PCE
Internet-Draft
Intended status: Standards Track
Expires: January 8, 2022

Q. Xiong
S. Peng
ZTE Corporation
F. Qin
China Mobile
July 7, 2021

# PCEP Extension for SR-MPLS Entropy Label Position
# draft-peng-pce-entropy-label-position-06

## Abstract

This document proposes a set of extensions for PCEP to configure the entropy label position for SR-MPLS networks.

## 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 January 8, 2022.

## Copyright Notice

Copyright (c) 2021 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 . . . . . 2
2. Conventions used in this document . . . . . 3
2.1. Terminology . . . . . 3
2.2. Requirements Language . . . . . 3
3. Entropy Labels in SR-MPLS Scenario with PCE . . . . . 3
4. PCEP Extensions . . . . . 5
4.1. The OPEN Object . . . . . 5
4.2. The LSP-EXTENDED-FLAG TLV . . . . . 5
4.3. The PATH-MINIMUM-ERLD TLV . . . . . 6
4.4. The SR-ERO Object . . . . . 6
5. Operations . . . . . 7
6. Security Considerations . . . . . 7
7. Acknowledgements . . . . . 7
8. IANA Considerations . . . . . 7
8.1. New SR PCE Capability Flag Registry . . . . . 7
8.2. New LSP-EXTENDED-FLAG Flag Registry . . . . . 8
8.3. The PATH-MINIMUM-ERLD TLV . . . . . 8
8.4. New SR-ERO Flag Registry . . . . . 8
9. Normative References . . . . . 9
Authors' Addresses . . . . . 10

## 1. Introduction

[RFC5440] describes the Path Computation Element Protocol (PCEP) which is used between a Path Computation Element (PCE) and a Path Computation Client (PCC) (or other PCE) to enable computation of Multi-protocol Label Switching (MPLS) for Traffic Engineering Label Switched Path (TE LSP). PCEP Extensions for the Stateful PCE Model [RFC8231] describes a set of extensions to PCEP to enable active control of MPLS-TE and Generalized MPLS (GMPLS) tunnels. [RFC8281] describes the setup and teardown of PCE-initiated LSPs under the active stateful PCE model, without the need for local configuration on the PCC, thus allowing for dynamic centralized control of a network.

Segment Routing (SR) leverages the source routing paradigm. Segment Routing can be instantiated on MPLS data plane which is referred to as SR-MPLS [RFC8660]. SR-MPLS leverages the MPLS label stack to construct the SR path. PCEP Extensions for Segment Routing [RFC8664] specifies extensions to the PCEP that allow a stateful PCE to compute and initiate TE paths, as well as a PCC to request a path subject to certain constraint(s) and optimization criteria in SR networks.

Entropy label (EL) [RFC6790] is a technique used in the MPLS data plane to improve load-balancing. Entropy Label Indicator (ELI) can be immediately preceding an EL in the MPLS label stack. The idea

behind the EL is that the ingress router computes a hash based on several fields from a given packet and places the result in an additional label, named "entropy label".  Then, this entropy label can be used as part of the hash keys used by an LSR.  Using the entropy label as part of the hash keys reduces the need for deep packet inspection in the LSR while keeping a good level of entropy in the load-balancing.  When the entropy label is used, the keys used in the hashing functions are still a local configuration matter and an LSR may use solely the entropy label or a combination of multiple fields from the incoming packet.

[RFC8662] proposes to use entropy labels for SR-MPLS networks and mutiple <ELI, EL> pairs SHOULD be inserted in the SR-MPLS label stack.  The ingress node may decide the number and place of the ELI/ELs which need to be inserted into the label stack.  But in some cases, the the controller (e.g.  PCE) could be used to perform the TE path computation as well as the Entropy Label Position (ELP) which is useful for inter-domain scenarios.  This document proposes a set of extensions for PCEP to configure the ELP information for SR-MPLS networks.

## 2.  Conventions used in this document

### 2.1.  Terminology

The terminology is defined as [RFC5440], [RFC6790], [RFC8664] and [RFC8662].

### 2.2.  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.

## 3.  Entropy Labels in SR-MPLS Scenario with PCE

[RFC8662] proposes to use entropy labels for SR-MPLS networks.  The Entropy Readable Label Depth (ERLD) is defined as the number of labels which means that the router will perform load-balancing using the ELI/EL.  An appropriate algorithm should consider the following criteria:

- o  a limited number of <ELI, EL> pairs SHOULD be inserted in the SR-MPLS label stack;

- o the inserted positions SHOULD be whithin the ERLD of a maximize number of transit LSRs;

- o a minimum number of <ELI, EL> pairs SHOULD be inserted while satisfying the above criteria.

As described in [RFC8662] section 7, the ERLD value is important for inserting ELI/EL and the ingress node need to evaluate the minimum ERLD value along the node segment path.  But it will add complexity in the ELI/EL insertion process.  Moreover, the ingress node cannot find the minimum ERLD along the path and does not support the computation of the minimum ERLD especilly in inter-domain scenarios. As the Figure 1 shown, in SR-MPLS inter-domain scenario, the ingress node of the first domain could not get the ERLD information of other nodes of other domains.

```
       +-----+                +-----+                 +-----+
       |PCE-1|                |PCE-2|                 |PCE-3|
       +--+--+                +--+--+                 +--+--+
          |                      |                       |
 .........+..........   .........+..........    .........+...........
 .                  .   .                  .    .                   .
 .+---+       +---+ .   . +---+      +---+ .    .+---+      +----+   .
 .| A |-------| B |------ | C |------| X |-------| Y |------| Z  |   .
 .+---+       +---+ .   . +---+      +---+ .    .+---+      +----+   .
 .    SR-AS 1       .   .   SR-AS 2        .    .     SR-AS 3       .
 ....................   ....................    .....................
```

Figure 1: Entropy Labels in SR-MPLS Inter-Domain Scenario

The PCEs could get the information of all nodes such as MSD, ELC and ERLD through IGP/BGP-LS protocol and can compute the minimum ERLD along the end-to-end path.  Moreover, the PCEs also can compute the Entropy Label Position (ELP) including the number and the places of the ELI/ELs.  Then the ingress nodes MAY be required to support the capabilities of inserting multiple ELI/ELs and need to advertise the capabilities to the PCEs.

This document provides an option for the ingress nodes to get the path minimum ERLD from PCE as defined in section 4.3 and the ingress nodes can perform the insertion of ELI/ELs based on the minimum ERLD value.  For another option, the PCE can perform the computation of the end-to-end path as well as the positions of entropy labels defined in section 4.4.  The ingress nodes can directly insert the ELI/ELs based on the positions.

## 4. PCEP Extensions

### 4.1. The OPEN Object

As defined in [RFC8664], PCEP speakers use SR PCE Capability sub-TLV to exchange information about their SR capability when PST=1 in the PST List of the PATH-SETUP-TYPE-CAPABILITY TLV carried in Open object. This document defined a new flag (E-flag) for SR PCE Capability sub-TLV as shown in Figure 2.

```
 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=4           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Reserved              |   Flags |E|N|X|      MSD      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: E-flag in SR-PCE-CAPABILITY sub-TLV

E (Entropy Label Configuration is supported) : A PCE sets this flag bit to 1 carried in Open message to indicate that it supports the computation of SR path with ELP information. A PCC sets this flag to 1 to indicate that it supports the capability of inserting multiple ELI/EL pairs and and supports the results of SR path with ELP from PCE.

### 4.2. The LSP-EXTENDED-FLAG TLV

The LSP Object is defined in Section 7.3 of [RFC8231]. This document defiend a new flag (E-flag) for the LSP-EXTENDED-FLAG TLV carried in LSP Object as defined in [I-D.ietf-pce-lsp-extended-flags]. The format is shown as Figure 3:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type=TBD            |          Length               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                             |E|
//                 LSP Extended Flags                          //
|                                                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: E-flag in LSP-EXTENDED-FLAG TLV

E (Request for ELP Configuration) : If the bit is set to 1, it indicates that the PCC requests PCE to compute the SR path with ELP information.  A PCE would also set this bit to 1 to indicate that the ELP information is included by PCE and encoded in the PCRep, PCUpd or PCInitiate message.

### 4.3. The PATH-MINIMUM-ERLD TLV

The PATH-MINIMUM-ERLD TLV is an optional TLV for use in the LSP Object for the path minimum ERLD configuration.  The type of this TLV is to be allocated by IANA.  The format is as shown below.

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Reserved               |Path Minimum ERLD|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: The PATH-MINIMUM-ERLD TLV

Path Minimum ERLD: 8 bits, indicates the minimum ERLD of the nodes along the path.

### 4.4. The SR-ERO Object

SR-ERO subobject is used for SR-TE path which consists of one or more SIDs as defined in [RFC8664].  This document defiend a new flag (E-flag) for the SR-ERO subobject as Figure 4 shown:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|   Type=36   |     Length    |  NT   |     Flags   |E|F|S|C|M|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         SID (optional)                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                    NAI (variable, optional)                  //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: E-flag in SR-ERO subobject

E (ELP Configuration) : If this flag is set, it means that the position after this SR-ERO subobject is the position to insert <ELI, EL>, otherwise it cannot insert <ELI, EL> after this segment.

## 5. Operations

The SR path is initiated by PCE or PCC with PCReq, PCInitiated or PCUpd messages and the E bit is set to 1 in LSP object to request the ELP configuration. The SR-TE path being recieved by PCC with SR-ERO segment list, for example, <S1, S2, S3, S4, S5, S6>, especially S3 and S6 with E-flag set. It indicates that two <ELI, EL> pairs MUST be inserted into the label stack of the SR-TE forwarding entry, repectively after the label for S3 and label for S6. With EL information, the label stack for SR-MPLS would be <label1, label2, label3, ELI, EL, label4, label5, label6, ELI, EL>.

## 6. Security Considerations

TBA

## 7. Acknowledgements

TBA

## 8. IANA Considerations

### 8.1. New SR PCE Capability Flag Registry

SR PCE Capability TLV is defined in [RFC8664], and the registry to manage the Flag field of the SR PCE Capability TLV is requested in [RFC8664]. IANA is requested to make allocations from the registry, as follows:

| Value | Name | Reference |
|---|---|---|
| TBD11 | Entropy Label Configuration is supported (E) | [this document] |

Table 1

## 8.2. New LSP-EXTENDED-FLAG Flag Registry

[I-D.ietf-pce-lsp-extended-flags] defines the LSP-EXTENDED-FLAG TLV. IANA is requested to make allocations from the Flag field registry, as follows:

| Value | Name | Reference |
|---|---|---|
| TBD | Request for ELP Configuration (E) | [this document] |

Table 2

## 8.3. The PATH-MINIMUM-ERLD TLV

This document requests that a new sub-registry named "PATH-MINIMUM-ERLD TLV" carried in LSP object.

| Value | Name | Reference |
|---|---|---|
| TBD | PATH-MINIMUM-ERLD TLV | [this document] |

Table 3

## 8.4. New SR-ERO Flag Registry

SR-ERO subobject is defined in [RFC8664], and the registry to manage the Flag field of SR-ERO is requested in [RFC8664].  IANA is requested to make allocations from the registry, as follows:

| Value | Name | Reference |
|---|---|---|
| 36 | ELP Configuration (E) | [this document] |

Table 4

## 9. Normative References

[I-D.ietf-pce-lsp-extended-flags]
Xiong, Q., "LSP Object Flag Extension of Stateful PCE", draft-ietf-pce-lsp-extended-flags-00 (work in progress), March 2021.

[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>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[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>.

[RFC8623] Palle, U., Dhody, D., Tanaka, Y., and V. Beeram, "Stateful Path Computation Element (PCE) Protocol Extensions for Usage with Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 8623, DOI 10.17487/RFC8623, June 2019, <https://www.rfc-editor.org/info/rfc8623>.

[RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <https://www.rfc-editor.org/info/rfc8660>.

[RFC8662] Kini, S., Kompella, K., Sivabalan, S., Litkowski, S., Shakir, R., and J. Tantsura, "Entropy Label for Source Packet Routing in Networking (SPRING) Tunnels", RFC 8662, DOI 10.17487/RFC8662, December 2019, <https://www.rfc-editor.org/info/rfc8662>.

[RFC8664] Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing", RFC 8664, DOI 10.17487/RFC8664, December 2019, <https://www.rfc-editor.org/info/rfc8664>.

## Authors' Addresses

Quan Xiong
ZTE Corporation
No.6 Huashi Park Rd
Wuhan, Hubei  430223
China

Email: xiong.quan@zte.com.cn

Shaofu Peng
ZTE Corporation
No.50 Software Avenue
Nanjing, Jiangsu  210012
China

Email: peng.shaofu@zte.com.cn

Fengwei Qin
China Mobile
Beijing
China

Email: qinfengwei@chinamobile.com