IS-IS Protocol Extensions for Path Computation Element (PCE) Discovery

Status of This Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Abstract

There are various circumstances where it is highly desirable for a Path Computation Client (PCC) to be able to dynamically and automatically discover a set of Path Computation Elements (PCEs), along with information that can be used by the PCC for PCE selection. When the PCE is a Label Switching Router (LSR) participating in the Interior Gateway Protocol (IGP), or even a server participating passively in the IGP, a simple and efficient way to announce PCEs consists of using IGP flooding. For that purpose, this document defines extensions to the Intermediate System to Intermediate System (IS-IS) routing protocol for the advertisement of PCE Discovery information within an IS-IS area or within the entire IS-IS routing domain.
1. Introduction

[RFC4655] describes the motivations and architecture for a Path Computation Element (PCE)-based path computation model for Multi-Protocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineered Label Switched Paths (TE LSPs). The model allows for the separation of the PCE from a Path Computation Client (PCC) (also referred to as a non co-located PCE) and allows for cooperation between PCEs (where one PCE acts as a PCC to make requests of the other PCE). This relies on a communication protocol between a PCC and PCE, and also between PCEs. The requirements for such a communication protocol can be found in [RFC4657], and the communication protocol is defined in [PCEP].

The PCE architecture requires that a PCC be aware of the location of one or more PCEs in its domain, and, potentially, of PCEs in other domains, e.g., in the case of inter-domain TE LSP computation.
A network may contain a large number of PCEs, each with potentially distinct capabilities. In such a context, it is highly desirable to have a mechanism for automatic and dynamic PCE discovery that allows PCCs to automatically discover a set of PCEs, along with additional information about each PCE that may be used by a PCC to perform PCE selection. Additionally, it is valuable for a PCC to dynamically detect new PCEs, failed PCEs, or any modification to the PCE information. Detailed requirements for such a PCE discovery mechanism are provided in [RFC4674].

Note that the PCE selection algorithm applied by a PCC is out of the scope of this document.

When PCCs are LSRs participating in the IGP (OSPF or IS-IS), and PCEs are either LSRs or servers also participating in the IGP, an effective mechanism for PCE discovery within an IGP routing domain consists of utilizing IGP advertisements.

This document defines extensions to IS-IS [ISO] to allow a PCE in an IS-IS routing domain to advertise its location, along with some information useful to a PCC for PCE selection, so as to satisfy dynamic PCE discovery requirements set forth in [RFC4674].

Generic capability advertisement mechanisms for IS-IS are defined in [RFC4971]. These allow a router to advertise its capabilities within an IS-IS area or an entire IS-IS routing domain. This document leverages this generic capability advertisement mechanism to fully satisfy the dynamic PCE discovery requirements.

This document defines a new sub-TLV (named the PCE Discovery (PCED)) to be carried within the IS-IS Router Capability TLV ([RFC4971]).

The PCE information advertised is detailed in Section 3. Protocol extensions and procedures are defined in Sections 4 and 5.

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

This document defines a set of sub-TLVs that are nested within each other. When the degree of nesting TLVs is 2 (a TLV is carried within another TLV) the TLV carried within a TLV is called a sub-TLV. Strictly speaking, when the degree of nesting is 3, a sub-sub-TLV is carried within a sub-TLV that is itself carried within a TLV. For the sake of terminology simplicity, a TLV carried within another TLV is called a sub-TLV regardless of the degree of nesting.
2. Terminology

ABR: IS-IS Area Border Router.

AS: Autonomous System.

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

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

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

Inter-area TE LSP: A TE LSP whose path transits two or more IGP areas. That is, a TE LSP that crosses at least one IGP area boundary.

Inter-AS TE LSP: A TE LSP whose path transits two or more ASes or sub-ASes (BGP confederations). That is, a TE LSP that crosses at least one AS boundary.

IS-IS LSP: Link State PDU.

LSR: Label Switching 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.

PCED: PCE Discovery.

PCE-Domain: In a PCE context, this refers to any collection of network elements within a common sphere of address management or path computational responsibility (referred to as a "domain" in [RFC4655]). Examples of PCE-Domains include IGP areas and ASes. This should be distinguished from an IS-IS routing domain as defined by [ISO].

PCEP: Path Computation Element communication Protocol.

TE LSP: Traffic Engineered Label Switched Path.

TLV: Type-Length-Variable data encoding.
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].

3. Overview

3.1. PCE Discovery Information

The PCE discovery information is composed of:

- The PCE location: an IPv4 and/or IPv6 address that is used to reach the PCE. It is RECOMMENDED to use an address that is always reachable if there is any connectivity to the PCE;

- The PCE path computation scope (i.e., intra-layer, inter-area, inter-AS, or inter-layer);

- The set of one or more PCE-Domain(s) into which the PCE has visibility and for which the PCE can compute paths;

- The set of zero, one, or more neighbor PCE-Domain(s) toward which the PCE can compute paths;

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

PCE discovery information is, by nature, fairly static and 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.

3.2. Flooding Scope

The flooding scope for PCE information advertised through IS-IS can be a single L1 area, an L1 area and the L2 sub-domain, or the entire IS-IS routing domain.

4. The IS-IS PCED Sub-TLV

The IS-IS PCED sub-TLV contains a non-ordered set of sub-TLVs.

The format of the IS-IS PCED sub-TLV and its sub-TLVs is identical to the TLV format used by the Traffic Engineering Extensions to IS-IS [RFC3784]. That is, the TLV is comprised of 1 octet for the type, 1 octet specifying the TLV length, and a value field. The Length field defines the length of the value portion in octets.
The IS-IS PCED sub-TLV has the following format:

<table>
<thead>
<tr>
<th>Sub-TLV type</th>
<th>Length</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>variable</td>
<td>PCE-ADDRESS sub-TLV</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>PATH-SCOPE sub-TLV</td>
</tr>
<tr>
<td>3</td>
<td>variable</td>
<td>PCE-DOMAIN sub-TLV</td>
</tr>
<tr>
<td>4</td>
<td>variable</td>
<td>NEIG-PCE-DOMAIN sub-TLV</td>
</tr>
<tr>
<td>5</td>
<td>variable</td>
<td>PCE-CAP-FLAGS sub-TLV</td>
</tr>
</tbody>
</table>

Five sub-TLVs are defined:

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

- The PCE-ADDRESS sub-TLV is mandatory; it MUST be present within the PCED sub-TLV. It MAY appear twice, when the PCE has both an IPv4 and IPv6 address. It MUST NOT appear more than once for the same address type. If it appears more than once for the same address type, only the first occurrence is processed and any others MUST be ignored.

- The PCE-DOMAIN and NEIG-PCE-DOMAIN sub-TLVs are optional. They MAY be present in the PCED sub-TLV to facilitate selection of inter-domain PCEs.

- The PCE-CAP-FLAGS sub-TLV is optional and MAY be present in the PCED sub-TLV to facilitate the PCE selection process.

- Any unrecognized sub-TLV MUST be silently ignored.

- The PCED sub-TLV is carried within an IS-IS CAPABILITY TLV defined in [RFC4971].

- No additional sub-TLVs will be added to the PCED TLV in the future. If a future application requires the advertisement of additional PCE information in IS-IS, this will not be carried in the CAPABILITY TLV.

- The following sub-sections describe the sub-TLVs that may be carried within the PCED sub-TLV.

4.1. PCE-ADDRESS Sub-TLV

The PCE-ADDRESS sub-TLV specifies an IP address that can be used to reach the PCE. It is RECOMMENDED to make use of an address that is always reachable, provided the PCE is alive and reachable.
The PCE-ADDRESS sub-TLV has the following format:

- **TYPE:** 1
- **LENGTH:** 5 for an IPv4 address or 17 for an 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

### 4.2. The PATH-SCOPE Sub-TLV

The PATH-SCOPE sub-TLV indicates the PCE path computation scope, which refers to the PCE’s ability to compute or take part in the computation of paths for intra-area, inter-area, inter-AS, or inter-layer TE LSPs.

The PATH-SCOPE sub-TLV is mandatory; it MUST be present within the PCED sub-TLV. There MUST be exactly one instance of the PATH-SCOPE sub-TLV within each PCED sub-TLV. If it appears more than once only the first occurrence is processed and any others MUST be ignored.

The PATH-SCOPE sub-TLV contains a set of bit flags indicating the supported path scopes, and four fields indicating PCE preferences.

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

- **TYPE:** 2
- **LENGTH:** 3
- **VALUE:** This comprises a 1-octet flags field where each flag represents a supported path scope, followed by a 2-octet preferences field indicating PCE preferences.

Here is the structure of the flags field:

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

```
+---------------+
```

Bit     Path Scope
0      L bit: Can compute intra-area paths.
1      R bit: Can act as PCE for inter-area TE LSP computation.
2      Rd bit: Can act as a default PCE for inter-area TE LSP computation.
3      S bit: Can act as PCE for inter-AS TE LSP computation.
4      Sd bit: Can act as a default PCE for inter-AS TE LSP computation.
5      Y bit: Can act as PCE for inter-layer TE LSP computation.
6-7   Reserved for future use.

Here is the structure of the preferences field:

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

PrefL field: PCE’s preference for intra-area TE LSP computation.
PrefR field: PCE’s preference for inter-area TE LSP computation.
PrefS field: PCE’s preference for inter-AS TE LSP computation.
Pref-Y field: PCE’s preference for inter-layer TE LSP computation.
Res: Reserved for future use.

The L, R, S, and Y bits are set when the PCE can act as a PCE for intra-area, inter-area, inter-AS, or inter-layer TE LSP 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 LSP computation (that is, the PCE can compute a path toward any neighbor area). Similarly, when set, the Sd bit indicates that the PCE can act as a default PCE for inter-AS TE LSP computation (the PCE can compute a path toward any neighbor AS).

When the Rd and Sd bit are set, the PCED sub-TLV MUST NOT contain a NEIG-PCE-DOMAIN sub-TLV (see Section 4.4).

When the R bit is clear, the Rd bit SHOULD be clear on transmission and MUST be ignored on receipt. When the S bit is clear, the Sd bit SHOULD be clear on transmission and MUST be ignored on receipt.

The PrefL, PrefR, PrefS and PrefY fields are each three bits long and allow the PCE to specify a preference for each computation scope,
where 7 reflects the highest preference. Such preferences can be used for weighted load balancing of path computation requests. An operator may decide to configure a preference for each computation scope at each PCE so as to balance the path computation load among them. The algorithms used by a PCC to balance its path computation requests according to such PCE preferences are out of the scope of this document and are a matter for local or network-wide policy. The same or different preferences may be used for each scope. For instance, an operator that wants a PCE capable of both inter-area and inter-AS computation to be preferred for use for inter-AS computations may configure PrefS higher than PrefR.

When the L, R, S, or Y bits are cleared, the PrefL, PrefR, PrefS, and PrefY fields SHOULD respectively be set to 0 on transmission and MUST be ignored on receipt.

Both reserved fields SHOULD be set to zero on transmission and MUST be ignored on receipt.

4.3. PCE-DOMAIN Sub-TLV

The PCE-DOMAIN sub-TLV specifies a PCE-Domain (area and/or AS) where the PCE has topology visibility and through which the PCE can compute paths.

The PCE-DOMAIN sub-TLV SHOULD be present when PCE-Domains for which the PCE can operate 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 (see Section 5 for a discussion of how the flooding scope is set and interpreted).

A PCED sub-TLV may include multiple PCE-DOMAIN sub-TLVs when the PCE has visibility into multiple PCE-Domains.

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

- **TYPE:** 3
- **LENGTH:** Variable
- **VALUE:** This is composed of one octet indicating the domain-type (area ID or AS Number) and a variable length IS-IS area ID or a 32-bit AS number, identifying a PCE-Domain where the PCE has visibility and can compute paths.

Two domain types are defined:

1. Area ID
2. AS Number
The Area ID is the area address as defined in [ISO].

When the AS number is coded in two octets, the AS Number field MUST have its first two octets set to 0.

4.4. NEIG-PCE-DOMAIN Sub-TLV

The NEIG-PCE-DOMAIN sub-TLV specifies a neighbor PCE-Domain (area or AS) toward which a PCE can compute paths. It means that the PCE can take part in the computation of inter-domain TE LSPs with paths that transit this neighbor PCE-Domain.

A PCED sub-TLV may include several NEIG-PCE-DOMAIN sub-TLVs when the PCE can compute paths towards several neighbor PCE-Domains.

The NEIG-PCE-DOMAIN sub-TLV has the same format as the PCE-DOMAIN sub-TLV:

- **TYPE:** 4
- **LENGTH:** Variable
- **VALUE:** This comprises one octet indicating the domain-type (area ID or AS Number) and a variable length IS-IS area ID or a 32-bit AS number, identifying a PCE-Domain toward which the PCE can compute paths.

Two domain types are defined:

1. Area ID
2. AS Number

The Area ID is the area address as defined in [ISO].

When the AS number is coded in two octets, the AS Number field MUST have its first two octets set to 0.

The NEIG-PCE-DOMAIN sub-TLV MUST be present at least once with domain-type set to 1 if the R bit is set and the Rd bit is cleared, and MUST be present at least once with domain-type set to 2 if the S bit is set and the Sd bit is cleared.

4.5. PCE-CAP-FLAGS Sub-TLV

The PCE-CAP-FLAGS sub-TLV is an optional sub-TLV used to indicate PCE capabilities. It MAY be present within the PCED sub-TLV. It MUST NOT be present more than once. If it appears more than once, only the first occurrence is processed and any others MUST be ignored.
The value field of the PCE-CAP-FLAGS sub-TLV is made up of an array of units of 32-bit flags numbered from the most significant bit as bit zero, where each bit represents one PCE capability.

The PCE-CAP-FLAGS sub-TLV has the following format:

- **TYPE:** 5
- **LENGTH:** Multiple of 4
- **VALUE:** This contains an array of units of 32-bit flags numbered from the most significant as bit zero, where each bit represents one PCE capability.

The PCE capability registry is managed by IANA; it is common with OSPF and defined in [RFC5088].

Reserved bits SHOULD be set to zero on transmission and MUST be ignored on receipt.

5. Elements of Procedure

The PCED sub-TLV is advertised within an IS-IS Router Capability TLV defined in [RFC4971]. As such, elements of procedures are inherited from those defined in [RFC4971].

The flooding scope is controlled by the S flag in the IS-IS Router Capability TLV (see [RFC4971]). When the scope of the PCED sub-TLV is area local, it MUST be carried within an IS-IS Router Capability TLV having the S bit cleared. When the scope of the PCED sub-TLV is the entire IS-IS routing domain, it MUST be carried within an IS-IS Router Capability TLV having the S bit set. Note that when only the L bit of the PATH-SCOPE sub-TLV is set, the flooding scope MUST be area local.

Note that an L1L2 node may include a PCED TLV in a Router Capability TLV with the S bit cleared in both in its L1 and L2 LSPs. This allows the flooding scope to be restricted to the L1 area and the L2 sub-domain.

When the PCE function is deactivated, the IS-IS speaker advertising this PCE MUST originate a new IS-IS LSP that no longer includes the corresponding PCED TLV.

The PCE address (i.e., the address indicated within the PCE-ADDRESS sub-TLV) SHOULD be reachable via some prefixes advertised by IS-IS.

The PCED sub-TLV information regarding a specific PCE is only considered current and useable when the router advertising this
information is itself reachable via IS-IS calculated paths at the level of the LSP in which the PCED sub-TLV appears.

A change in the state of a PCE (activate, deactivate, parameter change) MUST result in a corresponding change in the PCED sub-TLV information advertised by an IS-IS router (inserted, removed, updated) in its LSP. The way PCEs determine the information they advertise, and how that information is made available to IS-IS, is out of the scope of this document. Some information may be configured (e.g., address, preferences, scope) and other information may be automatically determined by the PCE (e.g., areas of visibility).

A change in information in the PCED sub-TLV MUST NOT trigger any SPF computation at a receiving router.

6. Backward Compatibility

The PCED sub-TLV defined in this document does not introduce any interoperability issues.

An IS-IS router not supporting the PCED sub-TLV will just silently ignore the sub-TLV as specified in [RFC4971].

7. IANA Considerations

IANA has defined a registry for the sub-TLVs carried in the IS-IS Router Capability TLV defined in [RFC4971]. IANA has assigned a new sub-TLV codepoint for the PCED sub-TLV carried within the Router Capability TLV.

<table>
<thead>
<tr>
<th>Value</th>
<th>Sub-TLV</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>PCED sub-TLV</td>
<td>(this document)</td>
</tr>
</tbody>
</table>

8. Security Considerations

This document defines IS-IS extensions for PCE discovery within an administrative domain. Hence the security of the PCE discovery relies on the security of IS-IS.

Mechanisms defined to ensure authenticity and integrity of IS-IS LSPs [RFC3567] and their TLVs, can be used to secure the PCED sub-TLV as well.

IS-IS provides no encryption mechanism for protecting the privacy of LSPs and, in particular, the privacy of the PCE discovery information.
9. Manageability Considerations

Manageability considerations for PCE Discovery are addressed in Section 4.10 of [RFC4674].

9.1. Control of Policy and Functions

Requirements for the configuration of PCE discovery parameters on PCCs and PCEs are discussed in Section 4.10.1 of [RFC4674].

In particular, a PCE implementation SHOULD allow the following parameters to be configured on the PCE:

- The PCE IPv4/IPv6 address(es) (see Section 4.1).
- The PCE Scope, including the inter-domain functions (inter-area, inter-AS, inter-layer), the preferences, and whether the PCE can act as default PCE (see Section 4.2).
- The PCE-Domains (see Section 4.3).
- The neighbor PCE-Domains (see Section 4.4).
- The PCE capabilities (see Section 4.5).

9.2. Information and Data Model

A MIB module for PCE Discovery is defined in [PCED-MIB].

9.3. Liveness Detection and Monitoring

This document specifies the use of IS-IS as a PCE Discovery Protocol. The requirements specified in [RFC4674] include the ability to determine liveness of the PCE Discovery protocol. Normal operation of the IS-IS protocol meets these requirements.

9.4. Verify Correct Operations

The correlation of information advertised against information received can be achieved by comparing the information in the PCED sub-TLV received by the PCC with that stored at the PCE using the PCED MIB [PCED-MIB]. The number of dropped, corrupt, and rejected information elements are available through the PCED MIB.

9.5. Requirements on Other Protocols and Functional Components

The IS-IS extensions defined in this document do not imply any requirements on other protocols.
9.6. Impact on Network Operations

Frequent changes in PCE information advertised in the PCED sub-TLV may have a significant impact on IS-IS and might destabilize the operation of the network by causing the PCCs to swap between PCEs.

As discussed in Section 4.10.4 of [RFC4674], it MUST be possible to apply at least the following controls:

- Configurable limit on the rate of announcement of changed parameters at a PCE.
- Control of the impact on PCCs, such as through rate-limiting the processing of PCED sub-TLVs.
- Configurable control of triggers that cause a PCC to swap to another PCE.

10. Acknowledgments

We would like to thank Lucy Wong, Adrian Farrel, Les Ginsberg, Mike Shand, Lou Berger, David Ward, Ross Callon, and Lisa Dusseault for their useful comments and suggestions.
11. References

11.1. Normative References


11.2. Informative References


Authors’ Addresses

Jean-Louis Le Roux (Editor)
France Telecom
2, avenue Pierre-Marzin
22307 Lannion Cedex
FRANCE
EMail: jeanlouis.leroux@orange-ftgroup.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
2160 E. Grand Ave.
El Segundo, CA 90025
USA
EMail: raymond.zhang@bt.com
Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein 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 HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF 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 this 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. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR 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 this standard. Please address the information to the IETF at ietf-ipr@ietf.org.