Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: November 1, 2020

H. Chen
Futurewei
M. Toy
Verizon
A. Wang
China Telecom
Z. Li
China Mobile
L. Liu
Fujitsu
X. Liu
Volta Networks
April 30, 2020

SR Path Ingress Protection
draft-chen-pce-sr-ingress-protection-03

Abstract

This document describes extensions to Path Computation Element (PCE) communication Protocol (PCEP) for protecting the ingress node of a Segment Routing (SR) tunnel or path.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

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 November 1, 2020.

Table of Contents

1. Introduction

The fast protection of a transit node of a Segment Routing (SR) path or tunnel is described in [I-D.bashandy-rtgwg-segment-routing-ti-lfa] and [I-D.hu-spring-segment-routing-proxy-forwarding]. [RFC8424] presents extensions to RSVP-TE for the fast protection of the ingress node of a traffic engineering (TE) Label Switching Path (LSP). However, these documents do not discuss any protocol extensions for the fast protection of the ingress node of an SR path or tunnel.

This document fills that void and specifies protocol extensions to Path Computation Element (PCE) communication Protocol (PCEP) for the fast protection of the ingress node of an SR path or tunnel. Ingress

node and ingress, fast protection and protection as well as SR path and SR tunnel will be used exchangeably in the following sections.

2. Terminologies

The following terminologies are used in this document.

SR: Segment Routing

SRv6: SR for IPv6

SRH: Segment Routing Header

SID: Segment Identifier

CE: Customer Edge

PE: Provider Edge

LFA: Loop-Free Alternate

TI-LFA: Topology Independent LFA

TE: Traffic Engineering

BFD: Bidirectional Forwarding Detection

VPN: Virtual Private Network

L3VPN: Layer 3 VPN

FIB: Forwarding Information Base

PLR: Point of Local Repair

BGP: Border Gateway Protocol

IGP: Interior Gateway Protocol

OSPF: Open Shortest Path First

IS-IS: Intermediate System to Intermediate System

3. SR Path Ingress Protection Example

Figure 1 shows an example of protecting ingress PE1 of a SR path, which is from ingress PE1 to egress PE3.

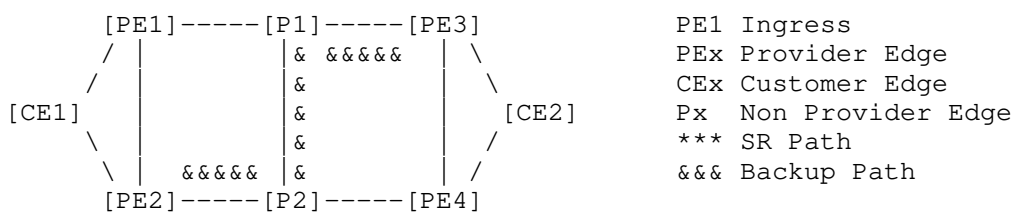


Figure 1: Protecting Ingress PE1 of SR Path

In normal operations, CE1 sends the traffic with destination PE3 to ingress PE1, which imports the traffic into the SR path.

When CE1 detects the failure of ingress PE1, it switches the traffic to backup ingress PE2, which imports the traffic from CE1 into a backup SR path. The backup path is from the backup ingress PE2 to the egress PE3. When the traffic is imported into the backup path, it is sent to the egress PE3 along the path.

4. Behavior after Ingress Failure

After failure of the ingress of an SR path happens, there are a couple of different ways to detect the failure. In each way, there may be some specific behavior for the traffic source (e.g., CE1) and the backup ingress (e.g., PE2).

In one way, the traffic source (e.g., CE1) is responsible for fast detecting the failure of the ingress (e.g., PE1) of an SR path. Fast detecting the failure means detecting the failure in a few or tens of milliseconds. The backup ingress (e.g., PE2) is ready to import the traffic from the traffic source into the backup SR path installed.

In normal operations, the source sends the traffic to the ingress of the SR path. When the source detects the failure of the ingress, it switches the traffic to the backup ingress, which delivers the traffic to the egress of the SR path via the backup SR path.

In another way, both the backup ingress and the traffic source are concurrently responsible for fast detecting the failure of the ingress of an SR path.

In normal operations, the source (e.g., CE1) sends the traffic to the ingress (e.g., PE1). It switches the traffic to the backup ingress (e.g., PE2) when it detects the failure of the ingress.

The backup ingress does not import any traffic from the source into the backup SR path in normal operations. When it detects the failure

of the ingress, it imports the traffic from the source into the backup SR path.

5. Extensions to PCEP

PCC runs on each of the edge nodes of a network normally. PCE runs on a server as a controller to communicate with PCCs. PCE and PCCs work together to support protection for the ingress of a SR path.

5.1. Capability for SR Path Ingress Protection

When a PCE and a PCC establish a PCEP session between them, they exchange their capabilities of supporting protection for the ingress node of an SR path/tunnel.

A new sub-TLV called SR_INGRESS_PROTECTION_CAPABILITY is defined. It is included in the PATH_SETUP_TYPE_CAPABILITY TLV with PST = TBD1 (suggested value 2 for backup SR path/tunnel) in the OPEN object, which is exchanged in Open messages when a PCC and a PCE establish a PCEP session between them. Its format is illustrated below.

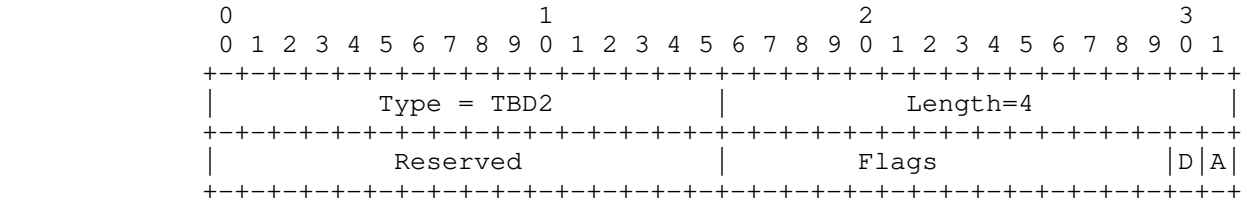


Figure 2: SR_INGRESS_PROTECTION_CAPABILITY sub-TLV

Type: TBD2 is to be assigned by IANA.

Length: 4.

Reserved: 2 octets. Must be set to zero in transmission and ignored on reception.

Flags: 2 octets. Two flags are defined.

- o D flag: A PCC sets this flag to 1 to indicate that it is able to detect its adjacent node's failure quickly.
- o A flag: A PCE sets this flag to 1 to request a PCC to let the forwarding entry for the backup SR path/tunnel be Active.

A PCC, which supports ingress protection for a SR tunnel/path, sends a PCE an Open message containing SR_INGRESS_PROTECTION_CAPABILITY

sub-TLV. This sub-TLV indicates that the PCC is capable of supporting the ingress protection for a SR tunnel/path.

A PCE, which supports ingress protection for a SR tunnel/path, sends a PCC an Open message containing SR_INGRESS_PROTECTION_CAPABILITY sub-TLV. This sub-TLV indicates that the PCE is capable of supporting the ingress protection for a SR tunnel/path.

Assume that both a PCC and a PCE support SR_PCE_CAPABILITY, that is that each of the Open messages sent by the PCC and PCE contains PATH-SETUP-TYPE-CAPABILITY TLV with a PST list containing PST=1 and a SR-PCE-CAPABILITY sub-TLV.

If a PCE receives an Open message without a SR_INGRESS_PROTECTION_CAPABILITY sub-TLV from a PCC, then the PCE MUST not send the PCC any request for ingress protection of a SR path/tunnel.

If a PCC receives an Open message without a SR_INGRESS_PROTECTION_CAPABILITY sub-TLV from a PCE, then the PCC MUST ignore any request for ingress protection of a SR path/tunnel from the PCE.

If a PCC sets D flag to zero, then the PCE SHOULD send the PCC an Open message with A flag set to one. When the PCE sends the PCC a message for initiating a backup SR path/tunnel, the PCC SHOULD let the forwarding entry for the backup SR path/tunnel be Active.

5.2. SR Path Ingress Protection

A new sub-TLV called SR_INGRESS_PROTECTION is defined. When a PCE sends a PCC a PCInitiate message for initiating a backup SR path/tunnel to protect the primary ingress node of a primary SR path/tunnel, the message contains this TLV in the RP/SRP object. Its format is illustrated below.

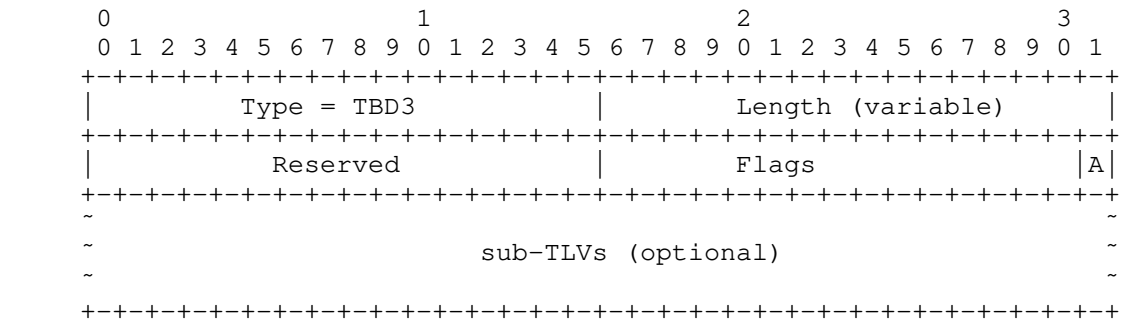


Figure 3: SR_INGRESS_PROTECTION sub-TLV

Type: TBD3 is to be assigned by IANA.

Length: Variable.

Reserved: 2 octets. Must be set to zero in transmission and ignored on reception.

Flags: 2 octets. One flag is defined.

- o A flag: A PCE sets this flag to 1 to request a PCC to let the forwarding entry for the backup SR path/tunnel be Active.

Three optional sub-TLVs are defined.

5.2.1. Traffic-Description sub-TLV

A Traffic-Description sub-TLV describes the traffic to be imported into a backup SR path/tunnel. Its format is illustrated below.

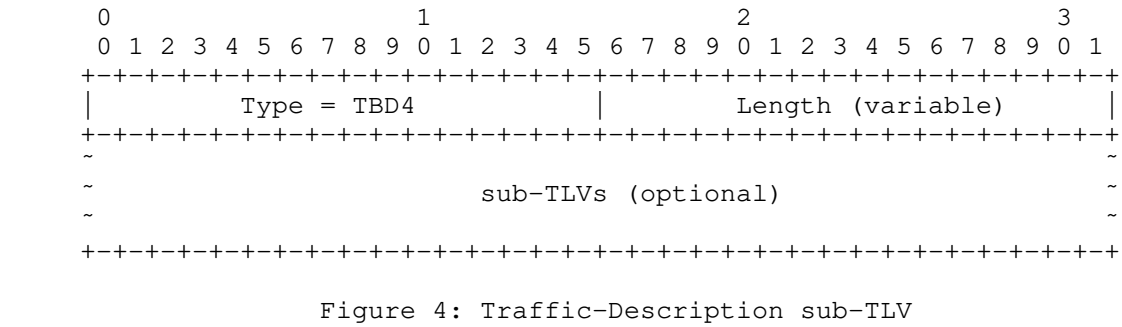


Figure 4: Traffic-Description sub-TLV

Type: TBD4 is to be assigned by IANA.

Length: Variable.

Two optional sub-TLVs are defined. One is FEC sub-TLV and the other interface sub-TLV.

A FEC sub-TLV describes the traffic to be imported into the backup SR path/tunnel. It is an IP prefix with an optional virtual network ID. It has two formats: one for IPv4 and the other for IPv6, which are illustrated below.

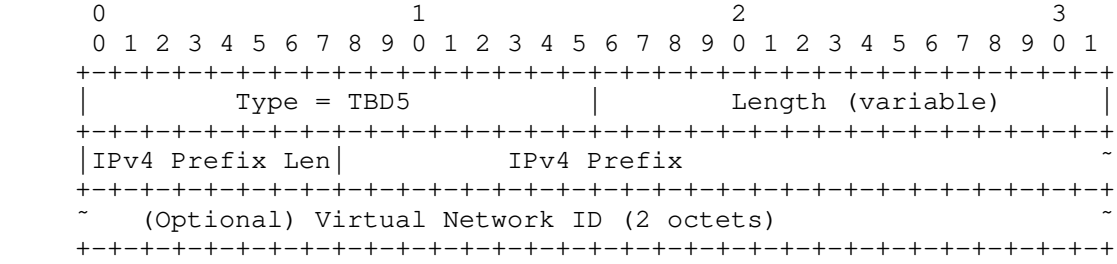


Figure 5: IPv4 FEC sub-TLV

Type: TBD5 is to be assigned by IANA.

Length: Variable.

IPv4 Prefix Len: Indicates the length of the IPv4 Prefix.

IPv4 Prefix: IPv4 Prefix rounded to octets.

Virtual Network ID: 2 octets. This is optional. It indicates the ID of a virtual network.

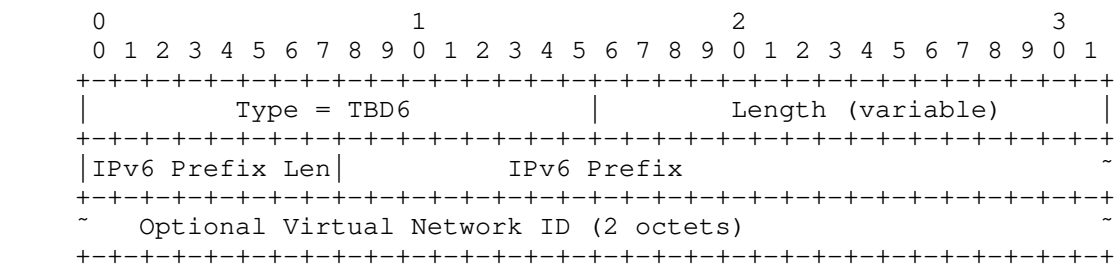


Figure 6: IPv6 FEC sub-TLV

Type: TBD6 is to be assigned by IANA.

Length: Variable.

IPv6 Prefix Len: Indicates the length of the IPv6 Prefix.

IPv6 Prefix: IPv6 Prefix rounded to octets.

Virtual Network ID: 2 octets. This is optional. It indicates the ID of a virtual network.

An Interface sub-TLV indicates the interface from which the traffic is received and imported into the backup SR path/tunnel. It has three formats: one for interface index, the other two for IPv4 and IPv6 address, which are illustrated below.

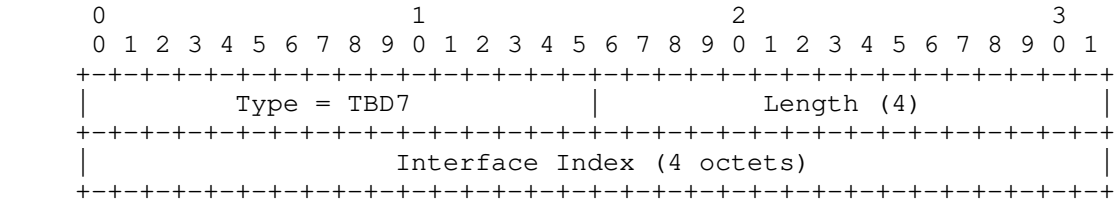


Figure 7: Interface Index sub-TLV

Type: TBD7 is to be assigned by IANA.

Length: 4.

Interface Index: 4 octets. It indicates the index of an interface.

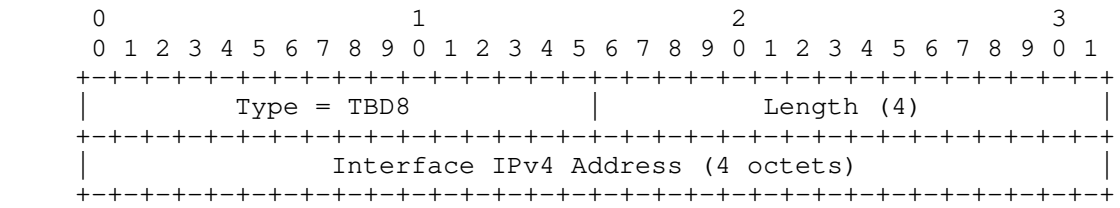


Figure 8: Interface IPv4 Address sub-TLV

Type: TBD8 is to be assigned by IANA.

Length: 4.

Interface IPv4 Address: 4 octets. It represents the IPv4 address of an interface.

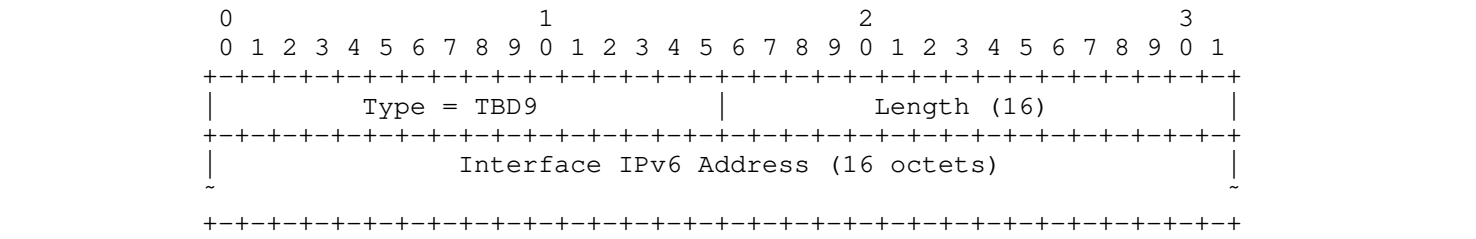


Figure 9: Interface IPv6 Address sub-TLV

Type: TBD9 is to be assigned by IANA.

Length: 16.

Interface IPv6 Address: 16 octets. It represents the IPv6 address of an interface.

5.2.2. Primary-Ingress sub-TLV

A Primary-Ingress sub-TLV indicates the IP address of the primary ingress node of a primary SR path/tunnel. It has two formats: one for primary ingress node IPv4 address and the other for primary ingress node IPv6 address, which are illustrated below.

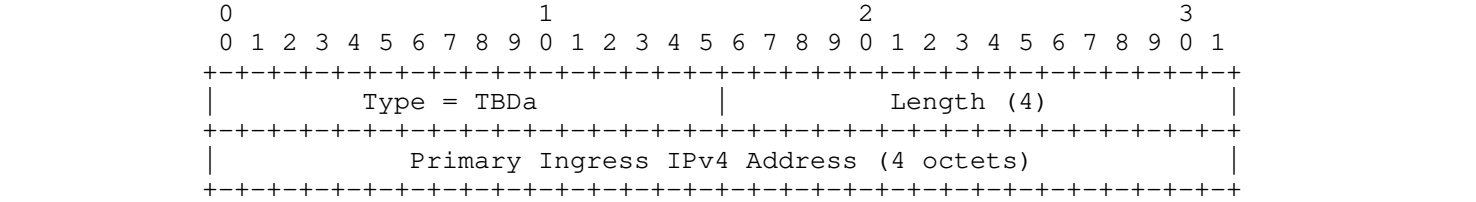


Figure 10: Primary Ingress IPv4 Address sub-TLV

Type: TBDa is to be assigned by IANA.

Length: 4.

Primary Ingress IPv4 Address: 4 octets. It represents an IPv4 host address of the primary ingress node of a SR path/tunnel.

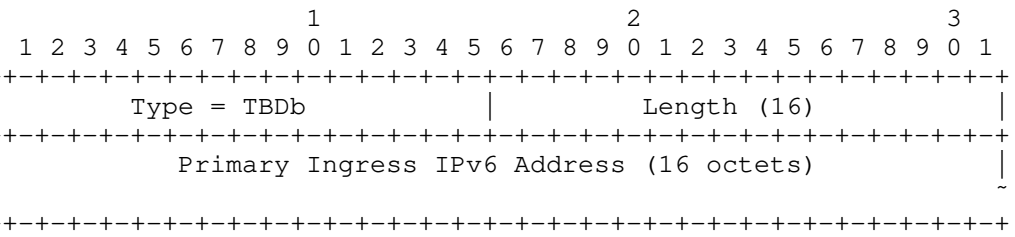


Figure 11: Primary Ingress IPv6 Address sub-TLV

Type: TBDb is to be assigned by IANA.

Length: 16.

Primary Ingress IPv6 Address: 16 octets. It represents an IPv6 host address of the primary ingress node of a SR path/tunnel.

5.2.3. Service sub-TLV

A Service sub-TLV contains a service ID or label to be added into a packet to be carried by a SR path/tunnel. It has two formats: one for the service identified by a label and the other for the service identified by a service identifier (ID) of 32 or 128 bits, which are illustrated below.

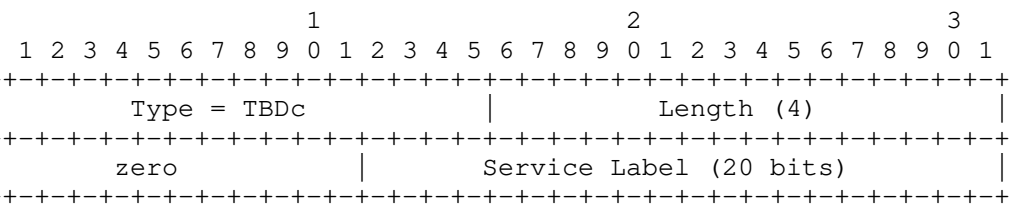


Figure 12: Service Label sub-TLV

Type: TBDc is to be assigned by IANA.

Length: 4.

Service Label: the least significant 20 bits. It represents a label of 20 bits.

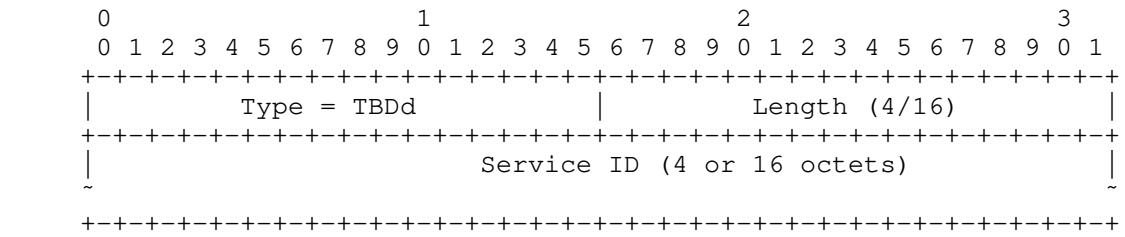


Figure 13: Service ID sub-TLV

Type: TBDd is to be assigned by IANA.

Length: 4 or 16.

Service ID: 4 or 16 octets. It represents Identifier (ID) of a service in 4 or 16 octets.

6. Security Considerations

TBD

7. Acknowledgements

The authors of this document would like to thank Dhruv Dhody for the review and comments.

8. IANA Considerations

TBD

9. References

9.1. Normative References

[I-D.bashandy-isis-srv6-extensions]
Psenak, P., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extensions to Support Routing over IPv6 Dataplane", draft-bashandy-isis-srv6-extensions-05 (work in progress), March 2019.

[I-D.hu-spring-segment-routing-proxy-forwarding]
Hu, Z., Chen, H., Yao, J., Bowers, C., and Y. Zhu, "SR-TE Path Midpoint Protection", draft-hu-spring-segment-routing-proxy-forwarding-07 (work in progress), January 2020.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-25 (work in progress), May 2019.

[I-D.ietf-ospf-segment-routing-extensions]
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", draft-ietf-ospf-segment-routing-extensions-27 (work in progress), December 2018.

[I-D.li-ospf-ospfv3-srv6-extensions]
Li, Z., Hu, Z., Cheng, D., Talaulikar, K., and P. Psenak, "OSPFv3 Extensions for SRv6", draft-li-ospf-ospfv3-srv6-extensions-07 (work in progress), November 2019.

[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>.

[RFC7356] Ginsberg, L., Previdi, S., and Y. Yang, "IS-IS Flooding Scope Link State PDUs (LSPs)", RFC 7356, DOI 10.17487/RFC7356, September 2014, <https://www.rfc-editor.org/info/rfc7356>.

[RFC8424] Chen, H., Ed. and R. Torvi, Ed., "Extensions to RSVP-TE for Label Switched Path (LSP) Ingress Fast Reroute (FRR) Protection", RFC 8424, DOI 10.17487/RFC8424, August 2018, <https://www.rfc-editor.org/info/rfc8424>.

9.2. Informative References

[I-D.bashandy-rtgwg-segment-routing-ti-lfa]
Bashandy, A., Filsfils, C., Decraene, B., Litkowski, S., Francois, P., daniel.voyer@bell.ca, d., Clad, F., and P. Camarillo, "Topology Independent Fast Reroute using Segment Routing", draft-bashandy-rtgwg-segment-routing-ti-lfa-05 (work in progress), October 2018.

[I-D.hegde-spring-node-protection-for-sr-te-paths]
Hegde, S., Bowers, C., Litkowski, S., Xu, X., and F. Xu, "Node Protection for SR-TE Paths", draft-hegde-spring-node-protection-for-sr-te-paths-05 (work in progress), July 2019.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-06 (work in progress), December 2019.

[I-D.sivabalan-pce-binding-label-sid]
Sivabalan, S., Filsfils, C., Tantsura, J., Hardwick, J., Previdi, S., and C. Li, "Carrying Binding Label/Segment-ID in PCE-based Networks.", draft-sivabalan-pce-binding-label-sid-07 (work in progress), July 2019.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <https://www.rfc-editor.org/info/rfc5462>.

Authors' Addresses

Huaimo Chen
Futurewei
Boston, MA
USA

Email: Huaimo.chen@futurewei.com

Mehmet Toy
Verizon
USA

Email: mehmet.toy@verizon.com

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

Email: wangaj3@chinatelecom.cn

Zhenqiang Li
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing 100053
China

Email: lizhengqiang@chinamobile.com

Lei Liu
Fujitsu
USA

Email: liulei.kddi@gmail.com

Xufeng Liu
Volta Networks
McLean, VA
USA

Email: xufeng.liu.ietf@gmail.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 4, 2021

H. Chen
Futurewei
M. Toy
Verizon
A. Wang
China Telecom
Z. Li
China Mobile
L. Liu
Fujitsu
X. Liu
Volta Networks

October 31, 2020

SR Path Ingress Protection
draft-chen-pce-sr-ingress-protection-04

Abstract

This document describes extensions to Path Computation Element (PCE) communication Protocol (PCEP) for protecting the ingress node of a Segment Routing (SR) tunnel or path.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

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 May 4, 2021.

Table of Contents

1. Introduction

The fast protection of a transit node of a Segment Routing (SR) path or tunnel is described in [I-D.bashandy-rtgwg-segment-routing-ti-lfa] and [I-D.hu-spring-segment-routing-proxy-forwarding]. [RFC8424] presents extensions to RSVP-TE for the fast protection of the ingress node of a traffic engineering (TE) Label Switching Path (LSP). However, these documents do not discuss any protocol extensions for the fast protection of the ingress node of an SR path or tunnel.

This document fills that void and specifies protocol extensions to Path Computation Element (PCE) communication Protocol (PCEP) for the fast protection of the ingress node of an SR path or tunnel. Ingress

node and ingress, fast protection and protection as well as SR path and SR tunnel will be used exchangeably in the following sections.

2. Terminologies

The following terminologies are used in this document.

SR: Segment Routing

SRv6: SR for IPv6

SRH: Segment Routing Header

SID: Segment Identifier

CE: Customer Edge

PE: Provider Edge

LFA: Loop-Free Alternate

TI-LFA: Topology Independent LFA

TE: Traffic Engineering

BFD: Bidirectional Forwarding Detection

VPN: Virtual Private Network

L3VPN: Layer 3 VPN

FIB: Forwarding Information Base

PLR: Point of Local Repair

BGP: Border Gateway Protocol

IGP: Interior Gateway Protocol

OSPF: Open Shortest Path First

IS-IS: Intermediate System to Intermediate System

3. SR Path Ingress Protection Example

Figure 1 shows an example of protecting ingress PE1 of a SR path, which is from ingress PE1 to egress PE3.

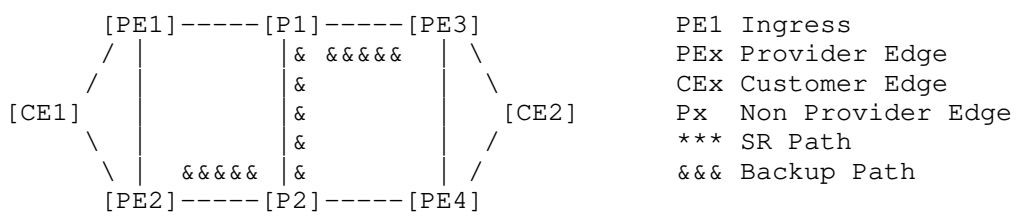


Figure 1: Protecting Ingress PE1 of SR Path

In normal operations, CE1 sends the traffic with destination PE3 to ingress PE1, which imports the traffic into the SR path.

When CE1 detects the failure of ingress PE1, it switches the traffic to backup ingress PE2, which imports the traffic from CE1 into a backup SR path. The backup path is from the backup ingress PE2 to the egress PE3. When the traffic is imported into the backup path, it is sent to the egress PE3 along the path.

4. Behavior after Ingress Failure

After failure of the ingress of an SR path happens, there are a couple of different ways to detect the failure. In each way, there may be some specific behavior for the traffic source (e.g., CE1) and the backup ingress (e.g., PE2).

In one way, the traffic source (e.g., CE1) is responsible for fast detecting the failure of the ingress (e.g., PE1) of an SR path. Fast detecting the failure means detecting the failure in a few or tens of milliseconds. The backup ingress (e.g., PE2) is ready to import the traffic from the traffic source into the backup SR path installed.

In normal operations, the source sends the traffic to the ingress of the SR path. When the source detects the failure of the ingress, it switches the traffic to the backup ingress, which delivers the traffic to the egress of the SR path via the backup SR path.

In another way, both the backup ingress and the traffic source are concurrently responsible for fast detecting the failure of the ingress of an SR path.

In normal operations, the source (e.g., CE1) sends the traffic to the ingress (e.g., PE1). It switches the traffic to the backup ingress (e.g., PE2) when it detects the failure of the ingress.

The backup ingress does not import any traffic from the source into the backup SR path in normal operations. When it detects the failure

of the ingress, it imports the traffic from the source into the backup SR path.

5. Extensions to PCEP

PCC runs on each of the edge nodes of a network normally. PCE runs on a server as a controller to communicate with PCCs. PCE and PCCs work together to support protection for the ingress of a SR path.

5.1. Capability for SR Path Ingress Protection

When a PCE and a PCC establish a PCEP session between them, they exchange their capabilities of supporting protection for the ingress node of an SR path/tunnel.

A new sub-TLV called SR_INGRESS_PROTECTION_CAPABILITY is defined. It is included in the PATH_SETUP_TYPE_CAPABILITY TLV with PST = TBD1 (suggested value 2 for backup SR path/tunnel) in the OPEN object, which is exchanged in Open messages when a PCC and a PCE establish a PCEP session between them. Its format is illustrated below.

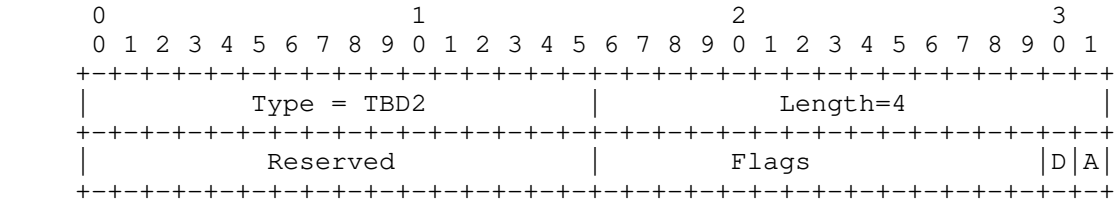


Figure 2: SR_INGRESS_PROTECTION_CAPABILITY sub-TLV

Type: TBD2 is to be assigned by IANA.

Length: 4.

Reserved: 2 octets. Must be set to zero in transmission and ignored on reception.

Flags: 2 octets. Two flags are defined.

- D flag: A PCC sets this flag to 1 to indicate that it is able to detect its adjacent node's failure quickly.

- A flag: A PCE sets this flag to 1 to request a PCC to let the forwarding entry for the backup SR path/tunnel be Active.

A PCC, which supports ingress protection for a SR tunnel/path, sends a PCE an Open message containing SR_INGRESS_PROTECTION_CAPABILITY

sub-TLV. This sub-TLV indicates that the PCC is capable of supporting the ingress protection for a SR tunnel/path.

A PCE, which supports ingress protection for a SR tunnel/path, sends a PCC an Open message containing SR_INGRESS_PROTECTION_CAPABILITY sub-TLV. This sub-TLV indicates that the PCE is capable of supporting the ingress protection for a SR tunnel/path.

Assume that both a PCC and a PCE support SR_PCE_CAPABILITY, that is that each of the Open messages sent by the PCC and PCE contains PATH-SETUP-TYPE-CAPABILITY TLV with a PST list containing PST=1 and a SR-PCE-CAPABILITY sub-TLV.

If a PCE receives an Open message without a SR_INGRESS_PROTECTION_CAPABILITY sub-TLV from a PCC, then the PCE MUST not send the PCC any request for ingress protection of a SR path/tunnel.

If a PCC receives an Open message without a SR_INGRESS_PROTECTION_CAPABILITY sub-TLV from a PCE, then the PCC MUST ignore any request for ingress protection of a SR path/tunnel from the PCE.

If a PCC sets D flag to zero, then the PCE SHOULD send the PCC an Open message with A flag set to one. When the PCE sends the PCC a message for initiating a backup SR path/tunnel, the PCC SHOULD let the forwarding entry for the backup SR path/tunnel be Active.

5.2. SR Path Ingress Protection

A new sub-TLV called SR_INGRESS_PROTECTION is defined. When a PCE sends a PCC a PCInitiate message for initiating a backup SR path/tunnel to protect the primary ingress node of a primary SR path/tunnel, the message contains this TLV in the RP/SRP object. Its format is illustrated below.

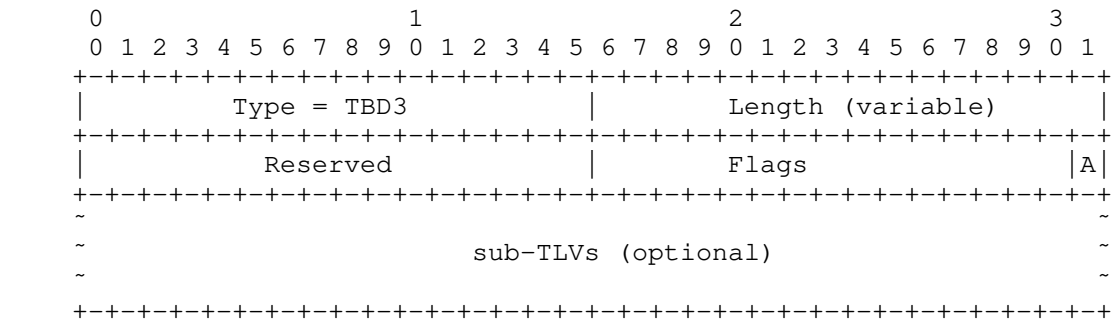


Figure 3: SR_INGRESS_PROTECTION sub-TLV

Type: TBD3 is to be assigned by IANA.

Length: Variable.

Reserved: 2 octets. Must be set to zero in transmission and ignored on reception.

Flags: 2 octets. One flag is defined.

- o A flag: A PCE sets this flag to 1 to request a PCC to let the forwarding entry for the backup SR path/tunnel be Active.

Three optional sub-TLVs are defined.

5.2.1. Traffic-Description sub-TLV

A Traffic-Description sub-TLV describes the traffic to be imported into a backup SR path/tunnel. Its format is illustrated below.

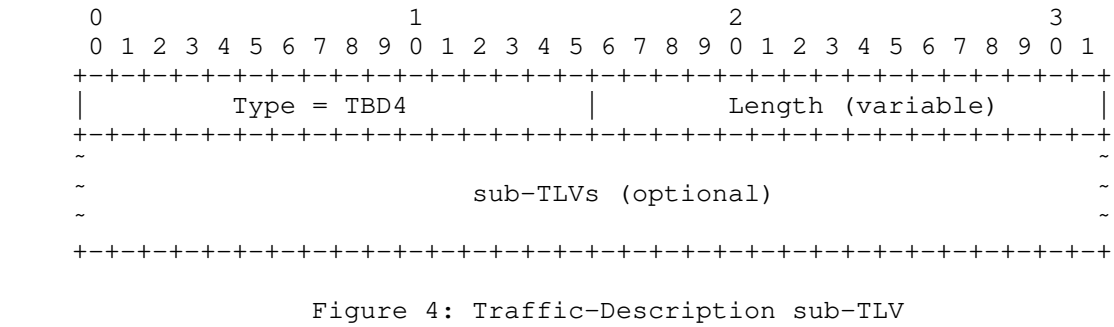


Figure 4: Traffic-Description sub-TLV

Type: TBD4 is to be assigned by IANA.

Length: Variable.

Two optional sub-TLVs are defined. One is FEC sub-TLV and the other interface sub-TLV.

A FEC sub-TLV describes the traffic to be imported into the backup SR path/tunnel. It is an IP prefix with an optional virtual network ID. It has two formats: one for IPv4 and the other for IPv6, which are illustrated below.

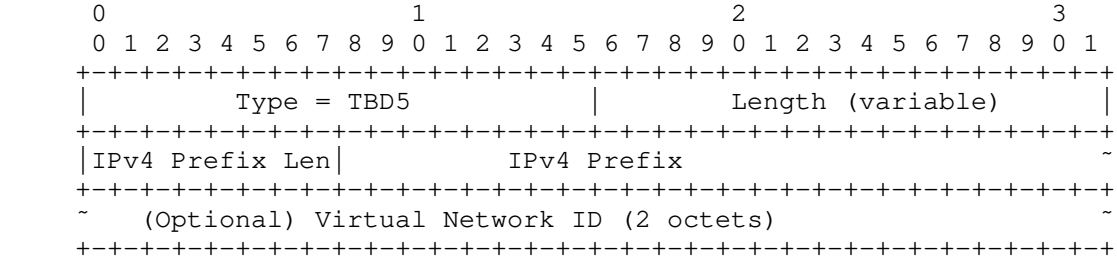


Figure 5: IPv4 FEC sub-TLV

Type: TBD5 is to be assigned by IANA.

Length: Variable.

IPv4 Prefix Len: Indicates the length of the IPv4 Prefix.

IPv4 Prefix: IPv4 Prefix rounded to octets.

Virtual Network ID: 2 octets. This is optional. It indicates the ID of a virtual network.

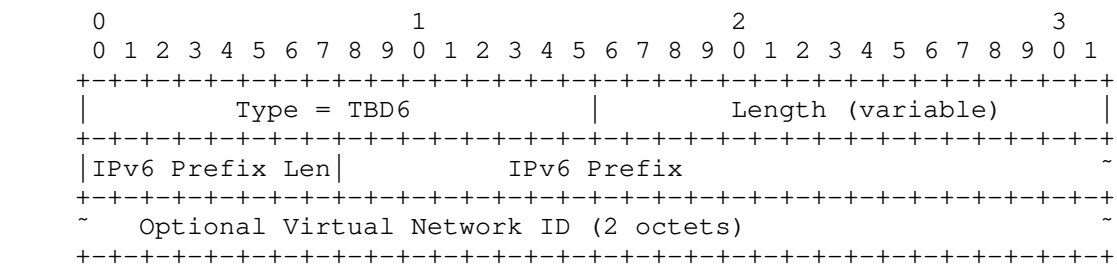


Figure 6: IPv6 FEC sub-TLV

Type: TBD6 is to be assigned by IANA.

Length: Variable.

IPv6 Prefix Len: Indicates the length of the IPv6 Prefix.

IPv6 Prefix: IPv6 Prefix rounded to octets.

Virtual Network ID: 2 octets. This is optional. It indicates the ID of a virtual network.

An Interface sub-TLV indicates the interface from which the traffic is received and imported into the backup SR path/tunnel. It has three formats: one for interface index, the other two for IPv4 and IPv6 address, which are illustrated below.

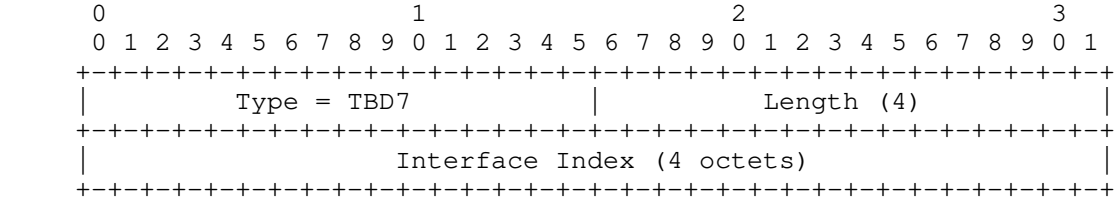


Figure 7: Interface Index sub-TLV

Type: TBD7 is to be assigned by IANA.

Length: 4.

Interface Index: 4 octets. It indicates the index of an interface.

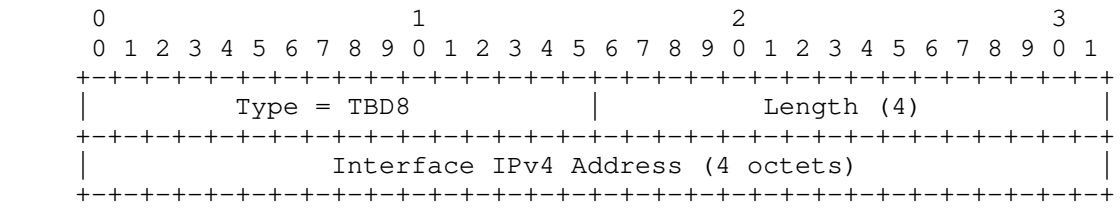


Figure 8: Interface IPv4 Address sub-TLV

Type: TBD8 is to be assigned by IANA.

Length: 4.

Interface IPv4 Address: 4 octets. It represents the IPv4 address of an interface.

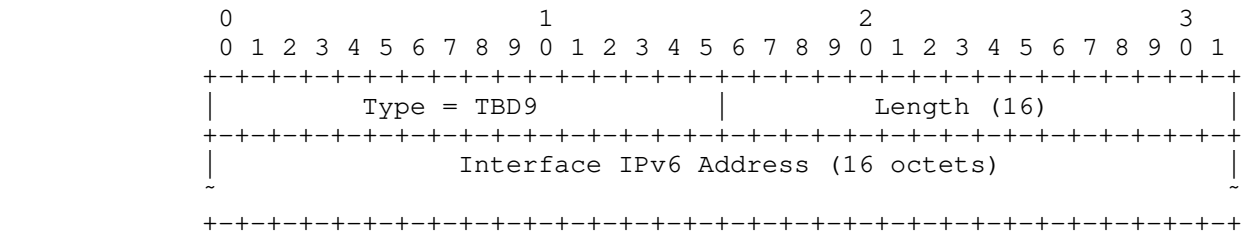


Figure 9: Interface IPv6 Address sub-TLV

Type: TBD9 is to be assigned by IANA.

Length: 16.

Interface IPv6 Address: 16 octets. It represents the IPv6 address of an interface.

5.2.2. Primary-Ingress sub-TLV

A Primary-Ingress sub-TLV indicates the IP address of the primary ingress node of a primary SR path/tunnel. It has two formats: one for primary ingress node IPv4 address and the other for primary ingress node IPv6 address, which are illustrated below.

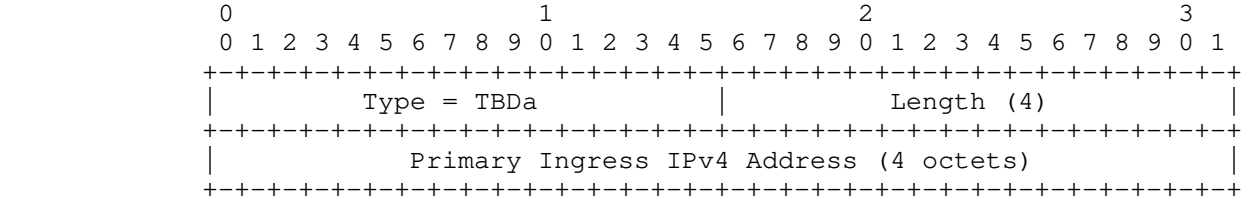


Figure 10: Primary Ingress IPv4 Address sub-TLV

Type: TBDa is to be assigned by IANA.

Length: 4.

Primary Ingress IPv4 Address: 4 octets. It represents an IPv4 host address of the primary ingress node of a SR path/tunnel.

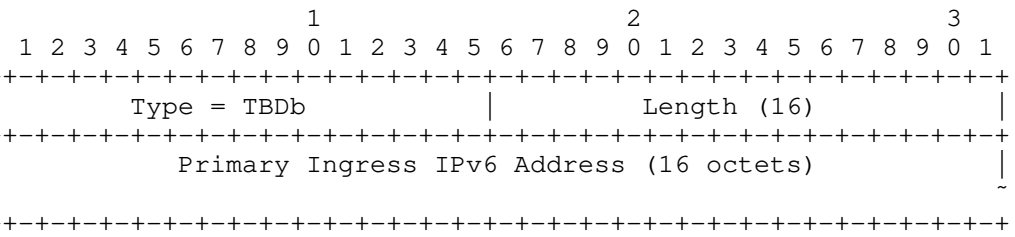


Figure 11: Primary Ingress IPv6 Address sub-TLV

Type: TBDb is to be assigned by IANA.

Length: 16.

Primary Ingress IPv6 Address: 16 octets. It represents an IPv6 host address of the primary ingress node of a SR path/tunnel.

5.2.3. Service sub-TLV

A Service sub-TLV contains a service ID or label to be added into a packet to be carried by a SR path/tunnel. It has two formats: one for the service identified by a label and the other for the service identified by a service identifier (ID) of 32 or 128 bits, which are illustrated below.

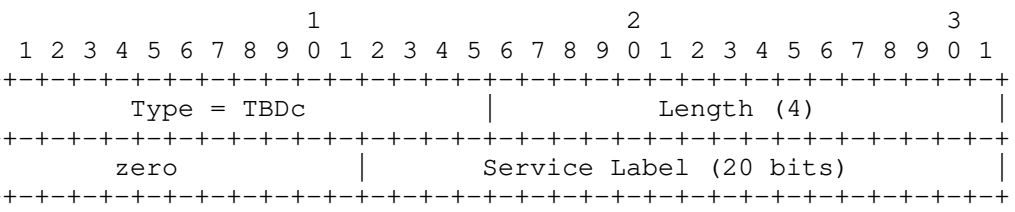


Figure 12: Service Label sub-TLV

Type: TBDc is to be assigned by IANA.

Length: 4.

Service Label: the least significant 20 bits. It represents a label of 20 bits.

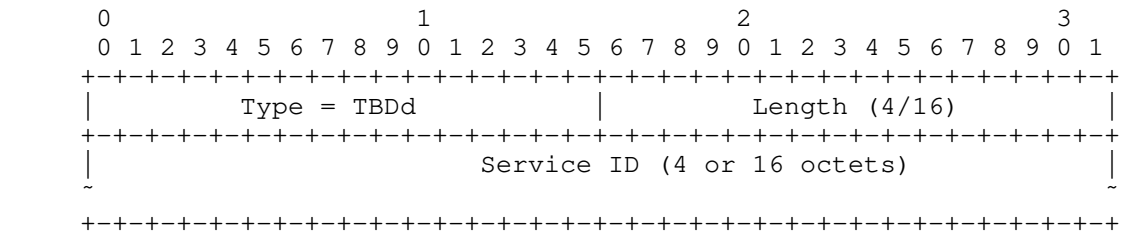


Figure 13: Service ID sub-TLV

Type: TBDd is to be assigned by IANA.

Length: 4 or 16.

Service ID: 4 or 16 octets. It represents Identifier (ID) of a service in 4 or 16 octets.

6. Security Considerations

TBD

7. Acknowledgements

The authors of this document would like to thank Dhruv Dhody for the review and comments.

8. IANA Considerations

TBD

9. References

9.1. Normative References

[I-D.bashandy-isis-srv6-extensions]
Psenak, P., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extensions to Support Routing over IPv6 Dataplane", draft-bashandy-isis-srv6-extensions-05 (work in progress), March 2019.

[I-D.hu-spring-segment-routing-proxy-forwarding]
Hu, Z., Chen, H., Yao, J., Bowers, C., and Y. Zhu, "SR-TE Path Midpoint Protection", draft-hu-spring-segment-routing-proxy-forwarding-12 (work in progress), October 2020.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-25 (work in progress), May 2019.

[I-D.ietf-ospf-segment-routing-extensions]
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", draft-ietf-ospf-segment-routing-extensions-27 (work in progress), December 2018.

[I-D.li-ospf-ospfv3-srv6-extensions]
Li, Z., Hu, Z., Cheng, D., Talaulikar, K., and P. Psenak, "OSPFv3 Extensions for SRv6", draft-li-ospf-ospfv3-srv6-extensions-07 (work in progress), November 2019.

[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>.

[RFC7356] Ginsberg, L., Previdi, S., and Y. Yang, "IS-IS Flooding Scope Link State PDUs (LSPs)", RFC 7356, DOI 10.17487/RFC7356, September 2014, <https://www.rfc-editor.org/info/rfc7356>.

[RFC8424] Chen, H., Ed. and R. Torvi, Ed., "Extensions to RSVP-TE for Label Switched Path (LSP) Ingress Fast Reroute (FRR) Protection", RFC 8424, DOI 10.17487/RFC8424, August 2018, <https://www.rfc-editor.org/info/rfc8424>.

9.2. Informative References

[I-D.bashandy-rtgwg-segment-routing-ti-lfa]
Bashandy, A., Filsfils, C., Decraene, B., Litkowski, S., Francois, P., daniel.voyer@bell.ca, d., Clad, F., and P. Camarillo, "Topology Independent Fast Reroute using Segment Routing", draft-bashandy-rtgwg-segment-routing-ti-lfa-05 (work in progress), October 2018.

[I-D.hegde-spring-node-protection-for-sr-te-paths]
Hegde, S., Bowers, C., Litkowski, S., Xu, X., and F. Xu, "Node Protection for SR-TE Paths", draft-hegde-spring-node-protection-for-sr-te-paths-07 (work in progress), July 2020.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-08 (work in progress), July 2020.

[I-D.sivabalan-pce-binding-label-sid]
Sivabalan, S., Filsfils, C., Tantsura, J., Hardwick, J., Previdi, S., and C. Li, "Carrying Binding Label/Segment-ID in PCE-based Networks.", draft-sivabalan-pce-binding-label-sid-07 (work in progress), July 2019.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic Class" Field", RFC 5462, DOI 10.17487/RFC5462, February 2009, <https://www.rfc-editor.org/info/rfc5462>.

Authors' Addresses

Huaimo Chen
Futurewei
Boston, MA
USA

Email: Huaimo.chen@futurewei.com

Mehmet Toy
Verizon
USA

Email: mehmet.toy@verizon.com

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

Email: wangaj3@chinatelecom.cn

Zhenqiang Li
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, 100053
China

Email: lizhengqiang@chinamobile.com

Lei Liu
Fujitsu

USA

Email: liulei.kddi@gmail.com

Xufeng Liu
Volta Networks

McLean, VA
USA

Email: xufeng.liu.ietf@gmail.com

PCE
Internet-Draft
Intended status: Standards Track
Expires: January 11, 2021

H. Chen
China Telecom
H. Yuan
UnionPay
T. Zhou
W. Li
G. Fioccola
Y. Wang
Huawei

July 10, 2020

PCEP SR Policy Extensions to Enable IFIT
draft-chen-pce-sr-policy-ifit-02

Abstract

Segment Routing (SR) policy is a set of candidate SR paths consisting of one or more segment lists and necessary path attributes. It enables instantiation of an ordered list of segments with a specific intent for traffic steering. In-situ Flow Information Telemetry (IFIT) refers to network OAM applications that apply dataplane on-path telemetry techniques. This document defines extensions to PCEP to distribute SR policies carrying IFIT information. So that IFIT behavior can be enabled automatically when the SR policy is applied.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

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 11, 2021.

Table of Contents

1. Introduction

Segment Routing (SR) policy [I-D.ietf-spring-segment-routing-policy] is a set of candidate SR paths consisting of one or more segment lists and necessary path attributes. It enables instantiation of an ordered list of segments with a specific intent for traffic steering.

In-situ Flow Information Telemetry (IFIT) refers to network OAM applications that apply dataplane on-path telemetry techniques, including In-situ OAM (IOAM) [I-D.ietf-ippm-ioam-data] and Alternate Marking [RFC8321]. It can provide flow information on the entire forwarding path on a per- packet basis in real time.

An automatic network requires the Service Level Agreement (SLA) monitoring on the deployed service. So that the system can quickly detect the SLA violation or the performance degradation, hence to change the service deployment. The SR policy native IFIT can facilitate the closed loop control, and enable the automation of SR service.

This document defines extensions to PCEP to distribute SR policies carrying IFIT information. So that IFIT behavior can be enabled automatically when the SR policy is applied.

This PCEP extension allows to signal the IFIT capabilities together with the SR-policy. In this way IFIT methods are automatically activated and running. The flexibility and dynamicity of the IFIT applications are given by the use of additional functions on the controller and on the network nodes, but this is out of scope here.

It is to be noted the companion document [I-D.qin-idr-sr-policy-ifit] that proposes the BGP extension to enable IFIT applications for SR policy.

2. IFIT Attributes in SR Policy

SR Policy Association Group (SRPAG) is defined in [I-D.ietf-pce-segment-routing-policy-cp] to extend PCEP to support association among candidate paths of a given SR policy. SR Policy Identifiers TLV, SR Policy Name TLV, SR Policy Candidate Path Identifiers TLV, and SR Policy Candidate Path Preference TLV are introduced to construct the SR policy structure.

This document is to add IFIT attribute TLVs to the SRPAG. The following sections will describe the requirement and usage of different IFIT modes, and define the corresponding TLV encoding in PCEP.

Note that the IFIT attributes here described can also be generalized and included as TLVs for other Association Groups. In this regard RFC 8697 [RFC8697] defines the generic mechanism to associate sets of LSPs and a set of attributes, for example IFIT.

3. SR Policy for IOAM

In-situ Operations, Administration, and Maintenance (IOAM) [I-D.ietf-ippm-ioam-data] records operational and telemetry information in the packet while the packet traverses a path between two points in the network. In terms of the classification given in RFC 7799 [RFC7799] IOAM could be categorized as Hybrid Type 1. IOAM

mechanisms can be leveraged where active OAM do not apply or do not offer the desired results.

When SR policy enables the IOAM, the IOAM header will be inserted into every packet of the traffic that is steered into the SR paths.

This document aims to define the control plane. While a relevant document for the data plane is [I-D.ietf-ippm-ioam-ipv6-options] for Segment Routing over IPv6 data plane (SRv6).

3.1. IOAM Pre-allocated Trace Option TLV

The IOAM tracing data is expected to be collected at every node that a packet traverses to ensure visibility into the entire path a packet takes within an IOAM domain. The preallocated tracing option will create pre-allocated space for each node to populate its information.

The format of IOAM pre-allocated trace option TLV is defined as follows:

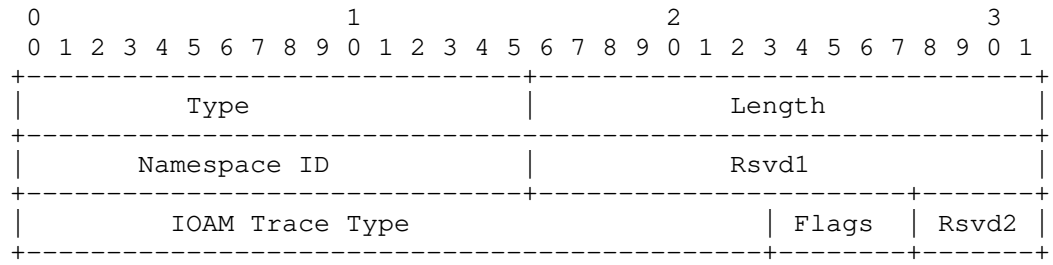


Fig. 1 IOAM Pre-allocated Trace Option TLV

Where:

Type: to be assigned by IANA.

Length: the total length of the value field not including Type and Length fields.

Namespace ID: A 16-bit identifier of an IOAM-Namespace. The definition is the same as described in section 4.4 of [I-D.ietf-ippm-ioam-data].

IOAM Trace Type: A 24-bit identifier which specifies which data types are used in the node data list. The definition is the same as described in section 4.4 of [I-D.ietf-ippm-ioam-data].

Flags: A 4-bit field. The definition is the same as described in [I-D.ietf-ippm-ioam-flags] and section 4.4 of [I-D.ietf-ippm-ioam-data].

Rsvd1: A 16-bit field reserved for further usage. It MUST be zero.

Rsvd2: A 4-bit field reserved for further usage. It MUST be zero.

3.2. IOAM Incremental Trace Option TLV

The incremental tracing option contains a variable node data fields where each node allocates and pushes its node data immediately following the option header.

The format of IOAM incremental trace option TLV is defined as follows:

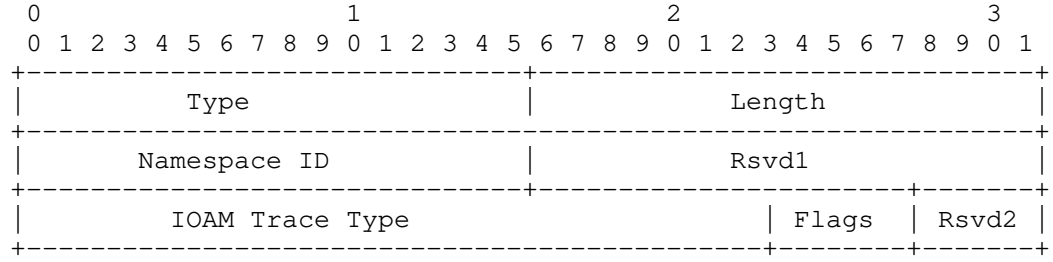


Fig. 2 IOAM Incremental Trace Option TLV

Where:

Type: to be assigned by IANA.

Length: the total length of the value field not including Type and Length fields.

All the other fields definition is the same as the pre-allocated trace option TLV in section 4.1.

3.3. IOAM Directly Export Option TLV

IOAM directly export option is used as a trigger for IOAM data to be directly exported to a collector without being pushed into in-flight data packets.

The format of IOAM directly export option TLV is defined as follows:

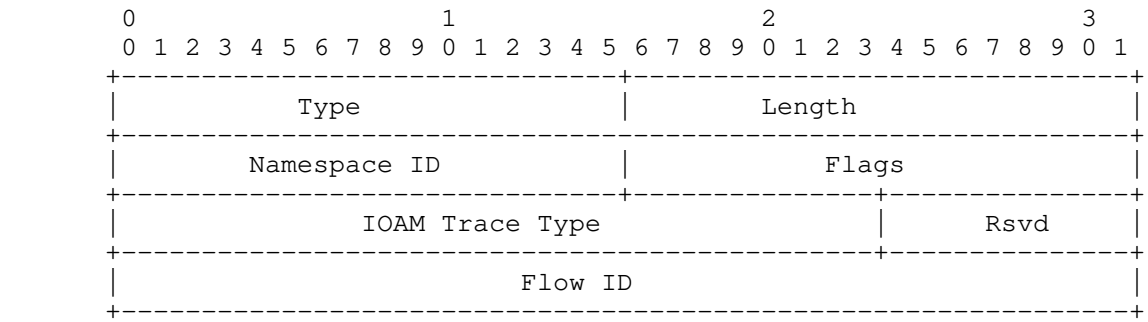


Fig. 3 IOAM Directly Export Option TLV

Where:

Type: to be assigned by IANA.

Length: the total length of the value field not including Type and Length fields.

Namespace ID: A 16-bit identifier of an IOAM-Namespace. The definition is the same as described in section 4.4 of [I-D.ietf-ippm-ioam-data].

IOAM Trace Type: A 24-bit identifier which specifies which data types are used in the node data list. The definition is the same as described in section 4.4 of [I-D.ietf-ippm-ioam-data].

Flags: A 16-bit field. The definition is the same as described in section 3.2 of [I-D.ietf-ippm-ioam-direct-export].

Flow ID: A 32-bit flow identifier. The definition is the same as described in section 3.2 of [I-D.ietf-ippm-ioam-direct-export].

Rsvd: A 4-bit field reserved for further usage. It MUST be zero.

3.4. IOAM Edge-to-Edge Option TLV

The IOAM edge to edge option is to carry data that is added by the IOAM encapsulating node and interpreted by IOAM decapsulating node.

The format of IOAM edge-to-edge option TLV is defined as follows:

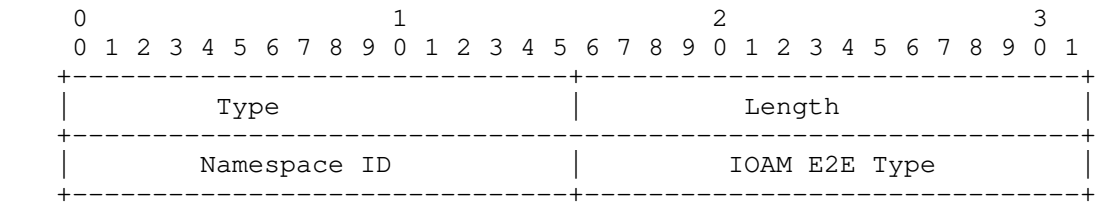


Fig. 4 IOAM Edge-to-Edge Option TLV

Where:

Type: to be assigned by IANA.

Length: the total length of the value field not including Type and Length fields.

Namespace ID: A 16-bit identifier of an IOAM-Namespace. The definition is the same as described in section 4.6 of [I-D.ietf-ippm-ioam-data].

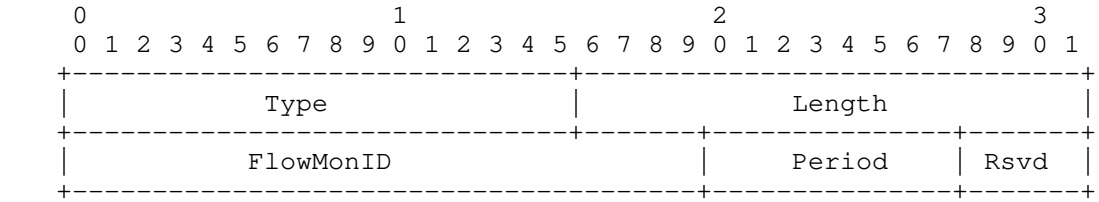
IOAM E2E Type: A 16-bit identifier which specifies which data types are used in the E2E option data. The definition is the same as described in section 4.6 of [I-D.ietf-ippm-ioam-data].

4. SR Policy for Enhanced Alternate Marking

The Alternate Marking [RFC8321]technique is an hybrid performance measurement method, per RFC 7799 [RFC7799] classification of measurement methods. Because this method is based on marking consecutive batches of packets. It can be used to measure packet loss, latency, and jitter on live traffic.

This document aims to define the control plane. While a relevant document for the data plane is [I-D.ietf-6man-ipv6-alt-mark] for Segment Routing over IPv6 data plane (SRv6).

The format of EAM TLV is defined as follows:



Where:

Type: to be assigned by IANA.

Length: the total length of the value field not including Type and Length fields.

FlowMonID: A 20-bit identifier to uniquely identify a monitored flow within the measurement domain. The definition is the same as described in section 5.3 of [I-D.ietf-6man-ipv6-alt-mark].

Period: Time interval between two alternate marking period. The unit is second.

Rsvd: A 4-bit field reserved for further usage. It MUST be zero.

5. Examples

5.1. PCE Initiated SR Policy

The interactions between the PCE and PCC is the same as described in [I-D.ietf-pce-segment-routing-policy-cp]. The only change is to take the additional optional IFIT TLVs within the SRPAG object.

PCE sends PCInitiate message, containing the SRPAG Association object. The Association Source is set to the IP address of the PCC and the Association ID is set to 0xFFFF.

PCC uses the color, endpoint, preference and IFIT option from the SRPAG object to create a new candidate path. If no SR policy exists to hold the candidate path, then a new SR policy is created to hold the new candidate-path. The Originator of the candidate path is set to be the address of the PCE that is sending the PCInitiate message.

PCC sends a PCRpt message back to the PCE to report the newly created Candidate Path. The PCRpt message contains the SRPAG Association object. The Association Source is set to the IP address of the PCC and the Association ID is set to a number that PCC locally chose to represent the SR Policy.

6. IANA Considerations

This document defines new IFIT TLVs for carrying additional information about SR policy and SR candidate paths. IANA is requested to make the assignment of a new value for the existing "PCEP TLV Type Indicators" registry as follows:

Codepoint	Description	Reference
TBD1	IOAM Pre-allocated Trace Option TLV	This document
TBD2	IOAM Incremental Trace Option TLV	This document
TBD3	IOAM Directly Export Option TLV	This document
TBD4	IOAM Edge-to-Edge Option TLV	This document
TBD5	Enhanced Alternate Marking TLV	This document

7. Security Considerations

TBD.

8. Acknowledgements

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>.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <https://www.rfc-editor.org/info/rfc7799>.

[RFC8321] Fioccola, G., Ed., Capello, A., Cociglio, M., Castaldelli, L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi, "Alternate-Marking Method for Passive and Hybrid Performance Monitoring", RFC 8321, DOI 10.17487/RFC8321, January 2018, <https://www.rfc-editor.org/info/rfc8321>.

[RFC8697] Minei, I., Crabbe, E., Sivabalan, S., Ananthakrishnan, H., Dhody, D., and Y. Tanaka, "Path Computation Element Communication Protocol (PCEP) Extensions for Establishing Relationships between Sets of Label Switched Paths (LSPs)", RFC 8697, DOI 10.17487/RFC8697, January 2020, <https://www.rfc-editor.org/info/rfc8697>.

9.2. Informative References

[I-D.ietf-6man-ipv6-alt-mark]
Fioccola, G., Zhou, T., Cociglio, M., Qin, F., and R. Pang, "IPv6 Application of the Alternate Marking Method", draft-ietf-6man-ipv6-alt-mark-01 (work in progress), June 2020.

[I-D.ietf-ippm-ioam-data]
Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., remy@barefootnetworks.com, r., daniel.bernier@bell.ca, d., and J. Lemon, "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data-09 (work in progress), March 2020.

[I-D.ietf-ippm-ioam-direct-export]
Song, H., Gafni, B., Zhou, T., Li, Z., Brockners, F., Bhandari, S., Sivakolundu, R., and T. Mizrahi, "In-situ OAM Direct Exporting", draft-ietf-ippm-ioam-direct-export-00 (work in progress), February 2020.

[I-D.ietf-ippm-ioam-flags]
Mizrahi, T., Brockners, F., Bhandari, S., Sivakolundu, R., Pignataro, C., Kfir, A., Gafni, B., Spiegel, M., and J. Lemon, "In-situ OAM Flags", draft-ietf-ippm-ioam-flags-01 (work in progress), January 2020.

[I-D.ietf-ippm-ioam-ipv6-options]
Bhandari, S., Brockners, F., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Kfir, A., Gafni, B., Lapukhov, P., Spiegel, M., Krishnan, S., and R. Asati, "In-situ OAM IPv6 Options", draft-ietf-ippm-ioam-ipv6-options-01 (work in progress), March 2020.

[I-D.ietf-pce-segment-routing-policy-cp]
Koldychev, M., Sivabalan, S., Barth, C., Peng, S., and H. Bidgoli, "PCEP extension to support Segment Routing Policy Candidate Paths", draft-ietf-pce-segment-routing-policy-cp-00 (work in progress), June 2020.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-08 (work in progress), July 2020.

[I-D.qin-idr-sr-policy-ifit]
Qin, F., Yuan, H., Zhou, T., Min, L., and G. Fioccola, "BGP SR Policy Extensions to Enable IFIT", draft-qin-idr-sr-policy-ifit-01 (work in progress), July 2020.

Appendix A.

Authors' Addresses

Huanan Chen
China Telecom
Guangzhou
China

Email: chenhuan6@chinatelecom.cn

Hang Yuan
UnionPay
1899 Gu-Tang Rd., Pudong
Shanghai
China

Email: yuanhang@unionpay.com

Tianran Zhou
Huawei
156 Beiqing Rd., Haidian District
Beijing
China

Email: zhoutianran@huawei.com

Weidong Li
Huawei
156 Beiqing Rd., Haidian District
Beijing
China

Email: poly.li@huawei.com

Giuseppe Fioccola
Huawei
Riesstrasse, 25
Munich
Germany

Email: giuseppe.fioccola@huawei.com

Yali Wang
Huawei
156 Beiqing Rd., Haidian District
Beijing
China

Email: wangyali11@huawei.com

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 14, 2020

C. Li
M. Chen
Huawei Technologies
W. Cheng
China Mobile
R. Gandhi
Cisco Systems, Inc.
Q. Xiong
ZTE Corporation
March 13, 2020

PCEP Extensions for Associated Bidirectional Segment Routing (SR) Paths
draft-ietf-pce-sr-bidir-path-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. Segment routing (SR) leverages the source routing and tunneling paradigms. The Stateful PCEP extensions allow stateful control of Segment Routing Traffic Engineering (TE) Paths. Furthermore, PCEP can be used for computing SR TE paths in the network.

This document defines PCEP extensions for grouping two unidirectional SR Paths (one in each direction in the network) into a single Associated Bidirectional SR Path. The mechanisms defined in this document can also be applied using a Stateful PCE for both PCE-Initiated and PCC-Initiated LSPs, as well as when using a Stateless 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 14, 2020.

Table of Contents

1. Introduction

Segment routing (SR) [RFC8402] leverages the source routing and tunneling paradigms. SR supports steering packets onto an explicit forwarding path at the ingress node. SR is specified for unidirectional paths. However, some applications require bidirectional paths in SR networks, for example, in mobile backhaul transport networks. The requirement for bidirectional SR Paths is specified in [I-D.ietf-spring-mpls-path-segment].

[RFC5440] describes the Path Computation Element (PCE) Communication Protocol (PCEP). PCEP enables the communication between a Path Computation Client (PCC) and a PCE, or between PCE and PCE, for the purpose of computation of Traffic Engineering (TE) Label Switched Paths (LSP). [RFC8231] specifies a set of extensions to PCEP to enable stateful control of TE LSPs within and across PCEP sessions. The mode of operation where LSPs are initiated from the PCE is described in [RFC8281].

[RFC8408] specifies extensions to the Path Computation Element Protocol (PCEP) [RFC5440] for SR networks, that allow a stateful PCE to compute and initiate SR TE paths, as well as a PCC to request, report or delegate them.

[RFC8697] introduces a generic mechanism to create a grouping of LSPs which can then be used to define associations between a set of LSPs and/or a set of attributes, and is equally applicable to the active and passive modes of a Stateful PCE [RFC8231] or a stateless PCE [RFC5440].

[I-D.ietf-pce-association-bidir] defines PCEP extensions for grouping two unidirectional RSVP-TE LSPs into an Associated Bidirectional LSP when using a Stateful PCE for both PCE-Initiated and PCC-Initiated LSPs as well as when using a Stateless PCE. It specifies the procedure for 'Double-sided Bidirectional LSP Association', where the PCE creates the association and provisions the forward LSPs at their ingress nodes. The RSVP-TE signals the forward LSPs to the egress nodes. Thus, both endpoints learn the reverse LSPs forming the bidirectional LSP association.

This document extends the bidirectional LSP association to SR by specifying PCEP extensions for grouping two unidirectional SR Paths into a bidirectional SR Path. For bidirectional SR, there are use-cases such as directed BFD [I-D.ietf-mpls-bfd-directed] and SR Performance Measurement (PM) [I-D.gandhi-spring-twamp-srpm] those

require PCC to be aware of the reverse direction SR Path. For such use-cases, the reverse SR Paths are also communicated to the ingress nodes using the PCEP extensions defined in this document. This allows both endpoints to be aware of the SR Paths in both directions, including their status and all other path related information. Associating an unidirectional SR Path with a reverse direction unidirectional RSVP-TE LSP to form a bidirectional LSP and vice versa, are outside the scope of this document.

2. Terminology

This document makes use of the terms defined in [RFC8408]. The reader is assumed to be familiar with the terminology defined in [RFC5440], [RFC8231], [RFC8281], [RFC8697], and [I-D.ietf-pce-association-bidir].

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. PCEP Extensions

As per [RFC8697], TE LSPs are associated by adding them to a common association group by a PCEP peer. [I-D.ietf-pce-association-bidir] uses the association group object and the procedures as specified in [RFC8697] to group two unidirectional RSVP-TE LSPs. Similarly, two SR Paths can also be associated using similar technique. This document extends these association mechanisms for bidirectional SR Paths. Two unidirectional SR Paths (one in each direction in the network) can be associated together by using the association group defined in this document for PCEP messages.

Note that the procedure for using the association group defined in this document is specific to the associated bidirectional SR Paths. The procedure for this association group is different than the bidirectional association groups defined in [I-D.ietf-pce-association-bidir] for associated bidirectional RSVP-TE LSPs.

[I-D.ietf-pce-sr-path-segment] defines a mechanism for communicating Path Segment Identifier (PSID) in PCEP for SR. The PSID is defined for SR-MPLS in [I-D.ietf-spring-mpls-path-segment]. The PSID can be used for identifying an SR Path of an associated bidirectional SR Path. The PATH-SEGMENT TLV MAY be included for each SR Path in the

LSP object to support required use-cases. The PATH-SEGMENT TLV MUST be handled as defined in [I-D.ietf-pce-sr-path-segment] and is not modified for associated bidirectional SR Path.

3.1. Double-sided Bidirectional with Reverse LSP Association Group

For associating two unidirectional SR Paths, this document defines a new Association Type called 'Double-sided Bidirectional with Reverse LSP Association Group' for Association Group Object (Class-Value 40) as follows:

- Association Type (TBD1 to be assigned by IANA) = Double-sided Bidirectional with Reverse LSP Association Group

Similar to RSVP-TE bidirectional LSP associations, this Association Type is also operator-configured in nature and statically created by the operator on the PCEP peers. 'Operator-configured Association Range' TLV (Value 29) [RFC8697] MUST NOT be sent for this Association Type, and MUST be ignored, so that the entire range of association ID can be used for it.

The handling of the Association ID, Association Source, optional Global Association Source and optional Extended Association ID in this association are set in the same way as [I-D.ietf-pce-association-bidir].

A member of the 'Double-sided Bidirectional with Reverse LSP Association Group' can take the role of a forward or reverse direction SR Path and follow the similar rules defined in [I-D.ietf-pce-association-bidir] for LSPs.

- An SR Path (forward or reverse) cannot be part of more than one 'Double-sided Bidirectional with Reverse LSP Association Group'.

- The endpoints of the SR Paths in 'Double-sided Bidirectional with Reverse LSP Association Group' cannot be different.

3.1.1. Bidirectional LSP Association Group TLV

In 'Double-sided Bidirectional with Reverse LSP Association Group', for properties such as forward and reverse direction and co-routed path, it uses the Bidirectional LSP Association Group TLV defined in [I-D.ietf-pce-association-bidir]. All fields and processing rules are as per [I-D.ietf-pce-association-bidir].

4. PCEP Procedures

For a Bidirectional SR Path, an ingress PCC is aware of the forward direction SR Path beginning from itself to the egress PCC using the existing PCEP procedures. For the use-cases which require the ingress PCC to be aware of the reverse direction SR Path, PCE informs the reverse SR Path to the ingress PCC. To achieve this, a PCInitiate message for the reverse SR Path is sent to the ingress PCC and a PCInitiate message for the forward SR Path is sent to the egress PCC (with the matching association group). These PCInitiate message MUST NOT trigger initiation of SR Paths on PCCs.

The SR Path can be identified by an LSP of an SR Policy as described in [I-D.barth-pce-segment-routing-policy-cp].

For a bidirectional LSP computation when using both direction LSPs on a node, the same LSP would need to be identified using 2 different PLSP-IDs based on the PCEP session to the ingress or the egress node. Note that the PLSP-ID space is independent at each PCC, the PLSP-ID allocated by the egress PCC cannot be used for the LSP at the ingress PCC (PLSP-ID conflict may occur). As per normal PCInitiate operations, PCC assigns the PLSP-IDs for the local LSPs. Hence, when the PCE notifies an ingress PCC of the reverse LSP, it does so by using PCInitiate operations and sets PLSP-ID to zero and sets the R bit in the Bidirectional LSP Association Group TLV in the association object to indicate that this PCInitiate LSP is a reverse LSP. The PCC upon receiving the PCInitiate MUST locally assign a new PLSP-ID and it MUST issue a PCRpt to PCE for this LSP containing the new PLSP-ID. This reverse direction LSP MUST NOT be instantiated on the PCC.

In other words, a given LSP will be identified by PLSP-ID A at the ingress node while it will be identified by PLSP-ID B at the egress node. The PCE will maintain two PLSP-IDs for the same LSP. For example, ingress PCC1 may report to PCE an LSP1 with PLSP-ID 100. Egress PCC2 may report to PCE an LSP2 with PLSP-ID 200. Both of these LSPs are part of a bidirectional association. When PCE notifies PCC1 of the reverse direction LSP2, it does so by sending a PCInitiate to PCC1 with PLSP-ID set to zero and R bit set in the Bidirectional LSP Association Group TLV. PCC1 upon reception of this generates a new PLSP-ID (example PLSP-ID 300) and issues a PCRpt to PCE. Thus there would two PLSP-ID associated for LSP2 (300 at PCC1 and 200 at PCC2).

4.1. PCE Initiated Associated Bidirectional SR Paths

As specified in [RFC8697], Associated Bidirectional SR Paths can be created by a Stateful PCE as shown in Figure 1.

- Stateful PCE can create and update the forward and reverse SR Paths independently for 'Double-sided Bidirectional with Reverse LSP Association Group'.
- Stateful PCE can establish and remove the association relationship on a per SR Path basis.
- Stateful PCE can create and update the SR Path and the association on a PCC via PCInitiate and PCUpd messages, respectively, using the procedures described in [RFC8697].
- The reverse direction SR Path (LSP2(R) at node S, LSP1(R) at node D as shown in Figure 1) SHOULD be informed by the PCE via PCInitiate message with the matching association group for the use-cases which require the PCC to be aware of the reverse direction SR Path.

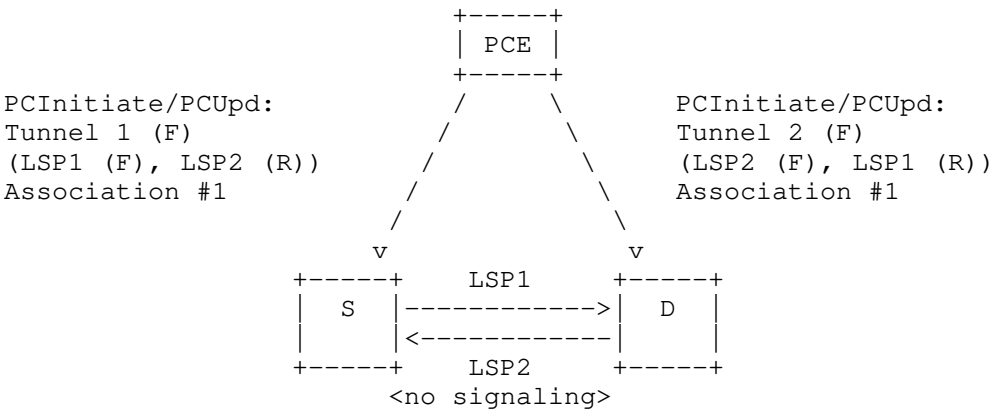


Figure 1: PCE-Initiated Associated Bidirectional SR Path with Forward and Reverse Direction SR Paths

4.2. PCC Initiated Associated Bidirectional SR Paths

As specified in [RFC8697], Associated Bidirectional SR Paths can also be created by a PCC as shown in Figure 2a and 2b.

- PCC can create and update the forward SR Path and update the reverse SR Path independently for a 'Double-sided Bidirectional with Reverse LSP Association Group'.

- PCC cannot instantiate a reverse SR Path in a bidirectional SR Path.

- PCC can establish and remove the association relationship on a per SR Path basis.

- PCC MUST report the change in the association group of an SR Path to PCE(s) via PCRpt message.

- PCC can report the forward and reverse SR Paths independently to PCE(s) via PCRpt message.

- PCC can delegate the forward and reverse SR Paths independently to a Stateful PCE, where PCE would control the SR Paths.

- Stateful PCE can update the SR Paths in the 'Double-sided Bidirectional with Reverse LSP Association Group' via PCUpd message, using the procedures described in [RFC8697].

- The reverse direction SR Path (LSP2(R) at node S, LSP1(R) at node D as shown in Figure 2b) SHOULD be informed by the PCE via PCInitiate message with the matching association group for the use-cases which require the PCC to be aware of the reverse direction SR Path.

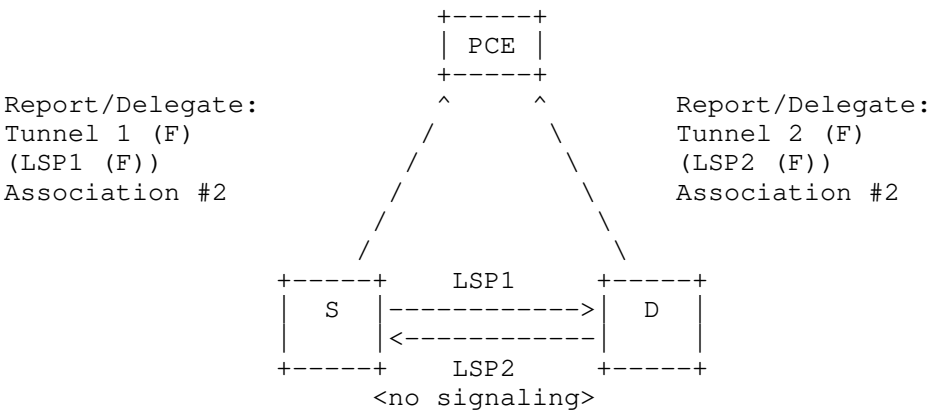


Figure 2a: Step 1: PCC-Initiated Associated Bidirectional SR Path with Forward Direction SR Paths

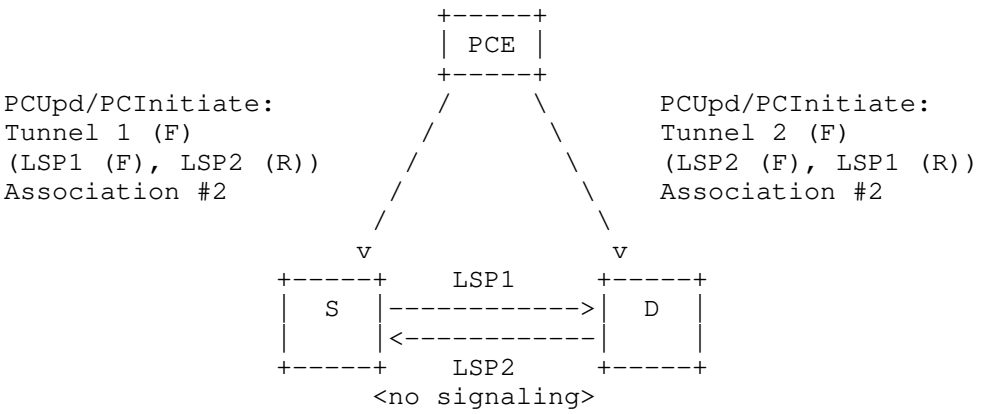


Figure 2b: Step 2: PCE-Updated/Initiated Associated Bidirectional SR Path with Reverse Direction SR Paths

4.3. Stateless PCE

As defined in [I-D.ietf-pce-association-bidir], for a stateless PCE, it might be useful to associate a path computation request to an association group, thus enabling it to associate a common set of configuration parameters or behaviors with the request. A PCC can request co-routed or non-co-routed forward and reverse direction paths from a stateless PCE for a bidirectional SR association group.

4.4. Bidirectional (B) Flag

The Bidirectional (B) flag in Request Parameters (RP) object [RFC5440] and Stateful PCE Request Parameter (SRP) object [I-D.ietf-pce-pcep-stateful-pce-gmpls] follow the procedure defined in [I-D.ietf-pce-association-bidir].

4.5. State Synchronization

During state synchronization, a PCC MUST report all the existing Bidirectional Association Groups to the Stateful PCE as per [RFC8697]. After the state synchronization, the PCE MUST remove all stale Bidirectional Association Groups.

4.6. Error Handling

The error handling as described in section 5.7 of [I-D.ietf-pce-association-bidir] continue to apply.

The PCEP Path Setup Type (PST) for SR is set to 'TE Path is Setup using Segment Routing' [RFC8408]. If a PCEP speaker receives a different PST value for 'Double-sided Bidirectional with Reverse LSP Association Group' and it does not support; it MUST send a PCErr message with Error-Type = 26 (Association Error) and Error-Value = 'Bidirectional LSP Association - Path Setup Type Not Supported' defined in [I-D.ietf-pce-association-bidir].

5. Implementation Status

[Note to the RFC Editor - remove this section before publication, as well as remove the reference to [RFC7942].

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 [RFC7942]. 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 [RFC7942], "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".

5.1. Huawei's Commercial Delivery

The feature is developing based on Huawei VRP8.

- Organization: Huawei
- Implementation: Huawei's Commercial Delivery implementation based on VRP8.
- Description: The implementation is under development.
- Maturity Level: Product
- Contact: tanren@huawei.com

5.2. ZTE's Commercial Delivery

- Organization: ZTE
- Implementation: ZTE's Commercial Delivery implementation based on Rosng v8.
- Description: The implementation is under development.
- Maturity Level: Product
- Contact: zhan.shuangping@zte.com.cn

6. Security Considerations

The security considerations described in [RFC5440], [RFC8231], [RFC8281], and [RFC8408] apply to the extensions defined in this document as well.

A new Association Type for the Association Object, 'Double-sided Bidirectional with Reverse LSP Association Group' is introduced in this document. Additional security considerations related to LSP associations due to a malicious PCEP speaker is described in [RFC8697] and apply to this Association Type. Hence, securing the PCEP session using Transport Layer Security (TLS) [RFC8253] is recommended.

7. Manageability Considerations

All manageability requirements and considerations listed in [RFC5440], [RFC8231], and [RFC8281] apply to PCEP protocol extensions defined in this document. In addition, requirements and considerations listed in this section apply.

7.1. Control of Function and Policy

The mechanisms defined in this document do not imply any control or policy requirements in addition to those already listed in [RFC5440], [RFC8231], and [RFC8281].

7.2. Information and Data Models

[RFC7420] describes the PCEP MIB, there are no new MIB Objects defined for 'Double-sided Bidirectional with Reverse LSP Association Groups'. The PCEP YANG module [I-D.ietf-pce-pcep-yang] defines data model for Associated Bidirectional SR Paths.

7.3. 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], [RFC8231], and [RFC8281].

7.4. Verify Correct Operations

Mechanisms defined in this document do not imply any new operation verification requirements in addition to those already listed in [RFC5440], [RFC8231], and [RFC8408].

7.5. Requirements On Other Protocols

Mechanisms defined in this document do not imply any new requirements on other protocols.

7.6. Impact On Network Operations

Mechanisms defined in [RFC5440], [RFC8231], and [RFC8408] also apply to PCEP extensions defined in this document.

8. IANA Considerations

8.1. Association Type

This document defines a new Association Type for the Association Object (Class Value 40) defined [RFC8697]. IANA is requested to make the assignment of a type for the sub-registry "ASSOCIATION Type" as follows:

Type	Name	Reference
TBD1	Double-sided Bidirectional with Reverse LSP Association Group	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>.

[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>.

[RFC8697] Minei, I., Crabbe, E., Sivabalan, S., Ananthakrishnan, H., Dhody, D., and Y. Tanaka, "Path Computation Element Communication Protocol (PCEP) Extensions for Establishing Relationships between Sets of Label Switched Paths (LSPs)", RFC 8697, DOI 10.17487/RFC8697, January 2020, <https://www.rfc-editor.org/info/rfc8697>.

[I-D.ietf-pce-association-bidir] Gandhi, R., Barth, C., and B. Wen, "PCEP Extensions for Associated Bidirectional Label Switched Paths (LSPs)", draft-ietf-pce-association-bidir-05 (work in progress), February 2020.

[I-D.ietf-pce-sr-path-segment] Li, C., Chen, M., Cheng, W., Gandhi, R., and Q. Xiong, "Path Computation Element Communication Protocol (PCEP) Extension for Path Segment in Segment Routing (SR)", draft-ietf-pce-sr-path-segment-00 (work in progress), October 2019.

9.2. Informative References

[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>.

[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>.

[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>.

[RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <https://www.rfc-editor.org/info/rfc7420>.

[RFC8408] Sivabalan, S., Tantsura, J., Minei, I., Varga, R., and J. Hardwick, "Conveying Path Setup Type in PCE Communication Protocol (PCEP) Messages", RFC 8408, DOI 10.17487/RFC8408, July 2018, <https://www.rfc-editor.org/info/rfc8408>.

[I-D.ietf-mpls-bfd-directed]
Mirsky, G., Tantsura, J., Varlashkin, I., and M. Chen, "Bidirectional Forwarding Detection (BFD) Directed Return Path", draft-ietf-mpls-bfd-directed-13 (work in progress), December 2019.

[I-D.gandhi-spring-twamp-srpm]
Gandhi, R., Filsfils, C., Voyer, D., Chen, M., and B. Janssens, "Performance Measurement Using TWAMP Light and STAMP for Segment Routing Networks", draft-gandhi-spring-twamp-srpm-07 (work in progress), March 2020.

[I-D.ietf-spring-mpls-path-segment]
Cheng, W., Li, H., Chen, M., Gandhi, R., and R. Zigler, "Path Segment in MPLS Based Segment Routing Network", draft-ietf-spring-mpls-path-segment-02 (work in progress), February 2020.

[I-D.ietf-pce-pcep-yang]
Dhody, D., Hardwick, J., Beeram, V., and J. Tantsura, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", draft-ietf-pce-pcep-yang-13 (work in progress), October 2019.

[I-D.ietf-pce-pcep-stateful-pce-gmpls]
Lee, Y., Zheng, H., Dios, O., Lopezalvarez, V., and Z. Ali, "Path Computation Element (PCE) Protocol Extensions for Stateful PCE Usage in GMPLS-controlled Networks", draft-ietf-pce-pcep-stateful-pce-gmpls-12 (work in progress), October 2019.

[I-D.barth-pce-segment-routing-policy-cp]
Koldychev, M., Sivabalan, S., Barth, C., Li, C., and H. Bidgoli, "PCEP extension to support Segment Routing Policy Candidate Paths", draft-barth-pce-segment-routing-policy-cp-04 (work in progress), October 2019.

Acknowledgments

Many thanks to Marina Fizgeer, Adrian Farrel, and Andrew Stone for the detailed review of this document and providing many useful comments.

Contributors

The following people have substantially contributed to this document:

Dhruv Dhody
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

Email: dhruv.ietf@gmail.com

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

Email: lizhenbin@huawei.com

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: jie.dong@huawei.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

Weiqiang Cheng
China Mobile
China

Email: chengweiqiang@chinamobile.com

Rakesh Gandhi
Cisco Systems, Inc.
Canada

Email: rgandhi@cisco.com

Quan Xiong
ZTE Corporation
China

Email: xiong.quan@zte.com.cn

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 31, 2021

C. Li
M. Chen
Huawei Technologies
W. Cheng
China Mobile
R. Gandhi
Cisco Systems, Inc.
Q. Xiong
ZTE Corporation
January 27, 2021

Path Computation Element Communication Protocol (PCEP) Extensions for Associated Bidirectional Segment Routing (SR) Paths
draft-ietf-pce-sr-bidir-path-05

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. Segment routing (SR) leverages the source routing and tunneling paradigms. The Stateful PCEP extensions allow stateful control of Segment Routing Traffic Engineering (TE) Paths. Furthermore, PCEP can be used for computing SR TE paths in the network.

This document defines PCEP extensions for grouping two unidirectional SR Paths (one in each direction in the network) into a single Associated Bidirectional SR Path. The mechanisms defined in this document can also be applied using a Stateful PCE for both PCE-Initiated and PCC-Initiated LSPs, as well as when using a Stateless PCE.

This Internet-Draft will expire on July 31, 2021.

Table of Contents

1. Introduction

Segment routing (SR) [RFC8402] leverages the source routing and tunneling paradigms. SR supports steering packets onto an explicit forwarding path at the ingress node. SR is specified for unidirectional paths. However, some applications require bidirectional paths in SR networks, for example, in mobile backhaul transport networks. The requirement for bidirectional SR Paths is specified in [I-D.ietf-spring-mpls-path-segment] and [I-D.ietf-spring-srv6-path-segment].

[RFC5440] describes the Path Computation Element (PCE) Communication Protocol (PCEP). PCEP enables the communication between a Path Computation Client (PCC) and a PCE, or between PCE and PCE, for the purpose of computation of Traffic Engineering (TE) Label Switched Paths (LSP). [RFC8231] specifies a set of extensions to PCEP to enable stateful control of TE LSPs within and across PCEP sessions. The mode of operation where LSPs are initiated from the PCE is described in [RFC8281].

[RFC8408] specifies extensions to the Path Computation Element Protocol (PCEP) [RFC5440] for SR networks, that allow a stateful PCE to compute and initiate SR TE paths, as well as a PCC to request, report or delegate them.

[RFC8697] introduces a generic mechanism to create a grouping of LSPs. This grouping can then be used to define associations between sets of LSPs or between a set of LSPs and a set of attributes, and it is equally applicable to the stateful PCE (active and passive modes) [RFC8231] and the stateless PCE [RFC5440].

For bidirectional SR paths, there are use-cases such as directed BFD [I-D.ietf-mpls-bfd-directed] and Performance Measurement (PM) [I-D.gandhi-spring-stamp-srpm] those require ingress node (PCC) to be aware of the reverse direction SR Path. For such use-cases, the reverse SR Paths need to be communicated to the ingress node (PCCs) using PCEP mechanisms. This allows both endpoint ingress nodes to be aware of the SR Paths in both directions, including their status and all other path related information.

[I-D.ietf-pce-association-bidir] defines PCEP extensions for grouping two unidirectional Resource Reservation Protocol - Traffic Engineering (RSVP-TE) LSPs into an Associated Bidirectional LSP when

using a Stateful PCE for both PCE-Initiated and PCC-Initiated LSPs as well as when using a Stateless PCE. Specifically, it defines the procedure for 'Double-sided Bidirectional LSP Association', where the PCE creates the association and provisions the forward LSPs at their ingress nodes. The RSVP-TE signals the forward LSPs to the egress nodes. Thus, both endpoints learn the reverse LSPs forming the bidirectional LSP association.

This document extends the Bidirectional LSP Association to SR paths by specifying PCEP extensions for grouping two unidirectional SR Paths into an Associated Bidirectional SR Path. Note that the procedure for using the association group defined in this document is specific to the Associated Bidirectional SR Paths. Associating an unidirectional SR Path with a reverse direction unidirectional RSVP-TE LSP to form a bidirectional LSP and vice versa, are outside the scope of this document.

An SR Policy may contain one or more Candidate-Paths (CPs), each Candidate-Path may contain one or more Segment Lists (SLs) [I-D.ietf-spring-segment-routing-policy]. Recall that in PCEP, an LSP identifies a Candidate-Path as described in [I-D.ietf-pce-segment-routing-policy-cp]. Two unidirectional Candidate-Paths containing a single Segment List (two unidirectional Segment Lists) are associated to form a bidirectional Candidate-Path using the procedures defined in this document. Association of two unidirectional Candidate-Paths containing multiple Segment Lists to form a bidirectional Candidate-Path are outside the scope of this document.

2. Terminology

This document makes use of the terms defined in [RFC8408]. The reader is assumed to be familiar with the terminology defined in [RFC5440], [RFC8231], [RFC8281], [RFC8697], and [I-D.ietf-pce-association-bidir].

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. PCEP Extensions

As per [RFC8697], TE LSPs are associated by adding them to a common association group by a PCEP peer. [I-D.ietf-pce-association-bidir] uses the association group object and the procedures as specified in [RFC8697] to group two unidirectional RSVP-TE LSPs. Similarly, two SR Paths can also be associated using similar technique. This document extends these association mechanisms for bidirectional SR Paths. Two unidirectional SR Paths (one in each direction in the network) can be associated together by using the association group defined in this document for PCEP messages.

[I-D.ietf-pce-sr-path-segment] defines a mechanism for communicating Path Segment Identifier (PSID) in PCEP for SR. The SR-MPLS PSID is defined in [I-D.ietf-spring-mpls-path-segment] and SRv6 PSID is defined in [I-D.ietf-spring-srv6-path-segment]. The PSID can be used for identifying an SR Path of an associated bidirectional SR Path. The PATH-SEGMENT TLV MAY be included for each SR Path in the LSP object to support required use-cases. The PATH-SEGMENT TLV MUST be handled as defined in [I-D.ietf-pce-sr-path-segment] and is not modified for associated bidirectional SR Path.

3.1. Double-sided Bidirectional with Reverse LSP Association

For associating two unidirectional SR Paths, this document defines a new Association Type called 'Double-sided Bidirectional with Reverse LSP Association' for Association Group Object (Class-Value 40) as follows:

- Association Type (TBD1 to be assigned by IANA) = Double-sided Bidirectional with Reverse LSP Association

The Bidirectional Association is considered to be both dynamic and operator-configured in nature. As per [RFC8697], the association group could be manually created by the operator on the PCEP peers, and the LSPs belonging to this association are conveyed via PCEP messages to the PCEP peer; alternately, the association group could be created dynamically by the PCEP speaker, and both the association group information and the LSPs belonging to the association group are conveyed to the PCEP peer. The Operator-configured Association Range MUST be set for this association-type to mark a range of Association Identifiers that are used for operator-configured associations to avoid any Association Identifier clash within the scope of the Association Source (Refer to [RFC8697]). Specifically, for the PCE Initiated Associated Bidirectional SR Paths, the Association Type is dynamically created by the PCE on the PCE peers.

The handling of the Association ID, Association Source, optional Global Association Source and optional Extended Association ID in this association are set in the same way as [I-D.ietf-pce-association-bidir].

[RFC8697] specifies the mechanism for the capability advertisement of the Association Types supported by a PCEP speaker by defining an ASSOC-Type-List TLV to be carried within an OPEN Object. This capability exchange for the Bidirectional Association MUST be done before using the Bidirectional Association Type. Thus, the PCEP speaker MUST include the Bidirectional Association Type in the ASSOC-Type-List TLV and MUST receive the same from the PCEP peer before using the Bidirectional Association in PCEP messages.

A member of the 'Double-sided Bidirectional with Reverse LSP Association' can take the role of a forward or reverse direction SR Path and follow the similar rules defined in [I-D.ietf-pce-association-bidir] for LSPs.

- An SR Path (forward or reverse) MUST NOT be part of more than one 'Double-sided Bidirectional with Reverse LSP Association'.

- The endpoint nodes of the SR Paths in 'Double-sided Bidirectional with Reverse LSP Association' MUST be matching in the reverse directions.

3.1.1. Bidirectional LSP Association Group TLV

In 'Double-sided Bidirectional with Reverse LSP Association', for properties such as forward and reverse direction and co-routed path, it uses the 'Bidirectional LSP Association Group TLV' defined in [I-D.ietf-pce-association-bidir]. All fields and processing rules are as per [I-D.ietf-pce-association-bidir].

4. PCEP Procedures

For an Associated Bidirectional SR Path, an ingress node PCC is aware of the forward direction SR Path beginning from itself to the egress node PCC using the existing PCEP procedures. For the use-cases which require the ingress node PCC to be aware of the reverse direction SR Path, PCE informs the reverse SR Path to the ingress node PCC. To achieve this, a PCInitiate message for the reverse SR Path is sent to the ingress node PCC and a PCInitiate message for the forward SR Path is sent to the egress node PCC (with the matching association group). These PCInitiate message MUST NOT trigger initiation of SR Paths on PCCs.

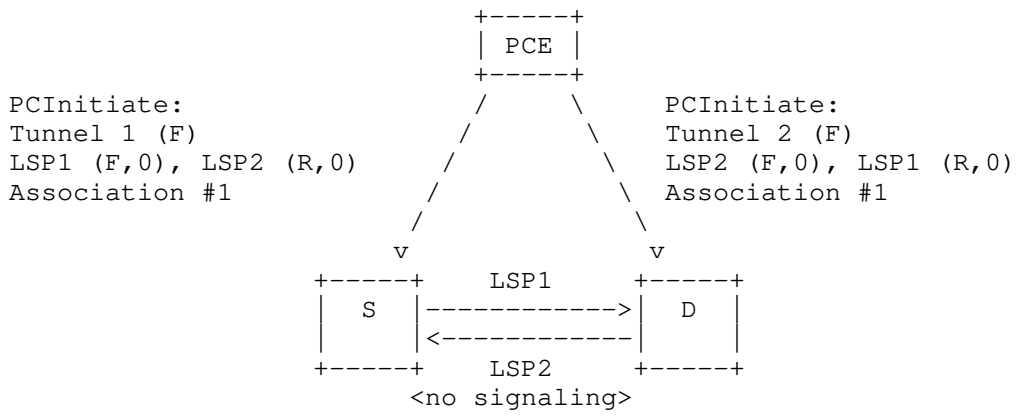
The PCEP procedure defined in this document is applicable to the following three scenarios:

- Neither unidirectional LSP exists, and both must be established.
- Both unidirectional LSPs exist, but the association must be established.
- One LSP exists, but the reverse associated LSP must be established.

4.1. PCE Initiated Associated Bidirectional SR Paths

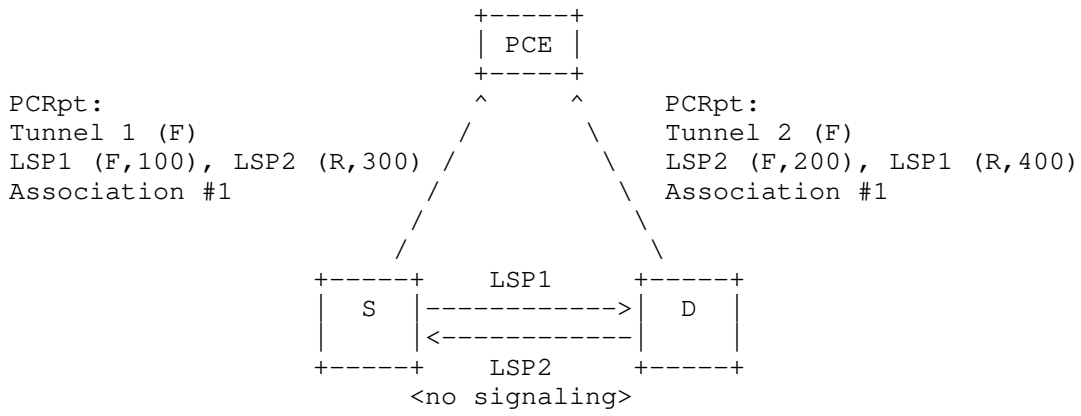
As specified in [RFC8697], Associated Bidirectional SR Paths can be created and updated by a Stateful PCE as shown in Figure 1.

- Stateful PCE MAY create and update the forward and reverse SR Paths independently for the 'Double-sided Bidirectional with Reverse LSP Association'.
- Stateful PCE MAY establish and remove the association relationship on a per SR Path basis.
- Stateful PCE MUST create and update the SR Path and the association on a PCC via PCInitiate and PCUpd messages, respectively, using the procedures described in [RFC8697].
- The reverse direction SR Path (LSP2(R) at node S, LSP1(R) at node D as shown in Figure 1) SHOULD be informed by the PCE via PCInitiate message with the matching association group for the use-cases which require the PCC to be aware of the reverse direction SR Path.



Legends: F = Forward LSP, R = Reverse LSP, (0) = PLSP-IDs

Figure 1a: PCE-Initiated Associated Bidirectional SR Path
with Forward and Reverse Direction SR Paths



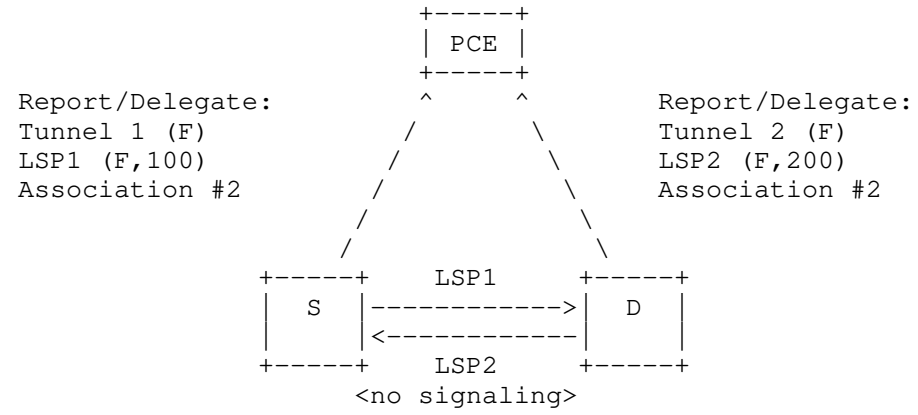
Legends: F=Forward LSP, R = Reverse LSP, (100,200,300,400)=PLSP-IDs

Figure 1b: PCC-Reported Bidirectional SR Path
with Forward and Reverse Direction SR Paths

4.2. PCC Initiated Associated Bidirectional SR Paths

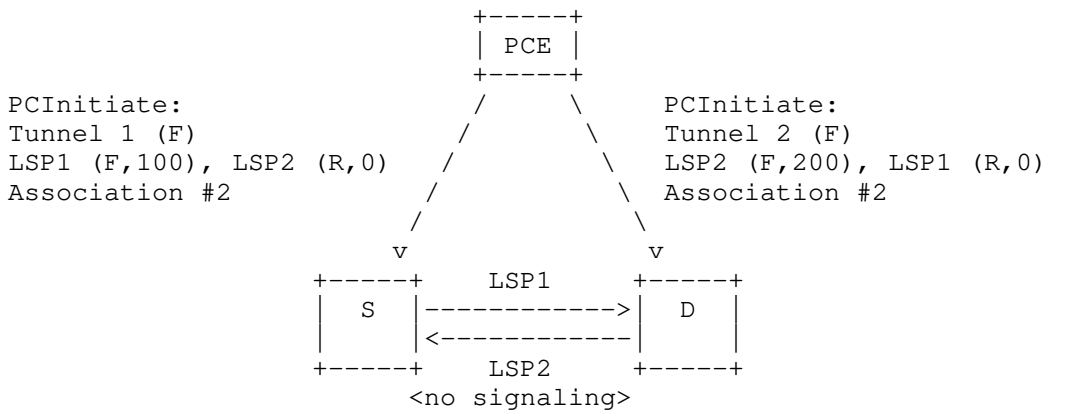
As specified in [RFC8697], Associated Bidirectional SR Paths can also be created and updated by a PCC as shown in Figure 2a and 2b.

- PCC MAY create and update the forward SR Path and update the reverse SR Path independently for the 'Double-sided Bidirectional with Reverse LSP Association'.
- PCC MUST NOT instantiate a reverse SR Path in a bidirectional SR Path.
- PCC MAY establish and remove the association relationship on a per SR Path basis.
- PCC MUST report the change in the association group of an SR Path to PCE(s) via PCRpt message.
- PCC reports the forward and reverse SR Paths independently to PCE(s) via PCRpt message.
- PCC MAY delegate the forward and reverse SR Paths independently to a Stateful PCE, where PCE would control the SR Paths.
- Stateful PCE updates the SR Paths in the 'Double-sided Bidirectional with Reverse LSP Association' via PCUpd message, using the procedures described in [RFC8697].
- The reverse direction SR Path (LSP2(R) at node S, LSP1(R) at node D as shown in Figure 2b) SHOULD be informed by the PCE via PCInitiate message with the matching association group for the use-cases which require the PCC to be aware of the reverse direction SR Path.



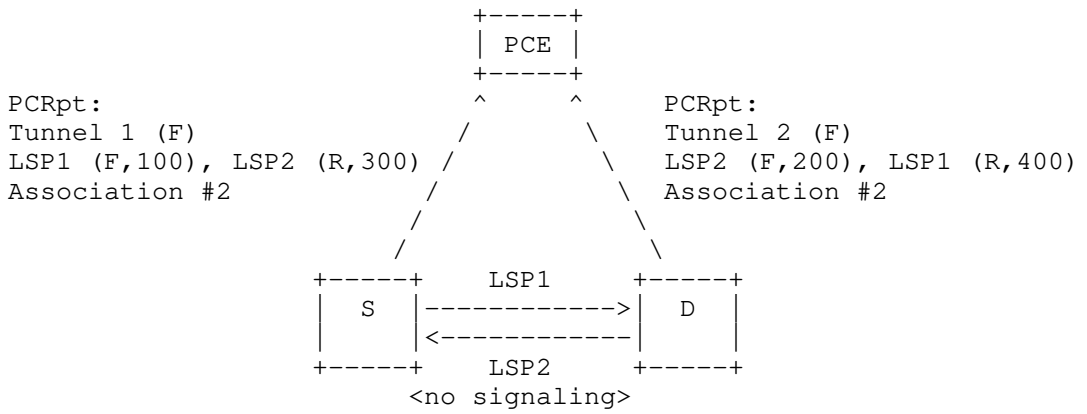
Legends: F = Forward LSP, R = Reverse LSP, (100,200) = PLSP-IDs

Figure 2a: Step 1: PCC-Initiated Associated Bidirectional SR Path with Forward Direction SR Paths



Legends: F = Forward LSP, R = Reverse LSP, (0,100,200) = PLSP-IDs

Figure 2b: Step 2: PCE-Initiated Associated Bidirectional SR Path with Reverse Direction SR Paths



Legends: F=Forward LSP, R = Reverse LSP, (100,200,300,400)=PLSP-IDs

Figure 2c: Step 3: PCC-Reported Associated Bidirectional SR Path with Reverse Direction SR Paths

4.3. Stateless PCE

As defined in [I-D.ietf-pce-association-bidir], for a stateless PCE, it might be useful to associate a path computation request to an association group, thus enabling it to associate a common set of configuration parameters or behaviors with the request [RFC8697]. A

PCC can request co-routed or non-co-routed forward and reverse direction paths from a stateless PCE for a Bidirectional SR Path.

4.4. Bidirectional (B) Flag

The Bidirectional (B) flag in Request Parameters (RP) object [RFC5440] and Stateful PCE Request Parameter (SRP) object [I-D.ietf-pce-pcep-stateful-pce-gmpls] follow the procedure defined in [I-D.ietf-pce-association-bidir].

4.5. PLSP-ID Usage

For a bidirectional LSP computation when using both direction LSPs on a node, the same LSP would need to be identified using 2 different PLSP-IDs based on the PCEP session to the ingress or the egress node. Note that the PLSP-ID space is independent at each PCC, the PLSP-ID allocated by the egress PCC cannot be used for the LSP at the ingress PCC (PLSP-ID conflict may occur). As per normal PCInitiate operations, PCC assigns the PLSP-IDs for the local LSPs. Hence, when the PCE notifies an ingress PCC of the reverse LSP, it does so by using PCInitiate operations and sets PLSP-ID to zero and sets the R bit in the 'Bidirectional LSP Association Group TLV' in the association object to indicate that this PCInitiate LSP is a reverse LSP. The PCC upon receiving the PCInitiate MUST locally assign a new PLSP-ID and it MUST issue a PCRpt to PCE for this LSP containing the new PLSP-ID. This reverse direction LSP MUST NOT be instantiated on the PCC.

In other words, a given LSP will be identified by PLSP-ID A at the ingress node while it will be identified by PLSP-ID B at the egress node. The PCE will maintain two PLSP-IDs for the same LSP. For example, ingress PCC1 may report to PCE an LSP1 with PLSP-ID 100. Egress PCC2 may report to PCE an LSP2 with PLSP-ID 200. Both of these LSPs are part of a bidirectional association. When PCE notifies PCC1 of the reverse direction LSP2, it does so by sending a PCInitiate to PCC1 with PLSP-ID set to zero and R bit set in the 'Bidirectional LSP Association Group TLV'. PCC1 upon reception of this generates a new PLSP-ID (example PLSP-ID 300) and issues a PCRpt to PCE. Thus there would two PLSP-ID associated for LSP2 (300 at PCC1 and 200 at PCC2).

For an Associated Bidirectional SR Path, LSP-IDENTIFIERS TLV [RFC8231] MUST be included in all forward and reverse LSPs.

4.6. State Synchronization

During state synchronization, a PCC MUST report all the existing Bidirectional Associations to the Stateful PCE as per [RFC8697]. After the state synchronization, the PCE MUST remove all stale Bidirectional Associations.

4.7. Error Handling

The error handling as described in section 5.7 of [I-D.ietf-pce-association-bidir] continue to apply.

The PCEP Path Setup Type (PST) for SR is set to 'TE Path is Setup using Segment Routing' [RFC8408] or 'Path is setup using SRv6' [I-D.ietf-pce-segment-routing-ipv6].

If a PCEP speaker receives a different PST value for the 'Double-sided Bidirectional with Reverse LSP Association', the PCE speaker MUST return a PCErr message with Error-Type = 26 (Association Error) and Error-Value = 'Bidirectional LSP Association - Path Setup Type Not Supported' defined in [I-D.ietf-pce-association-bidir].

5. Implementation Status

[Note to the RFC Editor - remove this section before publication, as well as remove the reference to [RFC7942].

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 [RFC7942]. 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 [RFC7942], "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".

5.1. Huawei's Commercial Delivery

The feature is developing based on Huawei VRP8.

- Organization: Huawei
- Implementation: Huawei's Commercial Delivery implementation based on VRP8.
- Description: The implementation is under development.
- Maturity Level: Product
- Contact: tanren@huawei.com

5.2. ZTE's Commercial Delivery

- Organization: ZTE
- Implementation: ZTE's Commercial Delivery implementation based on Rosng v8.
- Description: The implementation is under development.
- Maturity Level: Product
- Contact: zhan.shuangping@zte.com.cn

6. Security Considerations

The security considerations described in [RFC5440], [RFC8231], [RFC8281], and [RFC8408] apply to the extensions defined in this document as well.

A new Association Type for the Association Object, 'Double-sided Bidirectional with Reverse LSP Association' is introduced in this document. Additional security considerations related to LSP associations due to a malicious PCEP speaker is described in [RFC8697] and apply to this Association Type. Hence, securing the PCEP session using Transport Layer Security (TLS) [RFC8253] is recommended.

7. Manageability Considerations

All manageability requirements and considerations listed in [RFC5440], [RFC8231], and [RFC8281] apply to PCEP protocol extensions defined in this document. In addition, requirements and considerations listed in this section apply.

7.1. Control of Function and Policy

The mechanisms defined in this document do not imply any control or policy requirements in addition to those already listed in [RFC5440], [RFC8231], and [RFC8281].

7.2. Information and Data Models

[RFC7420] describes the PCEP MIB, there are no new MIB Objects defined for 'Double-sided Bidirectional with Reverse LSP Associations'. The PCEP YANG module [I-D.ietf-pce-pcep-yang] defines data model for Associated Bidirectional SR Paths.

7.3. 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], [RFC8231], and [RFC8281].

7.4. Verify Correct Operations

Mechanisms defined in this document do not imply any new operation verification requirements in addition to those already listed in [RFC5440], [RFC8231], and [RFC8408].

7.5. Requirements On Other Protocols

Mechanisms defined in this document do not imply any new requirements on other protocols.

7.6. Impact On Network Operations

Mechanisms defined in [RFC5440], [RFC8231], and [RFC8408] also apply to PCEP extensions defined in this document.

8. IANA Considerations

8.1. Association Type

This document defines a new Association Type, originally described in [RFC8697]. IANA is requested to assign the following new value in the "ASSOCIATION Type Field" subregistry [RFC8697] within the "Path Computation Element Protocol (PCEP) Numbers" registry:

Type	Name	Reference
TBD1	Double-sided Bidirectional with Reverse LSP Association	[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>.

[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>.

[RFC8697] Minei, I., Crabbe, E., Sivabalan, S., Ananthakrishnan, H., Dhody, D., and Y. Tanaka, "Path Computation Element Communication Protocol (PCEP) Extensions for Establishing Relationships between Sets of Label Switched Paths (LSPs)", RFC 8697, DOI 10.17487/RFC8697, January 2020, <https://www.rfc-editor.org/info/rfc8697>.

[I-D.ietf-pce-segment-routing-ipv6] Li, C., Negi, M., Sivabalan, S., Koldychev, M., Kaladharan, P., and Y. Zhu, "PCEP Extensions for Segment Routing leveraging the IPv6 data plane", draft-ietf-pce-segment-routing-ipv6-08 (work in progress), November 2020.

[I-D.ietf-pce-association-bidir]
Gandhi, R., Barth, C., and B. Wen, "Path Computation Element Communication Protocol (PCEP) Extensions for Associated Bidirectional Label Switched Paths (LSPs)", draft-ietf-pce-association-bidir-10 (work in progress), January 2021.

[I-D.ietf-pce-sr-path-segment]
Li, C., Chen, M., Cheng, W., Gandhi, R., and Q. Xiong, "Path Computation Element Communication Protocol (PCEP) Extension for Path Segment in Segment Routing (SR)", draft-ietf-pce-sr-path-segment-02 (work in progress), November 2020.

9.2. Informative References

[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>.

[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>.

[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>.

[RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <https://www.rfc-editor.org/info/rfc7420>.

[RFC8408] Sivabalan, S., Tantsura, J., Minei, I., Varga, R., and J. Hardwick, "Conveying Path Setup Type in PCE Communication Protocol (PCEP) Messages", RFC 8408, DOI 10.17487/RFC8408, July 2018, <https://www.rfc-editor.org/info/rfc8408>.

[I-D.ietf-mpls-bfd-directed]
Mirsky, G., Tantsura, J., Varlashkin, I., and M. Chen, "Bidirectional Forwarding Detection (BFD) Directed Return Path for MPLS Label Switched Paths (LSPs)", draft-ietf-mpls-bfd-directed-15 (work in progress), August 2020.

[I-D.gandhi-spring-stamp-srpm]
Gandhi, R., Filsfils, C., Voyer, D., Chen, M., and B. Janssens, "Performance Measurement Using Simple TWAMP (STAMP) for Segment Routing Networks", draft-gandhi-spring-stamp-srpm-04 (work in progress), January 2021.

[I-D.ietf-spring-mpls-path-segment]
Cheng, W., Li, H., Chen, M., Gandhi, R., and R. Zigler, "Path Segment in MPLS Based Segment Routing Network", draft-ietf-spring-mpls-path-segment-03 (work in progress), September 2020.

[I-D.ietf-spring-srv6-path-segment]
Li, C., Cheng, W., Chen, M., Dhody, D., and R. Gandhi, "Path Segment for SRv6 (Segment Routing in IPv6)", draft-ietf-spring-srv6-path-segment-00 (work in progress), November 2020.

[I-D.ietf-pce-pcep-yang]
Dhody, D., Hardwick, J., Beeram, V., and J. Tantsura, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", draft-ietf-pce-pcep-yang-15 (work in progress), October 2020.

[I-D.ietf-pce-pcep-stateful-pce-gmpls]
Lee, Y., Zheng, H., Dios, O., Lopez, V., and Z. Ali, "Path Computation Element (PCE) Protocol Extensions for Stateful PCE Usage in GMPLS-controlled Networks", draft-ietf-pce-pcep-stateful-pce-gmpls-14 (work in progress), December 2020.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-09 (work in progress), November 2020.

[I-D.ietf-pce-segment-routing-policy-cp]
Koldychev, M., Sivabalan, S., Barth, C., Peng, S., and H. Bidgoli, "PCEP extension to support Segment Routing Policy Candidate Paths", draft-ietf-pce-segment-routing-policy-cp-02 (work in progress), January 2021.

Acknowledgments

Many thanks to Marina Fizgeer, Adrian Farrel, Andrew Stone, and Tarek Saad for the detailed review of this document and providing many useful comments.

Contributors

The following people have substantially contributed to this document:

Dhruv Dhody
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

Email: dhruv.ietf@gmail.com

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

Email: lizhenbin@huawei.com

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: jie.dong@huawei.com

Authors' Addresses

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

Email: c.l@huawei.com

Mach(Guoyi) Chen
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: Mach.chen@huawei.com

Weiqiang Cheng
China Mobile
China

Email: chengweiqiang@chinamobile.com

Rakesh Gandhi
Cisco Systems, Inc.
Canada

Email: rgandhi@cisco.com

Quan Xiong
ZTE Corporation
China

Email: xiong.quan@zte.com.cn

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 7, 2021

M. Koldychev
Cisco Systems, Inc.
S. Sivabalan
Ciena Corporation
T. Saad
V. Beeram
Juniper Networks, Inc.
H. Bidgoli
Nokia
B. Yadav
Ciena
S. Peng
Huawei Technologies
July 06, 2020

PCEP Extensions for Signaling Multipath Information
draft-koldychev-pce-multipath-03

Abstract

Current PCEP standards allow only one intended and/or actual path to be present in a PCEP report or update. Applications that require multipath support such as SR Policy require an extension to allow signaling multiple intended and/or actual paths within a single PCEP message. This document introduces such an extension. Encoding of multiple intended and/or actual paths is done by encoding multiple Explicit Route Objects (EROs) and/or multiple Record Route Objects (RROs). A special separator object is defined in this document, to facilitate this. This mechanism is applicable to SR-TE and RSVP-TE and is dataplane agnostic.

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 7, 2021.

Table of Contents

1. Introduction

Path Computation Element (PCE) Communication Protocol (PCEP) [RFC5440] enables the communication between a Path Computation Client (PCC) and a Path Control Element (PCE), or between two PCEs based on the PCE architecture [RFC4655].

PCEP Extensions for the Stateful PCE Model [RFC8231] describes a set of extensions to PCEP that enable active control of Multiprotocol Label Switching Traffic Engineering (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.

PCEP Extensions for Segment Routing [RFC8664] specifies extensions to the Path Computation Element Protocol (PCEP) that allow a stateful PCE to compute and initiate Traffic Engineering (TE) paths, as well as for a PCC to request a path subject to certain constraint(s) and optimization criteria in SR networks.

Segment Routing Policy for Traffic Engineering [I-D.ietf-spring-segment-routing-policy] details the concepts of SR Policy and approaches to steering traffic into an SR Policy. In particular, it describes the SR candidate-path as a collection of one or more Segment-Lists. The current PCEP standards only allow for signaling of one Segment-List per Candidate-Path. PCEP extension to support Segment Routing Policy Candidate Paths [I-D.ietf-pce-segment-routing-policy-cp] specifically avoids defining how to signal multipath information, and states that this will be defined in another document.

This document defines the required extensions that allow the signaling of multipath information via PCEP.

2. Terminology

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.

2.1. Terms and Abbreviations

The following terms are used in this document:

PCEP Tunnel:

The object identified by the PLSP-ID, see [I-D.koldychev-pce-operational] for more details.

3. Motivation

This extension is motivated by the use-cases described below.

3.1. Signaling Multiple Segment-Lists of an SR Candidate-Path

The Candidate-Path of an SR Policy is the unit of report/update in PCEP, see [I-D.ietf-pce-segment-routing-policy-cp]. Each Candidate-Path can contain multiple Segment-Lists and each Segment-List is encoded by one ERO. However, each PCEP LSP can contain only a single ERO (containing multiple SR-ERO subobject), which prevents us from encoding multiple Segment- Lists within the same SR Candidate-Path.

With the help of the protocol extensions defined in this document, this limitation is overcome.

3.2. Splitting of Requested Bandwidth

A PCC may request a path with 80 Gbps of bandwidth, but all links in the network have only 50 Gbps capacity. The PCE can return two paths, that can together carry 80 Gbps. The PCC can then equally or unequally split the incoming 80 Gbps of traffic among the two paths. Section 4.3 introduces a new TLV that carries the path weight that allows for distribution of incoming traffic on to the multiple paths.

3.3. Providing Backup path for Protection

It is desirable for the PCE to compute and signal to the PCC a backup path that is used to protect a primary path within the multipaths in a given LSP.

Note that [RFC8745] specify the Path Protection association among LSPs. The use of [RFC8745] with multipath is out of scope of this document and is for future study.

When multipath is used, a backup path may protect one or more primary paths. For this reason, primary and backup path identifiers are needed to indicate which backup path(s) protect which primary path(s). Section 4.4 introduces a new TLV that carries the required information.

4. Protocol Extensions

4.1. Multipath Capability TLV

We define the MULTIPATH-CAP TLV that MAY be present in the OPEN object and/or the LSP object. The purpose of this TLV is two-fold:

1. From PCC: it tells how many multipaths per PCEP Tunnel, the PCC can install in forwarding.

2. From PCE: it tells that the PCE supports this standard and how many multipaths per PCEP Tunnel, the PCE can compute.

Only the first instance of this TLV can be processed, subsequent instances SHOULD be ignored.

Section 5 specify the usage of this TLV with Open message (within the OPEN object) and other PCEP messages (within the LSP object).

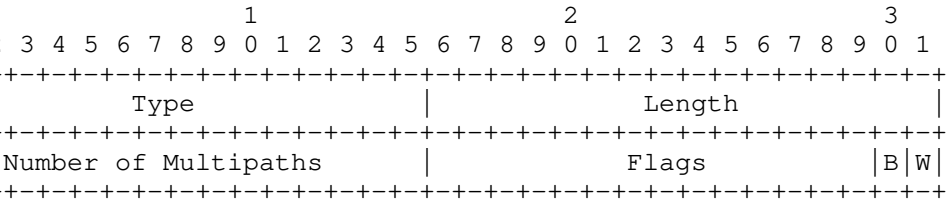


Figure 1: MULTIPATH-CAP TLV format

Type: TBD1 for "MULTIPATH-CAP" TLV.

Length: 4.

Number of Multipaths: the maximum number of multipaths per PCEP Tunnel. The value 0 indicates unlimited number.

Flags: Following bits are defined:

W-flag: whether MULTIPATH-WEIGHT-TLV is supported.

B-flag: whether MULTIPATH-BACKUP-TLV is supported.

Unassigned bits are for future use. They MUST be set to 0 on transmission and MUST be ignored on receipt.

4.2. Path Attributes Object

We define the PATH-ATTRIB object that is used to carry per-path information and to act as a separator between several ERO/RRO objects in the intended-path/actual-path RBNF element. The PATH-ATTRIB object always precedes the ERO/RRO that it applies to. If multiple ERO/RRO objects are present, then each ERO/RRO object MUST be preceded by an PATH-ATTRIB object that describes it.

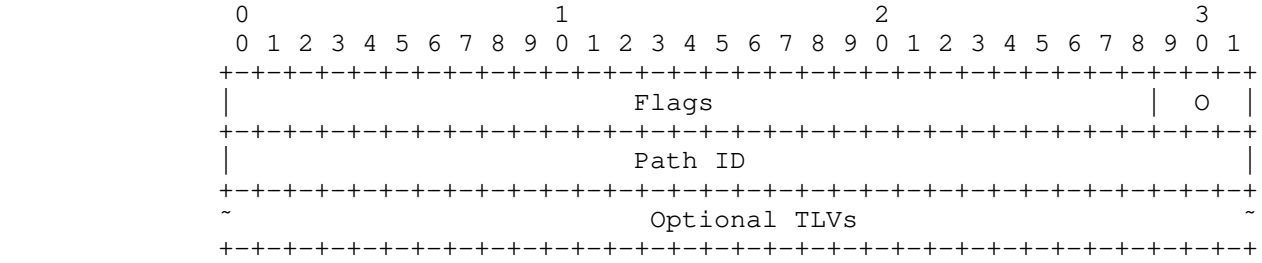


Figure 2: PATH-ATTRIB object format

Flags (32-bits): Following bits are assigned -

O (Operational - 3 bits): operational state of the path, same values as the identically named field in the LSP object {{RFC8231}}.

Unassigned bits are for future use. They MUST be set to 0 on transmission and MUST be ignored on receipt.

Path ID: 4-octet identifier that identifies a path in the set of multiple paths. It uniquely identifies a path (encoded in the ERO/RRO) within the set of multiple paths under the PCEP LSP. Once a path changes, a new Path ID is assigned.

TLVs that may be included in the PATH-ATTRIB object are described in the following sections. Other optional TLVs could be defined by future documents to be included within the PATH-ATTRIB object body.

4.3. Multipath Weight TLV

We define the MULTIPATH-WEIGHT TLV that MAY be present in the PATH-ATTRIB object.

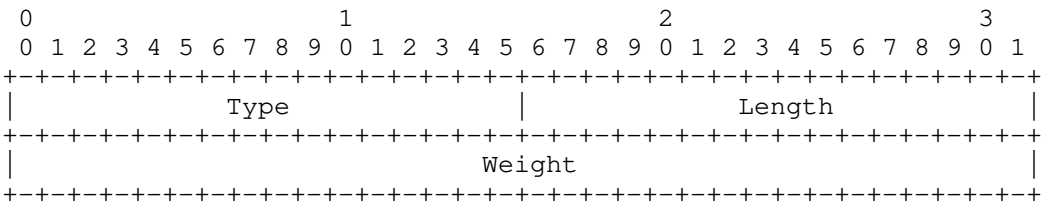


Figure 3: MULTIPATH-WEIGHT TLV format

Type: TBD2 for "MULTIPATH-WEIGHT" TLV.

Length: 4.

Weight: weight of this path within the multipath, if W-ECMP is desired. The fraction of flows a specific ERO/RRO carries is derived from the ratio of its weight to the sum of all other multipath ERO/RRO weights.

4.4. Multipath Backup TLV

This document introduces a new MULTIPATH-BACKUP TLV that is optional and can be present in the PATH-ATTRIB object.

This TLV is used to indicate the presence of a backup path that is used for protection in case of failure of the primary path. The format of the MULTIPATH-BACKUP TLV is:

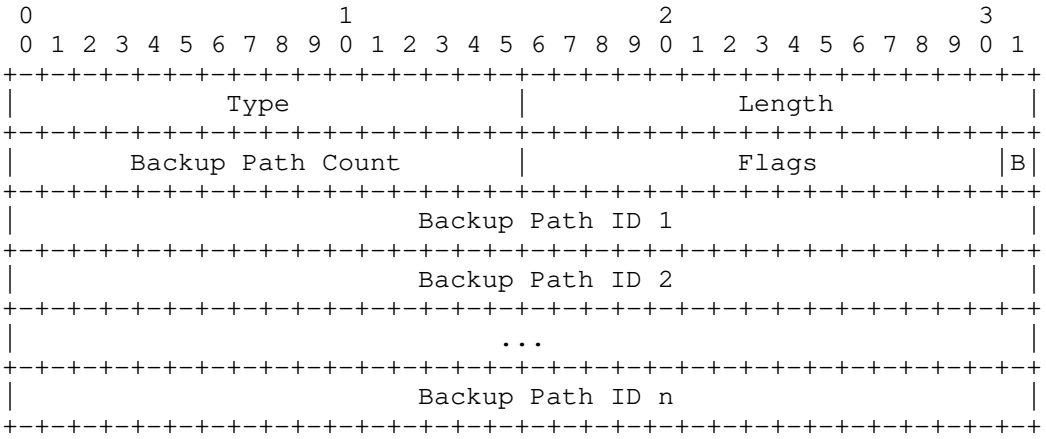


Figure 4: MULTIPATH-BACKUP TLV format

Type: TBD3 for "MULTIPATH-BACKUP" TLV

Length: 4 + (N * 4) (where N is the Backup Path Count)

Backup Path Count: Number of backup path(s).

Flags (16 bits): a flag field. Currently a single flag "B bit" is defined.
Unused flags MUST be set to zero while sending and ignored on receipt.

B: If set, indicates a pure backup path. This is a path that only carries rerouted traffic after the protected path fails. If this flag is not set, or if the MULTIPATH-BACKUP TLV is not carried in the PATH-ATTRIB object of an ERO or SERO, then the path is assumed to be primary that carries normal traffic.

Backup Path ID(s): a series of 4-octet identifier(s) that identify the backup path(s) in the set that protect this primary path.

5. Operation

When the PCC wants to indicate to the PCE that it wants to get multipaths for a PCEP Tunnel, instead of a single path, it can do (1) or both (1) and (2) of the following:

(1) Send the MULTIPATH-CAP TLV in the OPEN object during session establishment. This applies to all PCEP Tunnels on the PCC, unless overridden by PCEP Tunnel specific information.

(2) Additionally send the MULTIPATH-CAP TLV in the LSP object for a particular PCEP Tunnel in the PCRpt or PCReq message. This applies to the specified PCEP Tunnel and overrides the information from the OPEN object.

When PCE computes the path for a PCEP Tunnel, it MUST NOT return more multipaths than the corresponding value of "Number of Multipaths" from the MULTIPATH-CAP TLV. If this TLV is absent (from both OPEN and LSP objects), then the "Number of Multipaths" is assumed to be 1.

If the PCE supports this standard, then it MUST include the MULTIPATH-CAP TLV in the OPEN object. This tells the PCC that it can report multiple ERO/RRO objects per PCEP Tunnel to this PCE. If the PCE does not include the MULTIPATH-CAP TLV in the OPEN object, then the PCC MUST assume that the PCE does not support this standard and fall back to reporting only a single ERO/RRO. The PCE MUST NOT include MULTIPATH-CAP TLV in the LSP object in any other PCEP message towards the PCC and the PCC MUST ignore it if received.

The Path ID of each ERO/RRO MUST be unique within that LSP. If a PCEP speaker detects that there are two paths with the same Path ID, then the PCEP speaker SHOULD send PCError message with Error-Type = 1 ("Reception of an invalid object") and Error-Value = TBD4 ("Conflicting Path ID").

5.1. Signaling Multiple Paths for Loadbalancing

The PATH-ATTRIB object can be used to signal multiple path(s) and indicate (un)equal loadbalancing amongst the set of multipaths. In this case, the PATH-ATTRIB is populated for each ERO or SERO as follows:

1. The PCE assigns a unique Path ID to each ERO or SERO path and populates it inside the PATH-ATTRIB object. The Path ID is unique within the context of a PLSP or PCEP Tunnel.

2. The MULTIPATH-WEIGHT TLV MAY be carried inside the PATH-ATTRIB object. A weight is populated to reflect the relative loadshare that is to be carried by the path. If the MULTIPATH-WEIGHT is not carried inside a PATH-ATTRIB object, the default weight 1 MUST be assumed when computing the loadshare.

3. The fraction of flows carried by a specific primary path is derived from the ratio of its weight to the sum of all other multipath weights.

5.2. Signaling Multiple Paths for Protection

The PATH-ATTRIB object can be used to describe a set of backup path(s) protecting a primary path within a PCEP Tunnel. In this case, the PATH-ATTRIB is populated for each ERO or SERO as follows:

1. The PCE assigns a unique Path ID to each ERO or SERO path and populates it inside the PATH-ATTRIB object. The Path ID is unique within the context of a PLSP or PCEP Tunnel.

2. The MULTIPATH-BACKUP TLV MUST be added inside the PATH-ATTRIB object for each ERO or SERO that is protected. The backup path ID(s) are populated in the MULTIPATH-BACKUP TLV to reflect the set of backup path(s) protecting the primary path. The Length field and Backup Path Number in the MULTIPATH-BACKUP are updated according to the number of backup path ID(s) included.

3. The MULTIPATH-BACKUP TLV MAY be added inside the PATH-ATTRIB object for each ERO or SERO that is unprotected. In this case, MULTIPATH-BACKUP does not carry any backup path IDs in the TLV. If the path acts as a pure backup - i.e. the path only carries

rerouted traffic after the protected path(s) fail- then the B flag MUST be set.

Note that if a given path has the B-flag set, then there MUST be some other path within the same LSP that uses the given path as a backup. If this condition is violated, then the PCEP speaker SHOULD send a PCError message with Error-Type = 10 ("Reception of an invalid object") and Error-Value = TBD5 ("No primary path for pure backup").

Note that a given PCC may not support certain backup combinations, such as a backup path that is itself protected by another backup path, etc. If a PCC is not able to implement a requested backup scenario, the PCC SHOULD send a PCError message with Error-Type = 19 ("Invalid Operation") and Error-Value = TBD6 ("Not supported path backup").

6. PCEP Message Extensions

The RBNF of PCReq, PCRep, PCRpt, PCUpd and PCInit messages currently use intended-path and/or actual-path:

```
<intended-path> ::= (<ERO>|<SERO>)
                    [<intended-path>]

<actual-path> ::= (<RRO>|<SRRO>)
                  [<actual-path>]
```

In this standard, we extend these two elements:

```
<intended-path> ::= ((<ERO>|<SERO>)|
                     (<PATH-ATTRIB>(<ERO>|<SERO>)
                     [<intended-path>]))


<actual-path> ::= ((<RRO>|<SRRO>)|
                   (<PATH-ATTRIB>(<RRO>|<SRRO>)
                   [<actual-path>]))
```

7. Examples

7.1. SR Policy Candidate-Path with Multiple Segment-Lists

Consider how the following sample SR Policy, taken from [I-D.ietf-spring-segment-routing-policy], would be represented in a PCRpt message.

```
          SR policy POL1 <headend, color, endpoint>
              Candidate-path CP1 <protocol-origin = 20, originator =
      100:1.1.1.1, discriminator = 1>
                  Preference 200
                  Weight W1, SID-List1 <SID11...SID1i>
                  Weight W2, SID-List2 <SID21...SID2j>
              Candidate-path CP2 <protocol-origin = 20, originator =
      100:2.2.2.2, discriminator = 2>
                  Preference 100
                  Weight W3, SID-List3 <SID31...SID3i>
                  Weight W4, SID-List4 <SID41...SID4j>
```

As specified in [I-D.ietf-pce-segment-routing-policy-cp], CP1 and CP2 are signaled as separate state-report elements and each has a unique PLSP-ID, assigned by the PCC. Let us assign PLSP-ID 100 to CP1 and PLSP-ID 200 to CP2.

The state-report for CP1 can be encoded as:

```
<state-report> =
        <LSP PLSP_ID=100>
        <ASSOCIATION>
        <END-POINT>
        <PATH-ATTRIB Path_ID=1 <WEIGHT-TLV Weight=W1>>
        <ERO SID-List1>
        <PATH-ATTRIB Path_ID=2 <WEIGHT-TLV Weight=W2>>
        <ERO SID-List2>
```

The state-report for CP2 can be encoded as:

```
<state-report> =
        <LSP PLSP_ID=200>
        <ASSOCIATION>
        <END-POINT>
        <PATH-ATTRIB Path_ID=1 <WEIGHT-TLV Weight=W3>>
        <ERO SID-List3>
        <PATH-ATTRIB Path_ID=2 <WEIGHT-TLV Weight=W4>>
        <ERO SID-List4>
```

The above sample state-report elements only specify the minimum mandatory objects, of course other objects like SRP, LSPA, METRIC, etc., are allowed to be inserted.

Note that the syntax

```
<PATH-ATTRIB Path_ID=1 <WEIGHT-TLV Weight=W1>>
```

, simply means that this is PATH-ATTRIB object with Path ID field set to "1" and with a MULTIPATH-WEIGHT TLV carrying weight of "W1".

7.2. Two Primary Paths Protected by One Backup Path

Suppose there are 3 paths: A, B, C. Where A,B are primary and C is to be used only when A or B fail. Suppose the Path IDs for A, B, C are respectively 1, 2, 3. This would be encoded in a state-report as:

```
<state-report> =
        <LSP>
        <ASSOCIATION>
        <END-POINT>
        <PATH-ATTRIB Path_ID=1 <BACKUP-TLV B=0, Backup_Paths=[3]>>
        <ERO A>
        <PATH-ATTRIB Path_ID=2 <BACKUP-TLV B=0, Backup_Paths=[3]>>
        <ERO B>
        <PATH-ATTRIB Path_ID=3 <BACKUP-TLV B=1, Backup_Paths=[]>>
        <ERO C>
```

Note that the syntax

```
<PATH-ATTRIB Path_ID=1 <BACKUP-TLV B=0, Backup_Paths=[3]>>
```

, simply means that this is PATH-ATTRIB object with Path ID field set to "1" and with a MULTIPATH-BACKUP TLV that has B-flag cleared and contains a single backup path with Backup Path ID of 3.

8. IANA Considerations

IANA is requested to make the assignment of a new value for the existing "PCEP TLV Type Indicators" registry as follows:

TLV Type Value	TLV Name	Reference
TBD1	MULTIPATH-CAP	This document
TBD2	MULTIPATH-WEIGHT	This document
TBD3	MULTIPATH-BACKUP	This document

IANA is requested to make the assignment of a new value for the existing "PCEP-ERROR Object Error Types and Values" registry as follows:

Error-Type	Error-Value	Reference
10	TBD4 - Conflicting Path ID	This document
10	TBD5 - No primary path for pure backup	This document
19	TBD6 - Not supported path backup	This document

9. Security Considerations

None at this time.

10. Acknowledgement

Thanks to Dhruv Dhody for ideas and discussion.

11. Contributors

Andrew Stone
Nokia

Email: andrew.stone@nokia.com

12. References

12.1. Normative References

[I-D.ietf-pce-segment-routing-policy-cp]
Koldychev, M., Sivabalan, S., Barth, C., Peng, S., and H. Bidgoli, "PCEP extension to support Segment Routing Policy Candidate Paths", draft-ietf-pce-segment-routing-policy-cp-00 (work in progress), June 2020.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-07 (work in progress), May 2020.

[I-D.koldychev-pce-operational]
Koldychev, M., Sivabalan, S., Negi, M., Achaval, D., and H. Kotni, "PCEP Operational Clarification", draft-koldychev-pce-operational-01 (work in progress), February 2020.

[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>.

[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>.

12.2. Informative References

[RFC4655] Farrel, A., Vasseur, J., 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>.

[RFC8745] Ananthakrishnan, H., Sivabalan, S., Barth, C., Minei, I., and M. Negi, "Path Computation Element Communication Protocol (PCEP) Extensions for Associating Working and Protection Label Switched Paths (LSPs) with Stateful PCE", RFC 8745, DOI 10.17487/RFC8745, March 2020, <https://www.rfc-editor.org/info/rfc8745>.

Authors' Addresses

Mike Koldychev
Cisco Systems, Inc.

Email: mkoldych@cisco.com

Siva Sivabalan
Ciena Corporation

Email: ssivabal@ciena.com

Tarek Saad
Juniper Networks, Inc.

Email: tsaad@juniper.net

Vishnu Pavan Beeram
Juniper Networks, Inc.

Email: vbeeram@juniper.net

Hooman Bidgoli
Nokia

Email: hooman.bidgoli@nokia.com

Bhupendra Yadav
Ciena

Email: byadav@ciena.com

Shuping Peng
Huawei Technologies

Email: pengshuping@huawei.com

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 30, 2021

M. Koldychev
Cisco Systems, Inc.
S. Sivabalan
Ciena Corporation
T. Saad
V. Beeram
Juniper Networks, Inc.
H. Bidgoli
Nokia
B. Yadav
Ciena
S. Peng
Huawei Technologies
October 27, 2020

PCEP Extensions for Signaling Multipath Information
draft-koldychev-pce-multipath-04

Abstract

Current PCEP standards allow only one intended and/or actual path to be present in a PCEP report or update. Applications that require multipath support such as SR Policy require an extension to allow signaling multiple intended and/or actual paths within a single PCEP message. This document introduces such an extension. Encoding of multiple intended and/or actual paths is done by encoding multiple Explicit Route Objects (EROs) and/or multiple Record Route Objects (RROs). A special separator object is defined in this document, to facilitate this. This mechanism is applicable to SR-TE and RSVP-TE and is dataplane agnostic.

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 April 30, 2021.

Table of Contents

1. Introduction

Path Computation Element (PCE) Communication Protocol (PCEP) [RFC5440] enables the communication between a Path Computation Client (PCC) and a Path Control Element (PCE), or between two PCEs based on the PCE architecture [RFC4655].

PCEP Extensions for the Stateful PCE Model [RFC8231] describes a set of extensions to PCEP that enable active control of Multiprotocol Label Switching Traffic Engineering (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.

PCEP Extensions for Segment Routing [RFC8664] specifies extensions to the Path Computation Element Protocol (PCEP) that allow a stateful PCE to compute and initiate Traffic Engineering (TE) paths, as well as for a PCC to request a path subject to certain constraint(s) and optimization criteria in SR networks.

Segment Routing Policy for Traffic Engineering [I-D.ietf-spring-segment-routing-policy] details the concepts of SR Policy and approaches to steering traffic into an SR Policy. In particular, it describes the SR candidate-path as a collection of one or more Segment-Lists. The current PCEP standards only allow for signaling of one Segment-List per Candidate-Path. PCEP extension to support Segment Routing Policy Candidate Paths [I-D.ietf-pce-segment-routing-policy-cp] specifically avoids defining how to signal multipath information, and states that this will be defined in another document.

This document defines the required extensions that allow the signaling of multipath information via PCEP.

2. Terminology

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.

2.1. Terms and Abbreviations

The following terms are used in this document:

PCEP Tunnel:

The object identified by the PLSP-ID, see [I-D.koldychev-pce-operational] for more details.

3. Motivation

This extension is motivated by the use-cases described below.

3.1. Signaling Multiple Segment-Lists of an SR Candidate-Path

The Candidate-Path of an SR Policy is the unit of report/update in PCEP, see [I-D.ietf-pce-segment-routing-policy-cp]. Each Candidate-Path can contain multiple Segment-Lists and each Segment-List is encoded by one ERO. However, each PCEP LSP can contain only a single ERO (containing multiple SR-ERO subobject), which prevents us from encoding multiple Segment- Lists within the same SR Candidate-Path.

With the help of the protocol extensions defined in this document, this limitation is overcome.

3.2. Splitting of Requested Bandwidth

A PCC may request a path with 80 Gbps of bandwidth, but all links in the network have only 50 Gbps capacity. The PCE can return two paths, that can together carry 80 Gbps. The PCC can then equally or unequally split the incoming 80 Gbps of traffic among the two paths. Section 4.3 introduces a new TLV that carries the path weight that allows for distribution of incoming traffic on to the multiple paths.

3.3. Providing Backup path for Protection

It is desirable for the PCE to compute and signal to the PCC a backup path that is used to protect a primary path within the multipaths in a given LSP.

Note that [RFC8745] specify the Path Protection association among LSPs. The use of [RFC8745] with multipath is out of scope of this document and is for future study.

When multipath is used, a backup path may protect one or more primary paths. For this reason, primary and backup path identifiers are needed to indicate which backup path(s) protect which primary

path(s). Section 4.4 introduces a new TLV that carries the required information.

4. Protocol Extensions

4.1. Multipath Capability TLV

We define the MULTIPATH-CAP TLV that MAY be present in the OPEN object and/or the LSP object. The purpose of this TLV is two-fold:

1. From PCC: it tells how many multipaths per PCEP Tunnel, the PCC can install in forwarding.

2. From PCE: it tells that the PCE supports this standard and how many multipaths per PCEP Tunnel, the PCE can compute.

Only the first instance of this TLV can be processed, subsequent instances SHOULD be ignored.

Section 5 specify the usage of this TLV with Open message (within the OPEN object) and other PCEP messages (within the LSP object).

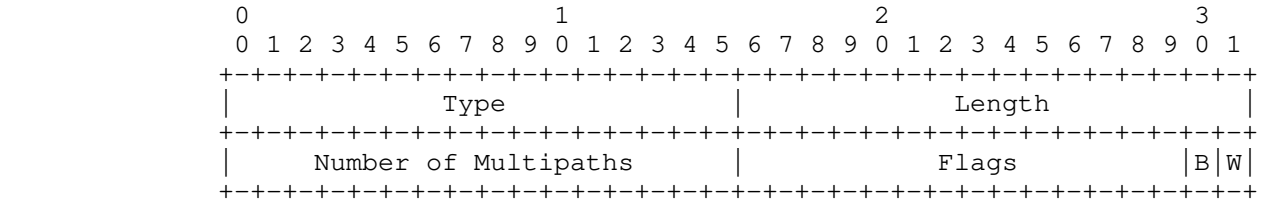


Figure 1: MULTIPATH-CAP TLV format

Type: TBD1 for "MULTIPATH-CAP" TLV.

Length: 4.

Number of Multipaths: the maximum number of multipaths per PCEP Tunnel. The value 0 indicates unlimited number.

Flags: Following bits are defined:

W-flag: whether MULTIPATH-WEIGHT-TLV is supported.

B-flag: whether MULTIPATH-BACKUP-TLV is supported.

Unassigned bits are for future use. They MUST be set to 0 on transmission and MUST be ignored on receipt.

4.2. Path Attributes Object

We define the PATH-ATTRIB object that is used to carry per-path information and to act as a separator between several ERO/RRO objects in the <intended-path>/<actual-path> RBNF element. The PATH-ATTRIB object always precedes the ERO/RRO that it applies to. If multiple ERO/RRO objects are present, then each ERO/RRO object MUST be preceded by an PATH-ATTRIB object that describes it.

The PATH-ATTRIB Object-Class value is TBD2.

The PATH-ATTRIB Object-Type value is 1.

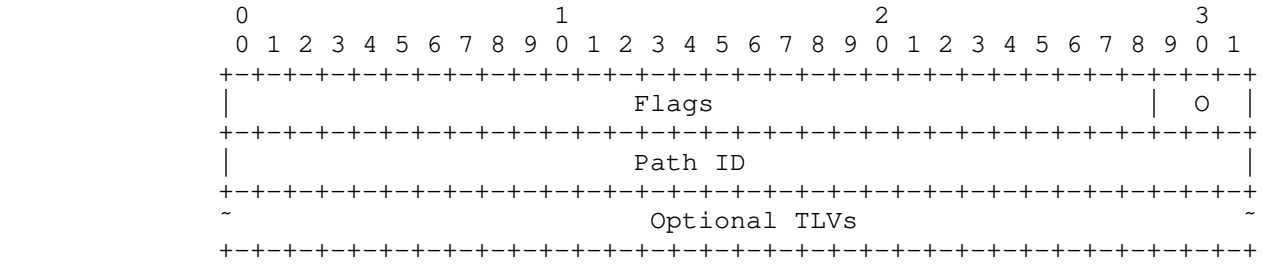


Figure 2: PATH-ATTRIB object format

Flags (32-bits): Following bits are assigned -

O (Operational - 3 bits): operational state of the path, same values as the identically named field in the LSP object {{RFC8231}}.

Unassigned bits are for future use. They MUST be set to 0 on transmission and MUST be ignored on receipt.

Path ID: 4-octet identifier that identifies a path in the set of multiple paths. It uniquely identifies a path (encoded in the ERO/RRO) within the set of multiple paths under the PCEP LSP. Once a path changes, a new Path ID is assigned.

TLVs that may be included in the PATH-ATTRIB object are described in the following sections. Other optional TLVs could be defined by future documents to be included within the PATH-ATTRIB object body.

4.3. Multipath Weight TLV

We define the MULTIPATH-WEIGHT TLV that MAY be present in the PATH-ATTRIB object.

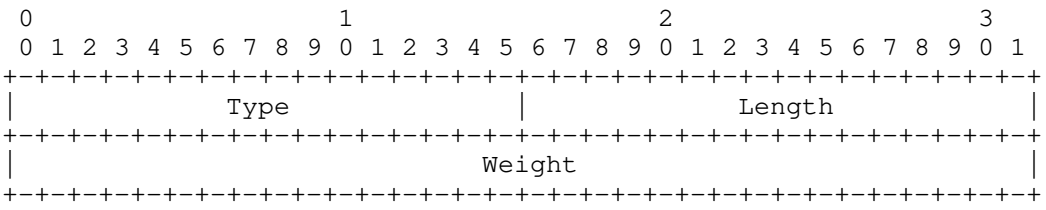


Figure 3: MULTIPATH-WEIGHT TLV format

Type: TBD3 for "MULTIPATH-WEIGHT" TLV.

Length: 4.

Weight: weight of this path within the multipath, if W-ECMP is desired. The fraction of flows a specific ERO/RRO carries is derived from the ratio of its weight to the sum of all other multipath ERO/RRO weights.

When the MULTIPATH-WEIGHT TLV is absent from the PATH-ATTRIB object, or the PATH-ATTRIB object is absent from the <intended-path>/<actual-path>, then the Weight of the corresponding path is taken to be "1".

4.4. Multipath Backup TLV

This document introduces a new MULTIPATH-BACKUP TLV that is optional and can be present in the PATH-ATTRIB object.

This TLV is used to indicate the presence of a backup path that is used for protection in case of failure of the primary path. The format of the MULTIPATH-BACKUP TLV is:

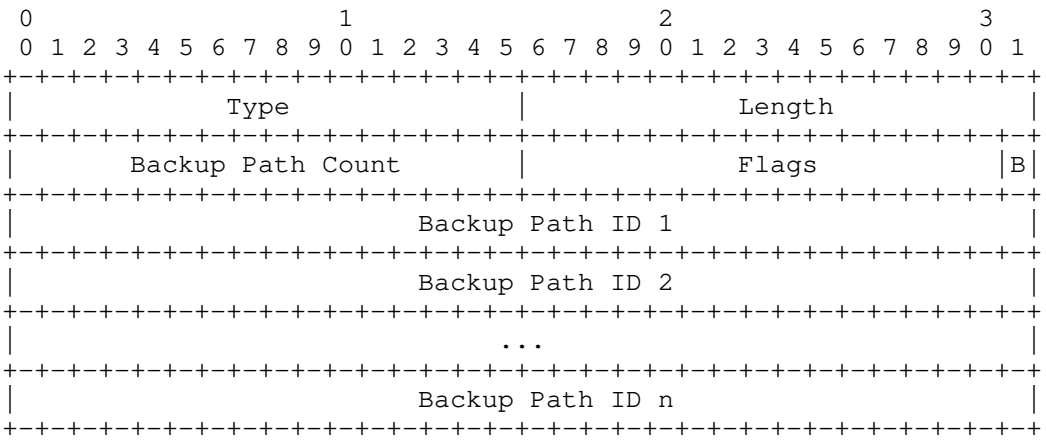


Figure 4: MULTIPATH-BACKUP TLV format

Type: TBD4 for "MULTIPATH-BACKUP" TLV

Length: 4 + (N * 4) (where N is the Backup Path Count)

Backup Path Count: Number of backup path(s).

Flags (16 bits): a flag field. Currently a single flag "B bit" is defined.
Unused flags MUST be set to zero while sending and ignored on receipt.

B: If set, indicates a pure backup path. This is a path that only carries rerouted traffic after the protected path fails. If this flag is not set, or if the MULTIPATH-BACKUP TLV is absent, then the path is assumed to be primary that carries normal traffic.

Backup Path ID(s): a series of 4-octet identifier(s) that identify the backup path(s) in the set that protect this primary path.

5. Operation

When the PCC wants to indicate to the PCE that it wants to get multipaths for a PCEP Tunnel, instead of a single path, it can do (1) or both (1) and (2) of the following:

(1) Send the MULTIPATH-CAP TLV in the OPEN object during session establishment. This applies to all PCEP Tunnels on the PCC, unless overridden by PCEP Tunnel specific information.

(2) Additionally send the MULTIPATH-CAP TLV in the LSP object for a particular PCEP Tunnel in the PCRpt or PCReq message. This applies to the specified PCEP Tunnel and overrides the information from the OPEN object.

When PCE computes the path for a PCEP Tunnel, it MUST NOT return more multipaths than the corresponding value of "Number of Multipaths" from the MULTIPATH-CAP TLV. If this TLV is absent (from both OPEN and LSP objects), then the "Number of Multipaths" is assumed to be 1.

If the PCE supports this standard, then it MUST include the MULTIPATH-CAP TLV in the OPEN object. This tells the PCC that it can report multiple ERO/RRO objects per PCEP Tunnel to this PCE. If the PCE does not include the MULTIPATH-CAP TLV in the OPEN object, then the PCC MUST assume that the PCE does not support this standard and fall back to reporting only a single ERO/RRO. The PCE MUST NOT include MULTIPATH-CAP TLV in the LSP object in any other PCEP message towards the PCC and the PCC MUST ignore it if received.

The Path ID of each ERO/RRO MUST be unique within that LSP. If a PCEP speaker detects that there are two paths with the same Path ID, then the PCEP speaker SHOULD send PCError message with Error-Type = 1 ("Reception of an invalid object") and Error-Value = TBD5 ("Conflicting Path ID").

5.1. Signaling Multiple Paths for Loadbalancing

The PATH-ATTRIB object can be used to signal multiple path(s) and indicate (un)equal loadbalancing amongst the set of multipaths. In this case, the PATH-ATTRIB is populated for each ERO as follows:

1. The PCE assigns a unique Path ID to each ERO path and populates it inside the PATH-ATTRIB object. The Path ID is unique within the context of a PLSP or PCEP Tunnel.

2. The MULTIPATH-WEIGHT TLV MAY be carried inside the PATH-ATTRIB object. A weight is populated to reflect the relative loadshare that is to be carried by the path. If the MULTIPATH-WEIGHT is not carried inside a PATH-ATTRIB object, the default weight 1 MUST be assumed when computing the loadshare.

3. The fraction of flows carried by a specific primary path is derived from the ratio of its weight to the sum of all other multipath weights.

5.2. Signaling Multiple Paths for Protection

The PATH-ATTRIB object can be used to describe a set of backup path(s) protecting a primary path within a PCEP Tunnel. In this case, the PATH-ATTRIB is populated for each ERO as follows:

1. The PCE assigns a unique Path ID to each ERO path and populates it inside the PATH-ATTRIB object. The Path ID is unique within the context of a PLSP or PCEP Tunnel.

2. The MULTIPATH-BACKUP TLV MUST be added inside the PATH-ATTRIB object for each ERO that is protected. The backup path ID(s) are populated in the MULTIPATH-BACKUP TLV to reflect the set of backup path(s) protecting the primary path. The Length field and Backup Path Number in the MULTIPATH-BACKUP are updated according to the number of backup path ID(s) included.

3. The MULTIPATH-BACKUP TLV MAY be added inside the PATH-ATTRIB object for each ERO that is unprotected. In this case, MULTIPATH-BACKUP does not carry any backup path IDs in the TLV. If the path acts as a pure backup - i.e. the path only carries rerouted traffic after the protected path(s) fail- then the B flag MUST be set.

Note that if a given path has the B-flag set, then there MUST be some other path within the same LSP that uses the given path as a backup. If this condition is violated, then the PCEP speaker SHOULD send a PCError message with Error-Type = 10 ("Reception of an invalid object") and Error-Value = TBD6 ("No primary path for pure backup").

Note that a given PCC may not support certain backup combinations, such as a backup path that is itself protected by another backup path, etc. If a PCC is not able to implement a requested backup scenario, the PCC SHOULD send a PCError message with Error-Type = 19 ("Invalid Operation") and Error-Value = TBD7 ("Not supported path backup").

6. PCEP Message Extensions

The RBNF of PCReq, PCRep, PCRpt, PCUpd and PCInit messages currently use a combination of <intended-path> and/or <actual-path>. As specified in Section 6.1 of [RFC8231], <intended-path> is represented by the ERO object and <actual-path> is represented by the RRO object:

```
<intended-path> ::= <ERO>

<actual-path> ::= <RRO>
```

In this standard, we extend these two elements to allow multiple ERO/ RRO objects to be present in the <intended-path>/<actual-path>:

```
<intended-path> ::= (<ERO>|
                     (<PATH-ATTRIB><ERO>)
                     [<intended-path>])


<actual-path> ::= (<RRO>|
                   (<PATH-ATTRIB><RRO>)
                   [<actual-path>])
```

7. Examples

7.1. SR Policy Candidate-Path with Multiple Segment-Lists

Consider how the following sample SR Policy, taken from [I-D.ietf-spring-segment-routing-policy], would be represented in a PCRpt message.

```
        SR policy POL1 <headend, color, endpoint>
            Candidate-path CP1 <protocol-origin = 20, originator =
    100:1.1.1.1, discriminator = 1>
                Preference 200
                Weight W1, SID-List1 <SID11...SID1i>
                Weight W2, SID-List2 <SID21...SID2j>
            Candidate-path CP2 <protocol-origin = 20, originator =
    100:2.2.2.2, discriminator = 2>
                Preference 100
                Weight W3, SID-List3 <SID31...SID3i>
                Weight W4, SID-List4 <SID41...SID4j>
```

As specified in [I-D.ietf-pce-segment-routing-policy-cp], CP1 and CP2 are signaled as separate state-report elements and each has a unique PLSP-ID, assigned by the PCC. Let us assign PLSP-ID 100 to CP1 and PLSP-ID 200 to CP2.

The state-report for CP1 can be encoded as:

```
<state-report> =
        <LSP PLSP_ID=100>
        <ASSOCIATION>
        <END-POINT>
        <PATH-ATTRIB Path_ID=1 <WEIGHT-TLV Weight=W1>>
        <ERO SID-List1>
        <PATH-ATTRIB Path_ID=2 <WEIGHT-TLV Weight=W2>>
        <ERO SID-List2>
```

The state-report for CP2 can be encoded as:

```
<state-report> =
        <LSP PLSP_ID=200>
        <ASSOCIATION>
        <END-POINT>
        <PATH-ATTRIB Path_ID=1 <WEIGHT-TLV Weight=W3>>
        <ERO SID-List3>
        <PATH-ATTRIB Path_ID=2 <WEIGHT-TLV Weight=W4>>
        <ERO SID-List4>
```

The above sample state-report elements only specify the minimum mandatory objects, of course other objects like SRP, LSPA, METRIC, etc., are allowed to be inserted.

Note that the syntax

```
<PATH-ATTRIB Path_ID=1 <WEIGHT-TLV Weight=W1>>
```

, simply means that this is PATH-ATTRIB object with Path ID field set to "1" and with a MULTIPATH-WEIGHT TLV carrying weight of "W1".

7.2. Two Primary Paths Protected by One Backup Path

Suppose there are 3 paths: A, B, C. Where A,B are primary and C is to be used only when A or B fail. Suppose the Path IDs for A, B, C are respectively 1, 2, 3. This would be encoded in a state-report as:

```
<state-report> =
        <LSP>
        <ASSOCIATION>
        <END-POINT>
        <PATH-ATTRIB Path_ID=1 <BACKUP-TLV B=0, Backup_Paths=[3]>>
        <ERO A>
        <PATH-ATTRIB Path_ID=2 <BACKUP-TLV B=0, Backup_Paths=[3]>>
        <ERO B>
        <PATH-ATTRIB Path_ID=3 <BACKUP-TLV B=1, Backup_Paths=[]>>
        <ERO C>
```

Note that the syntax

```
<PATH-ATTRIB Path_ID=1 <BACKUP-TLV B=0, Backup_Paths=[3]>>
```

, simply means that this is PATH-ATTRIB object with Path ID field set to "1" and with a MULTIPATH-BACKUP TLV that has B-flag cleared and contains a single backup path with Backup Path ID of 3.

8. IANA Considerations

8.1. PCEP Object

IANA is requested to make the assignment of a new value for the existing "PCEP Objects" registry as follows:

Object-Class Value	Name	Object-Type Value	Reference
TBD2	PATH-ATTRIB	1	This document

8.2. PCEP TLV

IANA is requested to make the assignment of a new value for the existing "PCEP TLV Type Indicators" registry as follows:

TLV Type Value	TLV Name	Reference
TBD1	MULTIPATH-CAP	This document
TBD3	MULTIPATH-WEIGHT	This document
TBD4	MULTIPATH-BACKUP	This document

8.3. PCEP-Error Object

IANA is requested to make the assignment of a new value for the existing "PCEP-ERROR Object Error Types and Values" sub-registry of the PCEP Numbers registry for the following errors:

Error-Type	Error-Value	Reference
10	TBD5 - Conflicting Path ID	This document
10	TBD6 - No primary path for pure backup	This document
19	TBD7 - Not supported path backup	This document

8.4. Flags in the Multipath Capability TLV

IANA is requested to create a new sub-registry to manage the Flag field of the MULTIPATH-CAP TLV, called "Flags in MULTIPATH-CAP TLV".

Following bits are defined:

Bit	Description	Reference
0-13	Unassigned	This document
14	B-flag: Backup support	This document
15	W-flag: Weighted ECMP support	This document

8.5. Flags in the Path Attribute Object

IANA is requested to create a new sub-registry to manage the Flag field of the PATH-ATTRIBUTE object, called "Flags in PATH-ATTRIBUTE Object".

Following bits are defined:

Bit	Description	Reference
0-12	Unassigned	This document
13-15	O-flag: Operational state	This document

8.6. Flags in the Multipath Backup TLV

IANA is requested to create a new sub-registry to manage the Flag field of the MULTIPATH-BACKUP TLV, called "Flags in MULTIPATH-BACKUP TLV".

Following bits are defined:

Bit	Description	Reference
0-14	Unassigned	This document
15	B-flag: Pure backup	This document

9. Security Considerations

None at this time.

10. Acknowledgement

Thanks to Dhruv Dhody for ideas and discussion.

11. Contributors

Andrew Stone
Nokia

Email: andrew.stone@nokia.com

12. References

12.1. Normative References

[I-D.ietf-pce-segment-routing-policy-cp]
Koldychev, M., Sivabalan, S., Barth, C., Peng, S., and H. Bidgoli, "PCEP extension to support Segment Routing Policy Candidate Paths", draft-ietf-pce-segment-routing-policy-cp-00 (work in progress), June 2020.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-08 (work in progress), July 2020.

[I-D.koldychev-pce-operational]
Koldychev, M., Sivabalan, S., Negi, M., Achaval, D., and H. Kotni, "PCEP Operational Clarification", draft-koldychev-pce-operational-02 (work in progress), August 2020.

[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>.

[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>.

12.2. Informative References

[RFC4655] Farrel, A., Vasseur, J., 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>.

[RFC8745] Ananthakrishnan, H., Sivabalan, S., Barth, C., Minei, I., and M. Negi, "Path Computation Element Communication Protocol (PCEP) Extensions for Associating Working and Protection Label Switched Paths (LSPs) with Stateful PCE", RFC 8745, DOI 10.17487/RFC8745, March 2020, <https://www.rfc-editor.org/info/rfc8745>.

Authors' Addresses

Mike Koldychev
Cisco Systems, Inc.

Email: mkoldych@cisco.com

Siva Sivabalan
Ciena Corporation

Email: ssivabal@ciena.com

Tarek Saad
Juniper Networks, Inc.

Email: tsaad@juniper.net

Vishnu Pavan Beeram
Juniper Networks, Inc.

Email: vbeeram@juniper.net

Hooman Bidgoli
Nokia

Email: hooman.bidgoli@nokia.com

Bhupendra Yadav
Ciena

Email: byadav@ciena.com

Shuping Peng
Huawei Technologies

Email: pengshuping@huawei.com

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 6, 2021

S. Peng
C. Li
Huawei Technologies
L. Han
China Mobile
July 5, 2020

Support for Path MTU (PMTU) in the Path Computation Element (PCE) communication Protocol (PCEP). draft-li-pce-pcep-pmtu-01

Abstract

The Path Computation Element (PCE) provides path computation functions in support of traffic engineering in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks.

The Source Packet Routing in Networking (SPRING) architecture describes how Segment Routing (SR) can be used to steer packets through an IPv6 or MPLS network using the source routing paradigm. A Segment Routed Path can be derived from a variety of mechanisms, including an IGP Shortest Path Tree (SPT), explicit configuration, or a Path Computation Element (PCE).

Since the SR does not require signaling, the path maximum transmission unit (MTU) information for SR path is not available. This document specify the extension to PCE communication protocol (PCEP) to carry path (MTU) in the PCEP messages.

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.

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/.

Table of Contents

1. Introduction

[RFC5440] describes the Path Computation Element (PCE) Communication Protocol (PCEP). PCEP enables the communication between a Path Computation Client (PCC) and a PCE, or between PCE and PCE, for the purpose of computation of Multiprotocol Label Switching (MPLS) as well as Generalzied MPLS (GMPLS) Traffic Engineering Label Switched Path (TE LSP) characteristics.

[RFC8231] specifies a set of extensions to PCEP to enable stateful control of TE LSPs within and across PCEP sessions in compliance with [RFC4657]. It includes mechanisms to effect LSP State Synchronization between PCCs and PCEs, delegation of control over LSPs to PCEs, and PCE control of timing and sequence of path computations within and across PCEP sessions. The model of operation where LSPs are initiated from the PCE is described in [RFC8281].

As per [RFC8402], with Segment Routing (SR), a node steers a packet through an ordered list of instructions, called segments. A segment can represent any instruction, topological or service-based. A segment can have a semantic local to an SR node or global within an SR domain. SR allows to enforce a flow through any path and service chain while maintaining per-flow state only at the ingress node of the SR domain. Segments can be derived from different components: IGP, BGP, Services, Contexts, Locators, etc. The SR architecture can be applied to the MPLS forwarding plane without any change, in which case an SR path corresponds to an MPLS Label Switching Path (LSP). The SR is applied to IPV6 forwarding plane using SRH. A SR path can be derived from an IGP Shortest Path Tree (SPT), but SR-TE paths may not follow IGP SPT. Such paths may be chosen by a suitable network planning tool, or a PCE and provisioned on the ingress node.

As per [RFC8664], it is possible to use a stateful PCE for computing one or more SR-TE paths taking into account various constraints and objective functions. Once a path is chosen, the stateful PCE can initiate an SR-TE path on a PCC using PCEP extensions specified in [RFC8281] using the SR specific PCEP extensions specified in [RFC8664]. [RFC8664] specifies PCEP extensions for supporting a SR-TE LSP for MPLS data plane. [I-D.ietf-pce-segment-routing-ipv6] extend PCEP to support SR for IPv6 data plane.

The maximum transmission unit (MTU) is the largest size packet or frame, in bytes, that can be sent in a network. An MTU that is too large might cause retransmissions. Too small an MTU might cause the router to send and handle relatively more header overhead and acknowledgments. When an LSP is created across a set of links with different MTU sizes, the ingress router need to know what the smallest MTU is on the LSP path. If this MTU is larger than the MTU of one of the intermediate links, traffic might be dropped, because MPLS packets cannot be fragmented. Also, the ingress router may not be aware of this type of traffic loss, because the control plane for the LSP would still function normally. [RFC3209] specify the mechanism of MTU signaling in RSVP.

Since the SR does not require signaling, the path MTU information for SR path is not available. This document specify the extension to PCEP to carry path MTU in the PCEP messages. It is assumed that the

PCE is aware of the link MTU as part of the Traffic Engineering Database (TED) population. This could be done via IGP, BGP-LS or some other means. Thus the PCE can find the path MTU at the time of path computation and include this information as part of the PCEP messages.

Though the key use case for path MTU is SR, the PCEP extension (as specified in this document) creates a new metric type for path MTU, making this a generic extension that can be used independent of SR.

2. PCEP Extention

2.1. Extensions to METRIC Object

The METRIC object is defined in Section 7.8 of [RFC5440], comprising metric-value and metric-type (T field), and a flags field, comprising a number of bit flags (B bit and C bit). This document defines a new type for the METRIC object for Path MTU.

- T = TBD: Path MTU.
- A network comprises of a set of N links {Li, (i=1...N)}.
- A path P of a LSP is a list of K links {Lpi,(i=1...K)}.
- A Link MTU of link L is denoted M(L).
- A Path MTU metric for the path P = Min {M(Lpi), (i=1...K)}.

The Path MTU metric type of the METRIC object in PCEP represents the minimum of the Link MTU of all links along the path.

When PCE computes the path, it can also find the Path MTU (based on the above criteria) and include this information in the METRIC object with the above metric type in the PCEP message when replying to the PCC. In a Path Computation Reply (PCRep) message, the PCE MAY insert the METRIC object with an Explicit Route Object (ERO) so as to provide the METRIC (path MTU) for the computed path. The PCE MAY also insert the METRIC object with a NO-PATH object to indicate that the metric constraint could not be satisfied.

Further, a PCC MAY use the Path MTU metric in a Path Computation Request (PCReq) message to request a path meeting the MTU requirement of the path. In this case, the B bit MUST be set to suggest a bound (a maximum) for the Path MTU metric that must not be exceeded for the PCC to consider the computed path as acceptable. The Path MTU metric must be less than or equal to the value specified in the metric-value field.

A PCC can also use this metric to ask PCE to optimize the path MTU during path computation. In this case, the B bit MUST be cleared.

The error handling and processing of the METRIC object is as specified in [RFC5440].

2.2. Stateful PCE and PCE Initiated LSPs

[RFC8231] specifies a set of extensions to PCEP to enable stateful control of MPLS-TE and GMPLS LSPs via PCEP and the maintaining of these LSPs at the stateful PCE. It further distinguishes between an active and a passive stateful PCE. A passive stateful PCE uses LSP state information learned from PCCs to optimize path computations but does not actively update LSP state. In contrast, an active stateful PCE utilizes the LSP delegation mechanism to update LSP parameters in those PCCs that delegated control over their LSPs to the PCE. [RFC8281] describes the setup, maintenance, and teardown of PCE-initiated LSPs under the stateful PCE model. The document defines the PCInitiate message that is used by a PCE to request a PCC to set up a new LSP.

The new metric type defined in this document can also be used with the stateful PCE extensions. The format of PCEP messages described in [RFC8231] and [RFC8281] uses <intended-attribute-list> and <attribute-list>, respectively, (where the <intended-attribute-list> is the attribute-list defined in Section 6.5 of [RFC5440].

A PCE MAY include the path MTU metric in PCInitiate or PCUpd message to inform the PCC of the path MTU calculated for the path. A PCC MAY include the path MTU metric as a bound constraint or to indicate optimization criteria (similar to PCReq).

2.3. Segement Routing

A Segment Routed path (SR path) can be derived from an IGP Shortest Path Tree (SPT). Segment Routed Traffic Engineering paths (SR-TE paths) may not follow IGP SPT. Such paths may be chosen by a suitable network planning tool and provisioned on the source node of the SR-TE path.

It is possible to use a PCE for computing one or more SR-TE paths taking into account various constraints and objective functions. Once a path is chosen, the PCE can inform an SR-TE path on a PCC using PCEP extensions specified in [RFC8664]. Further, [I-D.ietf-pce-segment-routing-ipv6] adds the support for IPv6 data plane in SR.

The new metric type for path MTU is applicable for the SR-TE path and require no additional extensions.

3. Security Considerations

This document defines a new METRIC type that do not add any new security concerns beyond those discussed in [RFC5440] in itself. Some deployments may find the path MTU information to be extra sensitive and could be used to influence path computation and setup with adverse effect. Additionally, snooping of PCEP messages with such data or using PCEP messages for network reconnaissance may give an attacker sensitive information about the operations of the network. Thus, such deployment should employ suitable PCEP security mechanisms like TCP Authentication Option (TCP-AO) [RFC5925] or Transport Layer Security (TLS) [RFC8253]. The procedure based on TLS is considered a security enhancement and thus is much better suited for the sensitive information.

4. IANA Considerations

This document makes following requests to IANA for action.

4.1. METRIC Types

IANA maintains the "Path Computation Element Protocol (PCEP) Numbers" registry. Within this registry, IANA maintains a subregistry for "METRIC Object T Field". IANA is requested to make the following allocation:

Value	Description	Reference
TBD	Path MTU.	This document

5. Acknowledgments

We would like to thank Dhruv Dhody for his contributions for this document.

6. References

6.1. 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>.

6.2. Informative References

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[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>.

[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>.

[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>.

[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>.

[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>.

[I-D.ietf-pce-segment-routing-ipv6] Li, C., Negi, M., Koldychev, M., Kaladharan, P., and Y. Zhu, "PCEP Extensions for Segment Routing leveraging the IPv6 data plane", draft-ietf-pce-segment-routing-ipv6-06 (work in progress), July 2020.

Authors' Addresses

Shuping Peng
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: pengshuping@huawei.com

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

Email: c.l@huawei.com

Liuyan Han
China Mobile
Beijing 100053
China

Email: hanliuyan@chinamobile.com

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 4, 2021

S. Peng
C. Li
Huawei Technologies
L. Han
China Mobile
L. Ndifor
MTN Cameroon
October 31, 2020

Support for Path MTU (PMTU) in the Path Computation Element (PCE) communication Protocol (PCEP). draft-li-pce-pcep-pmtu-03

Abstract

The Path Computation Element (PCE) provides path computation functions in support of traffic engineering in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks.

The Source Packet Routing in Networking (SPRING) architecture describes how Segment Routing (SR) can be used to steer packets through an IPv6 or MPLS network using the source routing paradigm. A Segment Routed Path can be derived from a variety of mechanisms, including an IGP Shortest Path Tree (SPT), explicit configuration, or a Path Computation Element (PCE).

Since the SR does not require signaling, the path maximum transmission unit (MTU) information for SR path is not available. This document specify the extension to PCE communication protocol (PCEP) to carry path (MTU) in the PCEP messages.

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.

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 May 4, 2021.

Table of Contents

1. Introduction

[RFC5440] describes the Path Computation Element (PCE) Communication Protocol (PCEP). PCEP enables the communication between a Path Computation Client (PCC) and a PCE, or between PCE and PCE, for the

purpose of computation of Multiprotocol Label Switching (MPLS) as well as Generalzied MPLS (GMPLS) Traffic Engineering Label Switched Path (TE LSP) characteristics.

[RFC8231] specifies a set of extensions to PCEP to enable stateful control of TE LSPs within and across PCEP sessions in compliance with [RFC4657]. It includes mechanisms to effect LSP State Synchronization between PCCs and PCEs, delegation of control over LSPs to PCEs, and PCE control of timing and sequence of path computations within and across PCEP sessions. The model of operation where LSPs are initiated from the PCE is described in [RFC8281].

As per [RFC8402], with Segment Routing (SR), a node steers a packet through an ordered list of instructions, called segments. A segment can represent any instruction, topological or service-based. A segment can have a semantic local to an SR node or global within an SR domain. SR allows to enforce a flow through any path and service chain while maintaining per-flow state only at the ingress node of the SR domain. Segments can be derived from different components: IGP, BGP, Services, Contexts, Locators, etc. The SR architecture can be applied to the MPLS forwarding plane without any change, in which case an SR path corresponds to an MPLS Label Switching Path (LSP). The SR is applied to IPV6 forwarding plane using SRH. A SR path can be derived from an IGP Shortest Path Tree (SPT), but SR-TE paths may not follow IGP SPT. Such paths may be chosen by a suitable network planning tool, or a PCE and provisioned on the ingress node.

As per [RFC8664], it is possible to use a stateful PCE for computing one or more SR-TE paths taking into account various constraints and objective functions. Once a path is chosen, the stateful PCE can initiate an SR-TE path on a PCC using PCEP extensions specified in [RFC8281] using the SR specific PCEP extensions specified in [RFC8664]. [RFC8664] specifies PCEP extensions for supporting a SR-TE LSP for MPLS data plane. [I-D.ietf-pce-segment-routing-ipv6] extend PCEP to support SR for IPv6 data plane.

The maximum transmission unit (MTU) is the largest size packet or frame, in bytes, that can be sent in a network. An MTU that is too large might cause retransmissions. Too small an MTU might cause the router to send and handle relatively more header overhead and acknowledgments. When an LSP is created across a set of links with different MTU sizes, the ingress router need to know what the smallest MTU is on the LSP path. If this MTU is larger than the MTU of one of the intermediate links, traffic might be dropped, because MPLS packets cannot be fragmented. Also, the ingress router may not be aware of this type of traffic loss, because the control plane for the LSP would still function normally. [RFC3209] specify the mechanism of MTU signaling in RSVP.

Since the SR does not require signaling, the path MTU information for SR path is not available. This document specify the extension to PCEP to carry path MTU in the PCEP messages. It is assumed that the PCE is aware of the link MTU as part of the Traffic Engineering Database (TED) population. This could be done via IGP, BGP-LS or some other means. Thus the PCE can find the path MTU at the time of path computation and include this information as part of the PCEP messages.

Though the key use case for path MTU is SR, the PCEP extension (as specified in this document) creates a new metric type for path MTU, making this a generic extension that can be used independent of SR.

2. Terminology

This draft refers to the terms defined in [RFC8201], [RFC4821] and [RFC3988].

MTU: Maximum Transmission Unit, the size in bytes of the largest IP packet, including the IP header and payload, that can be transmitted on a link or path. Note that this could more properly be called the IP MTU, to be consistent with how other standards organizations use the acronym MTU.

Link MTU: The Maximum Transmission Unit, i.e., maximum IP packet size in bytes, that can be conveyed in one piece over a link. Be aware that this definition is different from the definition used by other standards organizations.

For IETF documents, link MTU is uniformly defined as the IP MTU over the link. This includes the IP header, but excludes link layer headers and other framing that is not part of IP or the IP payload.

Be aware that other standards organizations generally define link MTU to include the link layer headers.

For the MPLS data plane, this size includes the IP header and data (or other payload) and the label stack but does not include any lower-layer headers. A link may be an interface (such as Ethernet or Packet-over-SONET), a tunnel (such as GRE or IPsec), or an LSP.

Path: The set of links traversed by a packet between a source node and a destination node.

Path MTU, or PMTU: The minimum link MTU of all the links in a path between a source node and a destination node.

For the MPLS data plane, it is the MTU of an LSP from a given LSR to the egress(es), over each valid (forwarding) path. This size includes the IP header and data (or other payload) and any part of the label stack that was received by the ingress LSR before it placed the packet into the LSP (this part of the label stack is considered part of the payload for this LSP). The size does not include any lower-level headers.

3. PCEP Extention

3.1. Extensions to METRIC Object

The METRIC object is defined in Section 7.8 of [RFC5440], comprising metric-value and metric-type (T field), and a flags field, comprising a number of bit flags (B bit and C bit). This document defines a new type for the METRIC object for Path MTU.

- T = TBD: Path MTU.
- A network comprises of a set of N links {Li, (i=1...N)}.
- A path P of a LSP is a list of K links {Lpi,(i=1...K)}.
- A Link MTU of link L is denoted M(L).
- A Path MTU metric for the path P = Min {M(Lpi), (i=1...K)}.

The Path MTU metric type of the METRIC object in PCEP represents the minimum of the Link MTU of all links along the path.

When PCE computes the path, it can also find the Path MTU (based on the above criteria) and include this information in the METRIC object with the above metric type in the PCEP message when replying to the PCC. In a Path Computation Reply (PCRep) message, the PCE MAY insert the METRIC object with an Explicit Route Object (ERO) so as to provide the METRIC (path MTU) for the computed path. The PCE MAY also insert the METRIC object with a NO-PATH object to indicate that the metric constraint could not be satisfied.

Further, a PCC MAY use the Path MTU metric in a Path Computation Request (PCReq) message to request a path meeting the MTU requirement of the path. In this case, the B bit MUST be set to suggest a bound (a maximum) for the Path MTU metric that must not be exceeded for the PCC to consider the computed path as acceptable. The Path MTU metric must be less than or equal to the value specified in the metric-value field.

A PCC can also use this metric to ask PCE to optimize the path MTU during path computation. In this case, the B bit MUST be cleared.

The error handling and processing of the METRIC object is as specified in [RFC5440].

3.2. Stateful PCE and PCE Initiated LSPs

[RFC8231] specifies a set of extensions to PCEP to enable stateful control of MPLS-TE and GMPLS LSPs via PCEP and the maintaining of these LSPs at the stateful PCE. It further distinguishes between an active and a passive stateful PCE. A passive stateful PCE uses LSP state information learned from PCCs to optimize path computations but does not actively update LSP state. In contrast, an active stateful PCE utilizes the LSP delegation mechanism to update LSP parameters in those PCCs that delegated control over their LSPs to the PCE. [RFC8281] describes the setup, maintenance, and teardown of PCE-initiated LSPs under the stateful PCE model. The document defines

the PCInitiate message that is used by a PCE to request a PCC to set up a new LSP.

The new metric type defined in this document can also be used with the stateful PCE extensions. The format of PCEP messages described in [RFC8231] and [RFC8281] uses <intended-attribute-list> and <attribute-list>, respectively, (where the <intended-attribute-list> is the attribute-list defined in Section 6.5 of [RFC5440].

A PCE MAY include the path MTU metric in PCInitiate or PCUpd message to inform the PCC of the path MTU calculated for the path. A PCC MAY include the path MTU metric as a bound constraint or to indicate optimization criteria (similar to PCReq).

3.3. Segment Routing

A Segment Routed path (SR path) can be derived from an IGP Shortest Path Tree (SPT). Segment Routed Traffic Engineering paths (SR-TE paths) may not follow IGP SPT. Such paths may be chosen by a suitable network planning tool and provisioned on the source node of the SR-TE path.

It is possible to use a PCE for computing one or more SR-TE paths taking into account various constraints and objective functions. Once a path is chosen, the PCE can inform an SR-TE path on a PCC using PCEP extensions specified in [RFC8664]. Further, [I-D.ietf-pce-segment-routing-ipv6] adds the support for IPv6 data plane in SR.

The new metric type for path MTU is applicable for the SR-TE path and require no additional extensions.

3.4. Path MTU Adjustment

The path MTU metric can be used for both primary and protection path.

The minimal value of the link MTU along the path is collected, based on which minor adjustment is made to cater for overhead introduced by the protection mechanisms such as TI-LFA. The path MTU is the value of the minimum link MTU minus the overhead. In this way, the ingress node can use the path MTU directly.

4. Security Considerations

This document defines a new METRIC type that do not add any new security concerns beyond those discussed in [RFC5440] in itself. Some deployments may find the path MTU information to be extra sensitive and could be used to influence path computation and setup

with adverse effect. Additionally, snooping of PCEP messages with such data or using PCEP messages for network reconnaissance may give an attacker sensitive information about the operations of the network. Thus, such deployment should employ suitable PCEP security mechanisms like TCP Authentication Option (TCP-AO) [RFC5925] or Transport Layer Security (TLS) [RFC8253]. The procedure based on TLS is considered a security enhancement and thus is much better suited for the sensitive information.

5. IANA Considerations

This document makes following requests to IANA for action.

5.1. METRIC Type

IANA maintains the "Path Computation Element Protocol (PCEP) Numbers" registry. Within this registry, IANA maintains a subregistry for "METRIC Object T Field". IANA is requested to make the following allocation:

Value	Description	Reference
TBD	Path MTU	This document

6. Acknowledgement

We would like to thank Dhruv Dhody for his contributions for this document.

7. References

7.1. 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>.

7.2. Informative References

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[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>.

[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>.

[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>.

[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>.

[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>.

[I-D.ietf-pce-segment-routing-ipv6]
Li, C., Negi, M., Koldychev, M., Kaladharan, P., and Y. Zhu, "PCEP Extensions for Segment Routing leveraging the IPv6 data plane", draft-ietf-pce-segment-routing-ipv6-06 (work in progress), July 2020.

Authors' Addresses

Shuping Peng
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: pengshuping@huawei.com

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

Email: c.l@huawei.com

Liuyan Han
China Mobile
Beijing 100053
China

Email: hanliuyan@chinamobile.com

Luc-Fabrice Ndifor
MTN Cameroon
Cameroon

Email: Luc-Fabrice.Ndifor@mtn.com

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 9, 2021

C. Li
Huawei Technologies
S. Sivabalan
Ciena Corporation
S. Peng
Huawei Technologies
M. Koldychev
Cisco Systems, Inc.
L. Ndifor
MTN Cameroon

July 8, 2020

A YANG Data Model for Segment Routing in IPv6 (SRv6) support in Path Computation Element Communications Protocol (PCEP) draft-li-pce-pcep-srv6-yang-01

Abstract

This document augments a YANG data model for the management of Path Computation Element communications Protocol (PCEP) for communications between a Path Computation Client (PCC) and a Path Computation Element (PCE), or between two PCEs in support for Segment Routing in IPv6. The data model includes configuration data and state data (status information and counters for the collection of statistics).

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 9, 2021.

Table of Contents

1. Introduction

The Path Computation Element (PCE) defined in [RFC4655] is an entity that is capable of computing a network path or route based on a network graph, and applying computational constraints. A Path Computation Client (PCC) may make requests to a PCE for paths to be computed.

PCEP is the communication protocol between a PCC and PCE and is defined in [RFC5440]. PCEP interactions include path computation requests and path computation replies as well as notifications of specific states related to the use of a PCE in the context of Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineering (TE). [RFC8231] specifies extensions to PCEP to enable stateful control of MPLS TE LSPs.

[I-D.ietf-pce-segment-routing-ipv6] extends [RFC8664] to support SR for IPv6 data plane.

[I-D.ietf-pce-pcep-yang] defines a YANG [RFC7950] data model for the management of PCEP speakers. This document contains a specification of the PCEP-SRv6 YANG module, "ietf-pcep-srv6" which provides the PCEP-SRv6 [I-D.ietf-pce-segment-routing-ipv6] data model.

The PCEP operational state is included in the same tree as the PCEP configuration consistent with Network Management Datastore Architecture [RFC8342]. The origin of the data is indicated as per the origin metadata annotation.

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. Terminology and Notation

This document also uses the following terms defined in [RFC7420]:

- PCEP entity: a local PCEP speaker.
- PCEP peer: to refer to a remote PCEP speaker.
- PCEP speaker: where it is not necessary to distinguish between local and remote.

Further, this document also uses the following terms defined in [RFC8231] :

- Stateful PCE, Passive Stateful PCE, Active Stateful PCE
- Delegation, Revocation, Redelegation
- LSP State Report, Path Computation Report message (PCRpt).
- LSP State Update, Path Computation Update message (PCUpd).

[RFC8281] :

- PCE-initiated LSP, Path Computation LSP Initiate Message (PCInitiate).

[RFC8408] :

- Path Setup Type (PST).

[RFC8664] :

- Segment Routing (SR).

[I-D.ietf-pce-segment-routing-ipv6] :

- Segment Routing in IPv6 (SRv6).

3.1. Tree Diagrams

A simplified graphical representation of the data model is used in this document. The meaning of the symbols in these diagrams is defined in [RFC8340].

3.2. Prefixes in Data Node Names

In this document, names of data nodes and other data model objects are often used without a prefix, as long as it is clear from the context in which YANG module each name is defined. Otherwise, names are prefixed using the standard prefix associated with the corresponding YANG module, as shown in Table 1.

Prefix	YANG module	Reference
te-types	ietf-te-types	[RFC8776]
pcep	ietf-pcep	[I-D.ietf-pce-pcep-yang]
srv6-types	ietf-srv6-types	[I-D.raza-spring-srv6-yang]

Table 1: Prefixes and corresponding YANG modules

4. The Design of PCEP-SRv6 Data Model

4.1. The Overview of PCEP SRv6 Data Model

The PCEP-SRv6 YANG module defined in this document has all the common building blocks for the PCEP-SRv6 extension.

```
module: ietf-pcep-srv6
  augment /pcep:pcep/pcep:entity/pcep:capability:
    +--rw srv6 {srv6}?
       +--rw enabled?     boolean
       +--rw msd-limit?   boolean
       +--rw srv6-msd* [msd-type]
          +--rw msd-type     uint8
          +--rw msd-value?   uint8
  augment /pcep:pcep/pcep:entity/pcep:peers/pcep:peer
            /pcep:capability:
    +--rw srv6 {srv6}?
       +--rw enabled?     boolean
       +--rw msd-limit?   boolean
       +--rw srv6-msd* [msd-type]
          +--rw msd-type     uint8
          +--rw msd-value?   uint8
  augment /pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp:
    +--ro srv6 {srv6}?
       +--ro segment-list
          +--ro segment* [index]
             +--ro index        uint32
             +--ro sid-value?   srv6-types:srv6-sid
```

5. PCEP-SRv6 YANG Modules

5.1. ietf-pcep-srv6 module

RFC Ed.: In this section, replace all occurrences of 'XXXX' with the actual RFC number and all occurrences of the revision date below with the date of RFC publication (and remove this note).

```
<CODE BEGINS> file "ietf-pcep-srv6@2020-07-08.yang"
module ietf-pcep-srv6 {
  yang-version 1.1;
  namespace "urn:ietf:params:xml:ns:yang:ietf-pcep-srv6";
  prefix pcep-srv6;

  import ietf-srv6-types {
    prefix srv6-types;
    reference
      "I-D.raza-spring-srv6-yang: YANG Data Model for SRv6
       Base and Static";
  }
  import ietf-te-types {
    prefix te-types;
    reference
```

```
        "RFC 8776: Common YANG Data Types for Traffic Engineering";
    }
    import ietf-pcep {
      prefix pcep;
      reference
        "I-D.ietf-pce-pcep-yang: A YANG Data Model for Path
         Computation Element Communications Protocol (PCEP)";
    }

    organization
      "IETF PCE (Path Computation Element) Working Group";
    contact
      "WG Web:   <http://tools.ietf.org/wg/pce/>
       WG List:  <mailto:pce@ietf.org>
       Editor:   Cheng Li
                 <mailto:c.l@huawei.com>";
    description
      "The YANG module augments the PCEP YANG operational
       model with SRv6.

       Copyright (c) 2020 IETF Trust and the persons identified as
       authors of the code.  All rights reserved.

       Redistribution and use in source and binary forms, with or
       without modification, is permitted pursuant to, and subject
       to the license terms contained in, the Simplified BSD License
       set forth in Section 4.c of the IETF Trust's Legal Provisions
       Relating to IETF Documents
       (https://trustee.ietf.org/license-info).

       This version of this YANG module is part of RFC XXXX; see the
       RFC itself for full legal notices.";

    revision 2020-07-08 {
      description
        "Initial revision.";
      reference
        "RFC XXXX:  A YANG Data Model for Path Computation
                    Element Communications Protocol
                    (PCEP) - Segement Routing in IPv6
                    (SRv6)";
    }

    /* Features */

    feature srv6 {
      description
        "Support Segment Routing in IPv6 (SRv6) for PCE.";
```

```
      reference
        "I-D.ietf-pce-segment-routing-ipv6: PCEP Extensions
         for Segment Routing leveraging the IPv6 data plane";
    }

    /* Identity */

    identity path-setup-srv6 {
      if-feature "srv6";
      base te-types:path-signaling-type;
      description
        "SRv6 path setup type";
    }

    /* Groupings */

    grouping srv6-msd {
      description
        "SRv6 MSD";
      leaf msd-type {
        type uint8;
        description
          "SRv6 Maximum Segment Depth (MSD) Type";
      }
      leaf msd-value {
        type uint8;
        description
          "SRv6 MSD value for the type";
      }
    }

    grouping srv6 {
      description
        "SRv6";
      container srv6 {
        if-feature "srv6";
        description
          "If SRv6 is supported";
        leaf enabled {
          type boolean;
          description
            "Enabled or Disabled";
        }
        leaf msd-limit {
          type boolean;
          default "false";
          description
            "True indicates no limit on MSD, the
```

```
            list srv6-msd is ignored";
        }
        list srv6-msd {
          key "msd-type";
          description
            "list of SRv6 MSD";
          uses srv6-msd;
        }
      }
    }

    grouping segment-list {
      description
        "Segment list grouping";
      container segment-list {
        description
          "Segments for given segment list";
        list segment {
          key "index";
          description
            "Configure Segment/hop at the index";
          uses segment-properties;
        }
      }
    }

    grouping segment-properties {
      description
        "Segment properties grouping";
      leaf index {
        type uint32;
        description
          "Segment index";
      }
      leaf sid-value {
        type srv6-types:srv6-sid;
        description
          "SRv6 SID value";
      }
    }

    /*
     * Augment modules to add SRv6
     */

    augment "/pcep:pcep/pcep:entity/pcep:capability" {
      description
        "Augmenting SRv6";
```

```
      uses srv6;
    }

    augment
      "/pcep:pcep/pcep:entity/pcep:peers/pcep:peer/pcep:capability" {
      description
        "Augmenting SRv6";
      uses srv6;
    }

    augment "/pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp" {
      description
        "Augmenting SRv6";
      container srv6 {
        when "/pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp/pcep:pst
              = 'path-setup-srv6'" {
          description
            "For SRv6 path";
        }
        if-feature "srv6";
        uses segment-list;
        description
          "SRv6";
      }
    }
  }
```

<CODE ENDS>

6. Security Considerations

The YANG module defined in this document is designed to be accessed via network management protocol such as NETCONF [RFC6241] or RESTCONF [RFC8040]. The lowest NETCONF layer is the secure transport layer and the mandatory-to-implement secure transport is SSH [RFC6242]. The lowest RESTCONF layer is HTTPS, and the mandatory-to-implement secure transport is TLS [RFC8446]

The NETCONF access control model [RFC8341] provides the means to restrict access for particular NETCONF or RESTCONF users to a pre-configured subset of all available NETCONF or RESTCONF protocol operations and content.

There are a number of data nodes defined in the YANG module which are writable/creatable/deletable (i.e., config true, which is the default). These data nodes may be considered sensitive or vulnerable in some network environments. Write operations (e.g., <edit-config>) to these data nodes without proper protection can have a negative

effect on network operations. These are the subtrees and data nodes and their sensitivity/vulnerability:

/pcep:pcep/pcep:entity/pcep:capability/pcep-srv6:srv6 - configure local SRv6 capability and parameters.

/pcep:pcep/pcep:entity/pcep:peers/pcep:peer/pcep:capability/pcep-srv6:srv6 - configure peer's SRv6 capability and parameters.

Unauthorized access to above list can adversely affect the PCEP session between the local entity and the peers. This may lead to inability to compute new paths, stateful operations on the delegated as well as PCE-initiated LSPs.

Some of the readable data nodes in this YANG module may be considered sensitive or vulnerable in some network environments. It is thus important to control read access (e.g., via get, get-config, or notification) to these data nodes. These are the subtrees and data nodes and their sensitivity/vulnerability:

/pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp/pcep-srv6:srv6 - The SRv6 SID in the network. Unauthorized access to this could provide the all path and network usage information.

7. IANA Considerations

This document registers a URI in the "IETF XML Registry" [RFC3688]. Following the format in RFC 3688, the following registration has been made.

URI: urn:ietf:params:xml:ns:yang:ietf-pcep-srv6

Registrant Contact: The PCE WG of the IETF.

XML: N/A; the requested URI is an XML namespace.

This document registers a YANG module in the "YANG Module Names" registry [RFC6020].

Name:	ietf-pcep-srv6
Namespace:	urn:ietf:params:xml:ns:yang:ietf-pcep-srv6
Prefix:	pcep-srv6
Reference:	This I-D

8. Acknowledgements

The authors would like to thank Dhruv Dhody for the initial YANG model.

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>.

[RFC3688] Mealling, M., "The IETF XML Registry", BCP 81, RFC 3688, DOI 10.17487/RFC3688, January 2004, <https://www.rfc-editor.org/info/rfc3688>.

[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>.

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, <https://www.rfc-editor.org/info/rfc6020>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

[RFC6242] Wasserman, M., "Using the NETCONF Protocol over Secure Shell (SSH)", RFC 6242, DOI 10.17487/RFC6242, June 2011, <https://www.rfc-editor.org/info/rfc6242>.

[RFC7950] Bjorklund, M., Ed., "The YANG 1.1 Data Modeling Language", RFC 7950, DOI 10.17487/RFC7950, August 2016, <https://www.rfc-editor.org/info/rfc7950>.

[RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <https://www.rfc-editor.org/info/rfc8040>.

[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>.

[RFC8340] Bjorklund, M. and L. Berger, Ed., "YANG Tree Diagrams", BCP 215, RFC 8340, DOI 10.17487/RFC8340, March 2018, <https://www.rfc-editor.org/info/rfc8340>.

[RFC8341] Bierman, A. and M. Bjorklund, "Network Configuration Access Control Model", STD 91, RFC 8341, DOI 10.17487/RFC8341, March 2018, <https://www.rfc-editor.org/info/rfc8341>.

[RFC8408] Sivabalan, S., Tantsura, J., Minei, I., Varga, R., and J. Hardwick, "Conveying Path Setup Type in PCE Communication Protocol (PCEP) Messages", RFC 8408, DOI 10.17487/RFC8408, July 2018, <https://www.rfc-editor.org/info/rfc8408>.

[RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <https://www.rfc-editor.org/info/rfc8446>.

[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>.

[RFC8776] Saad, T., Gandhi, R., Liu, X., Beeram, V., and I. Bryskin, "Common YANG Data Types for Traffic Engineering", RFC 8776, DOI 10.17487/RFC8776, June 2020, <https://www.rfc-editor.org/info/rfc8776>.

[I-D.raza-spring-srv6-yang] Raza, K., Agarwal, S., Liu, X., Hu, Z., Hussain, I., Shah, H., Voyer, D., Elmalky, H., Matsushima, S., Horiba, K., Abdelsalam, A., and J. Rajamanickam, "YANG Data Model for SRv6 Base and Static", draft-raza-spring-srv6-yang-05 (work in progress), October 2019.

[I-D.ietf-pce-pcep-yang]
Dhody, D., Hardwick, J., Beeram, V., and J. Tantsura, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", draft-ietf-pce-pcep-yang-13 (work in progress), October 2019.

[I-D.ietf-pce-segment-routing-ipv6]
Li, C., Negi, M., Koldychev, M., Kaladharan, P., and Y. Zhu, "PCEP Extensions for Segment Routing leveraging the IPv6 data plane", draft-ietf-pce-segment-routing-ipv6-06 (work in progress), July 2020.

9.2. Informative References

[RFC4655] Farrel, A., Vasseur, J., 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>.

[RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <https://www.rfc-editor.org/info/rfc7420>.

[RFC8342] Bjorklund, M., Schoenwaelder, J., Shafer, P., Watsen, K., and R. Wilton, "Network Management Datastore Architecture (NMDA)", RFC 8342, DOI 10.17487/RFC8342, March 2018, <https://www.rfc-editor.org/info/rfc8342>.

Authors' Addresses

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

EMail: c.l@huawei.com

Siva Sivabalan
Ciena Corporation

EMail: ssivabal@ciena.com

Shuping Peng
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

EMail: pengshuping@huawei.com

Mike Koldychev
Cisco Systems, Inc.

EMail: mkoldych@cisco.com

Luc-Fabrice Ndifor
MTN Cameroon
Cameroon

EMail: Luc-Fabrice.Ndifor@mtn.com

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 5, 2021

C. Li
Huawei Technologies
S. Sivabalan
Ciena Corporation
S. Peng
Huawei Technologies
M. Koldychev
Cisco Systems, Inc.
L. Ndifor
MTN Cameroon

November 1, 2020

A YANG Data Model for Segment Routing in IPv6 (SRv6) support in Path Computation Element Communications Protocol (PCEP) draft-li-pce-pcep-srv6-yang-02

Abstract

This document augments a YANG data model for the management of Path Computation Element Communications Protocol (PCEP) for communications between a Path Computation Client (PCC) and a Path Computation Element (PCE), or between two PCEs in support for Segment Routing in IPv6. The data model includes configuration data and state data (status information and counters for the collection of statistics).

Table of Contents

1. Introduction

The Path Computation Element (PCE) defined in [RFC4655] is an entity that is capable of computing a network path or route based on a network graph, and applying computational constraints. A Path Computation Client (PCC) may make requests to a PCE for paths to be computed.

PCEP is the communication protocol between a PCC and PCE and is defined in [RFC5440]. PCEP interactions include path computation requests and path computation replies as well as notifications of specific states related to the use of a PCE in the context of Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineering (TE). [RFC8231] specifies extensions to PCEP to enable stateful control of MPLS TE LSPs.

[I-D.ietf-pce-segment-routing-ipv6] extends [RFC8664] to support SR for IPv6 data plane.

[I-D.ietf-pce-pcep-yang] defines a YANG [RFC7950] data model for the management of PCEP speakers. This document contains a specification of the PCEP-SRv6 YANG module, "ietf-pcep-srv6" which provides the PCEP-SRv6 [I-D.ietf-pce-segment-routing-ipv6] data model.

The PCEP operational state is included in the same tree as the PCEP configuration consistent with Network Management Datastore Architecture [RFC8342]. The origin of the data is indicated as per the origin metadata annotation.

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. Terminology and Notation

This document also uses the following terms defined in [RFC7420]:

- PCEP entity: a local PCEP speaker.
- PCEP peer: to refer to a remote PCEP speaker.
- PCEP speaker: where it is not necessary to distinguish between local and remote.

Further, this document also uses the following terms defined in [RFC8231] :

- Stateful PCE, Passive Stateful PCE, Active Stateful PCE
- Delegation, Revocation, Redelegation
- LSP State Report, Path Computation Report message (PCRpt).
- LSP State Update, Path Computation Update message (PCUpd).

[RFC8281] :

- PCE-initiated LSP, Path Computation LSP Initiate Message (PCInitiate).

[RFC8408] :

- Path Setup Type (PST).

[RFC8664] :

- Segment Routing (SR).

[I-D.ietf-pce-segment-routing-ipv6] :

- Segment Routing in IPv6 (SRv6).

3.1. Tree Diagrams

A simplified graphical representation of the data model is used in this document. The meaning of the symbols in these diagrams is defined in [RFC8340].

3.2. Prefixes in Data Node Names

In this document, names of data nodes and other data model objects are often used without a prefix, as long as it is clear from the context in which YANG module each name is defined. Otherwise, names are prefixed using the standard prefix associated with the corresponding YANG module, as shown in Table 1.

Prefix	YANG module	Reference
te-types	ietf-te-types	[RFC8776]
pcep	ietf-pcep	[I-D.ietf-pce-pcep-yang]
srv6-types	ietf-srv6-types	[I-D.raza-spring-srv6-yang]
sr-policy	ietf-sr-policy	[I-D.ietf-spring-sr-policy-yang]
rt	ietf-routing	[RFC8349]

Table 1: Prefixes and corresponding YANG modules

4. The Design of PCEP-SRv6 Data Model

4.1. The Overview of PCEP SRv6 Data Model

The PCEP-SRv6 YANG module defined in this document has all the common building blocks for the PCEP-SRv6 extension.

```
module: ietf-pcep-srv6
  augment /pcep:pcep/pcep:entity/pcep:capability:
    +--rw srv6 {srv6}?
       +--rw enabled?     boolean
       +--rw msd-limit?   boolean
       +--rw srv6-msd* [msd-type]
          +--rw msd-type     uint8
          +--rw msd-value?   uint8
  augment /pcep:pcep/pcep:entity/pcep:peers/pcep:peer
            /pcep:capability:
    +--rw srv6 {srv6}?
       +--rw enabled?     boolean
       +--rw msd-limit?   boolean
       +--rw srv6-msd* [msd-type]
          +--rw msd-type     uint8
          +--rw msd-value?   uint8
  augment /pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp:
    +--ro srv6 {srv6}?
    |  +--ro segment-list
    |     +--ro segment* [index]
    |        +--ro index        uint32
    |        +--ro sid-value?   srv6-types:srv6-sid
    +--ro sr-policy
       +--ro color?             leafref
       +--ro endpoint?          leafref
       +--ro protocol-origin?   leafref
       +--ro originator?        leafref
       +--ro discriminator?     leafref
```

The sr-policy container is applicable for both SR-MPLS and SRv6.

5. PCEP-SRv6 YANG Modules

5.1. ietf-pcep-srv6 module

RFC Ed.: In this section, replace all occurrences of 'XXXX' with the actual RFC number and all occurrences of the revision date below with the date of RFC publication (and remove this note).

```
<CODE BEGINS> file "ietf-pcep-srv6@2020-10-31.yang"
module ietf-pcep-srv6 {
  yang-version 1.1;
  namespace "urn:ietf:params:xml:ns:yang:ietf-pcep-srv6";
  prefix pcep-srv6;

  import ietf-srv6-types {
```

```
      prefix srv6-types;
      reference
        "I-D.raza-spring-srv6-yang: YANG Data Model for SRv6
         Base and Static";
    }
    import ietf-te-types {
      prefix te-types;
      reference
        "RFC 8776: Common YANG Data Types for Traffic Engineering";
    }
    import ietf-pcep {
      prefix pcep;
      reference
        "I-D.ietf-pce-pcep-yang: A YANG Data Model for Path
         Computation Element Communications Protocol (PCEP)";
    }
    import ietf-sr-policy {
      prefix sr-policy;
      reference
        "I-D.ietf-spring-sr-policy-yang: YANG Data Model for
         Segment Routing Policy";
    }
    import ietf-routing {
      prefix rt;
      reference
        "RFC 8349: A YANG Data Model for Routing Management";
    }

    organization
      "IETF PCE (Path Computation Element) Working Group";
    contact
      "WG Web:   <http://tools.ietf.org/wg/pce/>
       WG List:  <mailto:pce@ietf.org>
       Editor:   Cheng Li
                 <mailto:c.l@huawei.com>";
    description
      "The YANG module augments the PCEP YANG operational
       model with SRv6.

       Copyright (c) 2020 IETF Trust and the persons identified as
       authors of the code.  All rights reserved.

       Redistribution and use in source and binary forms, with or
       without modification, is permitted pursuant to, and subject
       to the license terms contained in, the Simplified BSD License
       set forth in Section 4.c of the IETF Trust's Legal Provisions
       Relating to IETF Documents
       (https://trustee.ietf.org/license-info).
```

```
        This version of this YANG module is part of RFC XXXX; see the
        RFC itself for full legal notices.";

    revision 2020-10-31 {
      description
        "Initial revision.";
      reference
        "RFC XXXX:  A YANG Data Model for Path Computation
                    Element Communications Protocol
                    (PCEP) - Segement Routing in IPv6
                    (SRv6)";
    }

    /* Features */

    feature srv6 {
      description
        "Support Segment Routing in IPv6 (SRv6) for PCE.";
      reference
        "I-D.ietf-pce-segment-routing-ipv6: PCEP Extensions
         for Segment Routing leveraging the IPv6 data plane";
    }

    /* Identity */

    identity path-setup-srv6 {
      if-feature "srv6";
      base te-types:path-signaling-type;
      description
        "SRv6 path setup type";
    }

    /* Groupings */

    grouping srv6-msd {
      description
        "SRv6 MSD";
      leaf msd-type {
        type uint8;
        description
          "SRv6 Maximum Segment Depth (MSD) Type";
      }
      leaf msd-value {
        type uint8;
        description
          "SRv6 MSD value for the type";
      }
    }
```

```
    grouping srv6 {
      description
        "SRv6";
      container srv6 {
        if-feature "srv6";
        description
          "If SRv6 is supported";
        leaf enabled {
          type boolean;
          description
            "Enabled or Disabled";
        }
        leaf msd-limit {
          type boolean;
          default "false";
          description
            "True indicates no limit on MSD, the
             list srv6-msd is ignored";
        }
        list srv6-msd {
          key "msd-type";
          description
            "list of SRv6 MSD";
          uses srv6-msd;
        }
      }
    }

    grouping segment-list {
      description
        "Segment list grouping";
      container segment-list {
        description
          "Segments for given segment list";
        list segment {
          key "index";
          description
            "Configure Segment/hop at the index";
          uses segment-properties;
        }
      }
    }

    grouping segment-properties {
      description
        "Segment properties grouping";
      leaf index {
        type uint32;
```

```
        description
          "Segment index";
      }
      leaf sid-value {
        type srv6-types:srv6-sid;
        description
          "SRv6 SID value";
      }
    }

    grouping sr-policy {
      description
        "Segment Routing Policy grouping";
      // Editor's Note - headend is missig in SR Policy
      //  Yang mode
      leaf color {
        type leafref {
          path "/rt:routing/sr-policy:segment-routing/"
             + "sr-policy:traffic-engineering/sr-policy:"
             + "policies/sr-policy:policy/sr-policy:"
             + "color";
        }
        description
          "SR Policy Color";
        reference
          "I-D.ietf-spring-segment-routing-policy: Segment
           Routing Policy Architecture";
      }
      leaf endpoint {
        type leafref {
          path "/rt:routing/sr-policy:segment-routing/"
             + "sr-policy:traffic-engineering/sr-policy:"
             + "policies/sr-policy:policy/sr-policy:"
             + "endpoint";
        }
        description
          "SR Policy Endpoint";
        reference
          "I-D.ietf-spring-segment-routing-policy: Segment
           Routing Policy Architecture";
      }
      leaf protocol-origin {
        type leafref {
          path "/rt:routing/sr-policy:segment-routing/"
             + "sr-policy:traffic-engineering/sr-policy:"
             + "policies/sr-policy:policy/sr-policy:"
             + "candidate-paths/sr-policy:"
             + "candidate-path/sr-policy:protocol-origin";
```

```
        }
        must '(. = "pcep")' {
          error-message "The protocol origin must be PCEP";
        }
        description
          "SR Policy Candidate Path Protocol";
        reference
          "I-D.ietf-spring-segment-routing-policy: Segment
           Routing Policy Architecture";
      }
      leaf originator {
        type leafref {
          path "/rt:routing/sr-policy:segment-routing/"
             + "sr-policy:traffic-engineering/sr-policy:"
             + "policies/sr-policy:policy/sr-policy:"
             + "candidate-paths/sr-policy:"
             + "candidate-path/sr-policy:originator";
        }
        description
          "SR Policy Candidate Path Originator";
        reference
          "I-D.ietf-spring-segment-routing-policy: Segment
           Routing Policy Architecture";
      }
      leaf discriminator {
        type leafref {
          path "/rt:routing/sr-policy:segment-routing/"
             + "sr-policy:traffic-engineering/sr-policy:"
             + "policies/sr-policy:policy/sr-policy:"
             + "candidate-paths/sr-policy:"
             + "candidate-path/sr-policy:discriminator";
        }
        description
          "SR Policy Candidate Path Discriminator";
        reference
          "I-D.ietf-spring-segment-routing-policy: Segment
           Routing Policy Architecture";
      }
    }

    /*
     * Augment modules to add SRv6
     */

    augment "/pcep:pcep/pcep:entity/pcep:capability" {
      description
        "Augmenting SRv6";
      uses srv6;
```

```
    }

    augment
      "/pcep:pcep/pcep:entity/pcep:peers/pcep:peer/pcep:capability" {
      description
        "Augmenting SRv6";
      uses srv6;
    }

    augment "/pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp" {
      description
        "Augmenting SRv6";
      container srv6 {
        when "/pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp/pcep:pst
              = 'path-setup-srv6'" {
          description
            "For SRv6 path";
        }
        if-feature "srv6";
        uses segment-list;
        description
          "SRv6";
      }
      container sr-policy {
        when "/pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp/pcep:pst
              = 'te-types:path-setup-sr' or
              /pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp/pcep:pst
              = 'path-setup-srv6'" {
          description
            "Applicable for SR or SRv6";
        }
        uses sr-policy;
        description
          "SR Policy";
      }
    }
  }
```

<CODE ENDS>

6. Security Considerations

The YANG module defined in this document is designed to be accessed via network management protocol such as NETCONF [RFC6241] or RESTCONF [RFC8040]. The lowest NETCONF layer is the secure transport layer and the mandatory-to-implement secure transport is SSH [RFC6242].

The lowest RESTCONF layer is HTTPS, and the mandatory-to-implement secure transport is TLS [RFC8446]

The NETCONF access control model [RFC8341] provides the means to restrict access for particular NETCONF or RESTCONF users to a pre-configured subset of all available NETCONF or RESTCONF protocol operations and content.

There are a number of data nodes defined in the YANG module which are writable/creatable/deletable (i.e., config true, which is the default). These data nodes may be considered sensitive or vulnerable in some network environments. Write operations (e.g., <edit-config>) to these data nodes without proper protection can have a negative effect on network operations. These are the subtrees and data nodes and their sensitivity/vulnerability:

/pcep:pcep/pcep:entity/pcep:capability/pcep-srv6:srv6 - configure local SRv6 capability and parameters.

/pcep:pcep/pcep:entity/pcep:peers/pcep:peer/pcep:capability/pcep-srv6:srv6 - configure peer's SRv6 capability and parameters.

Unauthorized access to above list can adversely affect the PCEP session between the local entity and the peers. This may lead to inability to compute new paths, stateful operations on the delegated as well as PCE-initiated LSPs.

Some of the readable data nodes in this YANG module may be considered sensitive or vulnerable in some network environments. It is thus important to control read access (e.g., via get, get-config, or notification) to these data nodes. These are the subtrees and data nodes and their sensitivity/vulnerability:

/pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp/pcep-srv6:srv6 - The SRv6 SID in the network. Unauthorized access to this could provide entire path and network usage information.

/pcep:pcep/pcep:entity/pcep:lsp-db/pcep:lsp/pcep-srv6:sr-policy - The reference to SR Policy. Unauthorized access to this could provide SR Policy usage information.

7. IANA Considerations

This document registers a URI in the "IETF XML Registry" [RFC3688]. Following the format in RFC 3688, the following registration has been made.

URI: urn:ietf:params:xml:ns:yang:ietf-pcep-srv6

Registrant Contact: The PCE WG of the IETF.

XML: N/A; the requested URI is an XML namespace.

This document registers a YANG module in the "YANG Module Names" registry [RFC6020].

```
Name:         ietf-pcep-srv6
Namespace:    urn:ietf:params:xml:ns:yang:ietf-pcep-srv6
Prefix:       pcep-srv6
Reference:    This I-D
```

8. Acknowledgements

The authors would like to thank Dhruv Dhody for the initial YANG model.

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>.

[RFC3688] Mealling, M., "The IETF XML Registry", BCP 81, RFC 3688, DOI 10.17487/RFC3688, January 2004, <https://www.rfc-editor.org/info/rfc3688>.

[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>.

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, <https://www.rfc-editor.org/info/rfc6020>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

[RFC6242] Wasserman, M., "Using the NETCONF Protocol over Secure Shell (SSH)", RFC 6242, DOI 10.17487/RFC6242, June 2011, <https://www.rfc-editor.org/info/rfc6242>.

[RFC7950] Bjorklund, M., Ed., "The YANG 1.1 Data Modeling Language", RFC 7950, DOI 10.17487/RFC7950, August 2016, <https://www.rfc-editor.org/info/rfc7950>.

[RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <https://www.rfc-editor.org/info/rfc8040>.

[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>.

[RFC8340] Bjorklund, M. and L. Berger, Ed., "YANG Tree Diagrams", BCP 215, RFC 8340, DOI 10.17487/RFC8340, March 2018, <https://www.rfc-editor.org/info/rfc8340>.

[RFC8341] Bierman, A. and M. Bjorklund, "Network Configuration Access Control Model", STD 91, RFC 8341, DOI 10.17487/RFC8341, March 2018, <https://www.rfc-editor.org/info/rfc8341>.

[RFC8349] Lhotka, L., Lindem, A., and Y. Qu, "A YANG Data Model for Routing Management (NMDA Version)", RFC 8349, DOI 10.17487/RFC8349, March 2018, <https://www.rfc-editor.org/info/rfc8349>.

[RFC8408] Sivabalan, S., Tantsura, J., Minei, I., Varga, R., and J. Hardwick, "Conveying Path Setup Type in PCE Communication Protocol (PCEP) Messages", RFC 8408, DOI 10.17487/RFC8408, July 2018, <https://www.rfc-editor.org/info/rfc8408>.

[RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <https://www.rfc-editor.org/info/rfc8446>.

[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>.

[RFC8776] Saad, T., Gandhi, R., Liu, X., Beeram, V., and I. Bryskin, "Common YANG Data Types for Traffic Engineering", RFC 8776, DOI 10.17487/RFC8776, June 2020, <https://www.rfc-editor.org/info/rfc8776>.

[I-D.raza-spring-srv6-yang] Raza, K., Agarwal, S., Liu, X., Hu, Z., Hussain, I., Shah, H., Voyer, D., Matsushima, S., Horiba, K., Elmalky, H., Abdelsalam, A., and J. Rajamanickam, "YANG Data Model for SRv6 Base and Static", draft-raza-spring-srv6-yang-06 (work in progress), July 2020.

[I-D.ietf-pce-pcep-yang] Dhody, D., Hardwick, J., Beeram, V., and J. Tantsura, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", draft-ietf-pce-pcep-yang-14 (work in progress), July 2020.

[I-D.ietf-pce-segment-routing-ipv6] Li, C., Negi, M., Koldychev, M., Kaladharan, P., and Y. Zhu, "PCEP Extensions for Segment Routing leveraging the IPv6 data plane", draft-ietf-pce-segment-routing-ipv6-06 (work in progress), July 2020.

[I-D.ietf-spring-sr-policy-yang] Raza, K., Sawaya, R., Shunwan, Z., Voyer, D., Durrani, M., Matsushima, S., and V. Beeram, "YANG Data Model for Segment Routing Policy", draft-ietf-spring-sr-policy-yang-00 (work in progress), September 2020.

9.2. Informative References

[RFC4655] Farrel, A., Vasseur, J., 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>.

[RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <https://www.rfc-editor.org/info/rfc7420>.

[RFC8342] Bjorklund, M., Schoenwaelder, J., Shafer, P., Watsen, K., and R. Wilton, "Network Management Datastore Architecture (NMDA)", RFC 8342, DOI 10.17487/RFC8342, March 2018, <https://www.rfc-editor.org/info/rfc8342>.

Authors' Addresses

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

EMail: c.l@huawei.com

Siva Sivabalan
Ciena Corporation

EMail: ssivabal@ciena.com

Shuping Peng
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

EMail: pengshuping@huawei.com

Mike Koldychev
Cisco Systems, Inc.

EMail: mkoldych@cisco.com

Luc-Fabrice Ndifor
MTN Cameroon
Cameroon

EMail: Luc-Fabrice.Ndifor@mtn.com

PCE
Internet-Draft
Intended status: Standards Track
Expires: September 6, 2020

S. Peng
Q. Xiong
ZTE Corporation
F. Qin
China Mobile

March 5, 2020

PCEP Extension for SR-MPLS Entropy Label Position
draft-peng-pce-entropy-label-position-03

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 September 6, 2020.

Copyright Notice

Table of Contents

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:

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

- 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 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. The PCE MUST perform the computation of the end-to-end path as well as the the Entropy Label Position (ELP) including the number and the place of the ELI/ELs. The PCEs has the capability to get the ERLD information of all nodes in inter-domain scenarios.

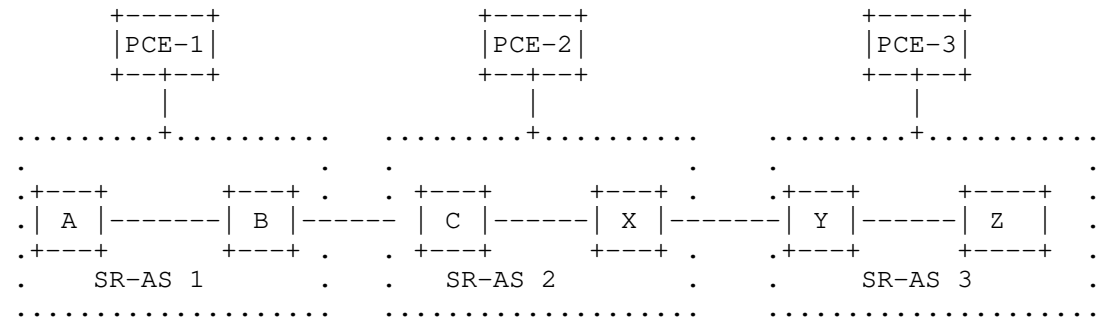


Figure 1: Entropy Labels in SR-MPLS Inter-Domain Scenario

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.

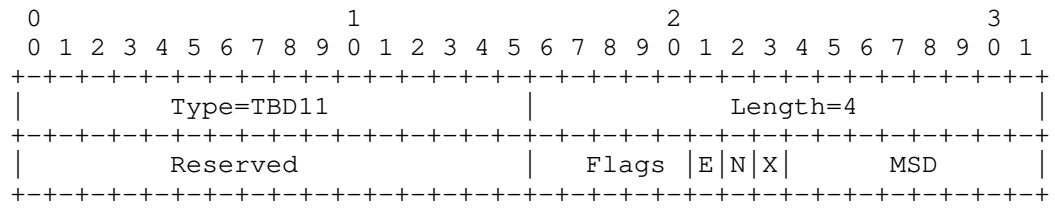


Figure 2: E-flag in SR-PCE-CAPABILITY sub-TLV

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 Object

The LSP Object is defined in Section 7.3 of [RFC8231]. This document defiend a new flag (E-flag) for the LSP Object as Figure 3 shown:

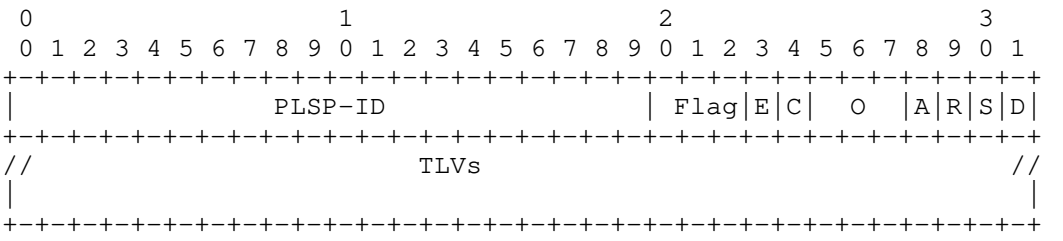


Figure 3: E-flag in LSP Object

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 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:

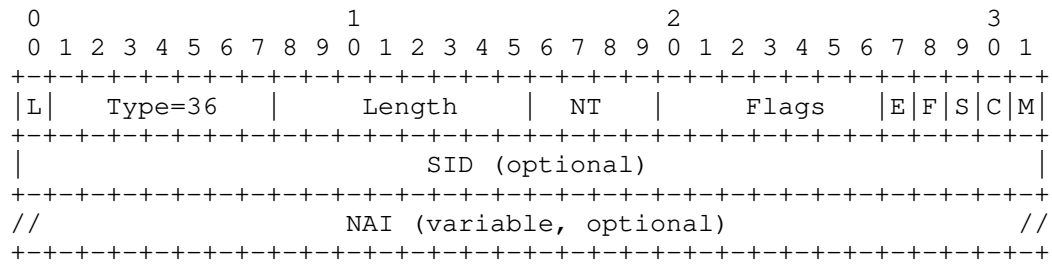


Figure 4: E-flag in SR-ERO subobject

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	ELP Configuration is supported (E)	[this document]

Table 1

8.2. New LSP Flag Registry

[RFC8231] defines the LSP object; per that RFC, IANA created a registry to manage the value of the LSP object's Flag field. IANA is requested to make allocations from the registry, as follows:

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

Table 2

8.3. 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
36	ELP Configuration (E)	[this document]

Table 3

9. 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>.

[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

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

Email: peng.shaofu@zte.com.cn

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

Email: xiong.quan@zte.com.cn

Fengwei Qin
China Mobile
Beijing
China

Email: qinfengwei@chinamobile.com

PCE
Internet-Draft
Intended status: Standards Track
Expires: February 13, 2021

S. Peng
Q. Xiong
ZTE Corporation
F. Qin
China Mobile

August 12, 2020

PCEP Extension for SR-MPLS Entropy Label Position
draft-peng-pce-entropy-label-position-04

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 February 13, 2021.

Copyright Notice

Table of Contents

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:

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

- 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 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. The PCE MUST perform the computation of the end-to-end path as well as the the Entropy Label Position (ELP) including the number and the place of the ELI/ELs. The PCEs has the capability to get the ERLD information of all nodes in inter-domain scenarios.

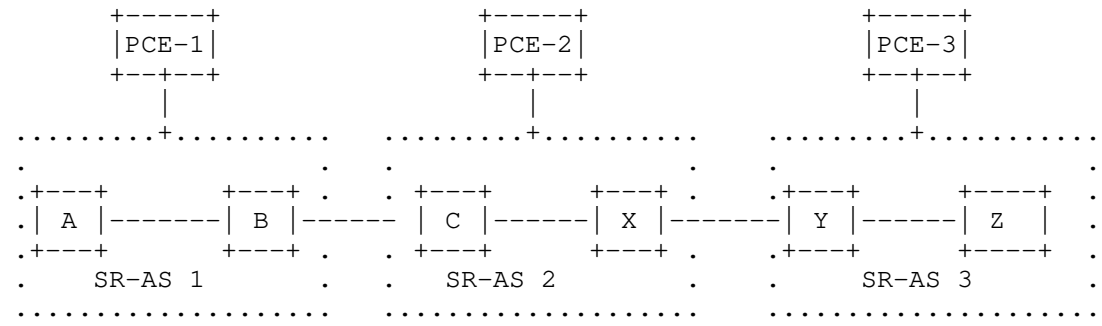


Figure 1: Entropy Labels in SR-MPLS Inter-Domain Scenario

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.

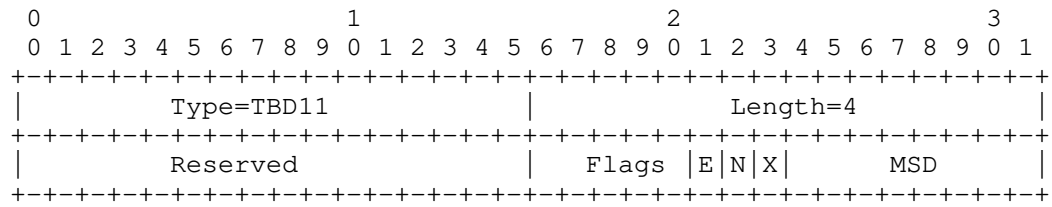


Figure 2: E-flag in SR-PCE-CAPABILITY sub-TLV

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.xiong-pce-lsp-flag]. The format is shown as Figure 3:

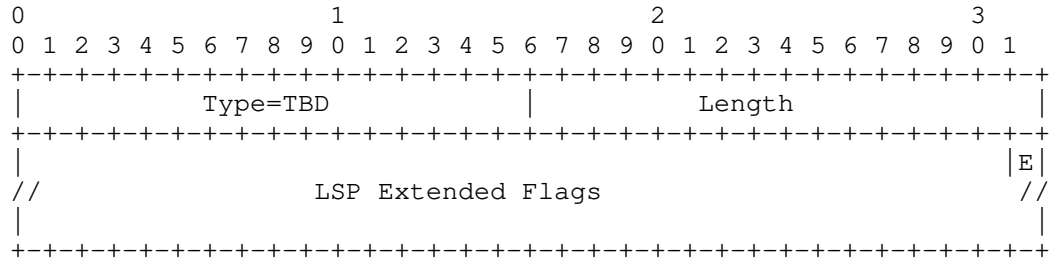


Figure 3: E-flag in LSP-EXTENDED-FLAG TLV

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 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:

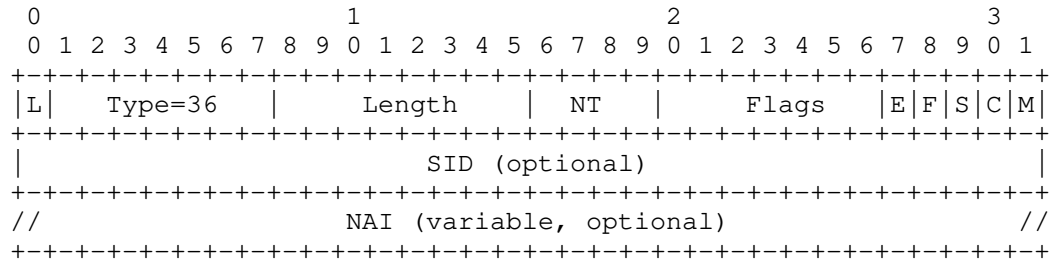


Figure 4: E-flag in SR-ERO subobject

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.xiong-pce-lsp-flag] 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. 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 3

9. Normative References

[I-D.xiong-pce-lsp-flag]
Xiong, Q., "LSP Object Flag Extension of Stateful PCE", draft-xiong-pce-lsp-flag-02 (work in progress), May 2020.

[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

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

Email: peng.shaofu@zte.com.cn

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

Email: xiong.quan@zte.com.cn

Fengwei Qin
China Mobile
Beijing
China

Email: qinfengwei@chinamobile.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 18, 2020

A. Stone
M. Aissaoui
Nokia
S. Sidor
Cisco Systems, Inc.
S. Sivabalan
Ciena Coroporation

June 16, 2020

Local Protection Enforcement in PCEP
draft-stone-pce-local-protection-enforcement-01

Abstract

This document aims to clarify existing usage of the local protection desired bit signalled in Path Computation Element Protocol (PCEP). This document also introduces a new flag for signalling protection strictness in PCEP.

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.

1. Introduction

Path Computation Element (PCE) Communication Protocol (PCEP) [RFC5440] enables the communication between a Path Computation Client (PCC) and a Path Control Element (PCE), or between two PCEs based on the PCE architecture [RFC4655].

PCEP [RFC5440] utilizes flags, values and concepts previously defined in RSVP-TE Extensions [RFC3209] and Fast Reroute Extensions to RSVP-TE [RFC4090]. One such concept in PCEP is the 'Local Protection Desired' (L-flag in the LSPA Object in RFC5440), which was originally defined in the SESSION-ATTRIBUTE Object in RFC3209. In RSVP, this flag signals to downstream routers that local protection is desired, which indicates to transit routers that they may use a local repair mechanism. The headend router calculating the path does not know whether a downstream router will or will not protect a hop during it's calculation. Therefore, a local protection desired does not require the transit router to satisfy protection in order to establish the RSVP signalled path. This flag is signalled in PCEP as an attribute of the LSP via the LSP Attributes object.

PCEP Extensions for Segment Routing (draft-ietf-pce-segment-routing) extends support in PCEP for Segment Routed LSPs (SR-LSPs) as defined in the Segment Routing Architecture [RFC8402]. As per the Segment Routing Architecture, Adjacency Segment Identifiers(Adj-SID) may be eligible for protection (using IPFRR or MPLS-FRR). The protection eligibility is advertised into IGP (draft-ietf-ospf-segment-routing-extensions and draft-ietf-isis-segment-routing-extensions) as the B-Flag part of the Adjacency SID sub-tlv and can be discovered by a PCE via BGP-LS [RFC7752] using the BGP-LS Segment Routing Extensions (draft-ietf-idr-bgp-ls-segment-routing-ext). An Adjacency SID may or may not have protection eligibility and for a given adjacency between two routers there may be multiple Adjacency SIDs, some of which are protected and some which are not.

A Segment Routed path calculated by PCE may contain various types of segments, as defined in [RFC8402] such as Adjacency, Node or Binding. The protection eligibility for Adjacency SIDs can be discovered by PCE, so therefore the PCE can take the protection eligibility into consideration as a path constraint. If a path is calculated to include other segment identifiers which are not applicable to having their protection state advertised, as they may only be locally significant for each router processing the SID such as Node SIDs, it

may not be possible for PCE to include the protection constraint as part of the path calculation.

It is desirable for an operator to define the enforcement, or strictness of the protection requirement when it can be applied.

2. Requirements Language

In this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in BCP 14, [RFC2119].

3. Terminology

This document uses the following terminology:

PROTECTION MANDATORY: path MUST have protection eligibility on all links.

UNPROTECTED MANDATORY: path MUST NOT have protection eligibility on all links.

PROTECTION PREFERRED: path SHOULD have protection eligibility on all links but MAY contain links which do not have protection eligibility.

UNPROTECTED PREFERRED: path SHOULD NOT have protection eligibility on all links but MAY contain links which have protection eligibility.

PCC: Path Computation Client. Any client application requesting a path computation to be performed by a Path Computation Element.

PCE: Path Computation Element. An entity (component, application, or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints.

PCEP: Path Computation Element Protocol.

4. Motivation

4.1. Implementation differences

As defined in [RFC5440] the mechanism to signal protection enforcement in PCEP is with the previously mentioned L-flag defined in the LSPA Object. The name of the flag uses the term "Desired", which by definition means "strongly wished for or intended" and is rooted in the RSVP use case. For RSVP, this is not within control of the PCE. However, [RFC5440] does state "When set, this means that

the computed path must include links protected with Fast Reroute as defined in [RFC4090]." Implementations of [RFC5440] have either interpreted the L-Flag as PROTECTION MANDATORY or PROTECTION PREFERRED, leading to operational differences.

4.2. SLA Enforcement

The boolean bit flag is unable to distinguish between the different options of PROTECTION MANDATORY, UNPROTECTED MANDATORY, PROTECTION PREFERRED and UNPROTECTED PREFERRED. The selection of the options are typically dependent on the service level agreement the operator wishes to impose on the LSP. When enforcement is used, the resulting shortest path calculation is impacted.

For example, PROTECTION MANDATORY is for use cases where an operator may need the LSP to follow a path which has local protection provided along the full path, ensuring that if there is anywhere along the path that traffic will be fast re-routed at the point of failure.

For another example, UNPROTECTED MANDATORY is when an operator may intentionally prefer an LSP to not be locally protected, and thus would rather local failures to cause the LSP to go down and/or rely on other protection mechanisms such as a secondary diverse path.

There are also use cases where there is simply no requirement to enforce protection or no protection along a path. This can be considered as "do not care to enforce". This is a relaxation of the protection constraint. The path calculation is permitted the use of any SID which is available along the calculated path. The SID backup availability does not impact the shortest path computation. Since links may have both protected and unprotected SIDs available, the option PROTECTION PREFERRED or UNPROTECTED PREFERRED is used to instruction PCE a preference on which SID to select, as the behaviour of the LSP would differ during a local failure depending on which SID is selected.

5. Protection Enforcement Flag (E-Flag)

Section 7.11 in Path Computation Element Protocol [RFC5440] describes the encoding of the Local Protection Desired (L-Flag). A new flag is proposed in this document in the LSP Attributes Object which extends the L-Flag to identify the protection enforcement.

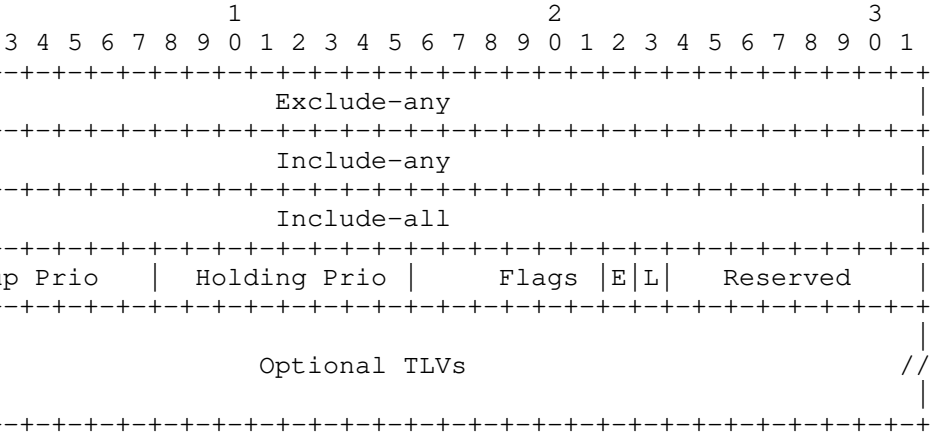
The flag bit is to be allocated by IANA following IETF Consensus.

This draft version proposes using bit 6.

Codespace of the Flag field (LSPA Object)

Bit	Description	Reference
7	Local Protection Desired	RFC5440
6	Local Protection Enforcement	This document

The format of the LSPA Object as defined in [RFC5440] is:



Flags (8 bits)

- L flag: As defined in [RFC5440] and further updated by this document. When set, protection is desired. When not set, protection is not desired. The enforcement of the protection is identified via the E-Flag.

- E flag (Protection Enforcement): When set, the value of the L-Flag MUST be treated as a MUST constraint where applicable, when protection state of a SID is known. When E flag is not set, the value of the L-Flag MUST be treated as a MAY constraint.

When L-flag is set and E-flag is set then PCE MUST consider the protection eligibility as PROTECTION MANDATORY constraint.

When L-flag is set and E-flag is not set then PCE MUST consider the protection eligibility as PROTECTION PREFERRED constraint.

When L-flag is not set and E-flag is not set then PCE SHOULD consider the protection eligibility as UNPROTECTED PREFERRED but MAY consider protection eligibility as UNPROTECTED MANDATORY constraint.

When L-flag is not set and E-flag is set then PCE MUST consider the protection eligibility as UNPROTECTED MANDATORY constraint.

For a PCC which does not yet support this draft, the E-flag bit is always set to zero as per [RFC5440]. Therefore, a PCE communicating with a PCC which does not support this draft would treat the L-Flag set as being PROTECTION PREFERRED.

The protection constraint can only be applied to resource selection in which the protection state is known to PCE. A PCE calculating a path that includes resources which does not support the protection state being known to PCE (such as Node SID), then the protection state MAY ignore the protection enforcement constraint.

UNPROTECTED PREFERRED and PROTECTED PREFERRED may seem similar but they indicate the preference of selection if PCE has an option of either protected or unprotected available for a link. When presented with either option, PCE SHOULD select the SID which has a protection state matching the state of the L-Flag.

6. Security Considerations

This document clarifies the behaviour of an existing flag and introduces a new flag to provide further control of that existing behaviour. The introduction of this new flag and behaviour clarification does not create any new sensitive information. No additional security measure is required.

Securing the PCEP session using Transport Layer Security (TLS) [RFC8253], as per the recommendations and best current practices in [RFC7525],, is RECOMMENDED.

7. IANA Considerations

8. LSP Attributes Protection Enforcement Flag

This document defines a new LSP Attribute Flag; IANA is requested to make the following bit allocation from the "LSPA Object" sub registry of the PCEP Numbers registry, as follows:

Value	Name	Reference
6	PROTECTION-ENFORCEMENT	This document

9. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, DOI 10.17487/RFC4090, May 2005, <https://www.rfc-editor.org/info/rfc4090>.

[RFC4655] Farrel, A., Vasseur, J., 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>.

[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>.

[RFC7525] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May 2015, <https://www.rfc-editor.org/info/rfc7525>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <https://www.rfc-editor.org/info/rfc7752>.

[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>.

[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>.

Authors' Addresses

Andrew Stone
Nokia

Email: andrew.stone@nokia.com

Mustapha Aissaoui
Nokia

Email: mustapha.aissaoui@nokia.com

Samuel Sidor
Cisco Systems, Inc.

Email: ssidor@cisco.com

Siva Sivabalan
Ciena Coroporation

Email: ssivabal@ciena.com

Network Working Group
Internet-Draft
Updates: 5440 (if approved)
Intended status: Standards Track
Expires: February 18, 2021

A. Stone
M. Aissaoui
Nokia
S. Sidor
Cisco Systems, Inc.
S. Sivabalan
Ciena Coroporation

August 17, 2020

Local Protection Enforcement in PCEP
draft-stone-pce-local-protection-enforcement-02

Abstract

This document updates [RFC5440] to clarify usage of the local protection desired bit signalled in Path Computation Element Protocol (PCEP). This document also introduces a new flag for signalling protection strictness in PCEP.

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 February 18, 2021.

Copyright Notice

Table of Contents

1. Introduction

Path Computation Element (PCE) Communication Protocol (PCEP) [RFC5440] enables the communication between a Path Computation Client (PCC) and a Path Control Element (PCE), or between two PCEs based on the PCE architecture [RFC4655].

PCEP [RFC5440] utilizes flags, values and concepts previously defined in RSVP-TE Extensions [RFC3209] and Fast Reroute Extensions to RSVP-TE [RFC4090]. One such concept in PCEP is the 'Local Protection Desired' (L-flag in the LSPA Object in RFC5440), which was originally defined in the SESSION-ATTRIBUTE Object in RFC3209. In RSVP, this flag signals to downstream routers that local protection is desired, which indicates to transit routers that they may use a local repair mechanism. The headend router calculating the path does not know whether a downstream router will or will not protect a hop during it's calculation. Therefore, a local protection desired does not require the transit router to satisfy protection in order to establish the RSVP signalled path. This flag is signalled in PCEP as an attribute of the LSP via the LSP Attributes object.

PCEP Extensions for Segment Routing (draft-ietf-pce-segment-routing) extends support in PCEP for Segment Routed LSPs (SR-LSPs) as defined in the Segment Routing Architecture [RFC8402]. As per the Segment Routing Architecture, Adjacency Segment Identifiers(Adj-SID) may be eligible for protection (using IPFRR or MPLS-FRR). The protection eligibility is advertised into IGP (draft-ietf-ospf-segment-routing-extensions and draft-ietf-isis-segment-routing-extensions) as the B-Flag part of the Adjacency SID sub-tlv and can be discovered by a PCE via BGP-LS [RFC7752] using the BGP-LS Segment Routing Extensions (draft-ietf-idr-bgp-ls-segment-routing-ext). An Adjacency SID may or may not have protection eligibility and for a given adjacency between two routers there may be multiple Adjacency SIDs, some of which are protected and some which are not.

A Segment Routed path calculated by PCE may contain various types of segments, as defined in [RFC8402] such as Adjacency, Node or Binding. The protection eligibility for Adjacency SIDs can be discovered by PCE, so therefore the PCE can take the protection eligibility into consideration as a path constraint. If a path is calculated to include other segment identifiers which are not applicable to having their protection state advertised, as they may only be locally significant for each router processing the SID such as Node SIDs, it may not be possible for PCE to include the protection constraint as part of the path calculation.

It is desirable for an operator to define the enforcement, or strictness of the protection requirement when it can be applied.

This document updates [RFC5440] by further describing the behaviour with Local Protection Desired Flag (L-Flag) and extends on it with the introduction of Enforcement Flag (E-Flag).

2. Requirements Language

In this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in BCP 14, [RFC2119].

3. Terminology

This document uses the following terminology:

PROTECTION MANDATORY: path MUST have protection eligibility on all links.

UNPROTECTED MANDATORY: path MUST NOT have protection eligibility on all links.

PROTECTION PREFERRED: path SHOULD have protection eligibility on all links but MAY contain links which do not have protection eligibility.

UNPROTECTED PREFERRED: path SHOULD NOT have protection eligibility on all links but MAY contain links which have protection eligibility.

PCC: Path Computation Client. Any client application requesting a path computation to be performed by a Path Computation Element.

PCE: Path Computation Element. An entity (component, application, or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints.

PCEP: Path Computation Element Protocol.

4. Motivation

4.1. Implementation differences

As defined in [RFC5440] the mechanism to signal protection enforcement in PCEP is with the previously mentioned L-flag defined in the LSPA Object. The name of the flag uses the term "Desired", which by definition means "strongly wished for or intended" and is rooted in the RSVP use case. For RSVP, this is not within control of the PCE. However, [RFC5440] does state "When set, this means that the computed path must include links protected with Fast Reroute as defined in [RFC4090]." Implementations of [RFC5440] have either interpreted the L-Flag as PROTECTION MANDATORY or PROTECTION PREFERRED, leading to operational differences.

4.2. SLA Enforcement

The boolean bit flag is unable to distinguish between the different options of PROTECTION MANDATORY, UNPROTECTED MANDATORY, PROTECTION PREFERRED and UNPROTECTED PREFERRED. The selection of the options are typically dependent on the service level agreement the operator wishes to impose on the LSP. When enforcement is used, the resulting shortest path calculation is impacted.

For example, PROTECTION MANDATORY is for use cases where an operator may need the LSP to follow a path which has local protection provided along the full path, ensuring that if there is anywhere along the path that traffic will be fast re-routed at the point of failure.

For another example, UNPROTECTED MANDATORY is when an operator may intentionally prefer an LSP to not be locally protected, and thus would rather local failures to cause the LSP to go down and/or rely on other protection mechanisms such as a secondary diverse path.

There are also use cases where there is simply no requirement to enforce protection or no protection along a path. This can be considered as "do not care to enforce". This is a relaxation of the protection constraint. The path calculation is permitted the use of any SID which is available along the calculated path. The SID backup availability does not impact the shortest path computation. Since links may have both protected and unprotected SIDs available, the option PROTECTION PREFERRED or UNPROTECTED PREFERRED is used to instruction PCE a preference on which SID to select, as the behaviour of the LSP would differ during a local failure depending on which SID is selected.

5. Protection Enforcement Flag (E-Flag)

Section 7.11 in Path Computation Element Protocol [RFC5440] describes the encoding of the Local Protection Desired (L-Flag). A new flag is proposed in this document in the LSP Attributes Object which extends the L-Flag to identify the protection enforcement.

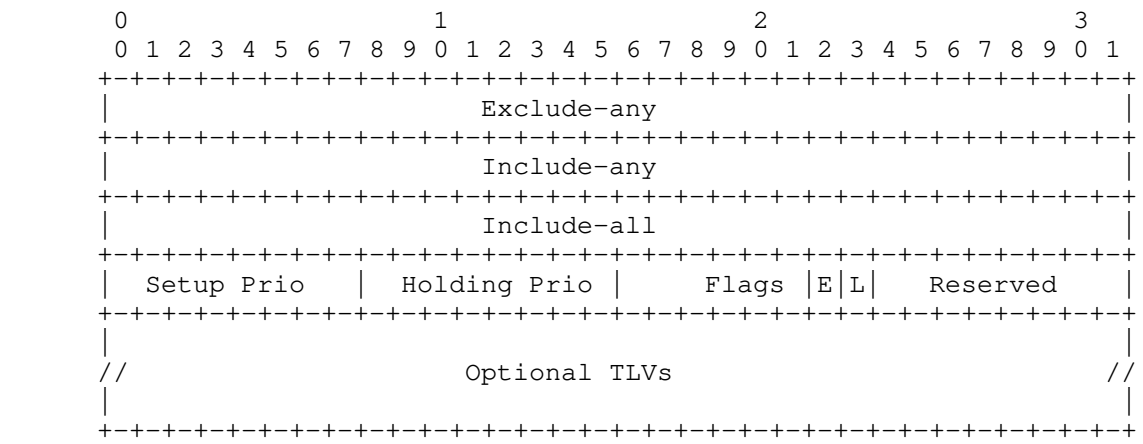
The flag bit is to be allocated by IANA following IETF Consensus.

This draft version proposes using the next available bit: {TBD} //Editor note, next available bit at time of writing is bit 6

Codespace of the Flag field (LSPA Object)

Bit	Description	Reference
7	Local Protection Desired	RFC5440
<TBD>	Local Protection Enforcement	This document

The format of the LSPA Object as defined in [RFC5440] is:



Flags (8 bits)

- L flag: As defined in [RFC5440] and further updated by this document. When set, protection is desired. When not set, protection is not desired. The enforcement of the protection is identified via the E-Flag.

- E flag (Protection Enforcement): When set, the value of the L-Flag MUST be treated as a MUST constraint where applicable, when protection state of a SID is known. When E flag is not set, the value of the L-Flag MUST be treated as a MAY constraint.

When L-flag is set and E-flag is set then PCE MUST consider the protection eligibility as PROTECTION MANDATORY constraint.

When L-flag is set and E-flag is not set then PCE MUST consider the protection eligibility as PROTECTION PREFERRED constraint.

When L-flag is not set and E-flag is not set then PCE SHOULD consider the protection eligibility as UNPROTECTED PREFERRED but MAY consider protection eligibility as UNPROTECTED MANDATORY constraint.

When L-flag is not set and E-flag is set then PCE MUST consider the protection eligibility as UNPROTECTED MANDATORY constraint.

UNPROTECTED PREFERRED and PROTECTED PREFERRED may seem similar but they indicate the preference of selection of a SID if PCE has an option of either protected or unprotected available on a link. When presented with either option, PCE SHOULD select the SID which has a protection state matching the state of the L-Flag.

The protection enforcement constraint can only be applied to resource selection in which the protection state is known to PCE. A PCE calculating a path that includes resources which does not support the protection state being known to PCE (such as Node SID), then the protection state MAY ignore the protection enforcement constraint.

5.1. Backwards Compatibility

Considerations in the message passing between PCC and PCE for the E-Flag bit which are not supported by the entity are outlined in this section, with requirements for PCE and PCC implementing this document described at the end.

For a PCC or PCE which does not yet support this document, the E-flag bit is ignored and set to zero in PCRpt and/or PCUpd as per [RFC5440] for PCC-initiated or as per ([RFC8281]) for PCE-initiated LSPs. It's important to note that [RFC8231] and [RFC8281] permit LSP Attribute Object to be included in PCUpd messages for PCC-initiated and PCE-initiated LSPs. For PCC-initiated LSPs, PCUpd E-Flag (and L-Flag) are an echo from the previous PCRpt however the bit value is ignored on PCE from the previous PCRpt, therefore the E-Flag value set in the PCUpd is zero. A PCE which does not support this document sends PCUpd messages with the E-Flag unset for PCC-initated LSPs even if set in the prior PCReq or PCRpt. A PCC which does not support this document sends PCRpt messages with the E-Flag unset for PCE-initiated LSPs even if set in the prior PCInitiate or PCUpd.

For a PCC which does support this document, it MAY set E-Flag bit depending on local configuration. If communicating with a PCE which does not yet support this document, the PCE follows the behaviour specified in [RFC5440] and will ignore the E-Flag bit thus it will not compute a path respecting the enforcement constraint.

For PCC-initiated LSPs, PCC SHOULD ignore the E-Flag value received from PCE in a PCUpd message.

For PCE-initiated LSPs, PCC MAY process the E-Flag value received from PCE in a PCUpd message. PCE SHOULD ignore the E-Flag value received from PCC in a PCRpt message.

6. Implementation Status

[Note to the RFC Editor - remove this section before publication, as well as remove the reference to RFC 7942.]

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 [RFC7942].

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 catalogue of available implementations or their features. Readers are advised to note that other implementations may exist.

According to [RFC7942], "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".

6.1. Nokia Implementation

- Organization: Nokia
- Implementation: NSP PCE and SROS PCC.
- Description: Implementation for calculation and conveying intention described in this document
- Maturity Level: Demo
- Coverage: Full
- Contact: andrew.stone@nokia.com

6.2. Cisco Implementation

- Organization: Cisco Systems, Inc.
- Implementation: IOS-XR PCE and PCC.
- Description: Implementation for calculation and conveying intention described in this document
- Maturity Level: Demo
- Coverage: Full
- Contact: ssidor@cisco.com

7. Security Considerations

This document clarifies the behaviour of an existing flag and introduces a new flag to provide further control of that existing behaviour. The introduction of this new flag and behaviour clarification does not create any new sensitive information. No additional security measure is required.

Securing the PCEP session using Transport Layer Security (TLS) [RFC8253], as per the recommendations and best current practices in [RFC7525] is RECOMMENDED.

8. IANA Considerations

8.1. LSP Attributes Protection Enforcement Flag

This document defines a new LSP Attribute Flag; IANA is requested to make the following bit allocation from the "LSPA Object" sub registry of the PCEP Numbers registry, as follows:

Value	Name	Reference
TBD	PROTECTION-ENFORCEMENT	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>.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, DOI 10.17487/RFC4090, May 2005, <https://www.rfc-editor.org/info/rfc4090>.

[RFC4655] Farrel, A., Vasseur, J., 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>.

[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>.

[RFC7525] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May 2015, <https://www.rfc-editor.org/info/rfc7525>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <https://www.rfc-editor.org/info/rfc7752>.

[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>.

[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>.

9.2. Informative References

[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>.

Acknowledgements

Thanks to Dhruv Dhody for comments and discussions on this document and Mike Koldychev for reviews.

Authors' Addresses

Andrew Stone
Nokia

Email: andrew.stone@nokia.com

Mustapha Aissaoui
Nokia

Email: mustapha.aissaoui@nokia.com

Samuel Sidor
Cisco Systems, Inc.

Email: ssidor@cisco.com

Siva Sivabalan
Ciena Coroporation

Email: ssivabal@ciena.com

PCE
Internet-Draft
Intended status: Standards Track
Expires: December 3, 2020

Q. Xiong
ZTE Corporation

June 1, 2020

LSP Object Flag Extension of Stateful PCE
draft-xiong-pce-lsp-flag-02

Abstract

RFC8231 describes a set of extensions to PCEP to enable stateful control of MPLS-TE and GMPLS Label Switched Paths(LSPs) via PCEP. One of the extensions is the LSP object which includes a Flag field and the length is 12 bits. However, 11 bits of the Flag field has been assigned in RFC8231, RFC8281 and RFC8623 respectively.

This document proposes to define a new LSP-EXTENDED-FLAG TLV for LSP object to extend the length of the flags.

Table of Contents

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. One of the extensions is the LSP object which contains a flag field indicating to a PCE that the LSP State Synchronization is in progress.

As defined in [RFC8231], the length of the flag field is 12 bits and the value from bit 5 to bit 11 is used for operational, administrative, remove, synchronize and delegate respectively. The bit value 4 is assigned in [RFC8281] for create. The bits from 1 to 3 is assigned in [RFC8623] for ERO-compression, fragmentation and P2MP respectively. Almost all bits of the Flag field has been assigned in RFC8231, RFC8281 and RFC8623 respectively. It is required to extend the length of the flag field for other cases.

This document proposes to define a new LSP-EXTENDED-FLAG TLV for LSP object to extend the length of the flag.

2. Conventions used in this document

2.1. Terminology

The terminology is defined as [RFC5440] and [RFC8231].

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. PCEP Extension

The LSP Object is defined in Section 7.3 of [RFC8231]. This document proposes to define a new LSP-EXTENDED-FLAG TLV for LSP Object to extend the length of the flag.

3.1. The LSP-EXTENDED-FLAG TLV

The format of the LSP-EXTENDED-FLAG TLV is as shown in the Figure 1.

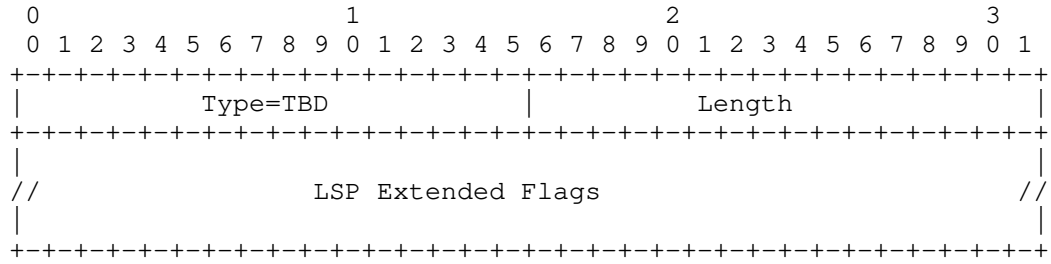


Figure 1: LSP-EXTENDED-FLAG TLV Format

Type (16 bits): the value is TBD1 by IANA.

Length (16 bits): multiple of 4 octets.

LSP Extended Flags: this contains an array of units of 32-bit flags numbered from the most significant as bit zero, where each bit represents one LSP capability or state.

3.2. Processing

The LSP Extended Flags field is an array of units of 32 flags and being used starting from the least significant bit. Any bit being assigned indicates a special LSP capability or state when the bit is set to 0. No bits are defined in this document and the bits of the LSP Extended Flags field MAY be assigned for future uses and IANA will manage the space of the LSP Extended Flags. Unassigned bits are reserved and SHOULD be set to zero on transmission and MUST be ignored on receipt.

The LSP-EXTENDED-FLAG TLV MUST be included in the LSP Object when the bits of the extended flag field need to be used. If the TLV is missing, the PCE will generate an error with Error-type=6 (Mandatory Object missing) and error-value TBD2 (LSP-EXTENDED-FLAG TLV missing) and close the session.

4. Backward Compatibility

The LSP-EXTENDED-FLAG TLV defined in this document does not introduce any interoperability issues.

A router not supporting the LSP-EXTENDED-FLAG TLV will just silently ignore the TLV as specified in section 3.2.

The LSP-EXTENDED-FLAG TLV MUST be defined as mandatory when a router supporting the LSP Object and needs to use the extended flag field.

5. IANA Considerations

5.1. LSP Object

5.1.1. LSP-EXTENDED-FLAG TLV

IANA has assigned a registry for TLVs carried in the LSP Object defined in [RFC8231]. IANA is requested to make allocations for the LSP-EXTENDED-FLAG TLV carried within LSP Object from the "PCEP TLV Type Indicators" subregistry of the "Path Computation Element Protocol (PCEP) Numbers" registry, as follows:

Value	Description	Reference
TBD1	LSP-EXTENDED-FLAG TLV	[This document]

Table 1

5.1.2. LSP Extended Flags Field

IANA is requested to create a new subregistry, named "LSP Extended Flags Field", from the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the LSP Extended Flags field of the LSP-EXTENDED-FLAG TLV. New values MUST request to be assigned by Standards Action [RFC8126] and IANA will manage the space of the bit flags numbering them in the usual IETF notation starting at zero and continuing at least through 31. Each bit should be tracked with the following qualities:

Bit number (counting from bit 0 as the most significant bit)

Capability description

Defining RFC

5.2. PCEP-Error Object

IANA is requested to register the following error types and error values within the "PCEP-ERROR Object Error Types and Values" subregistry of the "Path Computation Element Protocol (PCEP) Numbers" registry:

Error-Type	Meaning
6	Mandatory Object missing Error-value TBD2: LSP-EXTENDED-FLAG TLV missing

Table 2

6. Security Considerations

For LSP Object procssing security considerations, see [RFC8231].

No additional security issues are raised in this document beyond those that exist in the referenced documents.

7. Acknowledgements

Authors would like to thank the comments and suggestions from Dhruv Dhody and Farrel Adrian.

8. 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>.

[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>.

[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>.

Author's Address

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

Email: xiong.quan@zte.com.cn