Network Working Group   D. Dhody, Ed.
Internet-Draft   Huawei Technologies
Intended status: Standards Track   A. Farrel, Ed.
Expires: July 2, 2018   Juniper Networks
                                                              Z. Li
                                                 Huawei Technologies
                                                  December 29, 2017

                    PCEP Extension for Flow Specification
                        draft-li-pce-pcep-flowspec-03

Abstract

   The Path Computation Element (PCE) is a functional component capable
   of selecting the paths through a traffic engineered network. These
   paths may be supplied in response to requests for computation, or may
   be unsolicited directions issued by the PCE to network elements.
   Both approaches use the PCE Communication Protocol (PCEP) to convey
   the details of the computed path.

   Traffic flows may be categorized and described using "Flow
   Specifications". RFC 5575 defines the Flow Specification and
   describes how it may be distributed in BGP to allow specific traffic
   flows to be associated with routes.

   This document specifies a set of extensions to PCEP to support
   dissemination of Flow Specifications. This allows a PCE to indicate
   what traffic should be placed on each path that it is aware of.

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 July 2, 2018.

Table of Contents

1. Introduction

   [RFC4655] defines the Path Computation Element (PCE), a functional
   component capable of computing paths for use in traffic engineering
   networks. PCE was originally conceived for use in Multiprotocol
   Label Switching (MPLS) for Traffic Engineering (TE) networks to
   derive the routes of Label Switched Paths (LSPs). However, the scope
   of PCE was quickly extended to make it applicable to Generalized MPLS
   (GMPLS) networks, and more recent work has brought other traffic
   engineering technologies and planning applications into scope (for
   example, Segment Routing (SR) [I-D.ietf-pce-segment-routing]).

   [RFC5440] describes the Path Computation Element Communication
   Protocol (PCEP). PCEP defines the communication between a Path
   Computation Client (PCC) and a PCE, or between PCE and PCE, enabling
   computation of path for MPLS-TE LSPs.

   Stateful PCE [RFC8231] specifies a set of extensions to PCEP to
   enable control of TE-LSPs by a PCE that retains state about the the
   LSPs provisioned in the network (a stateful PCE). [RFC8281]
   describes the setup, maintenance, and teardown of LSPs initiated by a
   stateful PCE without the need for local configuration on the PCC,
   thus allowing for a dynamic network that is centrally controlled.
   [RFC8283] introduces the architecture for PCE as a central controller
   and describes how PCE can be viewed as a component that performs
   computation to place 'flows' within the network and decide how these
   flows are routed.

   Dissemination of traffic flow specifications (Flow Specifications)
   was introduced for BGP in [RFC5575]. A Flow Specification is
   comprised of traffic filtering rules and actions. The routers that
   receive a Flow Specification can classify received packets according
   to the traffic filtering rules and can direct packets based on the
   actions.

When a PCE is used to initiate tunnels (such as TE-LSPs or SR paths)
using PCEP, it is important that the head end of the tunnels
understands what traffic to place on each tunnel. The data flows
intended for a tunnel can be described using Flow Specifications, and
when PCEP is in use for tunnel initiation it makes sense for that
same protocol to be used to distribute the Flow Specifications that
describe what data is to flow on those tunnels.

This document specifies a set of extensions to PCEP to support
dissemination of Flow Specifications. The extensions include the
creation, update, and withdrawal of Flow Specifications via PCEP and
can be applied to tunnels initiated by the PCE or to tunnels where
control is delegated to the PCE by the PCC. Furthermore, a PCC
requesting a new path can include Flow Specifications in the request
to indicate the purpose of the tunnel allowing the PCE to factor this
in during the path computation.

Flow Specifications are carried in TLVs within a new Flow Spec Object
defined in this document. The flow filtering rules indicated by the
Flow Specifications are mainly defined by BGP Flow Specifications.

2. Terminology

   This document uses the following terms defined in [RFC5440]: PCC,
   PCE, PCEP Peer.

   The following term from [RFC5575] is used frequently throughout this
   document:

      Flow Specification (FlowSpec): A Flow Specification is an n-tuple
      consisting of several matching criteria that can be applied to IP
      traffic, including filters and actions. Each FlowSpec consists of
      a set of filters and a set of actions.

   This document uses the terms "stateful PCE" and "active PCE" as
   advocated in [RFC7399].

3. Procedures for PCE Use of Flow Specifications

   There are three elements of procedure:

   o A PCE and a PCC must be able to indicate whether or not they
      support the use of Flow Specifications.

   o A PCE or PCC must be able to include Flow Specifications in PCEP
      messages with clear understanding of the applicability of those
      Flow Specifications in each case including whether the use of such

information is mandatory, constrained, or optional, and how
overlapping Flow Specifications will be resolved..

o   Flow Specification information/state must be synchronized between
    PCEP peers so that, on recovery, the peers have the same
    understanding of which Flow Specifications apply.

The following subsections describe these points.

3.1. Capability Advertisement

3.1.1. PCEP OPEN Message

During PCEP session establishment, a PCC or PCE that supports the
procedures described in this document announces this fact by
including the "PCE FlowSpec Capability" TLV (described in Section 4)
in the OPEN Object carried in the PCEP Open message.

The presence of the PCE FlowSpec Capability TLV in the OPEN Object in
a PCE's OPEN message indicates that the PCE can support distribute
the FlowSpec to PCCs and can receive FlowSpecs in messages from the
PCCs.

The presence of the PCE FlowSpec Capability TLV in the OPEN Object in
a PCC's OPEN message indicates that the PCC supports the FlowSpec
functionality described in this document.

If either one of a pair of PCEP peers does not indicate support of
the functionality described in this document by not including the PCE
FlowSpec Capability TLV in the OPEN Object in its OPEN message, then
the other peer MUST NOT include a FlowSpec object in any PCEP message
sent to the peer that does not support the procedures. If a FlowSpec
object is received even though support has not been indicated, the
receiver will respond with a PCErr message reporting the objects
containing the FlowSpec as described in [RFC5440]: that is, it will
use 'Unknown Object' if it does not support this specification, and
'Not supported object' if it supports this specification but has not
chosen to support FlowSpec objects on this PCEP session.

3.1.2. IGP PCE Capabilities Advertisement

The ability to advertise support for PCEP and PCE features in IGP
advertisements is provided for OSPF in [RFC5088] and for IS-IS in
[RFC5089]. The mechanism uses the PCE Discovery TLV which has a PCE-
CAP-FLAGS sub-TLV containing bit-flags each of which indicates
support for a different feature.

This document defines a new PCE-CAP-FLAGS sub-TLV bit, the FlowSpec
Capable flag (bit number TBD1). Setting the bit indicates that an
advertising PCE supports the procedures defined in this document.

Note that while PCE FlowSpec Capability may be advertised during
discovery, PCEP speakers that wish to use Flow Specification in PCEP
MUST negotiate PCE FlowSpec Capability during PCEP session setup, as
specified in Section 3.1.1. A PCC MAY initiate PCE FlowSpec
Capability negotiation at PCEP session setup even if it did not
receive any IGP PCE capability advertisement.

3.2. Dissemination Procedures

This section describes the procedures to support Flow Specifications
in PCEP messages.

The primary purpose of distributing Flow Specification information is
to allow a PCE to indicate to a PCC what traffic it should place on a
path (such as an LSP or an SR path). This means that the Flow
Specification may be included in:

o PCInitiate messages so that an active PCE can indicate the traffic
   to place on a path at the time that the PCE instantiates the path.

o PCUpd messages so that an active PCE can indicate or change the
   traffic to place on a path that has already been set up.

o PCRpt messages so that a PCC can report the traffic that the PCC
   plans to place on the path.

o PCReq messages so that a PCC can indicate what traffic it plans to
   place on a path at the time it requests the PCE to perform a
   computation in case that information aids the PCE in its work.

o PCRep messages so that a PCE that has been asked to compute a path
   can suggest which traffic could be placed on a path that a PCC may
   be about to set up.

o PCErr messages so that issues related to paths and the traffic
   they carry can be reported to the PCE by the PCC, and so that
   problems with other PCEP messages that carry Flow Specifications
   can be reported.

To carry Flow Specifications in PCEP messages, this document defines
a new PCEP object called the PCEP Flow Spec Object. The object is
OPTIONAL in the messages described above and MAY appear more than
once in each message.

The PCEP Flow Spec Object carries zero or one Flow Filter TLV which
describes a traffic flow.

The inclusion of multiple PCEP Flow Spec Objects allows multiple
traffic flows to be placed on a single path.

Once a PCE and PCC have established that they can both support the
use of Flow Specifications in PCEP messages, such information may be
exchanged at any time for new or existing paths.

The application and prioritization of Flow Specifications is
described in Section 8.7.

3.3. Flow Specification Synchronization

The Flow Specifications are carried along with the LSP State
information as per [RFC8231] making the Flow Specifications part of
the LSP database (LSP-DB). Thus, the synchronization of the Flow
Specification information is done as part of LSP-DB synchronization.
This may be achieved using normal state synchronization procedures as
described in [RFC8231] or enhanced state synchronization procedures
as defined in [RFC8232].

The approach selected will be implementation and deployment specific
and will depend on issues such as how the databases are constructed
and what level of synchronization support is needed.

4. PCE FlowSpec Capability TLV

The PCE-FLOWSPEC-CAPABILITY TLV is an optional TLV that can be
carried in the OPEN Object [RFC5440] to exchange PCE FlowSpec
capabilities of PCEP speakers.

The format of the PCE-FLOWSPEC-CAPABILITY TLV follows the format of
all PCEP TLVs as defined in [RFC5440] and is shown in Figure 1.

```
 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=2           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Value=0            |            Padding            |
+---------------------------------------------------------------+
```
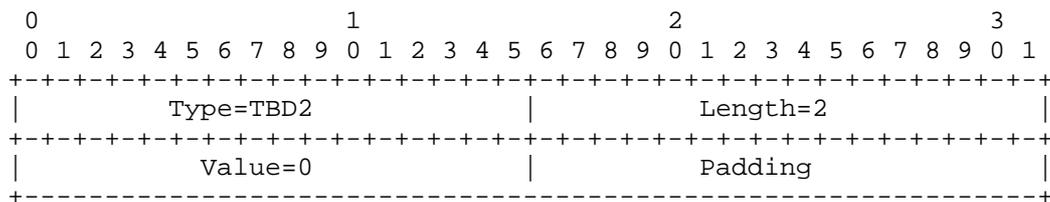
Figure 1: PCE-FLOWSPEC-CAPABILITY TLV format

The type of the PCE-FLOWSPEC-CAPABILITY TLV is TBD2 and it has a
fixed length of 2 octets. The Value field is set to default value 0.
The two bytes of padding MUST be set to zero and ignored on receipt.

The inclusion of this TLV in an OPEN object indicates that the sender
can perform FlowSpec handling as defined in this document.

5. PCEP Flow Spec Object

The PCEP Flow Spec object defined in this document is compliant with
the PCEP object format defined in [RFC5440]. It is OPTIONAL in the
PCReq, PCRep, PCErr, PCInitiate, PCRpt, and PCUpd messages and MAY be
present zero, one, or more times. Each instance of the object
specifies a traffic flow.

The PCEP Flow Spec object carries a FlowSpec filter rule encoded in a
TLV (as defined in Section 6.

The FLOW SPEC Object-Class is TBD3 (to be assigned by IANA).

The FLOW SPEC Object-Type is 1.

The format of the body of the PCEP Flow Spec object is shown in
Figure 2

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            FS-ID                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Reserved                           |R|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                              |
|                   Flow Filter TLV (variable)                 |
|                                                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
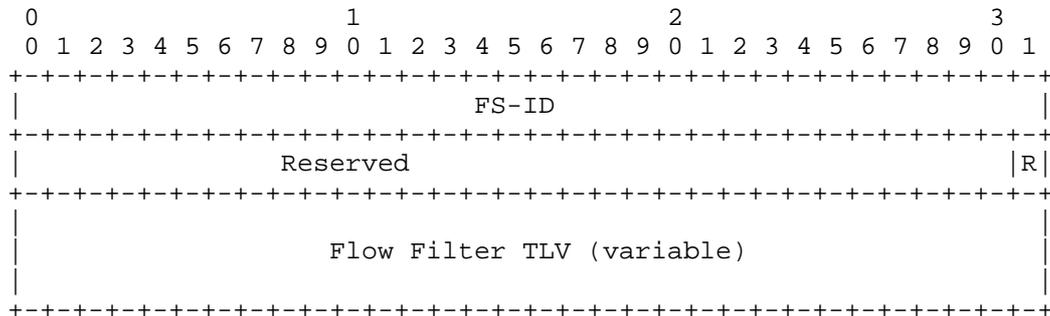
Figure 2: PCEP Flow Spec Object Body Format

FS-ID (32-bits): A PCEP-specific identifier for the FlowSpec
information. A PCE creates an FS-ID for each FlowSpec, the value is
unique within the scope of the PCE and is constant for the lifetime
of a PCEP session. All subsequent PCEP messages can identify the
FlowSpec using the FS-ID. The values 0 and 0xFFFFFFFF are reserved
and MUST NOT be used.

Reserved bits: MUST be set to zero on transmission and ignored on
receipt.

R bit: The Remove bit is set when a PCEP Flow Spec Object is included
in a PCEP message to indicate removal of the Flow Specification from
the associated tunnel. If the bit is clear, the Flow Specification
is being added or modified.

Flow Filter TLV (variable): One TLV MAY be included.

The Flow Filter TLV is OPTIONAL when the R bit is set. The TLV MUST
be present when the R bit is clear. If the TLV is missing when the R
bit is clear, the PCEP peer MUST respond with a PCErr message with
error-type TBD8 (FlowSpec Error), error-value 2 (Malformed FlowSpec).

6. Flow Filter TLV

A new PCEP TLV is defined to convey Flow Specification filtering
rules that specify what traffic is carried on a path. The TLV
follows the format of all PCEP TLVs as defined in [RFC5440]. The
Type field values come from the codepoint space for PCEP TLVs and has
the value TBD4.

The Value field contains one or more sub-TLVs (the Flow Specification
TLVs) as defined in Section 7. Only one Flow Filter TLV can be
present and represents the complete definition of a Flow
Specification for traffic to be placed on the tunnel indicated by the
PCEP message in which the PCEP Flow Spec Object is carried. The set
of Flow Specification TLVs in a single instance of a Flow Filter TLV
are combined to indicate the specific Flow Specification.

Further Flow Specifications can be included in a PCEP message by
including additional Flow Spec objects.

7. Flow Specification TLVs

Flow Filter TLV carries one or more Flow Specification sub-TLV. The
Flow Specification TLV also follows the format of all PCEP TLVs as
defined in [RFC5440], however, the Type values are selected from a
separate IANA registry (see Section 10) rather than from the common
PCEP TLV registry.

Type values are chosen so that there can be commonality with Flow
Specifications defined for use with BGP. This is possible because
the BGP Flow Spec encoding uses a single octet to encode the type
where PCEP uses two octets. Thus the space of values for the Type
field is partitioned as shown in Figure 3.

```
   Range            |
   -----------------+----------------------------------------------------
   0                | Reserved - must not be allocated.
                    |
   1 .. 255         | Per BGP registry defined by [RFC5575].
                    | Not to be allocated in this registry.
                    |
   256 ..    65535  | New PCEP Flow Specs allocated according to the
                    | registry defined in this document.
```

             Figure 3: Flow Specification TLV Type Ranges

The content of the Value field Flow in each TLV is specific to the
type and describes the parameters of the Flow Specification. The
definition of the format of many of these Value fields is inherited
from BGP specifications as shown in Figure 4. Specifically, the
inheritance is from [RFC5575] and [I-D.ietf-idr-flow-spec-v6], but
may also be inherited from future BGP specifications.

When multiple Flow Specification TLVs are present in a single Flow
Filter TLV they are combined to produce a more detailed description
of a flow. For examples and rules about how this is achieved, see
[RFC5575].

An implementation that receives a PCEP message carrying a Flow
Specification TLV with a type value that it does not recognize or
does not support MUST respond with a PCErr message with error-type
TBD8 (FlowSpec Error), error-value 1 (Unsupported FlowSpec) and MUST
NOT install the Flow Specification.

When used in other protocols (such as BGP) these Flow Specifications
are also associated with actions to indicate how traffic matching the
Flow Specification should be treated. In PCEP, however, the only
action is to associate the traffic with a tunnel and to forward
matching traffic on to that path, so no encoding of an action is
needed.

Section 8.7 describes how overlapping Flow Specifications are
prioritized and handled.

```
+-------+------------------------+---------------------------+
| Type  | Description            | Value defined in          |
|       |                        |                           |
+-------+------------------------+---------------------------+
|   *   | Destination IPv4 Prefix| [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | Source IPv4 Prefix     | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | IP Protocol            | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | Port                   | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | Destination port       | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | Source port            | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | ICMP type              | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | ICMP code              | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | TCP flags              | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | Packet length          | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | DSCP                   | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | Fragment               | [RFC5575]                 |
+-------+------------------------+---------------------------+
|   *   | Flow Label             | [I-D.ietf-idr-flow-spec-v6] |
+-------+------------------------+---------------------------+
|   *   | Destination IPv6 Prefix| [I-D.ietf-idr-flow-spec-v6] |
+-------+------------------------+---------------------------+
|   *   | Source IPv6 Prefix     | [I-D.ietf-idr-flow-spec-v6] |
+-------+------------------------+---------------------------+
|   *   | Next Header            | [I-D.ietf-idr-flow-spec-v6] |
+-------+------------------------+---------------------------+
| TBD5  | Route Distinguisher    | [I-D.dhodylee-pce-pcep-ls] |
+-------+------------------------+---------------------------+
| TBD6  | IPv4 Multicast Flow    | [This.I-D]                |
+-------+------------------------+---------------------------+
| TBD7  | IPv6 Multicast Flow    | [This.I-D]                |
+-------+------------------------+---------------------------+
```

   * Indicates that the TLV Type value comes from the value used
     in BGP.

                 Figure 4: Table of Flow Specification TLV Types

All Flow Specification TLVs with Types in the range 1 to 255 have
Values defined for use in BGP (for example in [RFC5575] and
[I-D.ietf-idr-flow-spec-v6]) and are set using the BGP encoding, but
without the type or length octets (the relevant information is in the
Type and Length fields of the TLV). The Value field is padded with
trailing zeros to achieve 4-byte alignment if necessary.

[I-D.dhodylee-pce-pcep-ls] defines a way to convey identification of
a VPN in PCEP via a Route Distinguisher (RD) [RFC4364] and encoded in
ROUTE-DISTINGUISHER TLV. A Flow Specification TLV with Type TBD5
carries a Value field matching that in the ROUTE-DISTINGUISHER TLV
and is used to identify that other flow filter information (for
example, an IPv4 destination prefix) is associated with a specific
VPN identified by the RD. See Section 8.6 for further discussion of
VPN identification.

Although it may be possible to describe a multicast Flow
Specification from the combination of other Flow Specification TLVs
with specific values, it is more convenient to use a dedicated Flow
Specification TLV. Flow Specification TLVs with Type values TBD6 and
TBD7 are used to identify a multicast flow for IPv4 and IPv6
respectively. The Value field is encoded as shown in Figure 5.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Rsvd    |S|W|R|    Rsvd    |B|Z| Src Mask Len | Grp Mask Len |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                        Source Address                         ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                   Group multicast Address                     ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
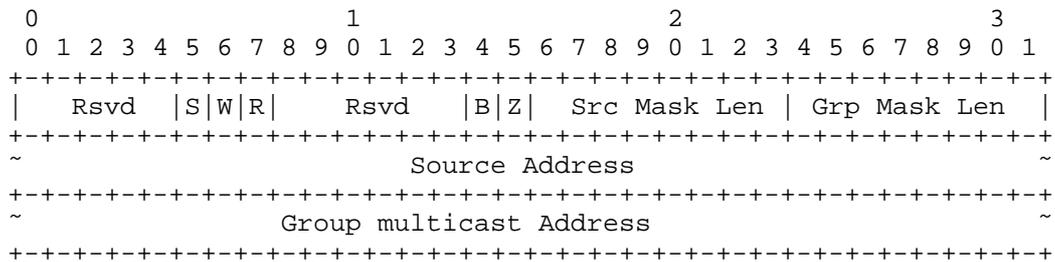
Figure 5: Multicast Flow Specification TLV Encoding

The fields of the two Multicast Flow Specification TLVs are as
described in Section 4.9.1 of [RFC7761] noting that the two address
fields are 32 bits for the IPv4 Multicast Flow and 128 bits for the
IPv6 Multicast Flow. Reserved fields MUST be set to zero and ignored
on receipt.

8. Detailed Procedures

This section outlines some specific detailed procedures for using the
protocol extensions defined in this document.

8.1. Default Behavior and Backward Compatibility

   The default behavior is that no Flow Specification is applied to a
   tunnel. That is, the default is that the Flow Spec object is not
   used as is the case in all systems before the implementation of this
   specification.

   In this case it is a local matter (such as through configuration) how
   tunnel head ends are instructed what traffic to place on a tunnel.

   [RFC5440]describes how receivers respond when they see unknown PCEP
   objects.

8.2. Composite Flow Specifications

   Flow Specifications may be represented by a single Flow Specification
   TLV or may require a more complex description using multiple Flow
   Specification TLVs. For example, a flow indicated by a source-
   destination pair of IPv6 addresses would be described by the
   combination of Destination IPv6 Prefix and Source IPv6 Prefix Flow
   Specification TLVs.

8.3. Modifying Flow Specifications

   A PCE may want to modify a Flow Specification associated with a
   tunnel, or a PCC may want to report a change to the Flow
   Specification it is using with a tunnel.

   It is important that the specific Flow Specification is identified so
   that it is clear that this is a modification of an existing flow and
   not the addition of a new flow as described in Section 8.4. The FS-
   ID field of the PCEP Flow Spec Object is used to identify a specific
   Flow Specification.

   When modifying a Flow Specification, all Flow Specification TLVs for
   the intended specification of the flow MUST be included in the PCEP
   Flow Spec Object and the FS-ID MUST be retained from the previous
   description of the flow.

8.4. Multiple Flow Specifications

   It is possible that multiple flows will be place on a single tunnel.
   In some cases it is possible to to define these within a single PCEP
   Flow Spec Object: for example, two Destination IPv4 Prefix TLVs could
   be included to indicate that packets matching either prefix are
   acceptable. PCEP would consider this as a single Flow Specification
   identified by a single FS-ID.

In other scenarios the use of multiple Flow Specification TLVs would
be confusing. For example, if flows from A to B and from C to D are
to be included then using two Source IPv4 Prefix TLVs and two
Destination IPv4 Prefix TLVs would be confusing (are flows from A to
D included?). In these cases, each Flow Specification is carried in
its own PCEP Flow Spec Object with multiple objects present on a
single PCEP message. Use of separate objects also allows easier
removal and modification of Flow Specifications.

8.5. Adding and Removing Flow Specifications

The Remove bit in the the PCEP Flow Spec Object is left clear when a
Flow Specification is being added or modified.

To remove a Flow Specification, a PCEP Flow Spec Object is included
with the FS-ID matching the one being removed, and the R bit set to
indicate removal. In this case it is not necessary to include any
Flow Specification TLVs.

If the R bit is set and Flow Specification TLVs are present an
implementation MAY ignore them. If the implementation checks the
Flow Specification TLVs against those recorded for the FS-ID of the
Flow Specification being removed and finds a mismatch, the Flow
Specification MUST still be removed and the implementation SHOULD
record a local exception or log.

8.6. VPN Identifiers

VPN instances are identified in BGP using Route Distinguishers (RDs)
[RFC4364]. These values are not normally considered to have any
meaning outside of the network, and they are not encoded in data
packets belonging to the VPNs. However, RDs provide a useful way of
identifying VPN instances and are often manually or automatically
assigned to VPNs as they are provisioned.

Thus the RD provides a useful way to indicate that traffic for a
particular VPN should be placed on a given tunnel. The tunnel head
end will need to interpret this Flow Specification not as a filter on
the fields of data packets, but using the other mechanisms that it
uses to identify VPN traffic. This could be based on the incoming
port (for port-based VPNs) or may leverage knowledge of the VRF that
is in use for the taffic.

8.7. Priorities and Overlapping Flow Specifications

TBD

An implementation that receives a PCEP message carrying a Flow
Specification that it cannot resolve against other Flow
Specifications already installed MUST respond with a PCErr message
with error-type TBD8 (FlowSpec Error), error-value 3 (Unresolvable
conflict) and MUST NOT install the Flow Specification.

9. PCEP Messages

The figures below use the notation defined in [RFC5511].

The FLOW SPEC Object is OPTIONAL and MAY be carried in the PCEP
messages.

The PCInitiate message is defined in [RFC8281] and updated as below:


```
<PCInitiate Message> ::= <Common Header>
                         <PCE-initiated-lsp-list>

Where:
   <PCE-initiated-lsp-list> ::= <PCE-initiated-lsp-request>
                                [<PCE-initiated-lsp-list>]

   <PCE-initiated-lsp-request> ::=
                                 ( <PCE-initiated-lsp-instantiation>|
                                  <PCE-initiated-lsp-deletion> )

   <PCE-initiated-lsp-instantiation> ::= <SRP>
                                         <LSP>
                                         [<END-POINTS>]
                                         <ERO>
                                         [<attribute-list>]
                                         [<flowspec-list>]

   Where:
      <flowspec-list> ::= <FLOWSPEC> [<flowspec-list>]
```


The PCUpd message is defined in [RFC8231] and updated as below:

```
   <PCUpd Message> ::= <Common Header>
                       <update-request-list>

   Where:
      <update-request-list> ::= <update-request>
                                  [<update-request-list>]

      <update-request> ::= <SRP>
                              <LSP>
                              <path>
                              [<flowspec-list>]

      Where:
         <path>::= <intended-path><intended-attribute-list>

         <flowspec-list> ::= <FLOWSPEC> [<flowspec-list>]
```

   The PCRpt message is defined in [RFC8231] and updated as below:

```
   <PCRpt Message> ::= <Common Header>
                       <state-report-list>

   Where:
      <state-report-list> ::= <state-report>[<state-report-list>]

      <state-report> ::= [<SRP>]
                            <LSP>
                            <path>
                            [<flowspec-list>]

       Where:
         <path>::= <intended-path>
                   [<actual-attribute-list><actual-path>]
                   <intended-attribute-list>

         <flowspec-list> ::= <FLOWSPEC> [<flowspec-list>]
```

   The PCReq message is defined in [RFC5440] and updated in [RFC8231],
   it is further updated below for flow specification:

```
   <PCReq Message>::= <Common Header>
                      [<svec-list>]
                      <request-list>

Where:
    <svec-list>::= <SVEC>[<svec-list>]

    <request-list>::= <request>[<request-list>]

    <request>::= <RP>
                 <END-POINTS>
                 [<LSP>]
                 [<LSPA>]
                 [<BANDWIDTH>]
                 [<metric-list>]
                 [<RRO>[<BANDWIDTH>]]
                 [<IRO>]
                 [<LOAD-BALANCING>]
                 [<flowspec-list>]

    Where:
        <flowspec-list> ::= <FLOWSPEC> [<flowspec-list>]
```

The PCRep message is defined in [RFC5440] and updated in [RFC8231],
it is further updated below for flow specification:

```
   <PCRep Message> ::= <Common Header>
                       <response-list>

Where:
    <response-list>::=<response>[<response-list>]

    <response>::=<RP>
                 [<LSP>]
                 [<NO-PATH>]
                 [<attribute-list>]
                 [<path-list>]
                 [<flowspec-list>]

    Where:
        <flowspec-list> ::= <FLOWSPEC> [<flowspec-list>]
```

10. IANA Considerations

   IANA maintains the "Path Computation Element Protocol (PCEP) Numbers"
   registry. This document requests IANA actions to allocate code
   points for the protocol elements defined in this document.

10.1. PCEP Objects

   Each PCEP object has an Object-Class and an Object-Type. IANA
   maintains a subregistry called "PCEP Objects". IANA is requested to
   make an assignment from this subregistry as follows:

Object-Class	Value Name	Object-Type	Reference
TBD3	FLOW SPEC	0 (Reserved)	[This.I-D]
		1	[This.I-D]

10.2. PCEP TLV Type Indicators

   IANA maintains a subregistry called "PCEP TLV Type Indicators". IANA
   is requested to make an assignment from this subregistry as follows:

Value	Meaning	Reference
TBD2	PCE-FLOWSPEC-CAPABILITY TLV	[This.I-D]
TBD4	FLOW FILTER TLV	[This.I-D]

10.3. Flow Specification TLV Type Indicators

   IANA is requested to create a new subregistry call the PCEP Flow
   Specification TLV Type Indicators registry.

   Allocations from this registry are to be made according to the
   following assignment policies [RFC8126]:

```
   Range           | Assignment policy
   ---------------+--------------------------------------------------
   0               | Reserved - must not be allocated.
                   |
   1 .. 255        | Reserved - must not be allocated.
                   | Usage mirrors the BGP FlowSpec registry [RFC5575].
                   |
   258 ..   64506  | Specification Required
                   |
   64507 .. 65531  | First Come First Served
                   |
   65532 .. 65535  | Experimental
```

   IANA is requested to pre-populate this registry with values defined
   in this document as follows:

```
   Value | Meaning
   ------+-----------------------
   TBD5  | Route Distinguisher
   TBD6  | IPv4 Multicast
   TBD7  | IPv6 Multicast
```

10.4. PCEP Error Codes

   IANA maintains a subregistry called "PCEP-ERROR Object Error Types
   and Values". Entries in this subregistry are described by Error-Type
   and Error-value. IANA is requested to make the following assignment
   from this subregistry:

```
   Error-| Meaning            | Error-value                | Reference
   Type  |                    |                            |
   ------+--------------------+----------------------------+-----------
   TBD8  | FlowSpec error     | 0: Unassigned              | [This.I-D]
         |                    | 1: Unsupported FlowSpec    | [This.I-D]
         |                    | 2: Malformed FlowSpec      | [This.I-D]
         |                    | 3: Unresolvable conflict   | [This.I-D]
         |                    | 4-255: Unassigned          | [This.I-D]
```

10.5. PCE Capability Flag

   IANA maintains a subregistry called "Open Shortest Path First v2
   (OSPFv2) Parameters" with a sub-registry called "Path Computation

   Element (PCE) Capability Flags". IANA is requested to assign a new
   capability bit from this registry as follows:

     Bit | Capability Description | Reference
     -------+-------------------------------+------------
     TBD1 | FlowSpec | [This.I-D]

11. Security Considerations

   We may assume that a system that utilizes a remote PCE is subject to
   a number of vulnerabilities that could allow spurious LSPs or SR
   paths to be established or that could result in existing paths being
   modified or torn down. Such systems, therefore, apply security
   considerations as described in [RFC5440], [RFC6952], and [RFC8253].

   The description of Flow Specifications associated with paths set up
   or controlled by a PCE add an further detail that could be attacked
   without tearing down LSPs or SR paths but causing traffic to be
   misrouted within the network. Therefore, the use of the security
   mechanisms for PCEP referenced above is important.

   Visibility into the information carried in PCEP does not have direct
   privacy concerns for end-users' data, however, knowledge of how data
   is routed in a network may make that data more vulnerable. Of
   course, the ability to interfere with the way data s routed also
   makes the data more vulnerable. Furthermore, knowledge of the
   connected end-points (such as multicast receivers or VPN sites) is
   usually considered private customer information. Therefore,
   implementations or deployments concerned to protect privacy MUST
   apply the mechanisms described in the documents referenced above.

   Experience with Flow Specifications in BGP systems indicates that
   they can become complex and that the overlap of Flow Specifications
   installed in different orders can lead to unexpected results.
   Although this is not directly a security issue per se, the confusion
   and unexpected forwarding behavior may be engineered or exploited by
   an attacker. Therefore, implementers and operators SHOULD pay
   careful attention to the Manageability Considerations described in
   Section 12.

12. Manageability Considerations

   TBD

13. Acknowledgements

   Thanks to Julian Lucek and Sudhir Cheruathur for useful discussions.

14. References

14.1. Normative References

   [I-D.dhodylee-pce-pcep-ls]
             Dhody, D., Lee, Y., and D. Ceccarelli, "PCEP Extension for
             Distribution of Link-State and TE Information.", draft-
             dhodylee-pce-pcep-ls-08 (work in progress), June 2017.

   [I-D.ietf-idr-flow-spec-v6]
             McPherson, D., Raszuk, R., Pithawala, B.,
             akarch@cisco.com, a., and S. Hares, "Dissemination of Flow
             Specification Rules for IPv6", draft-ietf-idr-flow-spec-
             v6-09 (work in progress), November 2017.

   [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>.

   [RFC5511] Farrel, A., "Routing Backus-Naur Form (RBNF): A Syntax
             Used to Form Encoding Rules in Various Routing Protocol
             Specifications", RFC 5511, DOI 10.17487/RFC5511, April
             2009, <https://www.rfc-editor.org/info/rfc5511>.

   [RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J.,
             and D. McPherson, "Dissemination of Flow Specification
             Rules", RFC 5575, DOI 10.17487/RFC5575, August 2009,
             <https://www.rfc-editor.org/info/rfc5575>.

   [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>.

   [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>.

14.2. Informative References

   [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-11 (work in progress),
             November 2017.

   [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private
             Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February
             2006, <https://www.rfc-editor.org/info/rfc4364>.

   [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>.

   [RFC5088] Le Roux, JL., Ed., Vasseur, JP., Ed., Ikejiri, Y., and R.
             Zhang, "OSPF Protocol Extensions for Path Computation
             Element (PCE) Discovery", RFC 5088, DOI 10.17487/RFC5088,
             January 2008, <https://www.rfc-editor.org/info/rfc5088>.

   [RFC5089] Le Roux, JL., Ed., Vasseur, JP., Ed., Ikejiri, Y., and R.
             Zhang, "IS-IS Protocol Extensions for Path Computation
             Element (PCE) Discovery", RFC 5089, DOI 10.17487/RFC5089,
             January 2008, <https://www.rfc-editor.org/info/rfc5089>.

   [RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of
             BGP, LDP, PCEP, and MSDP Issues According to the Keying
             and Authentication for Routing Protocols (KARP) Design
             Guide", RFC 6952, DOI 10.17487/RFC6952, May 2013,
             <https://www.rfc-editor.org/info/rfc6952>.

   [RFC7399] Farrel, A. and D. King, "Unanswered Questions in the Path
             Computation Element Architecture", RFC 7399,
             DOI 10.17487/RFC7399, October 2014,
             <https://www.rfc-editor.org/info/rfc7399>.

   [RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I.,
             Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent
             Multicast - Sparse Mode (PIM-SM): Protocol Specification
             (Revised)", STD 83, RFC 7761, DOI 10.17487/RFC7761, March
             2016, <https://www.rfc-editor.org/info/rfc7761>.

   [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>.

   [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>.

   [RFC8232] Crabbe, E., Minei, I., Medved, J., Varga, R., Zhang, X.,
              and D. Dhody, "Optimizations of Label Switched Path State
              Synchronization Procedures for a Stateful PCE", RFC 8232,
              DOI 10.17487/RFC8232, September 2017,
              <https://www.rfc-editor.org/info/rfc8232>.

   [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>.

   [RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An
              Architecture for Use of PCE and the PCE Communication
              Protocol (PCEP) in a Network with Central Control",
              RFC 8283, DOI 10.17487/RFC8283, December 2017,
              <https://www.rfc-editor.org/info/rfc8283>.

Appendix A. Contributors

   Shankara
   Huawei Technologies
   Divyashree Techno Park,
   Whitefield Bangalore,
   Karnataka
   560066
   India

   Email: shankara@huawei.com

   Qiandeng Liang
   Huawei Technologies
   101 Software Avenue,
   Yuhuatai District
   Nanjing
   210012
   China

   Email: liangqiandeng@huawei.com

   Cyril Margaria

      Juniper Networks
      200 Somerset Corporate Boulevard, Suite 4001
      Bridgewater, NJ
      08807
      USA

      Email: cmargaria@juniper.net

      Colby Barth
      Juniper Networks
      200 Somerset Corporate Boulevard, Suite 4001
      Bridgewater, NJ
      08807
      USA

      Email: cbarth@juniper.net

      Xia Chen
      Huawei Technologies
      Huawei Bld., No.156 Beiqing Rd.
      Beijing
      100095
      China

      Email: jescia.chenxia@huawei.com

      Shunwan Zhuang
      Huawei Technologies
      Huawei Bld., No.156 Beiqing Rd.
      Beijing
      100095
      China

      Email: zhuangshunwan@huawei.com

   Authors' Addresses

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

      Email: dhruv.ietf@gmail.com

Adrian Farrel (editor)
Juniper Networks

Email: afarrel@juniper.net

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

Email: lizhenbin@huawei.com