

PCE Working Group
Internet-Draft
Intended status: Experimental
Expires: 23 October 2025

D. Dhody
S. Peng
Huawei
Y. Lee
Samsung Electronics
D. Ceccarelli
Cisco
A. Wang
China Telecom
G. Mishra
Verizon Inc.
21 April 2025

PCEP extensions for Distribution of Link-State and TE Information
draft-ietf-pce-pcep-ls-03

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 the traffic engineering extensions.

This document extends the Path Computation Element Communication Protocol (PCEP) with Link-State and TE Information as an experimental extension to allow gathering more deployment and implementation feedback on the use of PCEP in this way.

Status of This Memo

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

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

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

This Internet-Draft will expire on 23 October 2025.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	3
1.1. Scope of the Experiment	5
1.1.1. Evaluation of the Experiment	6
1.2. Requirements Language	7
2. Terminology	7
3. Applicability	7
4. Requirements for PCEP Extensions	8
5. New Functions to distribute link-state (and TE) via PCEP . .	9
6. Overview of Extensions to PCEP	9
6.1. New Messages	9
6.2. Capability Advertisement	10
6.3. Initial Link-State (and TE) Synchronization	10
6.3.1. Optimizations for LS Synchronization	12
6.4. LS Report	13
7. Transport	13
8. PCEP Messages	13
8.1. LS Report Message	13
8.2. The PCErr Message	14
9. Objects and TLV	14
9.1. TLV Format	15
9.2. Open Object	15
9.2.1. LS Capability TLV	15
9.3. LS Object	16
9.3.1. Routing Universe TLV	18
9.3.2. Route Distinguisher TLV	18
9.3.3. Virtual Network TLV	19
9.3.4. Local Node Descriptors TLV	19
9.3.5. Remote Node Descriptors TLV	20
9.3.6. Node Descriptors Sub-TLVs	20
9.3.7. Link Descriptors TLV	21
9.3.8. Prefix Descriptors TLV	22
9.3.9. PCEP-LS Attributes	22

9.3.9.1. Node Attributes TLV	22
9.3.9.2. Link Attributes TLV	23
9.3.9.3. Prefix Attributes TLV	23
9.3.10. Removal of an Attribute	24
10. Other Considerations	24
10.1. Inter-AS Links	24
11. Security Considerations	24
12. Manageability Considerations	25
12.1. Control of Function and Policy	25
12.2. Information and Data Models	25
12.3. Liveness Detection and Monitoring	26
12.4. Verify Correct Operations	26
12.5. Requirements On Other Protocols	26
12.6. Impact On Network Operations	26
13. IANA Considerations	26
13.1. PCEP Messages	26
13.2. PCEP Objects	27
13.3. LS Object	27
13.4. PCEP-Error Object	28
13.5. PCEP TLV Type Indicators	28
13.6. PCEP-LS Sub-TLV Type Indicators	29
14. TLV Code Points Summary	29
15. Implementation Status	30
15.1. Hierarchical Transport PCE controllers	31
15.2. ONOS-based Controller (MDSC and PNC)	31
16. Acknowledgments	32
17. References	32
17.1. Normative References	32
17.2. Informative References	33
Appendix A. Examples	36
A.1. All Nodes	36
A.2. Designated Node	37
A.3. Between PCEs	38
Appendix B. Contributor Addresses	39
Authors' Addresses	40

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 the relevant information that a Path Computation Element (PCE) needs to perform its computations. The TED must be 'complete and accurate' each time the PCE performs a path computation.

In MPLS and GMPLS, Interior Gateway 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).

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 mechanism is offered by BGP-LS [I-D.ietf-idr-rfc7752bis] .

[RFC8231] 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 IGP, but also the set of active paths and their reserved resources for its computations. Path Computation Client (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/TED, 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 session with each network node and with an incremental update from each network node with only the attributes that get modified.

[RFC8281] 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 updates at the PCE.

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

This document describes how 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. This document is designated as "experimental" to allow for the collection of more deployment and implementation information regarding the use of PCEP for learning link state and traffic engineering (TE) topology information before attempting to standardize it (see Section 1.1).

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 [RFC8637], the Child PCE can also transport abstract Link-State and TE information from a Child PCE to a Parent PCE using the mechanism described in this document to build an abstract topology at the Parent PCE.

[RFC8231] describes LSP state synchronization between PCCs and PCEs in the case of stateful PCE. This document does not make any changes to the LSP state synchronization process. The mechanism described in this document is on top of the existing LSP state synchronization.

1.1. Scope of the Experiment

The procedures described in this document are experimental. The experiment is intended to enable research on the usage of PCEP to populate the Link-State and TE Information from a PCC to the PCE. For this purpose, this document specifies a new PCEP message, a new PCEP object, and new TLVs.

The new message introduced by this document will not be understood by legacy implementations. On receiving the message, a legacy implementation will behave according to the rules for an unknown message as per [RFC5440]. It is assumed that this experiment will be conducted only when both the PCE and PCC form part of the experiment. It is possible that a PCC or PCE can operate with peers, some of which form part of the experiment and some that do not. In this case, the capability exchange required before using this extension would take care of the mismatch. A PCEP speaker that offers this feature to its peer that does not support or does not wish to support the feature will not receive the indication of support in the Open message, and so is expected not to use the feature. Thus, this experimentation would not clash with or cause harm to existing deployments. Further, since a PCEP speaker would use the new message only after capability exchange, there is no danger of this experimentation "escaping" to the wider Internet. A PCEP speaker

that receives the new message (that is part of the feature when the use of the feature has not been agreed upon) will send an error message as described in Section 6.9 of [RFC5440]. A PCEP speaker that receives the new object that is part of the feature when the use of the feature has not been agreed, it will send an error message as described in Section 7.2 of [RFC5440].

1.1.1. Evaluation of the Experiment

For every implementation of this feature, the information as per [RFC7942] is collected and maintained in the Working Group (WG) wiki [WIKI].

For every deployment of this feature, the following information is collected and maintained in the wiki[WIKI]:

- * Network Type
- * Network Underlying Technology
- * Network Size (nodes and links)
- * Number of PCEP sessions with the feature enabled
- * Number of paths in Stateful PCE
- * The main reason for using this feature via PCEP over others
- * Any challenges (deployment, management, scalability...)

The information can also be shared directly with the WG chairs who can anonymise the deployment information if needed.

The experiment will end three years after the RFC is published. At that point, the WG will determine how widely this has been implemented and deployed. The WG will evaluate the success of this feature by comparing its applicability in various deployments in comparison to existing mechanisms. Key considerations include whether the feature adversely affects the performance of the PCE and if it is scalable. The Working Group will also monitor for any issues related to the implementation, deployment, and management of the feature.

When the results of implementation and deployment are available, this document (or part thereof) will be updated and refined, and then it could be moved from Experimental to Standards Track.

1.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Terminology

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

3. Applicability

The mechanism specified in this draft is applicable to deployments:

- * Where there is no IGP or BGP-LS running in the network.
- * Where there is no IGP or BGP-LS running at the PCE to learn link-state and TE information.
- * Where there is IGP or BGP-LS running, but with a need for a faster and direct 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 full (entire) information of the node/link).
 - A PCE may receive full information from both existing mechanisms (IGP or BGP-LS) and PCEP.
- * 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 Provisioning Network Controller (PNC) to Multi-Domain Service Coordinator (MDSC) in Abstraction and Control of TE Networks (ACTN) [RFC8453].
- * Where there is an existing PCEP session between all the nodes and the PCE-based central controller (PCECC) [RFC8283], and the operator would like to use PCEP as a direct southbound interface to all the nodes in the network. This enables the operator to use PCEP as a single direct protocol between the controller and all the nodes in the network. In this mode, all nodes send only the local information.

Based on the local policy and deployment scenario, a PCC chooses 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 the scope of this document.

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

[RFC8453] defines three types of topology abstraction - (1) Native/White Topology; (2) Black Topology; and (3) Grey Topology. Based on the local policy, the PNC (or Child PCE) would share the domain topology with the MDSC (or Parent PCE) based on the abstraction type. The protocol extensions defined in this document can carry any type of topology abstraction.

4. Requirements for PCEP Extensions

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

1. The PCEP speaker supporting this draft MUST have 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 MUST include self-originated (local) information and MAY also allow remote information learned via routing protocols. PCC MUST be capable of doing the initial bulk sync at the time of session initialization as well as any changes thereafter.
3. A PCE MAY learn link-state (and TE) from PCEP as well as from existing mechanisms like IGP/BGP-LS. PCEP extensions MUST have a mechanism to correlate 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 need not know about any OSPF or IS-IS or BGP-LS protocol specifics.
4. It SHOULD be possible to encode only the changes in link-state (and TE) properties (after the initial sync) in the PCEP messages. This leads to faster convergence.

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:

- * Capability advertisement (E-C,C-E): both the PCC and the PCE MUST announce during the PCEP session establishment that they support PCEP extensions for the distribution of link-state (and TE) information defined in this document.
- * 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 the case of stateful PCE, it is RECOMMENDED that this operation be done before LSP state synchronization.
- * Link-State (and TE) Report (C-E): a PCC sends an LS (and TE) report to a PCE whenever the Link-State and TE information changes.

6. Overview of Extensions to PCEP

6.1. New Messages

In this document, we define a new PCEP message 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 an LSRpt message can contain the node or link properties. A unique PCEP-specific LS identifier (LS-ID) is also carried in the message to identify a node or a link 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. In case 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 the 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 with local link-state (and TE) information. The presence of the LS Capability TLV in PCE's Open message indicates that the PCE is interested in receiving LS Reports with local link-state (and TE) information.

The PCEP 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 supports the extensions defined in this document but does not advertise this capability, then upon receipt of an 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) database in a PCE. State Synchronization is performed immediately after the Initialization phase (see [RFC5440]). In the case of stateful PCE ([RFC8231]) it is RECOMMENDED 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 an 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 the state it received from this PCC. The session re-establishment MUST be re-attempted per the procedures defined in [RFC5440], including the 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 the LS capability TLV.

The successful LS Synchronization sequence is shown in Figure 1.

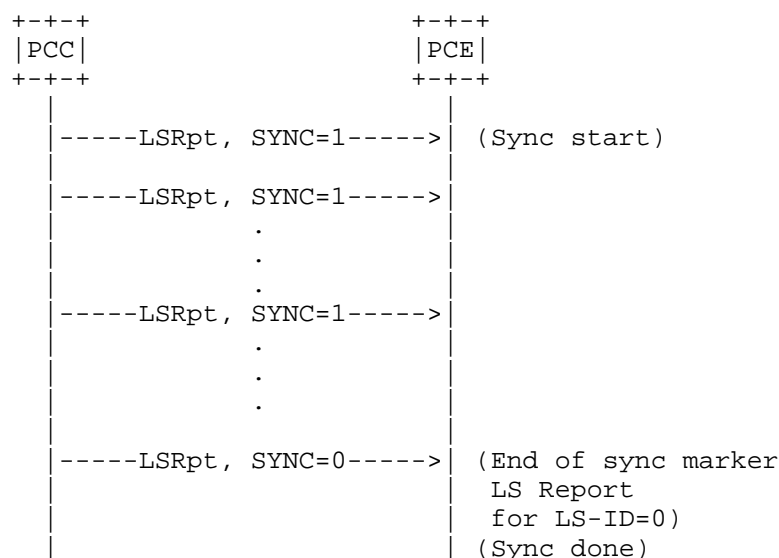


Figure 1: Successful LS synchronization

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

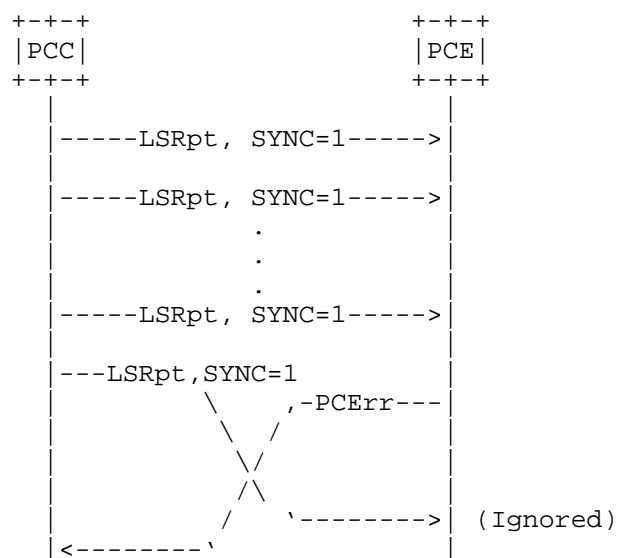


Figure 2: Failed LS synchronization (PCE failure)

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

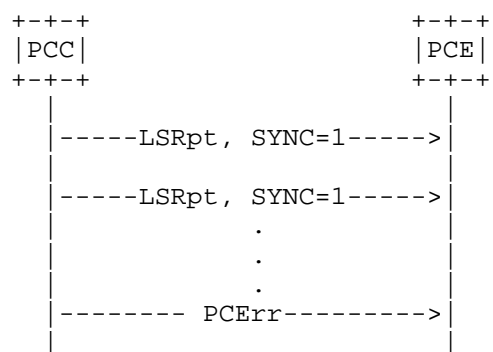


Figure 3: Failed LS synchronization (PCC failure)

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 an LS Report carried on an 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. It MAY also include the mapping of IGP or BGP-LS identifier to map the information populated via PCEP with IGP/BGP-LS identifiers.

More details about the LSRpt message are in Section 8.1.

7. Transport

A permanent PCEP session (section 4.2.8 of [RFC5440]) 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 is re-attempted as 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 specifies 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. An LSRpt message can carry more than one LS Reports (LS object). 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 that carries LS information of a node/prefix or a link. Each LS object has a 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 of a single PCC can populate. If an 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:

```
<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

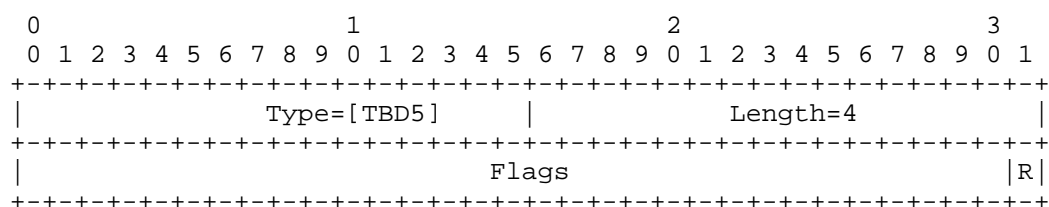
The TLV and the sub-TLV format (and padding) in this document are as per section 7.1 of [RFC5440].

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:



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):

- * R (remote allowed - 1 bit): if set to 1 by a PCC, the R Flag indicates that the PCC allows the 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 PCC and PCE for LSRpt messages to report remote as well as local LS information on a PCEP session. The TLVs related to the 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

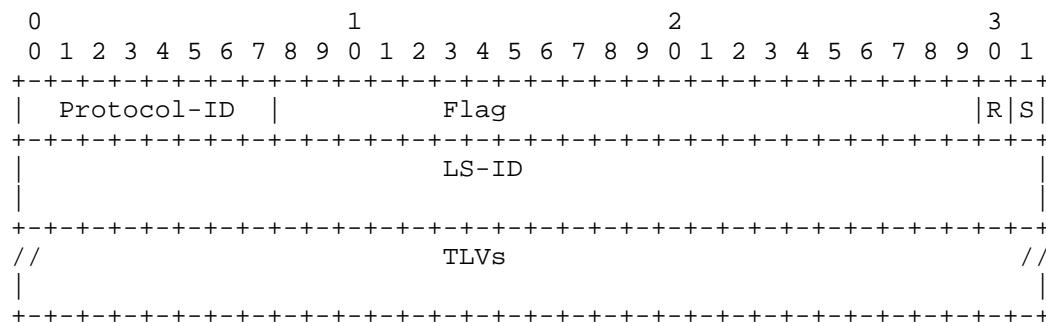
The LS (link-state) object MUST be carried within LSRpt messages and MAY be carried within PCErr messages. The LS object contains a set of fields used to specify the target node or link. It also contains a flag indicating to a PCE that the LS synchronization is in progress. The TLVs used with the LS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

- * LS Node: LS Object-Type is 1.
- * LS Link: LS Object-Type is 2.
- * LS IPv4 Topology Prefix: LS Object-Type is 3.
- * LS IPv6 Topology Prefix: LS Object-Type is 4.

The format of all types of LS objects are as follows:



Protocol-ID (8-bit): The field provides the source information. The protocol could be an IGP, BGP-LS, or an abstraction algorithm. In case the PCC only provides local information of the PCC, it MUST use Protocol-ID as Direct. The following values are defined (some of the initial values are the same as [I-D.ietf-idr-rfc7752bis]):

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
8	RSVP-TE
9	Segment Routing
10	PCEP
11	Abstraction

Flags (24-bit):

- * 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.
- * 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, link, or prefix information. A PCC creates a 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 time. All subsequent PCEP messages then address the node/link/prefix by the LS-ID. The values of 0 and 0xFFFFFFFFFFFFFFFF are reserved.

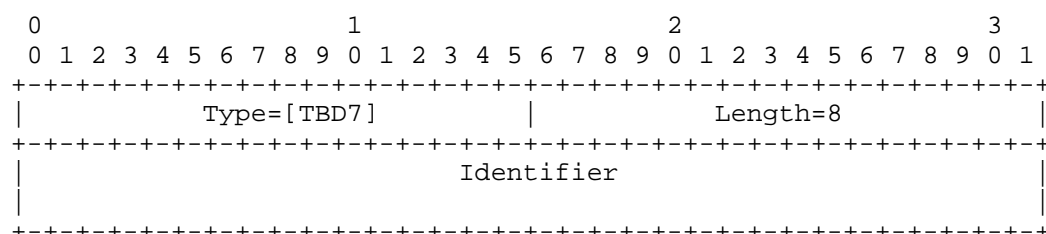
Unassigned bits are considered reserved. They MUST be set to 0 on transmission and MUST be ignored on receipt.

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

9.3.1. Routing Universe TLV

In the 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 [I-D.ietf-idr-rfc7752bis]. See [RFC8202] and [RFC6549] for more information. These instances define an independent "routing universe". The 64-bit 'Identifier' field is used to identify the "routing universe" where the LS object belongs. The LS objects representing IGP objects (nodes, links, or prefixes) 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:



The below table lists the 'Identifier' values that are defined as well-known in this draft (same as [I-D.ietf-idr-rfc7752bis]).

Identifier	Routing Universe
0	Default Layer 3 Routing topology

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 ROUTE-DISTINGUISHER TLV is defined in [RFC9168] as a Flow Specification TLVs with a separate registry. This document also adds the ROUTE-DISTINGUISHER TLV with TBD15 in the PCEP TLV registry to be used inside the LS object.

9.3.3. Virtual Network TLV

To realize ACTN, the MDSC needs to build a multi-domain topology. This topology is best served if this is an abstract view of the underlying network resources of each domain. It is also important to provide a customer view of the 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 the customer service-related knowledge of 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 [RFC9358] can be included in the 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 [I-D.ietf-idr-rfc7752bis], 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. In case 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 Descriptors TLV are globally unique. Some considerations for globally unique Node/Link/Prefix identifiers are described in [I-D.ietf-idr-rfc7752bis].

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 the 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 the specification of a flexible key for any given node/link/prefix information such that the 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 contain 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 the 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 Descriptors TLV (Local and Remote) carries one or more Node Descriptor Sub-TLV following the format of all PCEP TLVs as defined in [RFC5440], however, the Type values are selected from a new PCEP-LS sub-TLV IANA registry (see Section 13.6).

Type values are chosen so that there can be commonality with BGP-LS [I-D.ietf-idr-rfc7752bis]. This is possible because the "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" registry marks 0-255 as reserved. Thus the space of the sub-TLV values for the Type field can be partitioned as shown below -

Range	
0	Reserved - must not be allocated.
1 .. 255	New PCEP sub-TLV allocated according to the registry defined in this document.
256 .. 65535	Per BGP registry defined by [I-D.ietf-idr-rfc7752bis]. Not to be allocated in this registry.

All Node Descriptors TLVs defined for BGP-LS can then be used with PCEP-LS as well. One new PCEP sub-TLVs for Node Descriptor are defined in this document.

Sub-TLV	Description	Length	Value defined in
24	SPEAKER-ENTITY-ID	Variable	[RFC8232]

A new sub-TLV type (24) is allocated for SPEAKER-ENTITY-ID sub-TLV. The length and value fields are as per [RFC8232].

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 [I-D.ietf-idr-rfc7752bis].

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								
+-----																																							

All Link Descriptors TLVs defined for BGP-LS can then be used with PCEP-LS as well. No new PCEP sub-TLVs for Link Descriptor are defined in this document.

The format and semantics of the 'value' fields in most 'Link Descriptor' sub-TLVs correspond to the format and semantics of the 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 by IS-IS, 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 [I-D.ietf-idr-rfc7752bis].

9.3.8. Prefix Descriptors TLV

The Prefix Descriptors TLV contains Prefix Descriptors that 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 the 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) //|
|                                     |                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

All Prefix Descriptors TLVs defined for BGP-LS can then be used with PCEP-LS as well. No new PCEP sub-TLVs for Prefix Descriptor are defined in this document.

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)      //|
|                                     |                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

All Node Attributes TLVs defined for BGP-LS can then be used with PCEP-LS as well. No new PCEP sub-TLVs for Node Attributes are defined in this document.

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 the 'value' fields in IS-IS Extended IS Reachability sub-TLVs, defined in [RFC5305], [RFC5307] and [I-D.ietf-idr-rfc7752bis]. Although the encodings for 'Link Attribute' TLVs were originally defined for IS-IS, the TLVs can carry data sourced by IS-IS, OSPF, or direct.

```

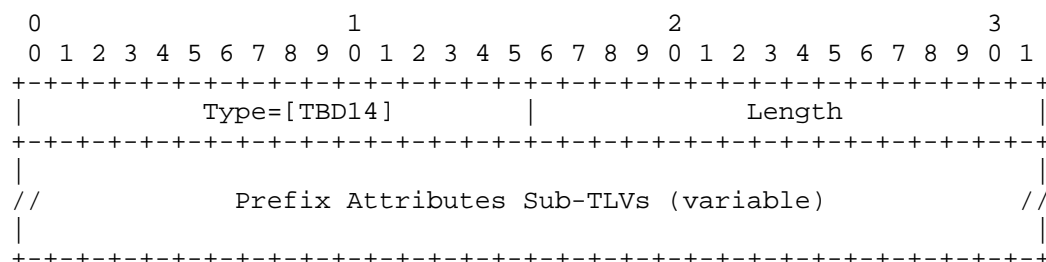
      0                   1                   2                   3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|      Type=[TBD13]                  |      Length                      |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                                     |
|                                     |                                     |
|//                               Link Attributes Sub-TLVs (variable)      //|
|                                     |                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

All Link Attributes TLVs defined for BGP-LS can then be used with PCEP-LS as well. No new PCEP sub-TLVs for Link Attributes are defined in this document.

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 metrics, 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.



All Prefix Attributes TLVs defined for BGP-LS can then be used with PCEP-LS as well. No new PCEP sub-TLVs for Prefix Attributes are defined in this document.

9.3.10. Removal of an Attribute

One of the key objectives of PCEP-LS is to encode and carry only the impacted attributes of a Node, a Link, or a Prefix. To accommodate this requirement, in case 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. An absence of a sub-TLV that was included earlier MUST be interpreted as no change.

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], [RFC9346]). 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 a new object and TLVs. Procedures and protocol extensions defined in this document do not affect the overall PCEP security model. See [RFC5440], [RFC8253]. 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 are activated on authenticated and encrypted sessions belonging to the same administrative authority.

Further, as stated in [RFC6952], PCEP implementations SHOULD support the TCP-AO [RFC5925] and not use TCP MD5 because of TCP MD5's known vulnerabilities and weaknesses. PCEP also support Transport Layer Security (TLS) [RFC8253] as per the recommendations and best current practices in [RFC9325].

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 identifier 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 following statistics:

- * Total number of LSRpt sent/received, as well as per neighbour
- * Number of errors received for LSRpt, per neighbour

* Total number of locally originated Link-State Information

These statistics should be recorded as absolute counts since the 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 peer" as well as 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 the 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	[This I-D]
	Object-Type=1 (LS Node)	
	Object-Type=2 (LS Link)	
	Object-Type=3 (LS IPv4 Prefix)	
	Object-Type=4 (LS IPv6 Prefix)	

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 "IETF Review" [RFC8126].

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	[This I-D]
8	RSVP-TE	[This I-D]
9	Segment Routing	[This I-D]
10	PCEP	[This I-D]
11	Abstraction	[This I-D]
12-255	Unassigned	

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 "IETF Review" [RFC8126]. Each bit should be tracked with the following qualities:

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

* Defining RFC

The following values are defined in this document:

Bit	Description	Reference
0-21	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)	[RFC8231] [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

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

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 a "PCEP-LS Sub-TLV Types" sub-registry 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).

Allocations from this registry are to be made according to the following assignment policies [RFC8126]:

Range	Assignment policy
0	Reserved - must not be allocated.
1 .. 251	IETF Review
252 .. 255	Experimental Use
256 .. 65535	Reserved - must not be allocated. Usage mirrors the BGP-LS TLV registry [I-D.ietf-idr-rfc7752bis]

IANA is requested to pre-populate this registry with values defined in this document as follows, taking the new values from the range 1 to 251:

Value	Meaning
24	SPEAKER-ENTITY-ID

14. TLV Code Points Summary

This section contains the global table of all TLVs in the 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	--	[ietf-pce-vn-association]
TBD8	Local Node Descriptors	256	[I-D.ietf-idr-rfc7752bis] /3.2.1.2
TBD9	Remote Node Descriptors	257	[I-D.ietf-idr-rfc7752bis] /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

Figure 4: TLV Table

15. Implementation Status

[Note to RFC Editor: Please remove this section.]

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in RFC 7942. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to RFC 7942, "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature. It is up to the individual working groups to use this information as they see fit".

The PCEP-LS protocol extensions 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 the IETF97 Hackathon and Bits-N-Bites demonstration. The use-case demonstrated was the 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 one of the high-level MDSC is 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 a quick report of failed network components.

15.2. ONOS-based Controller (MDSC and PNC)

Huawei (PNC, MDSC) and SKT (MDSC) implemented PCEP-LS during the 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 extensions) implemented PCEP-LS on its SBI as 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 (<https://wiki.onosproject.org/display/ONOS/PCEP+Protocol>) and the code at ONOS-PCEP-GITHUB (<https://github.com/opennetworkinglab/onos/tree/master/protocols/pcep>).

16. Acknowledgments

This document borrows some of the structure and text from the [I-D.ietf-idr-rfc7752bis].

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

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

17. References

17.1. Normative References

- [I-D.ietf-idr-rfc7752bis] Talaulikar, K., "Distribution of Link-State and Traffic Engineering Information Using BGP", Work in Progress, Internet-Draft, draft-ietf-idr-rfc7752bis-17, 25 August 2023, <<https://datatracker.ietf.org/doc/html/draft-ietf-idr-rfc7752bis-17>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <<https://www.rfc-editor.org/info/rfc5305>>.
- [RFC5307] Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 5307, DOI 10.17487/RFC5307, October 2008, <<https://www.rfc-editor.org/info/rfc5307>>.
- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <<https://www.rfc-editor.org/info/rfc5440>>.

- [RFC6119] Harrison, J., Berger, J., and M. Bartlett, "IPv6 Traffic Engineering in IS-IS", RFC 6119, DOI 10.17487/RFC6119, February 2011, <<https://www.rfc-editor.org/info/rfc6119>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC8232] Crabbe, E., Minei, I., Medved, J., Varga, R., Zhang, X., and D. Dhody, "Optimizations of Label Switched Path State Synchronization Procedures for a Stateful PCE", RFC 8232, DOI 10.17487/RFC8232, September 2017, <<https://www.rfc-editor.org/info/rfc8232>>.

17.2. Informative References

- [I-D.ietf-pce-pcep-yang] Dhody, D., Beeram, V. P., Hardwick, J., and J. Tantsura, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", Work in Progress, Internet-Draft, draft-ietf-pce-pcep-yang-30, 26 January 2025, <<https://datatracker.ietf.org/doc/html/draft-ietf-pce-pcep-yang-30>>.
- [I-D.ietf-teas-actn-requirements] Lee, Y., Ceccarelli, D., Miyasaka, T., Shin, J., and K. Lee, "Requirements for Abstraction and Control of TE Networks", Work in Progress, Internet-Draft, draft-ietf-teas-actn-requirements-09, 2 March 2018, <<https://datatracker.ietf.org/doc/html/draft-ietf-teas-actn-requirements-09>>.
- [I-D.kondreddy-pce-pcep-ls-sync-optimizations] Kondreddy, V. R. and M. S. Negi, "Optimizations of PCEP Link-State(LS) Synchronization Procedures", Work in Progress, Internet-Draft, draft-kondreddy-pce-pcep-ls-sync-optimizations-01, 5 April 2024, <<https://datatracker.ietf.org/doc/html/draft-kondreddy-pce-pcep-ls-sync-optimizations-01>>.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <<https://www.rfc-editor.org/info/rfc3630>>.

- [RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <<https://www.rfc-editor.org/info/rfc4203>>.
- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <<https://www.rfc-editor.org/info/rfc4364>>.
- [RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <<https://www.rfc-editor.org/info/rfc4655>>.
- [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, <<https://www.rfc-editor.org/info/rfc5392>>.
- [RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <<https://www.rfc-editor.org/info/rfc5925>>.
- [RFC6549] Lindem, A., Roy, A., and S. Mirtorabi, "OSPFv2 Multi-Instance Extensions", RFC 6549, DOI 10.17487/RFC6549, March 2012, <<https://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, <<https://www.rfc-editor.org/info/rfc6805>>.
- [RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", RFC 6952, DOI 10.17487/RFC6952, May 2013, <<https://www.rfc-editor.org/info/rfc6952>>.
- [RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", BCP 205, RFC 7942, DOI 10.17487/RFC7942, July 2016, <<https://www.rfc-editor.org/info/rfc7942>>.
- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <<https://www.rfc-editor.org/info/rfc8126>>.

- [RFC8202] Ginsberg, L., Previdi, S., and W. Henderickx, "IS-IS Multi-Instance", RFC 8202, DOI 10.17487/RFC8202, June 2017, <<https://www.rfc-editor.org/info/rfc8202>>.
- [RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", RFC 8231, DOI 10.17487/RFC8231, September 2017, <<https://www.rfc-editor.org/info/rfc8231>>.
- [RFC8253] Lopez, D., Gonzalez de Dios, O., Wu, Q., and D. Dhody, "PCEPS: Usage of TLS to Provide a Secure Transport for the Path Computation Element Communication Protocol (PCEP)", RFC 8253, DOI 10.17487/RFC8253, October 2017, <<https://www.rfc-editor.org/info/rfc8253>>.
- [RFC8281] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for PCE-Initiated LSP Setup in a Stateful PCE Model", RFC 8281, DOI 10.17487/RFC8281, December 2017, <<https://www.rfc-editor.org/info/rfc8281>>.
- [RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <<https://www.rfc-editor.org/info/rfc8283>>.
- [RFC8453] Ceccarelli, D., Ed. and Y. Lee, Ed., "Framework for Abstraction and Control of TE Networks (ACTN)", RFC 8453, DOI 10.17487/RFC8453, August 2018, <<https://www.rfc-editor.org/info/rfc8453>>.
- [RFC8637] Dhody, D., Lee, Y., and D. Ceccarelli, "Applicability of the Path Computation Element (PCE) to the Abstraction and Control of TE Networks (ACTN)", RFC 8637, DOI 10.17487/RFC8637, July 2019, <<https://www.rfc-editor.org/info/rfc8637>>.
- [RFC9168] Dhody, D., Farrel, A., and Z. Li, "Path Computation Element Communication Protocol (PCEP) Extension for Flow Specification", RFC 9168, DOI 10.17487/RFC9168, January 2022, <<https://www.rfc-editor.org/info/rfc9168>>.

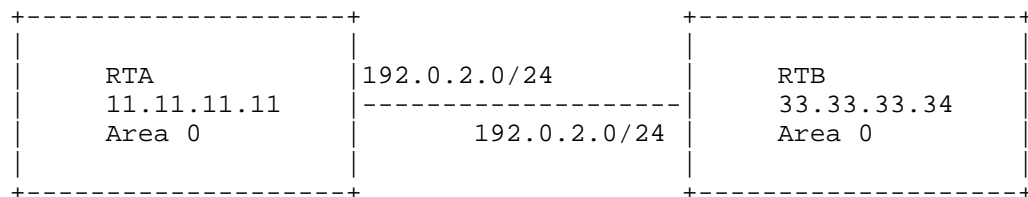
- [RFC9325] Sheffer, Y., Saint-Andre, P., and T. Fossati, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 9325, DOI 10.17487/RFC9325, November 2022, <<https://www.rfc-editor.org/info/rfc9325>>.
- [RFC9346] Chen, M., Ginsberg, L., Previdi, S., and D. Xiaodong, "IS-IS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 9346, DOI 10.17487/RFC9346, February 2023, <<https://www.rfc-editor.org/info/rfc9346>>.
- [RFC9358] Lee, Y., Zheng, H., and D. Ceccarelli, "Path Computation Element Communication Protocol (PCEP) Extensions for Establishing Relationships between Sets of Label Switched Paths and Virtual Networks", RFC 9358, DOI 10.17487/RFC9358, February 2023, <<https://www.rfc-editor.org/info/rfc9358>>.
- [WIKI] "PCE WG Wiki - Path Computation Element", <<https://wiki.ietf.org/group/pce/>>.

Appendix A. 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.

A.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

LS Node

TLV - Local Node Descriptors

Sub-TLV - 514: OSPF Area-ID: 0.0.0.0

Sub-TLV - 515: IGP Router-ID: 11.11.11.11

TLV - Node Attributes TLV

Sub-TLV(s)

LS Link

TLV - Local Node Descriptors
 Sub-TLV - 514: OSPF Area-ID: 0.0.0.0
 Sub-TLV - 515: IGP Router-ID: 11.11.11.11
TLV - Remote Node Descriptors
 Sub-TLV - 514: OSPF Area-ID: 0.0.0.0
 Sub-TLV - 515: IGP Router-ID: 22.22.22.22
TLV - Link Descriptors
 Sub-TLV - 259: IPv4 interface: 192.0.2.1
 Sub-TLV - 260: IPv4 neighbor: 192.0.2.2
TLV - Link Attributes TLV
 Sub-TLV(s)

RTB

LS Node

TLV - Local Node Descriptors
 Sub-TLV - 514: OSPF Area-ID: 0.0.0.0
 Sub-TLV - 515: IGP Router-ID: 22.22.22.22
TLV - Node Attributes TLV
 Sub-TLV(s)

LS Link

TLV - Local Node Descriptors
 Sub-TLV - 514: OSPF Area-ID: 0.0.0.0
 Sub-TLV - 515: IGP Router-ID: 22.22.22.22
TLV - Remote Node Descriptors
 Sub-TLV - 514: OSPF Area-ID: 0.0.0.0
 Sub-TLV - 515: IGP Router-ID: 11.11.11.11
TLV - Link Descriptors
 Sub-TLV - 259: IPv4 interface: 192.0.2.2
 Sub-TLV - 260: IPv4 neighbor: 192.0.2.1
TLV - Link Attributes TLV
 Sub-TLV(s)

A similar example with IPv6 address (say 2001:db8::1 and 2001:db8::2) for the links could be imagined with all other information as same and just IPv6 interface and neighbour TLVs.

A.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 A.1, the same LS Node and Link objects will be generated with a difference that it would be a designated router say RTA that generates all this information.

A.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 - 512: Autonomous System: 100 (Domain 1)

Sub-TLV - 515: IGP Router-ID: 11.11.11.11 (abstract)

LS Link

TLV - Local Node Descriptors

Sub-TLV - 512: Autonomous System: 100

Sub-TLV - 515: IGP Router-ID: 11.11.11.11 (abstract)

TLV - Remote Node Descriptors

Sub-TLV - 512: Autonomous System: 200 (Domain 2)

Sub-TLV - 515: IGP Router-ID: 22.22.22.22 (abstract)

TLV - Link Descriptors

Sub-TLV - 259: IPv4 interface: 192.0.2.1

Sub-TLV - 260: IPv4 neighbor: 192.0.2.2

TLV - Link Attributes TLV

Sub-TLV(s)

LS Link

TLV - Local Node Descriptors

Sub-TLV - 512: Autonomous System: 100

Sub-TLV - 515: IGP Router-ID: 11.11.11.11 (abstract)

TLV - Remote Node Descriptors

Sub-TLV - 512: Autonomous System: 200

Sub-TLV - 515: IGP Router-ID: 22.22.22.22 (abstract)

TLV - Link Descriptors

Sub-TLV - 259: IPv4 interface: 198.51.100.1

Sub-TLV - 260: IPv4 neighbor: 198.51.100.2

TLV - Link Attributes TLV

Sub-TLV(s)

LS Link

TLV - Local Node Descriptors

Sub-TLV - 512: Autonomous System: 100

Sub-TLV - 515: IGP Router-ID: 11.11.11.11 (abstract)
TLV - Remote Node Descriptors
Sub-TLV - 512: Autonomous System: 400 (Domain 4)
Sub-TLV - 515: IGP Router-ID: 44.44.44.44 (abstract)
TLV - Link Descriptors
Sub-TLV - 259: IPv4 interface: 203.0.113.1
Sub-TLV - 260: IPv4 neighbor: 203.0.113.2
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 the abstract internal path between
the border nodes may also be transported to the Parent PCE to enable
ACTN as described in [RFC8637] using the similar LS node and link
objects encodings.

Appendix B. Contributor Addresses

Gyan Mishra
Verizon Inc.

Email: gyan.s.mishra@verizon.com

Siva Sivabalan
Ciena Corporation

Email: ssivabal@ciena.com

Udayasree Palle

Email: udayasreereddy@gmail.com

Sergio Belotti
Nokia

Email: sergio.belotti@nokia.com

Satish Karunanithi
India

Email: satish.karunanithi@gmail.com

Cheng Li
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: c.l@huawei.com

Authors' Addresses

Dhruv Dhody
Huawei
India
Email: dhruv.ietf@gmail.com

Shuping Peng
Huawei
Huawei Bld., No.156 Beiqing Rd.
Beijing
100095
China
Email: pengshuping@huawei.com

Young Lee
Samsung Electronics
Seoul
South Korea
Email: younglee.tx@gmail.com

Daniele Ceccarelli
Cisco
Email: daniele.ietf@gmail.com

Aijun Wang
China Telecom
Beiqijia Town, Changping District
Beijing
102209
China
Email: wangaijun@tsinghua.org.cn

Gyan Mishra
Verizon Inc.
Email: gyan.s.mishra@verizon.com