

PCE Working Group
Internet-Draft
Intended status: Standard
Expires: December 31, 2011

D. Dhody
U. Palle
Huawei Technologies India Pvt Ltd
June 29, 2011

Encoding of Data Structure (DS) in the Path Computation Element
Communication Protocol (PCEP)
draft-dhody-pce-pcep-ds-00

Abstract

The ability to compute shortest constrained Traffic Engineering Label Switched Paths (TE LSPs) in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks across multiple domains has been identified as a key requirement for P2P and P2MP scenarios. Backward-Recursive PCE-Based Computation (BRPC) [RFC 5441] defines VSPT [Virtual Shortest Path Tree] as a default de-facto data structure for PCRep message in inter-domain scenarios.

This document defines extensions to the PCE communication Protocol (PCEP) to allow multiple data structures. Extensions are defined for PCE to indicate the set of Data Structure (DS) it supports; also PCC/PCE can indicate in a path computation request the required DS, and a PCE can report in a path computation reply the Data Structure that was used in the path reply message.

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

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

This Internet-Draft will expire on December 31, 2011.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This Internet-Draft will expire on December 31, 2011.

Table of Contents

1. Introduction	3
1.1. Requirements Language	3
2. Terminology	4
3. Need for multiple Data Structure	4
4. Discovery of PCE Data Structure	5
4.1. DS-List TLV	5
4.2. Elements of Procedure	6
5. Data Structure in PCEP Path Computation Request and Reply Messages	6
5.1. DS Object	7
5.1.1. Elements of Procedure	7
5.2. Carrying The DS Object In a PCEP Message	9
5.3. New RP Object Flag	11
5.3.1. Elements of Procedure	11
6. IANA Considerations	12
6.1. PCE Data Structure Sub-Registry	12
6.2. PCEP Code Points	12
6.2.1. DS Object	12
6.2.2. DS-List TLV	12
6.2.3. PCEP Error Values	13
6.2.4. RP Object Flag	13
7. Security Considerations	13
8. Manageability Considerations	14
8.1. Control of Function and Policy	14
8.2. Information and Data Models	14
8.3. Liveness Detection and Monitoring	14
8.4. Verify Correct Operations	14
8.5. Requirements On Other Protocols	15
8.6. Impact On Network Operations	15
9. Acknowledgments	15
10. References	15
10.1. Normative References	15
10.2. Informative References	15

1. Introduction

The Path Computation Element (PCE) architecture is defined in [RFC4655]. [RFC 5441] describe a PCE-based path computation procedure to compute optimal inter-domain constrained (G)MPLS TE LSPs. It also defines VSPT [Virtual Shortest Path Tree] which is the only data structure and is used in all inter-domain scenarios.

This document describes the need for multiple data structure (DS). It may be useful for a PCC/PCE to discover the set of Data Structure (DS) supported by a PCE. Furthermore, PCC/PCE requires the ability to indicate in a path computation request a required/desired Data Structure, as well as optional function parameters.

For these purposes, this document extends the PCE communication Protocol (PCEP). It defines PCEP extensions that allow a PCE to advertise a list of supported Data Structure (DS), as well as extensions to carry Data Structure (DS) in PCEP request and reply messages. It complements the PCEP base specification [RFC5440].

Note that OSPF- and IS-IS-based PCE discovery mechanisms are defined in [RFC5088] and [RFC5089]. These mechanisms are dedicated to the discovery of a few generic parameters, while more detailed PCE parameters should be discovered using the PCE communication Protocol.

Data Structure (DS) are in this second category; thus, the Data Structure discovery procedure is handled by PCEP.

A new PCEP TLV, named the DS-List TLV, is defined in Section 4. The DS-List TLV is carried in the PCEP OPEN object and allows a PCE to list, during PCEP session-setup phase, the Data Structure (DS) that it supports.

A new PCEP object, the DS object, is defined in Section 5. The DS object is carried within a PCReq (Path Computation Request) message to indicate the required/desired data structure to be applied by a PCE, or in a PCRep (Path Computation Reply) message to indicate the data structure that was used for path computation and the reply message.

1.1. 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.

2. Terminology

The following terminology is used in this document.

BRPC: Backward Recursive Path Computation.

DS: Data Structure.

H-PCE: Hierarchical PCE.

IGP: Interior Gateway Protocol. Either of the two routing protocols, Open Shortest Path First (OSPF) or Intermediate System to Intermediate System (IS-IS).

IS-IS: Intermediate System to Intermediate System.

OF: Objective Function.

OSPF: Open Shortest Path First.

PCC: Path Computation Client: any client application requesting a path computation to be performed by a Path Computation Element.

PCE: Path Computation Element. An entity (component, application, or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints.

P2MP: Point-to-Multipoint

P2P: Point-to-Point

TE LSP: Traffic Engineering Label Switched Path.

TLV: Type-Length-Variable data encoding.

VSPT: Virtual Shortest Path Tree as defined in [RFC 5441].

3. Need for multiple Data Structure

- o [PCE-P2MP-PROCEDURES] describes the need for an extended VSPT for computation of the best core-tree.
- o [RFC 6007] describes the need for disjoint VSPT in case of Synchronized Dependent Path Computations.
- o VSPT does not fit into hierarchical PCE paradigm described in [PCE-HIERARCHY-FWK].

- o VSPT does not work well with constraints like HOP-LIMIT.
- o Since PCEP allow multiple Objective Function (OF); it is natural to extend PCEP to support multiple Data Structure based on path computation scenario

4. Discovery of PCE Data Structure

This section defines PCEP extensions (see [RFC5440]) so as to support the advertisement of the Data Structure (DS) supported by a PCE.

A new PCEP DS-List (Data Structure list) TLV is defined. The PCEP DS-List TLV is carried within an OPEN object. This way, during PCEP session-setup phase, a PCE can advertise to a PCEP peer the list of data structure it supports.

4.1. DS-List TLV

The PCEP DS-List TLV is optional. It MAY be carried within an OPEN object sent by a PCE in an Open message to a PCEP peer so as to indicate the list of supported data structures.

The DS-List TLV format is compliant with the PCEP TLV format defined in [RFC5440]. That is, the TLV is composed of 2 octets for the type, 2 octets specifying the TLV length, and a Value field. The Length field defines the length of the value portion in octets. The TLV is padded to 4-octet alignment, and padding is not included in the Length field (e.g., a 3-octet value would have a length of three, but the total size of the TLV would be eight octets).

The PCEP DS-List TLV has the following format:

TYPE: 4

LENGTH: $N * 2$ (where N is the number of Data Structures)

VALUE: list of 2-byte data structure code points, identifying the data structures supported by the sender of the Open message.

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|               DS Code #1               |               DS Code #2               |
+-----+-----+-----+-----+-----+-----+-----+-----+
//                                         //
+-----+-----+-----+-----+-----+-----+-----+-----+
|               DS Code #N               |               padding               |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

DS Code (2 bytes): Data Structure code point identifier. IANA manages the "PCE Data Structure" code point registry (see Section 6).

4.2. Elements of Procedure

A PCE MAY include a DS-List TLV within an OPEN object in an Open message sent to a PCEP peer in order to advertise a set of one or more supported Data Structures. The DS-List TLV MUST NOT appear more than once in an OPEN object. If it appears more than once, the PCEP session MUST be rejected with error type 1 and error value 1 (PCEP session establishment failure / Reception of an invalid Open message). The absence of the DS-List TLV in an OPEN object MUST be interpreted as an absence of information on the list of supported data structures by the PCE, default data structure VSPT is always supported.

As specified in [RFC5440], a PCEP peer that does not recognize the DS-List TLV will silently ignore it.

5. Data Structure in PCEP Path Computation Request and Reply Messages

This section defines PCEP extensions [RFC5440] so as to support the communication of Data Structure (DS) in PCEP path computation request and reply messages. A new PCEP DS (Data Structure) object is defined, to be carried within a PCReq message in order for the PCC/PCE to indicate the required/desired data structure.

The PCEP DS object may also be carried within a PCRep message in order for the PCE to indicate the data structure that was used by the PCE and used in the reply message.

A new flag is defined in the RP (Request Parameters) object. The flag is used in a PCReq message to indicate that the PCE MUST include a DS object in the PCRep message to indicate the data structure that was used during path computation and encoded in the reply message.

Also, new PCEP error types and values are defined.

5.1. DS Object

The PCEP DS (Data Structure) object is optional. It MAY be carried within a PCReq message so as to indicate the desired/required data structure to be applied by the PCE during path computation or within a PCRep message so as to indicate the data structure that was used by the PCE during path computation and in the reply message.

The DS object format is compliant with the PCEP object format defined in [RFC5440].

The DS Object-Class is <TBA by IANA>.

The DS Object-Type is 1.

The format of the DS object body is:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
| DS Code |         Reserved         |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     |
//                               Optional TLV(s)                               //
|                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

DS Code (2 bytes): The identifier of the Data Structure. IANA manages the "PCE Data Structure" code point registry

Reserved (2 bytes): This field MUST be set to zero on transmission and MUST be ignored on receipt.

Optional TLVs may be defined in the future.

5.1.1. Elements of Procedure

To request the use of a specific data structure by the PCE, a PCC/PCE includes a DS object in the PCReq message.

[RFC5440] specifies a bit flag, referred to as the P bit, carried in

the common PCEP object header. The P bit is set by a PCC/PCE to mandate that a PCE must take the information carried in the object into account during the path computation.

If the P bit is set in the DS object, the data structure is mandatory (required data structure) and the PCE MUST use the data structure during path computation. If the P bit is clear in the DS object, the data structure is optional (desired data structure) and the PCE SHOULD apply the data structure if it is supported but MAY choose to apply a different data structure, according to local capabilities and policies.

On receipt of a PCReq message with a DS object, a PCE MUST proceed as follows:

- o If the DS object is unknown/unsupported, the PCE MUST follow procedures defined in [RFC5440]. That is, if the P bit is set, the PCE sends a PCErr message with error type 3 or 4 (Unknown / Not supported object) and error value 1 or 2 (unknown / unsupported object class / object type), and the related path computation request MUST be discarded. If the P bit is cleared, the PCE is free to ignore the object.
- o If the data structure is unknown/unsupported and the P bit is set, the PCE MUST send a PCErr message with error type 3 or 4 (Unknown / Not supported object) and error value 4 (Unrecognized/ Unsupported parameter), and the related path computation request MUST be discarded.
- o If the data structure is unknown/unsupported and the P bit is cleared, the PCE SHOULD apply another (default) data structure.
- o If the data structure is supported but policy does not permit applying it and if the P bit is set, the PCE MUST send a PCErr message with the PCEP error type "policy-violation" (type 5) and a new error value, "data structure not allowed", which is defined in this document.
- o If the data structure is supported but policy does not allow applying it and if the P bit is cleared, the PCE SHOULD apply another (default) data structure.
- o If the data structure is supported and policy allows applying it and if the P bit is set, the PCE MUST apply the requested data structure. Otherwise, if the P bit is cleared, the PCE is free to apply any other data structure.

The default data structure is VSPT or may be locally configured.

5.2. Carrying The DS Object In a PCEP Message

The DS object MAY be carried within a PCReq message. If a data structure is to be applied to a set of synchronized path computation requests, the DS object MUST be carried just after the corresponding SVEC (Synchronization VECTOR) object and MUST NOT be repeated for each elementary request.

A DS object specifying a data structure that applies to an individual path computation request (non-synchronized case) MUST follow the RP object for which it applies.

The format of the PCReq message is updated as follows.

```

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

where:
    <svec-list> ::= <SVEC>
                    [<OF>]
                    [<DS>]
                    [<metric-list>]
                    [<svec-list>]

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

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

and where:

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

```

The DS object MAY be carried within a PCRep message to indicate the data structure used by the PCE during path computation and in the reply message.

When the PCE wants to indicate to the PCC/PCE the data structure that was used for the synchronized computation of a set of paths, the PCRep message MUST include the corresponding SVEC object directly

followed by the DS object, which MUST NOT be repeated for each elementary request.

A DS object specifying a data structure used for an individual path computation (non-synchronized case) MUST follow the RP object for which it applies.

The format of the PCRep message is updated as follows.

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

where:

```
<svec-list> ::= <SVEC>
                [<OF>]
                [<DS>]
                [<metric-list>]
                [<svec-list>]

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

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

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

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

and where:

```
<attribute-list> ::= [<OF>]
                    [<DS>]
                    [<LSPA>]
                    [<BANDWIDTH>]
                    [<metric-list>]
                    [<IRO>]

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

Note: The DS object MAY be associated to a negative reply, i.e., a reply with a NO-PATH object.

5.3. New RP Object Flag

In some cases, where no data structure is specified in the request or an optional data structure is desired (P flag cleared in the DS object common header) but the PCE does not follow the request, the PCC/PCE may desire to know the data structure that was used by the PCE during path computation. To that end, a new flag is defined in the RP object, named the DS flag, allowing a PCC/PCE to request for the inclusion in the path computation reply of the data structure that was used by the PCE during path computation.

The following new bit flag of the RP object is defined: The Supply DS on response flag (bit number <TBA>). When set in a PCReq message, this indicates that the PCE MUST provide the applied data structure in the PCRep message. When set in a PCRep message, this indicates that the data structure that was used during path computation is included.

5.3.1. Elements of Procedure

If the PCC/PCE wants to know the data structure used by the PCE during path computation for a given request, it sets the DS flag in the RP object.

On receipt of a PCReq message with the DS flag in the RP object set, the PCE proceeds as follows:

- o If policy permits, it MUST include in the PCRep message a DS object indicating the data structure it used during path computation.
- o If policy does not permit, it MUST send a PCErr message with the PCEP error code "policy-violation" (type 5) and a new error value, "data structure indication not allowed", which is defined in this document.

Note that a legacy PCE might not recognize the DS flag in the RP object. According to the definition of the Flags field for the RP object (Section 7.4.1 of [RFC5440]), the legacy PCE will ignore the unknown flag, resulting in it sending a PCRep that does not contain a DS object. In this case, the PCC/PCE's behavior is an implementation choice. It might:

- o Discard the PCRep because it really wanted the DS object returned.
- o Accept the PCRep without the knowledge of the DS that was applied.

Note also that these procedures can give rise to the situation where

a PCC/PCE receives a PCRep that contains a DS object with a data structure identifier that the PCC/PCE does not recognize. In this situation, the PCC/PCE behavior is dependent on implementation and configuration. The PCC/PCE could choose any of the following (or some other action):

- o Ignore the DS object and use the computed path.
- o Add the data structure to its view of the PCE's repertoire for inclusion in future computation requests.
- o Discard the PCRep (i.e., the computed path) and send a PCReq to another PCE.
- o Discard the PCRep (i.e., the computed path) and send another PCReq to the same PCE explicitly requiring the use of some other data structure (i.e., by setting the P bit in the DS object).

6. IANA Considerations

6.1. PCE Data Structure Sub-Registry

This document defines a 16-bit PCE data structure identifier to be carried within the PCEP DS object, and also defines the PCEP DS-List TLV. IANA should create and manages the 16-bit "PCE Data Structure" code point registry. Values are TBD.

6.2. PCEP Code Points

6.2.1. DS Object

IANA manages the PCEP Objects code point registry (see [RFC5440]). This is maintained as the "PCEP Objects" sub-registry of the "Path Computation Element Protocol (PCEP) Numbers" registry. This document defines a new PCEP object, the DS object, to be carried in PCReq and PCRep messages.

IANA should make the following allocation:

Object Class	Name	Object Type	Name	Reference
TBA	DS	1	Data Structure	This Doc

6.2.2. DS-List TLV

IANA manages the PCEP TLV code point registry (see [RFC5440]). This is maintained as the "PCEP TLV Type Indicators" sub-registry of the

"Path Computation Element Protocol (PCEP) Numbers" registry. This document defines a new PCEP TLV, the DS-List TLV, to be carried in the OPEN object.

IANA should make the following allocation:

Type	TLV name	References

TBA	DS-List	This Doc

6.2.3. PCEP Error Values

IANA maintains a registry of Error-types and Error-values for use in PCEP messages. This is maintained as the "PCEP-ERROR Object Error Types and Values" sub-registry of the "Path Computation Element Protocol (PCEP) Numbers" registry.

Two new Error-values are defined for the Error-type "policy violation" (type 5):

Error-type	Meaning and error values	Reference

5	Policy violation	
	Error-value=TBA: data structure not allowed (request rejected)	This Doc
	Error-value=TBA: DS bit of the RP object set (request rejected)	This Doc

6.2.4. RP Object Flag

A new flag of the RP object (specified in [RFC5440]) is defined in this document. IANA maintains a registry of RP object flags in the "RP Object Flag Field" sub-registry of the "Path Computation Element Protocol (PCEP) Numbers" registry.

IANA should make the following allocation:

Bit	Description	Reference

TBA	Supply DS on response	This Doc

7. Security Considerations

PCEP security mechanisms are described in [RFC5440] and are used to secure entire PCEP messages. Nothing in this document changes the message flows or introduces any new messages, so the security

mechanisms set out in [RFC5440] continue to be applicable.

This document introduces a single new object that may optionally be carried on PCEP messages and will be automatically secured using the mechanisms described in [RFC5440].

If a PCEP message is vulnerable to attack (for example, because the security mechanisms are not used), then the DS object could be used as part of an attack; however, it is likely that other objects will provide far more significant ways of attacking a PCE or PCC in this case.

8. Manageability Considerations

8.1. Control of Function and Policy

It MUST be possible to configure the activation/deactivation of data structure discovery in PCEP. In addition to the parameters already listed in Section 8.1 of [RFC5440], a PCEP implementation SHOULD allow configuring a list of authorized data structure on a PCE. This may apply to any session the PCEP speaker participates in, to a specific session with a given PCEP peer, or to a specific group of sessions with a specific group of PCEP peers. Note that it is not mandatory for an implementation to support all data structure defined. It MUST be possible to configure a default data structure used for path computation when a path request is received that requests to use an optional data structure.

8.2. Information and Data Models

The PCEP MIB Module defined in [PCEP-MIB] could be extended to include data structure.

8.3. Liveness Detection and Monitoring

Mechanisms defined in this document do not imply any new liveness detection and monitoring requirements in addition to those already listed in [RFC5440].

8.4. Verify Correct Operations

Mechanisms defined in this document do not imply any new operation verification requirements in addition to those already listed in [RFC5440].

8.5. Requirements On Other Protocols

Mechanisms defined in this document do not imply any requirements on other protocols in addition to those already listed in [RFC5440].

8.6. Impact On Network Operations

Mechanisms defined in this document do not have any impact on network operations in addition to those already listed in [RFC5440].

9. Acknowledgments

We would like to thank Pradeep Shastry, Suresh babu, Quintin Zhao and Chen Huaimo for their useful comments and suggestions.

10. References

10.1. Normative References

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

10.2. Informative References

- [PCE-HIERARCHY-FWK] King, D. and A. Farrel, "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS", September 2010.
- [PCE-P2MP-PROCEDURES] Zhao, Q., Ali, Z., Saad,, T., and D. King, "PCE-based Computation Procedure To Compute Shortest Constrained P2MP Inter-domain Traffic Engineering Label Switched Paths", January 2011.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", August 2006.
- [RFC4657] Ash, J. and J. Le Roux, "Path Computation Element (PCE) Communication Protocol Generic Requirements", September 2006.
- [RFC4674] Le Roux, J., "Requirements for Path Computation Element (PCE) Discovery", October 2006.
- [RFC5088] Le Roux, J., Vasseur, JP., Ikejiri, Y., and R.

- Zhang, "OSPF Protocol Extensions for Path Computation Element (PCE) Discovery", January 2008.
- [RFC5089] Le Roux, J., Vasseur, JP., Ikejiri, Y., and R. Zhang, "IS-IS Protocol Extensions for Path Computation Element (PCE) Discovery", January 2008.
- [RFC5440] Ayyangar, A ., Farrel, A ., Oki, E., Atlas, A., Dolganow, A., Ikejiri, Y., Kumaki, K., Vasseur, J., and J. Roux, "Path Computation Element (PCE) communication Protocol (PCEP)", March 2009.
- [RFC5441] Vasseur, JP., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", April 2009.
- [RFC5541] Le Roux, J., Vasseur, JP., and Y. Ikejiri, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", June 2009.
- [RFC6007] Nishioka, I. and D. King, "Use of the Synchronization VECtor (SVEC) List for Synchronized Dependent Path Computations)", September 2010.

Authors' Addresses

Dhruv Dhody
Huawei Technologies India Pvt Ltd
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: dhruvd@huawei.com

Udayasree Palle
Huawei Technologies India Pvt Ltd
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: udayasreepalle@huawei.com

PCE Working Group
Internet-Draft
Intended status: Standard
Expires: December 31, 2011

D. Dhody
U. Palle
Huawei Technologies India Pvt Ltd
June 29, 2011

Supporting explicit-path per destination in Path Computation Element
Communication Protocol (PCEP) - P2MP Path Request.
draft-dhody-pce-pcep-p2mp-per-destination-00

Abstract

The ability to determine paths of point-to-multipoint (P2MP) Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineering Label Switched Paths (TE LSPs) is one the key requirements for Path Computation Element (PCE). [RFC 6006] and [PCE-P2MP-PROCEDURES] describes these mechanisms for intra and inter domain environment.

Explicit Path in this document refers to the configured list of network elements that MUST be traversed or MUST be excluded in the final path computation. This should not be confused with the RSVP terminology. Network elements can further be strict or loose hop.

This document describes extensions to the PCE communication Protocol (PCEP) to define explicit-path per destination in P2MP context.

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

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

This Internet-Draft will expire on December 31, 2011.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This Internet-Draft will expire on December 31, 2011.

Table of Contents

1. Introduction	3
1.1. Requirements Language	3
2. Terminology	3
3. Need to Define Explicit Path Per Destination	4
4. Detailed Description	4
4.1. Objective	4
4.2. Request Message Format	4
4.3. Ordering Destinations in END-POINTS Objects	5
4.3.1. Example	6
5. IANA Considerations	8
6. Security Considerations	8
7. Manageability Considerations	8
7.1. Control of Function and Policy	8
7.2. Information and Data Models	9
7.3. Liveness Detection and Monitoring	9
7.4. Verify Correct Operations	9
7.5. Requirements On Other Protocols	9
7.6. Impact On Network Operations	9
8. Acknowledgments	9
9. References	9
9.1. Normative References	9
9.2. Informative References	9

1. Introduction

The Path Computation Element (PCE) architecture is defined in [RFC4655]. [RFC 6006] describe a PCE-based path computation procedure to compute optimal constrained (G)MPLS P2MP TE LSPs. It also defines the format of path request message used in P2MP, which limits explicit path in form of <IRO> / <XRO> to be applied to full P2MP tunnel and thus to only the common path to all leaves.

This document describes the need for supporting explicit-path per destination in intra and inter-domain P2MP scenario. It further lists the path request format and mode of operations

1.1. 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.

2. Terminology

The following terminology is used in this document.

Explicit-Path: Set of network elements configured by the administrator that MUST be traversed or MUST be excluded.

IRO: Include Route Object.

PCC: Path Computation Client: any client application requesting a path computation to be performed by a Path Computation Element.

PCE: Path Computation Element. An entity (component, application, or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints.

P2MP: Point-to-Multipoint

P2P: Point-to-Point

RRO: Record Route Object

RSVP: Resource Reservation Protocol

TE LSP: Traffic Engineering Label Switched Path.

XRO: Exclude Route Object.

3. Need to Define Explicit Path Per Destination

- o [PCE-P2MP-PROCEDURES] defines inter-domain P2MP path computation procedure, since different destinations will have different domain paths within the domain tree, it requires domain-sequence encoded in form of <IRO> to be attached per destination. It cannot be encoded for all destinations.
- o Administrator at the source can exert stronger control by providing explicit path (include, exclude, loose etc) per destination.
- o Compatibility: Basic MPLS TE P2MP Tunnel configurations for various operators support the configuration of explicit-path per destination.

4. Detailed Description

4.1. Objective

[RFC 6006] defines Request Message Format and Objects, along with <end-point-rro-pair-list>. This section introduce the concept of <iro-list>, <xro-list> and <metric-list> which are added to the <end-point-rro-pair-list> to support 'per destination'.

Use of <iro-list>, <xro-list> to carry explicit-path per destination.

Use of <iro-list> to carry domain-sequence per destination in inter-domain scenario.

Use of <metric-list> to carry metric value of each calculated path encoded in <rro-list>.

4.2. Request Message Format

To carry explicit path for each destination, <END-POINTS> objects need to be ordered and grouped in a way such that IRO object, XRO object, RRO object and METRIC object can be associated with each destination.

The format of PCReq message is modified as follows:

```

<PCReq Message> ::= <Common Header>
                    <request>
where:
<request> ::= <RP>
              <end-point-iro-xro-rro-metric-list>
              [<OF>]
              [<LSPA>]
              [<BANDWIDTH>]
              [<metric-list>]
              [<IRO>]
              [<LOAD-BALANCING>]

where:
<end-point-iro-xro-rro-metric-list> ::=
              <END-POINTS>
              [<IRO-List>]
              [<XRO-List>]
              [<RRO-List>]
              [<metric-list>]
              [<end-point-iro-xro-rro-metric-list >]

<RRO-List> ::= <RRO> [<BANDWIDTH>] [<RRO-List>]
<metric-list> ::= <METRIC> [<metric-list>]
<IRO-List> ::= <IRO> [<IRO-List>]
<XRO-List> ::= <XRO> [<XRO-List>]

```

From [RFC 6006] usage of <end-point-rro-pair-list> is changed to <end-point-iro-xro-rro-metric-list> in this document.

Note that the new format is backward compatible to [RFC 6006] format.

4.3. Ordering Destinations in END-POINTS Objects

Multiple destinations are encoded into a single ENDPOINTS object, Each Endpoint maybe followed by multiple lists of IROs, XROs, RROs or METRICs. The first <IRO> object would belong to the first destination, the second <IRO> object to the second destination and hence forth. The first <XRO> object would belong to the first destination, the second <XRO> object to the second destination and hence forth...

Note that a destination (P2MP tree leaf) MAY have

- o both <IRO> and <XRO>
- o <IRO> only
- o <XRO> only
- o No explicit path

To maintain the ordering between the destination and objects in the list, there MAYBE a need to divide a set of destinations into multiple ENDPOINTS, this explained in below example.

4.3.1. Example

Destination 1 has include IRO1 and exclude XRO1

Destination 2 has only include IRO2

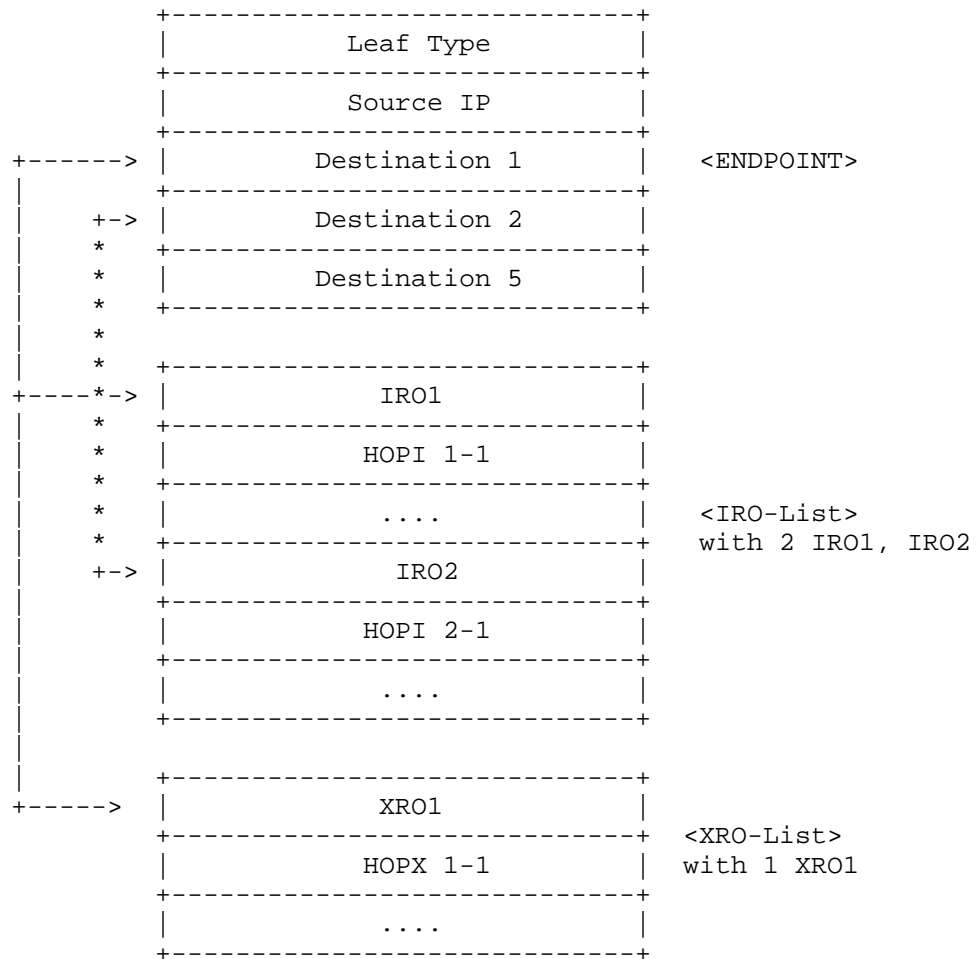
Destination 3 has only exclude XRO3

Destination 4 has only exclude XRO4

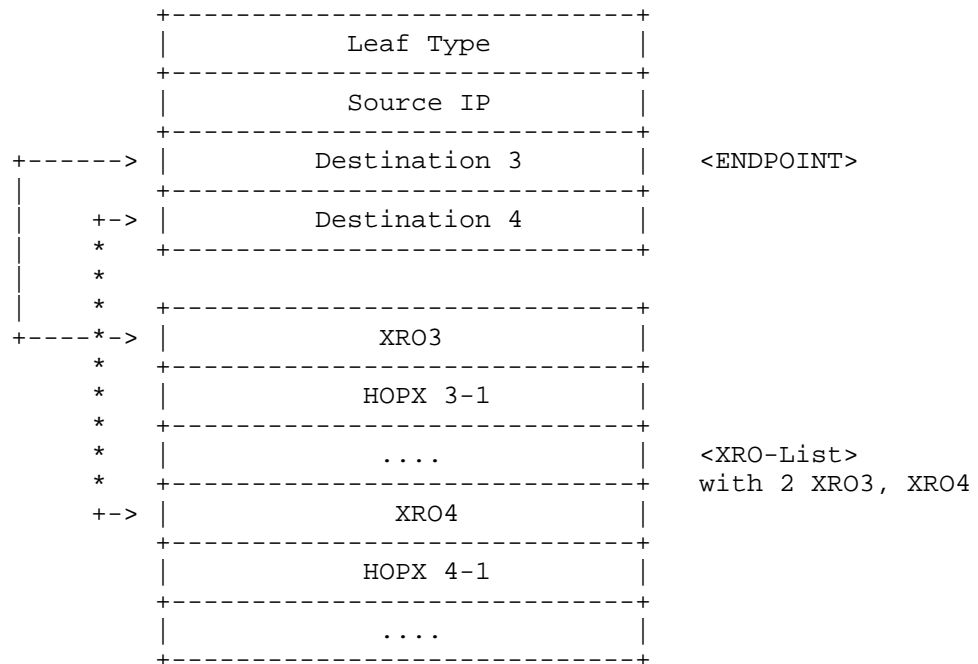
Destination 5 has none

Here if we try to encode all destinations in one <ENDPOINT> and objects in list, we will not map XRO3 to destination 3, the rule is to map sequentially and thus XRO3 will belong to destination 2.

To avoid this we must break the set of destinations into two sets as shown below



ENDPOINT1 carries destination 1, 2 and 5 and corresponding <iro-list> and <xro-list>. Here Destination 1 has IRO1 and XRO1; Destination 2 has IRO2; and Destination 5 has none.



ENDPOINT2 carries destination 3 and 4 and corresponding <xro-list> only. Here destination 3 maps to XRO3 and Destination 4 to XRO4.

5. IANA Considerations

TBD

6. Security Considerations

PCEP security mechanisms as described in [RFC6006] and [PCE-P2MP-PROCEDURES] are applicable for this document. This document does not add any new security threat.

7. Manageability Considerations

7.1. Control of Function and Policy

Mechanisms defined in this document do not add any new control function/policy requirements in addition to those already listed in [RFC6006].

7.2. Information and Data Models

Mechanisms defined in this document do not imply any new MIB requirements in addition to those already listed in [PCE-P2MP-MIB].

7.3. Liveness Detection and Monitoring

Mechanisms defined in this document do not imply any new liveness detection and monitoring requirements in addition to those already listed in [RFC6006].

7.4. Verify Correct Operations

Mechanisms defined in this document do not imply any new operation verification requirements in addition to those already listed in [RFC6006].

7.5. Requirements On Other Protocols

Mechanisms defined in this document do not imply any requirements on other protocols in addition to those already listed in [RFC6006].

7.6. Impact On Network Operations

Mechanisms defined in this document do not have any impact on network operations in addition to those already listed in [RFC6006].

8. Acknowledgments

We would like to thank Pradeep Shastry, Suresh babu, Quintin Zhao and Chen Huaimo for their useful comments and suggestions.

9. References

9.1. Normative References

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

9.2. Informative References

[PCE-P2MP-MIB] Zhao, Q., Dhody, D., Palle, U., and D. King, "Management Information Base for the PCE Communications Protocol (PCEP) When Requesting Point-to-Multipoint Services", March 2011.

[PCE-P2MP-PROCEDURES] Zhao, Q., Ali, Z., Saad,, T., and D. King, "PCE-based Computation Procedure To Compute

Shortest Constrained P2MP Inter-domain Traffic Engineering Label Switched Paths", January 2011.

- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", August 2006.
- [RFC4657] Ash, J. and J. Le Roux, "Path Computation Element (PCE) Communication Protocol Generic Requirements", September 2006.
- [RFC5440] Ayyangar, A ., Farrel, A ., Oki, E., Atlas, A., Dolganow, A., Ikejiri, Y., Kumaki, K., Vasseur, J., and J. Roux, "Path Computation Element (PCE) communication Protocol (PCEP)", March 2009.
- [RFC6006] Zhao, Q., Verhaeghe, F., Takeda, T., Ali, Z., Meuric , J., and D. King, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Point-to-Multipoint Traffic Engineering Label Switched Paths", September 2010.

Authors' Addresses

Dhruv Dhody
Huawei Technologies India Pvt Ltd
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: dhruvd@huawei.com

Udayasree Palle
Huawei Technologies India Pvt Ltd
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: udayasreepalle@huawei.com

Internet-Draft

PER-DEST

June 2011

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 9, 2012

C. Margaria, Ed.
Nokia Siemens Networks
O. Gonzalez de Dios, Ed.
Telefonica Investigacion y
Desarrollo
F. Zhang, Ed.
Huawei Technologies
July 08, 2011

PCEP extensions for GMPLS
draft-ietf-pce-gmpls-pcep-extensions-03

Abstract

This memo provides extensions for the Path Computation Element communication Protocol (PCEP) for the support of GMPLS control plane.

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 <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 January 9, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as

described in the Simplified BSD License.

Table of Contents

1. Introduction	3
1.1. Contributing Authors	3
1.2. PCEP requirements for GMPLS	3
1.3. PCEP existing objects related to GMPLS	4
1.4. Requirements Language	6
2. PCEP objects and extensions	7
2.1. RP object extension	8
2.2. Traffic parameters encoding, GENERALIZED-BANDWIDTH	9
2.3. Traffic parameters encoding, GENERALIZED-LOAD-BALANCING	11
2.4. END-POINTS Object extensions	14
2.4.1. Generalized Endpoint Object Type	15
2.4.2. END-POINTS TLVs extensions	18
2.5. LABEL-SET object	21
2.6. SUGGESTED-LABEL-SET object	22
2.7. LSPA extensions	22
2.8. NO-PATH Object Extension	23
2.8.1. Extensions to NO-PATH-VECTOR TLV	23
3. Additional Error Type and Error Values Defined	25
4. Manageability Considerations	27
5. IANA Considerations	28
5.1. PCEP Objects	28
5.2. END-POINTS object, Object Type Generalized Endpoint	29
5.3. New PCEP TLVs	30
5.4. RP Object Flag Field	31
5.5. New PCEP Error Codes	31
5.6. New NO-PATH-VECTOR TLV Fields	33
6. Security Considerations	34
7. Contributing Authors	35
8. Acknowledgments	37
9. References	38
9.1. Normative References	38
9.2. Informative References	39
Authors' Addresses	41

1. Introduction

PCEP RFCs [RFC5440], [RFC5521], [RFC5541], [RFC5520] are focused on path computation requests in MPLS networks. [RFC4655] defines the PCE framework also for GMPLS networks. This document complements these RFCs by providing some consideration of GMPLS applications and routing requests, for example for OTN and WSON networks.

The requirements on PCE extensions to support those characteristics are described in [I-D.ietf-pce-gmpls-aps-req] and [I-D.ietf-pce-wson-routing-wavelength].

1.1. Contributing Authors

Elie Sfeir, Franz Rambach (Nokia Siemens Networks) Francisco Javier Jimenez Chico (Telefonica Investigacion y Desarrollo) Suresh BR, Young Lee, SenthilKumar S, Jun Sun (Huawei Technologies), Ramon Casellas (CTTC)

1.2. PCEP requirements for GMPLS

This section provides a set of PCEP requirements to support GMPLS LSPs and assure signal compatibility in the path. When requesting a path computation (PCReq) to PCE, the PCC should be able to indicate, according to [I-D.ietf-pce-gmpls-aps-req] and to RSVP procedures like explicit label control (ELC), the following additional attributes:

(1) Switching capability: for instance PSC1-4, L2SC, TDM, LSC, FSC

(2) Encoding type: as defined in [RFC4202], [RFC4203], e.g., Ethernet, SONET/SDH, Lambda, etc.

(3) Signal Type: Indicates the type of elementary signal that constitutes the requested LSP. A lot of signal types with different granularity have been defined in SONET/SDH and G.709 ODUk, such as VC11, VC12, VC2, VC3 and VC4 in SDH, and ODU1, ODU2 and ODU3 in G.709 ODUk [RFC4606], [RFC4328] and other signal types like the one defined in [I-D.ceccarelli-ccamp-gmpls-ospf-g709] or [I-D.zhang-ccamp-gmpls-evolving-g709] .

(4) Concatenation Type: In SDH/SONET and G.709 OTN networks, two kinds of concatenation modes are defined: contiguous concatenation which requires co-route for each member signal and requires all the interfaces along the path to support this capability, and virtual concatenation which allows diverse routes for the member signals and only requires the ingress and egress interfaces to support this capability. Note that for the virtual concatenation, it also may specify co-routed or separated-routed. See [RFC4606]

and [RFC4328] about concatenation information.

(5) Concatenation Number: Indicates the number of signals that are requested to be contiguously or virtually concatenated. See also [RFC4606] and [RFC4328].

(6) Technology specific label(s) such as wavelength label as defined in [RFC6205]

(7) e2e Path protection type: as defined in [RFC4872], e.g., 1+1 protection, 1:1 protection, (pre-planned) rerouting, etc.

(8) Link Protection type: as defined in [RFC4203]

(9) Support for unnumbered interfaces: as defined in [RFC3477]

(10) Support for asymmetric bandwidth requests.

(11) Ability to indicate the requested granularity for the path ERO: node, link, label. This is to allow the use of the explicit label control of RSVP.

(12) In order to support the label control the Path computation response should provide label information matching signaling capabilities

(13) The PCC should be able to provide label restrictions similar to RSVP on the requests.

We describe in this document a proposal to fulfill those requirements.

1.3. PCEP existing objects related to GMPLS

PCEP as of [RFC5440], [RFC5521] and [I-D.ietf-pce-inter-layer-ext], supports the following information (in the PCReq and PCRep) related to the described requirements.

From [RFC5440]:

- o numbered endpoints
- o bandwidth (encoded as IEEE float)
- o ERO
- o LSP attributes (setup and holding priorities)

- o Request attribute (include some LSP attributes)

From [RFC5521], Extensions to PCEP for Route Exclusions, definition of a XRO object and a new semantic (F bit):

- o This object also allows to exclude (strict or not) resources; XRO include the diversity level (node, link, SRLG). The requested diversity is expressed in the XRO
- o This Object with the F bit set indicates that the existing route is failed and resources present in the RRO can be reused.

From [I-D.ietf-pce-inter-layer-ext]:

- o INTER-LAYER : indicates if inter-layer computation is allowed
- o SWITCH-LAYER : indicates which layer(s) should be considered, can be used to represent the RSVP-TE generalized label request
- o REQ-ADAP-CAP : indicates the adaptation capabilities requested, can also be used for the endpoints in case of mono-layer computation

The shortcomings of the existing PCEP information are:

The BANDWIDTH and LOAD-BALANCING objects do not describe the details of the traffic request (for example NVC, multiplier) in the context of GMPLS networks, for instance TDM or OTN networks.

The END-POINTS object does not allow specifying an unnumbered interface, nor the labels on the interface. Those parameters are of interest in case of switching constraints.

Current attributes do not allow to express the requested link level protection and end-to-end protection attributes.

The covered PCEP extensions are:

New objects are introduced (GENERALIZED-BANDWIDTH and GENERALIZED-LOAD-BALANCING) for flexible bandwidth encoding,

New Objects are introduced (LABEL-SET and SUGGESTED-LABEL-SET) in order to allow the PCC to restrict/influence the range of labels returned

A new object type is introduced for the END-POINTS object (generalized-endpoint),

A new TLV is added to the LSPA object.

In order to indicate the mandatory routing granularity in the response, a new flag in the RP object is added.

1.4. Requirements Language

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

2. PCEP objects and extensions

This section describes the required PCEP objects and extensions. The PCReq and PCRep messages are defined in [RFC5440]. The format of the request and response messages with the proposed extensions (GENERALIZED-BANDWIDTH, GENERALIZED-LOAD-BALANCING, SUGGESTED-LABEL-SET and LABEL-SET) is as follows:

```

<request> ::= <RP>
               <segment-computation> | <path-key-expansion>

<segment-computation> ::=
    <END-POINTS>
    [<LSPA>]
    [<BANDWIDTH>]
    [<GENERALIZED-BANDWIDTH>...]
    [<metric-list>]
    [<OF>]
    [<RRO> [<BANDWIDTH>] [<GENERALIZED-BANDWIDTH>...]]
    [<IRO>]
    [<SUGGESTED-LABEL-SET>]
    [<LABEL-SET>...]
    [<LOAD-BALANCING>]
    [<GENERALIZED-LOAD-BALANCING>...]
    [<XRO>]

<path-key-expansion> ::= <PATH-KEY>

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

<path-list> ::= <path> [<path-list>]
<path> ::= <ERO> <attribute-list>
<metric-list> ::= <METRIC> [<metric-list>]

```

Where:

```

<attribute-list> ::= [<LSPA>]
    [<BANDWIDTH>]
    [<LABEL-SET>...]
    [<SUGGESTED-LABEL-SET>...]
    [<GENERALIZED-BANDWIDTH>...]
    [<GENERALIZED-LOAD-BALANCING>...]
    [<metric-list>]
    [<IRO>]

```

For point-to-multipoint(P2MP) computations, the proposed grammar is:

```

<segment-computation> ::=
    <end-point-rro-pair-list>
    [<LSPA>]
    [<BANDWIDTH>]
    [<GENERALIZED-BANDWIDTH>...]
    [<metric-list>]
    [<IRO>]
    [<SUGGESTED-LABEL-SET>]
    [<LABEL-SET>]
    [<LOAD-BALANCING>]
    [<GENERALIZED-LOAD-BALANCING>...]
    [<XRO>]

<end-point-rro-pair-list> ::=
    <END-POINTS> [<RRO-List>] [<BANDWIDTH>]
    [<GENERALIZED-BANDWIDTH>...]
    [<end-point-rro-pair-list>]

<RRO-List> ::= <RRO> [<BANDWIDTH>]
    [< GENERALIZED-BANDWIDTH>...][<RRO-List>]

```

2.1. RP object extension

Explicit label control (ELC) is a procedure supported by RSVP-TE, where the outgoing label(s) is(are) encoded in the ERO. In consequence, the PCE may be able to provide such label(s) directly in the path ERO. The PCC, depending on policies or switching layer, may be required to use explicit label control or expect explicit link, thus it need to indicate in the PCReq which granularity it is expecting in the ERO. This correspond to requirement 11 of [I-D.ietf-pce-gmpls-aps-req] The possible granularities can be node, link, label. The granularities are inter-dependent, in the sense that link granularity imply the presence of node information in the ERO, similarly a label granularity imply that the ERO contain node, link and label information.

A new 2-bit routing granularity (RG) flag is defined in the RP object. The values are defined as follows

0 : node
1 : link
2 : label
3 : reserved

When the RP object appears in a request within a PCReq message the flag indicates the requested route granularity. The PCE MAY try to follow this granularity and MAY return a NO-PATH if the requested granularity cannot be provided. The PCE MAY return more details on the route based on its policy. The PCC can decide if the ERO is acceptable based on its content.

If a PCE did use the requested routing granularity in a PCReq is MUST indicate the routing granularity in the PCRep. The RG flag is backward-compatible with previous RFCs: the value sent by an implementation not supporting it will indicate a node granularity. This flag is optional for responses. A new capability flag in the PCE-CAP-FLAGS from [RFC5088] and [RFC5089] may be added.

2.2. Traffic parameters encoding, GENERALIZED-BANDWIDTH

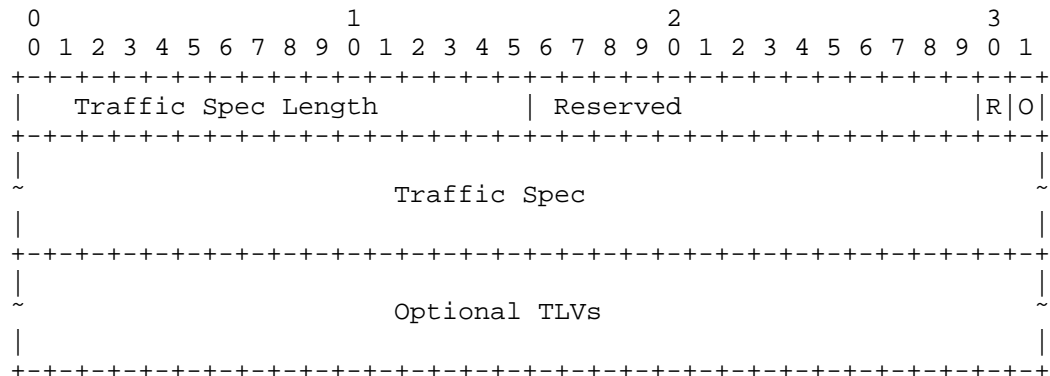
The PCEP BANDWIDTH does not describe the details of the signal (for example NVC, multiplier), hence the bandwidth information should be extended to use the RSVP Tspec object encoding. The PCEP BANDWIDTH object defines two types: 1 and 2. C-Type 2 is representing the existing bandwidth in case of re-optimization.

The following possibilities cannot be represented in the BANDWIDTH object:

- o Asymmetric bandwidth (different bandwidth in forward and reverse direction), as described in [RFC5467]
- o GMPLS (SDH/SONET, G.709, ATM, MEF etc) parameters are not supported.

This correspond to requirement 3,4,5 and 10 of [I-D.ietf-pce-gmpls-aps-req].

According to [RFC5440] the BANDWIDTH object has no TLV and has a fixed size of 4 bytes. This definition does not allow extending it with the required information. To express this information, a new object named GENERALIZED-BANDWIDTH having the following format is defined:



The GENERALIZED-BANDWIDTH has a variable length. The Traffic spec length field indicates the length of the Traffic spec field. The bits R and O have the following meaning:

O bit : when set the value refers to the previous bandwidth in case of re-optimization

R bit : when set the value refers to the bandwidth of the reverse direction

The Object type determines which type of bandwidth is represented by the object. The following object types are defined:

1. Intserv
2. SONET/SDH
3. G.709
4. Ethernet

The encoding of the field Traffic Spec is the same as in RSVP-TE, it can be found in the following references.

Object Type	Name	Reference
0	Reserved	
1	Reserved	
2	Intserv	[RFC2210]
3	Re served	
4	SONET/SDH	[RFC4606]
5	G.709	[RFC4328]
6	Ethernet	[RFC6003]

Traffic Spec field encoding

The GENERALIZED-BANDWIDTH MAY appear more than once in a PCReq message. If more than one GENERALIZED-BANDWIDTH have the same Object Type, Reserved, R and O values, only the first one is processed, the others are ignored.

a PCE MAY ignore GENERALIZED-BANDWIDTH objects, a PCC that requires a GENERALIZED-BANDWIDTH to be used can set the P (Processing) bit in the object header.

When a PCC needs to get a bi-directional path with asymmetric bandwidth, it SHOULD specify the different bandwidth in forward and reverse directions through two separate GENERALIZED-BANDWIDTH objects. If the PCC set the P bit on both object the PCE MUST compute a path that satisfies the asymmetric bandwidth constraint and return the path to PCC if the path computation is successful. If the P bit on the reverse GENERALIZED-BANDWIDTH object the PCE MAY ignore this constraint.

a PCE MAY include the GENERALIZED-BANDWIDTH objects in the response to indicate the GENERALIZED-BANDWIDTH of the path

Optional TLVs may be included within the object body to specify more specific bandwidth requirements. The specification of such TLVs is outside the scope of this document.

2.3. Traffic parameters encoding, GENERALIZED-LOAD-BALANCING

The LOAD-BALANCING object is used to request a set of maximum Max-LSP TE-LSP having in total the bandwidth specified in BANDWIDTH, each TE-LSP having a minimum of min-bandwidth bandwidth. The LOAD-BALANCING

follows the bandwidth encoding of the BANDWIDTH object, it does not describe enough details for the traffic specification expected by GMPLS. A PCC should be allowed to request a set of TE-LSP also in case of GMPLS traffic specification.

According to [RFC5440] the LOAD-BALANCING object has no TLV and has a fixed size of 8 bytes. This definition does not allow extending it with the required information. To express this information, a new Object named GENERALIZED-LOAD-BALANCING is defined.

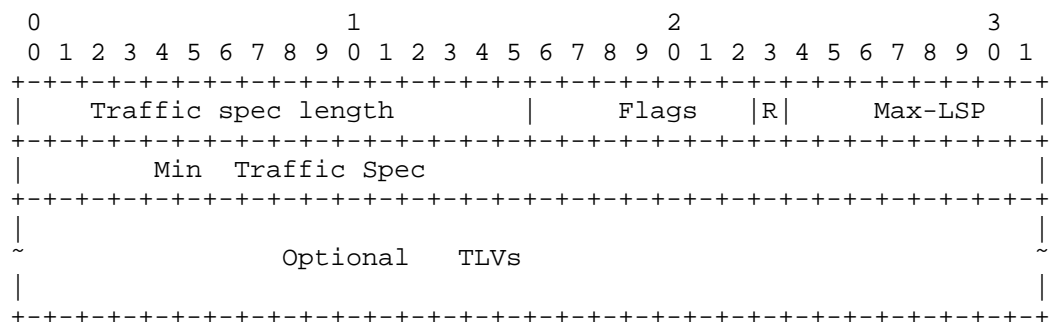
The GENERALIZED-LOAD-BALANCING object, as the LOAD-BALANCING object, allows the PCC to request a set of TE-LSP having in total the GENERALIZED-BANDWIDTH traffic specification with potentially Max-Lsp, each TE-LSP having a minimum of Min Traffic spec. The GENERALIZED-LOAD-BALANCING is optional.

GENERALIZED-LOAD-BALANCING Object-Class is to be assigned by IANA. The GENERALIZED-LOAD-BALANCING Object type determines which type of minimum bandwidth is represented by the object. The following object types are defined:

1. Intserv
2. SONET/SDH
3. G.709
4. Ethernet

The GENERALIZED-LOAD-BALANCING has a variable length.

The format of the GENERALIZED-LOAD-BALANCING object body is as follows:



Traffic spec length (16 bits): the total length of the min traffic specification. It should be noted that the RSVP traffic

specification may also include TLV different than the PCEP TLVs.

Flags (8 bits): The undefined Flags field MUST be set to zero on transmission and MUST be ignored on receipt. The following flag is defined:

R Flag : (1 bit) set when the value refer to the bandwidth of the reverse direction

Max-LSP (8 bits): maximum number of TE LSPs in the set.

Min-Traffic spec (variable): Specifies the minimum traffic spec of each element of the set of TE LSPs.

The encoding of the field Traffic Spec is the same as in RSVP-TE, it can be found in the following references.

Object Type	Name	Reference
2	Intserv	[RFC2210]
4	SONET/SDH	[RFC4606]
5	G.709	[RFC4328]
6	Ethernet	[RFC6003]

Traffic Spec field encoding

The GENERALIZED-LOAD-BALANCING MAY appear more than once in a PCReq message. If more than one GENERALIZED-LOAD-BALANCING have the same Object Type, and R Flag, only the first one is processed, the others are ignored.

a PCE MAY ignore GENERALIZED-LOAD-BALANCING objects. A PCC that requieres a GENERALIZED-LOAD-BALANCING to be used can set the P (Processing) bit in the object header.

When a PCC needs to get a bi-directional path with asymmetric bandwidth, it SHOULD specify the different bandwidth in forward and reverse directions through two separate GENERALIZED-LOAD-BALANCING objects with different R Flag. If the PCC set the P bit on both object the PCE MUST compute a path that satisfies the asymmetric bandwidth constraint and return the path to PCC if the path computation is successful. If the P bit on the reverse GENERALIZED-LOAD-BALANCING object the PCE MAY ignore this constraint.

Optional TLVs may be included within the object body to specify more

specific bandwidth requirements. The specification of such TLVs is outside the scope of this document.

The GENERALIZED-LOAD-BALANCING object has the same semantic as the LOAD-BALANCING object; If a PCC requests the computation of a set of TE LSPs so that the total of their generalized bandwidth is X, the maximum number of TE LSPs is N, and each TE LSP must at least have a bandwidth of B, it inserts a GENERALIZED-BANDWIDTH object specifying X as the required bandwidth and a GENERALIZED-LOAD-BALANCING object with the Max-LSP and Min-traffic spec fields set to N and B, respectively.

For example a request for one co-signaled n x VC-4 TE-LSP will not use the GENERALIZED-LOAD-BALANCING. In case the V4 components can use different paths, the GENERALIZED-BANDWIDTH will contain a traffic specification indicating the complete n x VC4 traffic specification and the GENERALIZED-LOAD-BALANCING the minimum co-signaled VC4. For a SDH network, a request to have a TE-LSP group with 10 VC4 container, each path using at minimum 2VC4 container, can be represented with a GENERALIZED-BANDWIDTH object with OT=4, the content of the Traffic specification is ST=6,RCC=0,NCC=0,NVC=10,MT=1. The GENERALIZED-LOAD-BALANCING, OT=4,R=0,Max-LSP=5, min Traffic spec is (ST=6,RCC=0,NCC=0,NVC=2,MT=1). The PCE can respond with a response with maximum 5 path, each of then having a GENERALIZED-BANDWIDTH OT=4,R=0, and traffic spec matching the minimum traffic spec from the GENERALIZED-LOAD-BALANCING object of the corresponding request.

2.4. END-POINTS Object extensions

The END-POINTS object is used in a PCReq message to specify the source and destination of the path for which a path computation is requested. From [RFC3471] the source IP address and the destination IP address are used to identify those. A new Object Type is defined to address the following possibilities:

- o Different endpoint types.
- o Label restrictions on the endpoint.
- o Specification of unnumbered endpoints type as seen in GMPLS networks.

The Object encoding is described in the following sections.

2.4.1. Generalized Endpoint Object Type

In GMPLS context the endpoints can:

- o Be unnumbered
- o Have label(s) associated to them
- o May have different switching capabilities

The IPv4 and IPv6 endpoints are used to represent the source and destination IP addresses. The scope of the IP address (Node or Link) is not explicitly stated. It should also be possible to request a Path between a numbered link and an unnumbered link, or a P2MP path between different type of endpoints.

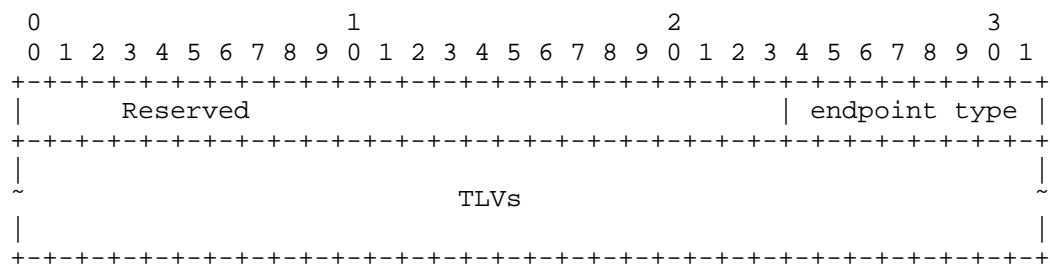
Since the PCEP END-POINTS object only support endpoints of the same type a new C-Type is proposed that support different endpoint types, including unnumbered. This new C-Type also supports the specification of constraints on the endpoint label to be use. The PCE might know the interface restrictions but this is not a requirement. On the path calculation request only the Tspec and switch layer need to be coherent, the endpoint labels could be different (supporting a different Tspec). Hence the label restrictions include a Generalized label request in order to interpret the labels. This correspond to requirement 6 and 9 of [I-D.ietf-pce-gmpls-aps-req].

The proposed object format consists of a body and a list of TLVs, which give the details of the endpoints and are described in Section 2.4.2. For each endpoint type, a different grammar is defined. The TLVs defined to describe an endpoint are:

1. IPv4 address.
2. IPv6 address.
3. Unnumbered endpoint.
4. Label request.
5. Label.
6. Upstream label.
7. Label set.

8. Suggested label set.

The labels TLV are used to restrict the label allocation in the PCE. They follow the set of restrictions provided by signaling with explicit value (label and upstream label), mandatory range restrictions (Label set) and optional range restriction (suggested label set). Single suggested value is using the suggested label set. The label range restriction are valid in GMPLS networks, either by PCC policy or depending on the switching technology used, for instance on given Ethernet or ODU equipment having limited hardware capabilities restricting the label range. Label set restriction also applies to WSON networks where the optical sender and receivers are limited in their frequency tunability ranges, restricting then in GMPLS the possible label ranges on the interface. The END-POINTS Object with Generalized Endpoint object type is encoded as follow:



Reserved bits should be set to 0 when a message is sent and ignored when the message is received

the endpoint type is defined as follow:

Value	Type	Meaning
0	Point-to-Point	
1	Point-to-Multipoint	New leaves to add
2		Old leaves to remove
3		Old leaves whose path can be modified/reoptimized
4		Old leaves whose path must be left unchanged
5-244	Reserved	
245-255	Experimental range	

The endpoint type is used to cover both point-to-point and different point-to-multipoint endpoint semantic. Endpoint type 0 MAY be accepted by the PCE, other endpoint type MAY be supported if the PCE implementation supports P2MP path calculation. A PCE not supporting a given endpoint type MUST respond with a PCErr with error code "Path computation failure", error type "Unsupported endpoint type in END-POINTS Generalized Endpoint object type". The TLVs present in the object body MUST follow the following grammar:

```

<generalized-endpoint-tlvs> ::=
  <p2p-endpoints> | <p2mp-endpoints>

<p2p-endpoints> ::=
  <source-endpoint>
  <destination-endpoint>

<source-endpoint> ::=
  <endpoint>
  [<endpoint-restriction-list>]

<destination-endpoint> ::=
  <endpoint>
  [<endpoint-restriction-list>]

<p2mp-endpoints> ::=
  <endpoint> [<endpoint-restriction-list>]
  [<endpoint> [<endpoint-restriction-list>]]...

```

For endpoint type Point-to-Multipoint several endpoint objects may be

present in the message and represent a leave, exact meaning depend on the endpoint type defined of the object.

An endpoint is defined as follows:

```

<endpoint> ::= <IPV4-ADDRESS> | <IPV6-ADDRESS> | <UNNUMBERED-ENDPOINT>
<endpoint-restriction-list> ::=
    <endpoint-restriction>
    [<endpoint-restriction-list>]

<endpoint-restriction> ::=
    <LABEL-REQUEST> <label-restriction-list>

<label-restriction-list> ::= <label-restriction>
    [<label-restriction-list>]
<label-restriction> ::= <LABEL> | <UPSTREAM-LABEL> |
    <LABEL-SET> |
    <SUGGESTED-LABEL-SET>

```

The different TLVs are described in the following sections. A PCE MAY support IPV4-ADDRESS, IPV6-ADDRESS or UNNUMBERED-ENDPOINT TLV. A PCE not supporting one of those TLV in a PCReq MUST respond with a PCRep with NO-PATH with the bit "Unknown destination" or "Unknown source" in the NO-PATH-VECTOR TLV, the PCRep MUST include the ENDPOINT object in the response with only the TLV it did not understood.

A PCE MAY support LABEL-REQUEST, LABEL, UPSTREAM-LABEL, LABEL-SET or SUGGESTED-LABEL-SET TLV. A PCE not supporting one of those TLV in a PCReq MUST respond with a PCRep with NO-PATH with the bit "No endpoint label resource" or "No endpoint label resource in range" in the NO-PATH-VECTOR TLV, the PCRep MUST include the ENDPOINT object in the response with only the TLV it did not understood or could not met the the constraint.

2.4.2. END-POINTS TLVs extensions

All endpoint TLVs have the standard PCEP TLV header as defined in [RFC5440] section 7.1

2.4.2.1. IPV4-ADDRESS

This TLV represent a numbered endpoint using IPv4 numbering, the format of the IPV4-ADDRESS TLV value (TLV-Type=TBA) is 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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     IPv4 address                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This TLV MAY be ignored, in which case a PCRep with NO-PATH should be responded, as described in Section 2.4.1.

2.4.2.2. IPV6-ADDRESS TLV

This TLV represent a numbered endpoint using IPV6 numbering, the format of the IPV6-ADDRESS TLV value (TLV-Type=TBA) is 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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     IPv6 address (16 bytes)                                     |
|-----|
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This TLV MAY be ignored, in which case a PCRep with NO-PATH should be responded, as described in Section 2.4.1.

2.4.2.3. UNNUMBERED-ENDPOINT TLV

This TLV represent an unnumbered interface. This TLV has the same semantic as in [RFC3477] The TLV value is encoded as follow (TLV-Type=TBA)

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     LSR's Router ID                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Interface ID (32 bits)                               |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This TLV MAY be ignored, in which case a PCRep with NO-PATH should be responded, as described in Section 2.4.1.

2.4.2.4. LABEL-REQUEST TLV

The LABEL-REQUEST TLV indicates the switching capability and encoding type of the label restriction list. Its format is the same as described in [RFC3471] Section 3.1 Generalized label request. The

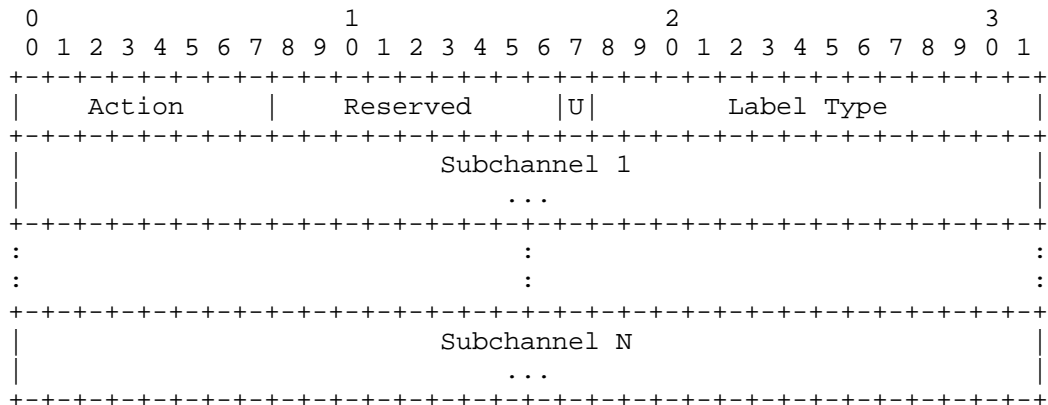
LABEL-REQUEST TLV use TLV-Type=TBA. The fields are encoded as in the RSVP-TE. The Encoding Type indicates the encoding type, e.g., SONET/SDH/GigE etc., that will be used with the data associated with the LSP. The Switching type indicates the type of switching that is being requested on the link. G-PID identifies the payload of the TE-LSP. This TLV and the following one are introduced to satisfy requirement 13 for the endpoint.

This TLV MAY be ignored, in which case a PCRep with NO-PATH should be responded, as described in Section 2.4.1.

2.4.2.5. Labels TLV

Label or label range restrictions may be specified for the TE-LSP endpoints. Those are encoded in the TLVs. The label value need to be interpreted with a description on the Encoding and switching type. The REQ-ADAP-CAP object from [I-D.ietf-pce-inter-layer-ext] can be used in case of mono-layer request, however in case of multilayer it is possible to have in the future more than one object, so it is better to have a dedicated TLV for the label and label request (the scope is then more clear). Those TLV MAY be ignored, in which case a PCRep with NO-PATH should be responded, as described in Section 2.4.1. TLVs are encoded as follow (following [RFC5440]) :

- o LABEL TLV, Type=TBA. The TLV Length is variable, the value is the same as [RFC3471] Section 3.2 Generalized label. This represent the downstream label
- o UPSTREAM-LABEL TLV, Type=TBA, The TLV Length is variable, the value is the same as [RFC3471] Section 3.2 Generalized label. This represent the upstream label
- o LABEL-SET TLV, Type=TBA. The TLV Length is variable, Encoding follow [RFC3471] Section 3.5 "Label set" with the addition of a U bit : the U bit is set for upstream direction in case of bidirectional LSP.



- o SUGGESTED-LABEL-SET TLV Set, Type=TBA. The TLV length is variable, Encoding is as LABEL-SET TLV.

A LABEL TLV represent the label used on the unnumbered interface, bit U is used to indicate which exact direction is considered. The label type indicates which type of label is carried. A LABEL-SET TLV represents a set of possible labels that can be used on the unnumbered interface. the label allocated on the first link SHOULD be within the label set range. The action parameter in the Label set indicates the type of list provided. Those parameters are described by [RFC3471] section 3.5.1 A SUGGESTED-LABEL-SET TLV has the same encoding as the LABEL-SET TLV, it indicates to the PCE a set of preferred (ordered) set of labels to be used. the PCE MAY use those labels for label allocation.

The U bit has the following meaning:

U: Upstream direction: set when the label or label set is in the reverse direction

2.5. LABEL-SET object

The LABEL-SET object is carried in a request within a PCReq message to restrict the set of labels to be assigned during the path computation. This is introduced to satisfy requirement 13.

When the P bit is set and the object accepted any label allocated by the PCE (and included in the ERO object on the response) MUST be in the range stated in the LABEL-SET. When no path satisfy this constraint a PCRep with a NO-PATH should be responded wit a NO-PATH-VECTOR TLV with the bit "No label resource in range" set and the

LABEL-SET object MAY be included to indicate the set of constraint that could not be satisfied.

When the P bit is not set a PCE MAY consider constraint, the PCC can verify that the constraint was applied by checking the ERO returned

The LABEL-SET Object encoding is defined as following

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                           |
|//                               TLVs                               //|
|                                           |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

where TLVs follow the following grammar

```
<label-set-tlvs> ::= <LABEL-REQUEST><LABEL-SET>[<LABEL-SET>]
```

The LABEL-REQUEST and LABEL-SET TLVs are as defined in Section 2.4.2.5, See also [RFC3471] and [RFC3473] for the definitions of the fields.

It is allowed to have more than one LABEL-SET object per request within a PCReq message (for example in case of multiple SWITCH-LAYER present).

2.6. SUGGESTED-LABEL-SET object

Similar to the endpoint restriction SUGGESTED-LABEL-SET TLV, but with end-to-end scope the SUGGESTED-LABEL-SET object indicate an optional set of label that the PCE MAY use when selecting the labels. The SUGGESTED-LABEL-SET object is carried within a PCReq or PCRep message to indicate the preferred set of label to be assigned during the path computation. The encoding is the same as the LABEL-SET object. It is allowed to have more than one SUGGESTED LABEL-SET object per PCReq (for example in case of multiple SWITCH-LAYER present).

This object is introduced similarly to the LABEL-SET to satisfy the requirement 6 and 13, more specifically the ability to indicate optional preference for the label selection support by RSVP using the SUGGESTED_LABEL.

2.7. LSPA extensions

The LSPA carries the LSP attributes. In the end-to-end protection context this also includes the protection state information. This

object is introduced to fulfill requirement 7 and is used as a policy input for route and label selection. The LSPA object can be extended by a protection TLV type: Type TBA: PROTECTION-ATTRIBUTE

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																													
S	P	N	O	Reserved						LSP Flags										Reserved										Link Flags									
I	R	Reserved						Seg.Flags												Reserved																			

The content is as defined in [RFC4872], [RFC4873].

LSP Flags can be considered for routing policy based on the protection type. The other attributes are only meaningful for a s_ateful PCE.

This TLV is optional and MAY be ignored by the PCE, in which case MUST NOT include the TLV in the LSPA, if present, of the PCRep. When the TLV is used by the PCE, a LSPA object and the PROTECTION-ATTRIBUTE TLV MUST be included in the PCRep. Fields that were not considered MUST be set to 0.

2.8. NO-PATH Object Extension

The NO-PATH object is used in PCRep messages in response to an unsuccessful path computation request (the PCE could not find a path satisfying the set of constraints). In this scenario, PCE MUST include a NO-PATH object in the PCRep message. The NO-PATH object may carries the NO-PATH-VECTOR TLV that specifies more information on the reasons that led to a negative reply. In case of GMPLS networks there could be some more additional constraints that led to the failure like protection mismatch, lack of resources, and so on. Few new flags have been introduced in the 32-bit flag field of the NO-PATH-VECTOR TLV and no modifications have been made in the NO-PATH object.

2.8.1. Extensions to NO-PATH-VECTOR TLV

The modified NO-PATH-VECTOR TLV carrying the additional information is as follows: New fields PM and NR are defined in the 23th and 22th bit of the Flags field respectively.

Bit number TBA - Protection Mismatch (1-bit). Specifies the mismatch of the protection type in the PROTECTION-ATTRIBUTE TLV in

the request.

Bit number TBA - No Resource (1-bit). Specifies that the resources are not currently sufficient to provide the path.

Bit number TBA - Granularity not supported (1-bit). Specifies that the PCE is not able to provide a route with the requested granularity.

Bit number TBA - No endpoint label resource (1-bit). Specifies that the PCE is not able to provide a route because of the endpoint label restriction.

Bit number TBA - No endpoint label resource in range (1-bit). Specifies that the PCE is not able to provide a route because of the endpoint label set restriction.

Bit number TBA - No label resource in range (1-bit). Specifies that the PCE is not able to provide a route because of the label set restriction.

3. Additional Error Type and Error Values Defined

A PCEP-ERROR object is used to report a PCEP error and is characterized by an Error-Type that specifies the type of error while Error-value that provides additional information about the error type. An additional error type and few error values are defined to represent some of the errors related to the newly identified objects related to SDH networks. For each PCEP error, an Error-Type and an Error-value are defined. Error-Type 1 to 10 are already defined in [RFC5440]. Additional Error- values are defined for Error-Type 10 and A new Error-Type is introduced (value TBA).

Error-Type Error-value

10	Reception of an invalid object	
	Error-value=TBA:	Bad Generalized Bandwidth Object value.
	Error-value=TBA:	Unsupported LSP Protection Type in PROTECTION-ATTRIBUTE TLV.
	Error-value=TBA:	Unsupported LSP Protection Flags in PROTECTION-ATTRIBUTE TLV.
	Error-value=TBA:	Unsupported Secondary LSP Protection Flags in PROTECTION-ATTRIBUTE TLV.
	Error-value=TBA:	Unsupported Link Protection Type in PROTECTION-ATTRIBUTE TLV.
	Error-value=TBA:	Unsupported Link Protection Type in PROTECTION-ATTRIBUTE TLV.
TBA	Path computation failure	
	Error-value=TBA:	Unacceptable request message.
	Error-value=TBA:	Generalized bandwidth object not supported.
	Error-value=TBA:	Label Set constraint could not be met.
	Error-value=TBA:	Label constraint could not be met.
	Error-value=TBA:	Unsupported endpoint type in END-POINTS Generalized Endpoint object type

Error-value=TBA: Unsupported TLV present in END-POINTS
Generalized Endpoint object type

Error-value=TBA: Unsupported granularity in the RP object
flags

4. Manageability Considerations

Liveness Detection and Monitoring This document makes no change to the basic operation of PCEP and so there are no changes to the requirements for liveness detection and monitoring set out in [RFC4657] and [RFC5440].

5. IANA Considerations

IANA assigns values to the PCEP protocol objects and TLVs. IANA is requested to make some allocations for the newly defined objects and TLVs introduced in this document. Also, IANA is requested to manage the space of flags that are newly added in the TLVs.

5.1. PCEP Objects

As described in Section 2.2 and Section 2.3 new Objects are defined IANA is requested to make the following Object-Type allocations from the "PCEP Objects" sub-registry.

Object Class to be assigned

Name GENERALIZED-BANDWIDTH

Object-Type 0 to 6

Reference This document (section Section 2.2)

Object Class to be assigned

Name GENERALIZED-LOAD-BALANCING

Object-Type 0 to 6

Reference This document (section Section 2.3)

Object Class to be assigned

Name LABEL-SET

Object-Type 0

Reference This document (section Section 2.5)

Object Class to be assigned

Name SUGGESTED-LABEL-SET

Object-Type 0

Reference This document (section Section 2.6)

As described in Section 2.4.1 a new Object type is defined IANA is requested to make the following Object-Type allocations from the "PCEP Objects" sub-registry. The values here are suggested for use by IANA.

Object Class 4

Name END-POINTS

Object-Type 5 : Generalized Endpoint

6-15 : unassigned

Reference This document (section Section 2.2)

5.2. END-POINTS object, Object Type Generalized Endpoint

IANA is requested to create a registry to manage the endpoint type field of the END-POINTS object, Object Type Generalized Endpoint and manage the code space.

New endpoint type in the Reserved range may be allocated by an IETF consensus action. Each endpoint type should be tracked with the following qualities:

- o endpoint type
- o Description
- o Defining RFC

New endpoint type in the Experimental range are for experimental use; these will not be registered with IANA and MUST NOT be mentioned by RFCs.

The following values have been defined by this document.
(Section 2.4.1, Table 4):

Value	Type	Meaning
0	Point-to-Point	
1	Point-to-Multipoint	New leaves to add
2		Old leaves to remove
3		Old leaves whose path can be modified/reoptimized
4		Old leaves whose path must be left unchanged
5-244	Reserved	
245-255	Experimental range	

5.3. New PCEP TLVs

IANA manages the PCEP TLV code point registry (see [RFC5440]). This is maintained as the "PCEP TLV Type Indicators" sub-registry of the "Path Computation Element Protocol (PCEP) Numbers" registry. This document defines new PCEP TLVs, to be carried in the END-POINTS object with Generalized Endpoint object Type. IANA is requested to do the following allocation. The values here are suggested for use by IANA.

Value	Meaning	Reference
7	IPv4 endpoint	This document (section Section 2.4.2.1)
8	IPv6 endpoint	This document (section Section 2.4.2.2)
9	Unnumbered endpoint	This document (section Section 2.4.2.3)
10	Label request	This document (section Section 2.4.2.4)
11	Requested GMPLS Label	This document (section Section 2.4.2.5)
12	Requested GMPLS Upstream Label	This document (section Section 2.4.2.5)

- | | | |
|----|----------------------------|--|
| 13 | Requested GMPLS Label Set | This document (section
Section 2.4.2.5) |
| 14 | Suggested GMPLS Label Set | This document (section
Section 2.4.2.5) |
| 15 | LSP Protection Information | This document (section Section 2.7) |

5.4. RP Object Flag Field

As described in Section 2.1 new flag are defined in the RP Object Flag IANA is requested to make the following Object-Type allocations from the "RP Object Flag Field" sub-registry. The values here are suggested for use by IANA.

Bit	Description	Reference
bit 17-16	routing granularity (RG)	This document, Section 2.1

5.5. New PCEP Error Codes

As described in Section Section 3, new PCEP Error-Type and Error Values are defined. IANA is requested to make the following allocation in the "PCEP-ERROR Object Error Types and Values" registry. The values here are suggested for use by IANA.

Error	name	Reference
Type=10	Reception of an invalid object	[RFC5440]
Value=2:	Bad Generalized Bandwidth Object value.	This Document
Value=3:	Unsupported LSP Protection Type in PROTECTION-ATTRIBUTE TLV.	This Document
Value=4:	Unsupported LSP Protection Flags in PROTECTION-ATTRIBUTE TLV.	This Document
Value=5:	Unsupported Secondary LSP Protection Flags in PROTECTION-ATTRIBUTE TLV.	This Document
Value=6:	Unsupported Link Protection Type in PROTECTION-ATTRIBUTE TLV.	This Document
Value=7:	Unsupported Link Protection Type in PROTECTION-ATTRIBUTE TLV.	This Document
Type=14	Path computation failure	This Document
Value=1:	Unacceptable request message.	This Document
Value=2:	Generalized bandwidth object not supported.	This Document
Value=3:	Label Set constraint could not be met.	This Document
Value=4:	Label constraint could not be met.	This Document
Value=5:	Unsupported endpoint type in END-POINTS Generalized Endpoint object type	This Document
Value=6:	Unsupported TLV present in END-POINTS Generalized Endpoint object type	This Document
Value=7:	Unsupported granularity in the RP object flags	This Document

5.6. New NO-PATH-VECTOR TLV Fields

As described in Section 2.8.1, new NO-PATH-VECTOR TLV Flag Fields have been defined. IANA is requested to do the following allocations in the "NO-PATH-VECTOR TLV Flag Field" sub-registry. The values here are suggested for use by IANA.

Bit number 23 - Protection Mismatch (1-bit). Specifies the mismatch of the protection type of the PROTECTION-ATTRIBUTE TLV in the request.

Bit number 22 - No Resource (1-bit). Specifies that the resources are not currently sufficient to provide the path.

Bit number 21 - Granularity not supported (1-bit). Specifies that the PCE is not able to provide a route with the requested granularity.

Bit number 20 - No endpoint label resource (1-bit). Specifies that the PCE is not able to provide a route because of the endpoint label restriction.

Bit number 19 - No endpoint label resource in range (1-bit). Specifies that the PCE is not able to provide a route because of the endpoint label set restriction.

Bit number 18 - No label resource in range (1-bit). Specifies that the PCE is not able to provide a route because of the label set restriction.

6. Security Considerations

None.

7. Contributing Authors

Nokia Siemens Networks:

Elie Sfeir
St Martin Strasse 76
Munich, 81541
Germany

Phone: +49 89 5159 16159
Email: elie.sfeir@nsn.com

Franz Rambach
St Martin Strasse 76
Munich, 81541
Germany

Phone: +49 89 5159 31188
Email: franz.rambach@nsn.com

Francisco Javier Jimenez Chico
Telefonica Investigacion y Desarrollo
C/ Emilio Vargas 6
Madrid, 28043
Spain

Phone: +34 91 3379037
Email: fjjc@tid.es

Huawei Technologies

Suresh BR
Shenzhen
China
Email: sureshbr@huawei.com

Young Lee
1700 Alma Drive, Suite 100
Plano, TX 75075
USA

Phone: (972) 509-5599 (x2240)
Email: ylee@huawei.com

SenthilKumar S
Shenzhen
China
Email: senthilkumars@huawei.com

Jun Sun
Shenzhen
China
Email: johnsun@huawei.com

CTTC - Centre Tecnologic de Telecomunicacions de Catalunya

Ramon Casellas
PMT Ed B4 Av. Carl Friedrich Gauss 7
08860 Castelldefels (Barcelona)
Spain
Phone: (34) 936452916
Email: ramon.casellas@cttc.es

8. Acknowledgments

The research of Ramon Casellas, Francisco Javier Jimenez Chico, Oscar Gonzalez de Dios, Cyril Margaria, and Franz Rambach leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no 247674.

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2210] Wroclawski, J., "The Use of RSVP with IETF Integrated Services", RFC 2210, September 1997.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [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.
- [RFC4202] Kompella, K. and Y. Rekhter, "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.
- [RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC4328] Papadimitriou, D., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.
- [RFC4606] Mannie, E. and D. Papadimitriou, "Generalized Multi-Protocol Label Switching (GMPLS) Extensions for Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH) Control", RFC 4606, August 2006.
- [RFC4872] Lang, J., Rekhter, Y., and D. Papadimitriou, "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872, May 2007.
- [RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, May 2007.
- [RFC5088] Le Roux, JL., Vasseur, JP., Ikejiri, Y., and R. Zhang,

"OSPF Protocol Extensions for Path Computation Element (PCE) Discovery", RFC 5088, January 2008.

- [RFC5089] Le Roux, JL., Vasseur, JP., Ikejiri, Y., and R. Zhang, "IS-IS Protocol Extensions for Path Computation Element (PCE) Discovery", RFC 5089, January 2008.
- [RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.
- [RFC5520] Bradford, R., Vasseur, JP., and A. Farrel, "Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Path-Key-Based Mechanism", RFC 5520, April 2009.
- [RFC5521] Oki, E., Takeda, T., and A. Farrel, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Route Exclusions", RFC 5521, April 2009.
- [RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", RFC 5541, June 2009.
- [RFC6003] Papadimitriou, D., "Ethernet Traffic Parameters", RFC 6003, October 2010.
- [RFC6205] Otani, T. and D. Li, "Generalized Labels for Lambda-Switch-Capable (LSC) Label Switching Routers", RFC 6205, March 2011.

9.2. Informative References

- [I-D.ceccarelli-ccamp-gmpls-ospf-g709]
Ceccarelli, D., Caviglia, D., Zhang, F., Li, D., Belotti, S., Grandi, P., Rao, R., Pithewan, K., and J. Drake, "Traffic Engineering Extensions to OSPF for Generalized MPLS (GMPLS) Control of Evolving G.709 OTN Networks", draft-ceccarelli-ccamp-gmpls-ospf-g709-06 (work in progress), June 2011.
- [I-D.ietf-pce-gmpls-aps-req]
Otani, T., Ogaki, K., Caviglia, D., and F. Zhang, "Requirements for GMPLS applications of PCE", draft-ietf-pce-gmpls-aps-req-04 (work in progress), June 2011.
- [I-D.ietf-pce-inter-layer-ext]
Oki, E., Takeda, T., Roux, J., Farrel, A., and F. Zhang,

"Extensions to the Path Computation Element communication Protocol (PCEP) for Inter-Layer MPLS and GMPLS Traffic Engineering", draft-ietf-pce-inter-layer-ext-05 (work in progress), June 2011.

[I-D.ietf-pce-wson-routing-wavelength]

Lee, Y., Bernstein, G., Martensson, J., Takeda, T., and T. Tsuritani, "PCEP Requirements for WSON Routing and Wavelength Assignment", draft-ietf-pce-wson-routing-wavelength-04 (work in progress), March 2011.

[I-D.zhang-ccamp-gmpls-evolving-g709]

Zhang, F., Zhang, G., Belotti, S., Ceccarelli, D., and K. Pithewan, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for the evolving G.709 Optical Transport Networks Control", draft-zhang-ccamp-gmpls-evolving-g709-08 (work in progress), July 2011.

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

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

[RFC5467] Berger, L., Takacs, A., Caviglia, D., Fedyk, D., and J. Meuric, "GMPLS Asymmetric Bandwidth Bidirectional Label Switched Paths (LSPs)", RFC 5467, March 2009.

Authors' Addresses

Cyril Margaria (editor)
Nokia Siemens Networks
St Martin Strasse 76
Munich, 81541
Germany

Phone: +49 89 5159 16934
Email: cyril.margaria@nsn.com

Oscar Gonzalez de Dios (editor)
Telefonica Investigacion y Desarrollo
C/ Emilio Vargas 6
Madrid, 28043
Spain

Phone: +34 91 3374013
Email: ogondio@tid.es

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

Email: zhangfatai@huawei.com

Network Working Group
Internet-Draft

Intended status: Standards Track
Expires: June 14, 2020

C. Margaria, Ed.
Juniper
O. Gonzalez de Dios, Ed.
Telefonica Investigacion y Desarrollo
F. Zhang, Ed.
Huawei Technologies
December 12, 2019

PCEP extensions for GMPLS
draft-ietf-pce-gmpls-pcep-extensions-16

Abstract

A Path Computation Element (PCE) provides path computation functions for Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks. Additional requirements for GMPLS are identified in RFC7025.

This memo provides extensions to the Path Computation Element communication Protocol (PCEP) for the support of the GMPLS control plane to address those requirements.

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 June 14, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
1.1. Terminology	3
1.2. PCEP Requirements for GMPLS	5
1.3. Requirements Applicability	5
1.3.1. Requirements on Path Computation Request	6
1.3.2. Requirements on Path Computation Response	7
1.4. Existing Support for GMPLS in Base PCEP Objects and its Limitations	7
2. PCEP Objects and Extensions	10
2.1. GMPLS Capability Advertisement	10
2.1.1. GMPLS Computation TLV in the Existing PCE Discovery Protocol	10
2.1.2. OPEN Object Extension GMPLS-CAPABILITY TLV	10
2.2. RP Object Extension	11
2.3. BANDWIDTH Object Extensions	12
2.4. LOAD-BALANCING Object Extensions	14
2.5. END-POINTS Object Extensions	16
2.5.1. Generalized Endpoint Object Type	17
2.5.2. END-POINTS TLV Extensions	20
2.6. IRO Extension	24
2.7. XRO Extension	24
2.8. LSPA Extensions	26
2.9. NO-PATH Object Extension	26
2.9.1. Extensions to NO-PATH-VECTOR TLV	27
3. Additional Error-Types and Error-Values Defined	27
4. Manageability Considerations	29
4.1. Control of Function through Configuration and Policy	29
4.2. Information and Data Models	29
4.3. Liveness Detection and Monitoring	29
4.4. Verifying Correct Operation	30
4.5. Requirements on Other Protocols and Functional Components	30
4.6. Impact on Network Operation	30
5. IANA Considerations	30
5.1. PCEP Objects	30
5.2. Endpoint type field in Generalized END-POINTS Object	31
5.3. New PCEP TLVs	32
5.4. RP Object Flag Field	32
5.5. New PCEP Error Codes	32
5.6. New NO-PATH-VECTOR TLV Fields	33

5.7. New Subobject for the Include Route Object	34
5.8. New Subobject for the Exclude Route Object	34
5.9. New GMPLS-CAPABILITY TLV Flag Field	35
6. Security Considerations	35
7. Contributing Authors	36
8. Acknowledgments	38
9. References	38
9.1. Normative References	38
9.2. Informative References	42
Appendix A. LOAD-BALANCING Usage for SDH Virtual Concatenation .	43
Authors' Addresses	43

1. Introduction

Although [RFC4655] defines the PCE architecture and framework for both MPLS and GMPLS networks, most preexisting PCEP RFCs [RFC5440], [RFC5521], [RFC5541], [RFC5520] are focused on MPLS networks, and do not cover the wide range of GMPLS networks. This document complements these RFCs by addressing the extensions required for GMPLS applications and routing requests, for example for Optical Transport Network (OTN) and Wavelength Switched Optical Network (WSN) networks.

The functional requirements to be addressed by the PCEP extensions to support these applications are fully described in [RFC7025] and [RFC7449].

1.1. Terminology

This document uses terminologies from the PCE architecture document [RFC4655], the PCEP documents including [RFC5440], [RFC5521], [RFC5541], [RFC5520], [RFC7025] and [RFC7449], and the GMPLS documents such as [RFC3471], [RFC3473] and so on. Note that it is expected the reader is familiar with these documents. The following abbreviations are used in this document

ODU ODU Optical Channel Data Unit [G.709-v3]
OTN Optical Transport Network [G.709-v3]
L2SC Layer-2 Switch Capable [RFC3471]
TDM Time-Division Multiplex Capable [RFC3471]
LSC Lambda Switch Capable [RFC3471]
SONET Synchronous Optical Networking

SDH Synchronous Digital Hierarchy

PCC Path Computation Client

RSVP-TE Resource Reservation Protocol - Traffic Engineering

LSP Label Switched Path

TE-LSP Traffic Engineering LSP

IRO Include Route Object

ERO Explicit Route Object

XRO eXclude Route Object

RRO Record Route Object

LSPA LSP Attribute

SRLG Shared Risk Link Group

NVC Number of Virtual Components [RFC4328] [RFC4606]

NCC Number of Contiguous Components [RFC4328] [RFC4606]

MT Multiplier [RFC4328] [RFC4606]

RCC Requested Contiguous Concatenation [RFC4606]

PCReq Path Computation Request [RFC5440]

PCRep Path Computation Reply [RFC5440]

MEF Metro Ethernet Forum

SSON Spectrum-Switched Optical Network

P2MP Point to Multi-Point

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.

1.2. PCEP Requirements for GMPLS

The document [RFC7025] describes the set of PCEP requirements to support GMPLS TE-LSPs. This document assumes a significant familiarity with [RFC7025] and existing PCEP extensions. As a short overview, those requirements can be broken down into the following categories.

- o Which data flow is switched by the LSP: a combination of Switching type (for instance L2SC or TDM), LSP Encoding type (e.g., Ethernet, SONET/SDH) and sometimes the Signal Type (e.g., in case of TDM/LSC switching capability).
- o Data flow specific traffic parameters, which are technology specific. For instance, in SDH/SONET and [G.709-v3] OTN networks the Concatenation Type and the Concatenation Number have an influence on the switched data and on which link it can be supported
- o Support for asymmetric bandwidth requests.
- o Support for unnumbered interface identifiers, as defined in [RFC3477]
- o Label information and technology specific label(s) such as wavelength labels as defined in [RFC6205]. A PCC should also be able to specify a label restriction similar to the one supported by RSVP-TE in [RFC3473].
- o Ability to indicate the requested granularity for the path ERO: node, link or label. This is to allow the use of the explicit label control feature of RSVP-TE.

The requirements of [RFC7025] apply to several objects conveyed by PCEP, this is described in Section 1.3. Some of the requirements of [RFC7025] are already supported in existing documents, as described in Section 1.4.

This document describes a set of PCEP extensions, including new object types, TLVs, encodings, error codes and procedures, in order to fulfill the aforementioned requirements not covered in existing RFCs.

1.3. Requirements Applicability

This section follows the organization of [RFC7025] Section 3 and indicates, for each requirement, the affected piece of information carried by PCEP and its scope.

1.3.1. Requirements on Path Computation Request

- (1) Switching capability/type: as described in [RFC3471] this piece of information is used with the Encoding Type and Signal Type to fully describe the switching technology and data carried by the TE-LSP. This is applicable to the TE-LSP itself and also to the TE-LSP endpoint (Carried in the END-POINTS object for MPLS networks in [RFC5440]) when considering multiple network layers. Inter-layer path computation requirements are addressed in [RFC8282] which addressing the TE-LSP itself, but the TE-LSP endpoints are not addressed.
- (2) Encoding type: see (1).
- (3) Signal type: see (1).
- (4) Concatenation type: this parameter and the Concatenation Number (5) are specific to some TDM (SDH and ODU) switching technology. They MUST be described together and are used to derive the requested resource allocation for the TE-LSP. It is scoped to the TE-LSP and is related to the [RFC5440] BANDWIDTH object in MPLS networks. See [RFC4606] and [RFC4328] about concatenation information.
- (5) Concatenation number: see (4).
- (6) Technology-specific label(s): as described in [RFC3471] the GMPLS Labels are specific to each switching technology. They can be specified on each link and also on the TE-LSP endpoints, in WSON networks for instance, as described in [RFC6163]. The label restriction can apply to endpoints and on each hop, the related PCEP objects are END-POINTS, IRO, XRO and RRO.
- (7) End-to-End (E2E) path protection type: as defined in [RFC4872], this is applicable to the TE-LSP. In MPLS networks the related PCEP object is LSPA (carrying local protection information).
- (8) Administrative group: as defined in [RFC3630], this information is already carried in the LSPA object.
- (9) Link protection type: as defined in [RFC4872], this is applicable to the TE-LSP and is carried in association with the E2E path protection type.
- (10) Support for unnumbered interfaces: as defined in [RFC3477]. Its scope and related objects are the same as labels

- (11) Support for asymmetric bandwidth requests: as defined [RFC6387], the scope is similar to (4)
- (12) Support for explicit label control during the path computation. This affects the TE-LSP and amount of information returned in the ERO.
- (13) Support of label restrictions in the requests/responses: This is described in (6).

1.3.2. Requirements on Path Computation Response

- (1) Path computation with concatenation: This is related to Path Computation request requirement (4). In addition there is a specific type of concatenation called virtual concatenation that allows different routes to be used between the endpoints. It is similar to the semantic and scope of the LOAD-BALANCING in MPLS networks.
- (2) Label constraint: The PCE should be able to include Labels in the path returned to the PCC, the related object is the ERO object.
- (3) Roles of the routes: as defined in [RFC4872], this is applicable to the TE-LSP and is carried in association with the E2E path protection type.

1.4. Existing Support for GMPLS in Base PCEP Objects and its Limitations

The support provided by specifications in [RFC8282] and [RFC5440] for the requirements listed in [RFC7025] is summarized in Table 1 and Table 2. In some cases the support may not be complete, as noted, and additional support need to be provided in this specification.

Req.	Name	Support
1	Switching capability/type	SWITCH-LAYER (RFC8282)
2	Encoding type	SWITCH-LAYER (RFC8282)
3	Signal type	SWITCH-LAYER (RFC8282)
4	Concatenation type	No
5	Concatenation number	No
6	Technology-specific label	(Partial) ERO (RFC5440)
7	End-to-End (E2E) path protection type	No
8	Administrative group	LSPA (RFC5440)
9	Link protection type	No
10	Support for unnumbered interfaces	(Partial) ERO (RFC5440)
11	Support for asymmetric bandwidth requests	No
12	Support for explicit label control during the path computation	No
13	Support of label restrictions in the requests/responses	No

Table 1: RFC7025 Section 3.1 requirements support

Req.	Name	Support
1	Path computation with concatenation	No
2	Label constraint	No
3	Roles of the routes	No

Table 2: RFC7025 Section 3.2 requirements support

As described in Section 1.3 PCEP as of [RFC5440], [RFC5521] and [RFC8282], supports the following objects, included in requests and responses, related to the described requirements.

From [RFC5440]:

- o END-POINTS: related to requirements (1, 2, 3, 6, 10 and 13). The object only supports numbered endpoints. The context specifies whether they are node identifiers or numbered interfaces.
- o BANDWIDTH: related to requirements (4, 5 and 11). The data rate is encoded in the bandwidth object (as IEEE 32 bit float). [RFC5440] does not include the ability to convey an encoding proper to all GMPLS-controlled networks.

- o ERO: related to requirements (6, 10, 12 and 13). The ERO content is defined in RSVP in [RFC3209][RFC3473][RFC3477][RFC7570] and supports all the requirements already.
- o LSPA: related to requirements (7, 8 and 9). The requirement 8 (setup and holding priorities) is already supported.

From [RFC5521]:

- o XRO:
 - * This object allows excluding (strict or not) resources and is related to requirements (6, 10 and 13). It also includes the requested diversity (node, link or SRLG).
 - * When the F bit is set, the request indicates that the existing path has failed and the resources present in the RRO can be reused.

From [RFC8282]:

- o SWITCH-LAYER: addresses requirements (1, 2 and 3) for the TE-LSP and indicates which layer(s) should be considered. The object can be used to represent the RSVP-TE generalized label request. It does not address the endpoints case of requirements (1, 2 and 3).
- o REQ-ADAP-CAP: indicates the adaptation capabilities requested, can also be used for the endpoints in case of mono-layer computation

The gaps in functional coverage of the base PCEP objects are:

The BANDWIDTH and LOAD-BALANCING objects do not describe the details of the traffic request (requirements 4 and 5, for example NVC, multiplier) in the context of GMPLS networks, for instance TDM or OTN networks.

The END-POINTS object does not allow specifying an unnumbered interface, nor potential label restrictions on the interface (requirements 6, 10 and 13). Those parameters are of interest in case of switching constraints.

The Include/exclude Route Objects (IRO/XRO) do not allow the inclusion/exclusion of labels (requirements 6, 10 and 13).

Base attributes do not allow expressing the requested link protection level and/or the end-to-end protection attributes.

The PCEP extensions defined later in this document to cover the gaps are:

Two new object types are defined for the BANDWIDTH object (Generalized bandwidth, Generalized bandwidth of existing TE-LSP for which a reoptimization is requested).

A new object type is defined for the LOAD-BALANCING object (Generalized Load Balancing).

A new object type is defined for the END-POINTS object (Generalized Endpoint).

A new TLV is added to the Open message for capability negotiation.

A new TLV is added to the LSPA object.

The Label TLV is now allowed in the IRO and XRO objects.

In order to indicate the used routing granularity in the response, a new flag in the RP object is added.

2. PCEP Objects and Extensions

This section describes the necessary PCEP objects and extensions. The PCReq and PCRep messages are defined in [RFC5440]. This document does not change the existing grammars.

2.1. GMPLS Capability Advertisement

2.1.1. GMPLS Computation TLV in the Existing PCE Discovery Protocol

IGP-based PCE Discovery (PCED) is defined in [RFC5088] and [RFC5089] for the OSPF and IS-IS protocols. Those documents have defined bit 0 in PCE-CAP-FLAGS Sub-TLV of the PCED TLV as "Path computation with GMPLS link constraints". This capability is optional and can be used to detect GMPLS-capable PCEs. PCEs that set the bit to indicate support of GMPLS path computation MUST follow the procedures in Section 2.1.2 to further qualify the level of support during PCEP session establishment.

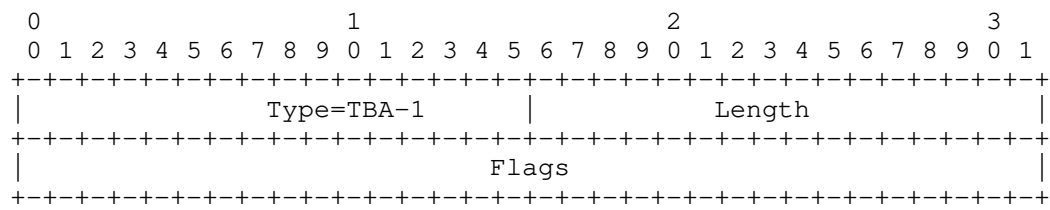
2.1.2. OPEN Object Extension GMPLS-CAPABILITY TLV

In addition to the IGP advertisement, a PCEP speaker MUST be able to discover the other peer GMPLS capabilities during the Open message exchange. This capability is also useful to avoid misconfigurations. This document defines a GMPLS-CAPABILITY TLV for use in the OPEN object to negotiate the GMPLS capability. The inclusion of this TLV

in the Open message indicates that the PCEP speaker support the PCEP extensions defined in the document. A PCEP speaker that is able to support the GMPLS extensions defined in this document MUST include the GMPLS-CAPABILITY TLV on the Open message. If one of the PCEP peers does not include the GMPLS-CAPABILITY TLV in the Open message, the peers MUST NOT make use of the objects and TLVs defined in this document.

If the PCEP speaker supports the extensions of this specification but did not advertise the GMPLS-CAPABILITY capability, upon receipt of a message from the PCE including an extension defined in this document, it MUST generate a PCEP Error (PCErr) with Error-Type=10 (Reception of an invalid object) and Error-value=TBA-42 (Missing GMPLS-CAPABILITY TLV), and it SHOULD terminate the PCEP session.

IANA has allocated value TBA-1 from the "PCEP TLV Type Indicators" sub-registry, as documented in Section 5.3 ("New PCEP TLVs"). The description is "GMPLS-CAPABILITY". Its format is shown in the following figure.



No Flags are defined in this document, they are reserved for future use.

2.2. RP Object Extension

Explicit label control (ELC) is a procedure supported by RSVP-TE, where the outgoing labels are encoded in the ERO. As a consequence, the PCE can provide such labels directly in the path ERO. Depending on policies or switching layer, it can be necessary for the PCC to use explicit label control or explicit link ids, thus it needs to indicate in the PCReq which granularity it is expecting in the ERO. This corresponds to requirement 12 of [RFC7025]. The possible granularities can be node, link or label. The granularities are inter-dependent, in the sense that link granularity implies the presence of node information in the ERO; similarly, a label granularity implies that the ERO contains node, link and label information.

A new 2-bit routing granularity (RG) flag (Bits TBA-13) is defined in the RP object. The values are defined as follows

0: reserved
1: node
2: link
3: label

Table 3: RG flag

The flag in the RP object indicates the requested route granularity. The PCE SHOULD follow this granularity and MAY return a NO-PATH if the requested granularity cannot be provided. The PCE MAY return any granularity on the route based on its policy. The PCC can decide if the ERO is acceptable based on its content.

If a PCE honored the requested routing granularity for a request, it MUST indicate the selected routing granularity in the RP object included in the response. Otherwise, the PCE MUST use the reserved RG to leave the check of the ERO to the PCC. The RG flag is backward-compatible with [RFC5440]: the value sent by an implementation (PCC or PCE) not supporting it will indicate a reserved value.

2.3. BANDWIDTH Object Extensions

From [RFC5440] the object carrying the requested size for the TE-LSP is the BANDWIDTH object. The object types 1 and 2 defined in [RFC5440] do not describe enough information to describe the TE-LSP bandwidth in GMPLS networks. The BANDWIDTH object encoding has to be extended to allow the object to express the bandwidth as described in [RFC7025]. RSVP-TE extensions for GMPLS provide a set of encodings allowing such representation in an unambiguous way, this is encoded in the RSVP-TE TSpec and FlowSpec objects. This document extends the BANDWIDTH object with new object types reusing the RSVP-TE encoding.

The following possibilities are supported by the extended encoding:

- o Asymmetric bandwidth (different bandwidth in forward and reverse direction), as described in [RFC6387]
- o GMPLS (SDH/SONET, G.709, ATM, MEF, etc.) parameters.

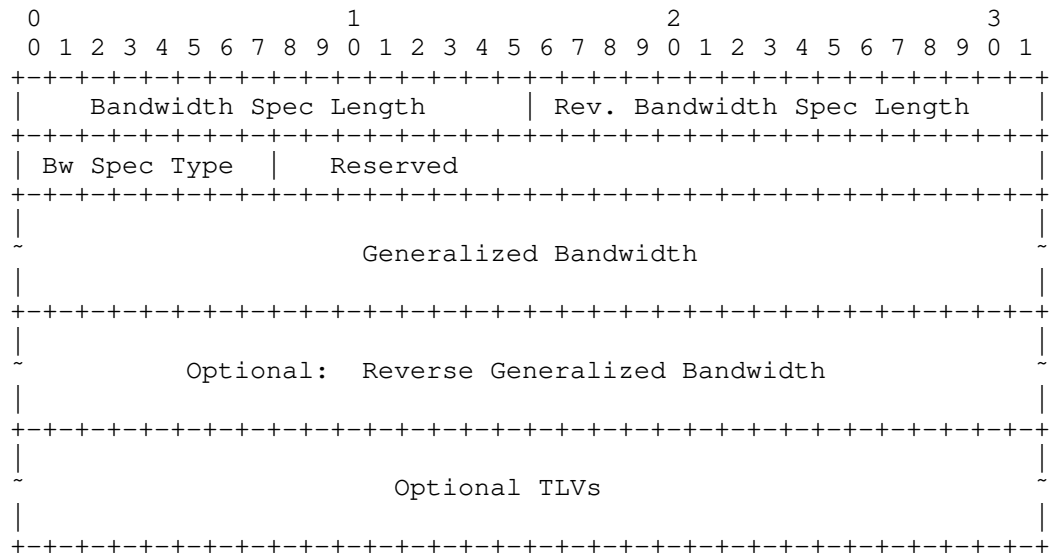
This corresponds to requirements 3, 4, 5 and 11 of [RFC7025] Section 3.1.

This document defines two Object Types for the BANDWIDTH object:

TBA-2 Generalized bandwidth

TBA-3 Generalized bandwidth of an existing TE-LSP for which a reoptimization is requested

The definitions below apply for Object Type TBA-2 and TBA-3. The body is as follows:



The BANDWIDTH object type TBA-2 and TBA-3 have a variable length. The 16-bit Bandwidth Spec Length field indicates the length of the Generalized Bandwidth field. The Bandwidth Spec Length MUST be strictly greater than 0. The 16-bit Reverse Bandwidth Spec Length field indicates the length of the Reverse Generalized Bandwidth field. The Reverse Bandwidth Spec Length MAY be equal to 0.

The Bw Spec Type field determines which type of bandwidth is represented by the object.

The Bw Spec Type corresponds to the RSVP-TE SENDER_TSPEC (Object Class 12) C-Types

The encoding of the fields Generalized Bandwidth and Reverse Generalized Bandwidth is the same as the Traffic Parameters carried in RSVP-TE, it can be found in the following references. It is to be noted that the RSVP-TE traffic specification MAY also include TLVs (e.g., [RFC6003] different from the PCEP TLVs).

Bw Spec	Type Name	Reference
2	Intserv	[RFC2210]
4	SONET/SDH	[RFC4606]
5	G.709	[RFC4328]
6	Ethernet	[RFC6003]
7	OTN-TDM	[RFC7139]
8	SSON	[RFC7792]

Table 4: Generalized Bandwidth and Reverse Generalized Bandwidth field encoding

When a PCC requests a bi-directional path with symmetric bandwidth, it SHOULD only specify the Generalized Bandwidth field, and set the Reverse Bandwidth Spec Length to 0. When a PCC needs to request a bi-directional path with asymmetric bandwidth, it SHOULD specify the different bandwidth in the forward and reverse directions with a Generalized Bandwidth and Reverse Generalized Bandwidth fields.

The procedure described in [RFC5440] for the PCRep is unchanged: a PCE MAY include the BANDWIDTH objects in the response to indicate the BANDWIDTH of the path.

As specified in [RFC5440] in the case of the reoptimization of a TE-LSP, the bandwidth of the existing TE-LSP MUST also be included in addition to the requested bandwidth if and only if the two values differ. The Object Type TBA-3 MAY be used instead of the previously specified object type 2 to indicate the existing TE-LSP bandwidth originally specified with object type TBA-2. A PCC that requested a path with a BANDWIDTH object of object type 1 MUST use object type 2 to represent the existing TE-LSP BANDWIDTH.

OPTIONAL TLVs MAY be included within the object body to specify more specific bandwidth requirements. No TLVs for the Object Type TBA-2 and TBA-3 are defined by this document.

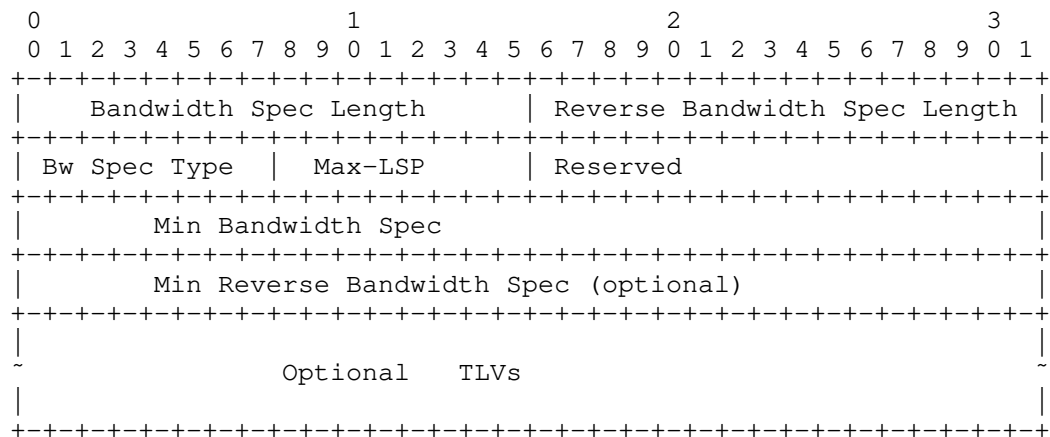
2.4. LOAD-BALANCING Object Extensions

The LOAD-BALANCING object [RFC5440] is used to request a set of at most Max-LSP TE-LSP having in total the bandwidth specified in BANDWIDTH, with each TE-LSP having at least a specified minimum bandwidth. The LOAD-BALANCING follows the bandwidth encoding of the BANDWIDTH object, and thus the existing definition from [RFC5440] does not describe enough details for the bandwidth specification expected by GMPLS.

Similarly to the BANDWIDTH object, a new object type is defined to allow a PCC to represent the bandwidth types supported by GMPLS networks.

This document defines the Generalized Load Balancing object type TBA-4 for the LOAD-BALANCING object. The Generalized Load Balancing object type has a variable length.

The format of the Generalized Load Balancing object type is as follows:



Bandwidth Spec Length (16 bits): the total length of the Min Bandwidth Spec field. The length MUST be strictly greater than 0.

Reverse Bandwidth Spec Length (16 bits): the total length of the Min Reverse Bandwidth Spec field. It MAY be equal to 0.

Bw Spec Type (8 bits): the bandwidth specification type, it corresponds to the RSVP-TE SENDER_TSPEC (Object Class 12) C-Types.

Max-LSP (8 bits): maximum number of TE-LSPs in the set.

Min Bandwidth Spec (variable): specifies the minimum bandwidth specification of each element of the TE-LSP set.

Min Reverse Bandwidth Spec (variable): specifies the minimum reverse bandwidth specification of each element of the TE-LSP set.

The encoding of the fields Min Bandwidth Spec and Min Reverse Bandwidth Spec is the same as in RSVP-TE SENDER_TSPEC object, it can be found in Table 4 from Section 2.3 from this document.

When a PCC requests a bi-directional path with symmetric bandwidth while specifying load balancing constraints it SHOULD specify the Min Bandwidth Spec field, and set the Reverse Bandwidth Spec Length to 0. When a PCC needs to request a bi-directional path with asymmetric bandwidth while specifying load balancing constraints, it MUST specify the different bandwidth in forward and reverse directions through a Min Bandwidth Spec and Min Reverse Bandwidth Spec fields.

OPTIONAL TLVs MAY be included within the object body to specify more specific bandwidth requirements. No TLVs for the Generalized Load Balancing object type are defined by this document.

The semantic of the LOAD-BALANCING object is not changed. If a PCC requests the computation of a set of TE-LSPs with at most N TE-LSPs so that it can carry generalized bandwidth X, each TE-LSP must at least transport bandwidth B, it inserts a BANDWIDTH object specifying X as the required bandwidth and a LOAD-BALANCING object with the Max-LSP and Min Bandwidth Spec fields set to N and B, respectively. When the BANDWIDTH and Min Bandwidth Spec can be summarized as scalars, the sum of all TE-LSPs bandwidth in the set is greater than X. The mapping of X over N path with (at least) bandwidth B is technology and possibly node specific. Each standard definition of the transport technology is defining those mappings and are not repeated in this document. A simplified example for SDH is described in Appendix A

In all other cases, including for technologies based on statistical multiplexing (e.g., InterServ, Ethernet), the exact bandwidth management (e.g., Ethernet's Excessive Rate) is left to the PCE's policies, according to the operator's configuration. If required, further documents may introduce a new mechanism to finely express complex load balancing policies within PCEP.

The BANDWIDTH and LOAD-BALANCING Bw Spec Type can be different depending on the endpoint nodes architecture. When the PCE is not able to handle those two Bw Spec Type, it MUST return a NO-PATH with the bit "LOAD-BALANCING could not be performed with the bandwidth constraints" set in the NO-PATH-VECTOR TLV.

2.5. END-POINTS Object Extensions

The END-POINTS object is used in a PCEP request message to specify the source and the destination of the path for which a path computation is requested. From [RFC5440], the source IP address and the destination IP address are used to identify those. A new Object Type is defined to address the following possibilities:

- o Different source and destination endpoint types.

- o Label restrictions on the endpoint.
- o Specification of unnumbered endpoints type as seen in GMPLS networks.

The Object encoding is described in the following sections.

In path computation within a GMPLS context the endpoints can:

- o Be unnumbered as described in [RFC3477].
- o Have labels associated to them, specifying a set of constraints on the allocation of labels.
- o Have different switching capabilities

The IPv4 and IPv6 endpoints are used to represent the source and destination IP addresses. The scope of the IP address (Node or numbered Link) is not explicitly stated. It is also possible to request a Path between a numbered link and an unnumbered link, or a P2MP path between different type of endpoints.

This document defines the Generalized Endpoint object type TBA-5 for the END-POINTS object. This new type also supports the specification of constraints on the endpoint label to be used. The PCE might know the interface restrictions but this is not a requirement. This corresponds to requirements 6 and 10 of [RFC7025].

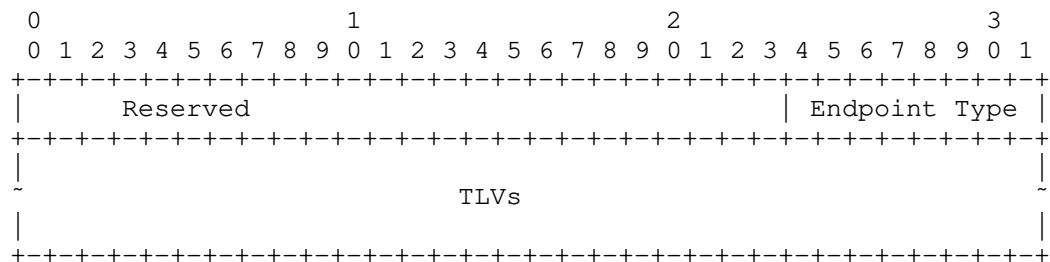
2.5.1. Generalized Endpoint Object Type

The Generalized Endpoint object type format consists of a body and a list of TLVs scoped to this object. The TLVs give the details of the endpoints and are described in Section 2.5.2. For each Endpoint Type, a different grammar is defined. The TLVs defined to describe an endpoint are:

1. IPv4 address endpoint.
2. IPv6 address endpoint.
3. Unnumbered endpoint.
4. Label request.
5. Label set.

The Label set TLV is used to restrict or suggest the label allocation in the PCE. This TLV expresses the set of restrictions which may

apply to signaling. Label restriction support can be an explicit or a suggested value (Label set describing one label, with the L bit respectively cleared or set), mandatory range restrictions (Label set with L bit cleared) and optional range restriction (Label set with L bit set). Endpoints label restriction may not be part of the RRO or IRO. They can be included when following [RFC4003] in signaling for egress endpoint, but ingress endpoint properties can be local to the PCC and not signaled. To support this case the label set allows indication which label are used in case of reoptimization. The label range restrictions are valid in GMPLS-controlled networks, either by PCC policy or depending on the switching technology used, for instance on given Ethernet or ODU equipment having limited hardware capabilities restricting the label range. Label set restriction also applies to WSON networks where the optical senders and receivers are limited in their frequency tunability ranges, consequently restricting the possible label ranges on the interface in GMPLS. The END-POINTS Object with Generalized Endpoint object type is encoded as follow:



Reserved bits SHOULD be set to 0 when a message is sent and ignored when the message is received.

The Endpoint Type is defined as follow:

Value	Type	Meaning
0	Point-to-Point	
1	Point-to-Multipoint	New leaves to add
2		Old leaves to remove
3		Old leaves whose path can be modified/reoptimized
4		Old leaves whose path has to be left unchanged
5-244	Reserved	
245-255	Experimental range	

Table 5: Generalized Endpoint endpoint types

The Endpoint Type is used to cover both point-to-point and different point-to-multipoint endpoints. A PCE may accept only Endpoint Type 0: Endpoint Types 1-4 apply if the PCE implementation supports P2MP path calculation. A PCE not supporting a given Endpoint Type SHOULD respond with a PCErr with Error-Type=4 (Not supported object), Error-value=TBA-15 (Unsupported endpoint type in END-POINTS Generalized Endpoint object type). As per [RFC5440], a PCE unable to process Generalized Endpoints may respond with Error-Type=3 (Unknown Object), Error-value=2 (Unrecognized object Type) or Error-Type=4 (Not supported object), Error-value=2 (Not supported object Type). The TLVs present in the request object body MUST follow the following [RFC5511] grammar:

```
<generalized-endpoint-tlvs> ::=
  <p2p-endpoints> | <p2mp-endpoints>

<p2p-endpoints> ::=
  <endpoint> [<endpoint-restriction-list>]
  <endpoint> [<endpoint-restriction-list>]

<p2mp-endpoints> ::=
  <endpoint> [<endpoint-restriction-list>]
  <endpoint> [<endpoint-restriction-list>]
  [<endpoint> [<endpoint-restriction-list>]]...
```

For endpoint type Point-to-Point, 2 endpoint TLVs MUST be present in the message. The first endpoint is the source and the second is the destination.

For endpoint type Point-to-Multipoint, several END-POINT objects MAY be present in the message and the exact meaning depending on the endpoint type defined for the object. The first endpoint TLV is the root and other endpoints TLVs are the leaves. The root endpoint MUST be the same for all END-POINTS objects for that P2MP tree request. If the root endpoint is not the same for all END-POINTS, a PCErr with Error-Type=17 (P2MP END-POINTS Error), Error-value=4 (The PCE cannot satisfy the request due to inconsistent END-POINTS) MUST be returned. The procedure defined in [RFC8306] Section 3.10 also apply to the Generalized Endpoint with Point-to-Multipoint endpoint types.

An endpoint is defined as follows:

```

<endpoint>::=<IPV4-ADDRESS>|<IPV6-ADDRESS>|<UNNUMBERED-ENDPOINT>
<endpoint-restriction-list> ::=                                <endpoint-restriction>
                                [<endpoint-restriction-list>]

<endpoint-restriction> ::=
                                [<LABEL-REQUEST>][<label-restriction-list>]

<label-restriction-list> ::= <label-restriction>
                                [<label-restriction-list>]
<label-restriction> ::= <LABEL-SET>

```

The different TLVs are described in the following sections. A PCE MAY support any or all of IPV4-ADDRESS, IPV6-ADDRESS, and UNNUMBERED-ENDPOINT TLVs. When receiving a PCReq, a PCE unable to resolve the identifier in one of those TLVs MUST respond using a PCRep with NO-PATH and set the bit "Unknown destination" or "Unknown source" in the NO-PATH-VECTOR TLV. The response SHOULD include the END-POINTS object with only the unsupported TLV(s).

A PCE MAY support either or both of the LABEL-REQUEST and LABEL-SET TLVs. If a PCE finds a non-supported TLV in the END-POINTS the PCE MUST respond with a PCErr message with Error-Type=4 (Not supported object) and Error-value=TBA-15 (Unsupported TLV present in END-POINTS Generalized Endpoint object type) and the message SHOULD include the END-POINTS object in the response with only the endpoint and endpoint restriction TLV it did not understand. A PCE supporting those TLVs but not being able to fulfil the label restriction MUST send a response with a NO-PATH object which has the bit "No endpoint label resource" or "No endpoint label resource in range" set in the NO-PATH-VECTOR TLV. The response SHOULD include an END-POINTS object containing only the TLV(s) related to the constraints the PCE could not meet.

2.5.2. END-POINTS TLV Extensions

All endpoint TLVs have the standard PCEP TLV header as defined in [RFC5440] Section 7.1. For the Generalized Endpoint Object Type the TLVs MUST follow the ordering defined in Section 2.5.1.

2.5.2.1. IPV4-ADDRESS TLV

This TLV represents a numbered endpoint using IPv4 numbering, the format of the IPV4-ADDRESS TLV value (TLV-Type=TBA-6) is 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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     IPv4 address                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This TLV MAY be ignored, in which case a PCRep with NO-PATH SHOULD be returned, as described in Section 2.5.1.

2.5.2.2. IPV6-ADDRESS TLV

This TLV represents a numbered endpoint using IPV6 numbering, the format of the IPV6-ADDRESS TLV value (TLV-Type=TBA-7) is 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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     IPv6 address (16 bytes)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This TLV MAY be ignored, in which case a PCRep with NO-PATH SHOULD be returned, as described in Section 2.5.1.

2.5.2.3. UNNUMBERED-ENDPOINT TLV

This TLV represents an unnumbered interface. This TLV has the same semantic as in [RFC3477]. The TLV value is encoded as follows (TLV-Type=TBA-8)

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     LSR's Router ID                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Interface ID (32 bits)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

This TLV MAY be ignored, in which case a PCRep with NO-PATH SHOULD be returned, as described in Section 2.5.1.

2.5.2.4. LABEL-REQUEST TLV

The LABEL-REQUEST TLV indicates the switching capability and encoding type of the following label restriction list for the endpoint. The value format and encoding is the same as described in [RFC3471]

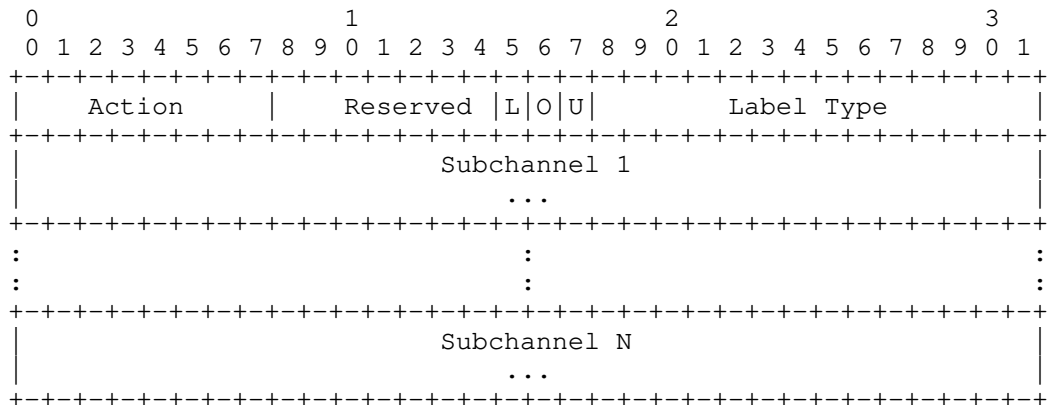
Section 3.1 Generalized label request. The LABEL-REQUEST TLV uses TLV-Type=TBA-9. The Encoding Type indicates the encoding type, e.g., SONET/SDH/GigE etc., of the LSP with which the data is associated. The Switching type indicates the type of switching that is being requested on the endpoint. G-PID identifies the payload. This TLV and the following one are defined to satisfy requirement 13 of [RFC7025] for the endpoint. It is not directly related to the TE-LSP label request, which is expressed by the SWITCH-LAYER object.

On the path calculation request only the GENERALIZED-BANDWIDTH and SWITCH-LAYER need to be coherent, the endpoint labels could be different (supporting a different LABEL-REQUEST). Hence the label restrictions include a Generalized label request in order to interpret the labels. This TLV MAY be ignored, in which case a PCRep with NO-PATH SHOULD be returned, as described in Section 2.5.1.

2.5.2.5. LABEL-SET TLV

Label or label range restrictions can be specified for the TE-LSP endpoints. Those are encoded using the LABEL-SET TLV. The label value need to be interpreted with a description on the Encoding and switching type. The REQ-ADAP-CAP object from [RFC8282] can be used in case of mono-layer request, however in case of multilayer it is possible to have more than one object, so it is better to have a dedicated TLV for the label and label request. These TLVs MAY be ignored, in which case a response with NO-PATH SHOULD be returned, as described in Section 2.5.1. TLVs are encoded as follows (following [RFC5440]):

- o LABEL-SET TLV, Type=TBA-10. The TLV Length is variable, Encoding follows [RFC3471] Section 3.5 "Label set" with the addition of a U bit, O bit and L bit. The L bit is used to represent a suggested set of labels, following the semantic of SUGGESTED_LABEL defined by [RFC3471].



A LABEL-SET TLV represents a set of possible labels that can be used on an interface. If the L bit is cleared, the label allocated on the first endpoint **MUST** be within the label set range. The action parameter in the Label set indicates the type of list provided. These parameters are described by [RFC3471] Section 3.5.1.

The U, O and L bits have the following meaning:

- U: Upstream direction: The U bit is set for upstream (revers) direction in case of bidirectional LSP.
- O: Old Label: set when the TLV represent the old (previously allocated) label in case of re-optimization. The R bit of the RP object **MUST** be set to 1. If the L bit is set, this bit **SHOULD** be set to 0 and ignored on receipt. When this bit is set, the Action field **MUST** be set to 0 (Inclusive List) and the Label Set **MUST** contain one subchannel.
- L: Loose Label: set when the TLV indicates to the PCE a set of preferred (ordered) labels to be used. The PCE **MAY** use those labels for label allocation.

Labels TLV bits

Several LABEL_SET TLVs **MAY** be present with the O bit cleared, LABEL_SET TLVs with L bit set can be combined with a LABEL_SET TLV with L bit cleared. There **MUST NOT** be more than two LABEL_SET TLVs present with the O bit set. If there are two LABEL_SET TLVs present, there **MUST NOT** be more than one with the U bit set, and there **MUST NOT** be more than one with the U bit cleared. For a given U bit value, if more than one LABEL_SET TLV with the O bit set is present, the first TLV **MUST** be processed and the following TLVs with the same U and O bit **MUST** be ignored.

A LABEL-SET TLV with the O and L bit set MUST trigger a PCErr message with Error-Type=10 (Reception of an invalid object) Error-value=TBA-25 (Wrong LABEL-SET TLV present with O and L bit set).

A LABEL-SET TLV with the O bit set and an Action Field not set to 0 (Inclusive list) or containing more than one subchannel MUST trigger a PCErr message with Error-Type=10 (Reception of an invalid object) Error-value=TBA-26 (Wrong LABEL-SET TLV present with O bit and wrong format).

If a LABEL-SET TLV is present with O bit set, the R bit of the RP object MUST be set, otherwise a PCErr message MUST be sent with Error-Type=10 (Reception of an invalid object) Error-value=TBA-24 (LABEL-SET TLV present with O bit set but without R bit set in RP).

2.6. IRO Extension

The IRO as defined in [RFC5440] is used to include specific objects in the path. RSVP-TE allows the inclusion of a label definition. In order to fulfill requirement 13 of [RFC7025] the IRO needs to support the new subobject type as defined in [RFC3473]:

Type	Sub-object
TBA-38	LABEL

The Label subobject MUST follow a subobject identifying a link, currently an IP address subobject (Type 1 or 2) or an interface ID (type 4) subobject. If an IP address subobject is used, then the given IP address MUST be associated with a link. More than one label subobject MAY follow each link subobject. The procedure associated with this subobject is as follows.

If the PCE is able to allocate labels (e.g., via explicit label control) the PCE MUST allocate one label from within the set of label values for the given link. If the PCE does not assign labels, then it sends a response with a NO-PATH object, containing a NO-PATH-VECTOR TLV with the bit 'No label resource in range' set.

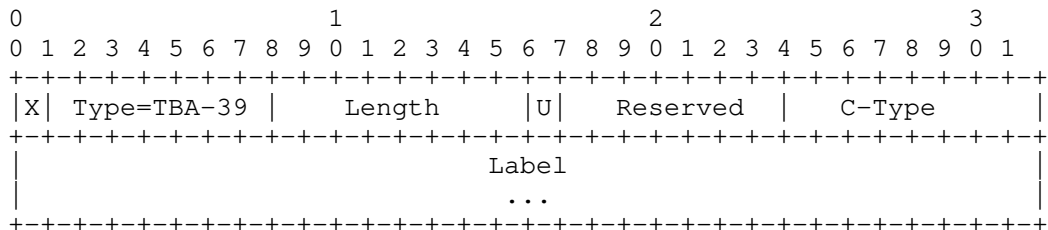
2.7. XRO Extension

The XRO as defined in [RFC5521] is used to exclude specific objects in the path. RSVP-TE allows the exclusion of certain labels ([RFC6001]). In order to fulfill requirement 13 of [RFC7025] Section 3.1, the PCEP's XRO needs to support a new subobject to enable label exclusion.

The encoding of the XRO Label subobject follows the encoding of the Label ERO subobject defined in [RFC3473] and XRO subobject defined in

[RFC5521]. The XRO Label subobject represent one Label and is defined as follows:

XRO Subobject Type TBA-39: Label Subobject.



X (1 bit): as per [RFC5521]. The X-bit indicates whether the exclusion is mandatory or desired. 0 indicates that the resource specified MUST be excluded from the path computed by the PCE. 1 indicates that the resource specified SHOULD be excluded from the path computed by the PCE, but MAY be included subject to PCE policy and the absence of a viable path that meets the other constraints and excludes the resource.

Type (7 bits): The Type of the XRO Label subobject is TBA-39.

Length (8 bits): see [RFC5521], the total length of the subobject in bytes (including the Type and Length fields). The Length is always divisible by 4.

U (1 bit): see [RFC3471] Section 6.1.

C-Type (8 bits): the C-Type of the included Label Object as defined in [RFC3473].

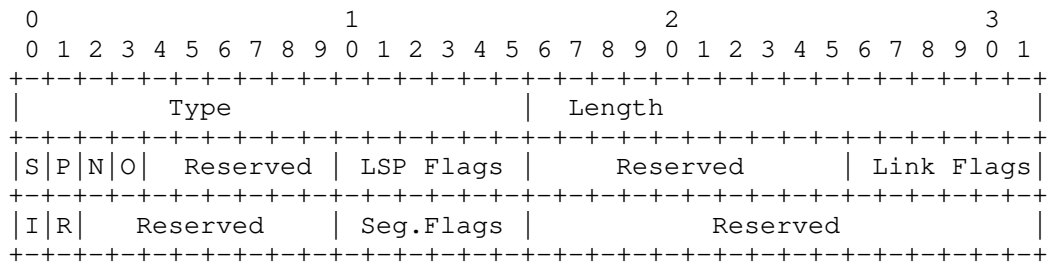
Label: see [RFC3471].

The Label subobject MUST follow a subobject identifying a link, currently an IP address subobject (Type 1 or 2) or an interface ID (type 4) subobject. If an IP address subobject is used, then the given IP address MUST be associated with a link. More than one label subobject MAY follow each link subobject.

Type Sub-object
3 LABEL

2.8. LSPA Extensions

The LSPA carries the LSP attributes. In the end-to-end recovery context, this also includes the protection state information. A new TLV is defined to fulfil requirement 7 of [RFC7025] Section 3.1 and requirement 3 of [RFC7025] Section 3.2. This TLV contains the information of the PROTECTION object defined by [RFC4872] and can be used as a policy input. The LSPA object MAY carry a PROTECTION-ATTRIBUTE TLV defined as: Type TBA-12: PROTECTION-ATTRIBUTE



The content is as defined in [RFC4872] Section 14, [RFC4873] Section 6.1.

LSP (protection) Flags or Link flags field can be used by a PCE implementation for routing policy input. The other attributes are only meaningful for a stateful PCE.

This TLV is OPTIONAL and MAY be ignored by the PCE. If ignored by the PCE, it MUST NOT include the TLV in the LSPA of the response. When the TLV is used by the PCE, a LSPA object and the PROTECTION-ATTRIBUTE TLV MUST be included in the response. Fields that were not considered MUST be set to 0.

2.9. NO-PATH Object Extension

The NO-PATH object is used in PCRep messages in response to an unsuccessful path computation request (the PCE could not find a path satisfying the set of constraints). In this scenario, PCE MUST include a NO-PATH object in the PCRep message. The NO-PATH object MAY carry the NO-PATH-VECTOR TLV that specifies more information on the reasons that led to a negative reply. In case of GMPLS networks there could be some additional constraints that led to the failure such as protection mismatch, lack of resources, and so on. Several new flags have been defined in the 32-bit flag field of the NO-PATH-VECTOR TLV but no modifications have been made in the NO-PATH object.

2.9.1. Extensions to NO-PATH-VECTOR TLV

The modified NO-PATH-VECTOR TLV carrying the additional information is as follows:

Bit number TBA-32 - Protection Mismatch (1-bit). Specifies the mismatch of the protection type in the PROTECTION-ATTRIBUTE TLV in the request.

Bit number TBA-33 - No Resource (1-bit). Specifies that the resources are not currently sufficient to provide the path.

Bit number TBA-34 - Granularity not supported (1-bit). Specifies that the PCE is not able to provide a path with the requested granularity.

Bit number TBA-35 - No endpoint label resource (1-bit). Specifies that the PCE is not able to provide a path because of the endpoint label restriction.

Bit number TBA-36 - No endpoint label resource in range (1-bit). Specifies that the PCE is not able to provide a path because of the endpoint label set restriction.

Bit number TBA-37 - No label resource in range (1-bit). Specifies that the PCE is not able to provide a path because of the label set restriction.

3. Additional Error-Types and Error-Values Defined

A PCEP-ERROR object is used to report a PCEP error and is characterized by an Error-Type that specifies the type of error while Error-value that provides additional information about the error. An additional error type and several error values are defined to represent some of the errors related to the newly identified objects related to GMPLS networks. For each PCEP error, an Error-Type and an Error-value are defined. Error-Type 1 to 10 are already defined in [RFC5440]. Additional Error-values are defined for Error-Types 4 and 10. A new Error-Type is defined (value TBA-27).

The Error-Type TBA-27 (path computation failure) is used to reflect constraints not understood by the PCE, for instance when the PCE is not able to understand the generalized bandwidth. If the constraints are understood, but the PCE is unable to find with those constraints, the NO-PATH is to be used.

Error-Type Error-value

4	Not supported object
	value=TBA-14: Bandwidth Object type TBA-2 or TBA-3 not supported
	value=TBA-15: Unsupported endpoint type in END-POINTS Generalized Endpoint object type
	value=TBA-16: Unsupported TLV present in END-POINTS Generalized Endpoint object type
	value=TBA-17: Unsupported granularity in the RP object flags
10	Reception of an invalid object
	value=TBA-18: Bad Bandwidth Object type TBA-2 (Generalized bandwidth) or TBA-3 (Generalized bandwidth of existing TE-LSP for which a reoptimization is requested)
	value=TBA-20: Unsupported LSP Protection Flags in PROTECTION-ATTRIBUTE TLV
	value=TBA-21: Unsupported Secondary LSP Protection Flags in PROTECTION-ATTRIBUTE TLV
	value=TBA-22: Unsupported Link Protection Type in PROTECTION-ATTRIBUTE TLV
	value=TBA-24: LABEL-SET TLV present with 0 bit set but without R bit set in RP
	value=TBA-25: Wrong LABEL-SET TLV present with 0 and L bit set
	value=TBA-26: Wrong LABEL-SET with 0 bit set and wrong format
	value=TBA-42: Missing GMPLS-CAPABILITY TLV
TBA-27	Path computation failure
	value=0: Unassigned
	value=TBA-28: Unacceptable request message
	value=TBA-29: Generalized bandwidth value not supported
	value=TBA-30: Label Set constraint could not be met
	value=TBA-31: Label constraint could not be met

4. Manageability Considerations

This section follows the guidance of [RFC6123].

4.1. Control of Function through Configuration and Policy

This document makes no change to the basic operation of PCEP and so the requirements described in [RFC5440] Section 8.1. also apply to this document. In addition to those requirements a PCEP implementation may allow the configuration of the following parameters:

- Accepted RG in the RP object.

- Default RG to use (overriding the one present in the PCReq)

- Accepted BANDWIDTH object type TBA-2 and TBA-3 parameters in request, default mapping to use when not specified in the request

- Accepted LOAD-BALANCING object type TBA-4 parameters in request.

- Accepted endpoint type and allowed TLVs in object END-POINTS with object type Generalized Endpoint.

- Accepted range for label restrictions in label restriction in END-POINTS, or IRO or XRO objects

- PROTECTION-ATTRIBUTE TLV acceptance and suppression.

The configuration of the above parameters is applicable to the different sessions as described in [RFC5440] Section 8.1 (by default, per PCEP peer, etc.).

4.2. Information and Data Models

This document makes no change to the basic operation of PCEP and so the requirements described in [RFC5440] Section 8.2. also apply to this document. This document does not introduce any new ERO sub objects, so that the, ERO information model is already covered in [RFC4802].

4.3. Liveness Detection and Monitoring

This document makes no change to the basic operation of PCEP and so there are no changes to the requirements for liveness detection and monitoring set out in [RFC4657] and [RFC5440] Section 8.3.

4.4. Verifying Correct Operation

This document makes no change to the basic operations of PCEP and considerations described in [RFC5440] Section 8.4. New errors defined by this document should satisfy the requirement to log error events.

4.5. Requirements on Other Protocols and Functional Components

No new Requirements on Other Protocols and Functional Components are made by this document. This document does not require ERO object extensions. Any new ERO subobject defined in the TEAS or CCAMP working group can be adopted without modifying the operations defined in this document.

4.6. Impact on Network Operation

This document makes no change to the basic operations of PCEP and considerations described in [RFC5440] Section 8.6. In addition to the limit on the rate of messages sent by a PCEP speaker, a limit MAY be placed on the size of the PCEP messages.

5. IANA Considerations

IANA assigns values to the PCEP objects and TLVs. IANA is requested to make some allocations for the newly defined objects and TLVs defined in this document. Also, IANA is requested to manage the space of flags that are newly added in the TLVs.

5.1. PCEP Objects

As described in Section 2.3, Section 2.4 and Section 2.5.1 new Objects types are defined. IANA is requested to make the following Object-Type allocations from the "PCEP Objects" sub-registry.

Object 5
Class
Name BANDWIDTH
Object-Type TBA-2: Generalized bandwidth
TBA-3: Generalized bandwidth of an existing TE-LSP for
which a reoptimization is requested
Reference This document (Section 2.3)

Object 14
Class
Name LOAD-BALANCING
Object-Type TBA-4: Generalized Load Balancing

Reference This document (Section 2.4)

Object 4
Class
Name END-POINTS
Object-Type TBA-5: Generalized Endpoint
Reference This document (Section 2.5)

5.2. Endpoint type field in Generalized END-POINTS Object

IANA is requested to create a registry to manage the Endpoint Type field of the END-POINTS object, Object Type Generalized Endpoint and manage the code space.

New endpoint type in the Reserved range are assigned by Standards Action [RFC8126]. Each endpoint type should be tracked with the following attributes:

- o Endpoint type
- o Description
- o Defining RFC

New endpoint type in the Experimental range are for experimental use; these will not be registered with IANA and MUST NOT be mentioned by RFCs.

The following values have been defined by this document.
(Section 2.5.1, Table 5):

Value	Type	Meaning
0	Point-to-Point	
1	Point-to-Multipoint	New leaves to add
2		Old leaves to remove
3		Old leaves whose path can be modified/reoptimized
4		Old leaves whose path has to be left unchanged
5-244	Unassigned	
245-255	Experimental range	

5.3. New PCEP TLVs

IANA manages the PCEP TLV code point registry (see [RFC5440]). This is maintained as the "PCEP TLV Type Indicators" sub-registry of the "Path Computation Element Protocol (PCEP) Numbers" registry. IANA is requested to do the following allocation. Note: TBA-11 is not used

Value	Meaning	Reference
TBA-6	IPV4-ADDRESS	This document (Section 2.5.2.1)
TBA-7	IPV6-ADDRESS	This document (Section 2.5.2.2)
TBA-8	UNNUMBERED-ENDPOINT	This document (Section 2.5.2.3)
TBA-9	LABEL-REQUEST	This document (Section 2.5.2.4)
TBA-10	LABEL-SET	This document (Section 2.5.2.5)
TBA-12	PROTECTION-ATTRIBUTE	This document (Section 2.8)
TBA-1	GMPLS-CAPABILITY	This document (Section 2.1.2)

5.4. RP Object Flag Field

As described in Section 2.2 new flag are defined in the RP Object Flag IANA is requested to make the following Object-Type allocations from the "RP Object Flag Field" sub-registry.

Bit	Description	Reference
TBA-13	routing granularity (2 bits) (RG)	This document, Section 2.2

5.5. New PCEP Error Codes

As described in Section 3, new PCEP Error-Types and Error-values are defined. IANA is requested to make the following allocation in the "PCEP-ERROR Object Error Types and Values" registry.

Error	name	Reference
Type=4	Not supported object	[RFC5440]
Value=TBA-14:	Bandwidth Object type TBA-2 or TBA-3 not supported	This Document
Value=TBA-15:	Unsupported endpoint type in END-POINTS Generalized Endpoint object type	This Document
Value=TBA-16:	Unsupported TLV present in END-POINTS Generalized Endpoint object type	This Document
Value=TBA-17:	Unsupported granularity in the RP object flags	This Document
Type=10	Reception of an invalid object	[RFC5440]
Value=TBA-18:	Bad Bandwidth Object type TBA-2 (Generalized bandwidth) or TBA-3 (Generalized bandwidth of existing TE-LSP for which a reoptimization is requested)	This Document
Value=TBA-20:	Unsupported LSP Protection Flags in PROTECTION-ATTRIBUTE TLV	This Document
Value=TBA-21:	Unsupported Secondary LSP Protection Flags in PROTECTION-ATTRIBUTE TLV	This Document
Value=TBA-22:	Unsupported Link Protection Type in PROTECTION-ATTRIBUTE TLV	This Document
Value=TBA-24:	LABEL-SET TLV present with 0 bit set but without R bit set in RP	This Document
Value=TBA-25:	Wrong LABEL-SET TLV present with 0 and L bit set	This Document
Value=TBA-26:	Wrong LABEL-SET with 0 bit set and wrong format	This Document
Value=TBA-42:	Missing GMPLS-CAPABILITY TLV	This Document
Type=TBA-27	Path computation failure	This Document
Value=0	Unassigned	This Document
Value=TBA-28:	Unacceptable request message	This Document
Value=TBA-29:	Generalized bandwidth value not supported	This Document
Value=TBA-30:	Label Set constraint could not be met	This Document
Value=TBA-31:	Label constraint could not be met	This Document

5.6. New NO-PATH-VECTOR TLV Fields

As described in Section 2.9.1, new NO-PATH-VECTOR TLV Flag Fields have been defined. IANA is requested to do the following allocations in the "NO-PATH-VECTOR TLV Flag Field" sub-registry.

Bit number TBA-32 - Protection Mismatch (1-bit). Specifies the mismatch of the protection type of the PROTECTION-ATTRIBUTE TLV in the request.

Bit number TBA-33 - No Resource (1-bit). Specifies that the resources are not currently sufficient to provide the path.

Bit number TBA-34 - Granularity not supported (1-bit). Specifies that the PCE is not able to provide a path with the requested granularity.

Bit number TBA-35 - No endpoint label resource (1-bit). Specifies that the PCE is not able to provide a path because of the endpoint label restriction.

Bit number TBA-36 - No endpoint label resource in range (1-bit). Specifies that the PCE is not able to provide a path because of the endpoint label set restriction.

Bit number TBA-37 - No label resource in range (1-bit). Specifies that the PCE is not able to provide a path because of the label set restriction.

Bit number TBA-40 - LOAD-BALANCING could not be performed with the bandwidth constraints (1 bit). Specifies that the PCE is not able to provide a path because it could not map the BANDWIDTH into the parameters specified by the LOAD-BALANCING.

5.7. New Subobject for the Include Route Object

The "PCEP Parameters" registry contains a subregistry "IRO Subobjects" with an entry for the Include Route Object (IRO).

IANA is requested to add a further subobject that can be carried in the IRO as follows:

Subobject type	Reference
TBA-38 Label subobject	This Document

5.8. New Subobject for the Exclude Route Object

The "PCEP Parameters" registry contains a subregistry "XRO Subobjects" with an entry for the XRO object (Exclude Route Object).

IANA is requested to add a further subobject that can be carried in the XRO as follows:

Subobject type	Reference
TBA-39 Label subobject	This Document

5.9. New GMPLS-CAPABILITY TLV Flag Field

IANA is requested to create a sub-registry to manage the Flag field of the GMPLS-CAPABILITY TLV within the "Path Computation Element Protocol (PCEP) Numbers" registry.

New bit numbers are to be assigned by Standards Action [RFC8126]. Each bit should be tracked with the following qualities:

- o Bit number (counting from bit 0 as the most significant bit)
- o Capability description
- o Defining RFC

The initial contents of the sub-registry are empty, with all bits marked unassigned

6. Security Considerations

GMPLS controls multiple technologies and types of network elements. The LSPs that are established using GMPLS, whose paths can be computed using the PCEP extensions to support GMPLS described in this document, can carry a high volume of traffic and can be a critical part of a network infrastructure. The PCE can then play a key role in the use of the resources and in determining the physical paths of the LSPs and thus it is important to ensure the identity of PCE and PCC, as well as the communication channel. In many deployments there will be a completely isolated network where an external attack is of very low probability. However, there are other deployment cases in which the PCC-PCE communication can be more exposed and there could be more security considerations. Three main situations in case of an attack in the GMPLS PCE context could happen:

- o PCE Identity theft: A legitimate PCC could request a path for a GMPLS LSP to a malicious PCE, which poses as a legitimate PCE. The answer can make that the LSP traverses some geographical place known to the attacker where confidentiality (sniffing), integrity (traffic modification) or availability (traffic drop) attacks could be performed by use of an attacker-controlled middlebox device. Also, the resulting LSP can omit constraints given in the requests (e.g., excluding certain fibers, avoiding some SRLGs) which could make that the LSP which will be later set-up can look perfectly fine, but will be in a risky situation. Also, the result can lead to the creation of an LSP that does not provide the desired quality and gives less resources than necessary.

- o PCC Identity theft: A malicious PCC, acting as a legitimate PCC, requesting LSP paths to a legitimate PCE can obtain a good knowledge of the physical topology of a critical infrastructure. It could get to know enough details to plan a later physical attack.
- o Message inspection: As in the previous case, knowledge of an infrastructure can be obtained by sniffing PCEP messages.

The security mechanisms can provide authentication and confidentiality for those scenarios where the PCC-PCE communication cannot be completely trusted. [RFC8253] provides origin verification, message integrity and replay protection, and ensures that a third party cannot decipher the contents of a message.

In order to protect against the malicious PCE case the PCC SHOULD have policies in place to accept or not the path provided by the PCE. Those policies can verify if the path follows the provided constraints. In addition, technology specific data plane mechanism can be used (following [RFC5920] Section 5.8) to verify the data plane connectivity and deviation from constraints.

The document [RFC8253] describes the usage of Transport Layer Security (TLS) to enhance PCEP security. The document describes the initiation of the TLS procedures, the TLS handshake mechanisms, the TLS methods for peer authentication, the applicable TLS ciphersuites for data exchange, and the handling of errors in the security checks. PCE and PCC SHOULD use [RFC8253] mechanism to protect against malicious PCC and PCE.

Finally, as mentioned by [RFC7025] the PCEP extensions to support GMPLS should be considered under the same security as current PCE work and this extension will not change the underlying security issues. However, given the critical nature of the network infrastructures under control by GMPLS, the security issues described above should be seriously considered when deploying a GMPLS-PCE based control plane for such networks. For more information on the security considerations on a GMPLS control plane, not only related to PCE/PCEP, [RFC5920] provides an overview of security vulnerabilities of a GMPLS control plane.

7. Contributing Authors

Elie Sfeir
Coriant
St Martin Strasse 76
Munich, 81541
Germany

Email: elie.sfeir@coriant.com

Franz Rambach
Nockherstrasse 2-4,
Munich 81541
Germany

Phone: +49 178 8855738
Email: franz.rambach@cgi.com

Francisco Javier Jimenez Chico
Telefonica Investigacion y Desarrollo
C/ Emilio Vargas 6
Madrid, 28043
Spain

Phone: +34 91 3379037
Email: fjjc@tid.es

Huawei Technologies

Suresh BR
Shenzhen
China
Email: sureshbr@huawei.com

Young Lee
1700 Alma Drive, Suite 100
Plano, TX 75075
USA

Phone: (972) 509-5599 (x2240)
Email: ylee@huawei.com

SenthilKumar S
Shenzhen
China
Email: senthilkumars@huawei.com

Jun Sun
Shenzhen
China
Email: johnsun@huawei.com

CTTC - Centre Tecnologic de Telecomunicacions de Catalunya

Ramon Casellas
PMT Ed B4 Av. Carl Friedrich Gauss 7
08860 Castelldefels (Barcelona)
Spain
Phone: (34) 936452916
Email: ramon.casellas@cttc.es

8. Acknowledgments

The research of Ramon Casellas, Francisco Javier Jimenez Chico, Oscar Gonzalez de Dios, Cyril Margaria, and Franz Rambach leading to these results has received funding from the European Community's Seventh Framework Program FP7/2007-2013 under grant agreement no 247674 and no 317999.

The authors would like to thank Julien Meuric, Lyndon Ong, Giada Lander, Jonathan Hardwick, Diego Lopez, David Sinicrope, Vincent Roca, Dhruv Dhody, Adrian Farrel and Tianran Zhou for their review and useful comments to the document.

Thanks to Alisa Cooper, Benjamin Kaduk, Elwun-davies, Martin Vigoureux, Roman Danyliw, and Suresh Krishnan for the IESG comments

9. References

9.1. Normative References

- [G.709-v3] ITU-T, "Interfaces for the optical transport network, Recommendation G.709/Y.1331", June 2016, <<https://www.itu.int/rec/T-REC-G.709-201606-I/en>>.
- [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>>.
- [RFC2210] Wroclawski, J., "The Use of RSVP with IETF Integrated Services", RFC 2210, DOI 10.17487/RFC2210, September 1997, <<https://www.rfc-editor.org/info/rfc2210>>.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <<https://www.rfc-editor.org/info/rfc3209>>.

- [RFC3471] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, DOI 10.17487/RFC3471, January 2003, <<https://www.rfc-editor.org/info/rfc3471>>.
- [RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, DOI 10.17487/RFC3473, January 2003, <<https://www.rfc-editor.org/info/rfc3473>>.
- [RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, DOI 10.17487/RFC3477, January 2003, <<https://www.rfc-editor.org/info/rfc3477>>.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <<https://www.rfc-editor.org/info/rfc3630>>.
- [RFC4003] Berger, L., "GMPLS Signaling Procedure for Egress Control", RFC 4003, DOI 10.17487/RFC4003, February 2005, <<https://www.rfc-editor.org/info/rfc4003>>.
- [RFC4328] Papadimitriou, D., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, DOI 10.17487/RFC4328, January 2006, <<https://www.rfc-editor.org/info/rfc4328>>.
- [RFC4606] Mannie, E. and D. Papadimitriou, "Generalized Multi-Protocol Label Switching (GMPLS) Extensions for Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH) Control", RFC 4606, DOI 10.17487/RFC4606, August 2006, <<https://www.rfc-editor.org/info/rfc4606>>.
- [RFC4802] Nadeau, T., Ed. and A. Farrel, Ed., "Generalized Multiprotocol Label Switching (GMPLS) Traffic Engineering Management Information Base", RFC 4802, DOI 10.17487/RFC4802, February 2007, <<https://www.rfc-editor.org/info/rfc4802>>.

- [RFC4872] Lang, J., Ed., Rekhter, Y., Ed., and D. Papadimitriou, Ed., "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872, DOI 10.17487/RFC4872, May 2007, <<https://www.rfc-editor.org/info/rfc4872>>.
- [RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, DOI 10.17487/RFC4873, May 2007, <<https://www.rfc-editor.org/info/rfc4873>>.
- [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>>.
- [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>>.
- [RFC5520] Bradford, R., Ed., Vasseur, JP., and A. Farrel, "Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Path-Key-Based Mechanism", RFC 5520, DOI 10.17487/RFC5520, April 2009, <<https://www.rfc-editor.org/info/rfc5520>>.
- [RFC5521] Oki, E., Takeda, T., and A. Farrel, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Route Exclusions", RFC 5521, DOI 10.17487/RFC5521, April 2009, <<https://www.rfc-editor.org/info/rfc5521>>.
- [RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", RFC 5541, DOI 10.17487/RFC5541, June 2009, <<https://www.rfc-editor.org/info/rfc5541>>.

- [RFC6001] Papadimitriou, D., Vigoureux, M., Shiomoto, K., Brungard, D., and JL. Le Roux, "Generalized MPLS (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 6001, DOI 10.17487/RFC6001, October 2010, <<https://www.rfc-editor.org/info/rfc6001>>.
- [RFC6003] Papadimitriou, D., "Ethernet Traffic Parameters", RFC 6003, DOI 10.17487/RFC6003, October 2010, <<https://www.rfc-editor.org/info/rfc6003>>.
- [RFC6205] Otani, T., Ed. and D. Li, Ed., "Generalized Labels for Lambda-Switch-Capable (LSC) Label Switching Routers", RFC 6205, DOI 10.17487/RFC6205, March 2011, <<https://www.rfc-editor.org/info/rfc6205>>.
- [RFC6387] Takacs, A., Berger, L., Caviglia, D., Fedyk, D., and J. Meuric, "GMPLS Asymmetric Bandwidth Bidirectional Label Switched Paths (LSPs)", RFC 6387, DOI 10.17487/RFC6387, September 2011, <<https://www.rfc-editor.org/info/rfc6387>>.
- [RFC7139] Zhang, F., Ed., Zhang, G., Belotti, S., Ceccarelli, D., and K. Pithewan, "GMPLS Signaling Extensions for Control of Evolving G.709 Optical Transport Networks", RFC 7139, DOI 10.17487/RFC7139, March 2014, <<https://www.rfc-editor.org/info/rfc7139>>.
- [RFC7570] Margaria, C., Ed., Martinelli, G., Balls, S., and B. Wright, "Label Switched Path (LSP) Attribute in the Explicit Route Object (ERO)", RFC 7570, DOI 10.17487/RFC7570, July 2015, <<https://www.rfc-editor.org/info/rfc7570>>.
- [RFC7792] Zhang, F., Zhang, X., Farrel, A., Gonzalez de Dios, O., and D. Ceccarelli, "RSVP-TE Signaling Extensions in Support of Flexi-Grid Dense Wavelength Division Multiplexing (DWDM) Networks", RFC 7792, DOI 10.17487/RFC7792, March 2016, <<https://www.rfc-editor.org/info/rfc7792>>.
- [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>>.
- [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>>.
- [RFC8282] Oki, E., Takeda, T., Farrel, A., and F. Zhang, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Inter-Layer MPLS and GMPLS Traffic Engineering", RFC 8282, DOI 10.17487/RFC8282, December 2017, <<https://www.rfc-editor.org/info/rfc8282>>.
- [RFC8306] Zhao, Q., Dhody, D., Ed., Palleti, R., and D. King, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Point-to-Multipoint Traffic Engineering Label Switched Paths", RFC 8306, DOI 10.17487/RFC8306, November 2017, <<https://www.rfc-editor.org/info/rfc8306>>.

9.2. Informative References

- [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>>.
- [RFC4657] Ash, J., Ed. and J. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, DOI 10.17487/RFC4657, September 2006, <<https://www.rfc-editor.org/info/rfc4657>>.
- [RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010, <<https://www.rfc-editor.org/info/rfc5920>>.
- [RFC6123] Farrel, A., "Inclusion of Manageability Sections in Path Computation Element (PCE) Working Group Drafts", RFC 6123, DOI 10.17487/RFC6123, February 2011, <<https://www.rfc-editor.org/info/rfc6123>>.
- [RFC6163] Lee, Y., Ed., Bernstein, G., Ed., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSOs)", RFC 6163, DOI 10.17487/RFC6163, April 2011, <<https://www.rfc-editor.org/info/rfc6163>>.

- [RFC7025] Otani, T., Ogaki, K., Caviglia, D., Zhang, F., and C. Margaria, "Requirements for GMPLS Applications of PCE", RFC 7025, DOI 10.17487/RFC7025, September 2013, <<https://www.rfc-editor.org/info/rfc7025>>.
- [RFC7449] Lee, Y., Ed., Bernstein, G., Ed., Martensson, J., Takeda, T., Tsuritani, T., and O. Gonzalez de Dios, "Path Computation Element Communication Protocol (PCEP) Requirements for Wavelength Switched Optical Network (WSO) Routing and Wavelength Assignment", RFC 7449, DOI 10.17487/RFC7449, February 2015, <<https://www.rfc-editor.org/info/rfc7449>>.

Appendix A. LOAD-BALANCING Usage for SDH Virtual Concatenation

For example a request for one co-signaled $n \times$ VC-4 TE-LSP will not use the LOAD-BALANCING. In case the VC-4 components can use different paths, the BANDWIDTH with object type TBA-2 will contain a traffic specification indicating the complete $n \times$ VC-4 traffic specification and the LOAD-BALANCING the minimum co-signaled VC-4. For an SDH network, a request to have a TE-LSP group with 10 VC-4 containers, each path using at minimum 2 \times VC-4 containers, can be represented with a BANDWIDTH object with OT=TBA-2, Bw Spec Type set to 4, the content of the Generalized Bandwidth is ST=6, RCC=0, NCC=0, NVC=10, MT=1. The LOAD-BALANCING, OT=TBA-4 with Bw Spec Type set to 4, Max-LSP=5, Min Bandwidth Spec is (ST=6, RCC=0, NCC=0, NVC=2, MT=1). The PCE can respond with a response with maximum 5 paths, each of them having a BANDWIDTH OT=TBA-2 and Generalized Bandwidth matching the Min Bandwidth Spec from the LOAD-BALANCING object of the corresponding request.

Authors' Addresses

Cyril Margaria (editor)
Juniper

Email: cmargaria@juniper.net

Oscar Gonzalez de Dios (editor)
Telefonica Investigacion y Desarrollo
C/ Ronda de la Comunicacion
Madrid 28050
Spain

Phone: +34 91 4833441
Email: oscar.gonzalezdedios@telefonica.com

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

Email: zhangfatai@huawei.com

Network Working Group
Internet-Draft
Intended status: Standards Track

E. Oki
UEC Tokyo
Tomonori Takeda
NTT
J-L Le Roux
France Telecom
A. Farrel
Old Dog Consulting
Fatai Zhang
Huawei
June 17, 2011

Expires: December 17, 2011

Extensions to the Path Computation Element communication Protocol
(PCEP) for Inter-Layer MPLS and GMPLS Traffic Engineering

draft-ietf-pce-inter-layer-ext-05.txt

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

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

This Internet-Draft will expire on December 17, 2011.

Abstract

The Path Computation Element (PCE) provides path computation functions in support of traffic engineering in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks.

MPLS and GMPLS networks may be constructed from layered service networks. It is advantageous for overall network efficiency to provide end-to-end traffic engineering across multiple network layers through a process called inter-layer traffic engineering. PCE is a candidate solution for such requirements.

The PCE communication Protocol (PCEP) is designed as a communication protocol between Path Computation Clients (PCCs) and PCEs. This document presents PCEP extensions for inter-layer traffic engineering.

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.....	3
2. Overview of PCE-Based Inter-Layer Path Computation.....	4
3. Protocol Extensions.....	4
3.1. INTER-LAYER Object.....	4
3.2. SWITCH-LAYER Object.....	7
3.3. REQ-ADAP-CAP Object.....	8
3.4. New Metric Types.....	9
3.5. ERO sub-object.....	10
4. Procedures.....	10
4.1. Path Computation Request.....	10
4.2. Path Computation Reply.....	11
5. Updated Format of PCEP Messages.....	12
6. Manageability Considerations.....	13
7. IANA Considerations.....	13
7.1. New PCEP Objects.....	13
7.2. New Registry for INTER-LAYER Object Flags.....	14
7.3. METRIC Type.....	14
8. Security Considerations.....	15
9. Acknowledgments.....	15
10. References.....	15
10.1. Normative References.....	15
10.2. Informative References.....	16
11. Authors' Addresses.....	16

1. Introduction

The Path Computation Element (PCE) defined in [RFC4655] is an entity that is capable of computing a network path or route based on a network graph, and applying computational constraints. A Path Computation Client (PCC) may make requests to a PCE for paths to be computed.

A network may comprise multiple layers. These layers may represent separations of technologies (e.g., packet switch capable (PSC), time division multiplex (TDM), lambda switch capable (LSC)) [RFC3945], separation of data plane switching granularity levels (e.g., PSC-1 and PSC-2, or VC4 and VC12) [RFC5212], or a distinction between client and server networking roles (e.g., commercial or administrative separation of client and server networks). In this multi-layer network, Label Switched Paths (LSPs) in lower layers are used to carry higher-layer LSPs. The network topology formed by lower-layer LSPs and advertised as traffic engineering links (TE links) in the higher layer is called a Virtual Network Topology (VNT) [RFC5212].

It is important to optimize network resource utilization globally, i.e., taking into account all layers, rather than optimizing resource utilization at each layer independently. This allows better network efficiency to be achieved. This is what we call inter-layer traffic engineering. This includes mechanisms allowing the computation of end-to-end paths across layers (known as inter-layer path computation), and mechanisms for control and management of the VNT by setting up and releasing LSPs in the lower layers [RFC5212].

PCE can provide a suitable mechanism for resolving inter-layer path computation issues. The framework for applying the PCE-based path computation architecture to inter-layer traffic engineering is described in [RFC5623].

The PCE communication protocol (PCEP) is designed as a communication protocol between PCCs and PCEs and is defined in [RFC5440]. A set of requirements for PCEP extensions to support inter-layer traffic engineering is described in [PCE-INTER-LAYER-REQ].

This document presents PCEP extensions for inter-layer traffic engineering that satisfy the requirements described in [PCE-INTER-LAYER-REQ].

2. Overview of PCE-Based Inter-Layer Path Computation

[RFC4206] defines a way to signal a higher-layer LSP which has an explicit route that includes hops traversed by LSPs in lower layers. The computation of end-to-end paths across layers is called Inter-Layer Path Computation.

A Label Switching Router (LSR) in the higher-layer might not have information on the lower-layer topology, particularly in an overlay or augmented model [RFC3945], and hence may not be able to compute an end-to-end path across layers.

PCE-based inter-layer path computation consists of using one or more PCEs to compute an end-to-end path across layers. This could be achieved by relying on a single PCE that has topology information about multiple layers and can directly compute an end-to-end path across layers considering the topology of all of the layers. Alternatively, the inter-layer path computation could be performed using multiple cooperating PCEs where each PCE has information about the topology of one or more layers (but not all layers) and where the PCEs collaborate to compute an end-to-end path.

As described in [RFC 5339], a hybrid nodes may advertise a single TE link with multiple switching capabilities. Those TE links exist at the layer/region boarder normally. In this case, PCE needs to be capable of specifying the server layer path information when the server layer path information is required to be returned to the PCC.

[RFC5623] describes models for inter-layer path computation in more detail.

3. Protocol Extensions

This section describes PCEP extensions for inter-layer path computation. Three new objects are defined: the INTER-LAYER object, the SWITCH-LAYER object, and the REQ-ADAP-CAP object. Also, two new metric types are defined.

3.1. INTER-LAYER Object

The INTER-LAYER object is optional and can be used in PCReq and PCRep messages.

In a PCReq message, the INTER-LAYER object indicates whether inter-layer path computation is allowed, the type of path to be computed, and whether triggered signaling (hierarchical LSPs per [RFC4206] or stitched LSPs per [RFC5150] depending on physical network

technologies) is allowed. When the INTER-LAYER object is absent from a PCReq message, the receiving PCE MUST process as though inter-layer path computation had been explicitly disallowed (I-bit set to zero - see below).

In a PCRep message, the INTER-LAYER object indicates whether inter-layer path computation has been performed, the type of path that has been computed, and whether triggered signaling is used.

When a PCReq message includes more than one request, an INTER-LAYER object is used per request. When a PCRep message includes more than one path per request that is responded to, an INTER-LAYER object is used per path.

INTER-LAYER Object-Class is to be assigned by IANA (recommended value=18)

INTER-LAYER Object-Type is to be assigned by IANA (recommended value=1)

The format of the INTER-LAYER object body is 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
+-----+-----+-----+-----+-----+-----+-----+-----+
|   Reserved                                         |T|M|I|
+-----+-----+-----+-----+-----+-----+-----+-----+

```

I flag (1 bit): The I flag is used by a PCC in a PCReq message to indicate to a PCE whether an inter-layer path is allowed. When the I flag is set (one), the PCE MAY perform inter-layer path computation and return an inter-layer path. When the flag is clear (zero), the path that is returned MUST NOT be an inter-layer path.

The I flag is used by a PCE in a PCRep message to indicate to a PCC whether the path returned is an inter-layer path. When the I flag is set (one), the path is an inter-layer path. When it is clear (zero), the path is contained within a single layer either because inter-layer path computation was not performed or because a mono-layer path (without any virtual TE link and without any loose hop that spans the lower-layer network) was found notwithstanding the use of inter-layer path computation.

M flag (1 bit): The M flag is used by a PCC in a PCReq message to indicate to a PCE whether mono-layer path or multi-layer path is

requested. When the M flag is set (one), multi-layer path is requested. When it is clear (zero), mono-layer path is requested.

The M flag is used by a PCE in a PCRep message to indicate to a PCC whether mono-layer path or multi-layer path is returned. When M flag is set (one), multi-layer path is returned. When M flag is set (zero), mono-layer path is returned.

If the I flag is clear (zero), the M flag has no meaning and MUST be ignored.

[PCE-INTER-LAYER-REQ] describes two sub-options for mono-layer path.

- A mono-layer path that is specified by strict hops. The path may include virtual TE links.
- A mono-layer path that includes loose hops that span the lower-layer network.

The choice of this sub-option can be specified by the use of O flag in the RP object specified in [RFC5440].

T flag (1 bit): The T flag is used by a PCC in a PCReq message to indicate to a PCE whether triggered signaling is allowed. When the T flag is set (one), triggered signaling is allowed. When it is clear (zero), triggered signaling is not allowed.

The T flag is used by a PCE in a PCRep message to indicate to a PCC whether triggered signaling is required to support the returned path. When the T flag is set (one), triggered signaling is required. When it is clear (zero), triggered signaling is not required.

Note that triggered signaling is used to support hierarchical [RFC4206] or stitched [RFC5150] LSPs according to the physical attributes of the network layers.

If the I flag is clear (zero), the T flag has no meaning and MUST be ignored.

Note that the I flag and M flag differ in the following ways. - When the I flag is clear (zero), virtual TE links must not be used in path computation. In addition, loose hops that span the lower-layer network must not be specified. Only regular TE links from the same layer may be used.

- When the I flag is set (one), the M flag is clear (zero), and the T flag is set (one), virtual TE links are allowed in path computation.

In addition, when the O flag of the RP object is set, loose hops that span the lower-layer network may be specified. This will initiate lower-layer LSP setup, thus inter-layer path is setup even though the path computation result from a PCE to a PCC include hops from the same layer only.

- However, when the I flag is set (one), the M flag is clear (zero), and the T flag is clear (zero), since triggered signaling is not allowed, virtual TE links must not be used in path computation. In addition, loose hops that span the lower-layer network must not be specified. Therefore, this is equivalent to the I flag being clear (zero).

Reserved bits of the INTER-LAYER object SHOULD be transmitted as zero and SHOULD be ignored on receipt. A PCE that forwards a path computation request to other PCEs SHOULD preserve the settings of reserved bits in the PCReq messages it sends and in the PCRep messages it forwards to PCCs.

3.2. SWITCH-LAYER Object

The SWITCH-LAYER object is optional on a PCReq message and specifies switching layers in which a path MUST, or MUST NOT, be established. A switching layer is expressed as a switching type and encoding type. The SWITCH-LAYER object MUST NOT be used on a PCReq unless an INTER-LAYER object is also present on the PCReq message.

The SWITCH-LAYER object is optional on a PCRep message, where it is used with the NO-PATH object in the case of unsuccessful path computation to indicate the set of constraints that could not be satisfied.

SWITCH-LAYER Object-Class is to be assigned by IANA (recommended value=19)

SWITCH-LAYER Object-Type is to be assigned by IANA (recommended value=1)

The format of the SWITCH-LAYER object body is 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
LSP Enc. Type								Switching Type								Reserved								I							
								.																							
//								.								//								//							
								.																							
LSP Enc. Type								Switching Type								Reserved								I							

Each row indicates a switching type and encoding type that must or must not be used for specified layer(s) in the computed path.

The format is based on [RFC3471], and has equivalent semantics.

LSP Encoding Type (8 bits): see [RFC3471] for a description of parameters.

Switching Type (8 bits): see [RFC3471] for a description of parameters.

I flag (1 bit): the I flag indicates whether a layer with the specified switching type and encoding type must or must not be used by the computed path. When the I flag is set (one), the computed path MUST traverse a layer with the specified switching type and encoding type. When the I flag is clear (zero), the computed path MUST NOT enter or traverse any layer with the specified switching type and encoding type.

When a combination of switching type and encoding type is not included in SWITCH-LAYER object, the computed path MAY traverse a layer with that combination of switching type and encoding type.

A PCC may want to specify only a Switching Type and not an LSP Encoding Type. In this case, the LSP Encoding Type is set to zero.

3.3. REQ-ADAP-CAP Object

The REQ-ADAP-CAP object is optional and is used to specify a requested adaptation capability for both ends of the lower layer LSP. The REQ-ADAP-CAP object is used in a PCReq message for inter-PCE communication, where the PCE that is responsible for computing higher layer paths acts as a PCC to request a path computation from a PCE that is responsible for computing lower layer paths.

The REQ-ADAP-CAP object is used in a PCRep message in case of unsuccessful path computation (in this case, the PCRep message also contains a NO-PATH object, and the REQ-ADAP-CAP object is used to indicate the set of constraints that could not be satisfied).

The REQ-ADAP-CAP object MAY be used in a PCReq message in a mono-layer network to specify a requested adaptation capability for both ends of the LSP. In this case, it MAY be carried without INTER-LAYER Object.

REQ-ADAP-CAP Object-Class is to be assigned by IANA (recommended value=20)

REQ-ADAP-CAP Object-Type is to be assigned by IANA (recommended value=1)

The format of the REQ-ADAP-CAP object body is 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
+-----+-----+-----+-----+-----+-----+-----+-----+
| Switching Cap |   Encoding   |   Reserved   |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

The format is based on [MLN-SOL] and has equivalent semantics.

Switching Capability (8 bits): see [RFC4203] for a description of parameters.

Encoding (8 bits): see [RFC3471] for a description of parameters.

A PCC may want to specify a Switching Capability, but not an Encoding. In this case, the Encoding MUST be set zero.

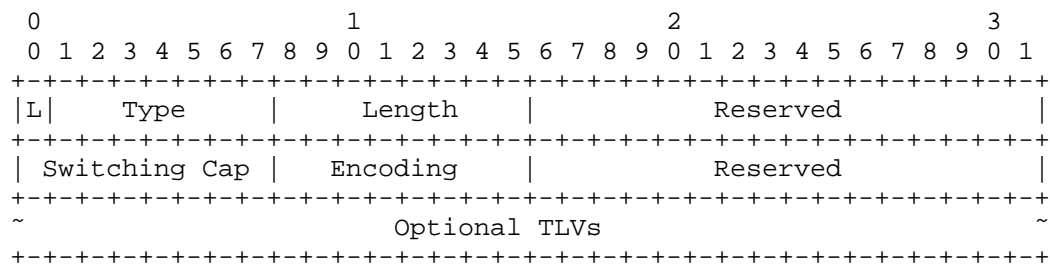
3.4. New Metric Types

Two new metric types are defined for the METRIC object in PCEP.

Type 11 (suggested value, to be assigned by IANA): Number of adaptations on a path.

Type 12 (suggested value, to be assigned by IANA): Number of layers to be involved on a path.

A new ERO sub-object named `SERVER_LAYER_INFO` sub-object is defined in this document. This sub-object can be used for specifying the detail server layer information of a server layer path, e.g. the switching capability and encoding of the server layer path.



This sub-object MAY be added immediately behind the node or link address sub-objects of the two edge nodes of a service path.

The SWITCH-LAYER object, which MUST NOT be present unless the INTER-LAYER object is also present, is optionally used to specify the

switching types and encoding types that define layers that must, or must not, be used in the computed path. When the SWITCH-LAYER object is used with the INTER-LAYER object I flag clear (zero), inter-layer path computation is not allowed, but constraints specified in the SWITCH-LAYER object apply. Example usage includes path computation in a single layer GMPLS network.

The REQ-ADAP-CAP object is optionally used to specify the interface switching capability of both ends of the lower layer LSP. The REQ-ADAP-CAP object is used in inter-PCE communication, where the PCE that is responsible for computing higher layer paths makes a request as a PCC to a PCE that is responsible for computing lower layer paths. Alternatively, the REQ-ADAP-CAP object may be used in the NMS-VNTM model, where the VNTM makes a request as a PCC to a PCE that is responsible for computing lower-layer paths.

The METRIC object is optionally used to specify metric types to be optimized or bounded. When metric type 11 (TBC by IANA) is used, it indicates that path computation MUST minimize or bound the number of adaptations on a path. When metric type 12 (TBC by IANA) is used, it indicates that path computation MUST minimize or bound the number of layers to be involved on a path.

Furthermore, in order to allow different objective functions to be applied within different network layers, multiple OF objects MAY be present. In such a case, the first OF object specifies an objective function for the higher-layer network, and subsequent OF objects specify objection functions of the subsequent lower-layer networks.

4.2. Path Computation Reply

In the case of successful path computation, the requested PCE replies to the requesting PCC for the inter-layer path computation result in a PCRep message that MAY include the INTER-LAYER object. When the INTER-LAYER object is included in a PCRep message, the I flag, M flag, and T flag indicate semantics of the path as described in Section 3.1. Furthermore, when the C flag of the METRIC object in a PCReq is set, the METRIC object MUST be included in the PCRep to provide the computed metric value, as specified in [RFC5440].

PCE MAY specify the server layer path information in the ERO. In this case, the server layer path information could be demarcated by two SERVER_LAYER_INFO sub-objects in the ERO object. The server layer path specified in the ERO could be loose or strict. Then the PCC (e.g. NMS, ingress node) can trigger the setup of the LSPs according to the computed path.

In the case of unsuccessful path computation, the PCRep message also contains a NO-PATH object, and the SWITCH-TYPE object and/or the REQ-ADAP-CAP MAY be used to indicate the set of constraints that could not be satisfied.

5. Updated Format of PCEP Messages

Message formats in this section, as those in [RFC5440] are presented using Backus-Naur Format as specified in [RFC5511].

The format of the PCReq message is updated as follows:

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

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

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

    <request> ::= <RP>
                <END-POINTS>
                [<of-list>]
                [<LSPA>]
                [<BANDWIDTH>]
                [<metric-list>]
                [<RRO> [<BANDWIDTH>]]
                [<IRO>]
                [<LOAD-BALANCING>]
                [<INTER-LAYER> [<SWITCH-LAYER>]]
                [<REQ-ADAP-CAP>]

    where:

    <of-list> ::= <OF> [<of-list>]
    <metric-list> ::= <METRIC> [<metric-list>]
```

The format of the PCRep message is updated as follows:

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

```

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

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

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

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

where:
  <attribute-list>::=[<of-list>]
    [<LSPA>]
    [<BANDWIDTH>]
    [<metric-list>]
    [<IRO>]
    [<INTER-LAYER>]
    [<SWITCH-LAYER>]
    [<REQ-ADAP-CAP>]
  <of-list>::=<OF>[<of-list>]
  <metric-list>::=<METRIC>[<metric-list>]

```

6. Manageability Considerations

TBD

Manageability of inter-layer traffic engineering with PCE must address the following consideration for section 5.1.

- need for a MIB module for control and monitoring
- need for built-in diagnostic tools
- configuration implication for the protocol

7. IANA Considerations

7.1. New PCEP Objects

Three new objects: the INTER-LAYER object, the SWITCH-LAYER object, and the REQ-ADAP-CAP object.

INTER-LAYER Object-Class is to be assigned by IANA (recommended value=18)

INTER-LAYER Object-Type is to be assigned by IANA (recommended value=1)

SWITCH-LAYER Object-Class is to be assigned by IANA (recommended value=19)

SWITCH-LAYER Object-Type is to be assigned by IANA (recommended value=1)

REQ-ADAP-CAP Object-Class is to be assigned by IANA (recommended value=20)

REQ-ADAP-CAP Object-Type is to be assigned by IANA (recommended value=1)

7.2. New Registry for INTER-LAYER Object Flags

IANA is requested to create a registry to manage the Flag field of the INTER-Layer object.

New bit numbers may be allocated only by an IETF Consensus action. Each bit should be tracked with the following qualities:

- o Bit number (counting from bit 0 as the most significant bit)
- o Capability Description
- o Defining RFC

Several bits are defined for the INTER-LAYER object flag fields in this document. The following values have been assigned:

Bit Number	Description	Reference
29	T flag	this document
30	M flag	this document
31	I flag	this document

7.3. METRIC Type

Two new metric types are defined in this document for the METRIC object (specified in [RFC5440]). The IANA is requested to make the following allocation (suggested value):

- Type 11 : Number of adaptations on a path
- Type 12 : Number of layers on a path

8. Security Considerations

TBD

Inter-layer traffic engineering with PCE may raise new security issues when PCE-PCE communication is done between different layer networks for inter-layer path computation. Security issues may also exist when a single PCE is granted full visibility of TE information that applies to multiple layers.

It is expected that solutions for inter-layer protocol extensions will address these issues in detail using security techniques such as authentication.

9. Acknowledgments

10. References

10.1. Normative References

- [RFC2119] S. Bradner, "Key words for use in RFCs to indicate requirements levels", RFC 2119, March 1997.
- [RFC3471] L. Burger, "Generalized Multi-Protocol Label Switching (GMPLS)", RFC 3471, January 2003.
- [RFC3945] E. Mannie, "Generalized Multi-Protocol Label Switching Architecture", RFC 3945, October 2004.
- [RFC4203] K. Kompella and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.
- [RFC4206] K. Kompella, and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.
- [RFC5440] JP. Vasseur et al, "Path Computation Element (PCE) Communication Protocol (PCEP)" RFC 5440, March 2009.

- [PCE-INTER-LAYER-REQ] E. Oki et al., "PCC-PCE Communication Requirements for Inter-Layer Traffic Engineering", draft-ietf-pce-inter-layer-req, work in progress.
- [RFC5623] E. Oki et al., "Framework for PCE-Based Inter-Layer MPLS and GMPLS Traffic Engineering", September 2009.
- RFC5339] JL. Le Roux et al, "Evaluation of Existing GMPLS Protocols against Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC5339, September 2008.

10.2. Informative References

- [RFC4655] A. Farrel, JP. Vasseur and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, September 2006.
- [RFC5212] K. Shiimoto et al., "Requirements for GMPLS-based multi-region and multi-layer networks (MRN/MLN)", RFC 5212, July 2008.
- [MLN-SOL] D. Papadimitriou et al., " Generalized Multi-Protocol Label Switching (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", draft-ietf-ccamp-gmpls-mln-extensions, work in progress.
- [RFC5150] A. Ayyangar et al., "Label Switched Path Stitching with Generalized Multiprotocol Label Switching Traffic Engineering (GMPLS TE)", RFC 5150, February 2008.
- [RFC5511] Farrel, A., "Reduced Backus-Naur Form (RBNF) A Syntax Used in Various Protocol Specifications", April 2009.

11. Authors' Addresses

Eiji Oki
University of Electro-Communications
Tokyo
Japan
Email: oki@ice.uec.ac.jp

Tomonori Takeda
NTT

3-9-11 Midori-cho,
Musashino-shi, Tokyo 180-8585, Japan
Email: takeda.tomonori@lab.ntt.co.jp

Jean-Louis Le Roux
France Telecom R&D,
Av Pierre Marzin,
22300 Lannion, France
Email: jeanlouis.leroux@orange-ftgroup.com

Adrian Farrel
Old Dog Consulting
Email: adrian@olddog.co.uk

Fatai Zhang
Huawei Technologies Co., Ltd.
F3-5-B R&D Center, Huawei Base,
Bantian, Longgang District
Shenzhen 518129 P.R.China
Phone: +86-755-28972912
Email: zhangfatai@huawei.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. 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.

Disclaimer of Validity

All IETF Documents and the information contained therein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Full Copyright Statement

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully,

as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 7, 2017

E. Oki
University of Electro-Communications
T. Takeda
NTT
A. Farrel
Juniper Networks
F. Zhang
Huawei Technologies Co., Ltd.
January 3, 2017

Extensions to the Path Computation Element communication Protocol (PCEP)
for Inter-Layer MPLS and GMPLS Traffic Engineering
draft-ietf-pce-inter-layer-ext-12

Abstract

The Path Computation Element (PCE) provides path computation functions in support of traffic engineering in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks.

MPLS and GMPLS networks may be constructed from layered service networks. It is advantageous for overall network efficiency to provide end-to-end traffic engineering across multiple network layers through a process called inter-layer traffic engineering. PCE is a candidate solution for such requirements.

The PCE communication Protocol (PCEP) is designed as a communication protocol between Path Computation Clients (PCCs) and PCEs. This document presents PCEP extensions for inter-layer traffic engineering.

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 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 July 7, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Overview of PCE-Based Inter-Layer Path Computation	4
3. Protocol Extensions	4
3.1. INTER-LAYER Object	4
3.2. SWITCH-LAYER Object	7
3.3. REQ-ADAP-CAP Object	9
3.4. New Metric Types	10
3.5. SERVER-INDICATION Object	10
4. Procedures	11
4.1. Path Computation Request	11
4.2. Path Computation Reply	12
4.3. Stateful PCE and PCE Initiated LSPs	13
5. Updated Format of PCEP Messages	13
6. Manageability Considerations	15
7. IANA Considerations	16
7.1. New PCEP Objects	16
7.2. New Registry for INTER-LAYER Object Flags	16
7.3. New Metric Types	17
8. Security Considerations	17
9. Acknowledgments	18
10. Contributors	18
11. References	18
11.1. Normative References	18
11.2. Informative References	19

Authors' Addresses	21
------------------------------	----

1. Introduction

The Path Computation Element (PCE) defined in [RFC4655] is an entity that is capable of computing a network path or route based on a network graph, and applying computational constraints. A Path Computation Client (PCC) may make requests to a PCE for paths to be computed, and a PCE may initiate or modify services in a network by supplying new paths ([I-D.ietf-pce-stateful-pce], [I-D.ietf-pce-pce-initiated-lsp]).

A network may comprise multiple layers. These layers may represent separations of technologies (e.g., packet switch capable (PSC), time division multiplex (TDM), lambda switch capable (LSC)) [RFC3945], separation of data plane switching granularity levels (e.g., VC4 and VC12) [RFC5212], or a distinction between client and server networking roles (e.g., commercial or administrative separation of client and server networks). In this multi-layer network, Label Switched Paths (LSPs) in lower layers are used to carry higher-layer LSPs. The network topology formed by lower-layer LSPs and advertised as traffic engineering links (TE links) in the higher layer is called a Virtual Network Topology (VNT) [RFC5212]. Discussion of other ways that network layering can be supported such that connectivity in a higher layer network can be provided by LSPs in a lower layer network is provided in [RFC7926].

It is important to optimize network resource utilization globally, i.e., taking into account all layers, rather than optimizing resource utilization at each layer independently. This allows better network efficiency to be achieved. This is what we call inter-layer traffic engineering. This includes mechanisms allowing the computation of end-to-end paths across layers (known as inter-layer path computation), and mechanisms for control and management of the VNT by setting up and releasing LSPs in the lower layers [RFC5212].

PCE can provide a suitable mechanism for resolving inter-layer path computation issues. The framework for applying the PCE-based path computation architecture to inter-layer traffic engineering is described in [RFC5623].

The PCE communication protocol (PCEP) is designed as a communication protocol between PCCs and PCEs and is defined in [RFC5440]. A set of requirements for PCEP extensions to support inter-layer traffic engineering is described in [RFC6457].

This document presents PCEP extensions for inter-layer traffic engineering that satisfy the requirements described in [RFC6457].

2. Overview of PCE-Based Inter-Layer Path Computation

[RFC4206] defines a way to signal a higher-layer LSP which has an explicit route that includes hops traversed by LSPs in lower layers. The computation of end-to-end paths across layers is called Inter-Layer Path Computation.

A Label Switching Router (LSR) in the higher-layer might not have information on the lower-layer topology, particularly in an overlay or augmented model [RFC3945], and hence may not be able to compute an end-to-end path across layers.

PCE-based inter-layer path computation consists of using one or more PCEs to compute an end-to-end path across layers. This could be achieved by relying on a single PCE that has topology information about multiple layers and can directly compute an end-to-end path across layers considering the topology of all of the layers. Alternatively, the inter-layer path computation could be performed using multiple cooperating PCEs where each PCE has information about the topology of one or more layers (but not all layers) and where the PCEs collaborate to compute an end-to-end path.

As described in [RFC5339], a hybrid node may advertise a single TE link with multiple switching capabilities. Those TE links exist at the layer/region border normally. In this case, a PCE needs to be capable of specifying the server layer path information when the server layer path information is required to be returned to the PCC.

[RFC5623] describes models for inter-layer path computation in more detail. It introduces the Virtual Network Topology Manager (VNTM), a functional element that controls the VNT, and sets out three distinct models (and a fourth hybrid model) for inter-layer control involving a PCE, triggered signalling, and a Network Management System (NMS).

3. Protocol Extensions

This section describes PCEP extensions for inter-layer path computation. Four new objects are defined: the INTER-LAYER object, the SWITCH-LAYER object, the REQ-ADAP-CAP object, and the SERVER-INDICATION object. Also, two new metric types are defined.

3.1. INTER-LAYER Object

The INTER-LAYER object is optional and can be used in PCReq and PCRep messages, and also in PCRpt, PCUpd, and PCInitiate messages.

In a PCReq message, the INTER-LAYER object indicates whether inter-layer path computation is allowed, the type of path to be computed,

and whether triggered signaling (hierarchical LSPs per [RFC4206] or stitched LSPs per [RFC5150] depending on physical network technologies) is allowed. When the INTER-LAYER object is absent from a PCReq message, the receiving PCE MUST process as though inter-layer path computation had been explicitly disallowed (I-bit set to zero - see below).

In a PCRep message, the INTER-LAYER object indicates whether inter-layer path computation has been performed, the type of path that has been computed, and whether triggered signaling is used.

When a PCReq message includes more than one request, an INTER-LAYER object is used per request. When a PCRep message includes more than one path per request that is responded to, an INTER-LAYER object is used per path.

The applicability of this object to PCRpt and PCUpd messages follows as the usage of other objects on those messages as described in [I-D.ietf-pce-stateful-pce]. The applicability of this object to the PCInitiate message follows as the usage of other objects on those messages as described in [I-D.ietf-pce-pce-initiated-lsp]. These messages use the <attribute-list> as defined in [RFC5440] and extended by further PCEP extensions, and so the <attribute-list> as extended in Section 5 can be used to include the INTER-LAYER object on these messages.

INTER-LAYER Object-Class TBD1 to be assigned by IANA.

INTER-LAYER Object-Type 1 to be assigned by IANA.

The format of the INTER-LAYER object body is shown in Figure 1.

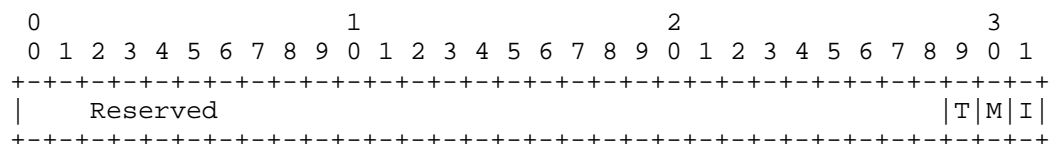


Figure 1: The INTER-LAYER Object

I flag (1 bit): The I flag is used by a PCC in a PCReq message to indicate to a PCE whether an inter-layer path is allowed. When the I flag is set (one), the PCE MAY perform inter-layer path computation and return an inter-layer path. When the flag is clear (zero), the path that is returned MUST NOT be an inter-layer path.

The I flag is used by a PCE in a PCRep message to indicate to a PCC whether the path returned is an inter-layer path. When the I flag is set (one), the path is an inter-layer path. When it is clear (zero), the path is contained within a single layer either because inter-layer path computation was not performed or because a mono-layer path (without any virtual TE link and without any loose hop that spans the lower-layer network) was found notwithstanding the use of inter-layer path computation.

M flag (1 bit): The M flag is used by a PCC in a PCReq message to indicate to a PCE whether mono-layer path or multi-layer path is requested. When the M flag is set (one), multi-layer path is requested. When it is clear (zero), mono-layer path is requested.

The M flag is used by a PCE in a PCRep message to indicate to a PCC whether mono-layer path or multi-layer path is returned. When M flag is set (one), multi-layer path is returned. When M flag is clear (zero), mono-layer path is returned.

If the I flag is clear (zero), the M flag has no meaning and MUST be ignored.

[RFC6457] describes two sub-options for mono-layer path.

- o A mono-layer path that is specified by strict hops. The path may include virtual TE links.
- o A mono-layer path that includes loose hops that span the lower-layer network.

The choice of this sub-option can be specified by the use of O flag in the RP object specified in [RFC5440].

T flag (1 bit): The T flag is used by a PCC in a PCReq message to indicate to a PCE whether triggered signaling is allowed. When the T flag is set (one), triggered signaling is allowed. When it is clear (zero), triggered signaling is not allowed.

The T flag is used by a PCE in a PCRep message to indicate to a PCC whether triggered signaling is required to support the returned path. When the T flag is set (one), triggered signaling is required. When it is clear (zero), triggered signaling is not required.

Note that triggered signaling is used to support hierarchical [RFC4206] or stitched [RFC5150] LSPs according to the physical attributes of the network layers.

If the I flag is clear (zero), the T flag has no meaning and MUST be ignored.

Note that the I flag and M flag differ in the following ways. When the I flag is clear (zero), virtual TE links must not be used in path computation. In addition, loose hops that span the lower-layer network must not be specified. Only regular TE links from the same layer may be used.

- o When the I flag is set (one), the M flag is clear (zero), and the T flag is set (one), virtual TE links are allowed in path computation. In addition, when the O flag of the RP object is set, loose hops that span the lower-layer network may be specified. This will initiate lower-layer LSP setup, thus inter-layer path is setup even though the path computation result from a PCE to a PCC include hops from the same layer only.
- o However, when the I flag is set (one), the M flag is clear (zero), and the T flag is clear (zero), since triggered signaling is not allowed, virtual TE links that have not been pre-signaled MUST NOT be used in path computation. In addition, loose hops that span the lower-layer network MUST NOT be specified. Therefore, this is equivalent to the I flag being clear (zero).

Reserved bits of the INTER-LAYER object sent between a PCC and PCE in the same domain MUST be transmitted as zero and SHOULD be ignored on receipt. A PCE that forwards a path computation request to other PCEs MUST preserve the settings of reserved bits in the PCReq messages it sends and in the PCRep messages it forwards to PCCs.

Note that the flags in the PCRpt message indicate the state of an LSP, whereas the flags in the PCUpd and the PCInitiate messages indicate the intended/desired state as determined by the PCE.

3.2. SWITCH-LAYER Object

The SWITCH-LAYER object is optional on a PCReq message and specifies switching layers in which a path MUST, or MUST NOT, be established. A switching layer is expressed as a switching type and encoding type.

When a SWITCH-LAYER object is used on a PCReq it is interpreted in the context of the INTER-LAYER object on the same message. If no INTER-LAYER object is present, the PCE MUST process the SWITCH-LAYER object as though inter-layer path computation had been explicitly disallowed. In such a case, the SWITCH-LAYER object MUST NOT have more than one LSP Encoding Type and Switching Type with the I flag set.

The SWITCH-LAYER object is optional on a PCRep message, where it is used with the NO-PATH object in the case of unsuccessful path computation to indicate the set of constraints that could not be satisfied.

The SWITCH-LAYER object may be used on a PCRpt message consistent with how properties of existing LSPs are reported on that message [I-D.ietf-pce-stateful-pce]. The PCRpt message uses the <attribute-list> as defined in [RFC5440] and extended by further PCEP extensions. This message can use the <attribute-list> as extended in Section 5 to carry the SWITCH-LAYER object. The SWITCH-LAYER object is not used on a PCUpd or PCInitiate messages.

SWITCH-LAYER Object-Class TBD2 is to be assigned by IANA.

SWITCH-LAYER Object-Type 1 is to be assigned by IANA.

The format of the SWITCH-LAYER object body is shown in Figure 2.

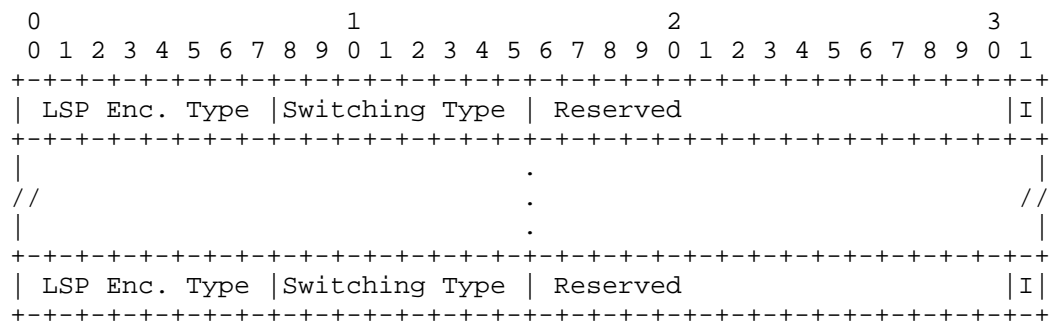


Figure 2: The SWITCH-LAYER Object

Each row indicates a switching type and encoding type that must or must not be used for specified layer(s) in the computed path.

The format is based on [RFC3471], and has equivalent semantics.

LSP Encoding Type (8 bits): see [RFC3471] for a description of parameters.

Switching Type (8 bits): see [RFC3471] for a description of parameters.

I flag (1 bit): the I flag indicates whether a layer with the specified switching type and encoding type must or must not be used

by the computed path. When the I flag is set (one), the computed path MUST traverse a layer with the specified switching type and encoding type. When the I flag is clear (zero), the computed path MUST NOT enter or traverse any layer with the specified switching type and encoding type.

When a combination of switching type and encoding type is not included in SWITCH-LAYER object, the computed path MAY traverse a layer with that combination of switching type and encoding type.

A PCC may want to specify only a Switching Type and not an LSP Encoding Type. In this case, the LSP Encoding Type is set to zero.

3.3. REQ-ADAP-CAP Object

The REQ-ADAP-CAP object is optional and is used to specify a requested adaptation capability for both ends of the lower layer LSP. The REQ-ADAP-CAP object is used in a PCReq message for inter-PCE communication, where the PCE that is responsible for computing higher layer paths acts as a PCC to request a path computation from a PCE that is responsible for computing lower layer paths.

The REQ-ADAP-CAP object is used in a PCRep message in case of unsuccessful path computation (in this case, the PCRep message also contains a NO-PATH object, and the REQ-ADAP-CAP object is used to indicate the set of constraints that could not be satisfied).

The REQ-ADAP-CAP object MAY be used in a PCReq message in a mono-layer network to specify a requested adaptation capability for both ends of the LSP. In this case, it MAY be carried without an INTER-LAYER Object.

The applicability of this object to PCRpt and PCUpd messages follows as the usage of other objects on those messages as described in [I-D.ietf-pce-stateful-pce]. The applicability of this object to the PCInitiate message follows as the usage of other objects on those messages as described in [I-D.ietf-pce-pce-initiated-lsp]. These messages use the <attribute-list> as defined in [RFC5440] and extended by further PCEP extensions. These messages can use the <attribute-list> as extended in Section 5 to carry the REQ-ADAP-CAP object.

REQ-ADAP-CAP Object-Class TBD3 is to be assigned by IANA.

REQ-ADAP-CAP Object-Type 1 is to be assigned by IANA.

The format of the REQ-ADAP-CAP object body is shown in Figure 3.

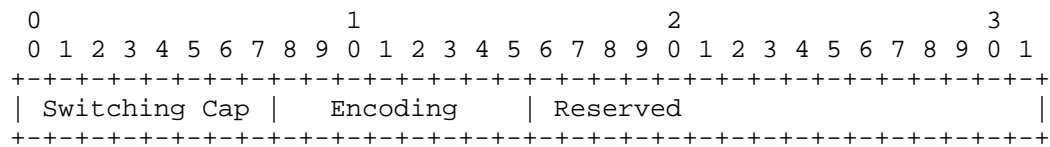


Figure 3: The REQ-ADAP-CAP Object

The format is based on [RFC6001] and has equivalent semantics as the Interface Adjustment Capability Descriptor (IACD) Upper Switching Capability and Lower Switching Capability fields.

Switching Capability (8 bits): see [RFC4203] for a description of parameters.

Encoding (8 bits): see [RFC3471] for a description of parameters.

A PCC may want to specify a Switching Capability, but not an Encoding. In this case, the Encoding MUST be set zero.

3.4. New Metric Types

This document defines two new metric types for use in the PCEP METRIC object.

IANA has assigned the value TBD5 to indicate the metric "Number of adaptations on a path."

IANA has assigned the value TBD6 to indicate the metric "Number of layers to be involved on a path."

See Section 4.1, Section 4.2, and Section 4.3 for a description of how these metrics are applied.

3.5. SERVER-INDICATION Object

The SERVER-INDICATION is optional and is used to indicate that path information included in the ERO is server layer information and specify the characteristics of the server layer, e.g., the switching capability and encoding of the server layer path.

The format of the SERVER-INDICATION object body is shown in Figure 4.

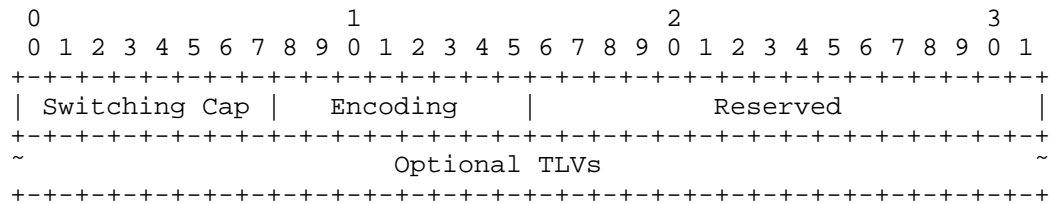


Figure 4: The SERVER-INDICATION Object

SERVER-INDICATION Object-Class TBD4 to be assigned by IANA.

SERVER-INDICATION Object-Type 1 to be assigned by IANA.

Switching Capability (8 bits): see [RFC4203] for a description of parameters.

Encoding (8 bits): see [RFC3471] for a description of parameters.

Optional TLVs: Optional TLVs MAY be included within the object to specify more specific server layer path information (e.g., traffic parameters). Such TLVs will be defined by other documents.

4. Procedures

4.1. Path Computation Request

A PCC requests or allows inter-layer path computation in a PCReq message by including the INTER-LAYER object with the I flag set. The INTER-LAYER object indicates whether inter-layer path computation is allowed, which path type is requested, and whether triggered signaling is allowed.

The SWITCH-LAYER object, which MUST NOT be present unless the INTER-LAYER object is also present, is optionally used to specify the switching types and encoding types that define layers that must, or must not, be used in the computed path. When the SWITCH-LAYER object is used with the INTER-LAYER object I flag clear (zero), inter-layer path computation is not allowed, but constraints specified in the SWITCH-LAYER object apply. Example usage includes path computation in a single layer GMPLS network.

The REQ-ADAP-CAP object is optionally used to specify the interface switching capability of both ends of the lower layer LSP. The REQ-ADAP-CAP object is used in inter-PCE communication, where the PCE that is responsible for computing higher layer paths makes a request as a PCC to a PCE that is responsible for computing lower layer

paths. Alternatively, the REQ-ADAP-CAP object may be used in the NMS-VNTM model, where the VNTM makes a request as a PCC to a PCE that is responsible for computing lower-layer paths.

The METRIC object is optionally used to specify metric types to be optimized or bounded. When metric type TBD5 is used, it indicates that path computation MUST minimize or bound the number of adaptations on a path. When metric type TBD6 is used, it indicates that path computation MUST minimize or bound the number of layers to be involved on a path.

Furthermore, in order to allow different objective functions to be applied within different network layers, multiple OF objects [RFC5541] MAY be present. In such a case, the first OF object specifies an objective function for the higher-layer network, and subsequent OF objects specify objection functions of the subsequent lower-layer networks.

4.2. Path Computation Reply

In the case of successful path computation, the requested PCE replies to the requesting PCC for the inter-layer path computation result in a PCRep message that MAY include the INTER-LAYER object. When the INTER-LAYER object is included in a PCRep message, the I flag, M flag, and T flag indicate semantics of the path as described in Section 3.1. Furthermore, when the C flag of the METRIC object in a PCReq is set, the METRIC object MUST be included in the PCRep to provide the computed metric value, as specified in [RFC5440].

The PCE MAY specify the server layer path information in the ERO. In this case, the requested PCE replies with a PCRep message that includes at least two sets of ERO information in the path-list, one is for the client layer path information, and another one is the server layer path information. When SERVER-INDICATION is included in a PCRep message, it indicates that the path in the ERO is the server layer path information. The server layer path specified in the ERO could be loose or strict. On receiving the replied path, the PCC (e.g., NMS, ingress node) can trigger the signaling to setup the LSPs according to the computed paths.

In the case of unsuccessful path computation, the PCRep message also contains a NO-PATH object, and the SWITCH-TYPE object and/or the REQ-ADAP-CAP MAY be used to indicate the set of constraints that could not be satisfied.

4.3. Stateful PCE and PCE Initiated LSPs

Processing for stateful PCEs is described in [I-D.ietf-pce-stateful-pce]. That document defines the PCRpt message to allow a PCC to report to a PCE an LSP that already exists in the network and to delegate control of that LSP to the PCE.

When the LSP is a multi-layer LSP (or a mono-layer LSP for which specific adaptations exist), the message objects defined in this document are used on the PCRpt to describe an LSP that is delegated to the PCE so that the PCE may process the LSP.

Furthermore, [I-D.ietf-pce-stateful-pce] defines the PCUpd message to allow a PCE to modify an LSP that has been delegated to it. When the LSP is a multi-layer LSP (or a mono-layer LSP for which specific adaptations exist), the message objects defined in this document are used on the PCUpd to describe the new attributes of the modified LSP.

Processing for PCE initiated LSPs is described in [I-D.ietf-pce-pce-initiated-lsp]. That document defines the PCInitiate message that is used by a PCE to request a PCC to set up a new LSP. When the LSP is a multi-layer LSP (or a mono-layer LSP for which specific adaptations exist), the message objects defined in this document are used on the PCInitiate to describe the attributes of the new LSP.

The new metric types defined in this document can also be used with the stateful PCE extensions. The format of PCEP messages described in [I-D.ietf-pce-stateful-pce] and [I-D.ietf-pce-pce-initiated-lsp] uses <attribute-list> (which is extended in Section 5 for the purpose of including the new metrics).

The stateful PCE implementation MAY use the extension of PCReq and PCRep messages as defined in Section 5 to enable the use of inter-layer parameters during passive stateful operations too, using the LSP object.

5. Updated Format of PCEP Messages

Message formats in this section, as those in [RFC5440] are presented using Routing Backus-Naur Format (RBNF) as specified in [RFC5511].

The format of the PCReq message is updated as shown in Figure 5

```

<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>]
                [<of-list>]
                [<RRO>[<BANDWIDTH>]]
                [<IRO>]
                [<LOAD-BALANCING>]
                [<INTER-LAYER> [<SWITCH-LAYER>]]
                [<REQ-ADAP-CAP>]

where:

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

```

Figure 5: The Updated PCReq Message

The format of the PCRep message is updated as shown in Figure 6

```

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

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

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

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

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

where:
  <attribute-list> ::= [<of-list>]
                     [<LSPA>]
                     [<BANDWIDTH>]
                     [<metric-list>]
                     [<IRO>]
                     [<INTER-LAYER>]
                     [<SWITCH-LAYER>]
                     [<REQ-ADAP-CAP>]
                     [<SERVER-INDICATION>]

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

```

Figure 6: The Updated PCRep Message

6. Manageability Considerations

Implementations of this specification should provide a mechanism to configure any optional features (such as whether a PCE supports inter-layer computation, and which metrics are supported).

A Management Information Base (MIB) module for modeling PCEP is described in [RFC7420]. Systems that already use a MIB module to manage their PCEP implementations might want to augment that module to provide controls and indicators for support of inter-layer features defined in this document, and to add counters of messages sent and received containing the objects defined here.

However, the preferred mechanism for configuration is through a YANG model. Work has started on a YANG model for PCEP

[I-D.ietf-pce-pcep-yang], and this could be enhanced as described for the MIB module, above.

Additional policy configuration might be provided to allow a PCE to discriminate between the computation services offered to different PCCs.

A set of monitoring tools for the PCE-based architecture are provided in [RFC5886]. Systems implementing this specification and PCE monitoring should consider defining extensions to the mechanisms defined in [RFC5886] to help monitor inter-layer path computation requests.

7. IANA Considerations

IANA maintains a registry called the "Path Computation Element Protocol (PCEP) Numbers". This document requests IANA to carry out actions on subregistries of that registry.

7.1. New PCEP Objects

IANA is requested to make the following assignments from the "PCEP Objects" subregistry.

Object-Class Value	Name	Object-Type	Reference
INTER-LAYER	TBD1	1: Inter-layer 2-15: Unassigned	[This.I-D]
SWITCH-LAYER	TBD2	1: Switch-layer 2-15: Unassigned	[This.I-D]
REQ-ADAP-CAP	TBD3	1: Req-Adap-Cap 2-15: Unassigned	[This.I-D]
SERVER-INDICATION	TBD4	1: Server-indication	[This.I-D]

Figure 7

7.2. New Registry for INTER-LAYER Object Flags

IANA is requested to create a new subregistry to manage the Flag field of the INTER-Layer object called the "Inter-Layer Object Path Property Bits" registry.

New bit numbers may be allocated only by an "IETF Review" action [RFC5226]. Each bit should be tracked with the following qualities:

- o Bit number (counting from bit 0 as the most significant bit up to a maximum of bit 31)
- o Capability Description
- o Defining RFC

IANA is requested to pre-populate the registry as follows:

Bit	Flag	Multi-Layer Path Property	Reference
0-28		Unassigned	
29	T	Triggered Signalling Allowed	[This.I-D]
30	M	Multi-Layer Requested	[This.I-D]
31	I	Inter-Layer Allowed	[This.I-D]

Figure 8

7.3. New Metric Types

Two new metric types are defined in this document for the METRIC object (specified in [RFC5440]). The IANA is requested to make the following allocations from the "Metric Object T Field" registry.

Value	Description	Reference
TBD5	Number of adaptations on a path	[This.I-D]
TBD6	Number of layers on a path	[This.I-D]

Figure 9

IANA is further requested to update the registry to show an assignment action of "IETF Consensus" as already documented in [RFC5440].

8. Security Considerations

Inter-layer traffic engineering with PCE may raise new security issues when PCE-PCE communication is done between different layer networks for inter-layer path computation. Security issues may also exist when a single PCE is granted full visibility of TE information that applies to multiple layers.

Path-Key-based mechanism defined in [RFC5520] MAY be applied to address the topology confidentiality between different layers.

9. Acknowledgments

The authors would like to thank Cyril Margaria for his valuable comments. Helpful comments and suggested text were offered by Dhruv Dhody who also fixed the RBNF. Jonathan Hardwick provided a helpful review as document shepherd.

10. Contributors

Jean-Louis Le Roux
France Telecom R&D
Av Pierre Marzin
Lannion
France
22300

Email: jeanlouis.leroux@orange.com

11. References

11.1. Normative References

- [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-07 (work in progress), July 2016.
- [I-D.ietf-pce-stateful-pce]
Crabbe, E., Minei, I., Medved, J., and R. Varga, "PCEP Extensions for Stateful PCE", draft-ietf-pce-stateful-pce-18 (work in progress), December 2016.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC3471] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, DOI 10.17487/RFC3471, January 2003, <<http://www.rfc-editor.org/info/rfc3471>>.

- [RFC3945] Mannie, E., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, DOI 10.17487/RFC3945, October 2004, <<http://www.rfc-editor.org/info/rfc3945>>.
- [RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <<http://www.rfc-editor.org/info/rfc4203>>.
- [RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, DOI 10.17487/RFC4206, October 2005, <<http://www.rfc-editor.org/info/rfc4206>>.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, DOI 10.17487/RFC5226, May 2008, <<http://www.rfc-editor.org/info/rfc5226>>.
- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <<http://www.rfc-editor.org/info/rfc5440>>.
- [RFC5520] Bradford, R., Ed., Vasseur, JP., and A. Farrel, "Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Path-Key-Based Mechanism", RFC 5520, DOI 10.17487/RFC5520, April 2009, <<http://www.rfc-editor.org/info/rfc5520>>.

11.2. Informative References

- [I-D.ietf-pce-pcep-yang] Dhody, D., Hardwick, J., Beeram, V., and j. jeffrant@gmail.com, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", draft-ietf-pce-pcep-yang-01 (work in progress), October 2016.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <<http://www.rfc-editor.org/info/rfc4655>>.

- [RFC5150] Ayyangar, A., Kompella, K., Vasseur, JP., and A. Farrel, "Label Switched Path Stitching with Generalized Multiprotocol Label Switching Traffic Engineering (GMPLS TE)", RFC 5150, DOI 10.17487/RFC5150, February 2008, <<http://www.rfc-editor.org/info/rfc5150>>.
- [RFC5212] Shiimoto, K., Papadimitriou, D., Le Roux, JL., Vigoureux, M., and D. Brungard, "Requirements for GMPLS-Based Multi-Region and Multi-Layer Networks (MRN/MLN)", RFC 5212, DOI 10.17487/RFC5212, July 2008, <<http://www.rfc-editor.org/info/rfc5212>>.
- [RFC5339] Le Roux, JL., Ed. and D. Papadimitriou, Ed., "Evaluation of Existing GMPLS Protocols against Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 5339, DOI 10.17487/RFC5339, September 2008, <<http://www.rfc-editor.org/info/rfc5339>>.
- [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, <<http://www.rfc-editor.org/info/rfc5511>>.
- [RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", RFC 5541, DOI 10.17487/RFC5541, June 2009, <<http://www.rfc-editor.org/info/rfc5541>>.
- [RFC5623] Oki, E., Takeda, T., Le Roux, JL., and A. Farrel, "Framework for PCE-Based Inter-Layer MPLS and GMPLS Traffic Engineering", RFC 5623, DOI 10.17487/RFC5623, September 2009, <<http://www.rfc-editor.org/info/rfc5623>>.
- [RFC5886] Vasseur, JP., Ed., Le Roux, JL., and Y. Ikejiri, "A Set of Monitoring Tools for Path Computation Element (PCE)-Based Architecture", RFC 5886, DOI 10.17487/RFC5886, June 2010, <<http://www.rfc-editor.org/info/rfc5886>>.
- [RFC6001] Papadimitriou, D., Vigoureux, M., Shiimoto, K., Brungard, D., and JL. Le Roux, "Generalized MPLS (GMPLS) Protocol Extensions for Multi-Layer and Multi-Region Networks (MLN/MRN)", RFC 6001, DOI 10.17487/RFC6001, October 2010, <<http://www.rfc-editor.org/info/rfc6001>>.

- [RFC6457] Takeda, T., Ed. and A. Farrel, "PCC-PCE Communication and PCE Discovery Requirements for Inter-Layer Traffic Engineering", RFC 6457, DOI 10.17487/RFC6457, December 2011, <<http://www.rfc-editor.org/info/rfc6457>>.
- [RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <<http://www.rfc-editor.org/info/rfc7420>>.
- [RFC7926] Farrel, A., Ed., Drake, J., Bitar, N., Swallow, G., Ceccarelli, D., and X. Zhang, "Problem Statement and Architecture for Information Exchange between Interconnected Traffic-Engineered Networks", BCP 206, RFC 7926, DOI 10.17487/RFC7926, July 2016, <<http://www.rfc-editor.org/info/rfc7926>>.

Authors' Addresses

Eiji Oki
University of Electro-Communications
Tokyo
Japan

Email: oki@ice.uec.ac.jp

Tomonori Takeda
NTT
3-9-11 Midori-cho
Musashino-shi, Tokyo
Japan

Email: tomonori.takeda@ntt.com

Adrian Farrel
Juniper Networks

Email: afarrel@juniper.net

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

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

Network Working Group
Internet Draft
Intended status: Informational
Expires: January 2012

Y. Lee
Huawei

G. Bernstein
Grotto Networking

Jonas Martensson
Acreo

T. Takeda
NTT

T. Tsuritani
KDDI

July 6, 2011

PCEP Requirements for WSON Routing and Wavelength Assignment

draft-ietf-pce-wson-routing-wavelength-05.txt

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/ietf/lid-abstracts.txt>

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

This Internet-Draft will expire on September 6, 2011.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Abstract

This memo provides application-specific requirements for the Path Computation Element communication Protocol (PCEP) for the support of Wavelength Switched Optical Networks (WSON). Lightpath provisioning in WSONs requires a routing and wavelength assignment (RWA) process. From a path computation perspective, wavelength assignment is the process of determining which wavelength can be used on each hop of a path and forms an additional routing constraint to optical light path computation. Requirements for Optical impairments will be addressed in a separate document.

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 0.

Table of Contents

1. Introduction.....	3
1.1. WSON RWA Processes.....	4
2. WSON PCE Architectures and Requirements.....	5
2.1. RWA PCC to PCE Interface.....	6
2.1.1. RWA Computation Type and Wavelength Assignment Option6	
2.1.2. Bulk RWA path request/reply.....	6
2.1.3. An RWA path re-optimization request/reply.....	7
2.1.4. Wavelength Range Constraint.....	7
2.1.5. Wavelength Policy Constraint.....	7

3. Manageability Considerations.....	8
3.1. Control of Function and Policy.....	8
3.2. Information and Data Models, e.g. MIB module.....	8
3.3. Liveness Detection and Monitoring.....	9
3.4. Verifying Correct Operation.....	9
3.5. Requirements on Other Protocols and Functional Components.....	9
3.6. Impact on Network Operation.....	9
4. Security Considerations.....	9
5. IANA Considerations.....	9
6. Acknowledgments.....	10
7. References.....	10
7.1. Normative References.....	10
7.2. Informative References.....	11
Authors' Addresses.....	11
Intellectual Property Statement.....	12
Disclaimer of Validity.....	13

1. Introduction

[RFC4655] defines the PCE based Architecture and explains how a Path Computation Element (PCE) may compute Label Switched Paths (LSP) in Multiprotocol Label Switching Traffic Engineering (MPLS-TE) and Generalized MPLS (GMPLS) networks at the request of Path Computation Clients (PCCs). A PCC is shown to be any network component that makes such a request and may be for instance an Optical Switching Element within a Wavelength Division Multiplexing (WDM) network. The PCE, itself, can be located anywhere within the network, and may be within an optical switching element, a Network Management System (NMS) or Operational Support System (OSS), or may be an independent network server.

The PCE communications Protocol (PCEP) is the communication protocol used between PCC and PCE, and may also be used between cooperating PCEs. [RFC4657] sets out the common protocol requirements for PCEP. Additional application-specific requirements for PCEP are deferred to separate documents.

This document provides a set of application-specific PCEP requirements for support of path computation in Wavelength Switched Optical Networks (WSON). WSON refers to WDM based optical networks in which switching is performed selectively based on the wavelength of an optical signal.

The path in WSON is referred to as a lightpath. A lightpath may span multiple fiber links and the path should be assigned a wavelength for each link. A transparent optical network is made up of optical devices that can switch but not convert from one wavelength to

another. In a transparent optical network, a lightpath operates on the same wavelength across all fiber links that it traverses. In such case, the lightpath is said to satisfy the wavelength-continuity constraint. Two lightpaths that share a common fiber link can not be assigned the same wavelength. To do otherwise would result in both signals interfering with each other. Note that advanced additional multiplexing techniques such as polarization based multiplexing are not addressed in this document since the physical layer aspects are not currently standardized. Therefore, assigning the proper wavelength on a lightpath is an essential requirement in the optical path computation process.

When a switching node has the ability to perform wavelength conversion the wavelength-continuity constraint can be relaxed, and a lightpath may use different wavelengths on different links along its route from origin to destination. It is, however, to be noted that wavelength converters may be limited due to their relatively high cost, while the number of WDM channels that can be supported in a fiber is also limited. As a WSON can be composed of network nodes that cannot perform wavelength conversion, nodes with limited wavelength conversion, and nodes with full wavelength conversion abilities, wavelength assignment is an additional routing constraint to be considered in all lightpath computation.

In this document we first review the processes for routing and wavelength assignment (RWA) used when wavelength continuity constraints are present and then specify requirements for PCEP to support RWA.

The remainder of this document uses terminology from [RFC4655].

1.1. WSON RWA Processes

In [WSON-Frame] three alternative process architectures were given for performing routing and wavelength assignment. These are shown schematically in Figure 1.

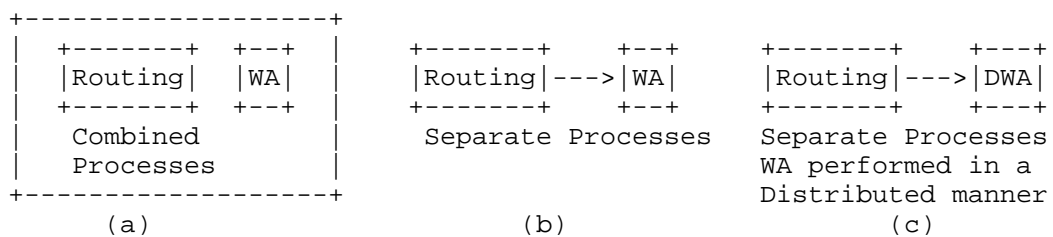


Figure 1 RWA process alternatives.

These alternatives have the following properties and impact on PCEP requirements in this document.

1. Combined Processes (R&WA) - Here path selection and wavelength assignment are performed as a single process. The requirements for PCC-PCE interaction with such a combined RWA process PCE is addressed in this document.
2. Routing separate from Wavelength Assignment (R+WA) - Here the routing process furnishes one or more potential paths to the wavelength assignment process that then performs final path selection and wavelength assignment. The requirements for PCE-PCE interaction with one PCE implementing the routing process and another implementing the wavelength assignment process are not addressed in this document.
3. Routing and distributed Wavelength Assignment (R+DWA) - Here a standard path computation (unaware of detailed wavelength availability) takes place, then wavelength assignment is performed along this path in a distributed manner via signaling (RSVP-TE). This alternative should be covered by existing or emerging GMPLS PCEP extensions and does not present new WSON specific requirements.

2. WSON PCE Architectures and Requirements

In the previous section various process architectures for implementing RWA have been reviewed. Figure 2 shows one typical PCE based implementation, which is referred to as Combined Process (R&WA). With this architecture, the two processes of routing and wavelength assignment are accessed via a single PCE. This architecture is the base architecture from which the requirements are specified in this document.

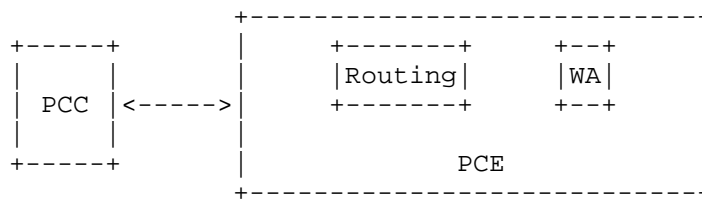


Figure 2 Combined Process (R&WA) architecture

2.1. RWA PCC to PCE Interface

The requirements for the PCC to PCE interface of Figure 2 are specified in this section.

2.1.1. RWA Computation Type and Wavelength Assignment Option

1. The PCReq Message MUST include the path computation type. This can be:

- (i) Both Routing and Wavelength Assignment (RWA), or
- (ii) Routing only.

This requirement is needed to differentiate between the currently supported routing with distributed wavelength assignment option and combined RWA. In case of distributed wavelength assignment option, wavelength assignment will be performed at each node of the route.

2. When the PCReq Message is RWA path computation type, the PCReq Message MUST further include the wavelength assignment options. At the minimum, the following option should be supported:

- (i) Explicit Label Control (ELC) [RFC4003]
- (ii) Non-Explicit labels in the form of Label Sets (This will allow Distributed WA at a node level where each node would select the wavelength from the Label Sets)

3. The PCRep Message MUST include the route, wavelengths assigned to the route and indication of which wavelength assignment option has been applied (ELC or Label Sets).
4. In the case where a valid path is not found, the PCRep Message MUST include why the path is not found (e.g., no route, wavelength not found, optical quality check failed, etc.)

2.1.2. Bulk RWA path request/reply

1. The PCReq Message MUST be able to specify an option for bulk RWA path request. Bulk path request is an ability to request a number of simultaneous RWA path requests.

2. The PCRep Message MUST include the route, wavelength assigned to the route for each RWA path request specified in the original bulk PCReq Message.

2.1.3. An RWA path re-optimization request/reply

1. For a re-optimization request, the PCReq Message MUST provide the path to be re-optimized and include the following options:
 - a. Re-optimize the path keeping the same wavelength(s)
 - b. Re-optimize wavelength(s) keeping the same path
 - c. Re-optimize allowing both wavelength and the path to change
2. The corresponding PCRep Message for the re-optimized request MUST provide the Re-optimized path and wavelengths.
3. In case that the path is not found, the PCRep Message MUST include why the path is not found (e.g., no route, wavelength not found, both route and wavelength not found, etc.)

2.1.4. Wavelength Range Constraint

For any PCReq Message that is associated with a request for wavelength assignment the requester (PCC) MUST be able to specify a restriction on the wavelengths to be used.

Note that the requestor (PCC) is NOT required to furnish any range restrictions. This restriction is to be interpreted by the PCE as a constraint on the tuning ability of the origination laser transmitter.

2.1.5. Wavelength Policy Constraint

The PCReq Message May include specific operator's policy information for WA (E.g., random assignment, descending order, ascending order, etc.)

2.1.6. Signal Processing Capability Restriction

The PCReq Message MUST be able to specify restrictions for signal compatibility either on the endpoint or any given link. The following signal processing capability should be supported at a minimum:

- o Modulation Type List
- o FEC Type List

3. Manageability Considerations

Manageability of WSON Routing and Wavelength Assignment (RWA) with PCE must address the following considerations:

3.1. Control of Function and Policy

In addition to the parameters already listed in Section 8.1 of [RFC5440], a PCEP implementation SHOULD allow configuring the following PCEP session parameters on a PCC:

- o The ability to send a WSON RWA request.

In addition to the parameters already listed in Section 8.1 of [RFC5440], a PCEP implementation SHOULD allow configuring the following PCEP session parameters on a PCE:

- o The support for WSON RWA.
- o The maximum number of bulk path requests associated with WSON RWA per request message.

These parameters may be configured as default parameters for any PCEP session the PCEP speaker participates in, or may apply to a specific session with a given PCEP peer or a specific group of sessions with a specific group of PCEP peers.

3.2. Information and Data Models, e.g. MIB module

As this document only concerns the requirements to support WSON RWA, no additional MIB module is defined in this document. However, the corresponding solution draft will list the information that should be added to the PCE MIB module defined in [PCEP-MIB].

3.3. Liveness Detection and Monitoring

Mechanisms defined in this document do not imply any new liveness detection and monitoring requirements in addition to those already listed in section 8.3 of [RFC5440].

3.4. Verifying Correct Operation

Mechanisms defined in this document do not imply any new verification requirements in addition to those already listed in section 8.4 of [RFC5440]

3.5. Requirements on Other Protocols and Functional Components

The PCE Discovery mechanisms ([RFC5089] and [RFC5088]) may be used to advertise WSON RWA path computation capabilities to PCCs.

3.6. Impact on Network Operation

Mechanisms defined in this document do not imply any new network operation requirements in addition to those already listed in section 8.6 of [RFC5440].

4. Security Considerations

This document has no requirement for a change to the security models within PCEP [RFC5440]. However the additional information distributed in order to address the RWA problem represents a disclosure of network capabilities that an operator may wish to keep private. Consideration should be given to securing this information.

5. IANA Considerations

A future revision of this document will present requests to IANA for codepoint allocation.

6. Acknowledgments

The authors would like to thank Adrian Farrel for many helpful comments that greatly improved the contents of this draft.

This document was prepared using 2-Word-v2.0.template.dot.

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.
- [RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.
- [RFC4657] Ash, J. and J. Le Roux, "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, September 2006.
- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) communication Protocol", RFC 5440, March 2009.
- [PCEP-MIB] Koushik, K, et al., "PCE communication protocol(PCEP) Management Information Base", draft-ietf-pce-pcep-mib, work in progress.

7.2. Informative References

- [WSON-IMP] Lee, Y. and Bernstein, G. (Editors), D. Li and G. Martinelli "A Framework for the Control and Measurement of Wavelength Switched Optical Networks (WSON) with Impairments, draft-ietf-ccamp-wson-impairments, work in progress.
- [WSON-Frame] Lee, Y. and Bernstein, G. (Editors), and W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks", draft-ietf-ccamp-rwa-wson-framework, work in progress.
- [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.

Authors' Addresses

Young Lee (Ed.)
Huawei Technologies
1700 Alma Drive, Suite 100
Plano, TX 75075, USA
Phone: (972) 509-5599 (x2240)
Email: ylee@huawei.com

Greg Bernstein (Ed.)
Grotto Networking
Fremont, CA, USA
Phone: (510) 573-2237
Email: gregb@grotto-networking.com

Jonas Martensson
Acreo
Email: Jonas.Martensson@acreo.se

Tomonori Takeda
NTT Corporation
3-9-11, Midori-Cho
Musashino-Shi, Tokyo 180-8585, Japan
Email: takeda.tomonori@lab.ntt.co.jp

Takehiro Tsuritani
KDDI R&D Laboratories, Inc.
2-1-15 Ohara Kamifukuoka Saitama, 356-8502. Japan
Phone: +81-49-278-7357
Email: tsuri@kddilabs.jp

Intellectual Property Statement

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.

Disclaimer of Validity

All IETF Documents and the information contained therein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.

Network Working Group
Internet Draft
Intended status: Informational
Expires: April 2015

Y. Lee
Huawei
G. Bernstein
Grotto Networking
Jonas Martensson
Acreo
T. Takeda
NTT
T. Tsuritani
KDDI
O. G. de Dios
Telefonica

October 28, 2014

PCEP Requirements for WSON Routing and Wavelength Assignment

draft-ietf-pce-wson-routing-wavelength-15.txt

Abstract

This memo provides application-specific requirements for the Path Computation Element communication Protocol (PCEP) for the support of Wavelength Switched Optical Networks (WSON). Lightpath provisioning in WSONs requires a routing and wavelength assignment (RWA) process. From a path computation perspective, wavelength assignment is the process of determining which wavelength can be used on each hop of a path and forms an additional routing constraint to optical light path computation. Requirements for PCEP extensions in support of optical impairments will be addressed in a separate document.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/ietf/lid-abstracts.txt>

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

This Internet-Draft will expire on April 28, 2009.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

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 [RFC2119].

Table of Contents

1. Introduction.....	3
2. WSON RWA Processes & Architecture.....	4
3. Requirements.....	6
3.1. Path Computation Type Option.....	6
3.2. RWA Processing.....	6
3.3. Bulk RWA Path Request/Reply.....	7
3.4. RWA Path Re-optimization Request/Reply.....	7
3.5. Wavelength Range Constraint.....	8
3.6. Wavelength Assignment Preference.....	8
3.7. Signal Processing Capability Restriction.....	8
4. Manageability Considerations.....	9
4.1. Control of Function and Policy.....	9
4.2. Information and Data Models, e.g. MIB module.....	9
4.3. Liveness Detection and Monitoring.....	10
4.4. Verifying Correct Operation.....	10

4.5. Requirements on Other Protocols and Functional Components	10
4.6. Impact on Network Operation	10
5. Security Considerations	10
6. IANA Considerations	11
7. Acknowledgments	11
8. References	11
8.1. Normative References	11
8.2. Informative References	11
Authors' Addresses	12
Intellectual Property Statement	13
Disclaimer of Validity	13

1. Introduction

[RFC4655] defines the PCE-based architecture and explains how a Path Computation Element (PCE) may compute Label Switched Paths (LSP) in Multiprotocol Label Switching Traffic Engineering (MPLS-TE) and Generalized MPLS (GMPLS)-controlled networks at the request of Path Computation Clients (PCCs). A PCC is shown to be any network component that makes such a request and may be for instance an optical switching element within a Wavelength Division Multiplexing (WDM) network. The PCE, itself, can be located anywhere within the network, and may be within an optical switching element, a Network Management System (NMS) or Operational Support System (OSS), or may be an independent network server.

The PCE communication Protocol (PCEP) is the communication protocol used between PCC and PCE, and may also be used between cooperating PCEs. [RFC4657] sets out the common protocol requirements for PCEP. Additional application-specific requirements for PCEP are deferred to separate documents.

This document provides a set of application-specific PCEP requirements for support of path computation in Wavelength Switched Optical Networks (WSON). WSON refers to WDM-based optical networks in which switching is performed selectively based on the wavelength of an optical signal.

The path in WSON is referred to as a lightpath. A lightpath may span multiple fiber links and the path should be assigned a wavelength for each link.

A transparent optical network is made up of optical devices that can switch but not convert from one wavelength to another. In a transparent optical network, a lightpath operates on the same wavelength across all fiber links that it traverses. In such case, the lightpath is said to satisfy the wavelength-continuity

constraint. Two lightpaths that share a common fiber link cannot be assigned the same wavelength. To do otherwise would result in both signals interfering with each other. Note that advanced additional multiplexing techniques such as polarization based multiplexing are not addressed in this document since the physical layer aspects are not currently standardized. Therefore, assigning the proper wavelength on a lightpath is an essential requirement in the optical path computation process.

When a switching node has the ability to perform wavelength conversion the wavelength-continuity constraint can be relaxed, and a lightpath may use different wavelengths on different links along its path from origin to destination. It is, however, to be noted that wavelength converters may be limited for cost reasons, while the number of WDM channels that can be supported in a fiber is also limited. As a WSON can be composed of network nodes that cannot perform wavelength conversion, nodes with limited wavelength conversion, and nodes with full wavelength conversion abilities, wavelength assignment is an additional routing constraint to be considered in all lightpath computations.

In this document we first review the processes for routing and wavelength assignment (RWA) used when wavelength continuity constraints are present and then specify requirements for PCEP to support RWA. Requirements for optical impairments will be addressed in a separate document.

The remainder of this document uses terminology from [RFC4655].

2. WSON RWA Processes & Architecture

In [RFC6163] three alternative process architectures were given for performing routing and wavelength assignment. These are shown schematically in Figure 1. R stands for Routing, WA for Wavelength Assignment, and DWA for Distributed Wavelength Assignment.

Figure 1. RWA process alternatives

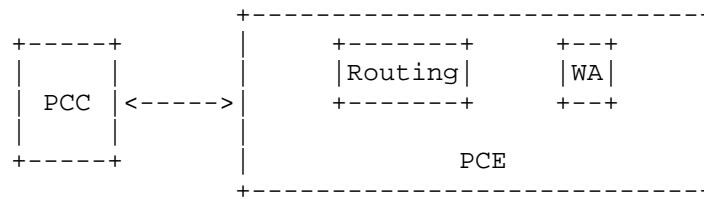


Figure 2. Combined Process (R&WA) architecture

3. Requirements

The requirements for the PCC to PCE interface of Figure 2 are specified in this section.

3.1. Path Computation Type Option

A PCEP request MAY include the path computation type. This can be:

- (i) Both Routing and Wavelength Assignment (RWA),
- (ii) Routing only.

This requirement is needed to differentiate between the currently supported routing with distributed wavelength assignment option and combined RWA. In case of distributed wavelength assignment option, wavelength assignment will be performed at each node of the route.

3.2. RWA Processing

- (a) When the request is a RWA path computation type, the request MUST further include the wavelength assignment options. At the minimum, the following option should be supported:

- (i) Explicit Label Control (ELC) [RFC3473]
- (ii) A set of recommended labels for each hop. The PCC can select the label based on local policy.

Note that option (ii) may also be used in R+WA or R+DWA.

- (b) In case of a RWA computation type, the response MUST include the wavelength(s) assigned to the path and an indication of which label assignment option has been applied (ELC or label set).

- (c) In the case where a valid path is not found, the response MUST include why the path is not found (e.g., network disconnected, wavelength not found, or both, etc.). Note that 'wavelength not found' may include several sub-cases such as wavelength continuity not met, unsupported FEC/Modulation type, etc.

3.3. Bulk RWA Path Request/Reply

Sending simultaneous path requests for "routing only" computation is supported by PCEP specification [RFC5440]. To remain consistent the following requirements are added.

- (a) A PCEP request MUST be able to specify an option for bulk RWA path request. Bulk path request is an ability to request a number of simultaneous RWA path requests.
- (b) The PCEP response MUST include the path and the assigned wavelength assigned for each RWA path request specified in the original bulk request.

3.4. RWA Path Re-optimization Request/Reply

1. For a re-optimization request, the request MUST provide both the path and current wavelength to be re-optimized and MAY include the following options:
 - a. Re-optimize the path keeping the same wavelength(s)
 - b. Re-optimize wavelength(s) keeping the same path
 - c. Re-optimize allowing both the wavelength and the path to change
2. The corresponding response to the re-optimized request MUST provide the re-optimized path and wavelengths even when the request asked for the path or the wavelength to remain unchanged.
3. In case that the new path is not found, the response MUST include why the path is not found (e.g., network disconnected, wavelength not found, or both, etc.). Note that 'wavelength not found' may include several sub-cases such as wavelength continuity not met, unsupported FEC/Modulation type, etc.

3.5. Wavelength Range Constraint

For any RWA computation type request, the requester (PCC) MUST be allowed to specify a restriction on the wavelengths to be used. The requester MAY use this option to restrict the assigned wavelength for explicit label or label set. This restriction may for example come from the tuning ability of a laser transmitter, any optical element, or a policy-based restriction.

Note that the requester (e.g., PCC) is not required to furnish any range restrictions.

3.6. Wavelength Assignment Preference

1. A RWA computation type request MAY include the requester preference for, e.g., random assignment, descending order, ascending order, etc. A response SHOULD follow the requestor preference unless it conflicts with operator's policy.
2. A request for two or more paths MUST allow the requester to include an option constraining the paths to have the same wavelength(s) assigned. This is useful in the case of protection with single transponder (e.g., 1+1 link disjoint paths).

In a network with wavelength conversion capabilities (e.g. sparse 3R regenerators), a request SHOULD be able to indicate whether a single, continuous wavelength should be allocated or not. In other words, the requesting PCC SHOULD be able to specify the precedence of wavelength continuity even if wavelength conversion is available.

3.7. Signal Processing Capability Restriction

Signal processing compatibility is an important constraint for optical path computation. The signal type for an end-to-end optical path must match at source and at destination.

The PCC MUST be allowed to specify the signal type at the endpoints (i.e., at source and at destination). The following signal processing capabilities should be supported at a minimum:

- o Modulation Type List
- o FEC Type List

The PCC MUST also be allowed to state whether transit modification is acceptable for the above signal processing capabilities.

4. Manageability Considerations

Manageability of WSON Routing and Wavelength Assignment (RWA) with PCE must address the following considerations:

4.1. Control of Function and Policy

In addition to the parameters already listed in Section 8.1 of [RFC5440], a PCEP implementation SHOULD allow configuring the following PCEP session parameters on a PCC:

- o The ability to send a WSON RWA request.

In addition to the parameters already listed in Section 8.1 of [RFC5440], a PCEP implementation SHOULD allow configuring the following PCEP session parameters on a PCE:

- o The support for WSON RWA.
- o The maximum number of bulk path requests associated with WSON RWA per request message.

These parameters may be configured as default parameters for any PCEP session the PCEP speaker participates in, or may apply to a specific session with a given PCEP peer or a specific group of sessions with a specific group of PCEP peers.

4.2. Information and Data Models, e.g. MIB module

As this document only concerns the requirements to support WSON RWA, no additional MIB module is defined in this document. However, the corresponding solution draft will list the information that should be added to the PCE MIB module defined in [PCEP-MIB].

4.3. Liveness Detection and Monitoring

No new mechanism is defined in this document that implies any new liveness detection and monitoring requirements in addition to those already listed in section 8.3 of [RFC5440].

4.4. Verifying Correct Operation

No new mechanism is defined in this document that implies any new verification requirements in addition to those already listed in section 8.4 of [RFC5440]

4.5. Requirements on Other Protocols and Functional Components

If PCE discovery mechanisms ([RFC5089] and [RFC5088]) were to be extended for technology-specific capabilities, advertising WSON RWA path computation capability should be considered.

4.6. Impact on Network Operation

No new mechanism is defined in this document that implies any new network operation requirements in addition to those already listed in section 8.6 of [RFC5440].

5. Security Considerations

This document has no requirement for a change to the security models within PCEP [RFC5440]. However the additional information distributed in order to address the RWA problem represents a disclosure of network capabilities that an operator may wish to keep private. Consideration should be given to securing this information.

Solutions that address the requirements in this document need to verify that existing PCEP security mechanisms adequately protect the additional network capabilities and must include new mechanisms as necessary.

6. IANA Considerations

This informational document does not make any requests for IANA action.

7. Acknowledgments

The authors would like to thank Adrian Farrel, Cycil Margaria and Ramon Casellas for many helpful comments that greatly improved the contents of this draft.

This document was prepared using 2-Word-v2.0.template.dot.

8. References

8.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.
- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) communication Protocol", RFC 5440, March 2009.

8.2. Informative References

- [RFC3473] L. Berger, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC4657] Ash, J. and J. Le Roux, "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, September 2006.

- [RFC6163] Y. Lee, G. Bernstein, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks", RFC 6163, April 2011.
- [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.
- [PCEP-MIB] Koushik, K, et al., "PCE communication protocol(PCEP) Management Information Base", draft-ietf-pce-pcep-mib, work in progress.

Authors' Addresses

Young Lee (Ed.)
Huawei Technologies
5340 Legacy Drive, Building 3
Plano, TX 75245, USA
Phone: (469)277-5838
Email: leeyoung@huawei.com

Greg Bernstein (Ed.)
Grotto Networking
Fremont, CA, USA
Phone: (510) 573-2237
Email: gregb@grotto-networking.com

Jonas Martensson
Acreo
Email:Jonas.Martensson@acreo.se

Tomonori Takeda
NTT Corporation
3-9-11, Midori-Cho
Musashino-Shi, Tokyo 180-8585, Japan
Email: takeda.tomonori@lab.ntt.co.jp

Takehiro Tsuritani
KDDI R&D Laboratories, Inc.
2-1-15 Ohara Kamifukuoka Saitama, 356-8502. Japan
Phone: +81-49-278-7357
Email: tsuri@kddilabs.jp

Oscar Gonzalez de Dios
Telefonica Investigacion y Desarrollo
C/ Emilio Vargas 6
Madrid, 28043
Spain
Phone: +34 91 3374013
Email: ogondio@tid.es

OSPF
Internet-Draft
Intended status: Standards Track
Expires: January 10, 2012

W. Lu
Ericsson
July 9, 2011

OSPF TE Extension for Area IDs
draft-lu-ospf-area-tlv-01

Abstract

For multi-area path computation, it is desirable to have knowledge of area boundaries and the corresponding border routers which are capable of processing inter-area TE traffic. This memo defines a TLV to the OSPF TE extensions to meet such need.

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 <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 January 10, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
1.1. Background	3
1.2. Current Solutions	3
1.2.1. Global TED	3
1.2.2. Stitch	3
1.2.3. Crankback	4
1.2.4. Distributed Path Computation	4
1.3. Requirements Language	4
1.4. Acronyms	4
2. Definitions	5
2.1. Exit Area	5
2.2. TE-ABR	5
3. Motivation	6
3.1. The importance of ABR whereabouts	6
3.2. Topic not Yet Addressed	6
3.3. Benefits	7
4. Scope of the proposal	7
5. Area ID TLV	7
5.1. TLV Origination Point	8
5.2. TLV encoding	9
5.3. Example	9
6. Applications	10
6.1. Use-Case 1	10
6.1.1. Crankback approach	10
6.1.2. BRPC approach	10
6.2. Use-Case 2	11
7. Acknowledgements	12
8. IANA Considerations	12
9. Security Considerations	12
10. References	12
10.1. Normative References	12
10.2. Informative References	12
Author's Address	13

1. Introduction

1.1. Background

Traffic Engineering (TE) based networks are widely used by network operators. The provision and setup mechanisms work fine in a single IGP area thanks to the well defined TE extensions to the corresponding protocols, namely RSVP-TE [RFC3209], OSPF-TE [RFC3630], and ISIS-TE [RFC3784]. From the single area TE database, LSPs can be derived to meet various TE constraints using some Path Computation Element (PCE) methods such as CSPF.

The mechanisms however cannot be applied directly to multi-area networks, for which the path computation is one of the key applications of the PCE-based architecture [RFC4655].

It is highly desirable to compute inter-area shortest paths that satisfy some bandwidth constraints or any other constraints, with little manual intervention, as is possible within a single IGP area.

1.2. Current Solutions

Listed below are a few existing inter-area path computation mechanisms. As can be seen the ABR whereabouts are indispensable in computing inter-area LSPs.

1.2.1. Global TED

A single TE database that contains all TE information of each and every area/domain is called the global TED. This certainly makes it easy to compute shortest path LSPs that meet all constraint requirements. The drawbacks nevertheless are apparent:

- a. IGP hierarchy enables improved IGP scalability by dividing the IGP domain into areas and limiting the flooding scope of topology information to within area boundaries. The global TED goes against this principle.
- b. Even if one is willing to compromise this principle, the LSPs created upon this global TED would lack of area information which may be required for enforcing path selection policies.

1.2.2. Stitch

This method uses per-area path computation based on ERO expansion on the head-end LSR and on ABRs. ABRs can be selected through either:

- a. Static configuration of ABRs as loose hops at the head-end LSR;
- b. Dynamic ABR selection - Proprietary, knowledge of ABRs may be acquired through non-standard protocol modification.

1.2.3. Crankback

Crankback method defined in [RFC4920] allows an LSP to be constructed to beyond the area scope provided some intermediate nodes, i.e. ABRs, are known. Crankback can probe the ABRs one after another till a viable path is found if it exists.

Note that this method does not allow computing an optimal path but just a feasible path. It may also have some non-negligible setup delay. These issues nevertheless are beyond the scope of this document.

1.2.4. Distributed Path Computation

PCE architecture document [RFC4655] outlines a distributed PCE architecture. The idea is that various PCEs which have partial information of the topologies work together to conclude the best paths that meet the computation constraint requirements.

Individual PCEs may not have any visibility beyond the areas they are servicing. For inter-area path computation purpose, the knowledge of ABRs are essential for the neighboring PCEs to find out the overall best path over the combined areas.

BRPC as defined in [RFC5441] provides one of such methods. OSPF-CAP [RFC5088] on the other hand defines methods to make known distributed PCEs using OSPF capability TLVs.

Note that although [RFC5088] contains PCE-DOMAIN sub-TLV for OSPF Area ID, it however cannot be used for identifying area border LSRs. It is for locating PCEs, not LSRs.

1.3. Requirements Language

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

1.4. Acronyms

ABR - Area Border Router
BRPC - Backward-Recursive Path Computation
CSPF - Constrained Shortest Path First
IGP - Interior Gateway Protocol
LSA - Link State Advertisement
LSDB - Link-State DataBase
LSP - Label Switched Path
LSR - Label Switching Router
OSPF - Open Shortest Path First
PCC - Path Computation Client
PCE - Path Computation Element
TE - Traffic Engineering
TED - Traffic Engineering Database
TLV - Type Length Value
VSPT - Virtual Shortest Path Tree

2. Definitions

The following definitions are under the context where TE is enabled.

2.1. Exit Area

A neighboring OSPF area to which an inter-area path can possibly be extended is called Exit Area.

The ABR which connects the current area to an Exit Area is called exit ABR.

2.2. TE-ABR

An exit ABR is also referred to as TE-ABR.

3. Motivation

3.1. The importance of ABR whereabouts

All solutions listed in Section 1.2, except the global TED approach, have to use the knowledge of exit ABRs to accomplish their inter-area path computation tasks.

Some methods require manual configuration which is costly and error prone. Others may have to use proprietary means to acquire the information.

It is desirable to have the exit ABR information available and make it conveniently accessible to the relevant PCEs without adding lot of complexity to the protocols nor too much burden to the participating routers.

3.2. Topic not Yet Addressed

Apart from the manual provision, currently the exit ABR information is difficult to acquire. The reasons are:

1. OSPF TE Database (TED) does not contain information on exit ABRs;
2. CSPF operates on the TED and is therefore limited to the information the latter provides. Unless the critical information of the exit ABRs becomes available, the CSPF cannot operate optimally by seeing beyond the area scope.
3. One may argue that CSPF can dig into OSPF's other repositories, such as LSDB, to find out the ABR whereabouts. This is not advisable because it negates the purpose of keeping TED opaque and independent from the normal OSPF operation. This is also technically difficult because usually CSPF and OSPF are two different processes that they may not be running in the same nodes.
4. Even if one is willing for CSPF to intrude into OSPF space, and use the ABR bit (B bit) information for locating border routers, these ABRs are not necessarily the exit ABRs. In other words, even if CSPF learns which routers sit on area borders, it is still unable to ascertain whether the ABRs are supporting TE on the other side of the borders. OSPF's hierarchical design limits the topology sharing to within area boundaries.
5. And even if one is willing to jump across this limit and somehow manages to acquire the ABRs' TE ability onto other areas, there is still the need for one more key information, the Area IDs.

6. The Area IDs are critical to the distributed PCE architecture. They are essential in enforcing path computation area sequences and PCE policies. They are also useful for consolidating path computation jobs. Section 6.2 provides an example of ABR consolidation.
7. It is important to understand that the proposed TLV also implies TE capabilities. In other words the TLV signifies that the advertising router is not only an ABR, but also an LSR capable of handling TE traffic. Unless tied with TE knowledge, methods of discovering of ABRs will not be useful in locating TE-ABRs, i.e. the LSRs that can transit TE traffic across areas.
8. At last, since the TLV is TE based, it should be defined in the OSPF TE extensions and maintained similarly with its counterparts, Router Address TLV and Link TLV.

3.3. Benefits

The benefits of having an OSPF Area ID TLV are listed below:

1. It fits into OSPF TE architecture, preserves the protocol's hierarchy, and adds little burden to the OSPF process;
2. It automates the TE-ABR discovery, and eliminates the need of manual provisioning;
3. Very often an LSR is also a PCE. If this LSR is also an ABR, it can compute a two-area LSP effortlessly. The PCC only needs to send the request to one TE-ABR (if there are multiple TE-ABRs sharing the same border), provided it has knowledge of the TE-ABR whereabouts.

4. Scope of the proposal

This document describes solutions for inter-area path computation and does not address inter-domain scenarios.

It is also specific to OSPF as an IGP protocol, though the concepts apply to ISIS which are to be defined in a separate document.

5. Area ID TLV

[RFC3630] section 2.4 defines two TLVs. This memo adds a third TLV called the Area ID TLV.

5.2. TLV encoding

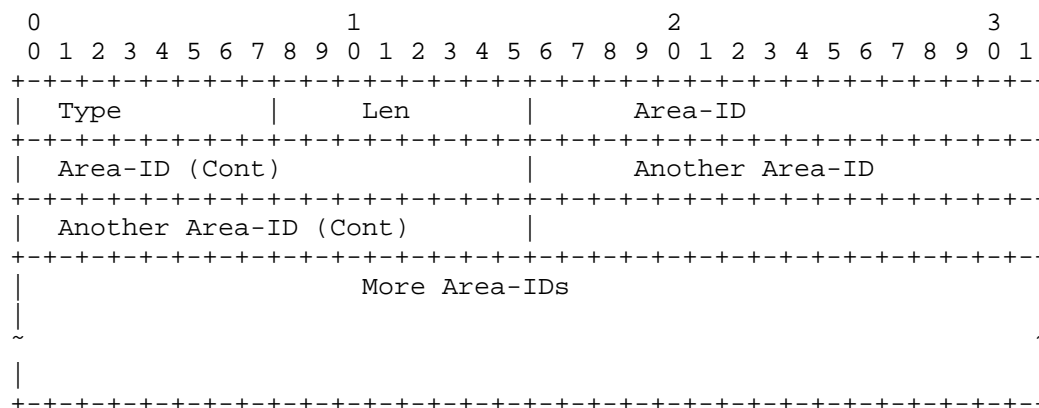


Figure 2: TLV Format

- Type** The Area ID TLV consists of a 1-octet type value which will be allocated by the IANA (suggested value 3).
- Len** 1-octet unsigned integer value 4xN to represent N exit areas, not including the originating area.
- Value** The value fields contains one or more Area IDs. Each Area ID is a four-octet OSPF Area ID associated with an exit area for which the ABR has TE enabled. An ABR may connect to multiple areas. Therefore it may generate 1 TLV with N IDs, where N is the total number of exit areas, not including the originating area. Since the Len field is of 1-octet, this TLV can hold upto 63 IDs. For non-ABR routers this Area-ID TLV SHOULD not occur.

5.3. Example

Consider Figure 1 again. Router ABR3 connects to four areas Area0, Area2, Area3, and Area4. Assuming all ABRs and areas are TE enabled except Area4 which is not TE-enabled. ABR3 originates and floods following Area-ID TLVs as shown in Figure 3, assuming type is 3.

To	Type	Len	Area-IDs	...
Area0	3	8	0 0 0 2	0 0 0 3
Area2	3	8	0 0 0 0	0 0 0 3
Area3	3	8	0 0 0 0	0 0 0 2
Area4	None			

Figure 3: Sample TLV

6. Applications

6.1. Use-Case 1

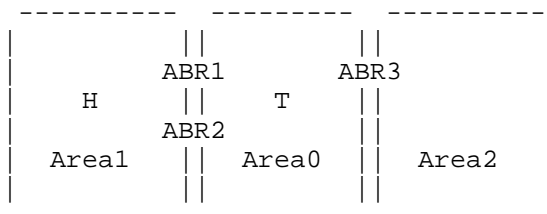


Figure 4: Use-Case 1

The topology is shown in Figure 4. The headend "H" is in Area1. The tailend "T" is in Area0. LSRs are running OSPF-TE and RSVP-TE.

6.1.1. Crankback approach

The crankback method requires a list of ABRs for tryout. The list usually has to be provisioned manually.

With the proposed Area-ID TLVs, the ABR information is made available through the OSPF TE database. Therefore "H" learns the ABR list dynamically from its OSPF TED, which is ABR1 and ABR2. "H" starts the crankback procedure with "ABR1" from which it reaches "T". The LSP thus is established successfully, though not necessarily optimal.

6.1.2. BRPC approach

With the method defined in [RFC5441], the VSPT has to be first built from "T" in Area0. Given that the area sequence is Area0->Area1, the initial VSPT has the candidate exit ABRs ABR1, ABR2, and ABR3.

Now with the proposed OSPF Area ID TLVs the PCE of Area0 knows that ABR3 connects Area2, not Area1, therefore ABR3 is not considered as an exit ABR.

Subsequently the PCE of Area1 takes the VSPT passed by the PCE of Area0 as input and concludes the end-to-end path computation.

6.2. Use-Case 2

Very often an LSR is also a PCE. In that case, the Area ID TLVs can make the path computation job more efficient.

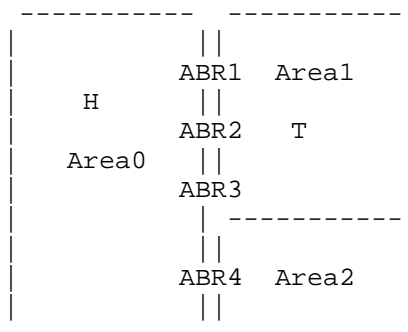


Figure 5: Use-Case 2

As shown in Figure 5, if the headend "H" knows that the tailend "T", or some intermediate node to reach, is in Area1, it sends a PCReq to one of the ABRs that connect to Area1. The draft proposal makes the ABR list conveniently available, which is [ABR1, ABR2, ABR3]. Note that ABR4 is not listed since it is not an exit ABR to Area1.

Assuming by some local policy ABR1 is chosen. Since ABR1 sits across Area0 and Area1, it has visibility to TEDs of both areas. ABR1 which is also a PCE performs the path computation job in the same way as an intra-area path computation.

Note that the resulting path, if existent, does not necessarily go through ABR1. It can be for example H->ABR3->T.

Note also that the ABR selection is a local decision. One can use some criteria, for example the highest te-router-id, to select the ABR. However, the candidate ABRs have to share the common border such as the one between Area0 and Area1. This can be achieved by grouping ABRs according to their exit Area IDs in the proposed OSPF Area ID TLVs.

7. Acknowledgements

The author would like to thank Sriganesh Kini, Meral Shirazipour, and Dimitri Papadimitriou for their reviews and comments.

8. IANA Considerations

This document defines the following TLV to the OSPF TE Extensions under TE LSA:

Type	Name	Source
TBD (recommend 3)	Area ID TLV	This document

9. Security Considerations

There are no specific security considerations within the scope of this document.

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

10.2. 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.
- [RFC3784] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, June 2004.
- [RFC4105] Le Roux, J., Vasseur, J., and J. Boyle, "Requirements for

Inter-Area MPLS Traffic Engineering", RFC 4105, June 2005.

- [RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.
- [RFC4674] Le Roux, J., "Requirements for Path Computation Element (PCE) Discovery", RFC 4674, October 2006.
- [RFC4920] Farrel, A., Satyanarayana, A., Iwata, A., Fujita, N., and G. Ash, "Crankback Signaling Extensions for MPLS and GMPLS RSVP-TE", RFC 4920, July 2007.
- [RFC4927] Le Roux, J., "Path Computation Element Communication Protocol (PCECP) Specific Requirements for Inter-Area MPLS and GMPLS Traffic Engineering", RFC 4927, June 2007.
- [RFC5088] Le Roux, JL., Vasseur, JP., Ikejiri, Y., and R. Zhang, "OSPF Protocol Extensions for Path Computation Element (PCE) Discovery", RFC 5088, January 2008.
- [RFC5441] Vasseur, JP., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", RFC 5441, April 2009.

Author's Address

Wenhu Lu
Ericsson
300 Holger Way
San Jose, California 95134
USA

Phone: 408 750-5436
Email: wenhu.lu@ericsson.com

