

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: September 9, 2019

C. Li
M. Chen
J. Dong
Z. Li
Huawei Technologies
A. Wang
China Telecom
C. Zhou
Cisco System
March 8, 2019

PCE Controlled ID Space
draft-li-pce-controlled-id-space-02

Abstract

The Path Computation Element Communication Protocol (PCEP) provides mechanisms for Path Computation Elements (PCEs) to perform path computations in response to Path Computation Clients (PCCs) requests. The Stateful PCE extensions allow stateful control of Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Label Switched Paths (LSPs) using PCEP. Furthermore, PCEP can be used for computing paths in SR networks.

Stateful PCE provide active control of MPLS-TE LSPs via PCEP, for a model where the PCC delegates control over one or more locally configured LSPs to the PCE. Further, stateful PCE could also create and delete PCE-initiated LSPs itself. A PCE-based central controller (PCECC) simplify the processing of a distributed control plane by integrating with elements of Software-Defined Networking (SDN).

In some use cases, such as PCECC or Binding Segment Identifier (SID) for Segment Routing (SR), there are requirements for a stateful PCE to make allocation of labels, SIDs, etc. These use cases require PCE aware of various identifier spaces where to make allocations on behalf of PCC. This document describes a mechanism for PCC to inform the PCE of the identifier space under its control via PCEP. The identifier could be MPLS label, SID or any other to-be-defined identifier to be allocated by a PCE.

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 September 9, 2019.

Copyright Notice

Copyright (c) 2019 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 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
2. Terminology	4
2.1. Requirements Language	4
3. Use cases	4
3.1. PCE-based Central Control	4
3.1.1. PCECC for MPLS/SR-MPLS	4
3.1.2. PCECC for SRv6	5
3.2. Binding SID Allocation	5
4. Overview	5
5. Objects	7
5.1. Open Object	7
5.1.1. LABEL-CONTROL-SPACE TLV	7
5.1.2. FUNCT-ID-CONTROL-SPACE TLV	8
6. Other Considerations	10
7. IANA Considerations	11
8. Security Considerations	11
9. Acknowledgements	11
10. References	11
10.1. Normative References	11
10.2. Informative References	12
Appendix A. Contributors	14

Authors' Addresses	14
------------------------------	----

1. Introduction

[RFC5440] defines the stateless Path Computation Element communication Protocol (PCEP) for Path Computation Elements (PCEs) to perform path computations in response to Path Computation Clients (PCCs) requests. For supporting stateful operations, [RFC8231] specifies a set of extensions to PCEP to enable stateful control of LSPs within and across PCEP sessions in compliance with [RFC4657]. Furthermore, [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.

[RFC8283] introduces the architecture for PCE as a central controller, it examines the motivations and applicability for PCEP as a control protocol in this environment, and introduces the implications for the protocol. Also, [I-D.ietf-pce-pcep-extension-for-pce-controller] specifies the procedures and PCEP protocol extensions for using the PCE as the central controller, where LSPs are calculated/setup/initiated and label forwarding entries are downloaded through extending PCEP. However, the document assumes that label range to be used by a PCE is known and set on both PCEP peers. This extension adds the capability to advertise the range via a PCEP extension.

Similarly, [I-D.zhao-pce-pcep-extension-pce-controller-sr] specifies the procedures and PCEP protocol extensions when a PCE-based controller is also responsible for configuring the forwarding actions on the routers (SR SID distribution in this case), in addition to computing the paths for packet flows in a segment routing network and telling the edge routers what instructions to attach to packets as they enter the network. However, the document assumes that label range to be used by a PCE is known and set on both PCEP peers. This extension adds the capability to advertise the range (from SRGB or SRLB of the node) via a PCEP extension.

In addition, [I-D.dhody-pce-pcep-extension-pce-controller-srv6] specifies the procedures and PCEP protocol extensions of PCECC for SRv6. An SRv6 SID is represented as LOC:FUNCT where LOC is the L most significant bits and FUNCT is the 128-L least significant bits. The FUNCT part of the SID is an opaque identification of a local function bound to the SID. This extension adds the capability to advertise the range of Function ID (FUNCT part) via a PCEP extension.

Once the PCC/node has given control over an ID space (for example labels), the PCC/node MUST NOT allocate the ID from this ID space.

For example, a PCC/node MUST NOT use this labels from the PCE controlled label space to make allocation for VPN Prefix distributed via BGP or labels used for LDP/RSVP-TE signalling. This is done to make sure that the PCE control over ID space does not conflict with the existing node allocation.

The use case are described in Section 3. The ID space range information can be advertised via the TLVs in the Open message. The detailed procedures are described in Section 4, and the objects' format is specified in Section 5.

2. Terminology

This memo makes use of the terms defined in [RFC5440], [RFC8231], [RFC8283] and [RFC8402].

2.1. 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.

3. Use cases

3.1. PCE-based Central Control

A PCE-based central controller (PCECC) can simplify the processing of a distributed control plane by integrating with elements of SDN. Thus, the LSP/SR path can be calculated/setup/initiated and the label/SID forwarding entries can also be downloaded through a centralized PCE server to each network devices along the path while leveraging the existing PCE technologies as much as possible.

3.1.1. PCECC for MPLS/SR-MPLS

[I-D.ietf-pce-pcep-extension-for-pce-controller] describes a mode where LSPs are provisioned as explicit label instructions at each hop on the end-to-end path. Each router along the path must be told what label forwarding instructions to program and what resources to reserve. The controller uses PCEP to communicate with each router along the path of the end-to-end LSP. For this to work, the PCE-based controller will take responsibility for managing some part of the MPLS label space for each router that it controls as described in section 3.1.2. of [RFC8283]. A mechanism for a PCC to inform the PCE of such a label space to control is needed within PCEP.

[I-D.ietf-pce-segment-routing] specifies extensions to PCEP that allow a stateful PCE to compute, update or initiate SR-TE paths. [I-D.zhao-pce-pcep-extension-pce-controller-sr] describes the mechanism for PCECC to allocate and provision the node/prefix/adjacency label (SID) via PCEP. To make such allocation, PCE needs to be aware of the label space from Segment Routing Global Block (SRGB) or Segment Routing Local Block (SRLB) [RFC8402] of the node that it controls. A mechanism for a PCC to inform the PCE of such label space to control is needed within PCEP. The full SRGB/SRLB of a node could be learned via existing IGP or BGP-LS mechanism.

3.1.2. PCECC for SRv6

[I-D.dhody-pce-pcep-extension-pce-controller-srv6] describes the mechanism for PCECC to allocate and provision the SRv6 SID via PCEP. An SRv6 SID is represented as LOC:FUNCT ([I-D.filsfils-spring-srv6-network-programming]) where LOC is the L most significant bits and FUNCT is the 128-L least significant bits. The FUNCT part of the SID is an opaque identification of a local function bound to the SID. To make such allocation, PCE needs to be aware of the Function ID space (FUNCT part) of the node that it controls. A mechanism for a PCC to inform the PCE of such a Function ID space to control is needed within PCEP.

3.2. Binding SID Allocation

The headend of an SR Policy binds a Binding SID to its policy [I-D.ietf-spring-segment-routing-policy]. The instantiation of which may involve a list of SIDs. Currently Binding SID are allocated by the node, but there is an inherent advantage in the Binding SID to be allocated by a PCE to allow SR policies to be dynamically created, updated according to the network status and operations. This is described in [I-D.ietf-pce-pcep-extension-for-pce-controller]. Therefore, a PCE needs to obtain the authority and control to allocate Binding SID actively from the PCC's label space as described in above use case.

4. Overview

During PCEP Initialization Phase, Open messages are exchanged between PCCs and PCEs. The OPEN object may also contain a set of TLVs used to convey capabilities in the Open message. The term 'ID' in this document, could be a MPLS label, SRv6 Function ID or any other future ID space for PCE to control and allocate from. A PCC can include a corresponding ID-CONTROL-SPACE TLVs in the OPEN Object to inform the corresponding ID space information that it wants the PCE to control. This TLV MUST NOT be included by the PCE and MUST be ignored on

receipt by a PCC. This is an optional TLV, the PCE could be aware of the ID space from some other means outside of PCEP.

For delegating multiple types of ID space, multiple TLVs corresponding to each ID type MUST be included in an Open message. The ID type can be MPLS label or other type of ID. The following ID-CONTROL-SPACE TLV is defined in this document -

- o LABEL-CONTROL-SPACE TLV - for MPLS Labels (including for SR-MPLS)
- o FUNCTION-ID-CONTROL-SPACE TLV - for SRv6 SID Function ID

The procedure of ID space control to PCE is shown below:

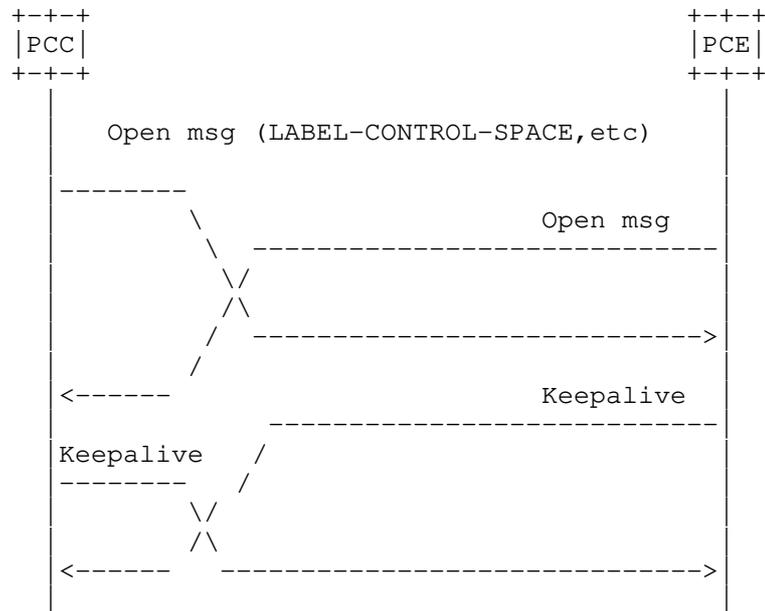


Figure 1: ID space control to PCE

If the ID space control procedure is successful, the PCE will return a KeepAlive message to the PCC. If there is any error in processing the corresponding TLV, an Error (PCErr) message will be sent to the PCC with Error-Type=1 (PCEP session establishment failure) and Error-value=TBD (ID space control failure).

After this process, a stateful PCE can learn the PCE controlled ID spaces of a node (PCC) under its control. A PCE can then allocate IDs within the control ID space. For example, a PCE can actively

allocate labels and download forwarding instructions for the PCECC LSP as described in [I-D.ietf-pce-pcep-extension-for-pce-controller]. A PCE can also allocate labels from the PCE controlled portion of the SRGB/SRLB for PCECC-SR [I-D.zhao-pce-pcep-extension-pce-controller-sr]. The full SRGB/SRLB of a node could be learned via existing IGP or BGP-LS mechanism.

5. Objects

5.1. Open Object

For advertising the PCE controlled ID space to a PCE, this document defines several TLVs within the OPEN object.

5.1.1. LABEL-CONTROL-SPACE TLV

For a PCC to inform the label space under the PCE control, this document defines a new LABEL-CONTROL-SPACE TLV.

The LABEL-CONTROL-SPACE TLV is an optional TLV in the OPEN object, and its format is shown in the following figure:

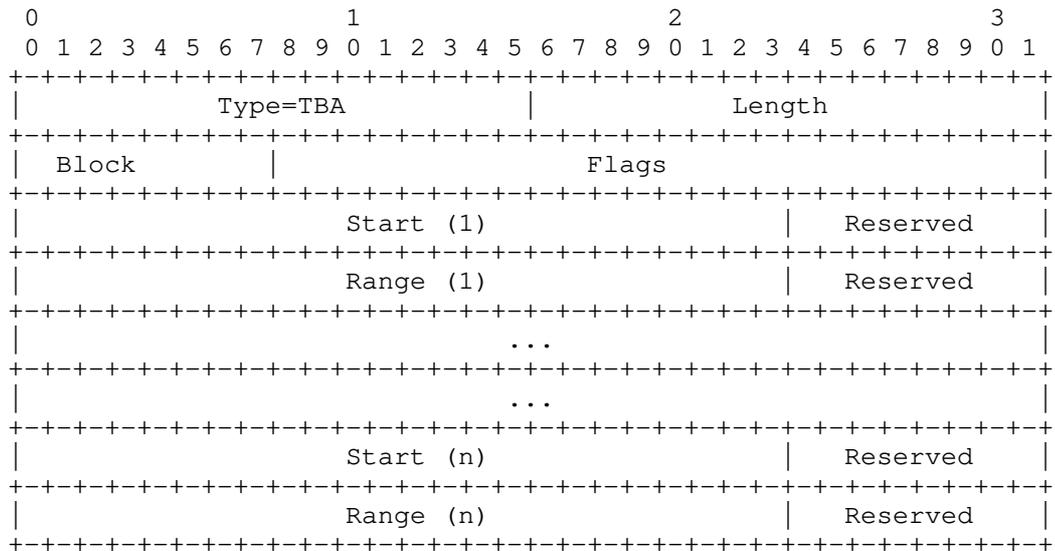


Figure 2: LABEL-CONTROL-SPACE TLV

The type (16 bits) of the TLV is TBA. The length field (16 bits) and has a variable value.

Block(8 bits): the number of ID blocks. The range of a block is described by a start field and a range field.

Flags (24 bits): No flag is currently defined. The unassigned bits of Flags field MUST be set to 0 on transmission and MUST be ignored on receipt.

Start(i) (24 bits): indicates the beginning of the label block i.

Range(i) (24 bits): indicates the range of the label block i.

Reserved: SHOULD be set to 0 on transmission and MUST be ignored on reception.

LABEL-CONTROL-SPACE TLV SHOULD be included only once in a Open Message. On receipt, only the first instance is processed and others MUST be ignored.

A stateful PCE can actively allocate labels and download forwarding instructions for the PCECC LSP as described in [I-D.ietf-pce-pcep-extension-for-pce-controller]. A PCE can also allocate labels from SRGB/SRLB for PCECC-SR [I-D.zhao-pce-pcep-extension-pce-controller-sr]. The Binding Segments can also be selected for the PCE controlled space [I-D.ietf-pce-pcep-extension-for-pce-controller].

5.1.2. FUNCT-ID-CONTROL-SPACE TLV

For a PCC to inform the SRv6 SID Function ID space under the PCE control, this document defines a new FUNCT-ID-CONTROL-SPACE TLV.

The FUNCT-ID-CONTROL-SPACE TLV is an optional TLV for use in the OPEN object, and its format is shown in the following figure:

The unassigned bits of Flags field MUST be set to 0 on transmission and MUST be ignored on receipt.

Start(i) (128 bits): indicates the beginning of the Function ID block i.

Range(i) (128 bits): indicates the range of the Function ID block i.

Loc size(8 bits): indicates the bit length of a Locator. Appears only when the L-flag is set.

Locator (variable length): the value of a Locator. The Function ID spaces specified in this TLV are associated with this locator.

As per [RFC5440], the value portion of the PCEP TLV needs to be 4-bytes aligned, so a FUNCT-ID-CONTROL-SPACE TLV is padded with trailing zeros to a 4-byte boundary.

Multiple FUNCT-ID-CONTROL-SPACE TLVs MAY be included in a OPEN object to specify Function ID space specific to each locator.

A stateful PCE can actively allocate SRv6 SID and download forwarding instructions for the PCECC SRv6 path as described in [I-D.dhody-pce-pcep-extension-pce-controller-srv6].

Note that SRv6 SID allocation involves LOC:FUNCT; the LOC is assumed to be known at PCE and FUNCT is allocated from the PCE controlled Function ID block.

6. Other Considerations

In case of multiple PCEs, a PCC MAY decide to give control over different ID space to each instance of the PCE. In case a PCC includes the same ID space to multiple PCEs, the PCE SHOULD use synchronization mechanism (such as [I-D.litkowski-pce-state-sync]) to avoid allocating the same ID.

The PCE would allocate ID from the PCE controlled ID space. The PCC would not allocate ID by itself from this space as long as it has an active PCEP session to a PCE to which it has given control over the ID space.

Note that if there is any change in the ID space, the PCC MUST bring the session down and re-establish the session with new TLVs. During state synchronization the PCE would need to consider the new ID space into consideration and SHOULD re-establish the LSP/SR-paths if needed.

The PCC can regain control of the ID space by closing the PCEP session and require new session without ID space TLVs specified in this document.

7. IANA Considerations

TBA.

8. Security Considerations

TBA.

9. Acknowledgements

TBD.

10. References

10.1. Normative References

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

10.2. Informative References

- [RFC4657] Ash, J., Ed. and J. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, DOI 10.17487/RFC4657, September 2006, <<https://www.rfc-editor.org/info/rfc4657>>.
- [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.
- [I-D.ietf-pce-segment-routing]
Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "PCEP Extensions for Segment Routing", draft-ietf-pce-segment-routing-16 (work in progress), March 2019.
- [I-D.ietf-pce-pcep-extension-for-pce-controller]
Zhao, Q., Li, Z., Negi, M., and C. Zhou, "PCEP Procedures and Protocol Extensions for Using PCE as a Central Controller (PCECC) of LSPs", draft-ietf-pce-pcep-extension-for-pce-controller-01 (work in progress), February 2019.
- [I-D.zhao-pce-pcep-extension-pce-controller-sr]
Zhao, Q., Li, Z., Negi, M., and C. Zhou, "PCEP Procedures and Protocol Extensions for Using PCE as a Central Controller (PCECC) of SR-LSPs", draft-zhao-pce-pcep-extension-pce-controller-sr-04 (work in progress), February 2019.
- [I-D.litkowski-pce-state-sync]
Litkowski, S., Sivabalan, S., and D. Dhody, "Inter Stateful Path Computation Element communication procedures", draft-litkowski-pce-state-sync-04 (work in progress), October 2018.

[I-D.ietf-spring-segment-routing-policy]

Filsfils, C., Sivabalan, S., daniel.voyer@bell.ca, d., bogdanov@google.com, b., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-02 (work in progress), October 2018.

[I-D.dhody-pce-pcep-extension-pce-controller-srv6]

Negi, M., Li, Z., and X. Geng, "PCEP Procedures and Protocol Extensions for Using PCE as a Central Controller (PCECC) for SRv6", draft-dhody-pce-pcep-extension-pce-controller-srv6-01 (work in progress), February 2019.

[I-D.filsfils-spring-srv6-network-programming]

Filsfils, C., Camarillo, P., Leddy, J., daniel.voyer@bell.ca, d., Matsushima, S., and Z. Li, "SRv6 Network Programming", draft-filsfils-spring-srv6-network-programming-07 (work in progress), February 2019.

Appendix A. Contributors

Dhruv Dhody

Huawei Technologies

Divyashree Techno Park, Whitefield

Bangalore, Karnataka 560066

India

EMail: dhruv.ietf@gmail.com

Authors' Addresses

Cheng Li

Huawei Technologies

Huawei Campus, No. 156 Beiqing Rd.

Beijing 100095

China

EMail: chengli13@huawei.com

Mach(Guoyi) Chen

Huawei Technologies

Huawei Campus, No. 156 Beiqing Rd.

Beijing 100095

China

EMail: Mach.chen@huawei.com

Jie Dong

Huawei Technologies

Huawei Campus, No. 156 Beiqing Rd.

Beijing 100095

China

EMail: jie.dong@huawei.com

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

EMail: lizhenbin@huawei.com

Aijun Wang
China Telecom
Beiqijia Town,
Beijing, Changping District 102209
China

EMail: wangaj.bri@chinatelecom.cn

Chao Zhou
Cisco System
San Jose
USA

EMail: chao.zhou@cisco.com