Network Working Group
Internet Draft
Category: Standard Track
Expires: December 2006

J.L. Le Roux (Editor)
France Telecom

J.P. Vasseur (Editor)
Cisco System Inc.

Yuichi Ikejiri
NTT Communications

Raymond Zhang
BT Infonet

June 2006

IGP protocol extensions for Path Computation Element (PCE) Discovery

draft-ietf-pce-disco-proto-igp-02.txt

Status of this Memo

Abstract

There are various circumstances in which it is highly desirable for a Path Computation Client (PCC) to be able to dynamically and automatically discover a set of Path Computation Element(s) (PCE), along with some of information that can be used for PCE selection. When the PCE is an LSR participating to the IGP, or even a server participating passively to the IGP, a simple and efficient way for

PCE discovery consists of relying on IGP flooding. For that purpose this document defines OSPF and ISIS extensions for the advertisement of PCE Discovery information within an IGP area or within the entire routing domain.

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.

Table of Contents

1. Note

This document specifies new TLVs and sub-TLVs to be carried within the OSPF Router information LSA ([OSPF-CAP]) and ISIS Router Capability TLV ([ISIS-CAP]) respectively. Because this document does not introduce any new element of procedure it will be discussed within the PCE Working Group with a review of the OSPF and ISIS Working Groups. Furthermore, once stabilized, it may be decided to split the document in two documents addressing the OSPF and ISIS aspects respectively.

2. Terminology

Terminology used in this document

ABR: IGP Area Border Router (OSPF ABR or ISIS L1L2 router).

AS: Autonomous System.

ASBR: AS Border Router.

Domain: any collection of network elements within a common sphere of address management or path computational responsibility. Examples of domains include IGP areas and Autonomous Systems.

IGP Area: OSPF Area or ISIS level.

Intra-area TE LSP: A TE LSP whose path does not cross IGP area boundaries.

Intra-AS TE LSP: A TE LSP whose path does not cross AS boundaries.

Inter-area TE LSP: A TE LSP whose path transits through two or more IGP areas.

Inter-AS MPLS TE LSP: A TE LSP whose path transits through two or more ASes or sub-ASes (BGP confederations).

LSR: Label Switch Router.

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.

PCECP: Path Computation Element Communication Protocol.

TE LSP: Traffic Engineered Label Switched Path.

3. Introduction

[PCE-ARCH] describes the motivations and architecture for a PCE-based path computation model for MPLS and GMPLS TE LSPs. The model allows the separation of PCE from PCC (also referred to as non co-located PCE) and allows cooperation between PCEs. This relies on a communication protocol between PCC and PCE, and between PCEs. The requirements for such communication protocol can be found in [PCECP-REQ] and the communication protocol is defined in [PCEP].

The PCE architecture requires, of course, that a PCC be aware of the location of one or more PCEs in its domain, and also potentially of some PCEs in other domains, e.g. in case of inter-domain TE LSP computation.

A network may comprise a large number of PCEs with potentially distinct capabilities. In such context it would be highly desirable to have a mechanism for automatic and dynamic PCE discovery, which would allow PCCs to automatically discover a set of PCEs, along with additional information required for PCE selection, and to dynamically detect new PCEs or any modification of PCE information. Detailed requirements for such a PCE discovery mechanism are described in [PCE-DISC-REQ].

Moreover, it may also be useful to discover when a PCE experiences some processing congestion state and exits such state, in order for the PCCs to take some appropriate actions (e.g. redirect to another PCE). Note that the PCE selection algorithm is out of the scope of this document.

When PCCs are LSRs participating to the IGP (OSPF, ISIS), and PCEs are LSRs or a servers also participating to the IGP, an efficient mechanism for PCE discovery within an IGP routing domain consists of relying on IGP advertisements.

This document defines OSPF and ISIS extensions allowing a PCE participating in the IGP to advertise its location along with some information useful for PCE selection so as to satisfy dynamic PCE discovery requirements set forth in [PCE-DISC-REQ]. This document also defines extensions allowing a PCE participating to the IGP to advertise its potential processing congestion state.

Generic capability mechanisms have been defined in [OSPF-CAP] and [ISIS-CAP] for OSPF and ISIS respectively the purpose of which is to allow a router to advertise its capability within an IGP area or an entire routing domain. Such IGP extensions fully satisfy the aforementioned dynamic PCE discovery requirements.

This document defines two new TLVs (named the PCE Discovery (PCED) TLV and the PCE Status (PCES) TLV) for ISIS and OSPF, to be carried within the ISIS Capability TLV ([ISIS-CAP]) and the OSPF Router Information LSA ([OSPF-CAP]).

The PCE information advertised is detailed in section 4. Protocol extensions and procedures are defined in section 5 for OSPF and 6 for ISIS.

This document does not define any new OSPF or ISIS element of procedure but how the procedures defined in [OSPF-CAP] and [ISIS-CAP] should be used.

The routing extensions defined in this document allow for PCE discovery within an IGP Routing domain. Solutions for PCE discovery across AS boundaries are beyond the scope of this document, and for further study.

4. Overview

4.1. PCE Information

PCE information advertised within the IGP includes PCE Discovery Information and PCE Status information.

4.1.1. PCE Discovery Information

The PCE Discovery information is comprised of:

- The PCE location: This an IPv4 and/or IPv6 address that must be used to reach the PCE. It is RECOMMENDED to use addresses always reachable;

- The PCE inter-domain functions: this refers to the PCE path computation scope (i.e. inter-area, inter-AS, inter-layer);

- The PCE domain(s): This is the set of one or more domain(s) where the PCE has visibility and can compute paths;

- The PCE Destination domain(s): This is the set of one or more destination domain(s) towards which a PCE can compute paths;

- A set of general PCECP capabilities (e.g. support for request prioritization) and path computation specific capabilities (e.g. supported constraints, supported objective functions).

It may also contain optional elements to describe more complex capabilities.

PCE Discovery information is by nature a static information that does not change with PCE activity. Changes in PCE Discovery information may occur as a result of PCE configuration updates, PCE deployment/activation, PCE deactivation/suppression or PCE failure. Hence, this information is not expected to change frequently.

4.1.2. PCE Status Information

The PCE Status is optional information that can be used to report a PCE processing congested state along with an estimated congestion duration. This is a dynamic information, which may change with PCE activity.

Procedures for a PCE to move from a processing congested state to a non congested state are beyond the scope of this document, but the rate at which a PCE Status change is advertised MUST not impact by any mean the IGP scalability. Particular attention should be given on procedures to avoid state oscillations.

4.2. Flooding scope

The flooding scope for PCE Discovery Information can be limited to one or more IGP areas the PCE belongs to or can be extended across the entire routing domain.
Note that some PCEs may belong to multiple areas, in which case the flooding scope may comprise these areas. This could be the case of an ABR for instance advertising its PCE information within the backbone area and/or a subset of its attached IGP area(s).

5. OSPF extensions

5.1. The OSPF PCED TLV

The OSPF PCE Discovery TLV (PCED TLV) is made of a set of non-ordered sub-TLVs.

The format of the OSPF PCED TLV and its sub-TLVs is the identical as the TLV format used by the Traffic Engineering Extensions to OSPF [OSPF-TE]. 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 four-octet alignment; padding is not included in the length field (so a three octet value would have a length of three, but the total size of the TLV would be eight octets). Nested TLVs are also 32-bit aligned. Unrecognized types are ignored. All types between 32768 and 65535 are reserved for vendor-specific extensions. All other undefined type codes are reserved for future assignment by IANA.

The OSPF PCED TLV has the following format:

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                            sub-TLVs                          //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+


      Type     To be defined by IANA (suggested value=2)
      Length   Variable
      Value    This comprises one or more sub-TLVs
```

Sub-TLVs types are under IANA control.

Currently five sub-TLVs are defined (type values to be assigned by IANA):

```
      Sub-TLV type  Length               Name
            1      variable     PCE-ADDRESS sub-TLV
            2         4         PATH-SCOPE sub-TLV
            3      variable     PCE-DOMAINS sub-TLV
            4      variable     PCE-DEST-DOMAINS sub-TLV
            5      variable     GENERAL-CAP sub-TLV
            6      variable     PATH-COMP-CAP sub-TLV
```

The sub-TLVs PCE-ADDRESS and PATH SCOPE MUST always be present within the PCED TLV.

The sub-TLVs PCE-DOMAINS and PCE-DEST-DOMAINS are optional. They MAY be present in some specific inter-domain cases.

The GENERAL-CAP and PATH-COMP-CAP sub-TLVs are optional and MAY be present in the PCED TLV to facilitate the PCE selection process.

Any non recognized sub-TLV MUST be silently ignored.

Additional sub-TLVs could be added in the future to advertise additional information.

The PCED TLV is carried within an OSPF Router Information LSA defined in [OSPF-CAP].

5.1.1. PCE-ADDRESS sub-TLV

The PCE-ADDRESS sub-TLV specifies the IP address(es) that MUST be used to reach the PCE. It is RECOMMENDED to make use of an address that is always reachable, provided that the PCE is alive.
The PCE-ADDRESS sub-TLV is mandatory; it MUST be present within the PCED TLV. It MAY appear twice, when the PCE has both an IPv4 and IPv6 address. It MUST not appear more than twice.

The format of the PCE-ADDRESS sub-TLV is as follows:

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     address-type              |           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                       PCE IP Address                        //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

PCE-ADDRESS sub-TLV format

```
  Type     To be assigned by IANA (suggested value =1)
  Length   4 (IPv4) or 16 (IPv6)

  Address-type:
                1   IPv4
                2   IPv6

  PCE IP Address: The IP address to be used to reach the PCE.
                  This is the address that will be used for
```

setting up PCC-PCE communication sessions.

5.1.2. PATH-SCOPE sub-TLV

The PATH-SCOPE sub-TLV indicates the PCE path computation scope(s), which refers to the PCE ability to compute or take part into the computation of intra-area, inter-area, inter-AS or inter-layer_TE LSP(s).

The PATH-SCOPE sub-TLV is mandatory; it MUST be present within the PCED TLV. There MUST be exactly one PATH-SCOPE sub-TLV within each PCED TLV.

The PATH-SCOPE sub-TLV contains a set of bit flags indicating the supported path scopes (intra-area, inter-area, inter-AS, inter-layer) and four fields indicating PCE preferences.

The PATH-SCOPE sub-TLV has the following format:

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|1|2|3|4|5|   Reserved        |PrefL|PrefR|PrefS|PrefY| Res   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
Type     To be defined by IANA (suggested value =3)
Length   Variable
Value    This comprises a 2 bytes flag where each bit
         represents a supported path scope, as well as four
         preference fields allowing to specify PCE preferences.
```

The following bits are defined:

```
Bit      Path Scope

 0      L bit: Can compute intra-area path
 1      R bit: Can act as PCE for inter-area TE LSPs
               computation
 2      Rd bit: Can act as a default PCE for inter-area TE LSPs
                computation
 3      S bit: Can act as PCE for inter-AS TE LSPs computation
 4      Sd bit: Can act as a default PCE for inter-AS TE LSPs
                computation
 5      Y bit: Can compute or take part into the computation of
               paths across layers.
```

Pref-L field: PCE's preference for intra-area TE LSPs computation.

Pref-R field: PCE s preference for inter-area TE LSPs computation.

Pref-S field: PCE s preference for inter-AS TE LSPs computation.

Pref-Y field: PCE's preference for inter-layer TE LSPs computation.

Res: Reserved for future usage.

The bits L, R, S and Y bits are set when the PCE can act as a PCE for intra-area, inter-area, inter-AS and inter-layer TE LSPs computation respectively. These bits are non exclusive.

When set the Rd bit indicates that the PCE can act as a default PCE for inter-area TE LSPs computation (the PCE can compute path for any destination area). Similarly, when set the Sd bit indicates that the PCE can act as a default PCE for inter-AS TE LSPs computation (the PCE can compute path for any destination AS).

When the Rd bit is set the PCE-DEST-DOMAIN TLV (see 5.1.4) does not contain any Area ID DOMAIN sub-TLV.

Similarly, when the Sd bit is set the PCE-DEST-DOMAIN TLV does not contain any AS DOMAIN sub-TLV.

The PrefL, PrefR, PrefS and PrefY fields are 3-bit long and allow the PCE to specify a preference for each computation scope, where 7 reflects the highest preference. Such preference can be used for weighted load balancing of requests. An operator may decide to configure a preference to each PCE so as to balance the path computation load among them, with respect to their respective CPU capacity. The algorithms used by a PCC to load balance its path computation requests according to such PCE s preference is out of the scope of this document. Same or distinct preferences may be used for different scopes. For instance an operator that wants a PCE capable of both inter-area and inter-AS computation to be used preferably for inter-AS computation may configure a PrefS higher than the PrefR.

When the L bit, R bit, S or Y bit are cleared, the PrefL, PrefR, PrefS, PrefY fields MUST respectively be set to 0.

5.1.3. PCE-DOMAINS sub-TLV

The PCE-DOMAINS sub-TLV specifies the set of domains (areas, AS) where the PCE has topology visibility and can compute paths. It contains a set of one or more sub-TLVs where each sub-TLV identifies a domain.

The PCED TLV MUST include zero or one PCE-DOMAINS sub-TLV.
The PCE-DOMAINS sub-TLV MUST be present when PCE domains cannot be inferred by other IGP information, for instance when the PCE is inter-area capable (i.e. when the R bit is set) and the flooding scope is the entire IGP routing domain.

The PCE-DOMAINS sub-TLV has the following format:

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                    DOMAIN sub-TLVs                          //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
      Type     To be defined by IANA (suggested value =3)
      Length   Variable
      Value    This comprises a set of one or more DOMAIN sub-TLVs
               where each DOMAIN sub-TLV identifies a domain where
               the PCE has topology visibility and can compute paths.
```

Sub-TLVs types are under IANA control.

Currently three DOMAIN sub-TLVs are defined (suggested type values to be assigned by IANA):

Sub-TLV type	Length	Name
1	variable	IPv4 area ID sub-TLV
2	variable	IPv6 area ID sub-TLV
3	variable	AS number sub-TLV

The PCE-DOMAINS sub-TLV MUST include at least one DOMAIN sub-TLV. Note than when the PCE visibility is an entire AS, the PCE-DOMAINS sub-TLV MUST uniquely include one AS number sub-TLV.

5.1.3.1. IPv4 area ID DOMAIN sub-TLV

The IPv4 area ID DOMAIN sub-TLV carries an IPv4 OSPF area identifier. It has the following format:

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         IPv4 Area ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
      Type     To be assigned by IANA (suggested value =1)
      Length   4
      IPv4 OSPF area ID: The IPv4 identifier of the OSPF area
```

5.1.3.2. IPv6 area ID DOMAIN sub-TLV

The IPv6 area ID sub-TLV carries an IPv6 OSPF area identifier. It has the following format:

```
     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            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                          IPv6 Area ID                         |
    |                                                               |
    |                                                               |
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
      Type     To be assigned by IANA (suggested value =2)
      Length   16
      IPv6 OSPF area ID: The IPv6 identifier of the OSPF area
```

5.1.3.3. AS Number sub-TLV

The AS Number sub-TLV carries an AS number. It has the following format:

```
     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            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                          AS Number                            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
      Type     To be assigned by IANA (suggested value =3)
      Length   4
      AS Number:  AS number identifying an AS. When coded on two
                  bytes (which is the current defined format as the
                  time of writing this document), the AS Number field
                  MUST have its left two bytes set to 0.
```

5.1.4. PCE-DEST-DOMAINS sub-TLV

The PCE-DEST-DOMAINS sub-TLV specifies the set of destination domains (areas, AS) toward which a PCE can compute path. It means that the PCE can compute or take part in the computation of inter-domain LSPs whose destinations are located within one of these domains. It contains a set of one or more sub-TLVs where each sub-TLV identifies a domain.

The PCE-DEST-DOMAINS sub-TLV has the following format:

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                   DOMAIN sub-TLVs                           //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
      Type     To be defined by IANA (suggested value =3)
      Length   Variable
      Value    This comprises a set of one or more Area and/or AS
               DOMAIN sub-TLVs where each DOMAIN sub-TLV identifies a
               domain toward which a PCE can compute paths.
```

The PCE-DEST-DOMAINS sub-TLV MUST be present if the R bit is set and the Rd bit is cleared, and/or, if the S bit is set and the Sd bit is cleared.

The PCE-DEST-DOMAINS sub-TLV MUST include at least one DOMAIN sub-TLV. It MUST include at least one area ID sub-TLV, if the R bit of the PATH-SCOPE TLV is set and the Rd bit of the PATH-SCOPE TLV is cleared. Similarly, it MUST include at least one AS number sub-TLV if the S bit of the PATH-SCOPE TLV is set and the Sd bit of the PATH-SCOPE TLV is cleared.

5.1.5. GENERAL-CAP sub-TLV

The GENERAL-CAP sub-TLV is an optional TLV used to indicate PCECP related capabilities. It MAY be present within the PCED TLV. It MUST not be present more than once.
The value field of the GENERAL-CAP sub-TLV is made of a 32-bit flag, where each bit corresponds to a general PCE capability. It MAY also include optional sub-TLVs to encode more complex capabilities.

The format of the GENERAL-CAP sub-TLV 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Type             |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 General Capabilities Flag                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
//                      Optional sub-TLVs                       //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
      Type      To be assigned by IANA (suggested value =1)
      Length    Variable.
      Value     This comprises a 32-bit flag. The bits are indexed
                from the most significant to the least significant,
                where each bit represents one general PCE capability.
                Optional TLVs may be added to specify more complex
                capabilities: there is no optional TLV currently
                defined.
```

IANA is requested to manage the space of the General Capabilities 32-bit flag.

The following bits are to be assigned by IANA:

```
  Bit       Capabilities

   0      P bit: Support for Request prioritization.
   1      M bit: Support for multiple messages within the same
                 request message.

  2-31    Reserved for future assignments by IANA.
```

5.1.6. The PATH-COMP-CAP sub-TLV

The PATH-COMP-CAP sub-TLV is an optional TLV used to indicate path computation specific capabilities. It MAY be present within the PCED TLV. It MUST not be present more than once.
It is made of a 32-bit flag, where each bit correspond to a path computation capability. It MAY also include optional sub-TLVs to encode more complex capabilities.

The format of the PATH-COMP-CAP sub-TLV 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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |              Type             |             Length            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |             Path Computation Capabilities  Flag               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   //                    Optional sub-TLVs                         //
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

      Type      To be assigned by IANA (suggested value =1)
      Length    Variable.
      Value     This comprises a 32 bit flag. Bits are indexed from
                the most significant to the least significant, where
                each bit represents one path computation capability.
```

Optional TLVs may be defined to specify more complex capabilities. Three optional sub-TLVs are currently defined.

IANA is requested to manage the space of the Path Commutation Capabilities 32-bit flag.

The following bits are to be assigned by IANA:

```
  Bit       Capabilities

   0      G bit: Capability to handle GMPLS link contraints
   1      B bit: Capability to compute bidirectional paths
   2      D bit: Capability to compute link/node/SRLG diverse paths
   3      L bit: Capability to compute load-balanced paths
   4      S bit: Capability to compute a set of paths in a
                 synchronized Manner
   5      O bit: Support for multiple objective functions
   6      P bit: Capability to handle path constraints (e.g. hop
          count, metric bound)

  7-31    Reserved for future assignments by IANA.
```

The G, B, D, L, S, O and P bits are not exclusive.

Three optional sub-TLVs are currently defined for the PATH-COMP-CAP TLV:

- The Objective Functions sub-TLV (type to be defined, suggested value =1) that carries a list of supported objective functions, where each objective function is identified by a 16 bit integer.
- The Opaque Objective Function sub-TLV (type to be defined, suggested value =2) that allows the user to encode a specific objective function in any appropriate language.
- The Switch Caps sub-TLV (type to be defined, suggested value =3) that carries a list of supported switching capabilities. It means that the PCE can compute path for the listed switching capabilities.

5.1.6.1. Objective Functions sub-TLV

The format of the Objective Functions sub-TLV 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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |              Type             |             Length            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |             function 1        |   function 2                  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

//   //

```
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |             function N        |                               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type	To be defined by IANA (suggested value =1)
Length	Variable (N*2), where N is the number of supported objective functions.
Value	This comprises a set of one or more 16 bit function ids, where each function id identifies a supported objective function.

Objectives functions and their identification will be defined in a separate document.

The Objective Functions sub-TLV is optional, it MAY be present with the PATH-COMP-CAP TLV. When present it MUST be present only once in the PATH-COMP-CAP TLV.

5.1.6.2. Opaque Objective Function sub-TLV

The format of the Opaque Objective Function sub-TLV 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
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |              Type             |             Length            |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |             Opaque objective function                         |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type	To be defined by IANA (suggested value =2)
Length	Variable
Value	This encode a specific objective function in any appropriate language.

The Opaque Objective Function sub-TLV is optional. The PATH-COMP-CAP TLV MAY comprise 0, one or more Opaque Objective Functions.

5.1.6.3. Switch Caps sub-TLV

The format of the Switch Caps sub-TLV 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
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |              Type             |             Length            |
```

```
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |   SC type     |   SC type     |   SC type     |               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      //                                                             //
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

         Type     To be defined by IANA (suggested value =3)
         Length   Variable = N, where N is the number of supported
                  switching capabilities
         Value    This comprises a set of one or more 8-bit switching
                  types, where each switching type identifies a
                  supported switching capability.
```

Switching type values are defined in [RFC4203].

The Objective function sub-TLV is optional, it MAY be present in the PATH-COMP-CAP TLV. When present it MUST be present only once in the PATH-COMP-CAP TLV.

5.2. The OSPF PCES TLV

The OSPF PCE Status TLV (PCES TLV) carries information related to PCE processing congestion state.
The PCES TLV is carried within an OSPF Router Information LSA which is defined in [OSPF-CAP].

The OSPF PCES TLV has the following format:

```
    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            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   //                      PCE ADDRESS sub-TLV                      //
   //                     CONGESTION sub-TLV                      //
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

         Type     To be defined by IANA (suggested value=3)
         Length   Variable
         Value    This comprises a PCE ADDRESS sub-TLV, identifying the
                  PCE and a CONGESTION sub-TLV that contains congestion
                  information.
```

Sub-TLV types are under IANA control.

Currently two sub-TLVs are defined (type values to be assigned by IANA):

```
         Sub-TLV type  Length               Name
               1      variable     PCE-ADDRESS sub-TLV
```

2   4   CONGESTION sub-TLV

The PCE-ADDRESS and CONGESTION sub-TLVs MUST be present once in a PCES TLV. The PCE-ADDRESS sub-TLV is defined in section 5.1.1. It carries one of the PCE IP addresses and is used to identify the PCE the processing congestion state information is applied to. This is required as the PCES and PCED TLVs may be carried in separate Router Information LSAs.

Any non recognized sub-TLV MUST be silently ignored.

Additional sub-TLVs could be added in the future to advertise additional congestion information.

5.2.1. The CONGESTION sub-TLV

The CONGESTION sub-TLV is used to indicate whether a PCE experiences a processing congestion state or not along with optionally the expected PCE congestion duration.
The CONGESTION sub-TLV is mandatory. It MUST be carried once within the PCES TLV.

The format of the CONGESTION sub-TLV 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Type             |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|C|       Reserved              |      Congestion Duration      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
Type     To be assigned by IANA (suggested value =2)
Length   4

Value
  -C bit: When set this indicates that the PCE experiences
          congestion and cannot support any new request. When
          cleared this indicates that the PCE does not
          experience congestion an can support a new request.

  -Congestion Duration: 2-bytes, the estimated PCE congestion
                        duration in seconds.
```

When C is set and the Congestion Duration field is equal to 0, this means that the Congestion Duration is unknown.
When C is cleared the Congestion Duration MUST be set to 0.

5.3. Elements of Procedure

The PCES and PCED TLV are advertised within an OSPFv2 Router Information LSA (Opaque type of 4 and Opaque ID of 0) or OSPFv3 Router information LSA (function code of 12) which are defined in [OSPF-CAP]. As such, elements of procedure are inherited from those defined in [OSPF-CAP].

As the PCES information is likely to change more frequently than the PCED information, it is RECOMMENDED to carry PCES and PCED TLVs in separate Router Information LSAs, so as not to carry all PCED information each time the PCE status changes.

In OSPFv2 the flooding scope is controlled by the opaque LSA type (as defined in [RFC2370]) and in OSPFv3 by the S1/S2 bits (as defined in [OSPF-v3]). If the flooding scope is local to an area then the PCED or PCES TLV MUST be carried within an OSPFv2 type 10 router information LSA or an OSPFV3 Router Information LSA with the S1 bit set and the S2 bit cleared. If the flooding scope is the entire domain then the PCED or PCES TLV MUST be carried within an OSPFv2 type 11 Router Information LSA or OSPFv3 Router Information LSA with the S1 bit cleared and the S2 bit set.
Note that when only the L bit of the PATH-SCOPE sub-TLV is set and the flooding scope MUST be local.
Note that the flooding scope of the PCED and PCES TLVs may be distinct, in which case they will be carried in separate LSA.

A router MUST originate a new OSPF router information LSA whenever the content of the PCED TLV or PCES TLV changes or whenever required by the regular OSPF procedure (LSA refresh (every LSRefreshTime)).

PCED and PCES sub-TLVs are OPTIONAL. When an OSPF LSA does not contain any PCED or PCES sub-TLV, this means that the PCE information of that node is unknown.

Note that a change in PCED information MUST not trigger any SPF computation.

The way PCEs retrieve their own information is out of the scope of this document. Some information may be configured on the PCE (e.g. address, preferences, scope) and other information may be automatically retrieved by the PCE (e.g. areas of visibility).

5.3.1. PCES TLV specific procedures

When a PCE enters into a processing congestion state, the conditions of which are implementation dependent, it SHOULD originate a Router Information LSA with a PCES TLV with the C bit set, and optionally a non-null expected congestion duration.

When a PCE leaves the processing congestion state, the conditions of which are implementation dependent, there are two cases:

- If the congestion duration in the previously originated PCES TLV was null, it SHOULD originate a PCES TLV with the C bit cleared and a null congestion duration;
- If the congestion duration in the previously originated PCES TLV was non null, it MAY originate a PCES TLV. Note that in some particular cases it may be desired to originate a PCES TLV with the C bit cleared if the saturation duration was over estimated.

The congestion duration allows reducing the amount of OSPF flooding, as only uncongested-congested state transitions are flooded.

An implementation SHOULD support an appropriate dampening algorithm so as to dampen OSPF flooding in order to not impact the OSPF scalability. It is RECOMMENDED to introduce some hysteresis for congestion state transition, so as to avoid state oscillations that may impact OSPF performances. For instance two thresholds MAY be configured: A resource saturation upper-threshold and a resource saturation lower-threshold. An LSR enters the congested state when the CPU load reaches the upper threshold and leaves the congested state when the CPU load goes under the lower threshold.

Upon receipt of an updated PCES TLV a PCC should take appropriate actions. In particular, the PCC SHOULD stop sending requests to a congested PCE, and SHOULD gradually start sending again requests to a no longer congested PCE.

6. ISIS extensions

6.1. IS-IS PCED TLV format

The IS-IS PCED TLV is made of various non ordered sub-TLVs.

The format of the IS-IS PCED TLV and its sub-TLVs is the same as the TLV format used by the Traffic Engineering Extensions to IS-IS [ISIS-TE]. That is, the TLV is composed of 1 octet for the type, 1 octet specifying the TLV length and a value field.

The IS-IS PCED TLV has the following format:

TYPE: To be assigned by IANA
LENGTH: Variable
VALUE: set of sub-TLVs

Sub-TLVs types are under IANA control.

Currently five sub-TLVs are defined (suggested type values to be assigned by IANA):

Sub-TLV type	Length	Name
1	variable	PCE-ADDRESS sub-TLV
2	3	PATH-SCOPE sub-TLV
3	variable	PCE-DOMAINS sub-TLV
4	variable	PCE-DEST-DOMAINS sub-TLV
5	variable	GENERAL-CAP sub-TLV
6	variable	PATH-COMP-CAP sub-TLV

The sub-TLVs PCE-ADDRESS and PATH-SCOPE MUST always be present within the PCED TLV.

The sub-TLVs PCE-DOMAINS and PCE-DEST-DOMAINS are optional. They MAY be present in some specific inter-domain cases.

The GENERAL-CAP and PATH-COMP-CAP are optional and MAY be present in the PCED TLV to facilitate the PCE selection process.

Any non recognized sub-TLV MUST be silently ignored.

Additional sub-TLVs could be added in the future to advertise additional PCE information.

The PCED TLV is carried within an ISIS CAPABILITY TLV defined in [ISIS-CAP], whose S bit is determined by the desired flooding scope.

6.1.1. PCE-ADDRESS sub-TLV

The PCE-ADDRESS sub-TLV specifies the IP address(es) that MUST be used to reach the PCE. It is RECOMMENDED to make use of an address that is always reachable, provided the PCE is alive.

The PCE-ADDRESS sub-TLV is mandatory; it MUST be present within the PCED TLV. It MAY appear twice, when the PCE has both an IPv4 and IPv6 address. It MUST not appear more than twice.

The PCE-ADDRESS sub-TLV has the following format:

```
   TYPE: To be assigned by IANA (Suggested value =1)
   LENGTH: 4 for IPv4 address and 16 for IPv6 address
   VALUE: This comprises one octet indicating the address-type and 4
          or 16 octets encoding the IPv4 or IPv6 address to be used
          to reach the PCE
```

```
Address-type:
               1   IPv4
               2   IPv6
```

The P in the PCED sub-LTV
originated by a PCE. It MAY appear multiple times, for instance when the PCE has both an IPv4 and IPv6 address. It MUST not appear more than two times.

6.1.2. The PATH-SCOPE sub-TLV

The PATH-SCOPE sub-TLV indicates the PCE path computation scope which refers to the PCE ability to compute or take part into the computation of intra-area, inter-area, inter-AS or inter-layer_TE LSP(s).

The PATH-SCOPE sub-TLV is mandatory; it MUST be present within the PCED TLV. There MUST be exactly one PATH-SCOPE sub-TLV within each PCED TLV.

The PATH-SCOPE sub-TLV contains a set of bit flags indicating the supported path scopes (intra-area, inter-area, inter-AS, inter-layer) and four fields indicating PCE preferences.

The PATH-SCOPE sub-TLV has the following format:

```
TYPE: To be assigned by IANA (Suggested value =2)
LENGTH: 3
VALUE: This comprises a one-byte flag of bits where each bit
```

represents a supported path scope, followed by a 2-bytes
preferences field indicating PCE preferences.

Here is the structure of the bit flag:

```
+-+-+-+-+-+-+-+-+
|0|1|2|3|4|5|Res|
+-+-+-+-+-+-+-+-+
```

```
Bit      Path Scope

0      L bit: Can compute intra-area path
1      R bit: Can act as PCE for inter-area TE LSPs
              computation
2      Rd bit: Can act as a default PCE for inter-area TE LSPs
               computation
3      S bit: Can act as PCE for inter-AS TE LSPs computation
4      Sd bit: Can act as a default PCE for inter-AS TE LSPs
               computation
5      Y bit: Can compute or take part into the computation of
              paths across layers
6-7          Reserved for future usage.
```

Here is the structure of the preferences field

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|PrefL|PrefR|PrefS|PrefY| Res   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Pref-L field: PCE's preference for intra-area TE LSPs computation.

Pref-R field: PCE s preference for inter-area TE LSPs computation.

Pref-S field: PCE s preference for inter-AS TE LSPs computation.

Pref-Y field: PCE's preference for inter-layer TE LSPs computation.

Res: Reserved for future usage.

The bits L, R, S and Y bits are set when the PCE can act as a PCE for intra-area, inter-area, inter-AS and inter-layer TE LSPs computation respectively. These bits are non exclusive.

When set the Rd bit indicates that the PCE can act as a default PCE for inter-area TE LSPs computation (the PCE can compute path for any destination area). Similarly, when set the Sd bit indicates that the PCE can act as a default PCE for inter-AS TE LSPs computation (the PCE can compute path for any destination AS).

When the Rd bit is set the PCE-DEST-DOMAIN TLV (see 5.1.4) does not contain any Area ID DOMAIN sub-TLV.

Similarly, when the Sd bit is set the PCE-DEST-DOMAIN TLV does not contain any AS DOMAIN sub-TLV.

The PrefL, PrefR, PrefS and PrefY fields are 3-bit long and allow the PCE to specify a preference for each computation scope, where 7 reflects the highest preference. Such preference can be used for weighted load balancing of requests. An operator may decide to configure a preference to each PCE so as to balance the path computation load among them, with respect to their respective CPU capacity. The algorithms used by a PCC to balance its path computation requests according to such PCE s preference is out of the scope of this document. Same or distinct preferences may be used for different scopes. For instance an operator that wants a PCE capable of both inter-area and inter-AS computation to be used preferably for inter-AS computation may configure a PrefS higher than the PrefR.

When the L bit, R bit, S or Y bit are cleared the PrefL, PrefR, PrefS, PrefY fields MUST respectively be set to 0.

6.1.3. PCE-DOMAINS sub-TLV

The PCE-DOMAINS sub-TLV specifies the set of domains (areas or AS) where the PCE has topology visibility and can compute paths. It contains a set of one or more sub-TLVs where each sub-TLV identifies a domain.

The PCE-DOMAINS sub-TLV MUST be present when PCE domains cannot be inferred by other IGP information, for instance when the PCE is inter-domain capable (i.e. when the R bit or S bit is set) and the flooding scope is the entire routing domain.

The PCE-DOMAINS sub-TLV has the following format:

```
TYPE: To be assigned by IANA (Suggested value =2)
LENGTH: Variable
VALUE: This comprises a set of one or more DOMAIN sub-TLVs where
       each DOMAIN sub-TLV identifies a domain where the PCE has
       topology visibility and can compute paths
```

DOMAIN Sub-TLVs types are under IANA control.

Currently two DOMAIN sub-TLVs are defined (suggested type values to be assigned by IANA):

Sub-TLV type	Length	Name
1	variable	Area ID sub-TLV
2	variable	AS number sub-TLV

At least one DOMAIN sub-TLV MUST be present in the PCE-DOMAINS sub-TLV.

6.1.3.1. Area ID DOMAIN sub-TLV

This sub-TLV carries an ISIS area ID. It has the following format

TYPE: To be assigned by IANA (Suggested value =1)
LENGTH: Variable
VALUE: This comprises a variable length ISIS area ID. This is the combination of an Initial Domain Part (IDP) and High Order part of the Domain Specific part (HO-DPS)

6.1.3.2. AS Number DOMAIN sub-TLV

The AS Number sub-TLV carries an AS number. It has the following format:

TYPE: To be assigned by IANA (Suggested value =2)
LENGTH: 4
VALUE: AS number identifying an AS. When coded on two bytes (which is the current defined format as the time of writing this document), the AS Number field MUST have its left two bytes set to 0.

6.1.4. PCE-DEST-DOMAINS sub-TLV

The PCE-DEST-DOMAINS sub-TLV specifies the set of destination domains (areas, AS) toward which a PCE can compute path. It means that the PCE can compute or take part in the computation of inter-domain LSPs whose destinations are located within one of these domains. It contains a set of one or more DOMAIN sub-TLVs where each DOMAIN sub-TLV identifies a domain.

The PCE-DEST-DOMAINS sub-TLV has the following format:

TYPE: To be assigned by IANA (Suggested value =3)
LENGTH: Variable
VALUE: This comprises a set of one or more Area or/and AS DOMAIN sub-TLVs where each sub-TLV identifies a destination domain toward which a PCE can compute path.

The PCE-DEST-DOMAINS sub-TLV MUST be present if the R bit is set and the Rd bit is cleared, and/or, if the S bit is set and the Sd bit is cleared.

The PCE-DEST-DOMAINS sub-TLV MUST include at least one DOMAIN sub-TLV. It MUST include at least one area ID sub-TLV, if the R bit of the PATH-SCOPE TLV is set and the Rd bit of the PATH-SCOPE TLV is cleared. Similarly, it MUST include at least one AS number sub-TLV if the S bit of the PATH-SCOPE TLV is set and the Sd bit of the PATH-

SCOPE TLV is cleared.

6.1.5. GENERAL-CAP sub-TLV

The GENERAL-CAP sub-TLV is an optional TLV used to indicate PCECP related capabilities. It carries a 32-bit flag, where each bit corresponds to a general PCE capability. It MAY also include optional sub-TLVs to encode more complex capabilities.

The GENERAL-CAP sub-TLV has the following format:

```
TYPE: To be assigned by IANA (Suggested value =4)
LENGTH: Variable
VALUE: This comprises a 32-bit General Capabilities flag where
       each bit corresponds to a general PCE capability, and
       optional sub-TLVs that may be defined to specify more
       complex capabilities. Currently no sub-TLVs are defined.
```

The following bits in the General Capabilities 32-bit flag are to be assigned by IANA:

```
 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

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|P|M|             Reserved                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

  Bit       Capabilities

   0      P bit: Support for Request prioritization.
   1      M bit: Support for multiple messages within the same
                 request message.

  2-31    Reserved for future assignments by IANA.
```

6.1.6. The PATH-COMP-CAP sub-TLV

The PATH-COMP-CAP sub-TLV is an optional TLV used to indicate path computation specific capabilities of a PCE.
This is a 32-bit flag, where each bit corresponds to a path computation capability. It MAY also include optional sub-TLVs to encode more complex capabilities.

The PATH-COMP-CAP sub-TLV has the following format:

```
TYPE: To be assigned by IANA (suggested value = 5)
LENGTH: Variable
VALUE: This comprises one 32 bit Path Computation Capabilities
       Flag, where each bit corresponds to a path computation
       capability, and optional sub-TLVs that may be defined to
```

specify more complex capabilities. Three optional sub-TLVs are currently defined.

The following bits in the Path Computation Capabilities 32-bit Flag are to be assigned by IANA:

```
 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

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
||M|G|D|L|S|0|P          Reserved                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
Bit       Capabilities

   0      G bit: Capability to handle GMPLS link contraints
   1      B bit: Capability to compute bidirectional paths
   2      D bit: Capability to compute link/node/SRLG diverse paths
   3      L bit: Capability to compute load-balanced paths
   4      S bit: Capability to compute a set of paths in a
                 synchronized Manner
   5      O bit: Support for multiple objective functions
   6      P bit: Capability to handle path constraints (e.g. hop
                 count, metric bound)

  7-31    Reserved for future assignments by IANA.
```

The G, B, D, L, S, O and P bits are not exclusive.

Three optional sub-TLVs are currently defined for the PATH-COMP-CAP TLV:

- The Objective Functions sub-TLV (type to be defined, suggested value =1) that carries a list of supported objective functions, where each objective function is identified by a 16 bit integer.
- The Opaque Objective Function sub-TLV (type to be defined, suggested value =2) that allows the user to encode a specific objective function in any appropriate language.
- The Switch Caps sub-TLV (type to be defined, suggested value =3) that carries a list of supported switching capabilities. This means that the PCE can compute paths for the listed switching capabilities.

6.1.6.1. Objective Functions sub-TLV

The format of the Objective Functions sub-TLV is as follows:

```
      TYPE:      To be defined by IANA (suggested value =1)
      LENGTH:    Variable (N*2)
      VALUE:     This comprises a set of one or more 16 bit function
                ids, where each function id identifies a supported
                objective functions.

   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |             function 1        |   function 2                  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   //                                                             //
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |             function N        |                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Objectives functions and their identification will be defined in a separate document.

The Objective Functions sub-TLV is optional. It MAY be present within the PATH-COMP-CAP TLV. When present it MUST be present only once in the PATH-COMP-CAP TLV.

6.1.6.2. Opaque Objective Function sub-TLV

The format of the Opaque Objective Function sub-TLV is as follows:

```
      TYPE:      To be defined by IANA (suggested value =2)
      LENGTH:    Variable
      VALUE:     This encode a specific objective function in any
                 appropriate language.

   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |             Opaque objective function                         |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The Opaque Objective function sub-TLV is optional. The PATH-COMP-CAP TLV MAY comprise 0, one or more Opaque Objective Functions.

6.1.6.3. Switch Caps sub-TLV

The format of the Switch Caps sub-TLV is as follows:

```
         TYPE     To be defined by IANA (suggested value =3)
         LENGTH   Variable = N, where N is the number of supported
                  switching capabilities
         VALUE    This comprises a set of one or more 8-bit switching
                  types, where each switching types identifies a
                  supported switching capability.

      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |   SC type     |   SC type     |   SC type     |               |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      //                                                              //
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Switching type values are defined in [RFC4205].

The Switch Caps sub-TLV is optional. It MAY be present in the PATH-COMP-CAP TLV. When present it MUST be present only once in the PATH-COMP-CAP TLV.

6.2. The ISIS PCES TLV

The ISIS PCE Status TLV (PCES TLV) carries information related to PCE processing congestion state.
The PCES TLV is carried within an ISIS Capability TLV which is defined in [ISIS-CAP].

The ISIS PCES TLV has the following format:

TYPE: To be assigned by IANA
LENGTH: Variable
VALUE: set of sub-TLVs

Sub-TLVs types are under IANA control.

Currently two sub-TLVs are defined (suggested type values to be assigned by IANA):

Sub-TLV type	Length	Name
1	variable	PCE-ADDRESS sub-TLV
2	3	CONGESTION sub-TLV

The PCE-ADDRESS and CONGESTION sub-TLVs MUST be present once in a PCES TLV. The PCE-ADDRESS sub-TLV is defined in section 6.1.1. It carries one of the PCE IP addresses and is used to identify the PCE the processing congestion state information is applied to. This is required as the PCES and PCED TLVs may be carried in separate ISIS Capability TLVs.

Any non recognized sub-TLV MUST be silently ignored.

Additional sub-TLVs could be added in the future to advertise additional congestion information.

6.2.1. The CONGESTION sub-TLV

The CONGESTION sub-TLV is used to indicate whether a PCE experiences a processing congestion state or not along with optionally the PCE expected congestion duration.
The CONGESTION sub-TLV is mandatory. It MUST be carried once within the PCES TLV.

The format of the CONGESTION sub-TLV is as follows:

TYPE: To be assigned by IANA (Suggested value =2)
LENGTH: 3
VALUE: This comprises a one-byte flag of bits indicating the congestion status, followed by a 2-bytes field indicating the congestion duration.

Here is the TLV structure

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|C|       Reserved|      Congestion Duration      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
      Value
        -C bit: When set this indicates that the PCE experiences
                congestion and cannot support any new request. When
                cleared this indicates that the PCE does not
                experience congestion an can support a new request.

        -Congestion Duration: 2-bytes, the estimated PCE congestion
                              duration in seconds.
```

When C is set and the Congestion Duration field is equal to 0, this means that the Congestion Duration is unknown.
When C is cleared the Congestion Duration MUST be set to 0.

6.3. Elements of Procedure

The PCED and PCES TLV are carried within an ISIS Capability TLV which is defined in [ISIS-CAP].

As PCES information is likely to change more frequently than the PCED information, it is RECOMMENDED to carry PCES and PCED TLVs in separate ISIS Capability TLVs, so as not to carry all PCED information each time the PCE status changes.

An ISIS router MUST originate a new ISIS LSP whenever the content of any of the PCED TLV or PCES TLV changes or whenever required by the regular ISIS procedure (LSP refresh).

When the scope of the PCED or PCES TLV is area local it MUST be carried within an ISIS CAPABILITY TLV having the S bit cleared.
When the scope of the PCED or PCES TLV is the entire IGP domain, the PCED TLV MUST be carried within an ISIS CAPABILITY TLV having the S bit set.
Note that when only the L bit of the PATH-SCOPE sub-TLV is set and the flooding scope MUST be local.
Note that the flooding scope of the PCED and PCES TLVs may be distinct, in which case they are carried in distinct ISIS Capability TLVs.

PCED and PCES sub-TLVs are OPTIONAL. When an ISIS LSP does not contain any PCED or PCES sub-TLV, this means that the PCE information of that node is unknown.

Note that a change in PCED information MUST not trigger any SPF computation.

The way PCEs retrieve their own information is out of the scope of this document. Some information may be configured (e.g. address, preferences, scope) and other information may be automatically retrieved (e.g. areas of visibility).

6.3.1. PCES TLV specific procedure

When a PCE enters into a processing congestion state, the conditions of which are implementation dependent, it SHOULD originate a new ISIS LSP with a Capability TLV carrying a PCES TLV with the C bit set and optionally a non-null expected congestion duration.

When a PCE leaves the processing congestion state, the conditions of which are implementation dependent, there are two cases:
- If the congestion duration in the previously originated PCES TLV was null, it SHOULD originate a PCES TLV with the C bit cleared and a null congestion duration;
- If the congestion duration in the previously originated PCES TLV was non null, it MAY originate a PCES TLV. Note that in some particular cases it may be desired to originate a PCES TLV with the C bit cleared if the saturation duration was over estimated.

The congestion duration allows reducing the amount of ISIS flooding, as only uncongested-congested state transitions are flooded.

An implementation SHOULD support an appropriate dampening algorithm so as to dampen ISIS flooding in order to not impact the ISIS scalability. It is RECOMMENDED to introduce some hysteresis for congestion state transition, so as to avoid state oscillations that may impact ISIS performances. For instance two thresholds MAY be configured: A resource saturation upper-threshold and a resource saturation lower-threshold. An LSR enters the congested state when the CPU load reaches the upper threshold and leaves the congested state when the CPU load goes under the lower threshold.

Upon receipt of an updated PCES TLV a PCC should take appropriate actions. In particular, the PCC SHOULD stop sending requests to a congested PCE, and SHOULD gradually start sending again requests to a no longer congested PCE.

7. Backward compatibility

The PCED and PCES TLVs defined in this document do not introduce any interoperability issue.
For OSPF, a router not supporting the PCED/PCES TLVs SHOULD just silently ignore the TLVs as specified in [OSPF-CAP].
For ISIS a router not supporting the PCED/PCES TLVs SHOULD just silently ignore the TLV as specified in [ISIS-CAP].

8. IANA considerations

8.1. OSPF TLVs

IANA will assign a new codepoint for the OSPF PCED TLV defined in this document and carried within the Router Information LSA.
IANA is requested to manage sub-TLV types for the PCED TLV.

Five sub-TLVs types are defined for this TLV and should be assigned by IANA:

- -PCE-ADDRESS sub-TLV (suggested value = 1)
- -PATH-SCOPE sub-TLV (suggested value = 2)
- -PCE-DOMAINS sub-TLV (suggested value = 3)
- -PCE-DEST-DOMAINS sub-TLV (suggested value =4)
- -GENERAL-CAP sub-TLV (suggested value = 5)
- -PATH-COMP-CAP sub-TLV (suggested value = 6)

Three sub-TLVs types are defined for the PCE-DOMAINS and PCE-DEST-DOMAINS TLVs and should be assigned by IANA:

- -IPv4 area ID sub-TLV (suggested value = 1)
- -IPv6 area ID sub-TLV (suggested value = 2)
- -AS number sub-TLV (suggested value = 3)

Three sub-TLV types are defined for the PATH-COMP-CAP TLV and should be assigned by IANA:

- -Objective Functions sub-TLV (suggested value =1)
- -Opaque Objective Function TLV (suggested value =2)

-Switch Caps sub-TLV (suggested value =3)

IANA will assign a new codepoint for the OSPF PCES TLV defined in this document and carried within the Router Information LSA.
IANA is requested to manage sub-TLV types for the PCES TLV. Two sub-TLVs types are defined for this TLV and should be assigned by IANA:

-PCE-ADDRESS sub-TLV (suggested value = 1)
-CONGESTION sub-TLV (suggested value = 2)

8.2. ISIS TLVs

IANA will assign a new codepoint for the PCED TLV defined in this document and carried within the ISIS CAPABILITY TLV.
IANA is requested to manage sub-TLV types for the PCED TLV.
Five sub-TLVs types are defined for the PCED TLV and should be assigned by IANA:

-PCE-ADDRESS sub-TLV (suggested value = 1)
-PATH-SCOPE sub-TLV (suggested value = 2)
-PCE-DEST-DOMAINS sub-TLV (suggested value = 3)
-PCE-DOMAINS sub-TLV (suggested value = 4)
-GENERAL-CAP sub-TLV (suggested value = 5)
-PATH-COMP-CAP sub-TLV (suggested value = 6)

Two sub-TLVs types are defined for the PCE-DOMAINS and PCE-DEST-DOMAINS TLVs and should be assigned by IANA:

-Area ID sub-TLV (suggested value = 1)
-AS number sub-TLV (suggested value = 2)

Three sub-TLV type are defined for the PATH-COMP-CAP TLV and should be assigned by IANA:

-Objective Functions TLV (suggested value =1)
-Opaque Objective Function TLV (suggested value =2)
-Switch Caps sub-TLV (suggested value =3)

IANA will assign a new codepoint for the ISIS PCES TLV defined in this document and carried within the ISIS CAPABILITY TLV.
IANA is requested to manage sub-TLV types for the PCES TLV. Two sub-TLVs types are defined for this TLV and should be assigned by IANA:

-PCE-ADDRESS sub-TLV (suggested value = 1)
-CONGESTION sub-TLV (suggested value = 2)

8.3. Capability bits

IANA is requested to manage the space of the General Capabilities 32-bit flag and the Path Computation Capabilities 32-bit flag defined in this document, numbering them in the usual IETF notation starting at zero and continuing through 31.
New bit numbers may be allocated only by an IETF Consensus action.
Each bit should be tracked with the following qualities:

- Bit number

- Defining RFC
- Name of bit

Currently two bits are defined in the General Capabilities flag. Here are the suggested values:

-0: Support for Request prioritization.
-1: Support for multiple messages within the same request message

Currently six bits are defined in the Path Computation Capabilities flag. Here are the suggested values:

-0: Capability to handle GMPLS Constraints
-1: Capability to compute bidirectional paths
-2: Capability to compute link/node/SRLG diverse paths
-3: Capability to compute load-balanced paths
-4: Capability to compute a set of paths in a synchronized Manner
-5: Support for multiple objective function
-6: Capability to handle path constraints (e.g. hop count, metric bound)

9. Security Considerations

To be completed in further revisions.

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.

[RFC3667] Bradner, S., "IETF Rights in Contributions", BCP 78, RFC 3667, February 2004.

[BCP79] Bradner, S., "Intellectual Property Rights in IETF Technology", RFC 3979, March 2005.

[OSPF-v2] Moy, J., "OSPF Version 2", RFC 2328, April 1998.

[OSPF-v3] Coltun, R., Ferguson, D., and J. Moy, "OSPF for IPv6", RFC 2740, December 1999.

[RFC2370] Coltun, R., The OSPF Opaque LSA Option , RFC 2370, July 1998.

[IS-IS] "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol " ISO 10589.

[IS-IS-IP] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.

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

[IS-IS-TE] Li, T., Smit, H., "IS-IS extensions for Traffic Engineering", RFC 3784, June 2004.

[OSPF-CAP] Lindem, A., Shen, N., Aggarwal, R., Shaffer, S., Vasseur, J.P., "Extensions to OSPF for advertising Optional Router Capabilities", draft-ietf-ospf-cap, work in progress.

[IS-IS-CAP] Vasseur, J.P. et al., "IS-IS extensions for advertising router information", draft-ietf-isis-caps, work in progress.

[PCE-ARCH] Farrel, A., Vasseur, J.P., Ash, J., "Path Computation Element (PCE) Architecture", draft-ietf-pce-architecture, work in progress.

[PCE-DISCO-REQ] Le Roux, J.L., et al. "Requirements for PCE discovery", draft-ietf-pce-discovery-reqs, work in progress

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

[RFC4205] Kompella, Rekhter, " IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC4205, October 2005.

10.2. Informative references

[PCECP-REQ] Ash, J., Le Roux, J.L., " PCE Communication Protocol Generic Requirements", draft-ietf-pce-comm-protocol-gen-reqs, work in progress.

[PCEP] Vasseur et al., Path Computation Element (PCE) communication Protocol (PCEP) - Version 1 , draft-ietf-pce-pcep, work in progress.

11. Authors' Addresses:

Jean-Louis Le Roux (Editor)
France Telecom
2, avenue Pierre-Marzin
22307 Lannion Cedex
FRANCE
Email: jeanlouis.leroux@francetelecom.com

Jean-Philippe Vasseur (Editor)
Cisco Systems, Inc.
1414 Massachusetts avenue
Boxborough , MA - 01719

USA
Email: jpv@cisco.com

Yuichi Ikejiri
NTT Communications Corporation
1-1-6, Uchisaiwai-cho, Chiyoda-ku
Tokyo 100-8019
JAPAN
Email: y.ikejiri@ntt.com

Raymond Zhang
BT Infonet
2160 E. Grand Ave.
El Segundo, CA 90025
USA
Email: raymond_zhang@infonet.com

12. Intellectual Property Statement