Network Working Group   Xian Zhang
Internet-Draft   Young Lee
Intended status: Standards Track   Fatai Zhang
                                                        Huawei
                                               Ramon Casellas
                                                          CTTC
                                       Oscar Gonzalez de Dios
                                               Telefonica I+D
                                                     Zafar Ali
                                                 Cisco Systems

Expires: April 21, 2014   October 21, 2013

Path Computation Element (PCE) Protocol Extensions for Stateful PCE
Usage in GMPLS-controlled Networks

draft-zhang-pce-pcep-stateful-pce-gmpls-03.txt

Status of this Memo

Abstract

   The Path Computation Element (PCE) facilitates Traffic Engineering
   (TE) based path calculation in large, multi-domain, multi-region, or
   multi-layer networks. [Stateful-PCE] provides the fundamental PCE
   communication Protocol (PCEP) extensions needed to support stateful
   PCE functions, without specifying the technology-specific extensions.
   This memo provides extensions required for PCEP so as to enable the
   usage of a stateful PCE capability in GMPLS-controlled networks.

Conventions used in this document

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in RFC-2119 [RFC2119].

Table of Contents

1. Introduction

[RFC 4655] presents the architecture of a Path Computation Element
(PCE)-based model for computing Multiprotocol Label Switching (MPLS)
and Generalized MPLS (GMPLS) Traffic Engineering Label Switched
Paths (TE LSPs). To perform such a constrained computation, a PCE
stores the network topology (i.e., TE links and nodes) and resource
information (i.e., TE attributes) in its TE Database (TED). To
request path computation services to a PCE, [RFC 5440] defines the
PCE communication Protocol (PCEP) for interaction between a Path
Computation Client (PCC) and a PCE, or between two PCEs. PCEP as
specified in [RFC 5440] mainly focuses on MPLS networks and the PCEP
extensions needed for GMPLS-controlled networks are provided in
[PCEP-GMPLS].

Stateful PCEs are shown to be helpful in many application scenarios,
in both MPLS and GMPLS networks, as illustrated in [Stateful-APP].
In order for these applications to able to exploit the capability of
stateful PCEs, extensions to the PCE communication protocol (i.e.,
PCEP) are required.

[Stateful-PCE] provides the fundamental extensions needed for
stateful PCE to support general functionality, but leaves out the
specification for technology-specific objects/TLVs. Complementarily,
this document focuses on the extensions that are necessary in order
for the deployment of stateful PCEs in GMPLS-controlled networks.

2. PCEP Extensions

2.1. Overview of Requirements

This section notes the main functional requirements for PCEP
extensions to support stateful PCE for use in GMPLS-controlled
networks, based on the description in [Stateful-APP]. Many
requirements are common across a variety of network types (e.g.,
MPLS-TE networks and GMPLS networks) and the protocol extensions to
meet the requirements are already described in [Stateful-PCE]. This
document does not repeat the description of those protocol
extensions. Other requirements that are also common across a variety
of network types do not currently have protocol extensions defined
in [Stateful-PCE]. In these cases, this document presents protocol
extensions for discussion by the PCE working group and potential
inclusion in [Stateful-PCE]. In addition, this document presents
protocol extensions for a set of requirements which are specific to
the use of a stateful PCE in a GMPLS-controlled network.

The basic requirements are as follows:

o Advertisement of the stateful PCE capability. This generic
   requirement is covered in Section 7.1.1 of [Stateful-PCE].
   Section 2.2 of this document discusses other potential extensions
   for this functionality.

o LSP delegation is already covered in Section 5.5 of [Stateful-PCE].
   Section 2.3 of this document provides extension for its
   application in GMPLS-controlled networks. Moreover, further
   discussion of some generic details that may need additional
   consideration is provided.

o LSP state synchronization. This is a generic requirement already
   covered in Section 5.4 of [Stateful-PCE]. However, there are
   further extensions required specifically for GMPLS-controlled
   networks and discussed in Section 2.4. Reference to LSPs by
   identifiers is discussed in Section 7.2 of [Stateful-PCE]. This
   feature can be applied to reduce the data carried in PCEP messages.
   Use cases and additional Error Codes are necessary, as described
   in Section 2.5 and 2.6.

2.2. Stateful PCE Capability Advertisement

Whether a PCE has stateful capability or not can be advertised
during the PCEP session establishment process. It can also be
advertised through routing protocols as described in [RFC5088]. In
either case, the following additional aspects should also be
considered.

2.2.1. PCE Capability Advertisement in Multi-layer Networks

In multi-layer network scenarios, such as an IP-over-optical network,
if there are dedicated PCEs responsible for each layer, then the
PCCs should be informed of which PCEs they should synchronize their
LSP states with, as well as send path computation requests to. The
Layer-Cap TLV defined in [INTER-LAYER] can be used to indicate which
layer a PCE is in charge of. (Editor's note: this change is
currently not included in the current version of the [INTER-LAYER]
draft. It is expected that it will be included in its next version.)
This TLV is optional and MAY be carried in the OPEN object. It is
RECOMMMENDED that a PCC synchronizes its LSP states with the same
PCEs that it can use for path computation in a multi-layer network.
In a single layer, this TLV MAY not be used. However, if the PCE
capability discovery depends on IGP and if an IGP instance spans
across multiple layers, this TLV is still needed.

Alternatively, the extension to current OSPF PCED TLV is needed. A
new domain-type denoting the layer information can be defined:

domain-type: T.B.D.

When it is carried in PCE-DOMAIN sub-TLV, it denotes the layer for
which a PCE is responsible for path computation as well as LSP state
synchronization. When carried in the PCE-NEIG-DOMAIN sub-TLV, it
denotes its adjacent layers for which a PCE can compute paths and
synchronize the LSP states. The DOMAIN-ID information can be
represented using the following format, to denote the layer
information:

```
 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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | LSP Enc. Type | Switching Type|            Reserved           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

2.3. LSP Delegation in GMPLS-controlled Networks

To enable the PCE to control an LSP, the PCUpd message is defined in
[Stateful-PCE]. However, However, the specification of technology
specific extensions is not covered. The following defines the
<path> descriptor, present in the PCUpd message, that should be
used in GMPLS-controlled networks:

<path>::=<ERO><attribute-list>

   Where:

      <attribute-list> ::= [<LSPA>]

                          [<BANDWIDTH>]

                          [<GENERALIZED-BANDWIDTH>...]

                          [<metric-list>]

      <metric-list>::= <METRIC>[<metric-list>]

As explained in [stateful-APP], LSP parameter update controlled by a
stateful PCE in a multi-domain network is complex and requires well-
defined operational procedures as well as protocol design.

[TBD: protocol extensions]

2.4. LSP Synchronization in GMPLS-controlled networks

   For LSP state synchronization of stateful PCEs in GMPLS networks,
   the LSP attributes, such as its bandwidth, associated route as well
   as protection information etc, should be updated by PCCs to PCE LSP
   database (LSP-DB). Note the LSP state synchronization described in
   this document denotes both the bulk LSP report at the initialization
   phase as well as the LSP state report afterwards described in
   [Stateful-PCE].

   As per [Stateful-PCE], it does not cover technology-specific
   specification for state synchronization. Therefore, extensions of
   PCEP for stateful PCE usage in GMPLS networks are required. For LSP
   state synchronization, the objects/TLVs that should be used for
   stateful PCE in GMPLS networks are defined in [PCEP-GMPLS] and are
   briefly summarized as below:

   o GENERALIZED BANDWIDTH

   o GENERALIZED ENDPOINTS

   o PROTECTION ATTRIBUTE

   o Use of IF_ID_ERROR_SPEC. [Stateful-PCE] section 7.2.2 only
   considers RSVP ERROR_SPEC TLVs. GMPLS extends this to also support
   IF_ID_ERROR_SPEC, for example, to report about failed unnumbered
   interfaces.

   o Extended objects to support the inclusion of the label and
   unnumbered links.

   Per [Stateful-PCE], the PCRpt message is defined for LSP state
   synchronization purposes. PCRpt is used by a PCC to report one or
   more of its LSPs to a stateful PCE. However, the <path> descriptor
   is technology-specific and left undefined.

   For LSP state synchronization in GMPLS-controlled networks, the
   encoding of the <path> descriptor is defined as follows:

   <path>::=<ERO><attribute-list>

      Where:

          <attribute-list> ::= [<LSPA>]

                               [<BANDWIDTH>]

                               [<GENERALIZED-BANDWIDTH>...]

                              [<IRO>]

                              [<XRO>]

                              [<metric-list>]

          <metric-list>::= <METRIC>[<metric-list>]

   The objects included in the <path> descriptor can be found in
   [RFC5440], [PCE-GMPLS] and [RFC5521].

   For all the objects presented in this section, the P and I bit MUST
   be set to 0 since they are only used by a PCC to report its LSP
   information.

   In GMPLS-controlled networks, the <ERO> object may include a list of
   the label sub-object for SDH/SONET, OTN and DWDM networks. It may
   also include a list of unnumbered interface IDs to denote the
   allocated resource. The <RRO>, <IRO> and <XRO> objects MAY include
   unnumbered interface IDs and labels for networks such as OTN and WDM
   networks.

   If the LSP being reported is a protecting LSP, the <PROTECTION-
   ATTRIBUTE> TLV MUST be included in the <LSPA> object to denote its
   attributes and restrictions. Moreover, if the status of the
   protecting LSP changes from non-operational to operational, this
   should be synchronized to the stateful PCE. For example, in 1:1
   protection, the combination of S=0, P=1 and O=0 denotes the
   protecting path is set up already but not used for carrying traffic.
   Upon the working path failure, the operational status of the
   aforementioned protecting LSP changes to in-use (i.e., O=1). This
   information should be synchronized with a stateful PCE through a
   PCRpt message.

   The O bit in the <GENERALIZED-BANDWIDTH> object has no meaning for
   LSP state synchronization and MUST be set to 0. Furthermore, this
   object MAY appear twice, one with R set to 1 and the other with R
   set to 0. This is to denote the asymmetric bandwidth property of the
   updated bi-directional LSP.

2.5. Modification of Existing PCEP Messages and Procedures

   One of the advantages mentioned in [Stateful-APP] is that the
   stateful nature of a PCE simplifies the information conveyed in PCEP
   messages, notably between PCC and PCE, since it is possible to refer
   to PCE managed state for active LSPs. To be more specific, with a

   stateful PCE, it is possible to refer to a LSP with a unique
   identifier in the scope of the PCC-PCEP session and thus use such
   identifier to refer to that LSP.

 2.5.1. Use cases

   Use Case 1: Assuming a stateful PCE's LSP-DB is up-to-date, a PCC
   (e.g. NMS) requesting for a re-optimization of one or several LSPs
   can send the request with "R" bit set and only provides the relevant
   LSP unique identifiers.

   Upon receiving the PCReq message, PCE should be able to correlate
   with one or multiple LSPs with their detailed state information and
   carry out optimization accordingly.

   The handling of RP object specified in [RFC5440] is stated as
   following:

   "The absence of an RRO in the PCReq message for a non-zero-bandwidth
   TE LSP (when the R bit of the RP object is set) MUST trigger the
   sending of a PCErr message with Error-Type="Required Object Missing"
   and Error-value="RRO Object missing for re-optimization."

   If a PCE has stateful capabilities, and such capabilities have been
   negotiated and advertised, specific rules given in [RFC5440] may
   need to be relaxed. In particular, the re-optimization case: if the
   re-optimization request refers to a given LSP state, and the RRO
   information is available, the PCE can proceed.

   Use Case 2: in order to set up a LSP which has a constraint that its
   route should not use resources used by one or more existing LSPs, a
   PCC can send a PCReq with the identifiers of these LSPs. A stateful
   PCE should be able to find the corresponding route and resource
   information so as to meet the constraints set by the requesting PCC.
   Hence, the LSP identifier TLV defined in [Stateful-PCE] can be used
   in XRO object for this purpose. Note that if the PCC is a node in
   the network, the constraint LSP ID information will be confined to
   the LSPs initiated by itself.

 2.5.2. Modification for LSP Re-optimization

   For re-optimization, upon receiving a path computation request and
   the "R" bit is set, the stateful PCE SHOULD still perform the re-
   optimization in the following two cases:

   Case 1: the existing bandwidth and route information of the to-be-
   optimized LSP is provided in the path computation request. This

information should be provided via <BANDWIDTH>, <GENERARLIZED-
BANDWIDTH>, <ERO> objects.

Case 2: the existing bandwidth and route information can be found
locally in its LSP-DB. In this case, the PCRep and PCReq messages
need to be modified to carry LSP identifiers. The stateful PCE can
find this information using the per-node LSP ID together with the
PCC's address.

If no LSP state information is available to carry out re-
optimization, the stateful PCE should report the error "LSP state
information unavailable for the LSP re-optimization" (Error Type =
T.B.D., Error value= T.B.D.).

2.5.3. Modification for Route Exclusion

A LSP identifier sub-object is defined and its format as follows:

```
  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 (T.B.D.)    |    Length     |        Reserved        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |          PLSP-ID                      |         Flag          |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 //                        Optional TLVs                        //
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   L bit:
      The L bit SHOULD NOT be set, so that the subobject represents
 a strict hop in the explicit route.

    Type:
      Subobject Type for a per-node LSP identifier.

    Length:
      The Length contains the total length of the subobject in bytes,
 including the Type and Length fields.

    PLSP-ID:
      This is the identifier given to a LSP and it is unique on a
 node basis. It is defined in [Stateful-PCE].

    Flags:
      This field is defined in [Stateful-PCE]. It is not used in
 this sub-object and should be ignored upon receipt.

   Optional TLVs:
        Additional TLVs can be defined in the future to provide
   further information to identify a LSP. In this document, no TLVs are
   defined.

   One or multiple of these sub-objects can be present in the XRO
   object. When a stateful PCE receives a path computation request
   carrying this sub-object, it should find relevant information of
   these LSPs and preclude the resource during the path computation
   process. If a stateful PCE cannot recognize one or more of the
   received LSP identifiers, it should reply PCErr saying "the LSP
   state information for route exclusion purpose cannot be found"
   (Error-type = T.B.D., Error-value= T.B.D.). Optionally, it may
   provide with the unrecognized identifier information to the
   requesting PCC.

2.6. Additional Error Type and Error Values Defined

   Error Type Meaning

   21(TBD)   LSP state information missing

             Error-value 1: LSP state information unavailable for the
             LSP re-optimization

             Error-value 2: the LSP state information for route
             exclusion purpose cannot be found

3. IANA Considerations

   IANA is requested to allocate new Types for the TLV/Object defined
   in this document.T.B.D.

4. Manageability Considerations

   The description and functionality specifications presented related
   to stateful PCEs should also comply with the manageability
   specifications covered in Section 8 of [RFC4655]. Furthermore, a
   further list of manageability issues presented in [Stateful-PCE]
   should also be considered.

   Additional considerations are presented in the next sections.

4.1. Requirements on Other Protocols and Functional Components

   When the detailed route information is included for LSP state
   synchronization (either at the initial stage or during LSP state

report process), this require the ingress node of an LSP carry the
RRO object in order to enable the collection of such information.

[5](5). Security Considerations

The security issues presented in [[RFC5440](RFC5440)] and [Stateful-PCE] apply
to this document.

[6](6). Acknowledgement

We would like to thank Adrian Farrel and Cyril Margaria for the
useful comments and discussions.

[7](7). References

[7.1](7.1). Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to indicate
          requirements levels", [RFC 2119](RFC 2119), March 1997.

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

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

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

[INTER-LAYER] Oki, E., Takeda, Tomonori, Le Roux, JL., Farrel, A.,
          Zhang, F., "Extensions to the Path Computation Element
          communication Protocol (PCEP) for Inter-Layer MPLS and
          GMPLS Traffic Engineering", draft-ietf-pce-inter-layer-ext,
          work in progress.

[7.2](7.2). Informative References

[Stateful-APP] Zhang, X., Minei, I., et al "Applicability of
          Stateful Path Computation Element (PCE) ", draft-ietf-pce-stateful-pce-app, , work in progress.

[Stateful-PCE]Crabbe, E., Medved, J., Varga, R., Minei, I., "PCEP
          Extensions for Stateful PCE", draft-ietf-pce-stateful-pce,
          work in progress.

   [PCE-IA-WSON] Lee, Y., Bernstein G., Takeda, T., Tsuritani, T.,
            "PCEP Extensions for WSON Impairments", draft-lee-pce-
            wson-impairments, work in progress.

   [PCEP-GMPLS] Margaria, C., Gonzalez de Dios, O., Zhang, F., "PCEP
            extensions for GMPLS", draft-ietf-pce-gmpls-pcep-
            extensions, work in progress.

8. Contributors' Address

   Dhruv Dhody
   Huawei Technology
   Leela Palace
   Bangalore, Karnataka 560008
   INDIA

   EMail: dhruvd@huawei.com

   Yi Lin
   Huawei Technologies
   F3-5-B R&D Center, Huawei Base
   Bantian, Longgang District
   Shenzhen 518129 P.R.China

   Phone: +86-755-28972914
   Email: yi.lin@huawei.com

Authors' Addresses

    Xian Zhang
    Huawei Technologies
    F3-5-B R&D Center, Huawei Base
    Bantian, Longgang District
    Shenzhen 518129 P.R.China

    Phone: +86-755-28972645
    Email: zhang.xian@huawei.com

    Young Lee
    Huawei
    1700 Alma Drive, Suite 100
    Plano, TX 75075
    US

    Phone: +1 972 509 5599 x2240
    Fax: +1 469 229 5397
    EMail: ylee@huawei.com

    Fatai Zhang
    Huawei
    F3-5-B R&D Center, Huawei Base
    Bantian, Longgang District
    P.R. China

    Phone: +86-755-28972912
    Email: zhangfatai@huawei.com

    Ramon Casellas
    CTTC
    Av. Carl Friedrich Gauss n7
    Castelldefels, Barcelona 08860
    Spain

    Phone:
    Email: ramon.casellas@cttc.es

    Oscar Gonzalez de Dios
    Telefonica Investigacion y Desarrollo
    Emilio Vargas 6
    Madrid,   28045
    Spain

      Phone: +34 913374013
      Email: ogondio@tid.es

      Zafar Ali
      Cisco Systems
      Email: zali@cisco.com

Intellectual Property

   The IETF Trust takes no position regarding the validity or scope of
   any Intellectual Property Rights or other rights that might be
   claimed to pertain to the implementation or use of the technology
   described in any IETF Document or the extent to which any license
   under such rights might or might not be available; nor does it
   represent that it has made any independent effort to identify any
   such rights.

   Copies of Intellectual Property disclosures made to the IETF
   Secretariat and any assurances of licenses to be made available, or
   the result of an attempt made to obtain a general license or
   permission for the use of such proprietary rights by implementers or
   users of this specification can be obtained from the IETF on-line
   IPR   repository at http://www.ietf.org/ipr

   The IETF invites any interested party to bring to its attention any
   copyrights, patents or patent applications, or other proprietary
   rights that may cover technology that may be required to implement
   any standard or specification contained in an IETF Document. Please
   address the information to the IETF at ietf-ipr@ietf.org.

   The definitive version of an IETF Document is that published by, or
   under the auspices of, the IETF. Versions of IETF Documents that are
   published by third parties, including those that are translated into
   other languages, should not be considered to be definitive versions
   of IETF Documents. The definitive version of these Legal Provisions
   is that published by, or under the auspices of, the IETF. Versions
   of   these Legal Provisions that are published by third parties,
   including   those that are translated into other languages, should
   not be   considered to be definitive versions of these Legal
   Provisions.

   For the avoidance of doubt, each Contributor to the IETF Standards
   Process licenses each Contribution that he or she makes as part of
   the IETF Standards Process to the IETF Trust pursuant to the
   provisions of [RFC 5378](RFC 5378). No language to the contrary, or terms,

conditions or rights that differ from or are inconsistent with the
rights and licenses granted under RFC 5378, shall have any effect
and   shall be null and void, whether published or posted by such
Contributor, or included with or in such Contribution.