PCE Working Group
Internet-Draft
Intended status: Experimental
Expires: January 6, 2013

D. Dhody
U. Palle
Huawei Technologies India Pvt Ltd
R. Casellas
CTTC - Centre Tecnologic de Telecomunicacions de Catalunya

July 5, 2012

Standard Representation Of Domain Sequence
draft-ietf-pce-pcep-domain-sequence-01

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 for P2P and P2MP scenarios. 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 IGP area or an Autonomous Systems. This document specifies a standard representation and encoding of a domain sequence, which is defined as an ordered sequence of domains traversed to reach the destination domain. This document also defines new sub-objects to be used to encode domain identifiers.

Table of Contents

1. Introduction

A PCE may be used to compute end-to-end paths across multi-domain environments using a per-domain path computation technique [RFC5152]. The so called backward recursive path computation (BRPC) mechanism [RFC5441] defines a PCE-based path computation procedure to compute inter-domain constrained (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. For inter-domain point-to-multi-point (P2MP) tree, [PCE-P2MP-PROCEDURES] assumes the domain-tree is known.

The list of domains in a point-to-point (P2P) path or a point-to-multi-point (P2MP) tree is usually a constraint in the path computation request. The PCE decouples the domain to determine the next PCE to forward the request.

According to BRPC mechanism the PCC MAY indicate the sequence of domains to be traversed using the Include Route Object (IRO) defined in [RFC5440].

This document proposes a standard way to represent and encode a domain sequence using IRO in various deployment scenarios including P2P, P2MP and Hierarchical PCE (HPCE) [PCE-HIERARCHY-FWK].

The domain sequence (the set of domains traversed to reach the destination domain) is either administratively predetermined or discovered by some means (H-PCE) that is outside of the scope of this document. Here the focus is only on a standard representation of the domain sequence in all possible scenarios.

1.1. 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: Any collection of network elements within a common sphere of address management or path computational responsibility. Examples of domains include Interior Gateway Protocol (IGP) areas and Autonomous Systems (ASs).

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

ERO: Explicit Route Object

H-PCE: Hierarchical PCE

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

IRO: Include Route Object

IS-IS: Intermediate System to Intermediate System.

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.

3. Detail Description

3.1. Domains

A domain can be defined as a separate administrative or geographic environment within the network. A domain may be further defined as a zone of routing or computational ability. Under these definitions a domain might be categorized as an Autonomous System (AS) or an Interior Gateway Protocol (IGP) area (as per [RFC4726] and [RFC4655]). To uniquely identify a domain in the domain sequence both AS and Area-id MAYBE important.

3.2. Domain-Sequence

A domain-sequence is an ordered sequence of domains traversed to reach the destination domain. In this context a Domain could be an Autonomous System (AS) or an IGP Area. Note that an AS can be further made of multiple Areas.

Domain Sequence can be applied as a constraint and carried in path computation request to PCE(s). In case of HPCE [PCE-HIERARCHY-FWK] Parent PCE MAY send the domain sequence as a result in path computation reply.

In this context, ordered sequence is important, 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 list of domain sequences.

One main goal of the Domain-Sequence is to enable a PCE to select the next PCE to forward the path computation request based on the domain information.

A PCC or PCE MAY add an additional constraints covering which Boundary Nodes (ABR or ASBR) or Border links (Inter-AS-link) MUST be traversed while defining a domain sequence.

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

- o AS Number
- o Area ID
- o Boundary Node ID
- o Inter-AS-Link Address

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 further be an input to BRPC procedure;

3. By a PCC (or PCE) to constraint the domains used in a H-PCE path computation, explicitly specifying which domains to be expanded;

3.3. Standard Representation

3.3.1. New Sub-Objects

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

3.3.1.1. Autonomous system

[RFC3209] already defines 2 octet AS number.

To support 4 octet AS number as per [RFC4893] following subobject is defined:

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

L: The L bit is an attribute of the subobject as define in [RFC3209].

Type: (TBA by IANA) indicating 4 octet AS Number.

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

Reserved: Zero at transmission, Ignored at receipt.

AS-ID: The 4 octet AS Number. Note that if 16-bit AS numbers are in use, the low order bits (16 through 31) should be used and the high order bits (0 through 15) should be set to zero.

3.3.1.2. 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 looks

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type     |     Length    |         Reserved              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    OSPF Area Id (4 bytes)                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

L: The L bit is an attribute of the subobject as define in [RFC3209].

Type: (TBA by IANA) indicating 4 octet OSPF Area ID.

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

Reserved: Zero at transmission, Ignored at receipt.

OSPF Area Id: The 4 octet 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 [ISO 10589].

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type     |     Length    |  Area-Len     |    Reserved   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                         IS-IS Area ID                       //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

L: The L bit is an attribute of the subobject as define in [RFC3209].

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

Length: Variable (Total length of the subobject in bytes including padding). The Length MUST be at least 4, 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 2 to 11 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-octet boundary.

3.3.2. Use in PCEP Objects

These sub-objects MAYBE used in -

- o Include Route Object (IRO): As per [RFC5440], used to specify set of network elements that MUST be traversed. These subobjects are used to specify the domain-sequence that MUST be traversed to reach the destination.

- o Exclude Route Object (XRO): As per [RFC5521], used to specify certain abstract nodes that MUST be excluded from whole path. These subobjects are used to specify certain domains that MUST be avoided to reach the destination.

- o Explicit Route Object (ERO): As per [RFC5440],used to specify a computed path in the network. These subobjects are used to specify the domain-sequence when computed by a Parent PCE ([PCE-HIERARCHY-FWK]).

- o Explicit Exclusion Route Sub-Object (EXRS): As per [RFC5521], used to specify exclusion of certain abstract nodes between a specific pair of nodes. EXRS are a sub-object inside the IRO. These subobjects are used to specify the domains that must be excluded between two abstract nodes.

3.3.2.1. Include Route Object

3.3.2.1.1. Option 1: New IRO Object Type

The IRO (Include Route Object) [RFC5440] is an optional object used to specify a set of specified network elements that the computed path MUST traverse. [RFC5440] in its description of IRO does not constrain the sub-objects to be in a given particular order. When considering a domain sequence, the domain relative ordering is a basic criterion and, as such, this document specifies a new IRO object type.

We define a new type of IRO Object to define Domain Sequence.

IRO Object-Class is 10.
IRO Object-Type is TBD. (2 suggested value to IANA)

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                        (Subobjects)                         //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Sub-objects: The IRO is made of sub-objects identical to the ones defined in [RFC3209], [RFC3473], and [RFC3477], where the IRO sub-object type is identical to the sub-object type defined in the related documents. Some new sub-objects related to Domain-Sequence are also added in this document.

The following sub-object types are used.

Type	Sub-object
1	IPv4 prefix
2	IPv6 prefix
4	Unnumbered Interface ID
32	Autonomous system number (2 Byte)
33	Explicit Exclusion (EXRS)
TBD	Autonomous system number (4 Byte)
TBD	OSPF Area id
TBD	ISIS Area id

[RFC3209] defines sub-objects for IPv4, IPv6 and unnumbered Interface ID, which in the context of domain-sequence is used to specify Boundary Node (ABR/ASBR) and Inter-AS-Links. The sub-objects for AS Number (2 or 4 Byte) and IGP Area is used to specify the domains in the domain-sequence.

The new IRO Object-Type used to define the domain-sequence would handle the L bit (Loose / Strict) in the sub-objects.

Note that PCReq message is free to carry both type of IRO with IRO Type 1 ([RFC5440]) used to specify the intra-domain abstract nodes and resources and the new IRO Type as described in this document to specify the domain-sequence.

All other rules of PCEP objects and message processing (ex. P bit handling of Common Object Header) is as per [RFC5440].

3.3.2.1.1.1. Mode of Operation

A domain sequence IRO object constraints or defines the domains involved in a multi-domain path computation, typically involving two or more collaborative PCEs.

A domain sequence can have varying degrees on granularity; it is possible to have a domain sequence composed of, uniquely, AS identifiers. It is also possible to list the involved 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).

A PCC builds a domain sequence IRO (new type) to encode the domain sequence, that is all domains that it wishes the cooperating PCEs to traverse in order to compute the end to end path.

For each inclusion, the PCC clears the L-bit to indicate that the PCE is required to include the domain, or sets the L-bit to indicate that the PCC simply desires that the domain be included in the domain-sequence.

When a PCE receives a PCReq message it looks for a domain sequence IRO (new type) to see if domain-sequence are required. If the PCE finds more than one domain sequence IRO (new type), it MUST use the first one in the message and MUST ignore subsequent instances.

If the PCE does not recognize the domain sequence IRO (new type), it MUST return a PCErr message with Error-Type "Unknown Object" and Error-value "Unrecognized object Type" as described in [RFC5440].

If the PCE is unwilling or unable to process the domain sequence IRO (new type), it MUST return a PCErr message with the Error-Type "Not supported object" and follow the relevant procedures described in [RFC5440].

If a PCE that supports the domain sequence IRO (new type) and encounters a subobject that it does not support or recognize, it MUST act according to the setting of the L-bit in the subobject. If the L-bit is clear, the PCE MUST respond with a PCErr with Error-Type "Unrecognized subobject" and set the Error-Value to the subobject type code. If the L-bit is set, the PCE MAY respond with a PCErr as already stated or MAY ignore the subobject: this choice is a local policy decision.

If a PCE parses a domain sequence IRO (new type) and encounters these subobject that it recognizes, it MUST act according to the requirements expressed in the subobject. That is, if the L-bit is clear, the PCE(s) MUST produce a path that follows domain-sequence nodes in order identified by the sub-objects in the path. If the L-bit is set, the PCE(s) SHOULD produce a path along the domain sequence unless it is not possible to construct a path complying with the other constraints expressed in the PCReq message.

A successful path computation reported in a PCRep message MUST include an ERO to specify the path that has been computed as specified in [RFC5440] following the domain-sequence.

When a PCE returns a path in a PCRep, it MAY also supply a domain sequence IRO (new type) in a PCRep message with the NO-PATH object indicates that the set of elements of the original domain sequence IRO prevented the PCE from finding a path.

Sub-Object types for AS and IGP Area guide the next domain selection and finding the PCE serving that domain.

Note that a particular domain in the domain-sequence can be identified by -

- o Just Area: Only the IGP (OSPF or ISIS) Area subobject is used to identify the next domain. (Refer Figure 1)

- o Just AS: Only the AS subobject is used to identify the next domain. (Refer Figure 2)

- o AS and IGP Area: Combination of both AS and Area are used to identify the next domain. In this case the order is AS Subobject followed by Area. (Refer Figure 3)

Sub-Object of other types representing Boundary Node or Inter-As-Link do not pay any role in selection of next domain and subsequently PCE selection in the domain-sequence. But they MUST be applied during the final path computation procedure as before.

3.3.2.1.2. Option 2: Same IRO Object Type

The IRO (Include Route Object) [RFC5440] is an optional object used to specify a set of specified network elements that the computed path MUST traverse.

The new sub-objects denoting the domain-sequence is carried in the same IRO Type 1, and all the rules of processing as specified in [RFC5440] are applied.

Note the following differences -

- o Order: Since there is no inherent order specified in the encoding of the subobjects in IRO Type 1 [RFC5440]. It is the job of PCE to figure out the order of the domains to be crossed to reach the destination domain. Note that in case of doubt, or when applicable, PCE can still apply the ordering as specified in the PCReq message.

- o Loose / Strict: [RFC5440] state that the L bit of the sub-objects within an IRO Type 1 [RFC5440] has no meaning. This is applicable for sub-objects denoting domain-sequence as well.

- o Scope: Sub-objects referring to domains and boundary nodes will mix with subobjects for internal network nodes of multiple domains. It is the job of PCE to figure out the scope and apply the processing rules accordingly. The PCE should distinguish between - the subobject is unknown (not in TED) or known but the computation fails. The PCE processing the IRO MAY include as many of the elements of the IRO as possible. If the PCE is passing the request onwards, it is OK for it to have unknown nodes, and it can assume that the next PCE might be able to satisfy the remaining elements of the IRO. On the other hand, if the PCE is making an end-to-end (or edge-to-edge, or end-to-edge) path and will return the response to a PCC (rather than pass it on) then the PCE must fail if it cannot satisfy the IRO. Ultimately, when the path segments are aggregated by a head-end PCE or by a parent PCE, that PCE can check to see whether any elements of the IRO are still missing and handle accordiangly.

3.3.2.2. Exclude Route Object

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

The following subobject types are defined to be used in XRO as defined in [RFC3209], [RFC3477], [RFC4874], and [RFC5521].

Type	Sub-object
1	IPv4 prefix
2	IPv6 prefix
4	Unnumbered Interface ID
32	Autonomous system number (2 Byte)
34	SRLG
64	IPv4 Path Key
65	IPv6 Path Key
TBD	Autonomous system number (4 Byte)
TBD	OSPF Area id
TBD	ISIS Area id

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

The X-bit indicates whether the exclusion is mandatory or desired. 0 indicates that the domain specified MUST be excluded from the path computed by the PCE(s). 1 indicates that the domain specified SHOULD be excluded 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 and excludes the domain. All other fields are consistent with the definition in Section 3.3.1.

4 Octet Autonomous system:

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

OSPF Area:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|X|    Type     |    Length     |         Reserved              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    OSPF Area Id (4 bytes)                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IS-IS Area:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|X|    Type     |     Length    |  Area-Len     |    Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                        IS-IS Area ID                        //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

If a PCE that supports XRO and encounters a subobject that it does not support or recognize, it MUST act according to the setting of the X-bit in the subobject. If the X-bit is clear, the PCE MUST respond with a PCErr with Error-Type "Unrecognized subobject" and set the Error-Value to the subobject type code. If the X-bit is set, the PCE MAY respond with a PCErr as already stated or MAY ignore the subobject: this choice is a local policy decision.

All the other processing rules are as per [RFC5521].

3.3.2.3. Explicit Route Object

The Explicit Route Object (ERO) [RFC5440] is used to specify a computed path in the network. PCEP ERO sub-object types correspond to RSVP-TE ERO sub-object types as defined in [RFC3209], [RFC3473], [RFC3477], [RFC4873], [RFC4874], and [RFC5520].

Type	Sub-object
1	IPv4 prefix
2	IPv6 prefix
3	Label
4	Unnumbered Interface ID
32	Autonomous system number (2 Byte)
33	Explicit Exclusion (EXRS)
37	Protection
64	IPv4 Path Key
65	IPv6 Path Key
TBD	Autonomous system number (4 Byte)
TBD	OSPF Area id
TBD	ISIS Area id

The new subobjects to support 4 octet AS and IGP (OSPF / ISIS) Area MAY 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, a Parent PCE ([PCE-HIERARCHY-FWK]) MAY be requested to find the domain-sequence. The Parent PCE MUST use ERO with AS and IGP Area subobjects to encode the computed domain-sequence. Refer example in Section 3.4.6.

3.3.2.4. Explicit Exclusion Route Sub-Object

Explicit Exclusion Route Sub-Object (EXRS) [RFC5521] is used to specify exclusion of certain abstract nodes between a specific pair of nodes.

The EXRS subobject may carry any of the subobjects defined for inclusion in the XRO, thus the new subobjects to support 4 octet AS and IGP (OSPF / ISIS) Area MAY 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.

All the processing rules are as per [RFC5521].

3.4. Other Considerations

3.4.1. Inter-Area Path Computation

In an inter-area path computation where ingress and egress belong to different IGP area, the domain sequence MAYBE represented using a ordered list of AREA sub-objects. AS number MAYBE skipped, as area information is enough to select the next PCE.

```
 +-------------------+                            +-------------------+
 |                   |                            |                   |
 |           +--+    |                            |     +--+          |
 |  +--+     |  |    |                            |     |  |          |
 |  |  |     +--+    |                            |     +--+   +--+   |
 |  +--*             +                            +            |  |   |
 |                   |                            |            +--+   |
 |         *--+      +                            +                   |
 |         |  |      |                            |     +--+          |
 |         +--+      |                            |     |  |          |
 |                   |+--------------------------+|     +--+          |
 |                  ++++                       +-++                   |
 |                  ||||         +--+          | ||                   |
 | Area 2           ++++         |  |          +-++  Area 4           |
 +-------------------+|          +--+           |+-------------------+
                      |                          |
                      |                +--+      |
                      |    +--+        |  |      |
                      |    |  |        +--+      |
                      |    +--+                  |
                      |                          |
                      |                          |
                      |                          |
                      |                          |
                      |           +--+           |
                      |           |  |           |
                      |           +--+           |
  +------------------+|                          |+--------------------+
  |                  ++-+                      +-++                    |
  |                  || |                      | ||                    |
  |                  ++-+    Area 0            +-++                    |
  |                  |+--------------------------+|     +--+           |
  |         +--+    |                            |     |  |           |
  |         |  |    |                            |     +--+           |
  | +--+     +--+    |                            |                    |
  | |  |             +                            +            +--+    |
  | +--+             |                            |            |  |    |
  |                  +                            +            +--+    |
  |      +--+        |                            |                    |
  |      |  |        |                            |     +--+           |
  |      +--+        |                            |     |  |           |
  |                  |                            |     +--+           |
  |                  |                            |                    |
  | Area 1           |                            |  Area 5            |
  +------------------+                            +--------------------+
```

AS Number is 100.

Figure 1: Inter-Area Path Computation

This could be represented as <IRO> as:

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


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

AS is optional and it MAY be skipped. PCE should be able to understand both notations.

3.4.2. Inter-AS Path Computation

In inter-AS path computation, where ingress and egress belong to different AS, the domain sequence is represented using an ordered list of AS sub-objects. The domain sequence MAY further include decomposed area information in AREA sub-objects.

3.4.2.1. Example 1

As shown in Figure 2, where AS to be made of a single area, the area subobject MAY be skipped in the domain sequence as AS is enough to uniquely identify the next domain and PCE.

```
                                     +---------------------------------+
                                     |AS 200                           |
                                     |            +------+             |
                                     |            |      |             |
      +------------------------+     |            |      |   +------+  |
      | AS 100                 |     |            +------+   |      |  |
      |               +------+ |     | +------+              |      |  |
      |               |      +-+-----+-+      |              +------+  |
      |               |      | |     | |      |                        |
      |               +------+ |     | +------+                        |
      | +------+               |     |              +------+           |
      | |      |               |     |              |      |           |
      | |      |               |     |              |      |           |
      | +------+               |     |              +------+           |
      |                        |     |                                 |
      |               +------+ |     | +------+                        |
      |               |      +-+-----+-+      |               +------+ |
      |               |      | |     | |      |               |      | |
      |               +------+ |     | +------+               |      | |
      |                        |     |                        +------+ |
      |                        |     |                                 |
      |                        |     |                                 |
      |       +------+         |     |              +------+           |
      |       |      |         |     |              |      |           |
      |       |PCE   |         |     |              |PCE   |           |
      |       +------+         |     |              +------+           |
      |                        |     |                                 |
      +------------------------+     |                                 |
                                     +---------------------------------+
```

Both AS are made of Area 0.

Figure 2: Inter-AS Path Computation

This could be represented as <IRO> as:

```
+---------+  +---------+  +---------+
|IRO      |  |Sub      |  |Sub      |
|Object   |  |Object As|  |Object As|
|Header   |  |100      |  |200      |
|         |  |         |  |         |
|         |  |         |  |         |
+---------+  +---------+  +---------+



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

Area is optional and it MAY be skipped. PCE should be able to understand both notations.

3.4.2.2. Example 2

As shown in Figure 3, where AS 200 is made up of multiple areas and multiple domain-sequence exist, PCE MAY include both AS and AREA subobject to uniquely identify the next domain and PCE.

```
                  |
                  |  +-------------+                +----------------+
                  |  |Area 2       |                |Area 4          |
                  |  |         +--+|                |          +--+  |
                  |  |         |  ||                |          |  |  |
                  |  |  +--+   +--+|                |   +--+   +--+  |
                  |  |  |  |       |                |   |  |         |
                  |  |  *--+       |                |   +--+         |
                  |  | /      +--+ |                |          +--+  |
                  |  |/       |  | |                |          |  |  |
                  |  /        +--+ |                |   +--+   +--+  |
                  | /|  +--+       |+--------------+|   |  |         |
                  |/ |  |  |       ++-+          +-++   +--+         |
   +-------------+/  |  +--+       || |          | ||                |
   |             /|  |             ++-+          +-++                |
   |         +--*||  +-------------+|              |+----------------+
   |         |  |||                 |     +--+     |
   |         +--+||                 |     |  |     |
   |    +--+     ||                 |     +--+     |
   |    |  |     ||                 |              |
   |    +--+     ||                 |              |
   |             ||                 |     +--+     |
   |+--+         ||                 |     |  |     |
   ||  |         ||                 |     +--+     |
   |+--+         ||                 |              |
   |             ||                 |     +--+     |
   |    +--+     ||  +------------+ |     |  |     |+----------------+
   |    |  |     ||  |Area 3      +-++    +--+   +-++ Area 5         |
   |    +--+     ||  |            | ||           | ||                |
   |             ||  |            +-++           +-++                |
   |         +--+||  |       +--+ | |  Area 0      ||   +--+         |
   |         |  |||  |       |  | | +--------------+|   |  |         |
   |         +--*||  |       +--+ |                 |   +--+         |
   |             \|  |            |                 |          +--+  |
   |Area 1       |\  |   +--+     |                 |   +--+   |  |  |
   +-------------+|\ |   |  |     |                 |   |  |   +--+  |
                  | \|   +--+  +--+                 |   +--+         |
                  |  \         |  |                 |                |
                  |  |\        +--+                 |          +--+  |
                  |  | \ +--+     |                 |          |  |  |
                  |  |  \|  |     |                 |          +--+  |
                  |  |   *--+     |                 |                |
                  |  |            |                 |                |
                  |  +------------+                 +----------------+
                  |
                  |
       As 100     |  AS 200
                  |
```

Figure 3: Inter-AS Path Computation

The domain sequence can be carried in IRO as shown below:

```
 +-------+ +-------+ +-------+ +-------+ +-------+ +-------+ +-------+
 |IRO    | |Sub    | |Sub    | |Sub    | |Sub    | |Sub    | |Sub    |
 |Object | |Object | |Object | |Object | |Object | |Object | |Object |
 |Header | |As 100 | |Area 1 | |AS 200 | |Area 3 | |Area 0 | |Area 4 |
 |       | |       | |       | |       | |       | |       | |       |
 +-------+ +-------+ +-------+ +-------+ +-------+ +-------+ +-------+
```

Combination of both AS and Area uniquely identify a domain in the domain sequence.

Note that an Area domain identifier always belongs to the previous AS that appear before it or, if no AS sub-objects are present, it is assumed to be the current AS.

If the area information cannot be provided, PCE MAY forward the path computation request to the next PCE based on AS only. If multiple PCEs of different area domain exist, PCE MAY apply local policy to select the next PCE. Furthermore the domain sequence (list of areas within AS) in the next PCE MAYBE pre-administered or MAYBE discovered via some mechanism (ex. HPCE).

3.4.3. Boundary Node and Inter-AS-Link

A PCC or PCE MAY add additional constraints covering which Boundary Nodes (ABR or ASBR) or Border links (Inter-AS-link) MUST be traversed while defining a domain sequence. In which case the Boundary Node or Link MAY be encoded as a part of the domain-sequence using the existing sub-objects.

Boundary Node (ABR / ASBR) can be encoded using the IPv4 or IPv6 prefix sub-objects. The Inter-AS link can be encoded using the IPv4 or IPv6 prefix or unnumbered interface sub-objects.

For Figure 1, an ABR to be traversed can be specified as:

```
+---------+ +---------+ +---------++---------+ +---------+
|IRO      | |Sub      | |Sub      ||Sub      | |Sub      |
|Object   | |Object   | |Object   ||Object   | |Object   |
|Header   | |Area 2   | |IPv4     ||Area 0   | |Area 4   |
|         | |         | |x.x.x.x  ||         | |         |
|         | |         | |         ||         | |         |
+---------+ +---------+ +---------++---------+ +---------+
```

For Figure 2, an inter-AS-link to be traversed can be specified as:

```
     +---------+  +---------+ +---------+ +---------+ +---------+
     |IRO      |  |Sub      | |Sub      | |Sub      | |Sub      |
     |Object   |  |Object As| |Object   | |Object   | |Object As|
     |Header   |  |100      | |IPv4     | |IPv4     | |200      |
     |         |  |         | |x.x.x.x  | |x.x.x.x  | |         |
     |         |  |         | |         | |         | |         |
     +---------+  +---------+ +---------+ +---------+ +---------+
```

3.4.4. PCE serving multiple domains

A single PCE MAYBE responsible for multiple domains; for example PCE function deployed on an ABR. Domain sequence should have no impact on this. PCE which can support 2 adjacent domains can internally handle this situation without any impact on the neighboring domains.

3.4.5. P2MP

In case of inter-domain P2MP path computation, (Refer [PCE-P2MP-PROCEDURES]) the path domain tree is nothing but a series of Domain Sequences, as shown in the below figure:

```
D1-D3-D6, D1-D3-D5 and D1-D2-D4.
            D1
           /  \
         D2   D3
        /    /  \
      D4   D5   D6
```

All rules of processing as applied to P2P can be applied to P2MP as well.

In case of P2MP, different destinations MAY have different domain sequence within the domain tree, it requires domain-sequence to be attached per destination. (Refer [PCE-P2MP-PER-DEST])

3.4.6. HPCE

As per [PCE-HIERARCHY-FWK], consider a case as shown in Figure 4 consisting of multiple child PCEs and a parent PCE.

```
                                +--------+
                                | Parent |
                                | PCE    |
                                +--------+

 +-------------------+                            +-------------------+
 |           +--+    |                            |     +--+          |
 |  +--+     |  |    |                            |     |  |          |
 |  |  |     +--+    |                            |     +--+   +--+   |
 |  +--*             +                            +            |  |   |
 |                   |                            |            +--+   |
 |         *--+      +                            +                   |
 |         |  |      |                            |     +--+          |
 |         +--+      |                            |     |  |          |
 |                   |+--------------------------+|     +--+          |
 |                  ++++                       +-++                   |
 |                  ||||         +--+          | ||                   |
 | Area 2           ++++         |  |          +-++  Area 4           |
 +-------------------+|          +--+            |+-------------------+
                      |                +--+      |
                      |    +--+        |  |      |
                      |    |  |        +--+      |
                      |    +--+                  |
                      |                          |
                      |           +--+           |
                      |           |  |           |
                      |           +--+           |
  +------------------+|                          |+--------------------+
  |                  ++-+                      +-++                    |
  |                  || |                      | ||                    |
  |                  ++-+    Area 0            +-++                    |
  |                  |+--------------------------+|     +--+           |
  |          +--+    |                            |     |  |           |
  |          |  |    |                            |     +--+           |
  | +--+     +--+    |                            |                    |
  | |  |             +                            +            +--+    |
  | +--+             |                            |            |  |    |
  |                  +                            +            +--+    |
  |       +--+       |                            |                    |
  |       |  |       |                            |     +--+           |
  |       +--+       |                            |     |  |           |
  |                  |                            |     +--+           |
  | Area 1           |                            |  Area 5            |
  +------------------+                            +--------------------+
```

Figure 4: Hierarchical PCE

In HPCE implementation the initiator PCE - PCE(1) can request the

parent PCE to determine the domain sequence and return in the path computation reply message (PCRep), using the ERO Object. The ERO can contain an ordered sequence of sub-object such as AS and Area (OSPF/ISIS). In this case, the PCRep would carry the domain sequence result as:

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

```
+---------+ +---------+ +---------+ +---------+ +---------+
|ERO      | |Sub      | |Sub      | |Sub      | |Sub      |
|Object   | |Object As| |Object   | |Object   | |Object   |
|Header   | |100      | |Area 2   | |Area 0   | |Area 4   |
|         | |         | |         | |         | |         |
|         | |         | |         | |         | |         |
+---------+ +---------+ +---------+ +---------+ +---------+
```

Note that, in the case of ERO objects, no new PCEP object type is required since the ordering constraint is assumed.

3.4.7. Relationship to PCE Sequence

[RFC5886] and [PCE-P2MP-PROCEDURES] along with Domain Sequence introduces the concept of PCE-Sequence, where a sequence of PCEs, based on the domain sequence, should be decided and attached in the PCReq at the very beginning of path computation.

An alternative would be to use domain sequences, note that PCE-Sequence can be used along with domain-sequence in which case PCE-Sequence SHOULD have higher precedence in selecting the next PCE in the inter-domain path computation procedures. Note that Domain-Sequence IRO constraints should still be checked as per the rules of processing IRO.

4. IANA Considerations

4.1. PCEP Objects

The "PCEP Parameters" registry contains a subregistry "PCEP Objects". IANA is requested to make the following allocations from this

registry.

Object Class	Name	Reference
10	IRO	[RFC5440]
	Object-Type (TBA): Domain Sequence	[This I.D.]

4.2. New Sub-Objects

The "PCEP Parameters" registry contains a subregistry "PCEP Objects" with an entry for the Include Route Object (IRO) and Exclude Route Object (XRO). IANA is requested to add further subobjects as follows:

Subobject	Type	Reference
TBA	4 octet AS number	[This I.D.]
TBA	OSPF Area ID	[This I.D.]
TBA	IS-IS Area ID	[This I.D.]

4.3. Error Object Field Values

The "PCEP Parameters" registry contains a subregistry "Error Types and Values". IANA is requested to make the following allocations from this subregistry

ERROR Type	Meaning	Reference
TBA	"Unrecognized subobject" Error-Value: type code	[This I.D.]

5. Security Considerations

This document specifies a standard representation of domain sequence, which MAYBE used in inter-domain PCE scenarios as explained in other RFC and drafts. The new sub-objects and domain sequence mechanisms defined in this document 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. Therefore, the security techniques described in [RFC5440] are considered more important.

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.

6. Manageability Considerations

6.1. Control of Function and Policy

Several local policy decisions should be made at the PCE. Firstly, the exact behavior with regard to desired inclusion and exclusion of domains must be available for examination by an operator and may be configurable. Second, the behavior on receipt of an unrecognized sub-objects with the L or X-bit set should be configurable and must be available for inspection. The inspection and control of these local policy choices may be part of the PCEP MIB module.

6.2. Information and Data Models

A MIB module for management of the PCEP is being specified in a separate document [PCEP-MIB]. That MIB module allows examination of individual PCEP messages, in particular requests, responses and errors. The MIB module MUST be extended to include the ability to view the domain-sequence extensions defined in this document.

6.3. 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].

6.4. Verify Correct Operations

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

6.5. Requirements On Other Protocols

The Sub-objects defined in this document SHOULD be supported by RSVP especially for per-domain path computation [RFC5152] where the domains need to encoded in the RSVP messages.

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

6.6. Impact On Network Operations

Mechanisms defined in this document do not have any impact on network operations in addition to those already listed in [RFC5440].

7. Acknowledgments

We would like to thank Adrian Farrel, Pradeep Shastry, Suresh Babu, Quintin Zhao, Fatai Zhang, Daniel King, Oscar Gonzalez, Chen Huaimo, Venugopal Reddy, Reeja Paul and Sandeep Boina for their useful comments and suggestions.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[ISO 10589] ISO, "Intermediate system to Intermediate system routeing information exchange protocol for use in conjunction with the Protocol for providing the Connectionless-mode Network Service (ISO 8473)", ISO/IEC 10589:2002.

8.2. Informative References

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.

[RFC4726] Farrel, A., Vasseur, J., and A. Ayyangar, "A

Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering", RFC 4726, November 2006.

[RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, May 2007.

[RFC4874] Lee, CY., Farrel, A., and S. De Cnodder, "Exclude Routes - Extension to Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)", RFC 4874, April 2007.

[RFC4893] Vohra, Q. and E. Chen, "BGP Support for Four-octet AS Number Space", RFC 4893, May 2007.

[RFC5152] Vasseur, JP., Ayyangar, A., and R. Zhang, "A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)", RFC 5152, February 2008.

[RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.

[RFC5441] Vasseur, JP., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", RFC 5441, April 2009.

[RFC5520] Bradford, R., Vasseur, JP., and A. Farrel, "Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Path-Key-Based Mechanism", RFC 5520, April 2009.

[RFC5521] Oki, E., Takeda, T., and A. Farrel, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Route Exclusions", RFC 5521, April 2009.

[RFC5886] Vasseur, JP., Le Roux, JL., and Y. Ikejiri, "A Set of Monitoring Tools for Path Computation Element (PCE)-Based Architecture", RFC 5886, June 2010.

[PCE-P2MP-PROCEDURES] Zhao, Q., Dhody, D., Ali, Z., Saad,, T., Sivabalan,, S., and R. Casellas, "PCE-based Computation Procedure To Compute Shortest Constrained P2MP Inter-domain Traffic Engineering Label Switched Paths (draft-ietf-pce-pcep-inter-domain-p2mp-procedures-02)", February 2012.

[PCE-HIERARCHY-FWK] King, D. and A. Farrel, "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS. (draft-ietf-pce-hierarchy-fwk-04)", June 2012.

[PCEP-MIB] Kiran Koushik, A S., Stephan, E., Zhao, Q., and D. King, "PCE communication protocol(PCEP) Management Information Base", July 2010.

[PCE-P2MP-PER-DEST] Dhody, D. and U. Palle, "Supporting explicit-path per destination in Path Computation Element Communication Protocol (PCEP) P2MP Path Request Message. (draft-dhody-pce-pcep-p2mp-per-destination-01)", Feb 2012.

Authors' Addresses

Dhruv Dhody
Huawei Technologies India Pvt Ltd
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: dhruv.dhody@huawei.com

Udayasree Palle
Huawei Technologies India Pvt Ltd
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: udayasree.palle@huawei.com

Ramon Casellas
CTTC - Centre Tecnologic de Telecomunicacions de Catalunya
Av. Carl Friedrich Gauss n7
Castelldefels, Barcelona 08860
SPAIN

EMail: ramon.casellas@cttc.es