

PCE Working Group  
Internet-Draft  
Updates: [8408](#) (if approved)  
Intended status: Standards Track  
Expires: January 9, 2020

Y. Qu  
H. Chen  
Futurewei  
July 8, 2019

PCEP Extensions for Preferred Path Routing  
draft-qc-pce-path-routing-00

## Abstract

This document specifies extensions to the Path Computation Element Protocol (PCEP) that allow a stateful PCE to initiate Preferred Path Routing (PPR) paths. It also specifies how PCC should react to 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 January 9, 2020.

## Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

|                        |  |                    |
|------------------------|--|--------------------|
| <a href="#">1.</a>     | Introduction . . . . .                               | <a href="#">2</a>  |
| <a href="#">2.</a>     | Terminology . . . . .                                | <a href="#">3</a>  |
| <a href="#">3.</a>     | Overview of PCEP Operation in PPR Networks . . . . . | <a href="#">4</a>  |
| <a href="#">4.</a>     | Extensions to PCEP . . . . .                         | <a href="#">4</a>  |
| <a href="#">4.1.</a>   | Capability for PPR . . . . .                         | <a href="#">4</a>  |
| <a href="#">4.2.</a>   | PPR TLV . . . . .                                    | <a href="#">6</a>  |
| <a href="#">4.2.1.</a> | PPR-Flags . . . . .                                  | <a href="#">7</a>  |
| <a href="#">4.2.2.</a> | PPR-Prefix Sub-TLV . . . . .                         | <a href="#">7</a>  |
| <a href="#">4.2.3.</a> | PPR-ID Sub-TLV . . . . .                             | <a href="#">8</a>  |
| <a href="#">4.2.4.</a> | PPR-PDE Sub-TLV . . . . .                            | <a href="#">9</a>  |
| <a href="#">4.2.5.</a> | PPR-Attributes Sub-TLV . . . . .                     | <a href="#">12</a> |
| <a href="#">5.</a>     | Procedures . . . . .                                 | <a href="#">12</a> |
| <a href="#">6.</a>     | Management Considerations . . . . .                  | <a href="#">13</a> |
| <a href="#">7.</a>     | Security Considerations . . . . .                    | <a href="#">13</a> |
| <a href="#">8.</a>     | IANA Considerations . . . . .                        | <a href="#">14</a> |
| <a href="#">9.</a>     | Acknowledgements . . . . .                           | <a href="#">14</a> |
| <a href="#">10.</a>    | References . . . . .                                 | <a href="#">14</a> |
| <a href="#">10.1.</a>  | Normative References . . . . .                       | <a href="#">14</a> |
| <a href="#">10.2.</a>  | Informative References . . . . .                     | <a href="#">15</a> |
|                        | Authors' Addresses . . . . .                         | <a href="#">16</a> |

## [1.](#) Introduction

Segment Routing (SR) [[RFC8402](#)] leverages the source routing paradigm. A packet is steered through a network using an ordered list of instructions, called "segments". A segment is often referred to by its Segment Identifier (SID), and can represent any instruction, topological or service based. The SR architecture can be implemented using either MPLS [[I-D.ietf-spring-segment-routing-mpls](#)] or IPv6 [[I-D.ietf-6man-segment-routing-header](#)] forwarding plane.

A Preferred Path Routing (PPR)  
[[I-D.chunduri-lsr-isis-preferred-path-routing](#)]  
[[I-D.chunduri-lsr-ospf-preferred-path-routing](#)] path is identified by  
a PPR ID and described as a list of Path Description Elements (PDEs).  
A PPR path could be used as a Traffic Engineering (TE) path, or a

Service Function Chaining (SFC) [[RFC7665](#)] path etc. PPR can help to  
reduce the data plane overhead and mitigate MTU issues.

[RFC5440] describes the Path Computation Element Protocol (PCEP) for  
communication between a Path Computation Client (PCC) and a Path  
Computation Element (PCE) or between a pair of PCEs. [[RFC8281](#)]  
specifies a mechanism to dynamically initiate LSPs on a PCC based on  
the requests from a stateful PCE or a controller using stateful PCE.  
A stateful PCE can compute one or more SR-TE paths, and initiate a  
SR-TE path on a PCC using the SR specific PCEP extensions specified  
in [[I-D.ietf-pce-segment-routing](#)].

It is possible to use a stateful PCE for computing PPR paths taking  
into account various constraints. This document specifies the PCEP  
extensions that can be used to for the stateful PCE to initiate a PPR  
path on a PCC.

## [2.](#) Terminology

The following terminologies are used in this document:

ERO: Explicit Route Object

IGP: Interior Gateway Protocol

IS-IS: Intermediate System to Intermediate System

LSR: Label Switching Router

NAI: Node or Adjacency Identifier

OSPF: Open Shortest Path First

PCC: Path Computation Client

PCE: Path Computation Element

PCEP: Path Computation Element Communication Protocol

RRO: Record Route Object

SID: Segment Identifier

SR: Segment Routing

PPR: Preferred Path Routing

PDE: Path Description Element

### [3.](#) Overview of PCEP Operation in PPR Networks

In a PPR network, the ingress node of a PPR path prepends a PPR header to all outgoing packets. Depending on the forwarding plane, different formats of header will be chosen. The header contains a PPR ID, in combination with the control plane information about the PPR ID, the packets will be routed through the network.

In PCEP messages, PPR path information is carried in the Explicit Route Object (ERO), which consists of a sequence of sub objects. PPR paths computed by a PCE can be represented in an ERO with a PPR ID followed by one of the following list:

- o An ordered set of IP addresses representing network nodes/links.
- o An ordered set of SIDs, with or without the corresponding IP addresses.
- o An ordered set of MPLS labels, with or without corresponding IP address.

After a PCC receives the PCEP messages, one of the following two methods can be used to program the control plane:

- o IGP can be used to populate the PPR path information as described in [[I-D.chunduri-lsr-isis-preferred-path-routing](#)] and [[I-D.chunduri-lsr-ospf-preferred-path-routing](#)].
- o If PCEP is used directly to program a PPR path, and a PCC sees

itself is part of a PPR path, it will install the corresponding information, such as PPR ID, next hop into forwarding table.

4. Extensions to PCEP

4.1. Capability for PPR

When a PCEP session is established between a PCE and a PCC, they exchange their capabilities of supporting PPR.

A new sub-TLV, which is called PPR\_CAPABILITY, is defined. It is included in the PATH\_SETUP\_TYPE\_CAPABILITY TLV with PST = TBD1 (suggested value 3 for PPR) in the OPEN object, which is exchanged in Open messages when the PCEP session is established. Its format is illustrated below.

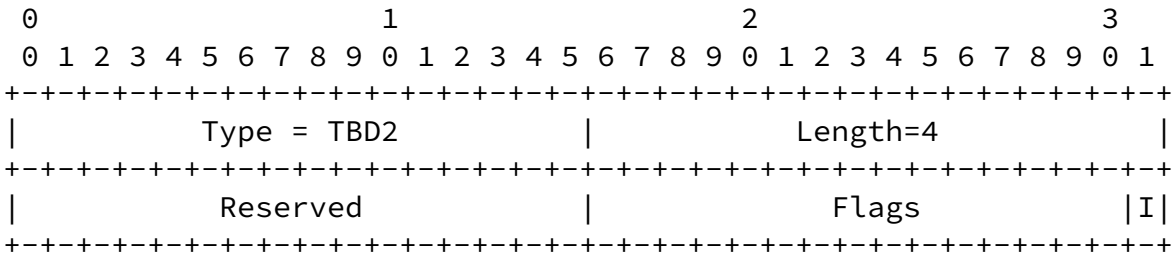


Figure 1: PPR\_CAPABILITY sub-TLV

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. This document defines the following flag bits. The other bits MUST be set to zero by the sender and MUST be ignored by the receiver.

- \* I: A PCC sets this flag bit to 1 to indicate that it is capable

| 0 |   |   |   |   |   |   |   |   |   | 1 |   |   |   |   |   |   |   |   |   | 2 |   |   |   |   |   |   |   |   |   | 3 |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - |



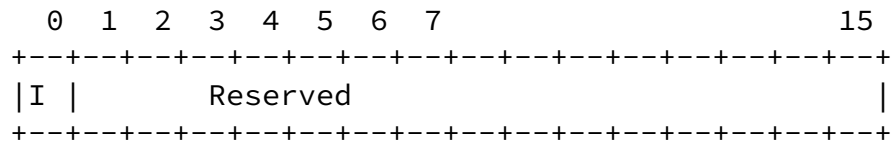


Figure 3: PPR-Flags Field

**I Flag:** IGP flag. If set, IGP will be used to flood this path as described in [[I-D.chunduri-lsr-isis-preferred-path-routing](#)] and [[I-D.chunduri-lsr-isis-preferred-path-routing](#)]. If not set, the PCC received this TLV will verify the path information, and if it sees itself as part of the PPR path, it will install the corresponding path information into its forwarding table.

**Reserved:** This field Must be set to zero in transmission and ignored on reception.

#### [4.2.2.](#) PPR-Prefix Sub-TLV

A PPR-Prefix Sub-TLV contains a prefix for the path described in a PPR TLV. Its format is illustrated below:

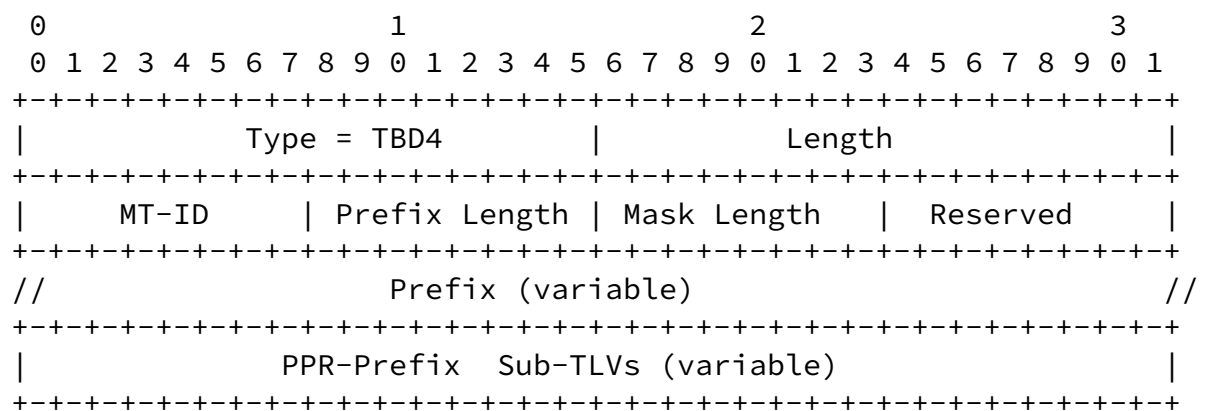


Figure 4: PPR-Prefix Sub-TLV Format

**Type:** TBD4 is to be assigned by IANA.

**Length:** Total length of the value field in bytes (variable).

**MT-ID:** 1 octet. Multi-Topology ID (as defined in [[RFC4915](#)]).



Prefix Length: 1 octet. It is the length of the prefix in bits.  
Only the most significant octets of the Prefix are encoded.

Mask Length: 1 octet. It is the valid length of the prefix in bits.  
Only the most significant octets of the Prefix are encoded.

Reserved: 1 octet. Must be set to zero in transmission and ignored on reception.

Prefix: Variable octets. It represents the prefix at the tail-end of the PPR. For the address family IPv4 unicast, the prefix itself is encoded as a 32-bit value. The default route is represented by a prefix of length 0.

PPR-Prefix Sub-TLVs: Variable octets. It has 1 octet type, 1 octet length and value field is defined per the type field.

#### 4.2.3. PPR-ID Sub-TLV

A PPR-ID Sub-TLV contains an identifier, which represents the actual data plane identifier in the packet and could be of any data plane as defined in PPR-ID-type field. Both Prefix and PPR-ID MUST belong to a same node in the network. The format of the PPR-ID Sub-TLV is illustrated below:

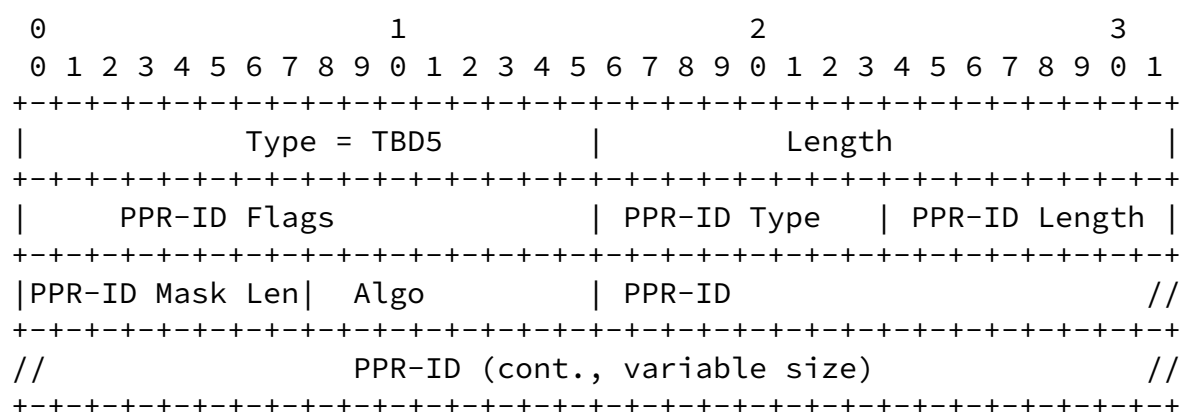


Figure 5: PPR-ID Sub-TLV Format

Type: TBD5 is to be assigned by IANA.

Length: Total length of the value field in bytes (variable).

PPR-ID Type: 1 octet. Data plane type of PPR-ID. This is a new registry (TBD IANA) for this Sub-TLV and the defined types are as follows:

- a. Type = 1: MPLS SID/Label.
- b. Type = 2: Native IPv4 Address/Prefix.
- c. Type = 3: Native IPv6 Address/Prefix.
- d. Type = 4: IPv6 SID in SRv6 with SRH.

PPR-ID Flags: 2 octets. Two flags are defined below:

```

      0  1  2  3  4  5  6  7                                15
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|   Reserved                                                    |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Reserved: Reserved bits for future use. They must be set to zero in transmission and ignored on reception.

PPR-ID Length: 1 octet. It is the length of the PPR-ID field in octets and this depends on the PPR-ID type. See PPR-ID below for the length of this field and other considerations.

PPR-ID Mask Len: 1 octet. It is applicable for only for PPR-ID Type 2. For Type 1 this value MUST be set to zero. It contains the length of the PPR-ID Prefix in bits. Only the most significant octets of the Prefix are encoded. This is needed, if PPR-ID followed is an IPv4 Prefix instead of 4 octet Address respectively.

Algo: 1 octet. It represents the SPF algorithm. Algorithm registry is as defined in [I-D.ietf-ospf-segment-routing-extensions].

PPR-ID: Variable octets. It represents the Preferred Path forwarding identifier that would be on the data packet. The value of this field is variable and it depends on the PPR-ID Type - for Type 1, this is and MPLS SID/Label. For Type 2 this is a 4 byte IPv4 address.

#### [4.2.4.](#) PPR-PDE Sub-TLV

This is a new Sub-TLV type in PPR TLV and is called as PPR Path Description Element (PDE). PPR-PDEs are used to describe the path in the form of set of contiguous and ordered Sub-TLVs, where first Sub-TLV represents (the top of the stack in MPLS data plane or) first node/segment of the path. These set of ordered Sub-TLVs can have both topological SIDs and non-topological SIDs (e.g., service

segments). The format of the PPR-PDE Sub-TLV is illustrated below:

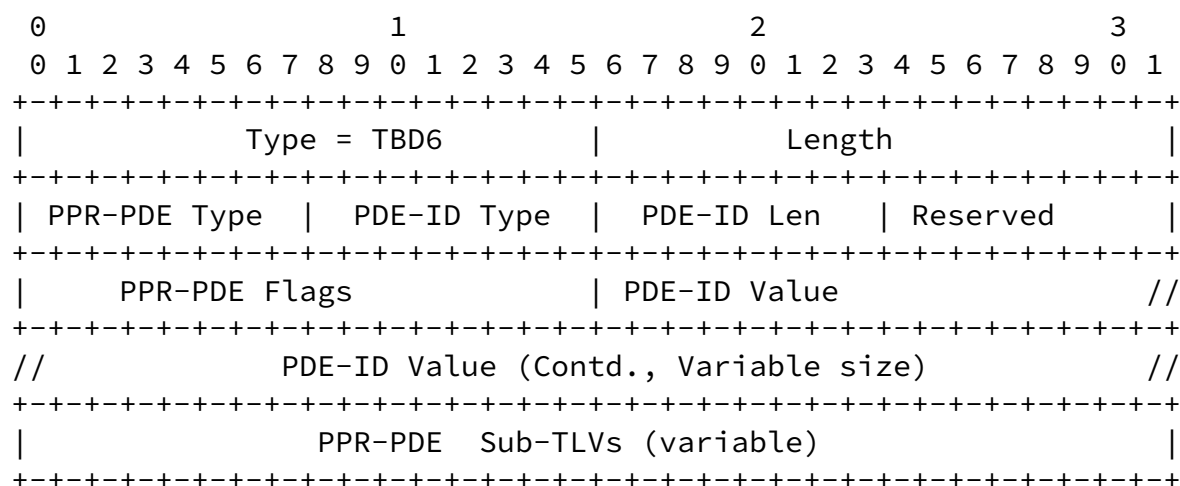


Figure 6: PPR-PDE Sub-TLV Format

Type: TBD6 is to be assigned by IANA.

Length: Total length of the value field in bytes (variable).

PPR-PDE Type: 1 octet. This is a new registry (TBD IANA) for this Sub-TLV and the defined types are as follows:

- a. Type = 1: Topological.
- b. Type = 2: Non-Topological.

PPR-ID Type: 1 octet. PDE-forwarding Identifier Type. This is a new registry (TBD IANA) for this Sub-TLV and the defined types and corresponding PDE-ID Len, PDE-ID Value are as follows:

Type = 1: SID/Label Sub-TLV as defined in [I-D.ietf-ospf-segment-routing-extensions]. PDE-ID Len and PDE-ID Value fields are defined as the above.

Type = 2: SR-MPLS Prefix SID. PDE-ID Len and PDE-ID Value are the same as Type 1.

Type = 3: SR-MPLS Adjacency SID. PDE-ID Len and PDE-ID Value are

Type = 4: IPv4 Node Address. PDE-ID Len is 4 bytes and PDE-ID Value is 4 bytes IPv4 address encoded similar to IPv4 Prefix described above.

Type = 5: IPv4 P2P interface Address. PDE-ID Len is 4 bytes and PDE-ID Value is 4 bytes IPv4 address encoded similar to IPv4 Prefix described above.

Type = 6: IPv4 LAN interface Address. PDE-ID Len is 4 bytes and PDE-ID Value is 4 bytes IPv4 address encoded similar to IPv4 Prefix described above.

Type = 7: IPv6 Node Address. PDE-ID Len is 16 bytes and PDE-ID Value is 16 bytes IPv6 address encoded similar to IPv6 Prefix described above.

Type = 8: IPv6 P2P interface Address. PDE-ID Len is 16 bytes and PDE-ID Value is 16 bytes IPv6 address encoded similar to IPv6 Prefix described above.

Type = 9: IPv6 LAN interface Address. PDE-ID Len is 16 bytes and PDE-ID Value is 16 bytes IPv6 address encoded similar to IPv6 Prefix described above.

Type = 10: SRv6 Node SID as defined in [\[I-D.ietf-lsr-isrv6-extensions\]](#).

Type = 11: SRv6 Adjacency-SID. PDE-ID Len and PDE-ID Values are similar to SRv6 Node SID above.

PPR-ID Len: 1 octet. It is the length of the PPR-ID field in octets.

Reserved: 1 octet. Reserved bits MUST be reset to zero on transmission and ignored on receive.

PPR-PDE Flags: 2 octets. Two flags are defined below:

0 1 2 3 4 5 6 7 15



## 5. Procedures

PPR\_CAPABILITY sub-TLV in the Open message is exchanged between a PCC and a PCE to indicate the support of PPR. When both the PCC and the PCE support PPR\_CAPABILITY, each of the Open messages sent by the PCC and PCE contains PATH\_SETUP\_TYPE\_CAPABILITY TLV with a PST list containing PST=TBD1 and a PPR\_CAPABILITY sub-TLV.

If a PCC sets the I bit to 1 in PPR\_CAPABILITY sub-TLV, it means this PCC is capable of flooding PPR path info across IGP domain. Otherwise it means it supports to install PPR path info into its forwarding table but can't flood the information.

After a PCC receives a PPR TLV, it needs to verify whether it's valid.

If a PCC receives a PPR TLV with flog I bit set to 1, however this PCC doesn't support IGP flooding of PPR info, it MUST consider the TLV invalid and MUST send a PCERR message with Error-Type = 10 ("Reception of an invalid object").

The PPR TLV contains a sequence of subobjects/sub TLVs, which defines the PPR path information. If a routing process/protocol is configured to flood PPR path, it interprets the sub TLVs and converts them into corresponding routing protocol TLVs and flood them.

Section 4 in [[I-D.chunduri-lsr-isis-preferred-path-routing](#)] describes how PCCs act after receiving path information from a controller.

## 6. Management Considerations

This document adds a new path setup type to PCEP to allow PPR paths to be set up on PCCs. This path setup type may be used with PCEP alongside other path setup types, such as RSVP-TE, Segment Routing Traffic Engineering, or it may be used exclusively.

The PATH-SETUP-TYPE-CAPABILITY TLV is used to indicate the path setup type supported. It requires both PCC and PCE to include PST=TBD1, also include the PPR-CAPABILITY sub-TLV.

A network operator can use the following steps to enable PCEP to setup paths using PPR as path setup type:

- o The operator enables the PPR PST on the PCE servers.
- o The operator enables the PPR PST on the PCCs.
- o The operator resets each PCEP session. The PCEP sessions come back up with PPR enabled.
- o If the operator detects a problem, they can roll the network back to its initial state by disabling the PPR PST on the PCEP speakers and resetting the PCEP sessions.

The data plane is unaffected if a PCEP session is reset. Any PPR path that was set up before the session reset will remain in active.

The PPR YANG module is defined in [[I-D.qct-lsr-ppr-yang](#)], and it is used to configure PPR path information on a router and it can coexist with the path set up method defined in this document.

## [7.](#) Security Considerations

The security considerations described in [[RFC5440](#)], [[RFC8231](#)], [[RFC8281](#)] and [[RFC8408](#)] are applicable to this specification. No additional security measure is required.

This draft enables a network controller to instantiate a path in the network, and that creates additional vulnerability.

## [8.](#) IANA Considerations

TBD.

## [9.](#) Acknowledgements

TBD.

## [10.](#) References

### [10.1.](#) Normative References

- [I-D.chunduri-lsr-isis-preferred-path-routing]  
Chunduri, U., Li, R., White, R., Tantsura, J., Contreras, L., and Y. Qu, "Preferred Path Routing (PPR) in IS-IS", [draft-chunduri-lsr-isis-preferred-path-routing-03](#) (work in progress), May 2019.
- [I-D.chunduri-lsr-ospf-preferred-path-routing]  
Chunduri, U., Qu, Y., White, R., Tantsura, J., and L. Contreras, "Preferred Path Routing (PPR) in OSPF", [draft-chunduri-lsr-ospf-preferred-path-routing-03](#) (work in progress), May 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>>.
- [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>>.

- [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>>.

## 10.2. Informative References

- [I-D.ietf-6man-segment-routing-header]  
Filsfils, C., Previdi, S., Leddy, J., Matsushima, S., and d. daniel.voyer@bell.ca, "IPv6 Segment Routing Header (SRH)", [draft-ietf-6man-segment-routing-header-18](#) (work in progress), April 2019.
- [I-D.ietf-lsr-isis-srv6-extensions]  
Psenak, P., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extensions to Support Routing over IPv6 Dataplane", [draft-ietf-lsr-isis-srv6-extensions-00](#) (work in progress), May 2019.
- [I-D.ietf-pce-segment-routing]  
Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "PCEP Extensions for Segment Routing", [draft-ietf-pce-segment-routing-16](#) (work in progress), March 2019.
- [I-D.ietf-spring-segment-routing]  
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", [draft-ietf-spring-segment-routing-15](#) (work in progress), January 2018.
- [I-D.ietf-spring-segment-routing-mpls]  
Bashandy, A., Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with MPLS data plane", [draft-ietf-spring-segment-routing-mpls-18](#) (work in progress), December 2018.
- [I-D.qct-lsr-ppr-yang]  
Qu, Y., Chunduri, U., and J. Tantsura, "YANG Data Model for Preferred Path Routing", [draft-qct-lsr-ppr-yang-01](#) (work in progress), March 2019.

- [RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", [RFC 4915](#), DOI 10.17487/RFC4915, June 2007, <<https://www.rfc-editor.org/info/rfc4915>>.
- [RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", [RFC 5120](#), DOI 10.17487/RFC5120, February 2008, <<https://www.rfc-editor.org/info/rfc5120>>.
- [RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", [RFC 7665](#), DOI 10.17487/RFC7665, October 2015, <<https://www.rfc-editor.org/info/rfc7665>>.
- [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

Yingzhen Qu  
Futurewei

EMail: [yingzhen.qu@futurewei.com](mailto:yingzhen.qu@futurewei.com)

Huaimo Chen  
Futurewei

EMail: [huaimo.chen@futurewei.com](mailto:huaimo.chen@futurewei.com)

