

PCE Working Group
Internet-Draft
Intended status: Experimental
Expires: July 11, 2014

D. Dhody
U. Palle
Huawei Technologies
R. Casellas
CTTC
January 7, 2014

Standard Representation Of Domain-Sequence
draft-ietf-pce-pcep-domain-sequence-04

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 Systems (AS). 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 to be used by Path Computation Elements (PCEs) to compute inter-domain shortest constrained 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 July 11, 2014.

Copyright Notice

Copyright (c) 2014 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. Requirements Language	4
2. Terminology	4
3. Detail Description	5
3.1. Domains	5
3.2. Domain-Sequence	5
3.3. Standard Representation	6
3.4. Include Route Object (IRO)	7
3.4.1. Subobjects	7
3.4.2. Option (A): New IRO Object Type	9
3.4.2.1. Handling of the Domain-Sequence IRO object	11
3.4.3. Option B: Existing IRO Object Type	12
3.4.4. Comparison	13
3.5. Exclude Route Object (XRO)	14
3.5.1. Subobjects	14
3.5.1.1. Autonomous system	14
3.5.1.2. IGP Area	15
3.6. Explicit Exclusion Route Subobject (EXRS)	16
3.7. Explicit Route Object (ERO)	17
4. Other Considerations	17
4.1. Inter-Area Path Computation	18
4.2. Inter-AS Path Computation	20
4.2.1. Example 1	20
4.2.2. Example 2	22
4.3. Boundary Node and Inter-AS-Link	24
4.4. PCE Serving multiple Domains	24
4.5. P2MP	25
4.6. Hierarchical PCE	25
4.7. Relationship to PCE Sequence	27
4.8. Relationship to RSVP-TE	27
5. IANA Considerations	28

5.1. PCEP Objects	28
5.2. New Subobjects	28
5.3. Error Object Field Values	28
6. Security Considerations	29
7. Manageability Considerations	29
7.1. Control of Function and Policy	29
7.2. Information and Data Models	29
7.3. Liveness Detection and Monitoring	29
7.4. Verify Correct Operations	29
7.5. Requirements On Other Protocols	30
7.6. Impact On Network Operations	30
8. Acknowledgments	30
9. References	30
9.1. Normative References	30
9.2. Informative References	30

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. Also for inter-domain point-to-multi-point (P2MP) tree computation, [PCE-P2MP-PROCEDURES] assumes the domain-tree is known in priori.

The list of domains (domain-sequence) 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 determines the next PCE to forward the request based on the domain-sequence. In a multi-domain path computation, a PCC MAY indicate the sequence of domains to be traversed using the Include Route Object (IRO) defined in [RFC5440].

When the sequence of domains is not known in advance, the Hierarchical PCE (H-PCE) [RFC6805] architecture and mechanisms can be used to determine the end-to-end Domain-Sequence.

This document defines a standard way to represent and encode a Domain-Sequence in various deployment scenarios including P2P, P2MP and H-PCE.

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.

[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 domain is defined in [RFC3209], this document specifies new subobjects to include or exclude domains such as an IGP area or an Autonomous Systems (4-Byte as per [RFC4893]).

Further, the domain identifier may simply act as delimiter to specify where the domain boundary starts and ends.

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

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: 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) areas and Autonomous Systems (ASs).

Domain-Sequence: 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

[RFC4726] and [RFC4655] define domain 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 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 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 path computation request to PCE(s). A domain-sequence can also be the result of a path computation. For example, in the case of H-PCE

[RFC6805] Parent PCE MAY send the Domain-Sequence as a result in a path computation reply.

In this context, the ordered nature of a domain-sequence is considered to be 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.

A domain-sequence enables 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 [RFC6805];
3. by a PCC (or PCE) to constraint the domains used in a H-PCE path computation, explicitly specifying which domains to be expanded;
4. by a PCE in per-domain path computation model [RFC5152] to identify the next domain(s);

3.3. Standard Representation

Domain-Sequence MAY appear in PCEP Messages, notably 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 Exclusion Route Subobject (EXRS): As per [RFC5521], used to specify exclusion of certain abstract nodes between a specific pair of nodes. EXRS are a subobject inside the IRO. These subobjects are used to specify the domains that must be excluded between two abstract nodes.
- o Explicit Route Object (ERO): As per [RFC5440], used to specify a computed path in the network. For example, in the case of H-PCE [RFC6805] Parent PCE MAY 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 MUST 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.

The following subobject types are used in IRO.

Type	Subobject
1	IPv4 prefix
2	IPv6 prefix
4	Unnumbered Interface ID
32	Autonomous system number (2 Byte)
33	Explicit Exclusion (EXRS)

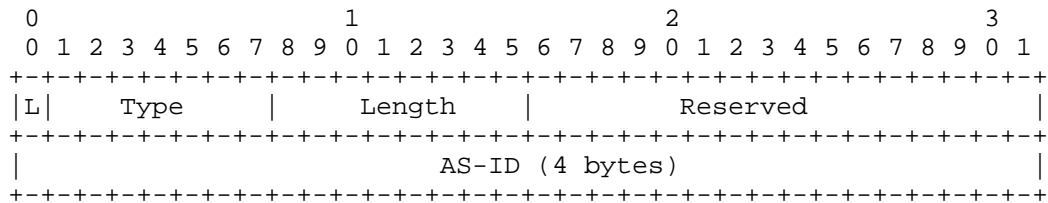
This document extends the above list to support 4-Byte AS numbers and IGP Areas.

Type	Subobject
TBD	Autonomous system number (4 Byte)
TBD	OSPF Area id
TBD	ISIS Area id

- Autonomous system

[RFC3209] already defines 2 byte AS number.

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



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

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

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

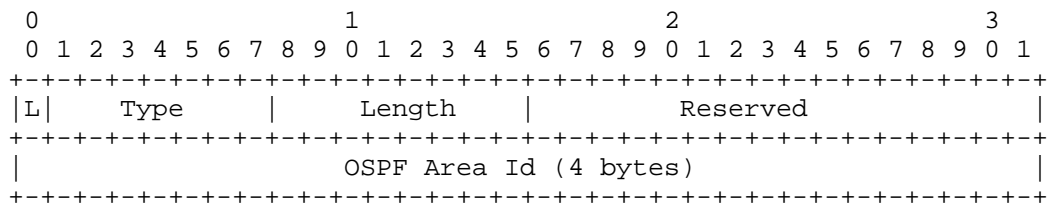
Reserved: Zero at transmission, ignored at receipt.

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

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

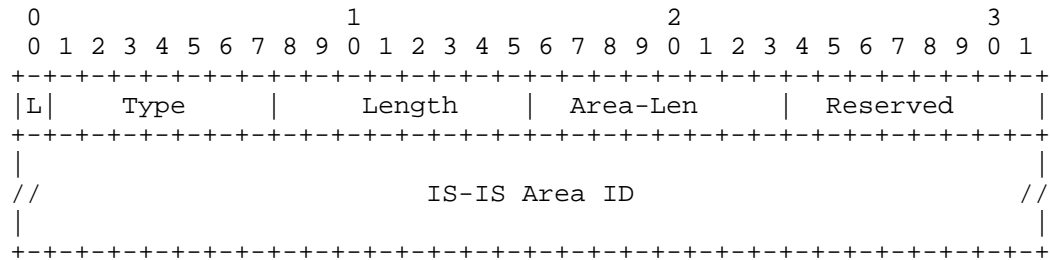
Type: (TBA by IANA) indicating a 4-Byte OSPF Area ID.

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

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

Length: Variable. As per [RFC3209], the total length of the subobject in bytes, including the L, Type and Length fields. 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-byte boundary.

3.4.2. Option (A): New IRO Object Type

[RFC5440] in its description of IRO does not require the subobjects to be in a given particular order. When considering a Domain-Sequence, the domain relative ordering is a basic criterion and, as such, this option suggest a new IRO object type.

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)                               //
  |                                                                    |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

```

Subobjects: The IRO is made of subobjects identical to the ones defined in [RFC3209], [RFC3473], and [RFC3477], where the IRO subobject type is identical to the subobject type defined in the related documents. Some new subobjects related to Domain-Sequence are also added in this document as mentioned in Section 3.4.

[RFC3209] defines subobjects 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 subobjects for AS Number (2 or 4 Byte) and IGP Area is used to specify the domain identifiers in the domain-sequence.

The new IRO Object-Type used to define the domain-sequence would handle the L bit (Loose / Strict) in the subobjects similar to [RFC3209].

Further we have following options:

- * Option (A.1): New IRO Object Type for Domain-Sequence object only. A new IRO Object Type is used to specify the ordered sequence of domains (Domain-Sequence) only. The PCReq message is modified to allow encoding of both types of IRO; one with IRO Type 1 [RFC5440] used to specify the intra-domain abstract nodes and resources and the second IRO Type with the new subobjects as described in this section 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].
- * Option (A.2): New IRO Object Type for both intra and inter-domain (domain-sequence). A new IRO Object Type is used to include both intra nodes and inter-domains nodes but the sequence of domain is strict. The intra-domain nodes can still be ordered. In case of inter-domain path computation, only the new IRO type is used which contains the specific intra domain network nodes as well as inter-domain abstract nodes or domains. The inter-domain abstract nodes are encoded in the sequence they must be traversed but the intra-

domain elements MAY be an unordered set. There is no need to change the PCReq message format.

3.4.2.1. Handling of the Domain-Sequence IRO object

An IRO object containing Domain-Sequence subobjects 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 of 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 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 PCEP Request message with an IRO, it looks for a Domain-Sequence IRO (new type) to see if a domain-sequence is specified. If the message contains more than one Domain-Sequence IRO (new type), it MUST use the first one in the message and MUST ignore subsequent instances.

If a 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 a 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), 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 subobjects 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 request.

A successful path computation reported in a PCEP response message MUST include an ERO to specify the path that has been computed as specified in [RFC5440] following the sequence of domains.

In a PCEP response message, PCE MAY also supply a Domain-Sequence IRO (new type) with the NO-PATH object indicating that the set of elements of the request's Domain-Sequence IRO prevented the PCE from finding a path.

Subobject types for AS and IGP Area affect the next domain selection as well as finding the PCE serving that domain.

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

- o A single IGP Area: Only the IGP (OSPF or ISIS) Area subobject is used to identify the next domain. (Refer Figure 1)
- o A single AS: Only the AS subobject is used to identify the next domain. (Refer Figure 2)
- o Both an AS and an 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)

Subobject representing Boundary Node or Inter-AS-Link MUST be applied during the final path computation procedure as before.

3.4.3. Option B: Existing IRO Object Type

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

The new subobjects denoting the domain-sequence are 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 the PCE to figure out the optimal order of the domains to be crossed to reach the destination domain. Note that in case of doubt, or when applicable, the PCE can still apply the ordering as specified in the request message. Further PCE may have to crankback and try multiple permutations before figuring out the correct sequence.
- o Loose / Strict (L-Bit): [RFC5440] state that the L bit of the subobjects within an IRO Type 1 [RFC5440] has no meaning. This will be applicable for subobjects denoting domain-sequence as well.
- o Scope: Coexistence of intra-domain nodes, boundary nodes and domain nodes in the same IRO List. It is the job of PCE to figure out the scope and apply the processing rules accordingly. The nodes in the IRO which are recognized by the PCE are handled locally and others are forwarded to next PCE hoping they would handle them, the aggregating PCE (Ingress PCE or Parent) would make sure that all nodes in IRO are handled correctly.

3.4.4. Comparison

	Option (A.1): New IRO Object Type for Domain-Sequence object only	Option (A.2): New IRO Object Type for both intra and inter-domain	Option B: Existing IRO Object Type
Order	Yes	Yes	No
L/X bit	Yes	Yes	No
Msg Format Unchanged	No	Yes	Yes
Seperation of scope	Yes	Yes*	No

* because of the ordered nature, intra-domain nodes would be first in the new IRO type

[Editor's Note: Based on our opinion and the feedback received so far, we feel the option A.2 should be selected.]

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

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

Type	Subobject
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

This document extends the above list to support 4-Byte AS numbers and IGP Areas.

Type	Subobject
TBD	Autonomous system number (4 Byte)
TBD	OSPF Area id
TBD	ISIS Area id

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 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.

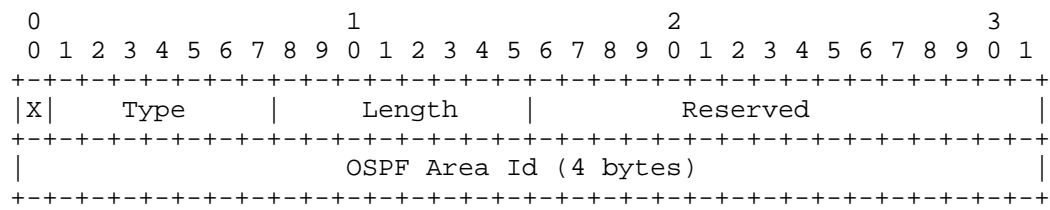
- 0: indicates that the AS specified MUST be excluded from the path computed by the PCE(s).
- 1: indicates that the AS specified SHOULD be avoided from the inter-domain path computed by the PCE(s), but MAY be included subject to PCE policy and the absence of a viable path that meets the other constraints.

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:



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

- 0: indicates that the OSFF Area specified MUST be excluded from the path computed by the PCE(s).
- 1: indicates that the OSFF 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																			
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																				//																			
+-----+										+-----+										+-----+										+-----+																			

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.

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].

3.6. Explicit Exclusion Route Subobject (EXRS)

Explicit Exclusion Route Subobject (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 byte 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].

3.7. 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].

Type	Subobject
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

This document extends the above list to support 4-Byte AS numbers and IGP Areas.

Type	Subobject
TBD	Autonomous system number (4 Byte)
TBD	OSPF Area id
TBD	ISIS Area id

The new subobjects to support 4 byte 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 [RFC6805], a Parent PCE MAY be requested to find the domain-sequence. Refer example in Section 4.6.

The format of the new ERO subobjects is similar to new IRO subobjects, refer Section 3.4.

4. Other Considerations

The examples in this section are for illustration purposes only; to show how the new subobjects may be encoded.

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 MAY be represented using a ordered list of Area subobjects. The AS number MAY be skipped, as area information is enough to select the next PCE.

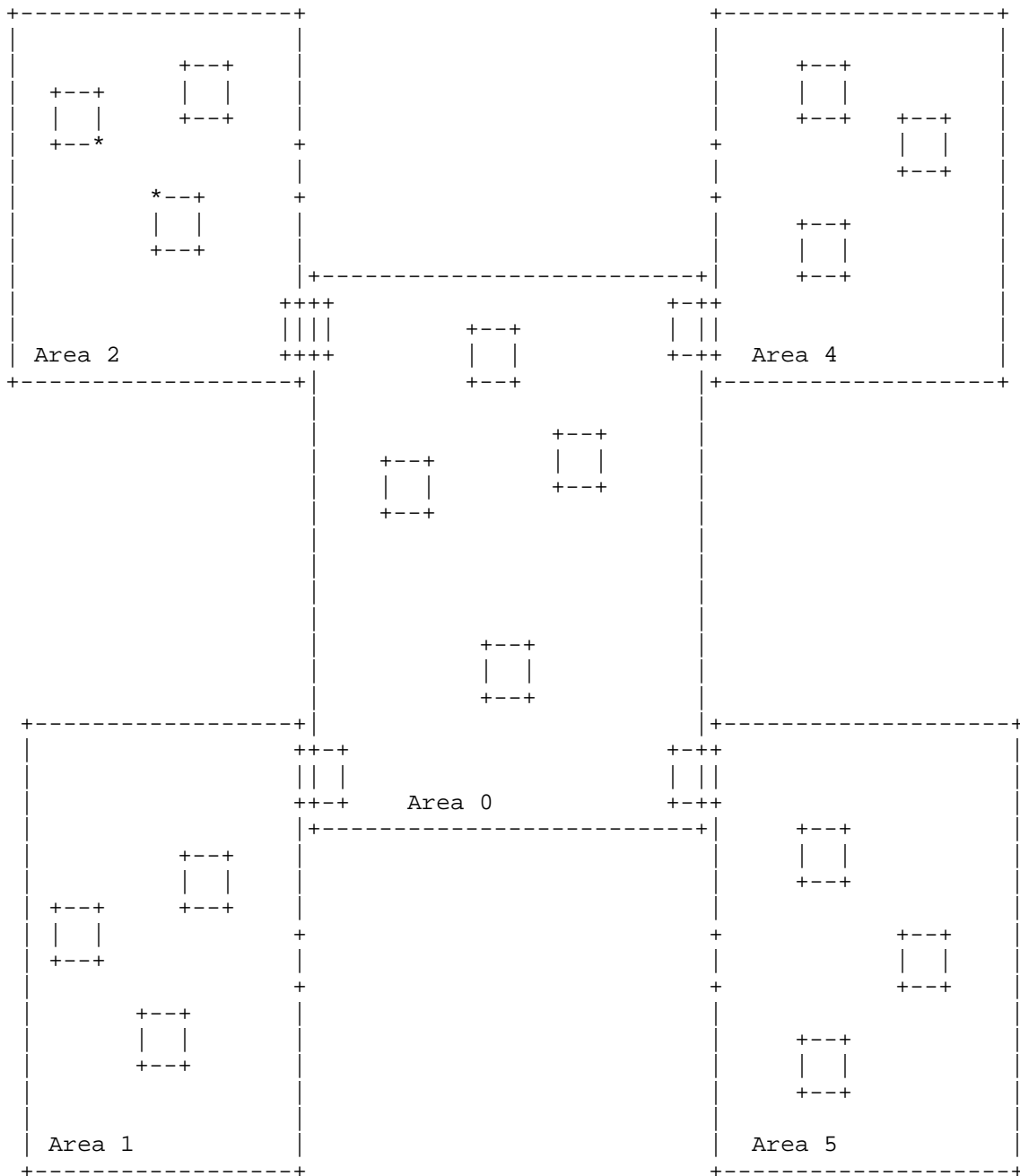


Figure 1: Inter-Area Path Computation

AS Number is 100.

This could be represented in the <IRO> as:

IRO Object Header	Sub Object Area 2	Sub Object Area 0	Sub Object Area 4
-------------------------	-------------------------	-------------------------	-------------------------

IRO Object Header	Sub Object AS 100	Sub Object Area 2	Sub Object Area 0	Sub Object Area 4
-------------------------	-------------------------	-------------------------	-------------------------	-------------------------

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

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 subobjects. The Domain-Sequence MAY further include decomposed area information in Area subobjects.

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.

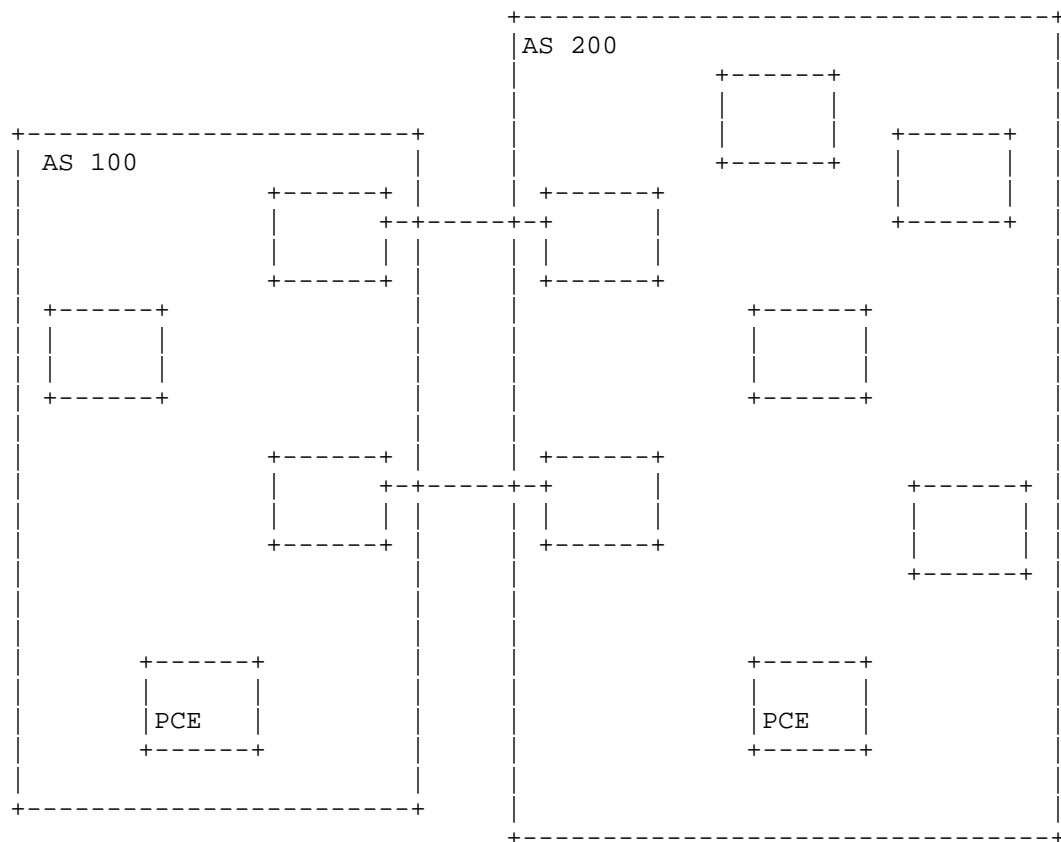
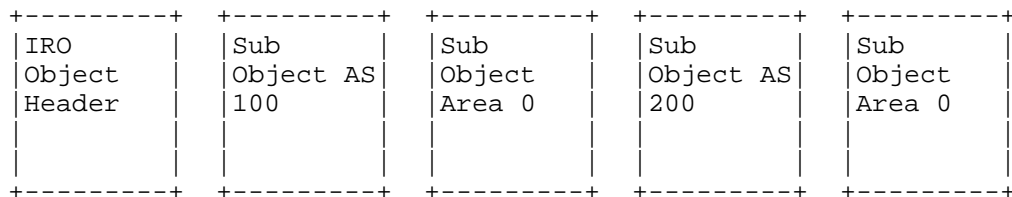
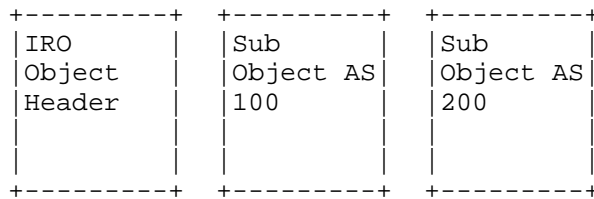


Figure 2: Inter-AS Path Computation

Both AS are made of Area 0.

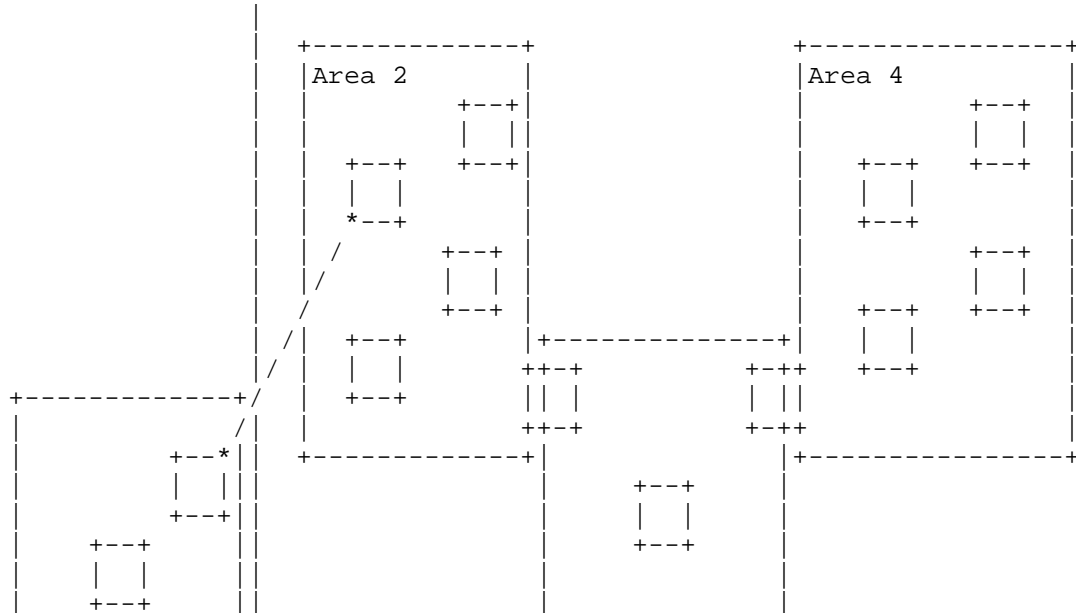
This could be represented in the <IRO> as:



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

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.



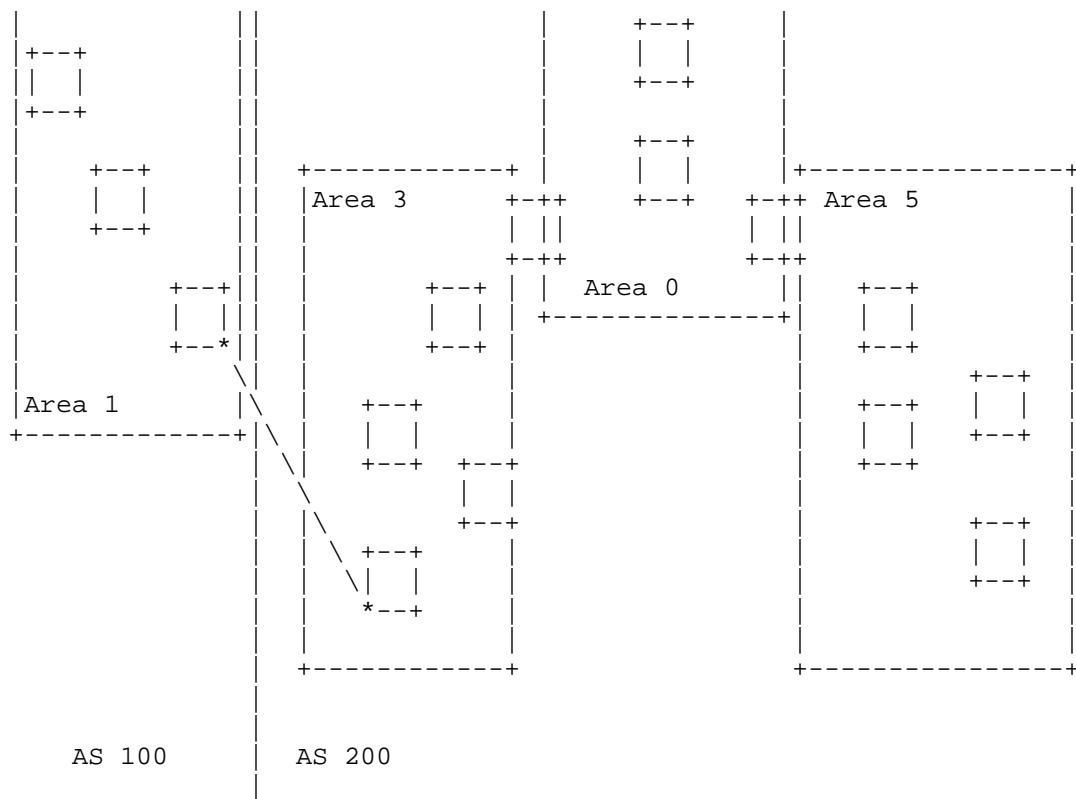


Figure 3: Inter-AS Path Computation

The Domain-Sequence can be carried in the 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
+	+	+	+	+	+	+

The combination of both an AS and an Area uniquely identify a domain in the Domain-Sequence.

Note that an Area domain identifier always belongs to the previous AS that appears before it or, if no AS subobjects 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 alone. If multiple PCEs are responsible, PCE MAY apply local policy to select the next PCE.

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 subobjects.

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 to be traversed can be specified as:

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

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

+-----+	+-----+	+-----+	+-----+	+-----+
IRO Object Header	Sub Object AS 100	Sub Object IPv4 x.x.x.x	Sub Object IPv4 x.x.x.x	Sub Object AS 200
+-----+	+-----+	+-----+	+-----+	+-----+

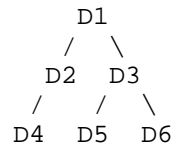
4.4. PCE Serving multiple Domains

A single PCE MAY be responsible for multiple domains; for example PCE function deployed on an ABR. A PCE which can support 2 adjacent domains can internally handle this situation without any impact on the neighboring domains.

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.



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])

4.6. Hierarchical PCE

As per [RFC6805], consider a case as shown in Figure 4 consisting of multiple child PCEs and a parent PCE.

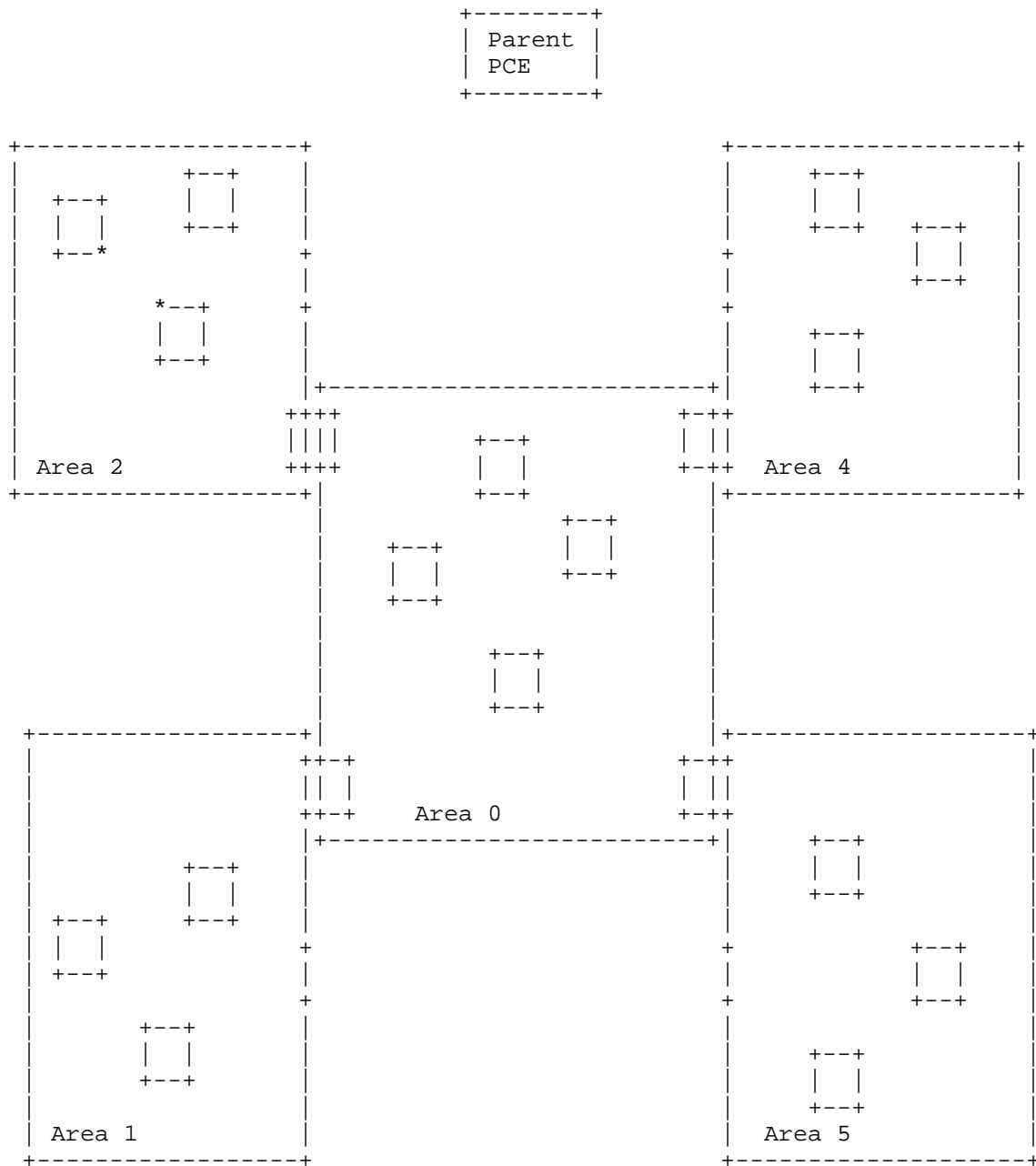


Figure 4: Hierarchical PCE

In H-PCE, the Ingress PCE PCE(1) can request the parent PCE to determine the Domain-Sequence and return it in the PCEP response, using the ERO Object. The ERO can contain an ordered sequence of subobjects such as AS and Area (OSPF/ISIS) subobjects. In this case, the Domain-Sequence appear as:

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

+-----+	+-----+	+-----+	+-----+	+-----+
ERO Object Header	Sub Object AS 100	Sub Object Area 2	Sub Object Area 0	Sub Object Area 4
+-----+	+-----+	+-----+	+-----+	+-----+

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

4.7. 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 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.

4.8. 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 MAY

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.

5. IANA Considerations

5.1. 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.]	

5.2. New Subobjects

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

7 ERO
10 IRO
17 XRO

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

5.3. 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.]

6. Security Considerations

This document specifies a standard representation of Domain-Sequence and new subobjects, which MAY be used in inter-domain PCE scenarios as explained in other RFC and drafts. The new subobjects 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.

7. Manageability Considerations

7.1. 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 subobjects 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.

7.2. 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.

7.3. 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].

7.4. Verify Correct Operations

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

7.5. 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, and signaling message may use 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].

7.6. 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].

8. Acknowledgments

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

9. References

9.1. Normative References

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

9.2. 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.
- [RFC6805] 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", RFC 6805, November 2012.

[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)", July 2013.

[PCEP-MIB]

Koushik, A., Emile, S., Zhao, Q., King, D., and J.
Hardwick, "PCE communication protocol(PCEP) Management
Information Base", July 2013.

[PCE-P2MP-PER-DEST]

Dhody, D., Palle, U., and V. Kondreddy, "Supporting
explicit inclusion or exclusion of abstract nodes for a
subset of P2MP destinations in Path Computation Element
Communication Protocol (PCEP). (draft-dhody-pce-pcep-p2mp-
per-destination)", October 2013.

[DOMAIN-SUBOBJ]

Dhody, D., Palle, U., Kondreddy, V., and R. Casellas,
"Domain Subobjects for Resource ReserVation Protocol -
Traffic Engineering (RSVP-TE). (draft-dhody-ccamp-rsvp-te-
domain-subobjects)", January 2014.

[ISO10589]

ISO, "Intermediate system to Intermediate system routing
information exchange protocol for use in conjunction with
the Protocol for providing the Connectionless-mode Network
Service (ISO 8473)", ISO/IEC 10589:2002, 1992.

Authors' Addresses

Dhruv Dhody
Huawei Technologies
Leela Palace
Bangalore, Karnataka 560008
INDIA

Email: dhruv.ietf@gmail.com

Udayasree Palle
Huawei Technologies
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: udayasree.palle@huawei.com

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

EMail: ramon.casellas@cttc.es