Domain Subobjects for Path Computation Element (PCE) Communication Protocol (PCEP).
draft-ietf-pce-pcep-domain-sequence-12

Abstract

The ability to compute shortest constrained Traffic Engineering Label Switched Paths (TE LSPs) in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks across multiple domains has been identified as a key requirement. In this context, a domain is a collection of network elements within a common sphere of address management or path computational responsibility such as an Interior Gateway Protocol (IGP) area or an Autonomous System (AS). This document specifies a representation and encoding of a Domain-Sequence, which is defined as an ordered sequence of domains traversed to reach the destination domain to be used by Path Computation Elements (PCEs) to compute inter-domain constrained shortest paths across a predetermined sequence of domains. This document also defines new subobjects to be used to encode domain identifiers.

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 June 9, 2016.
Copyright Notice

Copyright (c) 2015 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction ............................................. 3
   1.1. Scope ............................................ 4
   1.2. Requirements Language .............................. 4
2. Terminology ............................................. 5
3. Detail Description ...................................... 6
   3.1. Domains .......................................... 6
   3.2. Domain-Sequence ................................... 6
   3.3. Domain-Sequence Representation ..................... 7
   3.4. Include Route Object (IRO) ......................... 7
      3.4.1. Subobjects ................................... 8
         3.4.1.1. Autonomous system ....................... 8
         3.4.1.2. IGP Area .................................. 9
      3.4.2. Update in IRO specification ................... 10
      3.4.3. IRO for Domain-Sequence ...................... 10
         3.4.3.1. PCC Procedures ............................ 11
         3.4.3.2. PCE Procedures ............................ 11
   3.5. Exclude Route Object (XRO) ......................... 12
      3.5.1. Subobjects ................................... 13
         3.5.1.1. Autonomous system ....................... 13
         3.5.1.2. IGP Area .................................. 14
   3.6. Explicit Exclusion Route Subobject (EXRS) .......... 15
   3.7. Explicit Route Object (ERO) ......................... 16
4. Examples ............................................... 16
   4.1. Inter-Area Path Computation ....................... 16
   4.2. Inter-AS Path Computation .......................... 18
      4.2.1. Example 1 .................................... 19
      4.2.2. Example 2 .................................... 21
   4.3. Boundary Node and Inter-AS-Link .................... 23
   4.4. PCE Serving multiple Domains ....................... 24
   4.5. P2MP ............................................. 24
   4.6. Hierarchical PCE .................................. 26
A Path Computation Element (PCE) may be used to compute end-to-end paths across multi-domain environments using a per-domain path computation technique [RFC5152]. The backward recursive path computation (BRPC) mechanism [RFC5441] also defines a PCE-based path computation procedure to compute inter-domain constrained path for (G)MPLS TE LSPs. However, both per-domain and BRPC techniques assume that the sequence of domains to be crossed from source to destination is known, either fixed by the network operator or obtained by other means. Also for inter-domain point-to-multi-point (P2MP) tree computation, [RFC7334] assumes the domain-tree is known in priori.

The list of domains (Domain-Sequence) in point-to-point (P2P) or a domain tree in point-to-multipoint (P2MP) is usually a constraint in inter-domain path computation procedure.

The Domain-Sequence (the set of domains traversed to reach the destination domain) is either administratively predetermined or discovered by some means like H-PCE.

[RFC5440] defines the Include Route Object (IRO) and the Explicit Route Object (ERO). [RFC5521] defines the Exclude Route Object (XRO) and the Explicit Exclusion Route Subobject (EXRS). The use of Autonomous System (AS) (albeit with a 2-Byte AS number) as an abstract node representing a domain is defined in [RFC3209]. In the current document, we specify new subobjects to include or exclude domains including IGP area or an Autonomous Systems (4-Byte as per [RFC6793]).
Further, the domain identifier may simply act as delimiter to specify where the domain boundary starts and ends in some cases.

This is a companion document to Resource ReserVation Protocol - Traffic Engineering (RSVP-TE) extensions for the domain identifiers [DOMAIN-SUBOBJ].

1.1. Scope

The procedures described in this document are experimental. The experiment is intended to enable research for the usage of Domain-Sequence at the PCEs for inter-domain paths. For this purpose this document specifies new domain subobjects as well as how they incorporate with existing subobjects to represent a Domain-Sequence.

The experiment will end two years after the RFC is published. At that point, the RFC authors will attempt to determine how widely this has been implemented and deployed.

This document does not change the procedures for handling existing subobjects in PCEP.

The new subobjects introduced by this document will not be understood by legacy implementations. If a legacy implementation receives one of the subobjects that it does not understand in a PCEP object, the legacy implementation will behave as described in Section 3.4.3. Therefore, it is assumed that this experiment will be conducted only when both the PCE and the PCC form part of the experiment. It is possible that a PCC or PCE can operate with peers some of which form part of the experiment and some that do not. In this case, since no capabilities exchange is used to identify which nodes can use these extensions, manual configuration should be used to determine which peerings form part of the experiment.

When the result of implementation and deployment are available, this document will be updated and refined, and then be moved from Experimental to Standard Track.

1.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].
2. Terminology

The following terminology is used in this document.

ABR: OSPF Area Border Router. Routers used to connect two IGP areas.

AS: Autonomous System.

ASBR: Autonomous System Boundary Router.

BN: Boundary Node, Can be an ABR or ASBR.

BRPC: Backward Recursive Path Computation

Domain: As per [RFC4655], any collection of network elements within a common sphere of address management or path computational responsibility. Examples of domains include Interior Gateway Protocol (IGP) area and Autonomous System (AS).

Domain-Sequence: An ordered sequence of domains traversed to reach the destination domain.

ERO: Explicit Route Object

H-PCE: Hierarchical PCE


IRO: Include Route Object


OSPF: Open Shortest Path First.

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

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

P2MP: Point-to-Multipoint

P2P: Point-to-Point
RSVP: Resource Reservation Protocol

TE LSP: Traffic Engineering Label Switched Path.

XRO: Exclude Route Object

3. Detail Description

3.1. Domains

[RFC4726] and [RFC4655] define domain as a separate administrative or geographic environment within the network. A domain could be further defined as a zone of routing or computational ability. Under these definitions a domain might be categorized as an AS or an IGP area. Each AS can be made of several IGP areas. In order to encode a Domain-Sequence, it is required to uniquely identify a domain in the Domain-Sequence. A domain can be uniquely identified by area-id or AS number or both.

3.2. Domain-Sequence

A Domain-Sequence is an ordered sequence of domains traversed to reach the destination domain.

A Domain-Sequence can be applied as a constraint and carried in a path computation request to PCE(s). A Domain-Sequence can also be the result of a path computation. For example, in the case of Hierarchical PCE (H-PCE) [RFC6805], Parent PCE could send the Domain-Sequence as a result in a path computation reply.

In a P2P path, the domains listed appear in the order that they are crossed. In a P2MP path, the domain tree is represented as a list of Domain-Sequences.

A Domain-Sequence enables a PCE to select the next domain and the PCE serving that domain to forward the path computation request based on the domain information.

Domain-Sequence can include Boundary Nodes (ABR or ASBR) or Border links (Inter-AS-links) to be traversed as an additional constraint.

Thus a Domain-Sequence can be made up of one or more of -

- AS Number
- Area ID
- Boundary Node ID
These are encoded in the new subobjects defined in this document as well as the existing subobjects to represent a Domain-Sequence.

Consequently, a Domain-Sequence can be used:

1. by a PCE in order to discover or select the next PCE in a collaborative path computation, such as in BRPC [RFC5441];

2. by the Parent PCE to return the Domain-Sequence when unknown; this can then be an input to the BRPC procedure [RFC6805];

3. by a Path Computation Client (PCC) or a PCE, to constrain the domains used in inter-domain path computation, explicitly specifying which domains to be expanded or excluded;

4. by a PCE in the per-domain path computation model [RFC5152] to identify the next domain.

3.3. Domain-Sequence Representation

Domain-Sequence appears in PCEP messages, notably in -

- Include Route Object (IRO): As per [RFC5440], IRO can be used to specify a set of network elements to be traversed to reach the destination, which includes subobjects used to specify the Domain-Sequence.

- Exclude Route Object (XRO): As per [RFC5521], XRO can be used to specify certain abstract nodes, to be excluded from whole path, which includes subobjects used to specify the Domain-Sequence.

- Explicit Exclusion Route Subobject (EXRS): As per [RFC5521], EXRS can be used to specify exclusion of certain abstract nodes (including domains) between a specific pair of nodes. EXRS are a subobject inside the IRO.

- Explicit Route Object (ERO): As per [RFC5440], ERO can be used to specify a computed path in the network. For example, in the case of H-PCE [RFC6805], a Parent PCE can send the Domain-Sequence as a result, in a path computation reply using ERO.

3.4. Include Route Object (IRO)

As per [RFC5440], IRO (Include Route Object) can be used to specify that the computed path needs to traverse a set of specified network elements or abstract nodes.
3.4.1. Subobjects

Some subobjects are defined in [RFC3209], [RFC3473], [RFC3477] and [RFC4874], but new subobjects related to Domain-Sequence are needed.

This document extends the support for 4-Byte AS numbers and IGP Areas.

<table>
<thead>
<tr>
<th>Type</th>
<th>Subobject</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD1</td>
<td>Autonomous system number (4 Byte)</td>
</tr>
<tr>
<td>TBD2</td>
<td>OSPF Area id</td>
</tr>
<tr>
<td>TBD3</td>
<td>ISIS Area id</td>
</tr>
</tbody>
</table>

Note: The twins of these subobjects are carried in RSVP-TE messages as defined in [DOMAIN-SUBOBJ].

3.4.1.1. Autonomous system

[RFC3209] already defines 2 byte AS number.

To support 4 byte AS number as per [RFC6793] following subobject is defined:

```
+---------------+---------------+---------------+---------------+
| L  | Type    | Length | Reserved |
+---------------+---------------+---------------+---------------+
+---------------+---------------+---------------+---------------+
|                | AS-ID (4 bytes) |       |
```

L: The L bit is an attribute of the subobject as defined in [RFC3209] and usage in IRO subobject updated in [IRO-UPDATE].

Type: (TBD1 by IANA) indicating a 4-Byte AS Number.

Length: 8 (Total length of the subobject in bytes).

Reserved: Zero at transmission, ignored at receipt.

AS-ID: The 4-Byte AS Number. Note that if 2-Byte AS numbers are in use, the low order bits (16 through 31) MUST be used and the high order bits (0 through 15) MUST be set to zero.
3.4.1.2. IGP Area

Since the length and format of Area-id is different for OSPF and ISIS, following two subobjects are defined:

For OSPF, the area-id is a 32 bit number. The subobject is encoded as follows:

```
+------------------------------------------+
|     L        |     Type     |     Length    |         Reserved              |
|------------------------------------------|
| 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 |
```

L: The L bit is an attribute of the subobject as defined in [RFC3209] and usage in IRO subobject updated in [IRO-UPDATE].

Type: (TBD2 by IANA) indicating a 4-Byte OSPF Area ID.

Length: 8 (Total length of the subobject in bytes).

Reserved: Zero at transmission, ignored at receipt.

OSPF Area Id: The 4-Byte OSPF Area ID.

For IS-IS, the area-id is of variable length and thus the length of the Subobject is variable. The Area-id is as described in IS-IS by ISO standard [ISO10589]. The subobject is encoded as follows:

```
+------------------------------------------+
|     L        |     Type     |     Length    |  Area-Len     |  Reserved     |
|------------------------------------------|
|                                                |
+------------------------------------------+
```

L: The L bit is an attribute of the subobject as defined in [RFC3209] and usage in IRO subobject updated in [IRO-UPDATE].

Type: (TBD3 by IANA) indicating IS-IS Area ID.

Length: Variable. The Length MUST be at least 8, and MUST be a multiple of 4.
Area-Len: Variable (Length of the actual (non-padded) IS-IS Area Identifier in octets; Valid values are from 1 to 13 inclusive).

Reserved: Zero at transmission, ignored at receipt.

IS-IS Area Id: The variable-length IS-IS area identifier. Padded with trailing zeroes to a four-byte boundary.

3.4.2. Update in IRO specification

[RFC5440] describes IRO as an optional object used to specify network elements to be traversed by the computed path. It further state that the L bit of such subobject has no meaning within an IRO. It also did not mention if IRO is an ordered or un-ordered list of subobjects.

An update to IRO specification [IRO-UPDATE] makes IRO as an ordered list, as well as support for loose bit (L-bit) is added.

The use of IRO for Domain-Sequence, assumes the updated specification for IRO, as per [IRO-UPDATE].

3.4.3. IRO for Domain-Sequence

The subobject type for IPv4, IPv6, and unnumbered Interface ID can be used to specify Boundary Nodes (ABR/ASBR) and Inter-AS-Links. The subobject type for the AS Number (2 or 4 Byte) and the IGP Area are used to specify the domain identifiers in the Domain-Sequence.

The IRO can incorporate the new domain subobjects with the existing subobjects in a sequence of traversal.

Thus an IRO, comprising subobjects, that represents a Domain-Sequence, define the domains involved in an inter-domain path computation, typically involving two or more collaborative PCEs.

A Domain-Sequence can have varying degrees of granularity. It is possible to have a Domain-Sequence composed of, uniquely, AS identifiers. It is also possible to list the involved IGP areas for a given AS.

In any case, the mapping between domains and responsible PCEs is not defined in this document. It is assumed that a PCE that needs to obtain a "next PCE" from a Domain-Sequence is able to do so (e.g. via administrative configuration, or discovery).
3.4.3.1.  PCC Procedures

A PCC builds an IRO to encode the Domain-Sequence, so that the cooperating PCEs could compute an inter-domain shortest constrained path across the specified sequence of domains.

A PCC may intersperse Area and AS subobjects with other subobjects without change to the previously specified processing of those subobjects in the IRO.

3.4.3.2.  PCE Procedures

If a PCE receives an IRO in a Path Computation request (PCReq) message that contains the subobjects defined in this document, that it does not recognize, it will respond according to the rules for a malformed object as per [RFC5440]. The PCE MAY also include the IRO in the PCErr message as per [RFC5440].

The interpretation of Loose bit (L bit) is as per section 4.3.3.1 of [RFC3209] (as per [IRO-UPDATE]).

In a Path Computation reply (PCRep), PCE MAY also supply IRO (with Domain-Sequence information) with the NO-PATH object indicating that the set of elements (domains) of the request’s IRO prevented the PCEs from finding a path.

The following processing rules apply for Domain-Sequence in IRO –

- When a PCE parses an IRO, it interprets each subobject according to the AS number associated with the preceding subobject. We call this the "current AS". Certain subobjects modify the current AS, as follows.
  
  * The current AS is initialized to the AS number of the PCC.
  * If the PCE encounters an AS subobject, then it updates the current AS to this new AS number.
  * If the PCE encounters an Area subobject, then it assumes that the area belongs to the current AS.
  * If the PCE encounters an IP address that is globally routable, then it updates the current AS to the AS that owns this IP address. This document does not define how the PCE learns which AS owns the IP address.
  * If the PCE encounters an IP address that is not globally routable, then it assumes that it belongs to the current AS.
* If the PCE encounters an unnumbered link, then it assumes that it belongs to the current AS.

- When a PCE parses an IRO, it interprets each subobject according to the Area ID associated with the preceding subobject. We call this the "current Area". Certain subobjects modify the current Area, as follows.

- The current Area is initialized to the Area ID of the PCC.

- If the current AS is changed, the current Area is reset and need to be determined again by current or subsequent subobject.

- If the PCE encounters an Area subobject, then it updates the current Area to this new Area ID.

- If the PCE encounters an IP address that belongs to a different area, then it updates the current Area to the Area that has this IP address. This document does not define how the PCE learns which Area has the IP address.

- If the PCE encounters an unnumbered link that belongs to a different area, then it updates the current Area to the Area that has this link.

- Otherwise, it assumes that the subobject belongs to the current Area.

- In case the current PCE is not responsible for the path computation in the current AS or Area, then the PCE selects the "next PCE" in the domain-sequence based on the current AS and Area.

Note that it is advised that, PCC should use AS and Area subobject while building the domain-sequence in IRO and avoid using other mechanism to change the "current AS" and "current Area" as described above.

3.5. Exclude Route Object (XRO)

The Exclude Route Object (XRO) [RFC5521] is an optional object used to specify exclusion of certain abstract nodes or resources from the whole path.
3.5.1. Subobjects

Some subobjects to be used in XRO as defined in [RFC3209], [RFC3477], [RFC4874], and [RFC5520], but new subobjects related to Domain-Sequence are needed.

This document extends the support for 4-Byte AS numbers and IGP Areas.

<table>
<thead>
<tr>
<th>Type</th>
<th>Subobject</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD1</td>
<td>Autonomous system number (4 Byte)</td>
</tr>
<tr>
<td>TBD2</td>
<td>OSPF Area id</td>
</tr>
<tr>
<td>TBD3</td>
<td>ISIS Area id</td>
</tr>
</tbody>
</table>

Note: The twins of these subobjects are carried in RSVP-TE messages as defined in [DOMAIN-SUBOBJ].

3.5.1.1. Autonomous system

The new subobjects to support 4 byte AS and IGP (OSPF / ISIS) Area MAY also be used in the XRO to specify exclusion of certain domains in the path computation procedure.

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|X|    Type     |     Length    |         Reserved              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          AS-ID (4 bytes)                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The X-bit indicates whether the exclusion is mandatory or desired.

0: indicates that the AS specified MUST be excluded from the path computed by the PCE(s).

1: indicates that the AS specified SHOULD be avoided from the inter-domain path computed by the PCE(s), but MAY be included subject to PCE policy and the absence of a viable path that meets the other constraints.

All other fields are consistent with the definition in Section 3.4.
3.5.1.2. IGP Area

Since the length and format of Area-id is different for OSPF and ISIS, following two subobjects are defined:

For OSPF, the area-id is a 32 bit number. The subobject is encoded as follows:

```
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
|X|    Type     |     Length    |         Reserved              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
|                    OSPF Area Id (4 bytes)                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
```

The X-bit indicates whether the exclusion is mandatory or desired.

0: indicates that the OSPF Area specified MUST be excluded from the path computed by the PCE(s).

1: indicates that the OSPF Area specified SHOULD be avoided from the inter-domain path computed by the PCE(s), but MAY be included subject to PCE policy and the absence of a viable path that meets the other constraints.

All other fields are consistent with the definition in Section 3.4.

For IS-IS, the area-id is of variable length and thus the length of the subobject is variable. The Area-id is as described in IS-IS by ISO standard [ISO10589]. The subobject is encoded as follows:

```
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
|X|    Type     |     Length    |  Area-Len     |  Reserved     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
|                                                               |
//                        IS-IS Area ID                        //
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
```

The X-bit indicates whether the exclusion is mandatory or desired.

0: indicates that the ISIS Area specified MUST be excluded from the path computed by the PCE(s).
1: indicates that the ISIS Area specified SHOULD be avoided from the inter-domain path computed by the PCE(s), but MAY be included subject to PCE policy and the absence of a viable path that meets the other constraints.

All other fields are consistent with the definition in Section 3.4.

All the processing rules are as per [RFC5521].

Note that, if a PCE receives an XRO in a PCReq message that contains subobjects defined in this document, that it does not recognize, it will respond according to the rules for a malformed object as per [RFC5440].

IGP Area subobjects in the XRO are local to the current AS. In case of multi-AS path computation to exclude an IGP area in a different AS, IGP Area subobject should be part of Explicit Exclusion Route Subobject (EXRS) in the IRO to specify the AS in which the IGP area is to be excluded. Further policy may be applied to prune/ignore Area subobjects in XRO after "current AS" change during path computation.

3.6. Explicit Exclusion Route Subobject (EXRS)

EXRS [RFC5521] is used to specify exclusion of certain abstract nodes between a specific pair of nodes.

The EXRS subobject can carry any of the subobjects defined for inclusion in the XRO, thus the new subobjects to support 4 byte AS and IGP (OSPF / ISIS) Area can also be used in the EXRS. The meanings of the fields of the new XRO subobjects are unchanged when the subobjects are included in an EXRS, except that scope of the exclusion is limited to the single hop between the previous and subsequent elements in the IRO.

The EXRS subobject should be interpreted in the context of the current AS and current Area of the preceding subobject in the IRO. The EXRS subobject does not change the current AS or current Area. All other processing rules are as per [RFC5521].

Note that, if a PCE that supports the EXRS in an IRO, parses an IRO, and encounters an EXRS that contains subobjects defined in this document, that it does not recognize, it will act according to the setting of the X-bit in the subobject as per [RFC5521].
3.7. Explicit Route Object (ERO)

The Explicit Route Object (ERO) [RFC5440] is used to specify a computed path in the network. PCEP ERO subobject types correspond to RSVP-TE ERO subobject types as defined in [RFC3209], [RFC3473], [RFC3477], [RFC4873], [RFC4874], and [RFC5520]. The subobjects related to Domain-Sequence are further defined in [DOMAIN-SUBOBJ].

The new subobjects to support 4 byte AS and IGP (OSPF / ISIS) Area can also be used in the ERO to specify an abstract node (a group of nodes whose internal topology is opaque to the ingress node of the LSP). Using this concept of abstraction, an explicitly routed LSP can be specified as a sequence of domains.

In case of Hierarchical PCE [RFC6805], a Parent PCE can be requested to find the Domain-Sequence. Refer example in Section 4.6. The ERO in reply from parent PCE can then be used in Per-Domain path computation or BRPC.

If a PCC receives an ERO in a PCRep message that contains subobject defined in this document, that it does not recognize, it will respond according to the rules for a malformed object as per [RFC5440].

4. Examples

The examples in this section are for illustration purposes only; to highlight how the new subobjects could be encoded. They are not meant to be an exhaustive list of all possible usecases and combinations.

4.1. Inter-Area Path Computation

In an inter-area path computation where the ingress and the egress nodes belong to different IGP areas within the same AS, the Domain-Sequence could be represented using an ordered list of Area subobjects.
Figure 1: Inter-Area Path Computation
AS Number is 100.

If the ingress is in Area 2, egress in Area 4 and transit through Area 0. Some possible way a PCC can encode the IRO:

```
+---------+ +---------+ +---------+
| IRO     | | Sub     | | Sub     |
| Object  | | Object  | | Object  |
| Header  | | Area 0  | | Area 4  |
+---------+ +---------+ +---------+
```

or

```
+---------+ +---------+ +---------+ +---------+
| IRO     | | Sub     | | Sub     | | Sub     |
| Object  | | Object  | | Object  | | Object  |
| Header  | | Area 2  | | Area 0  | | Area 4  |
+---------+ +---------+ +---------+ +---------+
```

or

```
+---------+ +---------+ +---------+ +---------+ +---------+
| IRO     | | Sub     | | Sub     | | Sub     |
| Object  | | Object AS| | Object  | | Object  | | Object  |
| Header  | | 100      | | Area 2  | | Area 0  | | Area 4  |
+---------+ +---------+ +---------+ +---------+ +---------+
```

The Domain-Sequence can further include encompassing AS information in the AS subobject.

4.2. Inter-AS Path Computation

In inter-AS path computation, where ingress and egress belong to different AS, the Domain-Sequence could be represented using an ordered list of AS subobjects. The Domain-Sequence can further include decomposed area information in the Area subobject.
4.2.1. Example 1

As shown in Figure 2, where AS has a single area, AS subobject in the
domain-sequence can uniquely identify the next domain and PCE.

* All AS have one area (area 0)

Figure 2: Inter-AS Path Computation

If the ingress is in AS A, egress in AS C and transit through AS B.
Some possible way a PCC can encode the IRO:
Note that to get a domain disjoint path, the ingress could also request the backup path with -

```
+-------+ +-------+ +-------+ +-------+ +-------+ +-------+ +-------+
|IRO    | |Sub    | |Sub    | |Sub    | |Sub    | |Sub    |
|Object | |Object | |Object | |Object | |Object | |Object |
|Header | |AS B   | |AS C   | |       | |       | |       |
```

As described in Section 3.4.3, domain subobject in IRO changes the domain information associated with the next set of subobjects; till you encounter a subobject that changes the domain too. Consider the following IRO:

```
+-------+ +-------+ +-------+ +-------+ +-------+ +-------+ +-------+ +-------+
|IRO    | |Sub    | |Sub    | |Sub    | |Sub    | |Sub    | |Sub    |
|Object | |Object | |Object | |Object | |Object | |Object | |Object |
|Header | |AS B   | |IP     | |IP     | |AS C   | |IP     | |       |
|       | |B1     | |B3     | |       | |C1     | |       | |       |
```

On processing subobject "AS B", it changes the AS of the subsequent subobjects till we encounter another subobject "AS C" which changes the AS for its subsequent subobjects.

Consider another IRO:

```
+-------+ +-------+ +-------+ +-------+ +-------+
| IRO  | | Sub   | | Sub   | | Sub   | | Sub   |
| Object| | Object | | Object | | Object | | Object |
| Header| | AS D  | | IP    | | IP    | | IP    |
|       | |       | | D1    | | D3    | | C3    |
```

Here as well, on processing "AS D", it changes the AS of the subsequent subobjects till you encounter another subobject "C3" which belong in another AS and changes the AS for its subsequent subobjects.

Further description for the Boundary Node and Inter-AS-Link can be found in Section 4.3.

4.2.2. Example 2

In Figure 3, AS 200 is made up of multiple areas.

```
+-------------+                +----------------+
|  +-------------+                |          +--+  |
|  | Area 2       |                |          | B|  |
|  |         +--+|                |          |  |  |
|  |         |  ||                |          |  |  |
|  |  +--+   +--+|                |   +--+   +--+  |
|  |  |  |       |                |   |  |         |
|  |  +--+       |                |   +--+         |
|  |        +--+ |                |          +--+  |
|  |        |  | |                |          |  |  |
|  |        +--+ |                |   +--+   +--+  |
|  |  +--+       |+--------------+|   |  |         |
|  |  |  |       +--+          +--+   +--+         |
|  |  +--+  |   |----------------|  |            |
|  |    +--+  |   |          Inter-AS     +--+   +--+     |
|  |    |  |  +---+                +--+            |
|  |    +--+  |   |----------------|  |            |
|  |          +---+   Inter-AS     +--+   +--+     |
|  |+--+         ||    Links        |     |  |     |
+-------------+|  |  +--+       |  |          |  |                |
```

For LSP (A-B), where ingress A is in (AS 100, Area 0), egress B in (AS 200, Area 4) and transit through (AS 200, Area 0). Some possible way a PCC can encode the IRO:

```
+-------+ +-------+ +-------+ +-------+
| IRO   | | Sub    | | Sub    | | Sub    | | Sub    |
| Object | | Object | | Object | | Object | | Object |
| Header | | AS 200 | | Area 0 | | Area 4 |
+-------+ +-------+ +-------+ +-------+
```

or

```
+-------+ +-------+ +-------+ +-------+ +-------+ +-------+
| IRO   | | Sub    | | Sub    | | Sub    | | Sub    | | Sub    |
| Object | | Object | | Object | | Object | | Object | | Object |
| Header | | AS 100 | | Area 0 | | AS 200 | | Area 0 | | Area 4 |
+-------+ +-------+ +-------+ +-------+ +-------+ +-------+
```
For LSP (A-C), where ingress A is in (AS 100, Area 0), egress C in (AS 200, Area 5) and transit through (AS 200, Area 0). Some possible way a PCC can encode the IRO:

```
+-------+ +-------+ +-------+ +-------+
| IRO   | | Sub    | | Sub    | | Sub    |
| Object | | Object | | Object | | Object |
| Header | | AS 200 | | Area 0 | | Area 5 |
+-------+ +-------+ +-------+ +-------+
```

or

```
+-------+ +-------+ +-------+ +-------+ +-------+ +-------+
| IRO   | | Sub    | | Sub    | | Sub    | | Sub    |
| Object | | Object | | Object | | Object | | Object |
| Header | | AS 100 | | Area 0 | | AS 200 | | Area 0 | | Area 5 |
+-------+ +-------+ +-------+ +-------+ +-------+ +-------+
```

4.3. Boundary Node and Inter-AS-Link

A PCC or PCE can include additional constraints covering which Boundary Nodes (ABR or ASBR) or Border links (Inter-AS-link) to be traversed while defining a Domain-Sequence. In which case the Boundary Node or Link can be encoded as a part of the Domain-Sequence.

Boundary Nodes (ABR / ASBR) can be encoded using the IPv4 or IPv6 prefix subobjects usually the loopback address of 32 and 128 prefix length respectively. An Inter-AS link can be encoded using the IPv4 or IPv6 prefix subobjects or unnumbered interface subobjects.

For Figure 1, an ABR (say 203.0.113.1) to be traversed can be specified in IRO as:

```
+---------+ +---------+ +---------+ +---------+
| IRO     | | Sub     | | Sub     | | Sub     |
| Object   | | Object   | | Object   | | Object   |
| Header   | | Area 2   | | Object   | | IPv4     |
|          | |         | | Area 0   | | Area 0   |
|          | |         | |          | | Area 4   |
|          | |         | | 203.0.   | |          |
|          | |         | | 112.1    | |          |
+---------+ +---------+ +---------+ +---------+
```
For Figure 3, an inter-AS-link (say 198.51.100.1 - 198.51.100.2) to be traversed can be specified as:

```
+---------+  +---------+ +---------+ +---------+
|IRO      |  |Sub      | |Sub      | |Sub      |
|Object   |  |Object AS| |Object AS| |Object AS|
|Header   |  |100      | |IPv4     | |200      |
|         |  |         | |198.51.  | |         |
|         |  |         | |100.2    | |         |
+---------+  +---------+ +---------+ +---------+
```

4.4. PCE Serving multiple Domains

A single PCE can be responsible for multiple domains; for example PCE function deployed on an ABR could be responsible for multiple areas. A PCE which can support adjacent domains can internally handle those domains in the Domain-Sequence without any impact on the other domains in the Domain-Sequence.

4.5. P2MP

[RFC7334] describes an experimental inter-domain P2MP path computation mechanism where the path domain tree is described as a series of Domain-Sequences, an example is shown in the below figure:
The domain tree can be represented as a series of domain-sequence:

- Domain D1, Domain D3, Domain D6
- Domain D1, Domain D3, Domain D5
- Domain D1, Domain D2, Domain D4

The domain sequence handling described in this document could be applied to P2MP path domain tree.
4.6. Hierarchical PCE

In case of H-PCE [RFC6805], the parent PCE can be requested to determine the Domain-Sequence and return it in the path computation reply, using the ERO. For the example in section 4.6 of [RFC6805], the Domain-Sequence can possibly appear as:

```
+---------+ +---------+ +---------+ +---------+
|ERO      | |Sub      | |Sub      | |Sub      |
|Object   | |Object   | |Object   | |Object   |
|Header   | |Domain 1 | |Domain 2 | |Domain 3 |
|         | |         | |         | |         |
|         | |         | |         | |         |
+---------+ +---------+ +---------+ +---------+
```

or

```
+---------+ +---------+ +---------+
|ERO      | |Sub      | |Sub      |
|Object   | |Object   | |Object   |
|Header   | |BN 21    | |Domain 3 |
|         | |         | |         |
|         | |         | |         |
+---------+ +---------+ +---------+
```

5. Other Considerations

5.1. Relationship to PCE Sequence

Instead of a Domain-Sequence, a sequence of PCEs MAY be enforced by policy on the PCC, and this constraint can be carried in the PCReq message (as defined in [RFC5886]).

Note that PCE-Sequence can be used along with Domain-Sequence in which case PCE-Sequence MUST have higher precedence in selecting the next PCE in the inter-domain path computation procedures.

5.2. Relationship to RSVP-TE

[RFC3209] already describes the notion of abstract nodes, where an abstract node is a group of nodes whose internal topology is opaque to the ingress node of the LSP. It further defines a subobject for AS but with a 2-Byte AS Number.

[DOMAIN-SUBOBJ] extends the notion of abstract nodes by adding new subobjects for IGP Areas and 4-byte AS numbers. These subobjects can
be included in Explicit Route Object (ERO), Exclude Route object (XRO) or Explicit Exclusion Route Subobject (EXRS) in RSVP-TE.

In any case subobject type defined in RSVP-TE are identical to the subobject type defined in the related documents in PCEP.

6. IANA Considerations

6.1. New Subobjects

IANA maintains the "Path Computation Element Protocol (PCEP) Numbers" at <http://www.iana.org/assignments/pcep>. Within this registry IANA maintains two sub-registries:

o IRO Subobjects (see IRO Subobjects at http://www.iana.org/assignments/pcep)

o XRO Subobjects (see XRO Subobjects at http://www.iana.org/assignments/pcep)

Upon approval of this document, IANA is requested to make identical additions to these registries as follows:

<table>
<thead>
<tr>
<th>Subobject Type</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD1 4 byte AS number</td>
<td>[This I.D.][DOMAIN-SUBOBJ]</td>
</tr>
<tr>
<td>TBD2 OSPF Area ID</td>
<td>[This I.D.][DOMAIN-SUBOBJ]</td>
</tr>
<tr>
<td>TBD3 IS-IS Area ID</td>
<td>[This I.D.][DOMAIN-SUBOBJ]</td>
</tr>
</tbody>
</table>

Further upon approval of this document, IANA is requested to add a reference to this document to the new RSVP numbers that are registered by [DOMAIN-SUBOBJ].

7. Security Considerations

The protocol extensions defined in this document do not substantially change the nature of PCEP. Therefore, the security considerations set out in [RFC5440] apply unchanged. Note that further security considerations for the use of PCEP over TCP are presented in [RFC6952].

This document specifies a representation of Domain-Sequence and new subobjects, which could be used in inter-domain PCE scenarios as explained in [RFC5152], [RFC5441], [RFC6805], [RFC7334] etc. The security considerations set out in each of these mechanisms remain unchanged by the new subobjects and Domain-Sequence representation in this document.
But the new subobjects do allow finer and more specific control of the path computed by a cooperating PCE(s). Such control increases the risk if a PCEP message is intercepted, modified, or spoofed because it allows the attacker to exert control over the path that the PCE will compute or to make the path computation impossible. Consequently, it is important that implementations conform to the relevant security requirements of [RFC5440]. These mechanisms include:

- Securing the PCEP session messages using TCP security techniques (Section 10.2 of [RFC5440]). PCEP implementations SHOULD also consider the additional security provided by the TCP Authentication Option (TCP-AO) [RFC5925] or [PCEPS].

- Authenticating the PCEP messages to ensure the message is intact and sent from an authorized node (Section 10.3 of [RFC5440]).

- PCEP operates over TCP, so it is also important to secure the PCE and PCC against TCP denial-of-service attacks. Section 10.7.1 of [RFC5440] outlines a number of mechanisms for minimizing the risk of TCP-based denial-of-service attacks against PCEs and PCCs.

- In inter-AS scenarios, attacks may be particularly significant with commercial as well as service-level implications.

Note, however, that the Domain-Sequence mechanisms also provide the operator with the ability to route around vulnerable parts of the network and may be used to increase overall network security.

8. Manageability Considerations

8.1. Control of Function and Policy

The exact behaviour with regards to desired inclusion and exclusion of domains MUST be available for examination by an operator and MAY be configurable. Manual configurations is needed to identify which PCEP peers understand the new domain subobjects defined in this document.

8.2. Information and Data Models

A MIB module for management of the PCEP is being specified in a separate document [RFC7420]. This document does not imply any new extension to the current MIB module.
8.3. Liveness Detection and Monitoring

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

8.4. Verify Correct Operations

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

8.5. Requirements On Other Protocols

In case of per-domain path computation [RFC5152], where the full path of an inter-domain TE LSP cannot be, or is not determined at the ingress node, a signaling message can use the domain identifiers. The Subobjects defined in this document SHOULD be supported by RSVP-TE. [DOMAIN-SUBOBJ] extends the notion of abstract nodes by adding new subobjects for IGP Areas and 4-byte AS numbers.

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

8.6. Impact On Network Operations

The mechanisms described in this document can provide the operator with the ability to exert finer and more specific control of the path computation by inclusion or exclusion of domain subobjects. There may be some scaling benefit when a single domain subobject may substitute for many subobjects and can reduce the overall message size and processing.

Backward compatibility issues associated with the new subobjects arise when a PCE does not recognize them, in which case PCE responds according to the rules for a malformed object as per [RFC5440]. For successful operations the PCEs in the network would need to be upgraded.

9. Acknowledgments

Authors would like to especially thank Adrian Farrel for his detailed reviews as well as providing text to be included in the document.

Further, we would like to thank Pradeep Shastry, Suresh Babu, Quintin Zhao, Fatai Zhang, Daniel King, Oscar Gonzalez, Chen Huaimo,
Venugopal Reddy, Reeja Paul, Sandeep Boina, Avantika Sergio Belotti and Jonathan Hardwick for their useful comments and suggestions.

Thanks to Jonathan Hardwick for shepherding this document.

Thanks to Deborah Brungard for being the Responsible AD.

Thanks to Amanda Baber for IANA Review.

Thanks to Joel Halpern for Gen-ART Review.

Thanks to Klaas Wierenga for SecDir Review.

Thanks to Spencer Dawkins and Barry Leiba for comments during the IESG Review.

10. References

10.1. Normative References


10.2. Informative References


Authors’ Addresses

Dhruv Dhody
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka  560037
India

EMail: dhruv.ietf@gmail.com

Udayasree Palle
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka  560037
India

EMail: udayasree.palle@huawei.com

Ramon Casellas
CTTC
Av. Carl Friedrich Gauss n7
Castelldefels, Barcelona  08860
Spain

EMail: ramon.casellas@cttc.es