pce] describe LSP state synchronization between PCCs and PCEs in case of stateful PCE. This document does not make any change to the LSP state synchronization process. The mechanism described in this document are on top of the existing LSP state synchronization.

1.1. Requirements Language

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

2. Terminology

The terminology is as per [RFC4655] and [RFC5440].

3. Applicability

The mechanism specified in this draft is applicable to deployments:

- o Where there is no IGP or BGP-LS running in the network.

- o Where there is no IGP or BGP-LS running at the PCE to learn link-state and TE information.

- o Where there is IGP or BGP-LS running but with a need for a faster TE and link-state population and convergence at the PCE.

  - * A PCE may receive partial information (say basic TE, link-state) from IGP and other information (optical and impairment) from PCEP.

  - * A PCE may receive an incremental update (as opposed to the entire information of the node/link).

  - * A PCE may receive full information from both existing mechanism (IGP or BGP) and PCEP.

- o Where there is a need for transporting (abstract) Link-State and TE information from child PCE to a Parent PCE in H-PCE [RFC6805]; as well as for Physical Network Controller (PNC) to Multi-Domain

Service Coordinator (MDSC) in Abstraction and Control of TE Networks (ACTN) [I-D.ietf-teas-actn-framework].

A PCC may further choose to send only local information or both local and remote learned information.

How a PCE manages the link-state (and TE) information is implementation specific and thus out of scope of this document.

The prefix information in PCEP-LS can also help in determining the domain of the endpoints in H-PCE (and ACTN). Section 4.5 of [RFC6805] describe various mechanism and procedures that might be used, PCEP-LS provides a simple mechanism to exchange this information.

4. Requirements for PCEP extension

Following key requirements associated with link-state (and TE) distribution are identified for PCEP:

1. The PCEP speaker supporting this draft MUST be a mechanism to advertise the Link-State (and TE) distribution capability.

2. PCC supporting this draft MUST have the capability to report the link-state (and TE) information to the PCE. This includes self originated information and remote information learned via routing protocols. PCC MUST be capable to do the initial bulk sync at the time of session initialization as well as changes after.

3. A PCE MAY learn link-state (and TE) from PCEP as well as from existing mechanism like IGP/BGP-LS. PCEP extension MUST have a mechanism to link the information learned via other means. There MUST NOT be any changes to the existing link-state (and TE) population mechanism via IGP/BGP-LS. PCEP extension SHOULD keep the properties in a protocol (IGP or BGP-LS) neutral way, such that an implementation may not need to know about any OSPF or IS-IS or BGP protocol specifics.

4. It SHOULD be possible to encode only the changes in link-state (and TE) properties (after the initial sync) in PCEP messages.

5. The same mechanism should be used for both MPLS TE as well as GMPLS, optical and impairment aware properties.

6. The same mechanism should be used for PCE to PCE Link-state (and TE) synchronization.

5. New Functions to distribute link-state (and TE) via PCEP

Several new functions are required in PCEP to support distribution of link-state (and TE) information. 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:

- o Capability advertisement (E-C,C-E): both the PCC and the PCE must announce during PCEP session establishment that they support PCEP extensions for distribution of link-state (and TE) information defined in this document.

- o Link-State (and TE) synchronization (C-E): after the session between the PCC and a PCE is initialized, the PCE must learn Link-State (and TE) information before it can perform path computations. In case of stateful PCE it is RECOMENDED that this operation be done before LSP state synchronization.

- o Link-State (and TE) Report (C-E): a PCC sends a LS (and TE) report to a PCE whenever the Link-State and TE information changes.

6. Overview of Extension to PCEP

6.1. New Messages

In this document, we define a new PCEP messages called LS Report (LSRpt), a PCEP message sent by a PCC to a PCE to report link-state (and TE) information. Each LS Report in a LSRpt message can contain the node or link properties. An unique PCEP specific LS identifier (LS-ID) is also carried in the message to identify a node or link and that remains constant for the lifetime of a PCEP session. This identifier on its own is sufficient when no IGP or BGP-LS running in the network for PCE to learn link-state (and TE) information. Incase PCE learns some information from PCEP and some from the existing mechanism, the PCC SHOULD include the mapping of IGP or BGP-LS identifier to map the information populated via PCEP with IGP/BGP-LS. See Section 8.1 for details.

6.2. Capability Advertisement

During PCEP Initialization Phase, PCEP Speakers (PCE or PCC) advertise their support of LS (and TE) distribution via PCEP extensions. A PCEP Speaker includes the "LS Capability" TLV, described in Section 9.2.1, in the OPEN Object to advertise its support for PCEP-LS extensions. The presence of the LS Capability TLV in PCC's OPEN Object indicates that the PCC is willing to send LS Reports whenever local link-state (and TE) information changes. The presence of the LS Capability TLV in PCE's OPEN message indicates

that the PCE is interested in receiving LS Reports whenever local link-state (and TE) information changes.

The PCEP protocol extensions for LS (and TE) distribution MUST NOT be used if one or both PCEP Speakers have not included the LS Capability TLV in their respective OPEN message. If the PCE that supports the extensions of this draft but did not advertise this capability, then upon receipt of a LSRpt message from the PCC, it SHOULD generate a PCErr with error-type 19 (Invalid Operation), error-value TBD1 (Attempted LS Report if LS capability was not advertised) and it will terminate the PCEP session.

The LS reports sent by PCC MAY carry the remote link-state (and TE) information learned via existing means like IGP and BGP-LS only if both PCEP Speakers set the R (remote) Flag in the "LS Capability" TLV to 'Remote Allowed (R Flag = 1)'. If this is not the case and LS reports carry remote link-state (and TE) information, then a PCErr with error-type 19 (Invalid Operation) and error-value TBD1 (Attempted LS Report if LS remote capability was not advertised) and it will terminate the PCEP session.

6.3. Initial Link-State (and TE) Synchronization

The purpose of LS Synchronization is to provide a checkpoint-in- time state replica of a PCC's link-state (and TE) data base in a PCE. State Synchronization is performed immediately after the Initialization phase (see [RFC5440]]). In case of stateful PCE ([I-D.ietf-pce-stateful-pce]) it is RECOMENDED that the LS synchronization should be done before LSP state synchronization.

During LS Synchronization, a PCC first takes a snapshot of the state of its database, then sends the snapshot to a PCE in a sequence of LS Reports. Each LS Report sent during LS Synchronization has the SYNC Flag in the LS Object set to 1. The end of synchronization marker is a LSRpt message with the SYNC Flag set to 0 for an LS Object with LS-ID equal to the reserved value 0. If the PCC has no link-state to synchronize, it will only send the end of synchronization marker.

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 state it received from this PCC. The session re-establishment MUST be re-attempted per the procedures defined in [RFC5440], including use of a back-off timer.

If the PCC encounters a problem which prevents it from completing the LS synchronization, it MUST send a PCErr message with error-type TBD2

(LS Synchronization Error) and error-value 2 (indicating an internal PCC error) to the PCE and terminate the session.

The PCE does not send positive acknowledgements for properly received LS synchronization messages. It MUST respond with a PCErr message with error-type TBD2 (LS Synchronization Error) and error-value 1 (indicating an error in processing the LSRpt) if it encounters a problem with the LS Report it received from the PCC and it MUST terminate the session.

The LS reports can carry local as well as remote link-state (and TE) information depending on the R flag in LS capability TLV.

The successful LS Synchronization sequences is shown in Figure 1.

```
      +-+-+                     +-+-+
      |PCC|                     |PCE|
      +-+-+                     +-+-+
        |                         |
        |-----LSRpt, SYNC=1----->| (Sync start)
        |                         |
        |-----LSRpt, SYNC=1----->|
        |            .            |
        |            .            |
        |            .            |
        |-----LSRpt, SYNC=1----->|
        |            .            |
        |            .            |
        |            .            |
        |                         |
        |-----LSRpt, SYNC=0----->| (End of sync marker
        |                         |  LS Report
        |                         |  for LS-ID=0)
        |                         | (Sync done)
```

Figure 1: Successful LS synchronization

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

```
         +-+-+                    +-+-+
         |PCC|                    |PCE|
         +-+-+                    +-+-+
           |                        |
           |-----LSRpt, SYNC=1----->|
           |                        |
           |-----LSRpt, SYNC=1----->|
           |            .           |
           |            .           |
           |            .           |
           |-----LSRpt, SYNC=1----->|
           |                        |
           |---LSRpt,SYNC=1         |
           |         \    ,-PCErr---|
           |          \  /          |
           |           \/           |
           |           /\           |
           |          /   `-------->| (Ignored)
           |<--------`              |
```

Figure 2: Failed LS synchronization (PCE failure)

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

```
         +-+-+                    +-+-+
         |PCC|                    |PCE|
         +-+-+                    +-+-+
           |                        |
           |-----LSRpt, SYNC=1----->|
           |                        |
           |-----LSRpt, SYNC=1----->|
           |            .           |
           |            .           |
           |            .           |
           |-------- PCErr--------->|
           |                        |
```

Figure 3: Failed LS synchronization (PCC failure)

6.3.1. Optimizations for LS Synchronization

These optimizations are described in [I-D.kondreddy-pce-pcep-ls-sync-optimizations].

6.4. LS Report

The PCC MUST report any changes in the link-state (and TE) information to the PCE by sending a LS Report carried on a LSRpt message to the PCE. Each node and Link would be uniquely identified by a PCEP LS identifier (LS-ID). The LS reports may carry local as well as remote link-state (and TE) information depending on the R flag in LS capability TLV. In case R flag is set, It MAY also include the mapping of IGP or BGP-LS identifier to map the information populated via PCEP with IGP/BGP-LS.

More details about LSRpt message are in Section 8.1.

7. Transport

A permanent PCEP session MUST be established between a PCE and PCC supporting link-state (and TE) distribution via PCEP. In the case of session failure, session re-establishment MUST be re-attempted per the procedures defined in [RFC5440].

8. PCEP Messages

As defined in [RFC5440], a PCEP message consists of a common header followed by a variable-length body made of a set of objects that can be either mandatory or optional. An object is said to be mandatory in a PCEP message when the object must be included for the message to be considered valid. For each PCEP message type, a set of rules is defined that specify the set of objects that the message can carry. An implementation MUST form the PCEP messages using the object ordering specified in this document.

8.1. LS Report Message

A PCEP LS Report message (also referred to as LSRpt message) is a PCEP message sent by a PCC to a PCE to report the link-state (and TE) information. A LSRpt message can carry more than one LS Reports. The Message-Type field of the PCEP common header for the LSRpt message is set to [TBD3].

The format of the LSRpt message is as follows:

```
<LSRpt Message> ::= <Common Header>
                    <ls-report-list>
Where:

<ls-report-list> ::= <LS>[<ls-report-list>]
```

The LS object is a mandatory object which carries LS information of a node or a link. Each LS object has an unique LS-ID as described in Section 9.3. If the LS object is missing, the receiving PCE MUST send a PCErr message with Error-type=6 (Mandatory Object missing) and Error-value=[TBD4] (LS object missing).

A PCE may choose to implement a limit on the LS information a single PCC can populate. If a LSRpt is received that causes the PCE to exceed this limit, it MUST send a PCErr message with error-type 19 (invalid operation) and error-value 4 (indicating resource limit exceeded) in response to the LSRpt message triggering this condition and SHOULD terminate the session.

8.2. The PCErr Message

If a PCEP speaker has advertised the LS capability on the PCEP session, the PCErr message MAY include the LS object. If the error reported is the result of an LS report, then the LS-ID number MUST be the one from the LSRpt that triggered the error.

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

The format of the PCErr message is 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> | <ls-id-list>]
           <error-obj-list>

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

<ls-id-list>::=<LS>[<ls-id-list>]

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

9. Objects and TLV

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

9.1. TLV Format

The TLV and the sub-TLV format (and padding) in this document, is as per section 7.1 of [RFC5440].

9.2. Open Object

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

9.2.1. LS Capability TLV

The LS-CAPABILITY TLV is an optional TLV for use in the OPEN Object for link-state (and TE) distribution via PCEP capability advertisement. 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=[TBD5]     |            Length=4           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Flags                           |R|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The type of the TLV is [TBD5] and it has a fixed length of 4 octets.

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

o R (remote - 1 bit): if set to 1 by a PCC, the R Flag indicates that the PCC allows reporting of remote LS information learned via other means like IGP and BGP-LS; if set to 1 by a PCE, the R Flag indicates that the PCE is capable of receiving remote LS information (from the PCC point of view). The R Flag must be advertised by both a PCC and a PCE for LSRpt messages to report remote as well as local LS information on a PCEP session. The TLVs related to IGP/BGP-LS identifier MUST be encoded when both PCEP speakers have the R Flag set.

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

Advertisement of the LS capability implies support of local link-state (and TE) distribution, as well as the objects, TLVs and procedures defined in this document.

9.3. LS Object

The LS (link-state) object MUST be carried within LSRpt messages and MAY be carried within PCErr messages. The LS object contains a set of fields used to specify the target node or link. It also contains a flag indicating to a PCE that the LS synchronization is in progress. The TLVs used with the LS object correlate with the IGP/BGP-LS encodings.

LS Object-Class is [TBD6].

Four Object-Type values are defined for the LS object so far:

o LS Node: LS Object-Type is 1.

o LS Link: LS Object-Type is 2.

o LS IPv4 Topology Prefix: LS Object-Type is 3.

o LS IPv6 Topology Prefix: LS Object-Type is 4.

The format of all types of LS object is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Protocol-ID  |          Flag                             |R|S|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          LS-ID                                |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                         TLVs                                //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Protocol-ID (8-bit): The field provide the source information. The protocol could be an IGP, BGP-LS or an abstraction algorithm. Incase PCC only provides local information of the PCC, it MUST use Protocol-ID as Direct. The following values are defined (some of them are same as [RFC7752]):

Protocol-ID	Source protocol
1	IS-IS Level 1
2	IS-IS Level 2
3	OSPFv2
4	Direct
5	Static configuration
6	OSPFv3
7	BGP-LS
8	PCEP-LS
9	Abstraction
10	Unspecified

Flags (24-bit):

- o S (SYNC - 1 bit): the S Flag MUST be set to 1 on each LSRpt sent from a PCC during LS Synchronization. The S Flag MUST be set to 0 in other LSRpt messages sent from the PCC.

- o R (Remove - 1 bit): On LSRpt messages the R Flag indicates that the node/link/prefix has been removed from the PCC and the PCE SHOULD remove from its database. Upon receiving an LS Report with the R Flag set to 1, the PCE SHOULD remove all state for the node/link/prefix identified by the LS Identifiers from its database.

LS-ID(64-bit): A PCEP-specific identifier for the node or link or prefix information. A PCC creates an unique LS-ID for each node/link/prefix that is constant for the lifetime of a PCEP session. The PCC will advertise the same LS-ID on all PCEP sessions it maintains at a given times. All subsequent PCEP messages then address the node/link/prefix by the LS-ID. The values of 0 and 0xFFFFFFFFFFFFFFFF are reserved.

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

TLVs that may be included in the LS Object are described in the following sections.

9.3.1. Routing Universe TLV

In case of remote link-state (and TE) population when existing IGP/BGP-LS are also used, OSPF and IS-IS may run multiple routing protocol instances over the same link as described in [RFC7752]. See [RFC6822] and [RFC6549] for more information. These instances define

independent "routing universes". The 64-Bit 'Identifier' field is used to identify the "routing universe" where the LS object belongs. The LS objects representing IGP objects (nodes or links or prefix) from the same routing universe MUST have the same 'Identifier' value; LS objects with different 'Identifier' values MUST be considered to be from different routing universes.

The format of the optional ROUTING-UNIVERSE 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=[TBD7]         |           Length=8            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Identifier                         |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Below table lists the 'Identifier' values that are defined as well-known in this draft (same as [RFC7752]).

Identifier	Routing Universe
0	Default Layer 3 Routing topology
1-31	Reserved

If this TLV is not present the default value 0 is assumed.

9.3.2. Route Distinguisher TLV

To allow identification of VPN link, node and prefix information in PCEP-LS, a Route Distinguisher (RD) [RFC4364] is used. The LS objects from the same VPN MUST have the same RD; LS objects with different RD values MUST be considered to be from different VPNs.

The format of the optional ROUTE-DISTINGUISHER 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=[TBD15]        |           Length=8            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Route Distinguisher                       |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The format of RD is as per [RFC4364].

9.3.3. Virtual Network TLV

To realize ACTN, the MDSC needs to build an multi-domain topology. This topology is best served, if this is an abstracted view of the underlying network resources of each domain. It is also important to provide a customer view of network slice for each customer. There is a need to control the level of abstraction based on the deployment scenario and business relationship between the controllers.

Virtual service coordination function in ACTN incorporates customer service-related knowledge into the virtual network operations in order to seamlessly operate virtual networks while meeting customer's service requirements. [I-D.ietf-teas-actn-requirements] describes various VN operations initiated by a customer/application. In this context, there is a need for associating the abstracted link state and TE topology with a VN "construct" to facilitate VN operations in PCE architecture.

VIRTUAL-NETWORK-TLV as per [I-D.leedhody-pce-vn-association] can be included in LS object to identify the link, node and prefix information belongs to a particular VN.

9.3.4. Local Node Descriptors TLV

As described in [RFC7752], each link is anchored by a pair of Router-IDs that are used by the underlying IGP, namely, 48 Bit ISO System-ID for IS-IS and 32 bit Router-ID for OSPFv2 and OSPFv3. Incase of additional auxiliary Router-IDs used for TE, these MUST also be included in the link attribute TLV (see Section 9.3.9.2).

It is desirable that the Router-ID assignments inside the Node Descriptor are globally unique. Some considerations for globally unique Node/Link/Prefix identifiers are described in [RFC7752].

The Local Node Descriptors TLV contains Node Descriptors for the node anchoring the local end of the link. This TLV MUST be included in the LS Report when during a given PCEP session a node/link/prefix is

first reported to a PCE. A PCC sends to a PCE the first LS Report either during State Synchronization, or when a new node/link/prefix is learned at the PCC. The value contains one or more Node Descriptor Sub-TLVs, which allows specification of a flexible key for any given node/link/prefix information such that global uniqueness of the node/link/prefix is ensured.

This TLV is applicable for all LS Object-Type.

```
 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=[TBD8]         |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//              Node Descriptor Sub-TLVs (variable)            //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The value contains one or more Node Descriptor Sub-TLVs defined in Section 9.3.6.

9.3.5. Remote Node Descriptors TLV

The Remote Node Descriptors contains Node Descriptors for the node anchoring the remote end of the link. This TLV MUST be included in the LS Report when during a given PCEP session a link is first reported to a PCE. A PCC sends to a PCE the first LS Report either during State Synchronization, or when a new link is learned at the PCC. The length of this TLV is variable. The value contains one or more Node Descriptor Sub-TLVs defined in Section 9.3.6.

This TLV is applicable for LS Link Object-Type.

```
 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=[TBD9]         |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//              Node Descriptor Sub-TLVs (variable)            //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

9.3.6. Node Descriptors Sub-TLVs

The Node Descriptor Sub-TLV type Type and lengths are listed in the following table:

Sub-TLV	Description	Length	Value defined in
0	Reserved	-	-
1	Autonomous System	4	[RFC7752]
2	BGP-LS Identifier	4	/ section
3	OSPF Area-ID	4	3.2.1.4
4	Router-ID	Variable	

The sub-TLV values in Node Descriptor TLVs are defined as follows (similar to [RFC7752]):

- o Autonomous System: opaque value (32 Bit AS Number)

- o BGP-LS Identifier: opaque value (32 Bit ID). In conjunction with ASN, uniquely identifies the BGP-LS domain as described in [RFC7752]. This sub-TLV is present only if the node implements BGP-LS and the ID is set by the operator.

- o OSPF Area ID: It is used to identify the 32 Bit area to which the LS object belongs. Area Identifier allows the different LS objects of the same node to be discriminated.

- o Router ID: opaque value. Usage is described in [RFC7752] as IGP Router ID. In case this is not learned from IGP, it SHOULD contain the unique router ID, such as TE router ID.

9.3.7. Link Descriptors TLV

The Link Descriptors TLV contains Link Descriptors for each link. This TLV MUST be included in the LS Report when during a given PCEP session a link is first reported to a PCE. A PCC sends to a PCE the first LS Report either during State Synchronization, or when a new link is learned at the PCC. The length of this TLV is variable. The value contains one or more Link Descriptor Sub-TLVs.

The 'Link descriptor' TLVs uniquely identify a link among multiple parallel links between a pair of anchor routers similar to [RFC7752].

This TLV is applicable for LS Link Object-Type.

```
 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=[TBD10]        |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//              Link Descriptor Sub-TLVs (variable)            //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Link Descriptor Sub-TLV type and lengths are listed in the following table:

Sub-TLV	Description	IS-IS TLV /Sub-TLV	Value defined in:
6	Link Local/Remote Identifiers	22/4	[RFC5307]/1.1
7	IPv4 interface address	22/6	[RFC5305]/3.2
8	IPv4 neighbor address	22/8	[RFC5305]/3.3
9	IPv6 interface address	22/12	[RFC6119]/4.2
10	IPv6 neighbor address	22/13	[RFC6119]/4.3
5	Multi-Topology identifier	-	[RFC7752]/ 3.2.1.5

The format and semantics of the 'value' fields in most 'Link Descriptor' sub-TLVs correspond to the format and semantics of value fields in IS-IS Extended IS Reachability sub-TLVs, defined in [RFC5305], [RFC5307] and [RFC6119]. Although the encodings for 'Link Descriptor' TLVs were originally defined for IS-IS, the TLVs can carry data sourced either by IS-IS or OSPF or direct.

The information about a link present in the LSA/LSP originated by the local node of the link determines the set of sub-TLVs in the Link Descriptor of the link as described in [RFC7752].

9.3.8. Prefix Descriptors TLV

The Prefix Descriptors TLV contains Prefix Descriptors uniquely identify an IPv4 or IPv6 Prefix originated by a Node. This TLV MUST be included in the LS Report when during a given PCEP session a prefix is first reported to a PCE. A PCC sends to a PCE the first LS

Report either during State Synchronization, or when a new prefix is learned at the PCC. The length of this TLV is variable.

This TLV is applicable for LS Prefix Object-Types for both IPv4 and IPv6.

```
 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=[TBD11]        |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//             Prefix Descriptor Sub-TLVs (variable)           //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The value contains one or more Prefix Descriptor Sub-TLVs defined below -

TLV Code Point	Description	Length	Value defined in:
5	Multi-Topology Identifier	variable	[RFC7752] /3.2.1.5
11	OSPF Route Type	1	[RFC7752] /3.2.3.1
12	IP Reachability Information	variable	[RFC7752] /3.2.3.2

9.3.9. PCEP-LS Attributes

9.3.9.1. Node Attributes TLV

This is an optional attribute that is used to carry node attributes. This TLV is applicable for LS Node Object-Type.

```
 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=[TBD12]        |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//              Node Attributes Sub-TLVs (variable)            //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Node Attributes Sub-TLV type and lengths are listed in the following table:

Sub TLV	Description	Length	Value defined in:
5	Multi-Topology Identifier	variable	[RFC7752] /3.2.1.5
13	Node Flag Bits	1	[RFC7752] /3.3.1.1
14	Opaque Node Properties	variable	[RFC7752] /3.3.1.5
15	Node Name	variable	[RFC7752] /3.3.1.3
16	IS-IS Area Identifier	variable	[RFC7752] /3.3.1.2
17	IPv4 Router-ID of Local Node	4	[RFC5305]/4.3
18	IPv6 Router-ID of Local Node	16	[RFC6119]/4.1

9.3.9.2. Link Attributes TLV

This TLV is applicable for LS Link Object-Type. The format and semantics of the 'value' fields in some 'Link Attribute' sub-TLVs correspond to the format and semantics of value fields in IS-IS Extended IS Reachability sub-TLVs, defined in [RFC5305], [RFC5307] and [RFC7752]. Although the encodings for 'Link Attribute' TLVs were originally defined for IS-IS, the TLVs can carry data sourced either by IS-IS or OSPF or direct.

```
 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=[TBD13]        |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//              Link Attributes Sub-TLVs (variable)            //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The following 'Link Attribute' sub-TLVs are valid :

Sub-TLV	Description	IS-IS TLV /Sub-TLV	Defined in:

		BGP-LS TLV	
17	IPv4 Router-ID of Local Node	134/---	[RFC5305]/4.3
18	IPv6 Router-ID of Local Node	140/---	[RFC6119]/4.1
19	IPv4 Router-ID of Remote Node	134/---	[RFC5305]/4.3
20	IPv6 Router-ID of Remote Node	140/---	[RFC6119]/4.1
6	Link Local/Remote Identifiers	22/4	[RFC5307]/1.1
22	Administrative group (color)	22/3	[RFC5305]/3.1
23	Maximum link bandwidth	22/9	[RFC5305]/3.3
24	Max. reservable link bandwidth	22/10	[RFC5305]/3.5
25	Unreserved bandwidth	22/11	[RFC5305]/3.6
26	TE Default Metric	22/18	[RFC7752] /3.3.2.3
27	Link Protection Type	22/20	[RFC5307]/1.2
28	MPLS Protocol Mask	1094	[RFC7752] /3.3.2.2
29	IGP Metric	1095	[RFC7752] /3.3.2.4
30	Shared Risk Link Group	1096	[RFC7752] /3.3.2.5
31	Opaque link attributes	1097	[RFC7752] /3.3.2.6
32	Link Name attribute	1098	[RFC7752] /3.3.2.7
33	Unidirectional Link Delay	22/33	[RFC7810]/4.1
34	Min/Max Unidirectional Link Delay	22/34	[RFC7810]/4.2
35	Unidirectional Delay Variation	22/35	[RFC7810]/4.3
36	Unidirectional Link Loss	22/36	[RFC7810]/4.4
37	Unidirectional Residual Bandwidth	22/37	[RFC7810]/4.5
38	Unidirectional Available Bandwidth	22/38	[RFC7810]/4.6
39	Unidirectional	22/39	[RFC7810]/4.7

	Bandwidth Utilization		
40	Extended Admin Group (EAG)	22/14	[RFC7308]/2.1

9.3.9.3. Prefix Attributes TLV

This TLV is applicable for LS Prefix Object-Types for both IPv4 and IPv6. Prefixes are learned from the IGP (IS-IS or OSPF) or BGP topology with a set of IGP attributes (such as metric, route tags, etc.). This section describes the different attributes related to the IPv4/IPv6 prefixes. Prefix Attributes TLVs SHOULD be encoded in the LS Prefix Object.

```
 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=[TBD14]        |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//            Prefix Attributes Sub-TLVs (variable)            //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The following 'Prefix Attribute' sub-TLVs are valid :

Sub-TLV	Description	BGP-LS TLV	Defined in:
41	IGP Flags	1152	[RFC7752] /3.3.3.1
42	Route Tag	1153	[RFC7752] /3.3.3.2
43	Extended Tag	1154	[RFC7752] /3.3.3.3
44	Prefix Metric	1155	[RFC7752] /3.3.3.4
45	OSPF Forwarding Address	1156	[RFC7752] /3.3.3.5
46	Opaque Prefix Attribute	1157	[RFC7752] /3.3.3.6

9.3.10. Removal of an Attribute

One of a key objective of PCEP-LS is to encode and carry only the impacted attributes of a Node, a Link or a Prefix. To accommodate this requirement, incase of a removal of an attribute, the sub-TLV MUST be included with no 'value' field and length=0 to indicate that the attribute is removed. On receiving a sub-TLV with zero length, the receiver removes the attribute from the database.

10. Other Considerations

10.1. Inter-AS Links

The main source of LS (and TE) information is the IGP, which is not active on inter-AS links. In some cases, the IGP may have information of inter-AS links ([RFC5392], [RFC5316]). In other cases, an implementation SHOULD provide a means to inject inter-AS links into PCEP. The exact mechanism used to provision the inter-AS links is outside the scope of this document.

11. Security Considerations

This document extends PCEP for LS (and TE) distribution including a new LSRpt message with new object and TLVs. Procedures and protocol extensions defined in this document do not effect the overall PCEP security model. See [RFC5440], [I-D.ietf-pce-pceps]. Tampering with the LSRpt message may have an effect on path computations at PCE. It also provides adversaries an opportunity to eavesdrop and learn sensitive information and plan sophisticated attacks on the network infrastructure. The PCE implementation SHOULD provide mechanisms to prevent strains created by network flaps and amount of LS (and TE) information. Thus it is suggested that any mechanism used for securing the transmission of other PCEP message be applied here as well. As a general precaution, it is RECOMMENDED that these PCEP extensions only be activated on authenticated and encrypted sessions belonging to the same administrative authority.

12. Manageability Considerations

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

12.1. Control of Function and Policy

A PCE or PCC implementation MUST allow configuring the PCEP-LS capabilities as described in this document.

A PCC implementation SHOULD allow configuration to suggest if remote information learned via routing protocols should be reported or not.

An implementation SHOULD allow the operator to specify the maximum number of LS data to be reported.

An implementation SHOULD also allow the operator to create abstracted topologies that are reported to the peers and create different abstractions for different peers.

An implementation SHOULD allow the operator to configure a 64-bit Instance-ID for Routing Universe TLV.

12.2. Information and Data Models

An implementation SHOULD allow the operator to view the LS capabilities advertised by each peer. To serve this purpose, the PCEP YANG module [I-D.ietf-pce-pcep-yang]" can be extended to include advertised capabilities.

An implementation SHOULD also provide the statistics:

- o Total number of LSRpt sent/received, as well as per neighbor

- o Number of error received for LSRpt, per neighbor

- o Total number of locally originated Link-State Information

These statistics should be recorded as absolute counts since system or session start time. An implementation MAY also enhance this information by recording peak per-second counts in each case.

An operator SHOULD define an import policy to limit inbound LSRpt to "drop all LSRpt from a particular peers" as well provide means to limit inbound LSRpts.

12.3. 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]".

12.4. Verify Correct Operations

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

12.5. Requirements On Other Protocols

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

12.6. 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].

13. IANA Considerations

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

13.1. PCEP Messages

IANA created a registry for PCEP messages. Each PCEP message has a message type value. This document defines a new PCEP message value.

Value	Meaning	Reference
TBD3	LSRpt	[This I-D]

13.2. PCEP Objects

This document defines the following new PCEP Object-classes and Object-values:

Object-Class Value	Name	Reference
TBD6	LS Object Object-Type=1 (LS Node) Object-Type=2 (LS Link) Object-Type=3 (LS IPv4 Prefix) Object-Type=4 (LS IPv6 Prefix)	[This I-D]

13.3. LS Object

This document requests that a new sub-registry, named "LS Object Protocol-ID Field", is created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the Flag field of the LSP object. New values are to be assigned by Standards Action [RFC5226].

Value	Meaning	Reference
0	Reserved	[This I-D]
1	IS-IS Level 1	[This I-D]
2	IS-IS Level 2	[This I-D]
3	OSPFv2	[This I-D]
4	Direct	[This I-D]
5	Static configuration	[This I-D]
6	OSPFv3	[This I-D]
7	BGP-LS	[This I-D]
8	PCEP-LS	[This I-D]
9	Abstraction	[This I-D]
10	Unspecified	[This I-D]

Further, this document also requests that a new sub-registry, named "LS Object Flag Field", is created within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the Flag field of the LSP object.New values are to be assigned by Standards Action [RFC5226]. Each bit should be tracked with the following qualities:

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

The following values are defined in this document:

Bit	Description	Reference
0-21	Unassigned	
22	R (Remove bit)	[This I-D]
23	S (Sync bit)	[This I-D]

13.4. 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 Error-Value=TBD4 (LS object missing)	[RFC5440] [This I-D]
19	Invalid Operation Error-Value=TBD1 (Attempted LS Report if LS remote capability was not advertised)	[I-D.ietf-pce-stateful-pce] [This I-D]
TBD2	LS Synchronization Error Error-Value=1 (An error in processing the LSRpt) Error-Value=2 (An internal PCC error)	[This I-D]

13.5. PCEP TLV Type Indicators

This document defines the following new PCEP TLVs.

Value	Meaning	Reference
TBD5	LS-CAPABILITY TLV	[This I-D]
TBD7	ROUTING-UNIVERSE TLV	[This I-D]
TBD15	ROUTE-DISTINGUISHER TLV	[This I-D]
TBD8	Local Node Descriptors TLV	[This I-D]
TBD9	Remote Node Descriptors TLV	[This I-D]
TBD10	Link Descriptors TLV	[This I-D]
TBD11	Prefix Descriptors TLV	[This I-D]
TBD12	Node Attributes TLV	[This I-D]
TBD13	Link Attributes TLV	[This I-D]
TBD14	Prefix Attributes TLV	[This I-D]

13.6. PCEP-LS Sub-TLV Type Indicators

This document specifies the PCEP-LS Sub-TLVs. IANA is requested to create an "PCEP-LS Sub-TLV Types" sub-registry in the "PCEP TLV Type Indicators" for the sub-TLVs carried in the PCEP-LS TLV (Local and Remote Node Descriptors TLV, Link Descriptors TLV, Prefix Descriptors TLV, Node Attributes TLV, Link Attributes TLV and Prefix Attributes TLV. This document defines the following types:

Sub-TLV	Description	Ref	Value defined

		Sub-TLV	in:
1	Autonomous System	512	[RFC7752] /3.2.1.4
2	BGP-LS Identifier	513	[RFC7752] /3.2.1.4
3	OSPF Area-ID	514	[RFC7752] /3.2.1.4
4	Router-ID	515	[RFC7752] /3.2.1.4
5	Multi-Topology-ID	263	[RFC7752] /3.2.1.5
6	Link Local/Remote Identifiers	22/4	[RFC5307]/1.1
7	IPv4 interface address	22/6	[RFC5305]/3.2
8	IPv4 neighbor address	22/8	[RFC5305]/3.3
9	IPv6 interface address	22/12	[RFC6119]/4.2
10	IPv6 neighbor address	22/13	[RFC6119]/4.3
11	OSPF Route Type	264	[RFC7752] /3.2.3.1
12	IP Reachability Information	265	[RFC7752] /3.2.3.2
13	Node Flag Bits	1024	[RFC7752] /3.3.1.1
14	Opaque Node Properties	1025	[RFC7752] /3.3.1.5
15	Node Name	1026	[RFC7752] /3.3.1.3
16	IS-IS Area Identifier	1027	[RFC7752] /3.3.1.2
17	IPv4 Router-ID of Local Node	134/--	[RFC5305]/4.3
18	IPv6 Router-ID of Local Node	140/--	[RFC6119]/4.1
19	IPv4 Router-ID of Remote Node	134/--	[RFC5305]/4.3
20	IPv6 Router-ID of Remote Node	140/--	[RFC6119]/4.1
22	Administrative group (color)	22/3	[RFC5305]/3.1
23	Maximum link bandwidth	22/9	[RFC5305]/3.3
24	Max. reservable link bandwidth	22/10	[RFC5305]/3.5

25	Unreserved bandwidth	22/11	[RFC5305]/3.6
26	TE Default Metric	22/18	[RFC7752] /3.3.2.3
27	Link Protection Type	22/20	[RFC5307]/1.2
28	MPLS Protocol Mask	1094	[RFC7752] /3.3.2.2
29	IGP Metric	1095	[RFC7752] /3.3.2.4
30	Shared Risk Link Group	1096	[RFC7752] /3.3.2.5
31	Opaque link attributes	1097	[RFC7752] /3.3.2.6
32	Link Name attribute	1098	[RFC7752] /3.3.2.7
33	Unidirectional Link Delay	22/33	[RFC7810]/4.1
34	Min/Max Unidirectional Link Delay	22/34	[RFC7810]/4.2
35	Unidirectional Delay Variation	22/35	[RFC7810]/4.3
36	Unidirectional Link Loss	22/36	[RFC7810]/4.4
37	Unidirectional Residual Bandwidth	22/37	[RFC7810]/4.5
38	Unidirectional Available Bandwidth	22/38	[RFC7810]/4.6
39	Unidirectional Bandwidth Utilization	22/39	[RFC7810]/4.7
40	Extended Admin Group (EAG)	22/14	[RFC7308]/2.1
41	IGP Flags	1152	[RFC7752] /3.3.3.1
42	Route Tag	1153	[RFC7752] /3.3.3.2
43	Extended Tag	1154	[RFC7752] /3.3.3.3
44	Prefix Metric	1155	[RFC7752] /3.3.3.4
45	OSPF Forwarding Address	1156	[RFC7752] /3.3.3.5
46	Opaque Prefix Attribute	1157	[RFC7752] /3.3.3.6

New values are to be assigned by Standards Action [RFC5226].

14. TLV/Sub-TLV Code Points Summary

This section contains the global table of all TLVs/Sub-TLVs in LS object defined in this document.

TLV	Description	Ref TLV	Value defined in:
TBD7	Routing Universe	--	Sec 9.2.1
TBD15	Route Distinguisher	--	Sec 9.2.2
*	Virtual Network	--	[leedhody-pce-vn-association]
TBD8	Local Node Descriptors	256	[RFC7752] /3.2.1.2
TBD9	Remote Node Descriptors	257	[RFC7752] /3.2.1.3
TBD10	Link Descriptors	--	Sec 9.2.8
TBD11	Prefix Descriptors	--	Sec 9.2.9
TBD12	Node Attributes	--	Sec 9.2.10.1
TBD13	Link Attributes	--	Sec 9.2.10.2
TBD14	Prefix Attributes	--	Sec 9.2.10.3

* this TLV is defined in a different PCEP document

TLV Table

Refer Section 13.6 for the table of Sub-TLVs.

15. Implementation Status

The PCEP-LS protocol extension as described in this I-D were implemented and tested for a variety of applications. Apart from the below implementation, there exist other experimental implementations done for optical networks.

15.1. Hierarchical Transport PCE controllers

The PCEP-LS has been implemented as part of IETF97 Hackathon and Bits-N-Bites demonstration. The use-case demonstrated was DCI use-case of ACTN architecture in which to show the following scenarios:

- connectivity services on the ACTN based recursive hierarchical SDN/PCE platform that has the three tier level SDN controllers

(two-tier level MDSC and PNC) on the top of the PTN systems managed by EMS.

- Integration test of two tier-level MDSC: The SBI of the low level MDSC is the YANG based Korean national standards and the one of the high level MDSC the PCEP-LS based ACTN protocols.

- Performance test of three types of SDN controller based recovery schemes including protection, reactive and proactive restoration. PCEP-LS protocol was used to demonstrate quick report of failed network components.

15.2. ONOS-based Controller (MDSC and PNC)

Huawei (PNC, MDSC) and SKT (MDSC) implemented PCEP-LS during Hackathon and IETF97 Bits-N-Bites demonstration. The demonstration was ONOS-based ACTN architecture in which to show the following capabilities:

Both packet PNC and optical PNC (with optical PCEP-LS extension) implemented PCEP-LS on its SBI and well as its NBI (towards MDSC).

SKT orchestrator (acting as MDSC) also supported PCEP-LS (as well as RestConf) towards packet and optical PNCs on its SBI.

Further description can be found at <ONOS-PCEP> and the code at <ONOS-PCEP-GITHUB>.

16. Acknowledgments

This document borrows some of the structure and text from the [RFC7752].

Thanks to Eric Wu, Venugopal Kondreddy, Mahendra Singh Negi, Avantika, and Zhengbin Li for the reviews.

Thanks to Ramon Casellas for his comments and suggestions based on his implementation experience.

17. References

17.1. Normative References

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

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <http://www.rfc-editor.org/info/rfc5305>.

[RFC5307] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 5307, DOI 10.17487/RFC5307, October 2008, <http://www.rfc-editor.org/info/rfc5307>.

[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>.

[RFC6119] Harrison, J., Berger, J., and M. Bartlett, "IPv6 Traffic Engineering in IS-IS", RFC 6119, DOI 10.17487/RFC6119, February 2011, <http://www.rfc-editor.org/info/rfc6119>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <http://www.rfc-editor.org/info/rfc7752>.

[RFC7810] Previdi, S., Ed., Giacalone, S., Ward, D., Drake, J., and Q. Wu, "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 7810, DOI 10.17487/RFC7810, May 2016, <http://www.rfc-editor.org/info/rfc7810>.

17.2. Informative References

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <http://www.rfc-editor.org/info/rfc3630>.

[RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <http://www.rfc-editor.org/info/rfc4203>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <http://www.rfc-editor.org/info/rfc4364>.

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

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <http://www.rfc-editor.org/info/rfc5120>.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 5226, DOI 10.17487/RFC5226, May 2008, <http://www.rfc-editor.org/info/rfc5226>.

[RFC5316] Chen, M., Zhang, R., and X. Duan, "ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5316, DOI 10.17487/RFC5316, December 2008, <http://www.rfc-editor.org/info/rfc5316>.

[RFC5392] Chen, M., Zhang, R., and X. Duan, "OSPF Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5392, DOI 10.17487/RFC5392, January 2009, <http://www.rfc-editor.org/info/rfc5392>.

[RFC6549] Lindem, A., Roy, A., and S. Mirtorabi, "OSPFv2 Multi-Instance Extensions", RFC 6549, DOI 10.17487/RFC6549, March 2012, <http://www.rfc-editor.org/info/rfc6549>.

[RFC6805] King, D., Ed. and A. Farrel, Ed., "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS", RFC 6805, DOI 10.17487/RFC6805, November 2012, <http://www.rfc-editor.org/info/rfc6805>.

[RFC6822] Previdi, S., Ed., Ginsberg, L., Shand, M., Roy, A., and D. Ward, "IS-IS Multi-Instance", RFC 6822, DOI 10.17487/RFC6822, December 2012, <http://www.rfc-editor.org/info/rfc6822>.

[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-21 (work in progress), June 2017.

[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-10 (work in progress), June 2017.

[I-D.ietf-pce-pceps]
Lopez, D., Dios, O., Wu, Q., and D. Dhody, "Secure Transport for PCEP", draft-ietf-pce-pceps-14 (work in progress), May 2017.

[I-D.ietf-pce-pcep-yang]
Dhody, D., Hardwick, J., Beeram, V., and j. jefftant@gmail.com, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", draft-ietf-pce-pcep-yang-02 (work in progress), March 2017.

[I-D.ietf-pce-applicability-actn]
Dhody, D., Lee, Y., and D. Ceccarelli, "Applicability of Path Computation Element (PCE) for Abstraction and Control of TE Networks (ACTN)", draft-ietf-pce-applicability-actn-00 (work in progress), June 2017.

[I-D.ietf-teas-actn-framework]
Ceccarelli, D. and Y. Lee, "Framework for Abstraction and Control of Traffic Engineered Networks", draft-ietf-teas-actn-framework-06 (work in progress), June 2017.

[I-D.ietf-teas-actn-requirements]
Lee, Y., Dhody, D., Belotti, S., Pithewan, K., Ceccarelli, D., Miyasaka, T., and J. Shin, "Requirements for Abstraction and Control of TE Networks", draft-ietf-teas-actn-requirements-05 (work in progress), May 2017.

[I-D.kondreddy-pce-pcep-ls-sync-optimizations]
Kondreddy, V. and M. Negi, "Optimizations of PCEP Link-State(LS) Synchronization Procedures", draft-kondreddy-pce-pcep-ls-sync-optimizations-00 (work in progress), October 2015.

[I-D.leedhody-pce-vn-association]
Lee, Y., Dhody, D., Zhang, X., and D. Ceccarelli, "PCEP Extensions for Establishing Relationships Between Sets of LSPs and Virtual Networks", draft-leedhody-pce-vn-association-02 (work in progress), March 2017.

[ONOS-PCEP]
"Support for PCEP in ONOS",
<https://wiki.onosproject.org/display/ONOS/PCEP+Protocol>.

[ONOS-PCEP-GITHUB]
"Github for PCEP code in ONOS",
<https://github.com/opennetworkinglab/onos/tree/master/protocols/pcep>.

Appendix A. Relevant OSPF TLV and sub-TLV

This section list the relevant TLVs and sub-TLVs defined for OSPF.

Sub-TLV	Description	OSPF-TE Sub-TLV	Value defined in:
6	Link Local/Remote Identifiers	11	[RFC4203]/1.1
7	IPv4 interface address	3	[RFC3630]/2.5.3
8	IPv4 neighbor address	4	[RFC3630]/2.5.4
9	IPv6 interface address	19	[RFC5329]/4.3
10	IPv6 neighbor address	20	[RFC5329]/4.4
17	IPv4 Router-ID of Local Node	1	[RFC3630]/2.4.1
18	IPv6 Router-ID of Local Node	3	[RFC5329]/3
19	IPv4 Router-ID of Remote Node	1	[RFC3630]/2.4.1
20	IPv6 Router-ID of Remote Node	3	[RFC5329]/3
22	Administrative group (color)	9	[RFC3630]/2.5.9
23	Maximum link bandwidth	6	[RFC3630]/2.5.6
24	Max. reservable link bandwidth	7	[RFC3630]/2.5.7
25	Unreserved bandwidth	8	[RFC3630]/2.5.8
27	Link Protection Type	14	[RFC4203]/1.2
30	Shared Risk Link Group	16	[RFC4203]/1.3
33	Unidirectional Link Delay	27	[RFC7471]/4.1
34	Min/Max Unidirectional Link Delay	28	[RFC7471]/4.2
35	Unidirectional Delay Variation	29	[RFC7471]/4.3
36	Unidirectional Link Loss	30	[RFC7471]/4.4
37	Unidirectional	31	[RFC7471]/4.5

```
|           | Residual Bandwidth  |               |                 |
|    38     | Unidirectional      |      32       | [RFC7471]/4.6   |
|           | Available Bandwidth |               |                 |
|    39     | Unidirectional      |      33       | [RFC7471]/4.7   |
|           | Bandwidth           |               |                 |
|           | Utilization         |               |                 |
|    40     | Extended Admin      |      26       | [RFC7308]/2.1   |
|           | Group (EAG)         |               |                 |
+-----------+---------------------+---------------+-----------------+
```

Appendix B. Examples

These examples are for illustration purposes only to show how the new PCEP-LS message could be encoded. They are not meant to be an exhaustive list of all possible use cases and combinations.

B.1. All Nodes

Each node (PCC) in the network chooses to provide its own local node and link information, and in this way PCE can build the full link state and TE information.

```
+--------------------+                   +--------------------+
|                    |                   |                    |
|     RTA            |10.1.1.1           |     RTB            |
|     1.1.1.1        |-------------------|     2.2.2.2        |
|     Area 0         |           10.1.1.2|     Area 0         |
|                    |                   |                    |
+--------------------+                   +--------------------+
RTA
---
LS Node
   TLV - Local Node Descriptors
       Sub-TLV - 3: OSPF Area-ID: 0.0.0.0
       Sub-TLV - 4: Router-ID: 1.1.1.1
   TLV - Node Attributes TLV
       Sub-TLV(s)


LS Link
   TLV - Local Node Descriptors
       Sub-TLV - 3: OSPF Area-ID: 0.0.0.0
       Sub-TLV - 4: Router-ID: 1.1.1.1
   TLV - Remote Node Descriptors
       Sub-TLV - 3: OSPF Area-ID: 0.0.0.0
       Sub-TLV - 4: Router-ID: 2.2.2.2
   TLV - Link Descriptors
       Sub-TLV - 7: IPv4 interface: 10.1.1.1
```

```
            Sub-TLV - 8: IPv4 neighbor: 10.1.1.2
         TLV - Link Attributes TLV
            Sub-TLV(s)

   RTB
   ---
   LS Node
      TLV - Local Node Descriptors
            Sub-TLV - 3: OSPF Area-ID: 0.0.0.0
            Sub-TLV - 4: Router-ID: 2.2.2.2
      TLV - Node Attributes TLV
            Sub-TLV(s)

   LS Link
      TLV - Local Node Descriptors
            Sub-TLV - 3: OSPF Area-ID: 0.0.0.0
            Sub-TLV - 4: Router-ID: 2.2.2.2
      TLV - Remote Node Descriptors
            Sub-TLV - 3: OSPF Area-ID: 0.0.0.0
            Sub-TLV - 4: Router-ID: 1.1.1.1
      TLV - Link Descriptors
            Sub-TLV - 7: IPv4 interface: 10.1.1.2
            Sub-TLV - 8: IPv4 neighbor: 10.1.1.1
      TLV - Link Attributes TLV
            Sub-TLV(s)
```

B.2. Designated Node

A designated node(s) in the network will provide its own local node as well as all learned remote information, and in this way PCE can build the full link state and TE information.

As described in Appendix B.1, the same LS Node and Link objects will be generated with a difference that it would be a designated router say RTA that generate all this information.

B.3. Between PCEs

As per Hierarchical-PCE [RFC6805], Parent PCE builds an abstract domain topology map with each domain as an abstract node and inter-domain links as an abstract link. Each child PCE may provide this information to the parent PCE. Considering the example in figure 1 of [RFC6805], following LS object will be generated:

```
PCE1
----
LS Node
   TLV - Local Node Descriptors
       Sub-TLV - 1: Autonomous System: 100 (Domain 1)
       Sub-TLV - 4: Router-ID: 11.11.11.11 (abstract)

LS Link
   TLV - Local Node Descriptors
       Sub-TLV - 1: Autonomous System: 100
       Sub-TLV - 4: Router-ID: 11.11.11.11 (abstract)
   TLV - Remote Node Descriptors
       Sub-TLV - 1: Autonomous System: 200 (Domain 2)
       Sub-TLV - 4: Router-ID: 22.22.22.22 (abstract)
   TLV - Link Descriptors
       Sub-TLV - 7: IPv4 interface: 11.1.1.1
       Sub-TLV - 8: IPv4 neighbor: 11.1.1.2
   TLV - Link Attributes TLV
       Sub-TLV(s)

LS Link
   TLV - Local Node Descriptors
       Sub-TLV - 1: Autonomous System: 100
       Sub-TLV - 4: Router-ID: 11.11.11.11 (abstract)
   TLV - Remote Node Descriptors
       Sub-TLV - 1: Autonomous System: 200
       Sub-TLV - 4: Router-ID: 22.22.22.22 (abstract)
   TLV - Link Descriptors
       Sub-TLV - 7: IPv4 interface: 12.1.1.1
       Sub-TLV - 8: IPv4 neighbor: 12.1.1.2
   TLV - Link Attributes TLV
       Sub-TLV(s)

LS Link
   TLV - Local Node Descriptors
       Sub-TLV - 1: Autonomous System: 100
       Sub-TLV - 4: Router-ID: 11.11.11.11 (abstract)
   TLV - Remote Node Descriptors
       Sub-TLV - 1: Autonomous System: 400 (Domain 4)
       Sub-TLV - 4: Router-ID: 44.44.44.44 (abstract)
   TLV - Link Descriptors
       Sub-TLV - 7: IPv4 interface: 13.1.1.1
       Sub-TLV - 8: IPv4 neighbor: 13.1.1.2
   TLV - Link Attributes TLV
       Sub-TLV(s)

* similar information will be generated by other PCE
  to help form the abstract domain topology.
```

Further the exact border nodes and abstract internal path between the border nodes may also be transported to the Parent PCE to enable ACTN as described in [I-D.ietf-pce-applicability-actn] using the similar LS node and link objects encodings.

Appendix C. Contributor Addresses

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

EMail: udayasreereddy@gmail.com

Sergio Belotti
Alcatel-Lucent
Italy

EMail: sergio.belotti@alcatel-lucent.com

Veerendranatha Reddy Vallem
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

Email: veerendranatharv@huawei.com

Satish Karunanithi
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

Email: satishk@huawei.com

Authors' Addresses

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

EMail: dhruv.ietf@gmail.com

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

EMail: leeyoung@huawei.com

Daniele Ceccarelli
Ericsson
Torshamnsgatan,48
Stockholm
Sweden

EMail: daniele.ceccarelli@ericsson.com