PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 8, 2017

H. Chen
Huawei Technologies
M. Toy
Comcast
L. Liu
Fujitsu
Z. Li
China Mobile

July 7, 2016

Connections and Accesses for Hierarchical PCE
draft-chen-pce-h-connect-access-00

Abstract

This document presents extensions to the Path Computation Element Communication Protocol (PCEP) to distribute information about connections and access points for supporting a hierarchical PCE system.

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 8, 2017.

Table of Contents

1. Introduction 3
2. Terminology 3
3. Conventions Used in This Document 3
4. Connections and Accesses 4
4.1. Information on Inter-domain Link 4
4.2. Information on ABR 5
4.3. Information on Access Point 5
5. Extensions to PCEP 5
5.1. Extension to Existing Message 5
5.1.1. TLVs 6
5.1.2. Sub-TLVs 8
5.2. Procedures 8
5.2.1. Child Procedures 8
5.2.2. Parent Procedures 10
6. Security Considerations 11
7. IANA Considerations 11
8. Acknowledgement 11
9. References 11
9.1. Normative References 11
9.2. Informative References 12
Appendix A. New Message 12
A.1. CONNECTION and ACCESS Object 13
A.2. TLVs in CONNECTION and ACCESS Object 15

1. Introduction

A hierarchical PCE architecture is described in RFC 6805, in which a parent PCE maintains a domain topology containing its child domains (seen as vertices in the topology) and the connections among the domains.

For a domain for which a child PCE is responsible, connections attached to the domain may comprise inter-domain links and Area Border Routers (ABRs). For a parent PCE to have the domain topology, each of its child PCEs needs to advertise its connections to the parent PCE.

In addition to the connections attached to the domain, there may be some access points in the domain, which are the addresses in the domain to be accessible outside of the domain. For example, an address of a server in the domain that provides a number of services to users outside of the domain is an access point.

This document presents extensions to the Path Computation Element Communication Protocol (PCEP) for a child PCE to advertise the information about its connections and access points and for a parent PCE to build and maintain the domain topology as well as access points.

2. Terminology

The following terminology is used in this document.

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

ASBR: Autonomous System Border Router. Router used to connect together ASes of the same or different service providers via one or more inter-AS links.

TED: Traffic Engineering Database.

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. Connections and Accesses

A connection is an inter-domain link between two domains in general. An ABR is also a connection, which connects two special domains called areas in a same Autonomous System (AS).

An access point in a domain is an address in the domain to be accessible to the outside of the domain. An access point is simply called an access.

4.1. Information on Inter-domain Link

An inter-domain link connects two domains in two different ASes. Since there is no IGP running over an inter-domain link, we may not obtain the information about the link generated by an IGP. We may suppose that IP addresses are configured on inter-domain links.

For a point-to-point link connecting two ABSRs A and B in two different domains, from A's point of view, the following information about the link may be obtained:

1) Link Type: Point-to-point
2) Local IP address of the link
3) Remote IP address of the link
4) Traffic engineering metric of the link
5) Maximum bandwidth of the link
6) Maximum reservable bandwidth of the link
7) Unreserved bandwidth of the link
8) Administrative group of the link

Note that no link ID (i.e., the Router ID of the neighbor) may be obtained since no IGP adjacency over the link is formed.

For a broadcast link connecting multiple ASBRs in a number of domains, on each of the ASBRs X, the same information about the link as above may be obtained except for the followings:

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 ASBR X comprises a), b), 4), 5), 6), 7) and 8), but does not include any remote IP address or link ID. Note that no link ID (i.e., the interface address of the designated router for the link) may be obtained since no IGP selects it.

A parent PCE constructs an AS domain topology after receiving the information about each of the inter-domain links described above from its child PCEs plus the local node ID such as A's ID.

A PCE such as a child PCE can constructs a TED for the domain for which the PCE is responsible after receiving the information described above about each of the links attached to the nodes in the domain from the PCCs running on the nodes without running IGP.

4.2. Information on ABR

For an AS running IGP and containing multiple areas, an ABR connects two or more areas. For each area connected to the ABR, the PCE as a child responsible for the area sends its parent PCE the information about the ABR, which indicates the identifier (ID) of the ABR.

A parent PCE has the information about each of its child PCEs, which includes the domain such as the area for which the child PCE is responsible. The parent PCE knows all the areas to which each ABR connects after receiving the information on the ABR from each of its child PCEs.

4.3. Information on Access Point

For an IP address in a domain to be accessible outside of the domain, the PCE as a child responsible for the domain sends its parent PCE the information about the address, which indicates the address.

The parent PCE has all the access points (i.e., IP addresses) to be accessible outside of all its child PCEs' domains after receiving the information on the access points from each of its child PCEs.

5. Extensions to PCEP

This section describes the extensions to PCEP to distribute the information about inter-domain links, ABRs and access points. The information is sent from a child PCE to its parent PCE. A child PCE is simply called a child and a parent PCE is called a parent in the following sections.

5.1. Extension to Existing Message

An existing Notification message may be extended to advertise the information about connections and access points. 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): Connections and Accesses

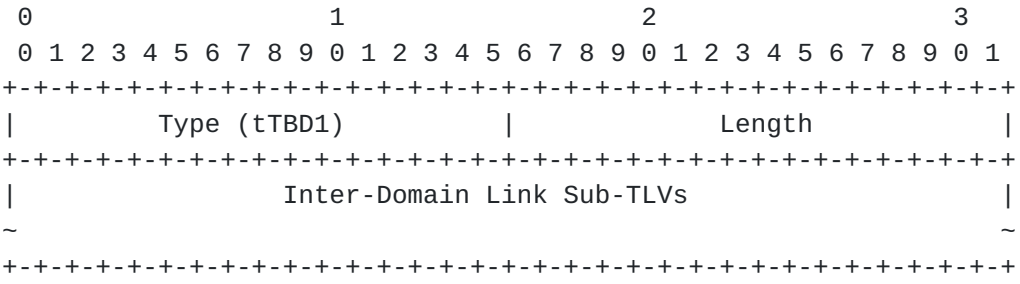
  - NV=1: Updates on Connections and Accesses. A NT=8 and NV=1 indicates that the child sends its parent updates on the information about Connections and Accesses, and TLVs containing the information are in the NOTIFICATION object. The format and contents of the TLVs are described below.

  - NV=2: Withdraw Connections and Accesses. A NT=8 and NV=2 indicates that the child asks its parent to remove Connections and Accesses indicated by the TLVs in the object.

5.1.1. TLVs

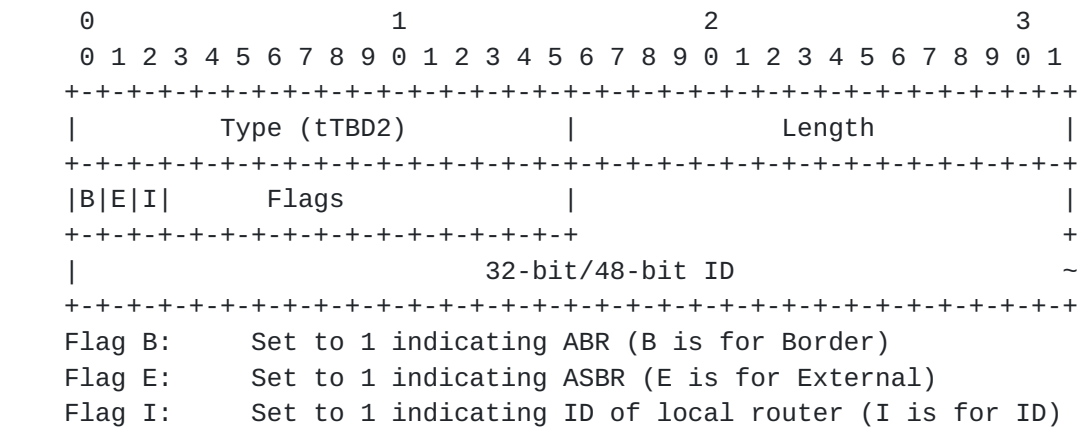
Four TLVs are defined for the information on connections and accesses. They are Inter-Domain link TLV, Router-ID TLV, Access IPv4 Prefix TLV and Access IPv6 Prefix TLV.

The format of the Inter-Domain link TLV is illustrated below. The Type=tTBD1 indicates an Inter-Domain link TLV Type. The Length indicates the size of the Inter-Domain Link Sub-TLVs.



An Inter-Domain link TLV describes a single inter-domain link. It comprises a number of inter-domain link sub-TLVs for the information described in section 4.1, which are the sub-TLVs defined in RFC 3630 or their equivalents except for the local IP address with mask length defined below.

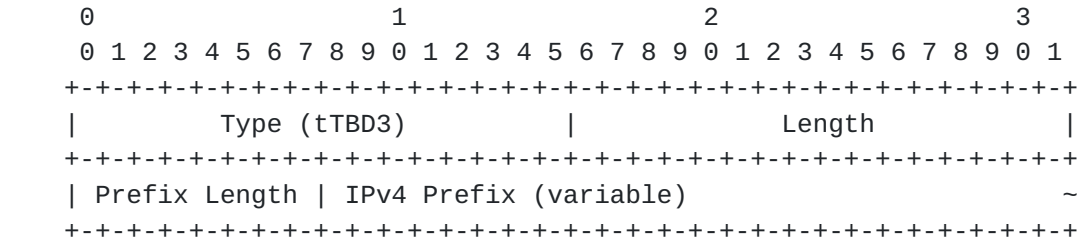
The format of the Router-ID TLV is shown below. The Type=tTBD2 indicates a Router-ID TLV Type. The Length indicates the size of the ID and flags field.



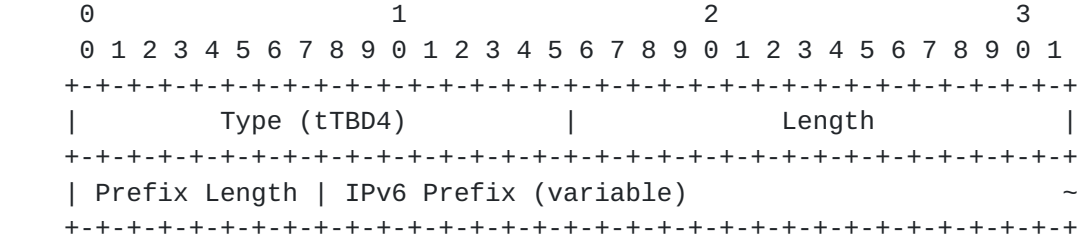
Flag B: Set to 1 indicating ABR (B is for Border)
Flag E: Set to 1 indicating ASBR (E is for External)
Flag I: Set to 1 indicating ID of local router (I is for ID)

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

The format of the Access IPv4 Prefix TLV is shown as follows. The Type=tTBD3 indicates an Access IPv4 Address Prefix TLV Type. The Length indicates the size of the IPv4 prefix and Prefix Length. The Prefix Length indicates the length of the IPv4 prefix. The IPv4 Prefix indicates an access IPv4 address prefix.

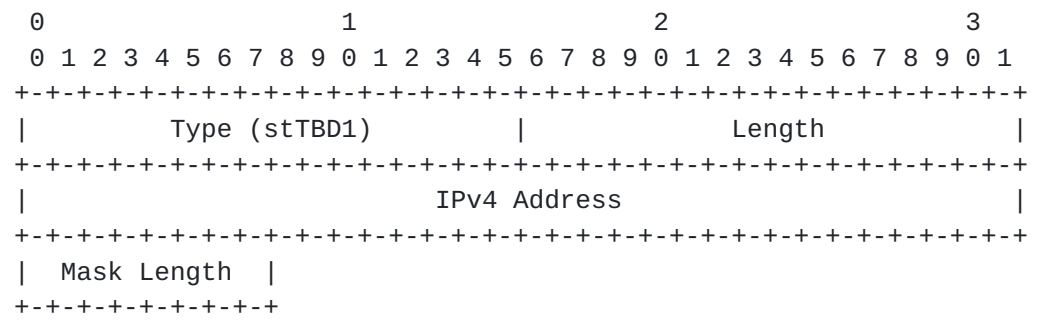


The format of the Access IPv6 Prefix TLV is illustrated below. The Type=tTBD4 indicates an Access IPv6 Address Prefix TLV Type. The Length indicates the size of the IPv6 prefix and Prefix Length. The Prefix Length indicates the length of the IPv6 prefix. The IPv6 Prefix indicates an access IPv6 address prefix.

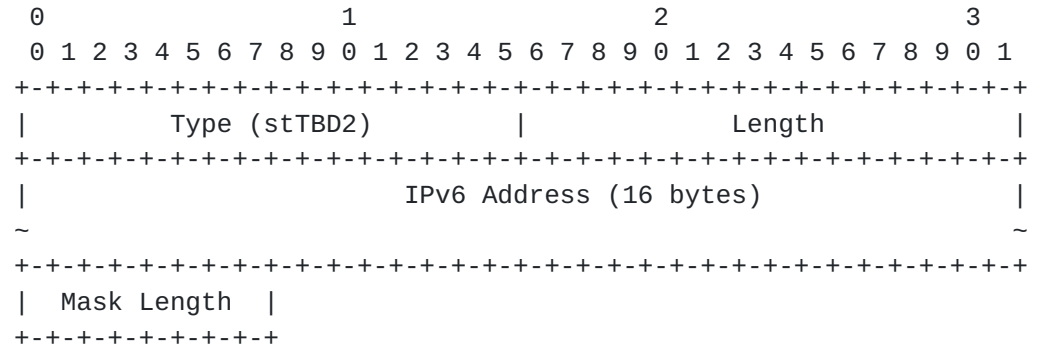


5.1.2. Sub-TLVs

The format of the Sub-TLV for a local IPv4 address with mask length is shown as follows. 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. Child Procedures

1. New or Changed Connections and Accesses

After a child discovers its parent, it sends the parent messages that contain the information about the connections (i.e., inter-domain links and ABRs) from its domain to its adjacent domains and the

access points in its domain.

When there are changes on the connections and the accesses of the domain, the child sends its parent messages for the changed connections and accesses.

For any new or changed inter-domain links, ABRs and access points in the domain for which a child is responsible, the child sends its parent messages containing the information about these links, ABRs and access points with indication of Updates on Connections and Accesses.

For example, for a new inter-domain point-to-point link from ASBR A in a child's domain to ASBR B in another domain, the child may send its parent a Notification message having a NOTIFICATION object with NT=8 and NV=1 (indicating Updates on Connections and Accesses), which contains a Router-ID TLV, followed by an Inter-domain link TLV. The Router-ID TLV comprises the ID of ASBR A and flag E and I set to 1. The Inter-domain link TLV comprises the Sub-TLVs for the information on 1) to 8) described in section 4.1.

For multiple new or changed inter-domain links from ASBR A, the child may send its parent a Notification message having a NOTIFICATION object with NT=8 and NV=1, which contains a Router-ID TLV for the ID of ASBR A, followed by multiple Inter-domain link TLVs for the links from ASBR A. If A is also an ABR, in addition to setting flag E and I to 1, the child sets flag B to 1 in the Router-ID TLV. Thus this one message contains the information about multiple inter-domains links and an ABR.

In another example, for a new or changed inter-domain broadcast link connected to ASBR X, an ABR Y and an access point 10.10.10.1/32 in a child's domain, the child may send its parent a Notification message having a NOTIFICATION object with NT=8 and NV=1, which contains a Router-ID TLV for the ID of ASBR X, followed by an Inter-domain link TLV for the link attached to ASBR X, another Router-ID TLV for the ID of ABR Y with flag B set to 1, and an Access IPv4 Prefix TLV with 10.10.10.1 and Prefix Length = 32. The Inter-domain link TLV comprises the Sub-TLVs for the information on a), b), 4), 5), 6), 7) and 8) described in section 4.1.

2. Connections and Accesses Down

For any inter-domain links, ABRs and access points down in the domain for which a child is responsible, the child sends its parent messages containing the information about these links, ABRs and access points with indication of Withdraw Connections and Accesses.

For example, for the inter-domain point-to-point link from ASBR A to ASBR B down, the child may send its parent a Notification message having a NOTIFICATION object with NT=8 and NV=2 (indicating Withdraw Connections and Accesses), which contains a Router-ID TLV, followed by an Inter-domain link TLV. The Router-ID TLV comprises the ID of ASBR A and flag E and I set to 1. The Inter-domain link TLV comprises the Sub-TLVs for the information on 1), 2) and 3) described in section 4.1.

For multiple inter-domain links from ASBR A down, the child may send its parent a Notification message having a NOTIFICATION object with NT=8 and NV=2, which contains a Router-ID TLV for the ID of ASBR A, followed by multiple Inter-domain link TLVs for the links from ASBR A. The TLV for a point-to-point link comprises the Sub-TLVs for the information on 1), 2) and 3) described in section 4.1. The TLV for a broadcast link comprises the Sub-TLVs for the information on a) and b) described in section 4.1.

3. Child as a Parent

If a parent P1 is also a child of another parent P2, P1 as a child sends its parent P2 a message containing the information about the connections and access points. P1 as a parent has the connections among its children's domains. But these connections are hidden from its parent P2. P1 may have connections from its children's domains to other domains. P1 as a child sends its parent P2 these connections.

P1 as a parent has the access points in its children's domains to be accessible outside of the domains. P1 as child may not send all of these to its parent P2. It sends its parent some of these access points according to some local policies.

From P2's point of view, its child P1 is responsible for one domain, which has some connections to its adjacent domains and some access points to be accessible.

5.2.2. Parent Procedures

A parent stores the connections and accesses for each of its children according to the messages for connections and accesses received. For a message containing updates on connections and accesses, it updates the connections and accesses accordingly. For a message containing withdraw connections and accesses, it removes the connections and accesses.

When a child is down, its parent removes the connections and accesses of the child's domain.

After receiving the messages for connections and accesses from its children, a parent builds and maintains a TED for the topology of its children's domains, in which each of the domains is seen as a cloud or a node. The information inside each of the domains is hidden from the parent. There are connections among the domains and the access points in the domains to be accessible in the topology.

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 Eric Wu and others 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, <http://www.rfc-editor.org/info/rfc2119>.

[RFC6805] King, D., Ed. and A. Farrel, Ed., "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS", RFC 6805, DOI 10.17487/RFC6805, November 2012, <http://www.rfc-editor.org/info/rfc6805>.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <http://www.rfc-editor.org/info/rfc5440>.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <http://www.rfc-editor.org/info/rfc3630>.

9.2. Informative References

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <http://www.rfc-editor.org/info/rfc5305>.

[RFC5392] Chen, M., Zhang, R., and X. Duan, "OSPF Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5392, DOI 10.17487/RFC5392, January 2009, <http://www.rfc-editor.org/info/rfc5392>.

[RFC5316] Chen, M., Zhang, R., and X. Duan, "ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5316, DOI 10.17487/RFC5316, December 2008, <http://www.rfc-editor.org/info/rfc5316>.

[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, <http://www.rfc-editor.org/info/rfc7752>.

Appendix A. New Message

A new message may be defined to advertise the connections and accesses from a child to its parent. The format of the message containing Connections and Access points (AC for short) is as follows:

```
  <AC Message> ::= <Common Header>
                   <NRP>
                   <Connection-List>
                   [<Access-Address-List>]
  where:
    <Connection-List> ::= <Connection>
                          [<Connection-List>]
    <Connection> ::= [<Inter-Domain-Link> | <ABR>]
    <Access-Address-List> ::= <Access-Address>
                              [<Access-Address-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 connections and access points are withdrawn. When flag W is set to one, the parent removes the connections and accesses

contained in the message after receiving it from its child. When flag W is set to zero, the parent adds/updates the connections and accesses in the message after receiving it.

An alternative to flag W in the NRP object is a similar flag in each CONNECTION and ACCESS object such as using one bit in Res flags for flag W. For example, when the flag is set to one in the object, the parent removes the connections and accesses in the object after receiving it. When the flag is set to zero in the object, the parent adds/updates the connections and accesses in the object after receiving it.

In another option, one byte in a CONNECTION and ACCESS Object is defined as flags field and one bit is used as flag W. The other undefined bits in the flags field MUST be set to zero.

The objects in the new message are defined below.

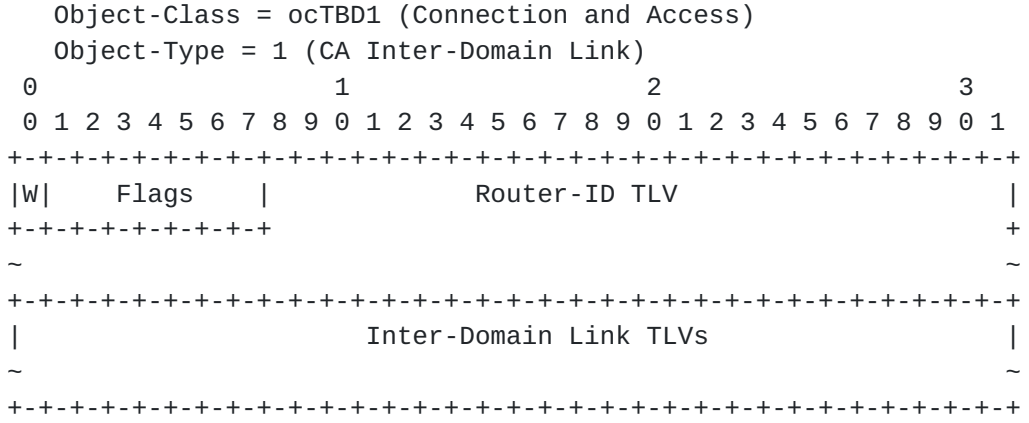
A.1. CONNECTION and ACCESS Object

A new object, called CONNECTION and ACCESS Object (CA for short), is defined. It has Object-Class ocTBD1. Four Object-Types are defined under CA object:

- CA Inter-Domain Link: CA Object-Type is 1.
- CA ABR: CA Object-Type is 2.
- CA Access IPv4 Prefix: CA Object-Type is 3.
- CA Access IPv6 Prefix: CA Object-Type is 4.

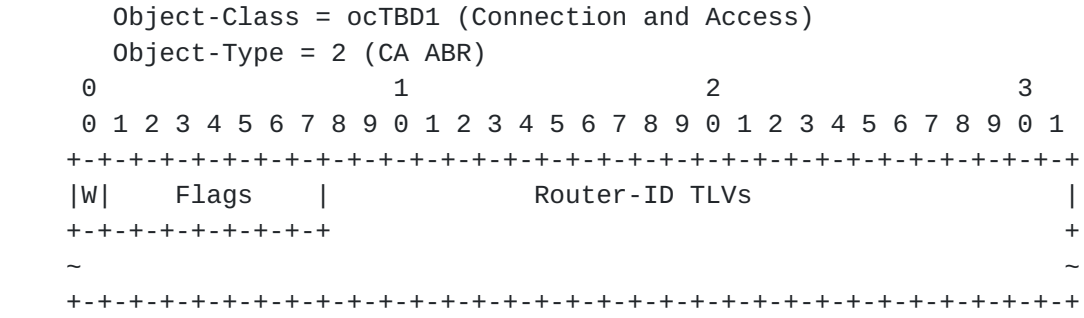
Each of these objects are described below.

The format of Inter-Domain Link object body is as follows:



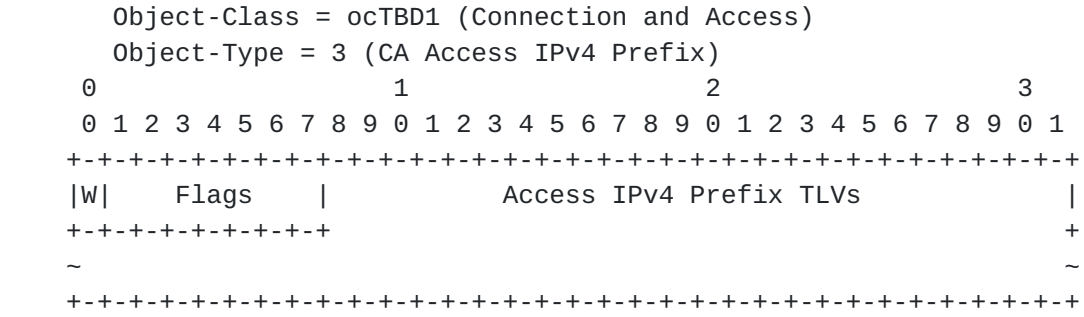
The Router-ID TLV indicates an ASBR in the domain, which is a local end of inter-domain links. Each of the Inter-Domain Link TLVs describes an inter-domain link and comprises a number of inter-domain link Sub-TLVs. Flag W=1 indicates withdraw the links. W=0 indicates new or changed links.

The format of ABR object body is illustrated below:



Each of the Router-ID TLVs indicates an ABR in the domain. Flag W=1 indicates withdraw the ABRs. W=0 indicates new ABRs.

The format of Access IPv4 Prefix object body is as follows:



Each of the Access IPv4 Prefix TLVs describes an access IPv4 address prefix in the domain, which is accessible to outside of the domain. Flag W=1 indicates withdraw the address prefixes. W=0 indicates new address prefixes.

The format of Access IPv6 Prefix object body is shown below:

```
   Object-Class = ocTBD1 (Connection and Access)
   Object-Type = 4 (CA Access IPv6 Prefix)
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|W|    Flags    |           Access IPv6 Prefix TLVs             |
+-+-+-+-+-+-+-+-+                                               +
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Each of the Access IPv6 Prefix TLVs describes an access IPv6 address prefix in the domain, which is accessible to outside. Flag W=1 indicates withdraw the prefixes. W=0 indicates new prefixes.

A.2. TLVs in CONNECTION and ACCESS Object

The format of the Router-ID TLV is illustrated below:

```
 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 (tTBD1)          |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           32-bit/48-bit ID                    ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

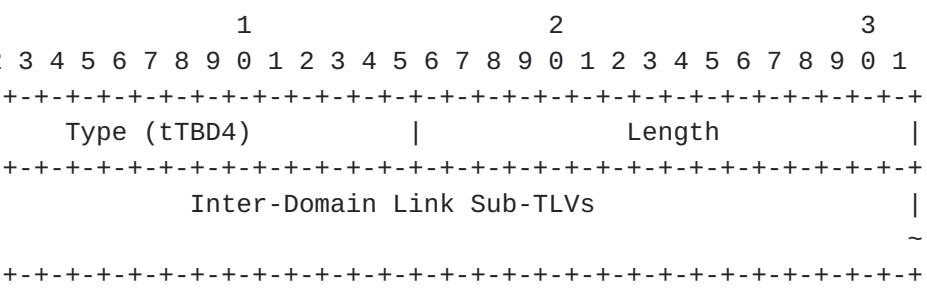
The format of the Access IPv4 Prefix TLV is shown as follows:

```
 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 (tTBD2)          |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Prefix Length | IPv4 Prefix (variable)                        ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The format of the Access IPv6 Prefix TLV is illustrated below:

```
 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 (tTBD3)          |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Prefix Length | IPv6 Prefix (variable)                        ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The format of the Inter-Domain link TLV is shown below:



The inter-domain link sub-TLVs are for the information on a link described in section 4.1, which are the sub-TLVs defined in RFC 3630 or their equivalents except for local IP address with mask length.

Authors' Addresses

Huaimo Chen
Huawei Technologies
Boston, MA,
USA

EMail: Huaimo.chen@huawei.com

Mehmet Toy
Comcast
1800 Bishops Gate Blvd.
Mount Laurel, NJ 08054
USA

EMail: mehmet_toy@cable.comcast.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