Workgroup: PCE
Internet-Draft:
draft-ietf-pce-entropy-label-position-01
Published: 25 February 2024
Intended Status: Standards Track
Expires: 28 August 2024
Authors: Q. Xiong, ZTE Corporation
S. Peng, ZTE Corporation
F. Qin, China Mobile
J. Zhao, CAICT

Path Computation Element Communication Protocol (PCEP) Extension for SR-MPLS Entropy Label Positions

Abstract

Entropy label (EL) can be used in the SR-MPLS data plane to improve load-balancing and multiple Entropy Label Indicator (ELI)/EL pairs may be inserted in the SR-MPLS label stack as per RFC8662.

This document proposes a set of extensions for Path Computation Element Communication Protocol (PCEP) to configure the Entropy Label Positions (ELP) for SR-MPLS networks.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

[RFC5440] describes the Path Computation Element Computation 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 Label Switch Router (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 multiple <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. The Entropy Label Position (ELP) is used to indicate the positions of the ELI/ELs which need to be inserted into the label stack as per [I-D.ietf-idr-bgp-srmpls-elp].

In some cases, the controller(e.g. PCE) could be used to perform the TE path computation as well as ELP information 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 in [RFC8662] section 4.

As described in [RFC8662] section 7.2.1, the ELRD value is an important consideration when inserting ELI/EL and the minimum ELRD must be evaluated for each node along a computed path. This necessary step adds additional complexity in the ELI/EL insertion process and it may not be feasible for an ingress router to compute the appropriate ERLD for each node in the path, since a SR-MPLS path may contain segments the ingress router can resolve such as 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

When computing the ELI/EL positions, the PCE MUST take into consideration Maximum SID Depth (MSD) imposition. The PCEs could get the information of all nodes such as MSD (e.g. Base MPLS Imposition MSD (BMI-MSD) or ERLD-MSD) through Interior Gateway Protocol (IGP) and can compute the minimum ERLD along the end-to-end path. IS-IS [RFC8491] and OSPF [RFC8476] provide examples of advertisement of the MSD. The ERLD value can be collected via IS-IS [RFC9088], and OSPF [RFC9089]. Moreover, the PCEs also can compute the ELP information 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 proposes the extensions for PCE to perform the computation of the end-to-end path as well as the positions of entropy labels in SR-MPLS networks. 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 defines a new flag (E-flag) for SR PCE Capability sub-TLV.

E (ELP 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 defines a new flag (E-flag) for the LSP-EXTENDED-FLAG TLV carried in LSP Object as defined in [RFC9357].

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. The PCE SHOULD set this bit to 1 to indicate that the ELP information is included by PCE and encoded in the Path Computation Reply (PCRep) message as per [RFC5440]. And in a stateful PCE model, it also CAN be carried in Path Computation Update Request (PCUpd) message as per [RFC8231] or LSP Initiate Request (PCInitiate) message as per [RFC8281].

4.3. 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 defined a new flag (E-flag) for the SR-ERO subobject.

E (ELP Configuration) : If this flag is set, the PCC SHOULD insert <ELI, EL> into the position after this SR-ERO subobject, otherwise it SHOULD not insert <ELI, EL> after this segment.

5. Operational Example

A PCC can request the computation of SR path and a PCE may respond with PCRep message. And the SR path can also be initiated by PCE with PCInitiate or PCUpd message in stateful PCE mode. When the E bit in LSP object is set to 1 within the message, it indicates to request the ELP configuration with the SR path. The SR path being received by PCC encoded in SR-ERO, for example, <S1, S2, S3, S4, S5, S6>, especially S3 and S6 with E-flag set. It indicates that two <ELI, EL>

pairs SHOULD be inserted into the label stack of the SR forwarding entry, respectively 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

This document defines a new E bit for entropy label, which do not introduce any new security considerations beyond those already listed in [RFC9357], [RFC8662] and [RFC8664].

7. IANA Considerations

7.1. 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
TBD1	ELP Configuration is supported (E)	[this document]

Table 1

7.2. New LSP-EXTENDED-FLAG Flag Registry

[RFC9357] defines the LSP-EXTENDED-FLAG TLV. IANA is requested to make allocations from the Flag field registry, as follows:

Value	Name	Reference
TBD2	Request for ELP Configuration (E)	[this document]

Table 2

7.3. 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
TBD3	ELP Configuration (E)	[this document]

Table 3

8. Acknowledgements

The authors would like to thank Stephane Litkowski, Dhruv Dhody, Tarek Saad, Zhenbin Li, Jeff Tantsura, Andrew Stone, Xuesong Geng, Ran Chen, Gyan Mishra, Weiqiang Cheng and Samuel Sidor for their review, suggestions and comments to this document.

9. References

9.1. 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>.

[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>.

[RFC8476] Tantsura, J., Chunduri, U., Aldrin, S., and P. Psenak, "Signaling Maximum SID Depth (MSD) Using OSPF", RFC 8476, DOI 10.17487/RFC8476, December 2018, <https://www.rfc-editor.org/info/rfc8476>.

[RFC8491] Tantsura, J., Chunduri, U., Aldrin, S., and L. Ginsberg, "Signaling Maximum SID Depth (MSD) Using IS-IS", RFC 8491, DOI 10.17487/RFC8491, November 2018, <https://www.rfc-editor.org/info/rfc8491>.

[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., Hendrickx, 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>.

[RFC9088] Xu, X., Kini, S., Psenak, P., Filsfils, C., Litkowski, S., and M. Bocci, "Signaling Entropy Label Capability and Entropy Readable Label Depth Using IS-IS", RFC 9088, DOI 10.17487/RFC9088, August 2021, <https://www.rfc-editor.org/info/rfc9088>.

[RFC9089] Xu, X., Kini, S., Psenak, P., Filsfils, C., Litkowski, S., and M. Bocci, "Signaling Entropy Label Capability and Entropy Readable Label Depth Using OSPF", RFC 9089, DOI 10.17487/RFC9089, August 2021, <https://www.rfc-editor.org/info/rfc9089>.

[RFC9357] Xiong, Q., "Label Switched Path (LSP) Object Flag Extension for Stateful PCE", RFC 9357, DOI 10.17487/RFC9357, February 2023, <https://www.rfc-editor.org/info/rfc9357>.

9.2. Informative References

[I-D.ietf-idr-bgp-srmpls-elp] Liu, Y. and S. Peng, "BGP Extension for SR-MPLS Entropy Label Position", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-srmpls-elp-01, 5 February 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-srmpls-elp-01>.

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

Junfeng Zhao
CAICT
Beijing
China

Email: zhaojunfeng@caict.ac.cn