PCE Working Group   Young Lee
                                                          Dhruv Dhody
Internet Draft   Huawei
Intended Status: Standard
Expires: November 2017   Daniele Ceccarelli
                                                             Ericsson

                                                      June 23, 2017

      **PCEP Extensions for Stitching LSPs in Hierarchical Stateful PCE
                                Model**

                    draft-lee-pce-lsp-stitching-00.txt

Abstract

   This document extends the Path Communication Element Communication
   Protocol (PCEP) to coordinate an end-to-end inter-domain tunnel
   setup over a multi-domain networks in the context of Hierarchical
   Stateful PCE environments. This document uses Stitching Label (SL)
   to stich per-domain LSPs.

The list of Internet-Draft Shadow Directories can be accessed at
http://www.ietf.org/shadow.html

This Internet-Draft will expire on December 23, 2017.

Copyright Notice

Table of Contents

1. Introduction

In Multiprotocol Label Switching (MPLS) and Generalized MPLS
(GMPLS), a Traffic Engineering Database (TED) is used in computing
paths for connection oriented packet services and for circuits. The
TED contains all relevant information that a Path Computation
Element (PCE) needs to perform its computations. It is important
that the TED should be complete and accurate anytime so that the PCE
can perform path computations.

In MPLS and GMPLS networks, Interior Gateway routing Protocols
(IGPs) have been used to create and maintain a copy of the TED at

each node. One of the benefits of the PCE architecture [RFC4655] is the use of computationally more sophisticated path computation algorithms and the realization that these may need enhanced processing power not necessarily available at each node participating in an IGP.

[Stateful-PCE] describes a set of extensions to PCEP to provide stateful control. A stateful PCE has access to not only the information carried by the network's Interior Gateway Protocol (IGP), but also the set of active paths and their reserved resources for its computations. PCC can delegate the rights to modify the LSP parameters to an Active Stateful PCE.

[RFC6805] describes a Hierarchical PCE (H-PCE) architecture which can be used for computing end-to-end paths for inter-domain MPLS Traffic Engineering (TE) and GMPLS Label Switched Paths (LSPs). Within the Hierarchical PCE (H-PCE) architecture [RFC6805], the Parent PCE (P-PCE) is used to compute a multi-domain path based on the domain connectivity information. A Child PCE (C-PCE) may be responsible for a single domain or multiple domains, it is used to compute the intra-domain path based on its domain topology information.

[Stateful H-PCE] presents general considerations for stateful PCE(s) in hierarchical PCE architecture. In particular, the behavior changes and additions to the existing stateful PCE mechanisms (including PCE-initiated LSP setup and active PCE usage) in the context of networks using the H-PCE architecture. Section 3.3.1 of [Stateful H-PCE] describe the per domain stitched LSP mode, where the individual per domain LSP are stitched together.

[PCE-CC] introduces the architecture for PCE as a central controller, and examines the motivations and applicability for PCEP as a southbound interface. Section 2.1.3 describes the approach with hierarchical controllers.

[BRPC-Stitch] describes how inter-domain labels over the inter-domain interfaces are determined in the multi-domain BRPC-based PCE environments. Further, the document introduces the concept of Stitching Label (SL) and Inter-domain Path Setup Type [PST]. This document also uses these concepts in the hierarchical Stateful PCE model.

This document extends the Path Communication Element Communication Protocol (PCEP) to coordinate an end-to-end tunnel for a virtual network over multi-domain networks in the context of Hierarchical Stateful PCE environments.

2. Network Settings and concepts

This section describes network settings for this draft. Figure 1
shows the context of Hierarchical Stateful PCE architecture where
multi-domain LSP stitching is required for an end-to-end tunnel
associated with a VN member.

```
                                +----------+
                                +  Parent  +
                                +   PCE    +
                                +----------+

                           .         .        .
                       .          .         .
                   .           .          .
               .            .          .
   Stateful      +-----+  .          +-----+          +-----+
   C-PCE for    + PCE +            + PCE +          + PCE +
   Domain A      +-----+   Domain B +-----+   Domain C +-----+
                 .                    .                    .
                 .                    .                    .
                 .                    .                    .
                 .                    .                    .
                 .                    .                    .
               ___                  ___                  ___
              (   )                (   )                (   )
             (  o  )              (  o  )              ( o--o)
            (  / \  )            (   |\  )            (  |  | )
    EP1-------(o-o---o-o)==========(o-o-o-o-o)==========(o--o--o-o)--------
EP2
            |( \ / )|            |( \ / )|            |( | / )|
            |( o )|            |( o )|            |( o-o )|
            | (___) |            | (___) |            | (___) |
            |       |            |       |            |       |
            |Domain |Inter-domain|Domain |Inter-domain|Domain |
            |  A    |    A-B     |  B    |    B-C     |  C    |
            |<----->|<---------->|<----->|<---------->|<------>

                         E2E Tunnel
            <-------------------------------------------------->
```

              Figure 1: Multi-domain LSP stitching for an end-to-end tunnel

   The draft provides PCE mechanisms to identify and isolate an end-to-
   end tunnel for a virtual network by concatenating a set of
   LSP/tunnel segments comprising an end-to-end tunnel. From Figure 1,
   there are a set of segments comprising an end-to-end tunnel: Per

   Domain LSP A, Inter-domain Link A-B, Per Domain LSP B, Inter-domain
   Link B-C, and Per Domain C.

   It is important to realize that this end-to-end tunnel for a virtual
   network should be identifiable from other tunnels in the networks so
   as to guarantee its performance objective associated with this
   particular tunnel. See Section 2.2 for ACTN applicability for
   detailed discussion on this aspect.

   As per [BRPC-Stitch], Stitching Label (SL) is defined as a dedicated
   label that is used to stitch two tunnels (RSVP-TE tunnels or Segment
   Routing paths). This label is exchanged between exit BN(i) and entry
   BN(i+1) via PCEP. In case of H-PCE, the SL is conveyed from entry
   BN(i+1) to the child PCE(i+1) to the parent PCE, and then to child
   PCE(i) to the entry BN(i). The exit BN(i) learns the SL via the per-
   domain LSp setup technique (RSVP-TE, SR, PCECC etc).

   [BRPC-Stitch] define new LSP setup types for BRPC mode, this
   document also uses the same LSP setup type for the Stateful H-PCE
   mode.

      o TBD1: Inter-Domain Traffic engineering end-to-end path is
   setup using H-PCE method. This new LSP-TYPE value MUST be set in a
   PCInitiate messages sends by a P-PCE (Parent PCE) to its C-PCE
   (child PCE) of transit and destination domains to initiate a new
   inter-domain LSP tunnel. In turn, the C-PCE MUST return a Stitching
   Label SL in the RRO of the PCRpt message to P-PCE.

      o TBD2: Inter-Domain Traffic engineering local path is setup
   using RSVP-TE. This new LSP-TYPE value MUST be set in the
   PCInitiate message sends by a C-PCE(i) requesting to a PCC of
   domain(i) to initiate a new local LSP tunnel(i) which is part of an
   inter-domain LSP tunnel. This LSP-TYPE value MUST be used by the C-
   PCE(i) only after receiving a PCInitiate message with an LSP-TYPE
   equal to TBD1 from a P-PCE. In turn, the PCC of domain(i) MUST
   return a Stitching Label SL in the RRO of the PCRpt message.

      o TBD3: Inter-Domain Traffic engineering local path is setup
   using Segment Routing (SR). This new LSP-TYPE value MUST be set in
   the PCInitiate message sends by a C-PCE(i) requesting to a PCC of
   domain(i) to initiate a new Segment Routing path which is part of
   an inter-domain Segment Routing path. This LSP-TYPE value MUST be
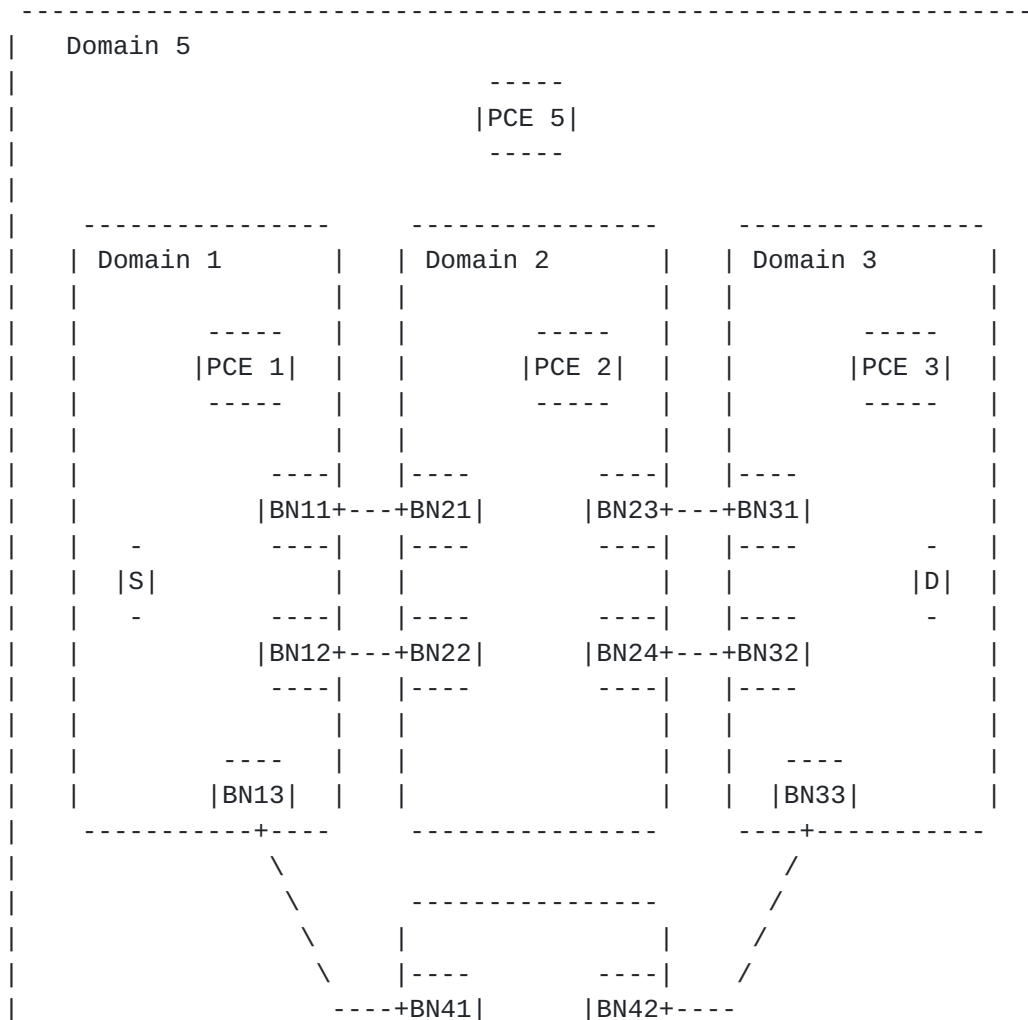   used by the C-PCE(i) only after receiving a PCInitiate message with

an LSP-TYPE equal to TBD1 from a P-PCE. In turn, the PCC MUST
return a Stitching Label SL in the RRO of the PCRpt message.

[Editor's Note - This draft authors plan to discuss with authors of
[BRPC-Stitch] to simplify this, as any new path setup type like
PCECC would require another path-setup type to be defined here.]
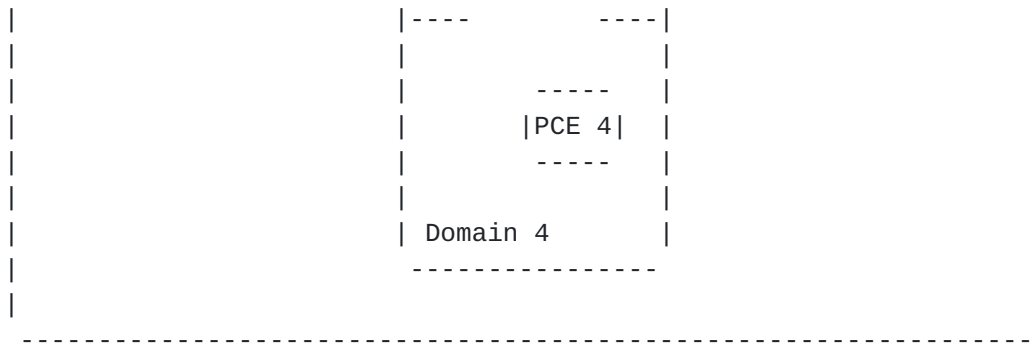
Thus, these LSP-TYPE value MUST be set in PCEP messages sends by a
Parent PCE to child PCE as well as between child PCE and the PCCs
when SL is used.

2.1. Stateful H-PCE Stitching Procedure

Taking the sample hierarchical domain topology example from
[RFC6805] as the reference topology for the entirety of this
document.

```
   -----------------------------------------------------------------------
  |    Domain 5                                                           |
  |                                -----                                  |
  |                               |PCE 5|                                 |
  |                                -----                                  |
  |                                                                       |
  |     ---------------      ---------------      ---------------         |
  |    | Domain 1       |    | Domain 2       |    | Domain 3       |   |  |
  |    |                |    |                |    |                |   |  |
  |    |      -----      |    |      -----      |    |      -----      |   |  |
  |    |     |PCE 1|     |    |     |PCE 2|     |    |     |PCE 3|     |   |  |
  |    |      -----      |    |      -----      |    |      -----      |   |  |
  |    |                |    |                |    |                |   |  |
  |    |        ----|   |----        ----|   |----                |   |  |
  |    |        |BN11+---+BN21|       |BN23+---+BN31|               |   |  |
  |    |   -    ----|   |----        ----|   |----       -        |   |  |
  |    |  |S|        |   |                |   |       |D|          |   |  |
  |    |   -    ----|   |----        ----|   |----       -        |   |  |
  |    |        |BN12+---+BN22|       |BN24+---+BN32|               |   |  |
  |    |        ----|   |----        ----|   |----                |   |  |
  |    |                |   |                |   |                |   |  |
  |    |     ----       |   |                |   |    ----         |   |  |
  |    |    |BN13|      |   |                |   |   |BN33|        |   |  |
  |     -----------+----      ---------------      ----+-----------    |
  |              \                                 /                      |
  |               \      ---------------         /                       |
  |                \    |                |      /                        |
  |                 \   |----        ----|    /                         |
  |                  ----+BN41|       |BN42+----                          |
```

```
          |                  |----       ----|                          |
          |                  |               |                          |
          |                  |      -----    |                          |
          |                  |     |PCE 4|   |                          |
          |                  |      -----    |                          |
          |                  |               |                          |
          |                  | Domain 4      |                          |
          |                   ---------------                           |
          |                                                             |
           -------------------------------------------------------------
```

Section 3.3.1 of [Stateful H-PCE] describes the per-domain stitched
LSP mode and list all the steps needed. To support SL based
stitching, the steps are modified as follows -

Using the reference architecture described in Figure above:

(1) The P-PCE (PCE5) is requested to initiate a LSP.

Steps 4 to 10 of section 4.6.2 of [RFC6805] are executed to
determine the end to end path, which are broken into per-domain LSPs
say -

   o S-BN41

   o BN41-BN33

   o BN33-D

   For LSP (BN33-D)

   (2) The P-PCE (PCE5) sends the initiate request to the child PCE
   (PCE3) via PCInitiate message for LSP (BN33-D) with ERO=(BN33..D)
   and LSP-TYPE=TBD1.

   (3) The PCE3 further propagates the initiate message to BN33 with
   the ERO and LSP-TYPE=TBD2/TBD3 based on setup type.

   (4) BN33 initiates the setup of the LSP as per the path and reports
   to the PCE3 the LSP status ("GOING-UP").

(5) The PCE3 further reports the status of the LSP to the P-PCE
(PCE5).

(6) The node BN33 notifies the LSP state to PCE3 when the state is
"UP" it also carry the stitching label (SL33) in RRO as
(SL33,BN33..D).

(7) The PCE3 further reports the status of the LSP to the P-PCE
(PCE5) as well as carry the stitching label (SL33) in RRO as
(SL33,BN33..D).

For LSP (BN41-BN33)

(8) The P-PCE (PCE5) sends the initiate request to the child PCE
(PCE4) via PCInitiate message for LSP (BN41-BN33) with
ERO=(BN41..BN42,SL33,BN33) and LSP-TYPE=TBD1.

(9) The PCE4 further propagates the initiate message to BN41 with
the ERO and LSP-TYPE=TBD2/TBD3 based on setup type. In case of
RSVP_TE, the node BN41 encode the stitching label SL33 as part of
the ERO to make sure the node BN42 uses the label SL33 towards node
BN33. In case of SR, the label SL33 is part of the label stack
pushed at node BN41.

(10) BN41 initiates the setup of the LSP as per the path and reports
to the PCE4 the LSP status ("GOING-UP").

(11) The PCE4 further reports the status of the LSP to the P-PCE
(PCE5).

(l2) The node BN41 notifies the LSP state to PCE4 when the state is
"UP" it also carry the stitching label (SL41) in RRO as
(SL41,BN41..BN33).

(13) The PCE4 further reports the status of the LSP to the P-PCE
(PCE5) as well as carry the stitching label (SL41) in RRO as
(SL41,BN41..BN33).

For LSP (S-BN41)

(14) The P-PCE (PCE5) sends the initiate request to the child PCE
(PCE1) via PCInitiate message for LSP (S-BN41) with
ERO=(S..BN13,SL41,BN41).

(15) The PCE1 further propagates the initiate message to node S
with the ERO. In case of RSVP_TE, the node S encode the stitching
label SL41 as part of the ERO to make sure the node BN13 uses the
label SL41 towards node BN41. In case of SR, the label SL41 is part
of the label stack pushed at node S.

(16) S initiates the setup of the LSP as per the path and reports
to the PCE1 the LSP status ("GOING-UP").

(17) The PCE1 further reports the status of the LSP to the P-PCE
(PCE5).

(18) The node S notifies the LSP state to PCE1 when the state
is"UP".

(19) The PCE1 further reports the status of the LSP to the P-PCE
(PCE5).

In this way, per-domain LSP are stitched together using the
stitching label (SL). The per-domain LSP MUST be setup from the
destination domain towards the source domain one after the other.

Once the per-domain LSP is setup, the entry BN chooses a free label for the Stitching Label SL and add a new entry in its MPLS LFIB with this SL label. The SL from the destination domain is propagated to adjacent transit domain, towards the source domain at each step. This happens through the entry BN to C-PCE to the P-PCE and vice-versa. In case of RSVP-TE, the entry BN further propagates the SL label to the exit BN via RSVP-TE. In case of SR, the SL label is pushed as part of the SR label stack.

2.2. Applicability to ACTN

[ACTN] describes framework for Abstraction and Control of TE Networks (ACTN), where each Physical Network Controller (PNC) is equivalent to C-PCE and P-PCE is the Multi-Domain Service Coordinator (MDSC). The Per domain stitched LSP as per the Hierarchical PCE architecture described in Section 3.3.1 and Section 4.1 of [Stateful H-PCE] is well suited for ACTN.

The stitching label (SL) mechanism as described in this document is well suited for ACTN when per domain LSP needs to be stitched to form an E2E tunnel or a VN Member. It is to be noted that certain VNs require isolation from other clients. The stitching label mechanism described in this document can be applicable to the VN isolation use-case by uniquely identifying the concatenated stitching labels across multi-domain only to a certain VN member or an E2E tunnel.

3. Security Considerations

Procedures and protocol extensions defined in this document do not effect the overall PCEP security model. See [RFC5440], [I-D.ietf-pce-pceps]. It is suggested that any mechanism used for securing the transmission of other PCEP message be applied here as well. As a general precaution, it is RECOMMENDED that these PCEP extensions only be activated on authenticated and encrypted sessions belonging to the same administrative authority.

4. IANA Considerations

This document requests IANA actions to allocate code points for the protocol elements defined in this document.

[5](). References

5.1. Normative References

[RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path
          Computation Element (PCE)-Based Architecture", RFC 4655,
          August 2006.

[RFC4674] Le Roux, J., Ed., "Requirements for Path Computation
          Element (PCE) Discovery", RFC 4674, October 2006.

[RFC5088] Le Roux, JL., Ed., Vasseur, JP., Ed., Ikejiri, Y., and R.
          Zhang, "OSPF Protocol Extensions for Path Computation
          Element (PCE) Discovery", RFC 5088, January 2008.

[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, January 2008.

[RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The
          OSPF Opaque LSA Option", RFC 5250, July 2008.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic
          Engineering", RFC 5305, October 2008.

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

5.2. Informative References

[JMS]   Java Message Service, Version 1.1, April 2002, Sun
         Microsystems.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic
          Engineering (TE) Extensions to OSPF Version 2", RFC 3630,
          September 2003.

[RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions
          in Support of Generalized Multi-Protocol Label Switching
          (GMPLS)", RFC 4203, October 2005.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation
           Element (PCE)-Based Architecture", RFC 4655, August 2006.

[BGP-LS] Gredler, H., Medved, J., Previdi, S., Farrel, A., and
           S.Ray, "North-Bound Distribution of Link-State and TE
           information using BGP", draft-ietf-idr-ls-distribution,
           work in progress.

[S-PCE-GMPLS] X. Zhang, et. al, "Path Computation Element (PCE)
           Protocol Extensions for Stateful PCE Usage in GMPLS-
           controlled Networks", draft-ietf-pce-pcep-stateful-pce-
           gmpls, work in progress.

[RFC7399] A. Farrel and D. king, "Unanswered Questions in the Path
           Computation Element Architecture", RFC 7399, October 2015.

[RFC7449] Y. Lee, G. Bernstein, "Path Computation Element
           Communication Protocol (PCEP) Requirements for Wavelength
           Switched Optical Network (WSON) Routing and Wavelength
           Assignment", RFC 7449, February 2015.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route
           Reflection: An Alternative to Full Mesh Internal BGP
           (IBGP)", RFC 4456, April 2006.

[RFC6163] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS
           and PCE Control of Wavelength Switched Optical Networks",
           RFC 6163,

[G.680] ITU-T Recommendation G.680, Physical transfer functions of
           optical network elements, July 2007.

[ACTN-Frame] D.Ceccarelli, and Y. Lee (Editors), "Framework for
           Abstraction and Control of TE Networks", draft-ietf-teas-
           actn-framework, work in progress.

[RFC6805] A. Farrel and D. King, "The Application of the Path
           Computation Element Architecture to the Determination of a
           Sequence of Domains in MPLS and GMPLS", RFC 6805, November
           2012.

[PCEP-LS-Arch] Y. Lee, D. Dhody and D. Ceccarelli, "Architecture and
           Requirement for Distribution of Link-State and TE
           Information via PCEP", draft-leedhody-teas-pcep-ls, work
           in progress.

[PCEP-LS] D. Dhody, Y. Lee and D. Ceccarelli "PCEP Extension for
          Distribution of Link-State and TE Information.", work in
          progress, September 21, 2015[Stateful-PCE] Crabbe, E.,
          Minei, I., Medved, J., and R. Varga, "PCEP Extensions for
          Stateful PCE", draft-ietf-pce-stateful-pce, work in
          progress.

[PCE-Initiated] 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, work in
          progress.

[Stateful H-PCE] D. Dhody, Y. Lee and D. Ceccarelli, "Hierarchical
          Stateful Path Computation Element (PCE)", draft-ietf-pce-
          stateful-hpce, work-in-progress.

[FlexOSPF] X. Zhang, H. Zheng, R. Casellas, O. Gonzalez de Dios, D.
          Ceccarelli, "GMPLS OSPF Extensions in support of Flexi-
          grid DWDM networks", draft-ietf-ccamp-flexible-grid-ospf-
          ext-05, work in progress.

[PST] Sivabalan, S., Medved, J., Minei, I., Crabbe, E., Varga,
          R., Tantsura, J., and J. Hardwick, "Conveying path setup
          type in PCEP messages", draft-ietf-pce-lsp-setup-type-03
          work in progress.

[Appendix A](). Contributor Addresses

Author's Addresses

    Young Lee
    Huawei Technologies
    5340 Legacy Drive, Building 3
    Plano, TX 75023, USA

    Email: leeyoung@huawei.com

    Dhruv Dhody
    Huawei Technologies
    Divyashree Techno Park, Whitefield
    Bangalore, Karnataka 560066
    India
    Email: dhruv.ietf@gmail.com

    Daniele Ceccarelli
    Ericsson
    Torshamnsgatan,48
    Stockholm
    Sweden

    Email: daniele.ceccarelli@ericsson.com