

PCE Working Group
Internet Draft
Intended Status: Standard track
Expires: March 2021

Young Lee
Samsung
Haomian Zheng
Huawei
Daniele Ceccarelli
Ericsson
Wei Wang
Beijing Univ. of Posts and Telecom
Peter Park
KT
Bin Young Yoon
ETRI

September 9, 2020

PCEP Extension for Distribution of Link-State and TE information for Optical Networks

[draft-lee-pce-pcep-ls-optical-10](#)

Abstract

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

This document extends the Path Communication Element Communication Protocol (PCEP) with Link-State and TE information for optical networks.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>

This Internet-Draft will expire on September 9, 2020.

Copyright Notice

Copyright (c) 2020 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	4
2.	Applicability	5
3.	Requirements for PCEP extension	6
3.1.	Reachable source-destination	6
3.2.	Optical latency	7
4.	PCEP-LS extension for Optical Networks	7
4.1.	Node Attributes TLV	7
4.2.	Link Attributes TLV	8
4.3.	PCEP-LS for Optical Network Extension	9
5.	Security Considerations	10
6.	IANA Considerations	10
6.1.	PCEP-LS Sub-TLV Type Indicators	10
7.	References	11
7.1.	Normative References	11
7.2.	Informative References	11

Appendix A. Contributor Addresses	13
Author's Addresses	13

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 should be complete and accurate anytime so that the PCE can perform path computations.

In MPLS and GMPLS networks, Interior Gateway routing Protocols (IGPs) have been used to create and maintain a copy of the TED at each node. 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 [[RFC7752](#)].

[RFC7399] touches upon this issue: "It has also been proposed that the PCE Communication Protocol (PCEP) [[RFC5440](#)] could be extended to serve as an information collection protocol to supply information from network devices to a PCE. The logic is that the network devices may already speak PCEP and so the protocol could easily be used to report details about the resources and state in the network, including the LSP state discussed in Sections [14](#) and [15](#)."

[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 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. [[S-PCE-GMPLS](#)] specified the extensions to apply stateful PCE to GMPLS-based networks.

[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 update at the PCE.

[PCEP-LS] proposes alternative architecture approaches for learning and maintaining the Link State (and TE) information directly on a PCE from network nodes as an alternative to IGPs and BGP transport and investigate the impact from the PCE, routing protocol, and network node perspectives.

[RFC6805] describes a Hierarchical PCE (H-PCE) architecture which can be used for computing end-to-end paths for inter-domain MPLS Traffic Engineering (TE) and GMPLS Label Switched Paths (LSPs). Within the Hierarchical PCE (H-PCE) architecture [[RFC6805](#)], the Parent PCE (P-PCE) is used to compute a multi-domain path based on the domain connectivity information. A Child PCE (C-PCE) may be responsible for a single domain or multiple domains, it is used to compute the intra-domain path based on its domain topology information.

[RFC8751] presents general considerations for stateful PCE(s) in hierarchical PCE architecture. In particular, the behavior changes and additions to the existing stateful PCE mechanisms (including PCE-initiated LSP setup and active PCE usage) in the context of networks using the H-PCE architecture.

[PCEP-LS] describes a mechanism by which Link State and TE information can be collected from packet networks and shared with PCE with the PCEP itself. This is achieved using a new PCEP message format.

This draft describes an optical extension of [[PCEP-LS](#)] and explains how encodings suggested by [[PCEP-LS](#)] can be used in the optical network contexts.

1.1. Requirements Language

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

2. Applicability

There are three main applicability of this alternative proposed by this draft:

- Case 1: Where there is IGP running in optical network but there is a need for a faster link-state and TE resource collection at the PCE directly from an optical node (PCC) via a PCC-PCE interface.
 - o A PCE may receive an incremental update (as opposed to the entire TE information of the node/link).

Note: A PCE may receive full information from IGP using existing mechanism. In some cases, the convergence of full link-state and TE resource information of the entire network may not be appropriate for certain applications. Incremental update capability will enhance the accuracy of the TE information at a given time.

- Case 2: Where there is no IGP running in the optical network and there is a need for link-state and TE resource collections at the PCE directly from an optical node (PCC) via a PCC-PCE interface.
- Case 3: Where there is a need for transporting abstract optical link-state and TE information from child PCE and to a parent PCE in H-PCE [[RFC6805](#)] and [[RFC8751](#)] as well as for Physical Network Controller (PNC) to Multi-Domain Service Coordinator (MDSC) in Abstraction and Control of TE Networks (ACTN) [[RFC8453](#)].

Note: The applicability for Case 3 may arise as a consequence of Case 1 and Case 2. When TE information changes occur in the optical network, this may also affect abstracted TE information and thus needs to be updated to Parent PCE/MSDC from each child PCE/PNC.

3. Requirements for PCEP extension

The key requirements associated with link-state (and TE) distribution are identified for PCEP and listed in Section 4 of [PCEP-LS]. These new functions required in PCEP to support distribution of link-state (and TE) information are described in Section 5 of [PCEP-LS]. Details of PCEP messages and related Objects/TLVs are specified in Sections 8 and 9 of [PCEP-LS]. The key requirements and new functions specified in [PCEP-LS] are equally applicable to optical networks.

Besides the generic requirements specified in [PCEP-LS], optical specific features also need to be considered in this document. As connection-based network, there are specific parameters such as reachable table, optical latency, wavelength consistency and some other parameters that need to be included during the topology collection. Without these restrictions, the path computation may be inaccurate or infeasible for deployment, therefore these information MUST be included in the PCEP.

The procedure on how the optical parameters are used is described in following sections.

3.1. Reachable source-destination

The reachable source-destination node pair indicates that there are a few number of OCh paths between two nodes. The reachability is restricted by impairment, wavelength consistency and so on. This information is necessary at PCE to promise the path computed between source node and destination node is reachable. In this scenario, the PCE should be responsible to compute how many OCh paths are available to set up connections between source and destination node. Moreover, if a set of optical wavelength is indicated in the path computation request, the PCE should also determine whether a wavelength of the set of preselected optical wavelength is available for the source-destination pair connection.

To enable PCE to complete the above functions, the reachable relationship and OMS link information need to be reported to PCE. Once PCE detect that any wavelength is available, the corresponding OMS link should be included in a lambda plane. Then this link can be used for path computation in future.

Moreover, in a hierarchical PCE architecture, the information above need to be reported from child PCE to parent PCE, who acts as a service coordinator.

3.2. Optical latency

It is a usual case that the PCC indicates the latency when requesting the path computation. In optical networks the latency is a very sensitive parameter and there is stricter requirement on latency. Given the maximal number of OCh paths between source-destination nodes, the PCE also need to determine how many OCh path satisfies the indicated latency threshold.

There is usually high-performance algorithm running on the PCE to guarantee the performance of the computed path. During the computation, the delay factor may be converted into a kind of link weight. After the algorithm provides a few candidate paths between the source and destination nodes, the PCE SHOULD be capable to selecting one shortest path by computing the total path propagation delay.

Optical PCEs are embedded with optimization algorithm, e.g., shortest path algorithm, to improve the performance of computed path.

4. PCEP-LS extension for Optical Networks

This section provides additional PCEP-LS extension necessary to support optical networks. All Objects/TLVs defined in [[PCEP-LS](#)] are applicable to optical networks.

4.1. Node Attributes TLV

Node-Attributed TLV is defined in Section 9.2.10.1 in [PCEP-LS] as follows. This TLV is applicable for LS Node Object-Type as defined in [PCEP-LS].

```

0      1      2      3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+--+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                               |
|           Type                      |             Length          |
+-+--+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |                               |
//                                Node Attributes Sub-TLVs (variable) //
|                                     |                               |
+-+--+-----+-----+-----+-----+-----+-----+-----+-----+

```


The following 'Node Attribute' sub-TLVs are valid for optical networks:

Sub-TLV	Description	TLV/Sub-TLV	Length	Reference
TBD	Connectivity Matrix	5/14	variable	[RFC7579]
TBD	Resource Block Information	6/1	variable	[RFC7688]
TBD	Resource Block Accessibility	6/2	variable	[RFC7688]
TBD	Resource Block Wavelength Const	6/3	variable	[RFC7688]
TBD	Resource Block Pool State	6/4	variable	[RFC7688]
TBD	Resource Block Shared Access	6/5	variable	[RFC7688]
	Wavelength Avail.			

4.2. Link Attributes TLV

Link-Attributes TLV is defined in Section 9.2.10.2 in [PCEP-LS] as follows. This TLV is applicable for LS Link Object-Type as defined in [PCEP-LS].

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

The following 'Link Attribute' sub-TLVs are valid for optical networks:

Sub-TLV	Description	TLV/Sub-TLV	Length	Reference
TBD	ISCD	15	Variable	[RFC4203]
TBD	OTN-TDM SCSI	15/1,2	Variable	[RFC4203]
				[RFC7138]
TBD	WSN-LSC SCSI	15/1,2	Variable	[RFC4203]
				[RFC7688]
TBD	Flexi-grid SCSI	15/1	Variable	[RFC8363]
TBD	Port Label	34	Variable	[RFC7579]
	Restriction			[RFC7580]
				[RFC8363]

4.3. PCEP-LS for Optical Network Extension

This section provides additional PCEP-LS extension necessary to support optical networks parameters discussed in Sections 3.1 and 3.2.

The link state information collection is usually done before the path computation processing. The procedure can be divided into 1) link state collection by receiving the corresponding topology information in a periodical style; 2) path computation on PCE, triggered by receiving the path computation request message from PCC, and completed by transmitting a path computation reply with the path computation result, per [RFC4655].

For OTN networks, max bandwidth available may be per ODU 0/1/2/3 switching level or aggregated across all ODU switching levels (i.e., ODUj/k).

For WSON networks, RWA information collected from NEs would be utilized to compute light paths. The list of information can be found in [RFC7688]. More specifically, the max bandwidth available may be per lambda/frequency level (OCh) or aggregated across all lambda/frequency level. Per OCh level abstraction gives more detailed data to the P-PCE at the expense of more information processing. Either the OCh-level or the aggregated level abstraction in the RWA constraint (i.e., wavelength continuity) needs to be taken into account by the PCE during path computation. Resource Block Accessibility (i.e., wavelength conversion information) in [RFC7688] needs to be taken into account in order to guarantee the reliability for optical path computation.

[Editor's Note: Encoding will be provided in the revision, including [RFC7688](#) RWA information]

5. Security Considerations

This document extends PCEP for LS (and TE) distribution including a set of TLVs. Procedures and protocol extensions defined in this document do not affect the overall PCEP security model. See [\[RFC5440\]](#), [\[RFC8253\]](#). 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.

6. IANA Considerations

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

6.1. PCEP-LS Sub-TLV Type Indicators

This document specifies a set of PCEP-LS Sub-TLVs. IANA is requested to create a "PCEP-LS Sub-TLV Types" sub-registry in the "PCEP TLV Type Indicators" for the sub-TLVs carried in the PCEP-LS TLV (Node Attributes TLV and Link Attributes TLV).

Sub-TLV	Description	Ref Sub-TLV	Reference
TBD	Connectivity Matrix	5/14	[RFC7579]
TBD	Resource Block Information	6/1	[RFC7688]
TBD	Resource Block Accessibility	6/2	[RFC7688]
TBD	Resource Block Wavelength Const	6/3	[RFC7688]
TBD	Resource Block Pool State	6/4	[RFC7688]
TBD	Resource Block Shared Access	6/5	[RFC7688]
	Wavelength Avail.		

	TBD	ISCD	15	[RFC4203]	
	TBD	OTN-TDM SCSI	15/1,2	[RFC4203]	
				[RFC7138]	
	TBD	WSON-LSC SCSI	15/1,2	[RFC4203]	
				[RFC7688]	
	TBD	Flexi-grid SCSI	15/1	[RFC8363]	
	TBD	Port Label	34	[RFC7579]	
		Restriction		[RFC7580]	
				[RFC8363]	
+-----+-----+-----+-----+					

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", [RFC 5305](#), October 2008.
- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", [RFC 5440](#), March 2009.
- [RFC7688] Lee, Y., Ed., and G. Bernstein, Ed., "GMPLS OSPF Enhancement for Signal and Network Element Compatibility for Wavelength Switched Optical Networks", [RFC 7688](#), November 2015.
- [RFC8174] B. Leiba, "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [RFC 8174](#), May 2017.

7.2. Informative References

- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", [RFC 3630](#), September 2003.

- [RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", [RFC 4203](#), October 2005.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", [RFC 4655](#), August 2006.
- [RFC5307] Kompella, K., Ed., and Y. Rekhter, Ed., "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", [RFC 5307](#), October 2008.
- [RFC7752] Gredler, H., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and TE information using BGP", [RFC 7752](#), March 2016.
- [S-PCE-GMPLS] X. Zhang, et. al, "Path Computation Element (PCE) Protocol Extensions for Stateful PCE Usage in GMPLS-controlled Networks", [draft-ietf-pce-pcep-stateful-pce-gmpls](#), work in progress.
- [RFC7399] A. Farrel and D. King, "Unanswered Questions in the Path Computation Element Architecture", [RFC 7399](#), October 2015.
- [RFC8453] D. Ceccarelli, and Y. Lee (Editors), "Framework for Abstraction and Control of TE Networks", [RFC 8453](#), August, 2018.
- [RFC6805] A. Farrel and D. King, "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS", [RFC 6805](#), November 2012.
- [PCEP-LS] D. Dhody, Y. Lee and D. Ceccarelli "PCEP Extension for Distribution of Link-State and TE Information.", [draft-dhodylee-pce-pcep-ls](#), work in progress, July, 2020
- [RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "PCEP Extensions for Stateful PCE", [RFC 8231](#), September 2017.
- [RFC8253] Lopez, D., Gonzalez de Dios, O., Wu, Q., Dhody, D., "PCEPS: Usage of TLS to Provide a Secure Transport for the Path Computation Element Communication Protocol (PCEP)", [RFC 8253](#), October 2017.
- [RFC8281] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", [RFC 8281](#), December 2017.

- [RFC8751] D. Dhody, Y. Lee and D. Ceccarelli, "Hierarchical Stateful Path Computation Element (PCE)", [RFC8751](#), March 2020.
- [RFC8363] X. Zhang, H. Zheng, R. Casellas, O. Gonzalez de Dios, D. Ceccarelli, "GMPLS OSPF Extensions in support of Flexi-grid DWDM networks", [RFC8363](#), May 2018.

[Appendix A](#). Contributor Addresses

Dhruv Dhody
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India
Email: dhruv.ietf@gmail.com

Author's Addresses

Young Lee
Samsung
Email: younglee.tx@gmail.com

Haomian Zheng
Huawei Technologies Co., Ltd.
Email: zhenghaomian@huawei.com

Daniele Ceccarelli
Ericsson
Torshamnsgatan, 48
Stockholm
Sweden

EMail: daniele.ceccarelli@ericsson.com

Wei Wang
Beijing University of Posts and Telecom
No. 10, Xitucheng Rd. Haidian District, Beijing 100876, P.R.China

Email: weiw@bupt.edu.cn

Peter Park
KT
Email: peter.park@kt.com

Bin Yeong Yoon
ETRI
Email: byyun@etri.re.kr