Network Working Group   S. Sivabalan
Internet-Draft   J. Medved
Intended status: Standards Track   C. Filsfils
Expires: November 20, 2015   Cisco Systems, Inc.
                                                           E. Crabbe

                                                           R. Raszuk
                                                         Mirantis Inc.
                                                            V. Lopez
                                                      Telefonica I+D
                                                         J. Tantsura
                                                            Ericsson
                                                       W. Henderickx
                                                       Alcatel Lucent
                                                         J. Hardwick
                                                  Metaswitch Networks
                                                        May 19, 2015

                   PCEP Extensions for Segment Routing
                   draft-ietf-pce-segment-routing-04.txt

Abstract

   Segment Routing (SR) enables any head-end node to select any path
   without relying on a hop-by-hop signaling technique (e.g., LDP or
   RSVP-TE). It depends only on "segments" that are advertised by Link-
   State Interior Gateway Protocols (IGPs). A Segment Routed Path can
   be derived from a variety of mechanisms, including an IGP Shortest
   Path Tree (SPT), explicit configuration, or a Path Computation
   Element (PCE). This document specifies extensions to the Path
   Computation Element Protocol (PCEP) that allow a stateful PCE to
   compute and initiate Traffic Engineering (TE) paths, as well as a PCC
   to request a path subject to certain constraint(s) and optimization
   criteria in SR networks.

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 [RFC2119].

Status of this Memo

   This Internet-Draft is submitted in full conformance with the
   provisions of BCP 78 and BCP 79.

   Internet-Drafts are working documents of the Internet Engineering

Sivabalan, et al.   Expires November 20, 2015   [Page 1]

Task Force (IETF). Note that other groups may also distribute
working documents as Internet-Drafts. The list of current Internet-
Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months
and may be updated, replaced, or obsoleted by other documents at any
time. It is inappropriate to use Internet-Drafts as reference
material or to cite them other than as "work in progress."

This Internet-Draft will expire on November 20, 2015.

Copyright Notice

Table of Contents

1. Introduction

   SR technology leverages the source routing and tunneling paradigms.
   A source node can choose a path without relying on hop-by-hop
   signaling protocols such as LDP or RSVP-TE. Each path is specified
   as a set of "segments" advertised by link-state routing protocols
   (IS-IS or OSPF). [I-D.filsfils-rtgwg-segment-routing] provides an
   introduction to SR architecture. The corresponding IS-IS and OSPF
   extensions are specified in
   [I-D.ietf-isis-segment-routing-extensions] and
   [I-D.ietf-ospf-segment-routing-extensions], respectively. SR
   architecture defines a "segment" as a piece of information advertised
   by a link-state routing protocols, e.g. an IGP prefix or an IGP
   adjacency. Several types of segments are defined. A Node segment
   represents an ECMP-aware shortest-path computed by IGP to a specific
   node, and is always global within SR/IGP domain. An Adjacency
   Segment represents unidirectional adjacency. An Adjacency Segment is
   local to the node which advertises it. Both Node segments and
   Adjacency segments can be used for SR Traffic Engineering (SR-TE).

   The SR architecture can be applied to the MPLS forwarding plane
   without any change, in which case an SR path corresponds to an MPLS
   Label Switching Path (LSP). This document is relevant to only MPLS
   forwarding plane, and assumes that a 32-bit Segment Identifier (SID)
   represents an absolute value of MPLS label entry. In this document,
   "Node-SID" and "Adjacency-SID" denote Node Segment Identifier and
   Adjacency Segment Identifier respectively.

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

   [RFC5440] describes Path Computation Element Protocol (PCEP) for
   communication between a Path Computation Client (PCC) and a Path
   Computation Element (PCE) or between one a pair of PCEs. A PCE or a
   PCC operating as a PCE (in hierarchical PCE environment) computes
   paths for MPLS Traffic Engineering LSPs (MPLS-TE LSPs) based on
   various constraints and optimization criteria.
   [I-D.ietf-pce-stateful-pce] specifies extensions to PCEP that allow a
   stateful PCE to compute and recommend network paths in compliance
   with [RFC4657] and defines objects and TLVs for MPLS-TE LSPs.
   Stateful PCEP extensions provide synchronization of LSP state between
   a PCC and a PCE or between a pair of PCEs, delegation of LSP control,
   reporting of LSP state from a PCC to a PCE, controlling the setup and
   path routing of an LSP from a PCE to a PCC. Stateful PCEP extensions
   are intended for an operational model in which LSPs are configured on
   the PCC, and control over them is delegated to the PCE.

A mechanism to dynamically initiate LSPs on a PCC based on the
requests from a stateful PCE or a controller using stateful PCE is
specified in [I-D.ietf-pce-pce-initiated-lsp]. Such mechanism is
useful in Software Driven Networks (SDN) applications, such as demand
engineering, or bandwidth calendaring.

It is possible to use a stateful PCE for computing one or more SR-TE
paths taking into account various constraints and objective
functions. Once a path is chosen, the stateful PCE can initiate an
SR-TE path on a PCC using PCEP extensions specified in
[I-D.ietf-pce-pce-initiated-lsp] using the SR specific PCEP
extensions described in this document. Additionally, using
procedures described in this document, a PCC can request an SR path
from either stateful or a stateless PCE. This specification relies
on the PATH-SETUP-TYPE TLV and procedures specified in
[I-D.ietf-pce-lsp-setup-type].

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

MSD: Maximum SID Depth

NAI: Node or Adjacency Identifier

OSPF: Open Shortest Path First

PCC: Path Computation Client

PCE: Path Computation Element

PCEP: Path Computation Element Protocol

RRO: Record Route Object

SID: Segment Identifier

   SR: Segment Routing

   SR-TE: Segment Routed Traffic Engineering

   TED: Traffic Engineering Database

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

   In SR networks, an ingress node of an SR path appends all outgoing
   packets with an SR header consisting of a list of SIDs (or MPLS
   labels in the context of this document). The header has all
   necessary information to guide the packets from the ingress node to
   the egress node of the path, and hence there is no need for any
   signaling protocol.

   In a PCEP session, LSP information is carried in the Explicit Route
   Object (ERO), which consists of a sequence of subobjects. Various
   types of ERO subobjects have been specified in [RFC3209], [RFC3473],
   and [RFC3477]. In SR networks, an ingress node of an SR path appends
   all outgoing packets with an SR header consisting of a list of SIDs
   (or MPLS labels in the context of this document). SR-TE LSPs
   computed by a PCE can be represented in one of the following forms:

   o An ordered set of IP address(es) representing network nodes/links:
      In this case, the PCC needs to convert the IP address(es) into the
      corresponding MPLS labels by consulting its Traffic Engineering
      Database (TED).

   o An ordered set of SID(s).

   o An ordered set of both MPLS label(s) and IP address(es): In this
      case, the PCC needs to convert the IP address(es) into the
      corresponding SID(s) by consulting its TED.

   This document defines a new ERO subobject denoted by "SR-ERO
   subobject" capable of carrying a SID as well as the identity of the
   node/adjacency represented by the SID. SR-capable PCEP speakers
   should be able to generate and/or process such ERO subobject. An ERO
   containing SR-ERO subobjects can be included in the PCEP Path
   Computation Reply (PCRep) message defined in [RFC5440], the PCEP LSP
   Initiate Request message (PCInitiate) defined in
   [I-D.ietf-pce-pce-initiated-lsp], as well as in the PCEP LSP Update
   Request (PCUpd) and PCEP LSP State Report (PCRpt) messages defined in
   defined in [I-D.ietf-pce-stateful-pce].

   When a PCEP session between a PCC and a PCE is established, both PCEP
   speakers exchange information to indicate their ability to support

SR-specific functionality. Furthermore, an LSP initially established
via RSVP-TE signaling can be updated with SR-TE path. This
capability is useful when a network is migrated from RSVP-TE to SR-TE
technology. Similarly, an LSP initially created with SR-TE path can
updated to signal the LSP using RSVP-TE if necessary.

A PCC MAY include an RRO object containing the recorded LSP in PCReq
and PCRpt messages as specified in [RFC5440] and
[I-D.ietf-pce-stateful-pce] respectively. This document defines a
new RRO subobject for SR networks. Methods used by a PCC to record
SR-TE LSP are outside the scope of this document.

In summary, this document:

o Defines a new PCEP capability, new ERO subobject, new RRO
   subobject, a new TLV, and new PCEP error codes.

o Specifies how two PCEP speakers can establish a PCEP session that
   can carry information about SR-TE paths.

o Specifies processing rules of ERO subobject.

o Defines a new path setup type carried in the PATH-SETUP-TYPE TLV
   for SR-TE LSP.

The extensions specified in this document complement the existing
PCEP specifications to support SR-TE path. As such, the PCEP
messages (e.g., Path Computation Request, Path Computation Reply,
Path Computation Report, Path Computation Update, Path Computation
Initiate, etc.,) MUST be formatted according to [RFC5440],
[I-D.ietf-pce-stateful-pce], [I-D.ietf-pce-pce-initiated-lsp], and
any other applicable PCEP specifications.

4. **SR-Specific PCEP Message Extensions**

As defined in [RFC5440], a PCEP message consists of a common header
followed by a variable length body made up of mandatory and/or
optional objects. This document does not require any changes in the
format of PCReq and PCRep messages specified in [RFC5440], PCInitiate
message specified in [I-D.ietf-pce-pce-initiated-lsp], and PCRpt and
PCUpd messages specified in [I-D.ietf-pce-stateful-pce]. However,
PCEP messages pertaining to SR-TE LSP MUST include PATH-SETUP-TYPE
TLV in the RP or SRP object to clearly identify that SR-TE LSP is
intended. In other words, a PCEP speaker MUST not infer whether or
not a PCEP message pertains to SR-TE LSP from any other object or
TLV.

5. Object Formats

5.1. The OPEN Object

   This document defines a new optional TLV for use in the OPEN Object.

5.1.1. The SR PCE Capability TLV

   The SR-PCE-CAPABILITY TLV is an optional TLV associated with the OPEN
   Object to exchange SR capability of PCEP speakers. The format of the
   SR-PCE-CAPABILITY TLV is shown in the following figure:

```
    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=TBD             |            Length=4           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |           Reserved             |     Flags     |     MSD       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
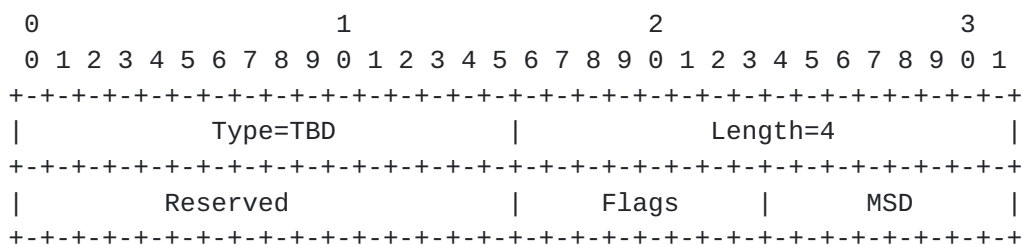
                   Figure 1: SR-PCE-CAPABILITY TLV format

   The code point for the TLV type is to be defined by IANA. The TLV
   length is 4 octets.

   The 32-bit value is formatted as follows. The "Maximum SID Depth" (1
   octet) field (MSD) specifies the maximum number of SIDs that a PCC is
   capable of imposing on a packet. The "Flags" (1 octet) and
   "Reserved" (2 octets) fields are currently unused, and MUST be set to
   zero on transmission and ignored on reception.

5.1.1.1. Exchanging SR Capability

   By including the SR-PCE-CAPABILITY TLV in the OPEN message destined
   to a PCE, a PCC indicates that it is capable of supporting the head-
   end functions for SR-TE LSP. By including the TLV in the OPEN
   message destined to a PCC, a PCE indicates that it is capable of
   computing SR-TE paths.

   The number of SIDs that can be imposed on a packet depends on PCC's
   data plane's capability. An MSD value of zero means that a PCC does
   not impose any default limitation on the number of SIDs included in
   any SR-TE path coming from PCE. Once an SR-capable PCEP session is
   established with a non-zero MSD value, the corresponding PCE MUST NOT
   send SR-TE paths with SIDs exceeding that MSD value. If a PCC needs
   to modify the MSD value, the PCEP session MUST be closed and re-
   established with the new MSD value. If a PCEP session is established

with a non-zero MSD value, and the PCC receives an SR-TE path
containing more SIDs than specified in the MSD value, the PCC MUST
send a PCErr message with Error-Type 10 (Reception of an invalid
object) and Error-Value 3 (Unsupported number of Segment ERO). If a
PCEP session is established with an MSD value of zero, then the PCC
MAY specify an MSD for each path computation request that it sends to
the PCE.

The SR Capability TLV is meaningful only in the OPEN message sent
from a PCC to a PCE. As such, a PCE does not need to set MSD value
in outbound message to a PCC. Similarly, a PCC ignores any MSD value
received from a PCE. If a PCE receives multiple SR-PCE-CAPABILITY
TLVs in an OPEN message, it processes only the first TLV is
processed.

5.2. The RP/SRP Object

In order to setup an SR-TE LSP using SR, RP or SRP object MUST PATH-
SETUP-TYPE TLV specified in [I-D.ietf-pce-lsp-setup-type]. This
document defines a new Path Setup Type (PST) for SR as follows:

o PST = 1: Path is setup using Segment Routing Traffic Engineering
   technique.

5.3. ERO Object

An SR-TE path consists of one or more SID(s) where each SID MAY be
associated with the identifier that represents the node or adjacency
corresponding to the SID. This identifier is referred to as the
'Node or Adjacency Identifier' (NAI). As described later, a NAI can
be represented in various formats (e.g., IPv4 address, IPv6 address,
etc). Furthermore, a NAI is used for troubleshooting purposes and,
if necessary, to derive SID value as described below.

The ERO object specified in [RFC5440] is used to carry SR-TE path
information. In order to carry SID and/or NAI, this document defines
a new ERO subobject referred to as "SR-ERO subobject" whose format is
specified in the following section. An ERO object carrying an SR-TE
path consists of one or more ERO subobject(s), and MUST carry only
SR-ERO subobject. Note that an SR-ERO subobject does not need to
have both SID and NAI. However, at least one of them MUST be
present.

When building the MPLS label stack from ERO, a PCC MUST assume that
SR-ERO subobjects are organized as a last-in-first-out stack. The
first subobject relative to the beginning of ERO contains the
information about the topmost label. The last subobject contains
information about the bottommost label.

5.3.1. SR-ERO Subobject

   An SR-ERO subobject consists of a 32-bit header followed by the SID
   and the NAI associated with the SID. The SID is a 32-bit number.
   The size of the NAI depends on its respective type, as described in
   the following sections.

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |L|   Type    |   Length    | ST  |     Flags     |F|S|C|M|
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                             SID                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   //                      NAI (variable)                        //
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
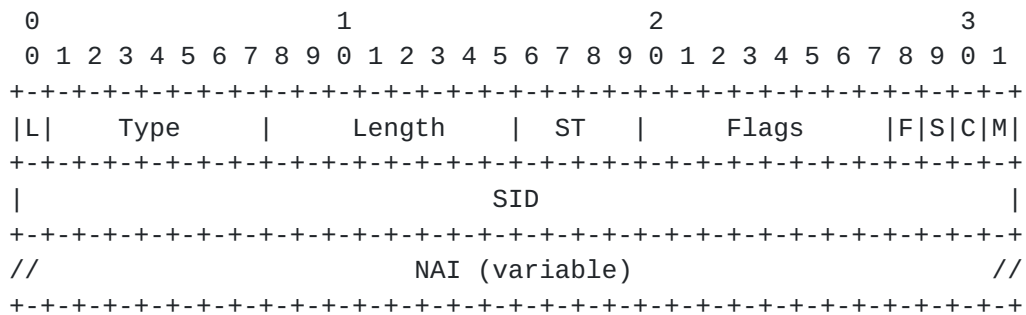
                    Figure 2: SR-ERO Subobject format

   The fields in the SR-ERO Subobject are as follows:

   The 'L' Flag indicates whether the subobject represents a loose-hop
      in the LSP [RFC3209]. If this flag is unset, a PCC MUST not
      overwrite the SID value present in the SR-ERO subobject.
      Otherwise, a PCC MAY expand or replace one or more SID value(s) in
      the received SR-ERO based on its local policy.

   Type is the type of the SR-ERO subobject. This document defines the
      SR-ERO subobject type, and requests a new codepoint from IANA.

   Length contains the total length of the subobject in octets,
      including the L, Type and Length fields. Length MUST be at least
      8, and MUST be a multiple of 4. As mentioned earlier, an SR-ERO
      subobject MUST have at least SID or NAI. The length should take
      into consideration SID or NAI only if they are not null. The
      flags described below used to indicate whether SID or NAI field is
      null.

   SID Type (ST) indicates the type of information associated with the
      SID contained in the object body. When ST value is 0, SID MUST
      NOT be null and NAI MUST be null. Other ST values are described
      later in this document.

Flags is used to carry any additional information pertaining to SID.
   Currently, the following flag bits are defined:

   * M: When this bit is set, the SID value represents an MPLS label
      stack entry as specified in [RFC5462] where only the label
      value is specified by the PCE. Other fields (TC, S, and TTL)
      fields MUST be considered invalid, and PCC MUST set these
      fields according to its local policy and MPLS forwarding rules.

   * C: When this bit as well as the M bit are set, then the SID
      value represents an MPLS label stack entry as specified in
      [RFC5462], where all the entry's fields (Label, TC, S, and TTL)
      are specified by the PCE. However, a PCC MAY choose to
      override TC, S, and TTL values according its local policy and
      MPLS forwarding rules.

   * S: When this bit is set, the SID value in the subobject body is
      null. In this case, the PCC is responsible for choosing the
      SID value, e.g., by looking up its TED using the NAI which, in
      this case, MUST be present in the subobject.

   * F: When this bit is set, the NAI value in the subobject body is
      null.

SID is the Segment Identifier.

NAI contains the NAI associated with the SID. Depending on the
   value of ST, the NAI can have different format as described in the
   following section.

5.3.2. NAI Associated with SID

This document defines the following NAIs:

'IPv4 Node ID' is specified as an IPv4 address. In this case, ST
   value is 1, and the Length is 8 or 12 depending on either SID or
   NAI or both are included in the subobject.

'IPv6 Node ID' is specified as an IPv6 address. In this case, ST
   and Length are 2, and Length is 8, 20, or 24 depending on either
   SID or NAI or both are included in the subobject.

'IPv4 Adjacency' is specified as a pair of IPv4 addresses. In this
   case, ST value is 3. The Length is 8, 12, or 16 depending on
   either SID or NAI or both are included in the subobject, and the
   format of the NAI is shown in the following figure:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Local IPv4 address                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Remote IPv4 address                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                 Figure 3: NAI for IPv4 Adjacency

'IPv6 Adjacency' is specified as a pair of IPv6 addresses. In this
   case, ST valie is 4. The Length is 8, 36 or 40 depending on
   whether SID or NAI or both included in the subobject,and the
   format of the NAI is shown in the following figure:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//            Local IPv6 address (16 bytes)                    //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//            Remote IPv6 address (16 bytes)                   //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                 Figure 4: NAI for IPv6 adjacenc y

'Unnumbered Adjacency with IPv4 NodeIDs' is specified as a pair of
   Node ID / Interface ID tuples. In this case, ST value is 5. The
   Length is 8, 20, or 24 depending on whether SID or NAI or both
   included in the subobject, and the format of the NAI is shown in
   the following figure:
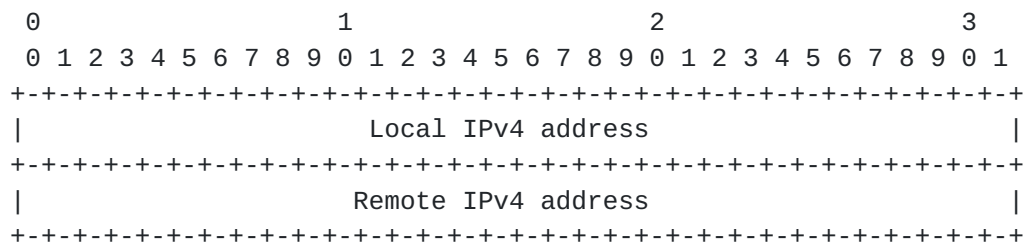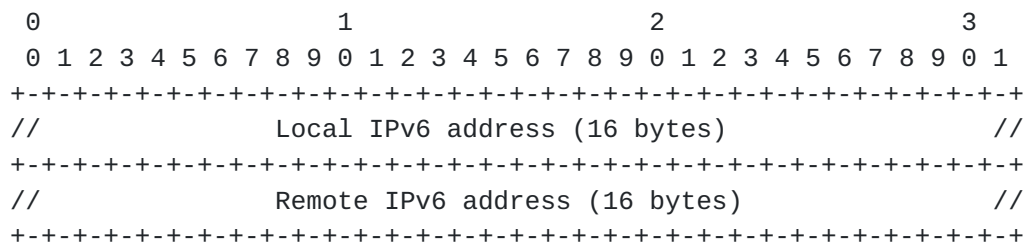
```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Local Node-ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Local Interface ID                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Remote Node-ID                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Remote Interface ID                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
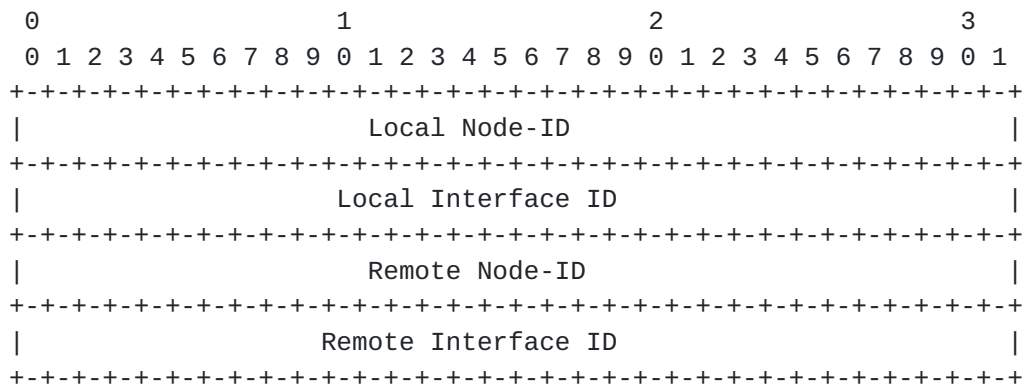
           Figure 5: NAI for Unnumbered adjacency with IPv4 Node IDs

   Editorial Note: We are yet to decide if another SID subobject is
   required for unnumbered adjacency with 128 bit node ID.

5.3.3. ERO Processing

   A PCEP speaker that does not recognize the SR-ERO subobject in PCRep,
   PCInitiate, PCUpd or PCRpt messages MUST reject the entire PCEP
   message and MUST send a PCErr message with Error-Type=3 ("Unknown
   Object") and Error-Value=2 ("Unrecognized object Type") or Error-
   Type=4 ("Not supported object") and Error-Value=2 ("Not supported
   object Type"), defined in [RFC5440].

   When the SID represents an MPLS label (i.e. the M bit is set), its
   value (20 most significant bits) MUST be larger than 15, unless it is
   special purpose label, such as an Entropy Label Indicator (ELI). If
   a PCEP speaker receives a label ERO subobject with an invalid value,
   it MUST send a PCErr message with Error-Type = 10 ("Reception of an
   invalid object") and Error Value = TBD ("Bad label value"). If both
   M and C bits of an ERO subobject are set, and if a PCEP speaker finds
   erroneous setting in one or more of TC, S, and TTL fields, it MUST
   send a PCErr message with Error-Type = 10 ("Reception of an invalid
   object") and Error-Value = TBD ("Bad label format").

   If a PCC receives a stack of SR-ERO subobjects, and the number of
   stack exceeds the maximum number of SIDs that the PCC can impose on
   the packet, it MAY send a PCErr message with Error-Type = 10
   ("Reception of an invalid object") and Error-Value = TBD
   ("Unsupported number of Segment ERO subobjects").

   When a PCEP speaker detects that all subobjects of ERO are not
   identical, and if it does not handle such ERO, it MUST send a PCErr
   message with Error-Type = 10 ("Reception of an invalid object") and
   Error-Value = TBD ("Non-identical ERO subobjects").

If a PCEP speaker receives an SR-ERO subobject in which both SID and
NAI are absent, it MUST consider the entire ERO object invalid and
send a PCErr message with Error-Type = 10 ("Reception of an invalid
object") and Error-Value = TBD ("Both SID and NAI are absent in ERO
subobject").

When a PCEP speaker receives an SR-ERO subobject in which ST is 0,
SID MUST be present and NAI MUST NOT be present(i.e., S-flag MUST be
0, F-flag MUST be 1, and the Length MUST be 8). Otherwise, it MUST
consider the entire ERO object invalid and send a PCErr message with
Error-Type = 10 ("Reception of an invalid object") and Error-Value =
TBD ("Malformed object"). The PCEP speaker MAY include the malformed
SR-ERO object in the PCErr message as well.

5.4. RRO Object

A PCC can record SR-TE LSP and report the LSP to a PCE via RRO. An
RRO object contains one or more subobjects called "SR-RRO subobjects"
whose format is shown 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      |     Length    |  ST   |      Flags    |F|S|C|M|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              SID                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                        NAI (variable)                       //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
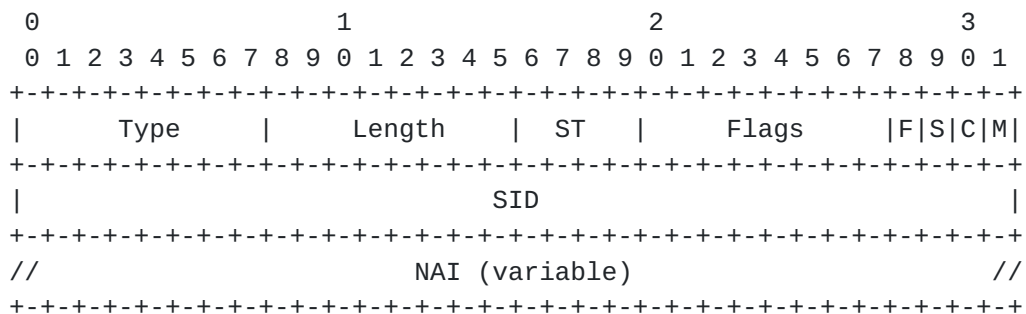
Figure 6: SR-RRO Subobject format

The format of SR-RRO subobject is the same as that of SR-ERO
subobject without L flag.

A PCC MUST assume that SR-RRO subobjects are organized such that the
first subobject relative to the beginning of RRO contains the
information about the topmost label, and the last subobject contains
information about the bottommost label of the SR-TE LSP.

5.4.1. RRO Processing

Processing rules of SR-RRO subobject are identical to those of SR-ERO
subobject.

If a PCEP speaker receives an SR-RRO subobject in which both SID and
NAI are absent, it MUST consider the entire RRO object invalid and
send a PCErr message with Error-Type = 10 ("Reception of an invalid

object") and Error-Value = TBD ("Both SID and NAI are absent in RRO
subobject").

If a PCE detects that all subobjects of RRO are not identical, and if
it does not handle such RRO, it MUST send a PCErr message with Error-
Type = 10 ("Reception of an invalid object") and Error-Value = TBD
("Non-identical RRO subobjects").

5.5. METRIC Object

If a PCEP session is established with an MSD value of zero, then the
PCC MAY specify the MSD for an individual path computation request
using the METRIC object defined in [RFC5440]. This document defines
a new type for the METRIC object to be used for this purpose as
follows:

o T = TBD (suggested value 11): Maximum SID Depth of the requested
   path.

The PCC sets the metric-value to the MSD for this path. The PCC MUST
set the B (bound) bit to 1 in the METRIC object, which specifies that
the SID depth for the computed path MUST NOT exceed the metric-value.

If a PCEP session is established with a non-zero MSD value, then the
PCC MUST NOT send an MSD METRIC object. If the PCE receives a path
computation request with an MSD METRIC object on a session with a
non-zero MSD value then it MUST consider the request invalid and send
a PCErr with Error-Type = 10 ("Reception of an invalid object") and
Error-Value TBD ("Default MSD is specified for the PCEP session").

6. Backward Compatibility

A PCEP speaker that does not support the SR PCEP capability cannot
recognize the SR-ERO or SR-RRO subobjects. As such, it MUST send a
PCEP error with Error-Type = 4 (Not supported object) and Error-Value
= 2 (Not supported object Type) as per [RFC5440].

7. Management Considerations

7.1. Policy

PCEP implementation:

o Can enable SR PCEP capability either by default or via explicit
   configuration.

   o May generate PCEP error due to unsupported number of SR-ERO or SR-
      RRO subobjects either by default or via explicit configuration.

7.2. The PCEP Data Model

   A PCEP MIB module is defined in [I-D.ietf-pce-pcep-mib] needs be
   extended to cover additional functionality provided by [RFC5440] and
   [I-D.ietf-pce-pce-initiated-lsp]. Such extension will cover the new
   functionality specified in this document.

8. Security Considerations

   The security considerations described in [RFC5440] and
   [I-D.ietf-pce-pce-initiated-lsp] are applicable to this
   specification. No additional security measure is required.

9. IANA Considerations

9.1. PCEP Objects

   This document defines a new sub-object type for the PCEP explicit
   route object (ERO), and a new sub-object type for the PCEP record
   route object (RRO). The code points for sub-object types of these
   objects is maintained in the RSVP parameters registry, under the
   EXPLICIT_ROUTE and ROUTE_RECORD objects. IANA is requested to
   allocate code points in the RSVP Parameters registry for each of the
   new sub-object types defined in this document, as follows:

Object	Sub-Object	Sub-Object Type
EXPLICIT_ROUTE	SR-ERO	TBD (recommended 5)
ROUTE_RECORD	SR-RRO	TBD (recommended 6)

9.2. PCEP-Error Object

   IANA is requested to allocate code-points in the PCEP-ERROR Object
   Error Types and Values registry for the following new error-values:

Error-Type	Meaning
10	Reception of an invalid object.
	Error-value = TBD (recommended 2): Bad label value

```
                    Error-value = TBD (recommended 3):  Unsupported number
                                                        of Segment ERO
                                                        subobjects
                    Error-value = TBD (recommended 4):  Bad label format
                    Error-value = TBD (recommended 5):  Non-identical ERO
                                                        subobjects
                    Error-value = TBD (recommended 6):  Both SID and NAI
                                                        are absent in ERO
                                                        subobject
                    Error-value = TBD (recommended 7):  Both SID and NAI
                                                        are absent in RRO
                                                        subobject
                    Error-value = TBD (recommended 9):  Default MSD is
                                                        specified for the
                                                        PCEP session
                    Error-value = TBD (recommended 10): Non-identical RRO
                                                        subobjects
                    Error-value = TBD (recommended 11): Malformed object
```

9.3. PCEP TLV Type Indicators

IANA is requested to allocate a new code point in the PCEP TLV Type
Indicators registry, as follows:

```
   Value                    Meaning                       Reference
   ------------------------ ---------------------------- --------------
   TBD (recommended 26)     SR-PCE-CAPABILITY             This document
```

9.4. New Path Setup Type

[I-D.ietf-pce-lsp-setup-type] defines the PATH-SETUP-TYPE TLV and
requests that IANA creates a registry to manage the value of the
PATH_SETUP_TYPE TLV's PST field. IANA is requested to allocate a new
code point in the PCEP PATH_SETUP_TYPE TLV PST field registry, as
follows:

```
   Value                    Description                   Reference
   ------------------------ ---------------------------- --------------
   1                        Traffic engineering path is   This document
                            setup using Segment Routing
                            technique.
```

9.5. New Metric Type

IANA is requested to allocate a new code point in the PCEP METRIC
object T field registry, as follows:

Value	Description	Reference
TBD (recommended 11)	Segment-ID (SID) Depth.	This document

10. Contributors

The following people contributed to this document:
   - Lakshmi Sharma (Cisco Systems)

11. Acknowledgements

We like to thank Ina Minei, George Swallow, Marek Zavodsky and Tomas
Janciga for the valuable comments.

12. References

12.1. Normative References

[I-D.filsfils-rtgwg-segment-routing]
          Filsfils, C., Previdi, S., Bashandy, A., Decraene, B.,
          Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R.,
          Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe,
          "Segment Routing Architecture",
          draft-filsfils-rtgwg-segment-routing-01 (work in
          progress), October 2013.

[I-D.ietf-isis-segment-routing-extensions]
          Previdi, S., Filsfils, C., Bashandy, A., Gredler, H.,
          Litkowski, S., and J. Tantsura, "IS-IS Extensions for
          Segment Routing",
          draft-ietf-isis-segment-routing-extensions-00 (work in
          progress), April 2014.

[I-D.ietf-ospf-segment-routing-extensions]
          Psenak, P., Previdi, S., Filsfils, C., Gredler, H.,
          Shakir, R., Henderickx, W., and J. Tantsura, "OSPF
          Extensions for Segment Routing",
          draft-ietf-ospf-segment-routing-extensions-00 (work in
          progress), June 2014.

[I-D.ietf-pce-lsp-setup-type]
          Sivabalan, S., Medved, J., Minei, I., Crabbe, E., and R.
          Varga, "Conveying path setup type in PCEP messages",
          draft-ietf-pce-lsp-setup-type-00 (work in progress),
          October 2014.

[I-D.ietf-pce-pce-initiated-lsp]
          Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "PCEP
          Extensions for PCE-initiated LSP Setup in a Stateful PCE
          Model", draft-ietf-pce-pce-initiated-lsp-01 (work in
          progress), June 2014.

[I-D.ietf-pce-pcep-mib]
          Koushik, K., Stephan, E., Zhao, Q., King, D., and J.
          Hardwick, "PCE communication protocol (PCEP) Management
          Information Base", draft-ietf-pce-pcep-mib-04 (work in
          progress), February 2013.

[I-D.ietf-pce-stateful-pce]
          Crabbe, E., Medved, J., Minei, I., and R. Varga, "PCEP
          Extensions for Stateful PCE",
          draft-ietf-pce-stateful-pce-05 (work in progress),
          July 2013.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
          Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element
          (PCE) Communication Protocol (PCEP)", RFC 5440,
          March 2009.

[RFC5462] Andersson, L. and R. Asati, "Multiprotocol Label Switching
          (MPLS) Label Stack Entry: "EXP" Field Renamed to "Traffic
          Class" Field", RFC 5462, February 2009.

12.2. Informative References

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V.,
          and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP
          Tunnels", RFC 3209, December 2001.

[RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching
          (GMPLS) Signaling Resource ReserVation Protocol-Traffic
          Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links
          in Resource ReSerVation Protocol - Traffic Engineering
          (RSVP-TE)", RFC 3477, January 2003.

[RFC4657] Ash, J. and J. Le Roux, "Path Computation Element (PCE)
          Communication Protocol Generic Requirements", RFC 4657,
          September 2006.

Authors' Addresses

    Siva Sivabalan
    Cisco Systems, Inc.
    2000 Innovation Drive
    Kanata, Ontario K2K 3E8
    Canada

    Email: msiva@cisco.com

    Jan Medved
    Cisco Systems, Inc.
    170 West Tasman Dr.
    San Jose, CA 95134
    US

    Email: jmedved@cisco.com

    Clarence Filsfils
    Cisco Systems, Inc.
    Pegasus Parc
    De kleetlaan 6a, DIEGEM BRABANT 1831
    BELGIUM

    Email: cfilsfil@cisco.com

    Edward Crabbe

    Robert Raszuk
    Mirantis Inc.
    100-615 National Ave.
    Mountain View, CA 94043
    US

    Email: robert@raszuk.net

Victor Lopez
Telefonica I+D
Don Ramon de la Cruz 82-84
Madrid 28045
Spain

Email: vlopez@tid.es

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, CA 95134
USA

Email: jeff.tantsura@ericsson.com

Wim Henderickx
Alcatel Lucent
Copernicuslaan 50
Antwerp 2018, CA 95134
BELGIUM

Email: wim.henderickx@alcatel-lucent.com

Jon Hardwick
Metaswitch Networks
100 Church Street
Enfield, Middlesex
UK

Email: jon.hardwick@metaswitch.com