Authors: H. Chen   M. McBride   M. Toy   G. Mishra
         Futurewei   Futurewei   Verizon Inc. Verizon Inc.
         A. Wang   Z. Li   Y. Liu
         China Telecom   China Mobile   China Mobile
         B. Khasanov   L. Liu   X. Liu
         Yandex LLC   Fujitsu   Volta Networks
                    Path Ingress Protections

Abstract

This document describes extensions to Path Computation Element (PCE)
communication Protocol (PCEP) for fast protecting the ingress nodes
of two types of paths or tunnels, which are Segment Routing (SR)
paths and Bit Index Explicit Replication Tree/Traffic Engineering
(BIER-TE) paths. The extensions comprise a foundation for protecting
the ingress nodes of different types of paths. Based on this, the
ingress protection of a new type of paths can be easily supported.

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

Table of Contents

1. Introduction

   The fast protection of a transit node in each type of paths or
   tunnels have been proposed. For example, the fast protection of a
   transit node in a Segment Routing (SR) path or tunnel is described
   in [I-D.ietf-rtgwg-segment-routing-ti-lfa]. The fast protection of a
   transit node of a "Bit Index Explicit Replication" (BIER) Traffic
   Engineering (BIER-TE) path or tunnel is described in

[I-D.chen-bier-te-frr]. [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/tunnel, a BIER-TE path/tunnel, or other type of paths/tunnels.

This document fills that void and specifies protocol extensions to Path Computation Element (PCE) communication Protocol (PCEP) [RFC5440] and [RFC9050] for fast protecting the ingress nodes of two types of paths: SR paths and BIER-TE paths. The extensions comprise a foundation for protecting the ingress nodes of different types of paths. Based on this, the ingress protection of a new type of paths can be easily supported.

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

1.1. Terminologies

The following terminologies are used in this document.

PCE: Path Computation Element or Path Computation Element server

PCEP: PCE communication Protocol

PCC: Path Computation Client

BIER: Bit Index Explicit Replication

BIFT: Bit Index Forwarding Table

CE: Customer Edge

PE: Provider Edge

TE: Traffic Engineering

SR: Segment Routing

LFA: Loop-Free Alternate

TI-LFA: Topology Independent LFA

BFD: Bidirectional Forwarding Detection

VPN: Virtual Private Network

L3VPN:

> Layer 3 VPN

FIB: Forwarding Information Base

2. Path Ingress Protection Examples

This section shows two examples of path ingress protection. One is SR path ingress protection, and the other is BIER-TE path ingress protection.

2.1. SR Path Ingress Protection Example

Figure 1 shows an example of protecting ingress PE1 (or say primary ingress) of a SR path (or say primary SR path), which is from ingress PE1 to egress PE3 via P1 and represented by *** in the figure. A PCE computes the primary SR path and sends the path to primary ingress PE1 (i.e., the PCC running on PE1) in a PCEP message after the PCE receives a request with primary ingress PE1, egress PE3 and constraints on the path.

```
        *******  *******
    [PE1]-----[P1]-----[PE3]          PE1 Primary Ingress
     / |        |# ##### | \          PEx Provider Edge
    /  |        |#       |  \         CEx Customer Edge
[CE1]  |        |#       |   [CE2]    Px  Non Provider Edge
    \  |        |#       |  /         *** Primary SR Path
     \ |  ##### |#       | /          ### Backup SR Path
    [PE2]-----[P2]-----[PE4]          PE2 Backup Ingress
```
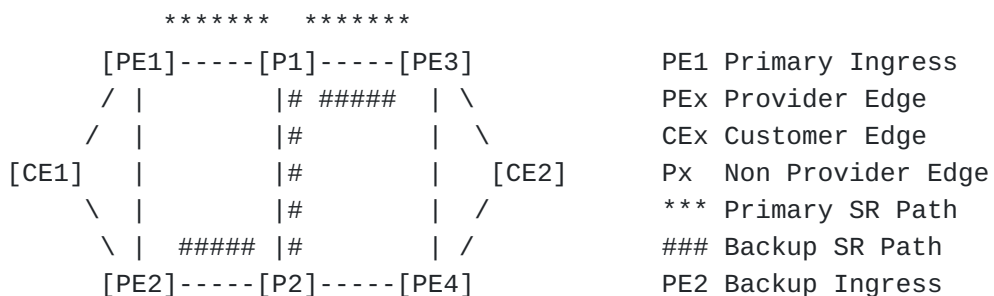
Figure 1: Protecting Ingress PE1 of SR Path

A backup SR path is from backup ingress PE2 to egress PE3 through P2 and P1, and represented by ### in the figure. The PCE computes the backup SR path and sends the backup path to backup ingress PE2 (i.e., the PCC running on PE2) in a PCEP message for protecting primary ingress PE1.

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

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

2.2. BIER-TE Path Ingress Protection Example

Figure 2 shows an example of protecting ingress PE1 (or say primary
ingress) of a primary BIER-TE path, which is from ingress PE1 to
egress nodes PE3 and PE4 via P1. This primary BIER-TE path is
represented by *** in the figure.

```
           *******  *******
     [PE1]-----[P1]-----[PE3]        PE1 Primary Ingress
      / |        #|*\#####  |        PEx Provider Edge
     /  |        #| *\__     |       CEx Customer Edge
[CE1]   |        #|  ***\    |       Px  Non Provider Edge
     \  |        #|     *\   |       *** Primary BIER-TE Path
      \ |        #|      *\  |       ### Backup BIER-TE Path
     [PE2]-----[P2]-----[PE4]        PE2 Backup Ingress
          #####    #####
```
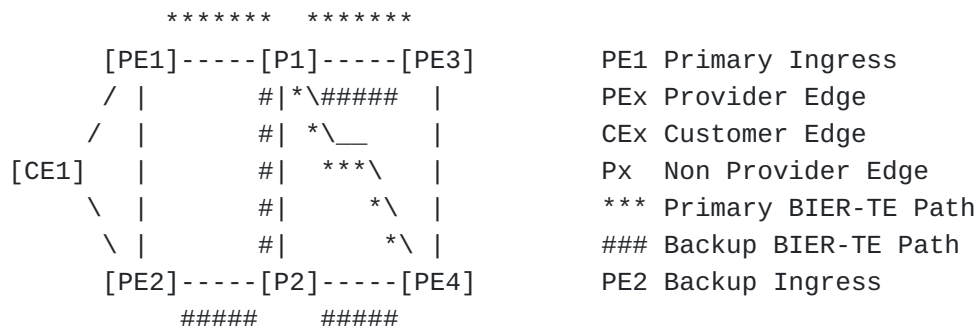
Figure 2: Protecting Ingress PE1 of BIER-TE Path

The backup BIER-TE path is from backup ingress PE2 to egress nodes
PE3 and PE4 through P2 and P1, which is represented by ### in the
figure.

In normal operations, CE1 sends the packets with a multicast group
and source to primary ingress PE1, which imports/encapsulates the
packets into the primary BIER-TE path through adding a BIER-TE
header. The header contains the primary BIER-TE path from primary
ingress PE1 to egress nodes PE3 and PE4. The packets are transmitted
to PE3 and PE4 along the primary BIER-TE path.

When CE1 detects the failure of primary ingress PE1 using a failure
detection mechanism such as BFD, it switches the traffic to backup
ingress PE2, which imports the traffic from CE1 into the backup
BIER-TE path. The traffic is sent to the egress nodes PE3 and PE4
along the backup BIER-TE path.

Given the traffic source (e.g., CE1), primary ingress (e.g., PE1)
and egresses (e.g., PE3 and PE4) of the primary BIER-TE path from
some PCEP messages, the PCE computes a backup ingress (e.g., PE2), a
backup BIER-TE path from the backup ingress to the egresses, and
sends the backup BIER-TE path to the PCC of the backup ingress in a
PCEP message. It also sends the information about the backup
ingress, the primary ingress and the traffic to the PCC of the
traffic source (e.g., CE1).

When the PCC of the traffic source receives the information about
the backup ingress, the primary ingress and the traffic, it sets up
the fast detection of the primary ingress failure and the switch
over target backup ingress. This setup lets the traffic source node

switch the traffic (to be sent to the primary ingress) to the backup
ingress when it detects the failure of the primary ingress.

When the PCC of the backup ingress receives the backup BIER-TE path,
it adds a forwarding entry into its BIFT. This entry encapsulates
the packets from the traffic source in the backup BIER-TE path. This
makes the backup ingress send the traffic received from the traffic
source to the egress nodes via the backup BIER-TE path.

3. Behavior around Ingress Failure

This section describes the behavior of some nodes connected to the
ingress before and after the ingress fails. These nodes are the
traffic source (e.g., CE1) and the backup ingress (e.g., PE2). It
presents three ways in which these nodes work together to protect
the ingress. The first way is called source detect, where the
traffic source is responsible for fast detecting the failure of the
ingress. The second way is called backup ingress detect, in which
the backup ingress is responsible for fast detecting the failure of
the ingress. The third way is called both detect, where both the
traffic source and the backup ingress are responsible for fast
detecting the failure of the ingress.

3.1. Source Detect

In normal operations, i.e., before the failure of the ingress of a
primary path such as a primary BIER-TE path, the traffic source
sends the traffic to the ingress of the primary path. The backup
ingress (e.g., PE2) is ready to import the traffic from the traffic
source into the backup path such as the backup BIER-TE path
installed.

When the traffic source detects the failure of the ingress, it
switches the traffic to the backup ingress, which delivers the
traffic to the egress nodes of the path via the backup path.

3.2. Backup Ingress Detect

The traffic source (e.g., CE1) always sends the traffic to both the
ingress (e.g., PE1) of the primary path such as the primary BIER-TE
path and the backup ingress (e.g., PE2).

The backup ingress does not import any traffic from the traffic
source into the backup path such as the backup BIER-TE path in
normal operations. When it detects the failure of the ingress of the
primary path, it imports the traffic from the source into the backup
path.

For the backup ingress to fast detect the failure of the primary
ingress, it SHOULD directly connect to the primary ingress. When a

PCE computes a backup ingress and a backup path, it SHOULD consider
this.

3.3. Both Detect

In normal operations, i.e., before the failure of the ingress, the
traffic source sends the traffic to the ingress of the primary path
such as the primary BIER-TE path. When it detects the failure of the
ingress, it switches the traffic to the backup ingress.

The backup ingress does not import any traffic from the traffic
source into the backup path such as the backup BIER-TE path in
normal operations. When it detects the failure of the ingress of the
primary path, it imports the traffic from the source into the backup
path.

4. Extensions to PCEP

A PCC runs on each of the edge nodes such as PEs of a network
normally. A 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 path. The path is a SR path, a BIER-TE path, or a path
of another type.

4.1. Capabilities for Ingress Protection

4.1.1. Capability for Ingress Protection with Backup Ingress

When a PCE and a PCC running on a backup ingress establish a PCEP
session between them, they exchange their capabilities of supporting
protection for the ingress node of each of different types of paths.

A new sub-TLV called INGRESS_PROTECTION_CAPABILITY is defined. It is
included in the PATH_SETUP_TYPE_CAPABILITY TLV with PST = TBD1
(suggested value 2 for path ingress protection) 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.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBD2          |            Length=4           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Reserved           | PathInd  |S|B|   Flags  |D|A|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
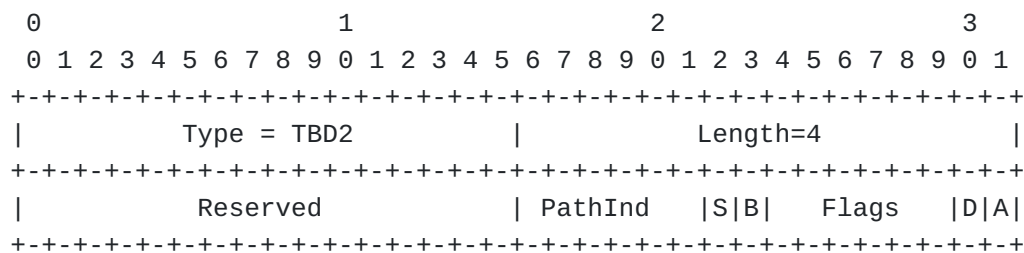
Figure 3: INGRESS_PROTECTION_CAPABILITY sub-TLV

Type: TBD2 is to be assigned by IANA.

Length:
4.

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

PathInd: 1 octet. Indicators for the types of paths whose ingress
   protections are supported. Two indicators are defined.

   o S : S = 1 indicating that the ingress protection of a SR path
      is supported.

   o B : B = 1 indicating that the ingress protection of a BIER-TE
      path is supported.

Flags: 1 octet. 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 path/tunnel be Active.

A PCC, which supports ingress protection for different types of
paths, sends a PCE an Open message containing
INGRESS_PROTECTION_CAPABILITY sub-TLV. This sub-TLV indicates that
the PCC is capable of supporting the ingress protection for the
types of paths.

For example, if a PCC supports ingress protection for SR path and
BIER-TE path, the PCC sends a PCE an Open message containing
INGRESS_PROTECTION_CAPABILITY sub-TLV with S = 1 and B = 1.

A PCE, which supports ingress protection for different types of
paths, sends a PCC an Open message containing
INGRESS_PROTECTION_CAPABILITY sub-TLV. This sub-TLV indicates that
the PCE is capable of supporting the ingress protection for the
types of paths.

If both a PCC and a PCE support INGRESS_PROTECTION_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 an
INGRESS_PROTECTION_CAPABILITY sub-TLV.

If a PCE receives an Open message from a PCC without a
INGRESS_PROTECTION_CAPABILITY sub-TLV indicating PCC's support for
the ingress protection of a type of paths, then the PCE MUST not
send the PCC any request for ingress protection of the type of
paths.

If a PCC receives an Open message from a PCE without a
INGRESS_PROTECTION_CAPABILITY sub-TLV indicating PCE's support for
the ingress protection of a type of paths, then the PCC MUST ignore
any request for ingress protection of the type of paths 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 and the fast detection of the
failure of the primary ingress MUST be done by the traffic source.
When the PCE sends the PCC a message for initiating a backup path,
the PCC MUST let the forwarding entry for the backup path be Active.

4.1.2. Capability for Ingress Protection with Traffic Source

When a PCE and a PCC running on a traffic source node establish a
PCEP session between them, they exchange their capabilities of
supporting ingress protection.

The PCECC-CAPABILITY sub-TLV defined in [RFC9050] is included in the
OPEN object in the PATH-SETUP-TYPE-CAPABILITY TLV, which is
exchanged in Open messages when a PCC and a PCE establish a PCEP
session between them.

A new flag bit P is defined in the Flags field of the PCECC-
CAPABILITY sub-TLV:

  *P flag (for Ingress Protection): if set to 1 by a PCEP speaker,
   the P flag indicates that the PCEP speaker supports and is
   willing to handle the PCECC based central controller instructions
   for ingress protection. The bit MUST be set to 1 by both a PCC
   and a PCE for the PCECC ingress protection instruction download/
   report on a PCEP session.

4.2. Extensions for Backup Ingress and Traffic Source

This section specifies the extensions to PCEP for the backup ingress
and the traffic source. The extensions let the traffic source

S1: fast detect the failure of the primary ingress and switch the
     traffic to the backup ingress when the traffic source detects the
     failure of the primary ingress, or

S2: always send the traffic to both the primary ingress and the
     backup ingress.

The extensions let the backup ingress

B1: always import the traffic received from the traffic source with
     possible service ID into the backup path, or

B2:
> import the traffic with possible service ID into the backup
> path when the backup ingress detects the failure of the primary
> ingress.

The following lists the combinations of Si and Bi (i = 1,2) for
different ways of failure detects.

Source Detect: S1 and B1.

Backup Ingress Detect: S2 and B2.

Both Detect: S1 and B2.

4.2.1. Extensions for Backup Ingress

For the packets from the traffic source, if the primary ingress
(i.e., the ingress of the primary path) encapsulates the packets
with a service ID or label into the path, the backup ingress MUST
have this service ID or label and encapsulates the packets with the
service ID or label into the backup path when the primary ingress
fails.

If the backup ingress is requested to detect the failure of the
primary ingress, it MUST have the information about the primary
ingress such as the address of the primary ingress.

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBD3          |         Length (variable)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Reserved           |            Flags            |A|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
~                      sub-TLVs (optional)                      ~
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
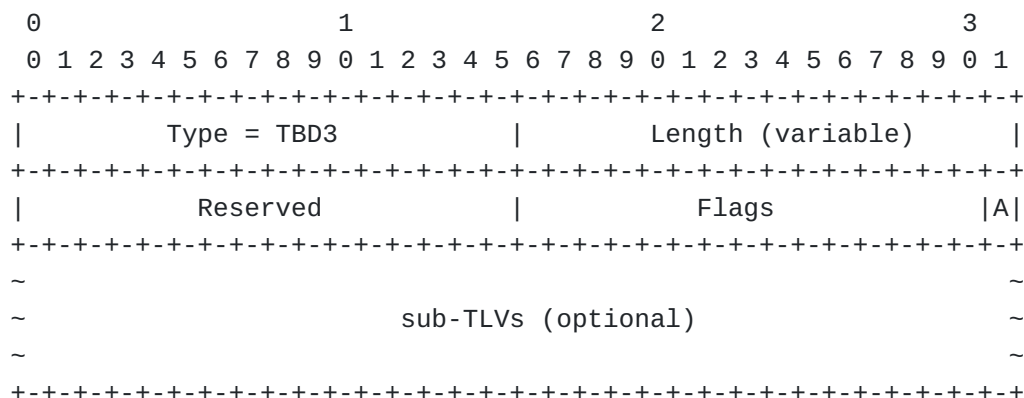
Figure 4: INGRESS_PROTECTION sub-TLV

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.

> A flag bit: it is set to 1 or 0 by PCE.
>
> o 1 is to request the backup ingress to let the forwarding entry for the backup path be Active always. In this case, the traffic source detects the failure of the primary ingress and switches the traffic to the backup ingress when it detects the failure.
>
> o 0 is to request the backup ingress to detect the failure of the primary ingress and let the forwarding entry for the backup path be Active when the primary ingress fails. In this case, the TLV includes the primary ingress address in a Primary-Ingress sub-TLV. The traffic source can send the traffic to both the primary ingress and the backup ingress. It may switch the traffic to the backup ingress from the primary ingress when it detects the failure of the primary ingress.

Three optional sub-TLVs are defined: Primary-Ingress sub-TLV, Service sub-TLV, and Traffic-Description sub-TLV. The Traffic-Description sub-TLV describes the traffic to be imported into the backup SR path. The Multicast Flow Specification TLV for IPv4 or IPv6, which is defined in [I-D.ietf-pce-pcep-flowspec], is used as a sub-TLV to indicate the traffic to be imported into the backup BIER-TE path.

4.2.1.1. Primary-Ingress sub-TLV

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Type = TBD4           |          Length (4)           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Primary Ingress IPv4 Address (4 octets)              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
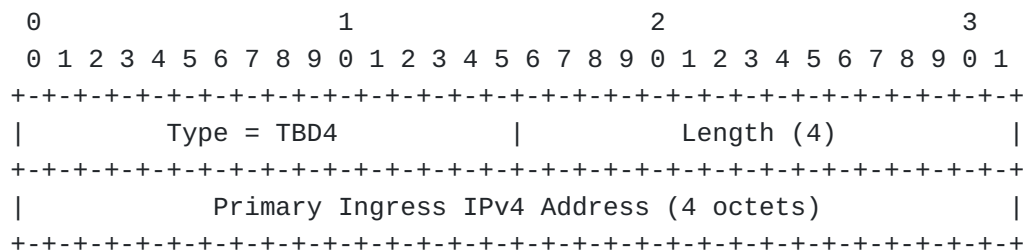
Figure 5: Primary Ingress IPv4 Address sub-TLV

Type:

TBD4 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 path.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type = TBD5         |           Length (16)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Primary Ingress IPv6 Address (16 octets)          |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
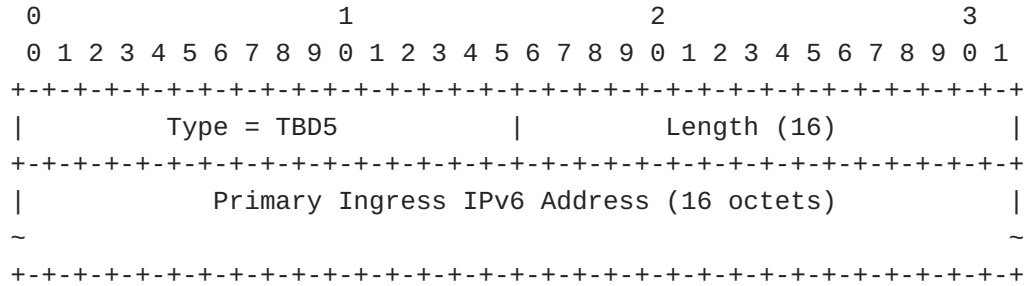
Figure 6: Primary Ingress IPv6 Address sub-TLV

Type: TBD5 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 path.

4.2.1.2. Service sub-TLV

A Service sub-TLV contains a service ID or label to be added into a
packet to be carried by a path. 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.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type = TBD6         |           Length (4)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          zero         |          Service Label (20 bits)      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
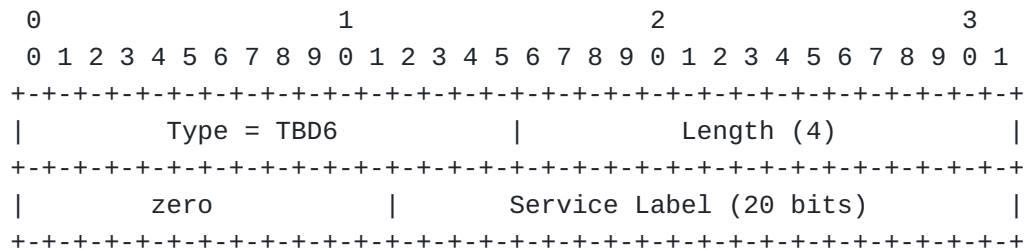
Figure 7: Service Label sub-TLV

Type:
       TBD6 is to be assigned by IANA.

Length: 4.

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBD7          |          Length (4/16)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Service ID (4 or 16 octets)                |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
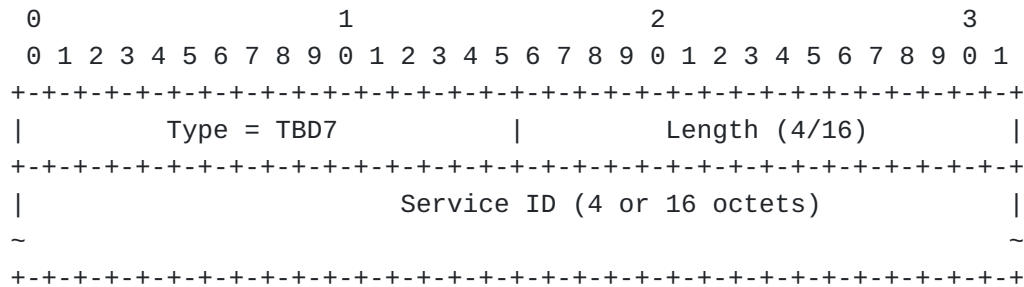
                  Figure 8: Service ID sub-TLV

Type: TBD7 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.

4.2.1.3. Traffic-Description sub-TLV

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBD8          |       Length (variable)       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
~                       sub-TLVs (optional)                     ~
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
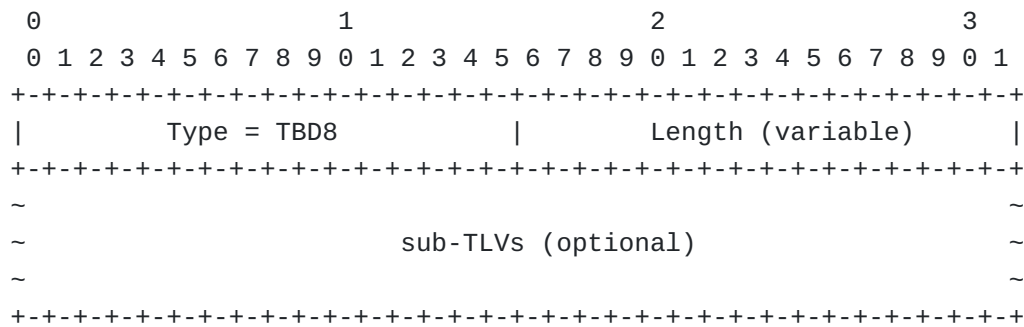
             Figure 9: Traffic-Description sub-TLV

Type: TBD8 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
path. 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.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Type = TBD9            |         Length (variable)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|IPv4 Prefix Len|       IPv4 Prefix                            ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~   (Optional) Virtual Network ID (2 octets)                   ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
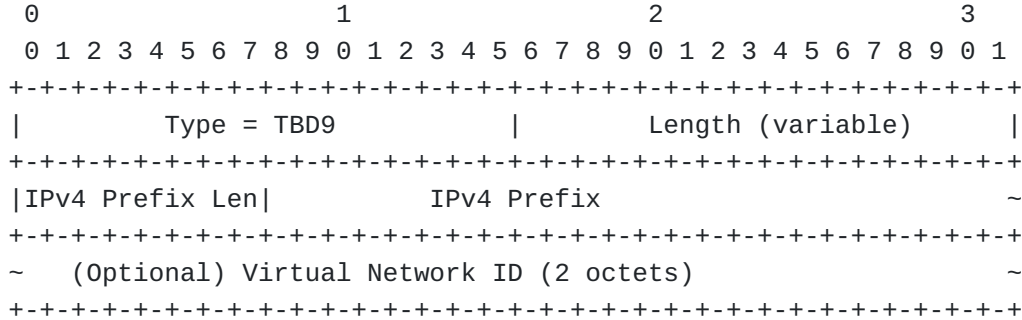
Figure 10: IPv4 FEC sub-TLV

Type: TBD9 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.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Type = TBDa            |         Length (variable)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|IPv6 Prefix Len|       IPv6 Prefix                            ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~   Optional Virtual Network ID (2 octets)                     ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
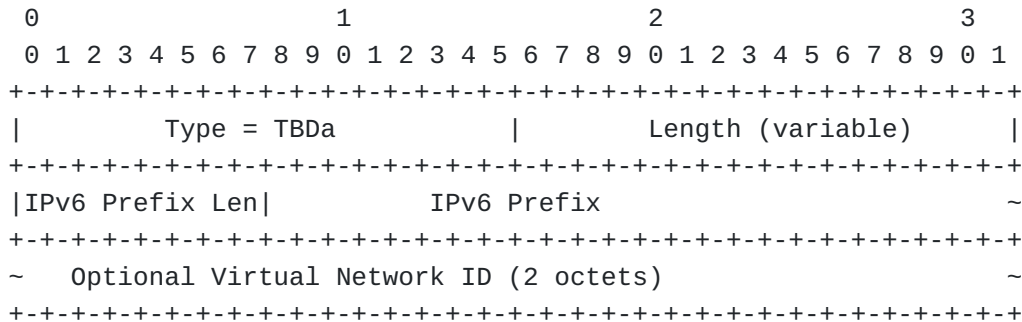
Figure 11: IPv6 FEC sub-TLV

Type: TBDa 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 path. It has three formats:
one for interface index, the other two for IPv4 and IPv6 address,
which are illustrated below.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBDb          |           Length (4)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Interface Index (4 octets)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
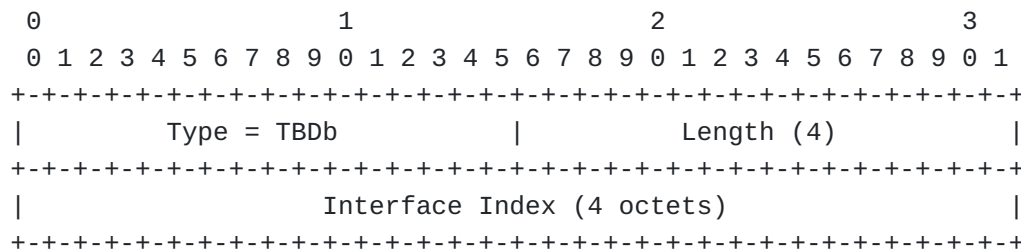
                Figure 12: Interface Index sub-TLV

Type: TBDb is to be assigned by IANA.

Length: 4.

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBDc          |           Length (4)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Interface IPv4 Address (4 octets)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
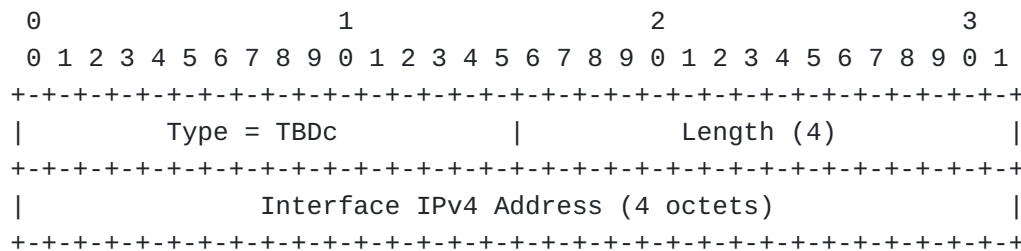
                Figure 13: Interface IPv4 Address sub-TLV

Type: TBDc is to be assigned by IANA.

Length: 4.

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type = TBDd          |          Length (16)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Interface IPv6 Address (16 octets)               |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
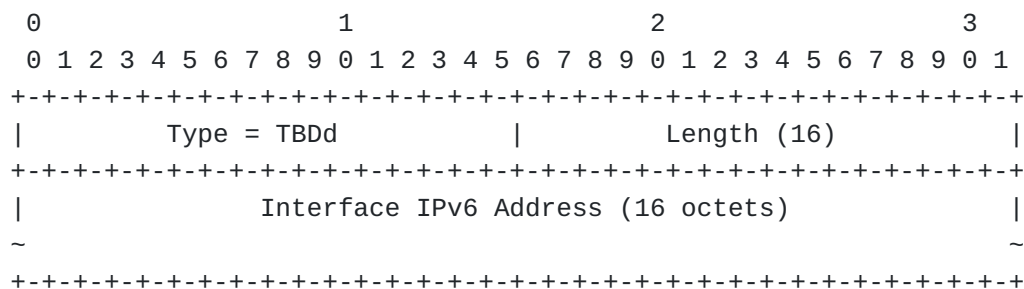
Figure 14: Interface IPv6 Address sub-TLV

Type: TBDd is to be assigned by IANA.

Length: 16.

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

4.2.2. Extensions for Traffic Source

If the traffic source is requested to detect the failure of the
primary ingress and switch the traffic (to be sent to the primary
ingress) to the backup ingress when the primary ingress fails, it
MUST have the information about the backup ingress, the primary
ingress and the traffic. This information may be transferred via a
CCI object for INGRESS-PROTECTION to the PCC of the traffic source
node from a PCE.

If the traffic source PCC does not accept the request from the PCE
or support the extensions, the PCE SHOULD have the information about
the behavior of the traffic source configured such as whether it
detects the failure of the primary ingress. Based on the
information, the PCE instructs the backup ingress accordingly.

The Central Control Instructions (CCI) Object is defined in
[RFC9050] for a PCE as a controller to send instructions for LSPs to
a PCC. This document defines a new object-type (TBDt) for ingress
protection based on the CCI object. The body of the object with the
new object-type is illustrated below. The object may be in PCRpt,
PCUpd, or PCInitiate message.

```
     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
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                            CC-ID                              |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |          Reserved             |           Flags        |B|D|
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    //                       Optional TLV                         //
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                  Figure 15: INGRESS-PROTECTION Object Body

   CC-ID: It is the same as described in [RFC9050].

   Flags: Two flag bits D and B are defined as follows:

      D: D = 1 instructs the PCC of the traffic source to Detect the
         failure of the primary ingress and switch the traffic to the
         backup ingress when it detects the failure.

      B: B = 1 instructs the PCC of the traffic source to send the
         traffic to Both the primary ingress and the backup ingress.

   Optional TLV: Primary ingress TLV, backup ingress TLV, Traffic-
      Description TLV or Multicast Flow Specification TLV.

   The primary ingress sub-TLV defined above is used as a TLV to
   contain the information about the primary ingress in the object. The
   Traffic-Description sub-TLV defined above is used as a TLV to
   contain the information about the traffic for a SR path in the
   object. The Multicast Flow Specification TLV for IPv4 or IPv6, which
   is defined in [I-D.ietf-pce-pcep-flowspec], is used to contain the
   information about the traffic for a BIER-TE path in the object. A
   new TLV, called backup ingress TLV, is defined to contain the
   information about the backup ingress in the object.

4.2.2.1. Backup-Ingress TLV

   A Backup-Ingress TLV indicates the IP address of the ingress node of
   a backup path. It has two formats: one for backup ingress node IPv4
   address and the other for backup ingress node IPv6 address, which
   are illustrated below. They have the same format as the Primary-
   Ingress sub-TLVs.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Type = TBDe          |           Length (4)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Backup Ingress IPv4 Address (4 octets)           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                 Figure 16: Backup Ingress IPv4 Address TLV

   Type: TBDe is to be assigned by IANA.

   Length: 4.

   Backup Ingress IPv4 Address: 4 octets. It represents an IPv4 host
      address of the backup ingress.


```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Type = TBDf          |           Length (16)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Backup Ingress IPv6 Address (16 octets)          |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                 Figure 17: Backup Ingress IPv6 Address TLV

   Type: TBDf is to be assigned by IANA.

   Length: 16.

   Backup Ingress IPv6 Address: 16 octets. It represents an IPv6 host
      address of the backup ingress node.

5. Security Considerations

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

   Note that this specification enables a network controller to
   instantiate a backup path in the network without the use of a hop-
   by-hop signaling protocol (such as RSVP-TE). This creates an
   additional vulnerability if the security mechanisms of [RFC5440],
   [RFC8231] and [RFC8281] are not used. If there is no integrity
   protection on the session, then an attacker could create a backup
   path which is not subjected to the further verification checks that
   would be performed by the signaling protocol.

6. Acknowledgements

The authors of this document would like to thank Dhruv Dhody and Robin Li for their reviews and comments.

7. IANA Considerations

TBD

8. References

8.1. 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>.

[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>.

[RFC9050]   Li, Z., Peng, S., Negi, M., Zhao, Q., and C. Zhou, "Path Computation Element Communication Protocol (PCEP) Procedures and Extensions for Using the PCE as a Central Controller (PCECC) of LSPs", RFC 9050, DOI 10.17487/RFC9050, July 2021, <https://www.rfc-editor.org/info/rfc9050>.

8.2. Informative References

[I-D.chen-bier-te-frr] Chen, H., McBride, M., Liu, Y., Wang, A., Mishra, G. S., Fan, Y., Liu, L., and X. Liu, "BIER-TE Fast ReRoute", Work in Progress, Internet-Draft, draft-chen-bier-te-frr-05, 30 July 2023, <https://datatracker.ietf.org/doc/html/draft-chen-bier-te-frr-05>.

[I-D.ietf-pce-pcep-flowspec] Dhody, D., Farrel, A., and Z. Li, "Path Computation Element Communication Protocol (PCEP)

Extension for Flow Specification", Work in Progress, Internet-Draft, draft-ietf-pce-pcep-flowspec-13, 14 October 2021, <https://datatracker.ietf.org/doc/html/draft-ietf-pce-pcep-flowspec-13>.

[I-D.ietf-rtgwg-segment-routing-ti-lfa]
Litkowski, S., Bashandy, A., Filsfils, C., Francois, P., Decraene, B., and D. Voyer, "Topology Independent Fast Reroute using Segment Routing", Work in Progress, Internet-Draft, draft-ietf-rtgwg-segment-routing-ti-lfa-11, 30 June 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-rtgwg-segment-routing-ti-lfa-11>.

[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>.

[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>.

Authors' Addresses

Huaimo Chen
Futurewei
Boston, MA,
United States of America

Email: Huaimo.chen@futurewei.com

Mike McBride
Futurewei

Email: michael.mcbride@futurewei.com

Mehmet Toy
Verizon Inc.
United States of America

Email: mehmet.toy@verizon.com

Gyan S. Mishra
Verizon Inc.
13101 Columbia Pike
Silver Spring, MD 20904
United States of America

Phone: 301 502-1347
Email: gyan.s.mishra@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

Yisong Liu
China Mobile

Email: liuyisong@chinamobile.com

Boris Khasanov
Yandex LLC
Moscow

Email: bhassanov@yahoo.com

Lei Liu
Fujitsu
United States of America

Email: liulei.kddi@gmail.com

Xufeng Liu
Volta Networks
McLean, VA
United States of America

Email: xufeng.liu.ietf@gmail.com