PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 15, 2018

H. Chen
Huawei Technologies
M. Toy
Verizon
X. Liu
Jabil
L. Liu
Fujitsu
Z. Li
China Mobile

September 11, 2017

Static PCEP Link State
draft-chen-pce-pcc-ted-03

Abstract

This document presents extensions to the Path Computation Element Communication Protocol (PCEP) for a PCC to advertise the information about the links without running IGP and for a PCE to build a TED based on the information received.

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 March 15, 2018.

Table of Contents

1. Introduction

A PCE architecture is described in RFC 4655, in which a Traffic Engineering Database (TED) for a PCE is constructed based on the link information from IGP (OSPF or IS-IS) running in the domain for which the PCE is responsible.

For a domain without running IGP, the PCE responsible for the domain may obtain the link information from a PCC running on each node in the domain.

This document presents extensions to the Path Computation Element Communication Protocol (PCEP) for a PCC to advertise the information about the links attached to the node running the PCC and for a PCE to build the TED based on the information received from the PCC.

2. Terminology

ABR: Area Border Router. Router used to connect two IGP areas (Areas in OSPF or levels in IS-IS).

ASBR: Autonomous System (AS) Border Router. Router used to connect together ASes via inter-AS links.

This document uses terminology defined in [RFC5440].

3. 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].

4. Link Information

Since no IGP runs over any link, we may not obtain any link information via IGP. But links are configured.

For a point-to-point (P2P) link between nodes A and B, from A's point of view, we have the following link information:

1) Link Type: P2P
2) Local IP address
3) Remote IP address
4) Traffic engineering metric
5) Maximum bandwidth
6) Maximum reservable bandwidth
7) Unreserved bandwidth
8) Administrative group

9) SRLG

A link ID for the link is obtained if a user configures it; otherwise, no link ID (i.e., the Router ID of A's neighbor) may be obtained since no IGP adjacency over the link is formed.

For a broadcast link connecting multiple nodes, on each of the nodes X, we have the same link information as above except for:

a) Link Type: Multi-access,
b) Local IP address with mask length, and
c) No Remote IP address.

In other words, the information about the broadcast link obtained by node X comprises a), b), 4) to 9), but does not include any remote IP address or link ID. A link ID for the link is obtained if a user configures it; otherwise, no link ID (i.e., the interface address of the designated router for the link) may be obtained since no IGP selects it.

A PCE constructs a TED for its responsible domain after receiving the link information from the PCC running on every node in the domain.

5. Extensions to PCEP

5.1. Extension to Existing Message

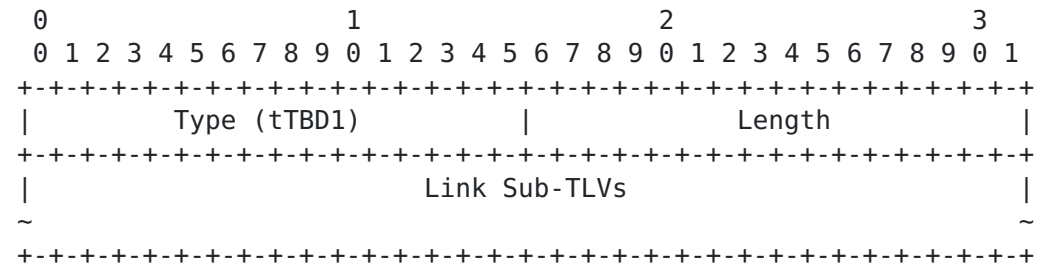
An existing Notification message may be extended to advertise the information about links. Alternatively, a new message can be used (refer to Appendix A).

The following new Notification-type (NT) and Notification-value (NV) of a NOTIFICATION object in a Notification message are defined:

- NT=8 (TBD): Links
  - NV=1: Update Links. NT=8 and NV=1 indicates that the PCC requests the PCE to update the link information based on the TLVs in the object, which are described below.
  - NV=2: Withdraw Links. NT=8 and NV=2 indicates that the PCC asks the PCE to remove the Links indicated by the TLVs in the object.

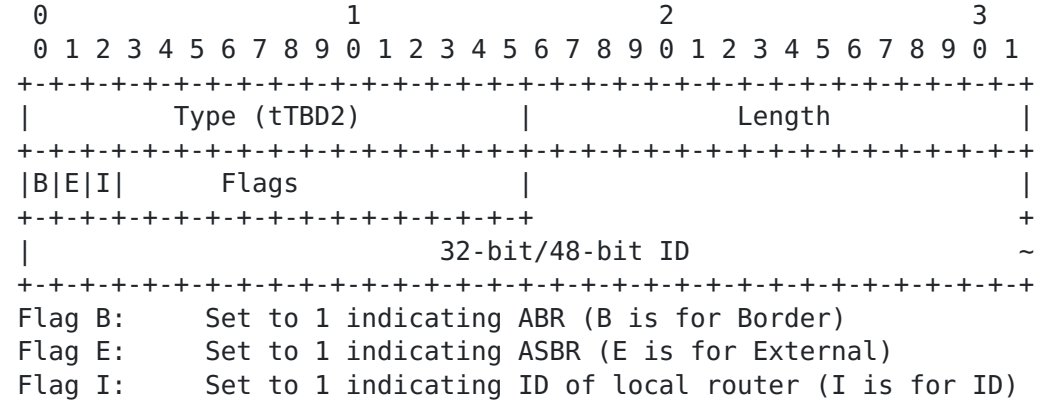
5.1.1. TLVs

A link TLV and a Router-ID TLV are defined. The format of the link TLV is illustrated below. The Type=tTBD1 indicates a link TLV Type. The Length indicates the size of the Link Sub-TLVs.



A link TLV describes a single link. It comprises a number of link sub-TLVs for the information described in section 4, which are the sub-TLVs defined in RFC 3630 or their equivalents except for the local IP address with mask length defined below.

The format of the Router-ID TLV is shown below. The Type=tTBD2 indicates a Router-ID TLV Type. The Length indicates the size of the ID and flags field.



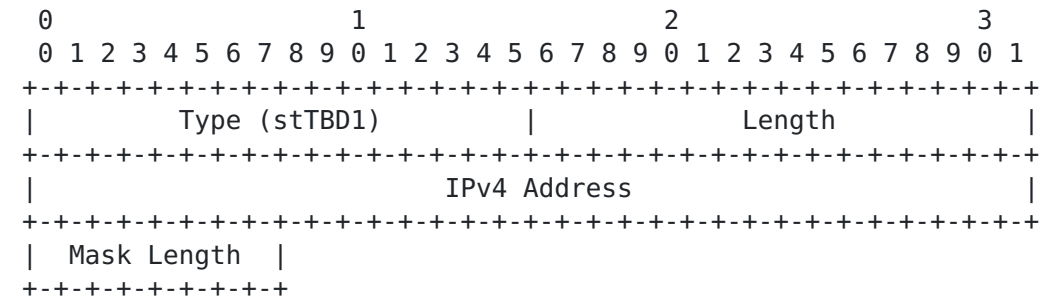
Flag B: Set to 1 indicating ABR (B is for Border)
Flag E: Set to 1 indicating ASBR (E is for External)
Flag I: Set to 1 indicating ID of local router (I is for ID)

Undefined flags MUST be set to zero. The ID indicates the ID of a router. For a router not running IGP, the ID may be the 32-bit or 48-bit ID of the router configured.

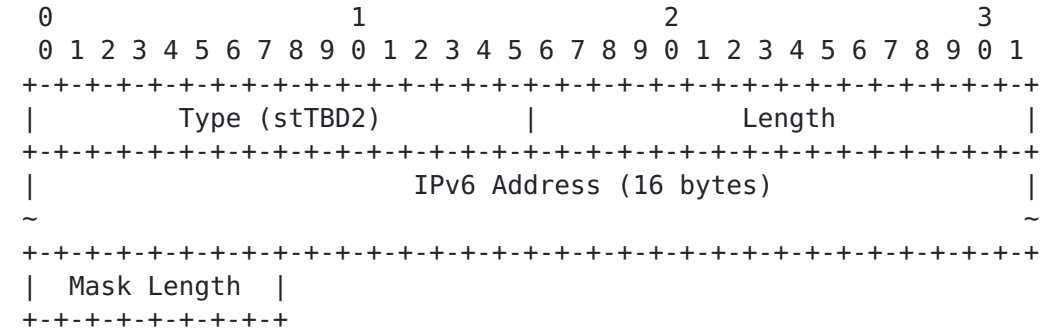
5.1.2. Sub-TLVs

The format of the Sub-TLV for a local IPv4 address with mask length is shown below. The Type=stTBD1 indicates a local IPv4 Address with mask length. The Length indicates the size of the IPv4 address and

Mask Length. The IPv4 Address indicates the local IPv4 address of a link. The Mask Length indicates the length of the IPv4 address mask.



The format of the Sub-TLV for a local IPv6 address with mask length is illustrated below. The Type=stTBD2 indicates a local IPv6 Address with mask length. The Length indicates the size of the IPv6 address and Mask Length. The IPv6 Address indicates the local IPv6 address of a link. The Mask Length indicates the length of the IPv6 address mask.



5.2. Procedures

5.2.1. PCC Procedures

1. New or Changed Links

After the session between a PCC and a PCE is established, the PCC sends the PCE a message containing the information about the links attached to the node running the PCC.

For any new or changed links, the PCC sends the PCE a message containing the information about these links with indication of Update Links.

For example, for a new P2P link from node A, the PCC running on A

sends the PCE a Notification message having a NOTIFICATION object with NT=8 and NV=1 (indicating Update Links), which contains a Router-ID TLV, followed by a link TLV. The former comprises A's ID and flag I set to 1. The latter comprises the Sub-TLVs for the information described in section 4.

For multiple new or changed links from node A, the PCC running on A sends the PCE a Notification message having a NOTIFICATION object with NT=8 and NV=1, which contains a Router-ID TLV for A's ID, followed by multiple link TLVs for the links.

2. Links Down

For links down, the PCC sends the PCE a message containing the information about these links with indication of Withdraw Links.

For example, for multiple links from node A down, the PCC running on A sends the PCE a Notification message having a NOTIFICATION object with NT=8 and NV=2 (indicating Withdraw Links), which contains a Router-ID TLV for A's ID, followed by multiple link TLVs for the links. The TLV for a P2P link comprises the Sub-TLVs for the information on 1), 2) and 3) described in section 4. The TLV for a broadcast link comprises the Sub-TLVs for the information on a) and b) described in section 4.

3. Simplified Message

Alternatively, the messages may be simplified. For each node, the source IP address of the PCC running on the node may be used as the ID of the node. The PCE knows the address after the session between the PCE and the PCC is up. Thus, a message containing the information about links does not need include any router-ID TLV.

For example, for a new P2P link attached to node A, the PCC running on A sends the PCE a Notification message having a NOTIFICATION object with NT=8 and NV=1 (indicating Update Links), which contains a link TLV comprising the Sub-TLVs for the information on 1) to 9) described in section 4. The object does not contain any Router-ID TLV for node A.

5.2.2. PCE Procedures

A PCE stores into its TED the links for each node according to the messages for the links received from the PCC running on the node. For a message containing Update Links, it updates the links accordingly. For a message containing Withdraw Links, it removes the links. When a node is down, the PCE removes the links attached to the node.

For a new P2P link between node A and B with no link ID configured, when receiving a message containing the link from the PCC running on A, the PCE stores the link for A (i.e., the link from A) into its TED. It will find the link's remote end B using the remote IP address of the link. After finding B, it associates the link for A with B and the link for B with A. This creates a bidirectional connection between A and B.

For a new broadcast link connecting multiple nodes with no link ID configured, when receiving a message containing the link from the PCC running on each of the nodes X, the PCE stores the link for X (i.e., the link from X) into its TED. It will find the link's remote end P using the link's local IP address with network mask. P is a Pseudo node identified by the local IP address of the designated node selected from the nodes connected to the link. After finding P, it associates the link for X with P and the link for P with X. This creates a bidirectional connection between X and P.

The first node and second node from which the PCE receives a message containing the link is selected as the designed node and backup designed node respectively. After the designed node is down, the backup designed node becomes the designed node and the node other than the designed node with the largest local IP address connecting to the link is selected as the backup designed node.

When the old designed node is down and the backup designed node becomes the new designed node, the PCE updates its TED through removing the link between each of nodes X and old P (the Pseudo node corresponding to the old designed node) and adding a link between each of nodes X (still connecting to the broadcast link) and new P (the Pseudo node corresponding to the new designed node).

6. Security Considerations

The mechanism described in this document does not raise any new security issues for the PCEP protocols.

7. IANA Considerations

This section specifies requests for IANA allocation.

8. Acknowledgement

The authors would like to thank Jescia Chen, and Eric Wu for their valuable comments on this draft.

9. References

9.1. Normative References

[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>.

[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>.

[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>.

[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>.

9.2. Informative References

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <https://www.rfc-editor.org/info/rfc7752>.

Appendix A. New Message

A new message may be defined to advertise the information on links. The format of the message for the information on Links (IL for short) is as follows:

```
<IL Message> ::= <Common Header> <NRP> <Link-List>
where:
  <Link-List> ::= <LINK> [<Link-List>]
```

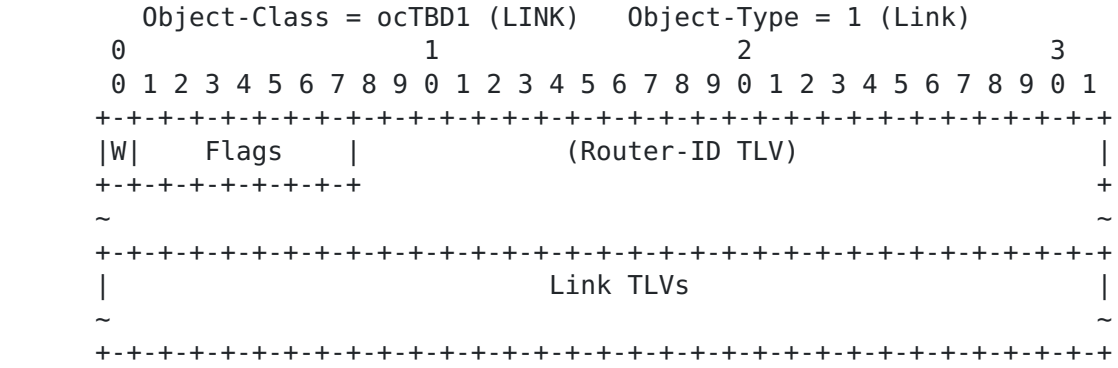
Where the value of the Message-Type in the Common Header indicates the new message type. The exact value is to be assigned by IANA. A new RP (NRP) object will be defined, which follows the Common Header.

A new flag W (Withdraw) in the NRP object is defined to indicate whether the links are withdrawn. When flag W is set to one, the PCE removes the links in the message after receiving it from the PCC.

When flag W is set to zero, the PCE adds/updates the links in the message.

An alternative to flag W in the NRP object is a similar flag W in each LINK object. For example, when the flag is set to one in the LINK object, the PCE removes the links in the object. When the flag is set to zero, the PCE adds/updates the links in the object.

The format of a LINK object body is as follows:



Flag W=1 indicates Withdraw links. W=0 indicates Updated links. Router-ID TLV is optional. Link TLVs are mandatory. They are the same as described in section 5.

Authors' Addresses

Huaimo Chen
Huawei Technologies
Boston, MA,
USA

EMail: Huaimo.chen@huawei.com

Mehmet Toy
Verizon
USA

EMail: mehmet.toy@verizon.com

Xufeng Liu
Jabil
McLean, VA
USA

EMail: Xufeng_Liu@jabil.com

Lei Liu
Fujitsu
USA

EMail: lliu@us.fujitsu.com

Zhenqiang Li
China Mobile
No.32 Xuanwumenxi Ave., Xicheng District
Beijing 100032
P.R. China

EMail: li_zhenqiang@hotmail.com