Internet Engineering Task Force (IETF)
Request for Comments: 7897
Category: Experimental
ISSN: 2070-1721

D. Dhody
U. Palle
Huawei Technologies
R. Casellas
CTTC

June 2016

Domain Subobjects for the Path Computation Element Communication Protocol (PCEP)

Abstract

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

Status of This Memo

This document is not an Internet Standards Track specification; it is published for examination, experimental implementation, and evaluation.

This document defines an Experimental Protocol for the Internet community. This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7897.

Table of Contents

1. Introduction

A Path Computation Element (PCE) may be used to compute end-to-end paths across multi-domain environments using a per-domain path computation technique [RFC5152]. The Backward-Recursive PCE-Based Computation (BRPC) mechanism [RFC5441] also defines a PCE-based path computation procedure to compute an inter-domain constrained path for (G)MPLS TE LSPs. However, both per-domain and BRPC techniques assume that the sequence of domains to be crossed from source to destination is known and is either fixed by the network operator or obtained by other means. Also, for inter-domain point-to-multipoint (P2MP) tree computation, it is assumed per [RFC7334] that the domain tree is known a priori.

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

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

[RFC5440] defines the Include Route Object (IRO) and the Explicit Route Object (ERO). [RFC5521] defines the Exclude Route Object (XRO) and the Explicit Exclusion Route subobject (EXRS). The use of an Autonomous System (albeit with a 2-byte AS number) as an abstract node representing a domain is defined in [RFC3209]. In the current document, we specify new subobjects to include or exclude domains including an IGP area or an AS (4 bytes as per [RFC6793]).

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

This is a companion document to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) extensions for the domain identifiers [RFC7898].

1.1. Scope

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

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

This document does not change the procedures for handling existing subobjects in the PCE Communication Protocol (PCEP).

The new subobjects introduced by this document will not be understood by legacy implementations. If a legacy implementation receives one of the subobjects that it does not understand in a PCEP object, the legacy implementation will behave according to the rules for a malformed object as per [RFC5440]. Therefore, it is assumed that this experiment will be conducted only when both the PCE and the Path Computation Client (PCC) form part of the experiment. It is possible that a PCC or PCE can operate with peers, some of which form part of the experiment and some that do not. In this case, since no capabilities exchange is used to identify which nodes can use these extensions, manual configuration should be used to determine which peerings form part of the experiment.

When the results of implementation and deployment are available, this document will be updated and refined, and then it could be moved from Experimental to Standards Track.

1.2. Requirements Language

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

2. Terminology

The following terminology is used in this document.

ABR: Area Border Router. Routers used to connect two IGP areas (Open Shortest Path First (OSPF) or Intermediate System to Intermediate System (IS-IS).

AS: Autonomous System

ASBR: Autonomous System Border Router

BN: Boundary node; can be an ABR or ASBR.

BRPC: Backward-Recursive PCE-Based 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 IGP area and AS.

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: OSPF or 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

XRO: Exclude Route Object

3. Detail Description

3.1. Domains

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

3.2. Domain Sequence

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

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

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

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

A domain sequence can include boundary nodes (ABR or ASBR) or border links (inter-AS links) to be traversed as an additional constraint.

Thus, a domain sequence can be made up of one or more of the following:

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

These are encoded in the new subobjects defined in this document as well as in the existing subobjects that represent a domain sequence.

Consequently, a domain sequence can be used by:

1. a PCE in order to discover or select the next PCE in a collaborative path computation, such as in BRPC [RFC5441];
2. the parent PCE to return the domain sequence when unknown; this can then be an input to the BRPC procedure [RFC6805];
3. a PCC or a PCE to constrain the domains used in inter-domain path computation, explicitly specifying which domains to be expanded or excluded; and
4. a PCE in the per-domain path computation model [RFC5152] to identify the next domain.

3.3. Domain Sequence Representation

A domain sequence appears in PCEP messages, notably in:

- Include Route Object (IRO): As per [RFC5440], IRO can be used to specify a set of network elements to be traversed to reach the destination, which includes subobjects used to specify the domain sequence.
- Exclude Route Object (XRO): As per [RFC5521], XRO can be used to specify certain abstract nodes, to be excluded from the whole path, which include subobjects used to specify the domain sequence.
- Explicit Exclusion Route Subobject (EXRS): As per [RFC5521], EXRS can be used to specify exclusion of certain abstract nodes (including domains) between a specific pair of nodes. EXRS is a subobject inside the IRO.

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

3.4. Include Route Object (IRO)

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

3.4.1. Subobjects

Some subobjects are defined in [RFC3209], [RFC3473], [RFC3477], and [RFC4874], but new subobjects related to domain sequence are needed.

This document extends the support for 4-byte AS numbers and IGP areas.

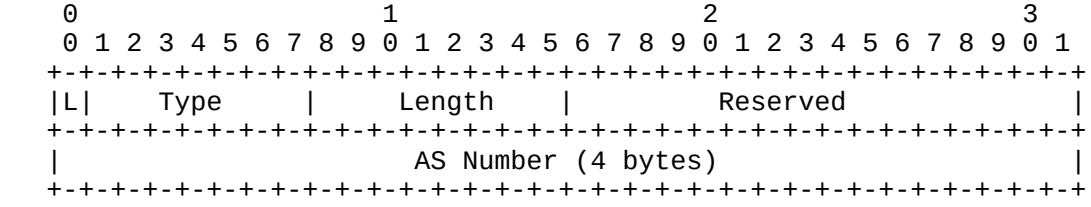
Value	Description
5	4-byte AS number
6	OSPF Area ID
7	IS-IS Area ID

Note: Identical subobjects are carried in RSVP-TE messages as defined in [RFC7898].

3.4.1.1. Autonomous System

[RFC3209] already defines 2-byte AS numbers.

To support 4-byte AS numbers as per [RFC6793], the following subobject is defined:



L: The L bit is an attribute of the subobject as defined in [RFC3209], and its usage in the IRO subobject is defined in [RFC7896].

Type: 5 (indicating a 4-byte AS number).

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

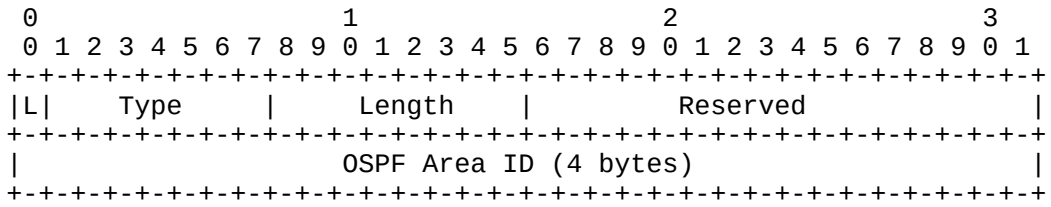
Reserved: Zero at transmission; ignored at receipt.

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

3.4.1.2. IGP Area

Since the length and format of Area ID is different for OSPF and IS-IS, the following two subobjects are defined below:

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], and its usage in the IRO subobject is defined in [RFC7896].

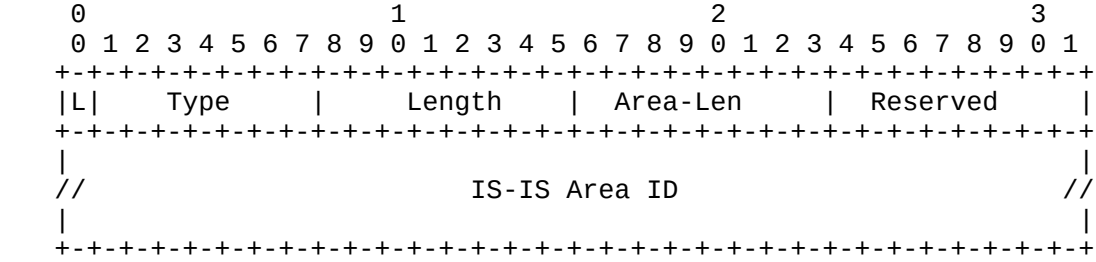
Type: 6 (indicating a 4-byte OSPF Area ID).

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



L: The L bit is an attribute of the subobject as defined in [RFC3209], and its usage in the IRO subobject is defined in [RFC7896].

Type: 7 (indicating the IS-IS Area ID).

Length: Variable. The length MUST be at least 8 and MUST be a multiple of 4.

Area-Len: Variable (length of the actual (non-padded) IS-IS area identifier in octets; valid values are from 1 to 13, inclusive).

Reserved: Zero at transmission; ignored at receipt.

IS-IS Area ID: The variable-length IS-IS area identifier. Padded with trailing zeroes to a 4-byte boundary.

3.4.2. Update in IRO Specification

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

An update to the IRO specification [RFC7896] makes IRO as an ordered list and includes support for the L bit.

The use of IRO for the domain sequence assumes the updated specification is being used for IRO, as per [RFC7896].

3.4.3. IRO for Domain Sequence

The subobject type for IPv4, IPv6, and unnumbered Interface IDs can be used to specify boundary nodes (ABR/ASBR) and inter-AS links. The subobject type for the AS Number (2 or 4 bytes) and the IGP area are used to specify the domain identifiers in the domain sequence.

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

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

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

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

3.4.3.1. PCC Procedures

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

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

3.4.3.2. PCE Procedures

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

The interpretation of the L bit is as per Section 4.3.3.1 of [RFC3209] (as per [RFC7896]).

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

The following processing rules apply for a domain sequence in IRO:

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

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

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

      * Otherwise, it assumes that the subobject belongs to the current area.

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

   Note that it is advised that PCC should use AS and area subobjects while building the domain sequence in IRO and avoid using other mechanisms to change the "current AS" and "current area" as described above.

3.5. Exclude Route Object (XRO)

   XRO [RFC5521] is an optional object used to specify exclusion of certain abstract nodes or resources from the whole path.

3.5.1. Subobjects

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

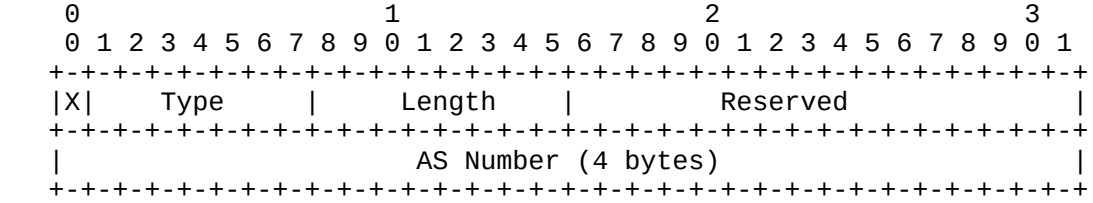
   This document extends the support for 4-byte AS numbers and IGP areas.

Value	Description
5	4-byte AS number
6	OSPF Area ID
7	IS-IS Area ID

   Note: Identical subobjects are carried in RSVP-TE messages as defined in [RFC7898].

3.5.1.1. Autonomous System

The new subobjects to support 4-byte AS numbers and the IGP (OSPF/IS-IS) 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 AS specified MUST be excluded from the path computed by the PCE(s).

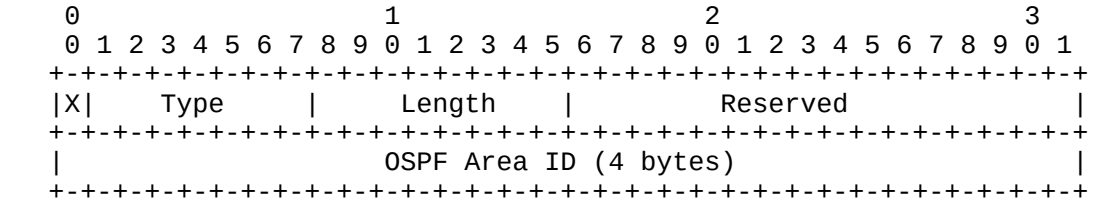
1: indicates that the AS specified SHOULD be avoided from the inter-domain path computed by the PCE(s), but it MAY be included subject to PCE policy and the absence of a viable path that meets the other constraints.

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

3.5.1.2. IGP Area

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

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



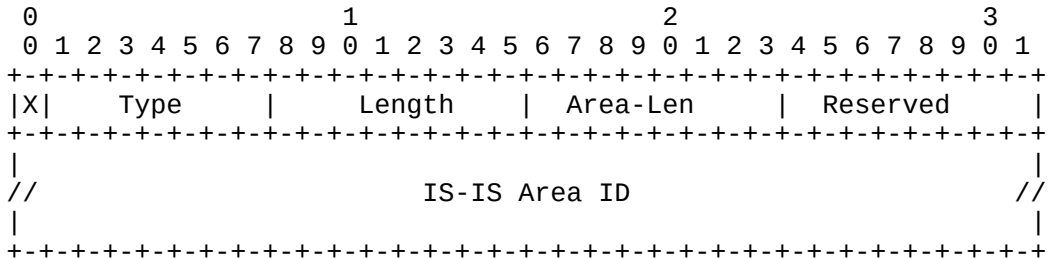
The X-bit indicates whether the exclusion is mandatory or desired.

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

1: indicates that the OSPF area specified SHOULD be avoided from the inter-domain path computed by the PCE(s), but it 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; thus, the length of the subobject is variable. The Area ID is as described in IS-IS by the ISO standard [ISO10589]. The subobject is encoded as follows:



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

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

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

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

All the processing rules are as per [RFC5521].

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

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

3.6. Explicit Exclusion Route Subobject (EXRS)

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

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

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

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

3.7. Explicit Route Object (ERO)

ERO [RFC5440] is used to specify a computed path in the network. PCEP ERO subobject types correspond to RSVP-TE ERO subobject types as defined in [RFC3209], [RFC3473], [RFC3477], [RFC4873], [RFC4874], and [RFC5520]. The subobjects related to the domain sequence are further defined in [RFC7898].

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

In case of H-PCE [RFC6805], a parent PCE can be requested to find the domain sequence. Refer to the example in Section 4.6 of this document. The ERO in reply from the parent PCE can then be used in per-domain path computation or BRPC.

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

4. Examples

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

4.1. Inter-Area Path Computation

In an inter-area path computation where the ingress and the egress nodes belong to different IGP areas within the same AS, the domain sequence could be represented using an ordered list of area subobjects.

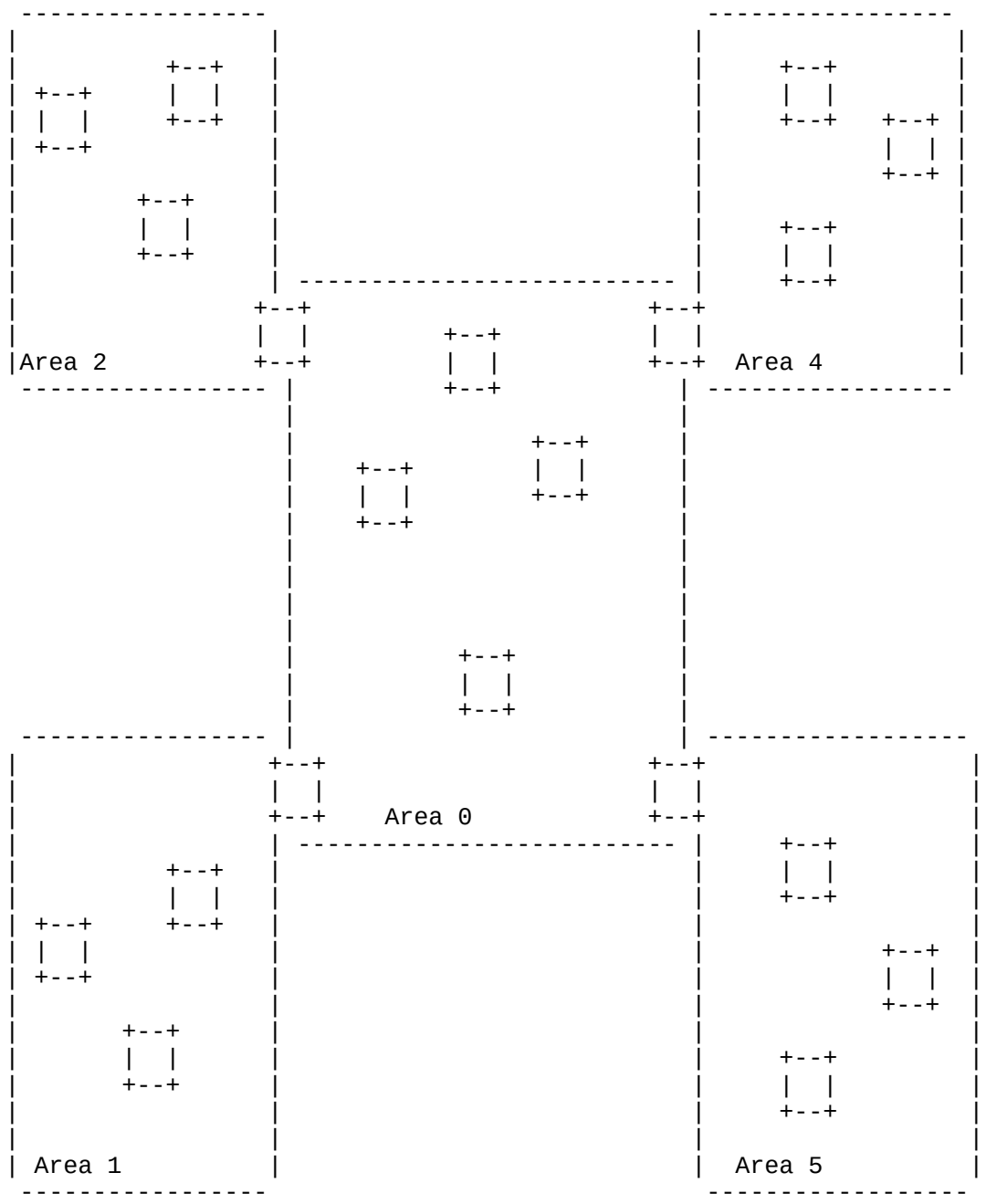
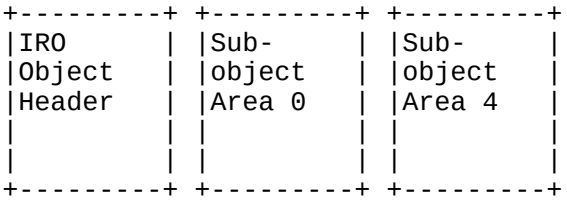


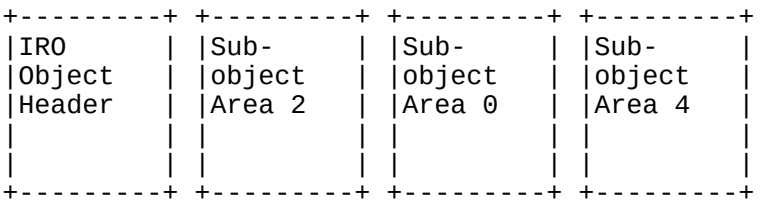
Figure 1: Inter-Area Path Computation

The AS Number is 100.

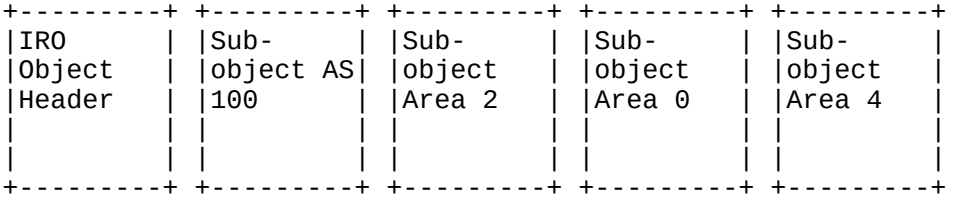
If the ingress is in area 2, the egress is in area 4, and transit is through area 0, here are some possible ways a PCC can encode the IRO:



or



or



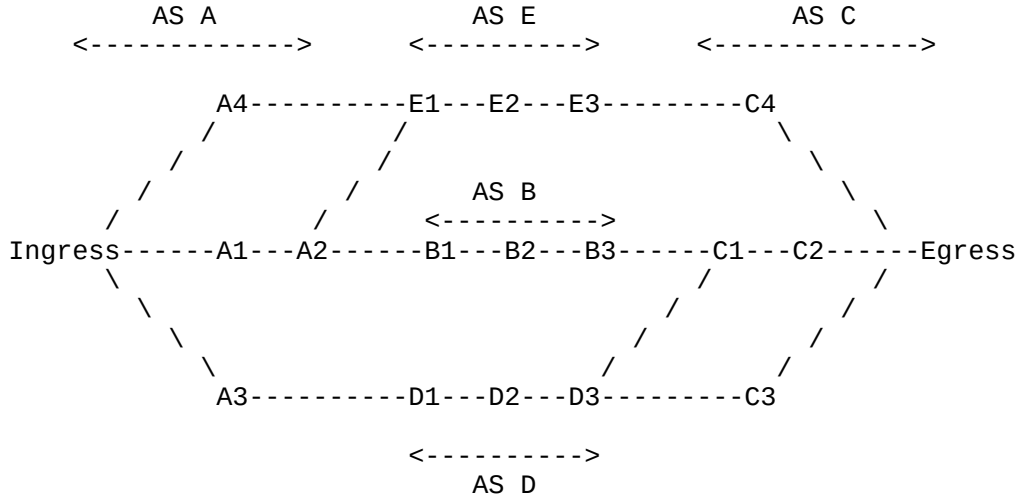
The domain sequence can further include encompassing AS information in the AS subobject.

4.2. Inter-AS Path Computation

In inter-AS path computation, where the ingress and egress belong to different ASes, the domain sequence could be represented using an ordered list of AS subobjects. The domain sequence can further include decomposed area information in the area subobject.

4.2.1. Example 1

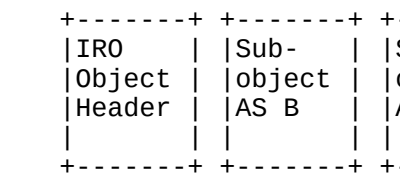
As shown in Figure 2, where AS has a single area, the AS subobject in the domain sequence can uniquely identify the next domain and PCE.



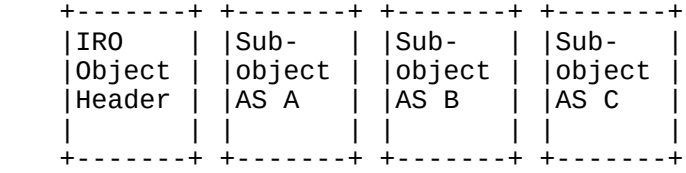
* All ASes have one area (area 0)

Figure 2: Inter-AS Path Computation

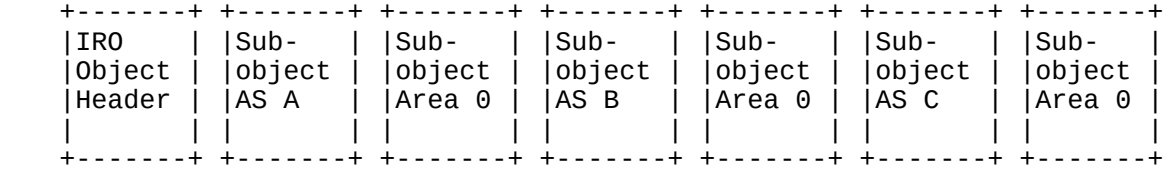
If the ingress is in AS A, the egress is in AS C, and transit is through AS B, here are some possible ways a PCC can encode the IRO:



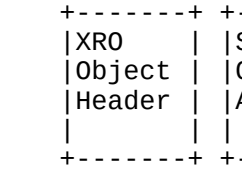
or



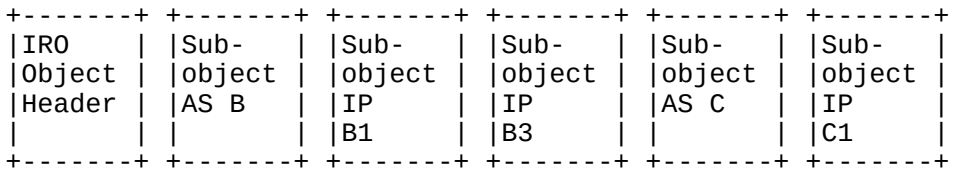
or



Note that to get a domain disjoint path, the ingress could also request the backup path with:

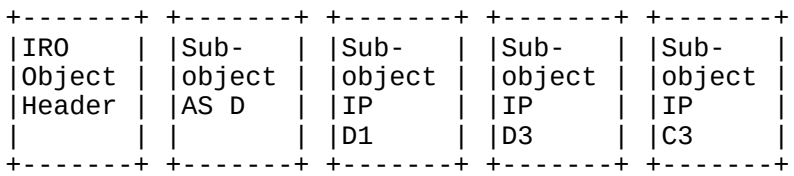


As described in Section 3.4.3, a domain subobject in IRO changes the domain information associated with the next set of subobjects till you encounter a subobject that changes the domain too. Consider the following IRO:



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

Consider another IRO:



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

Further description for the boundary node and inter-AS link can be found in Section 4.3.

4.2.2. Example 2

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

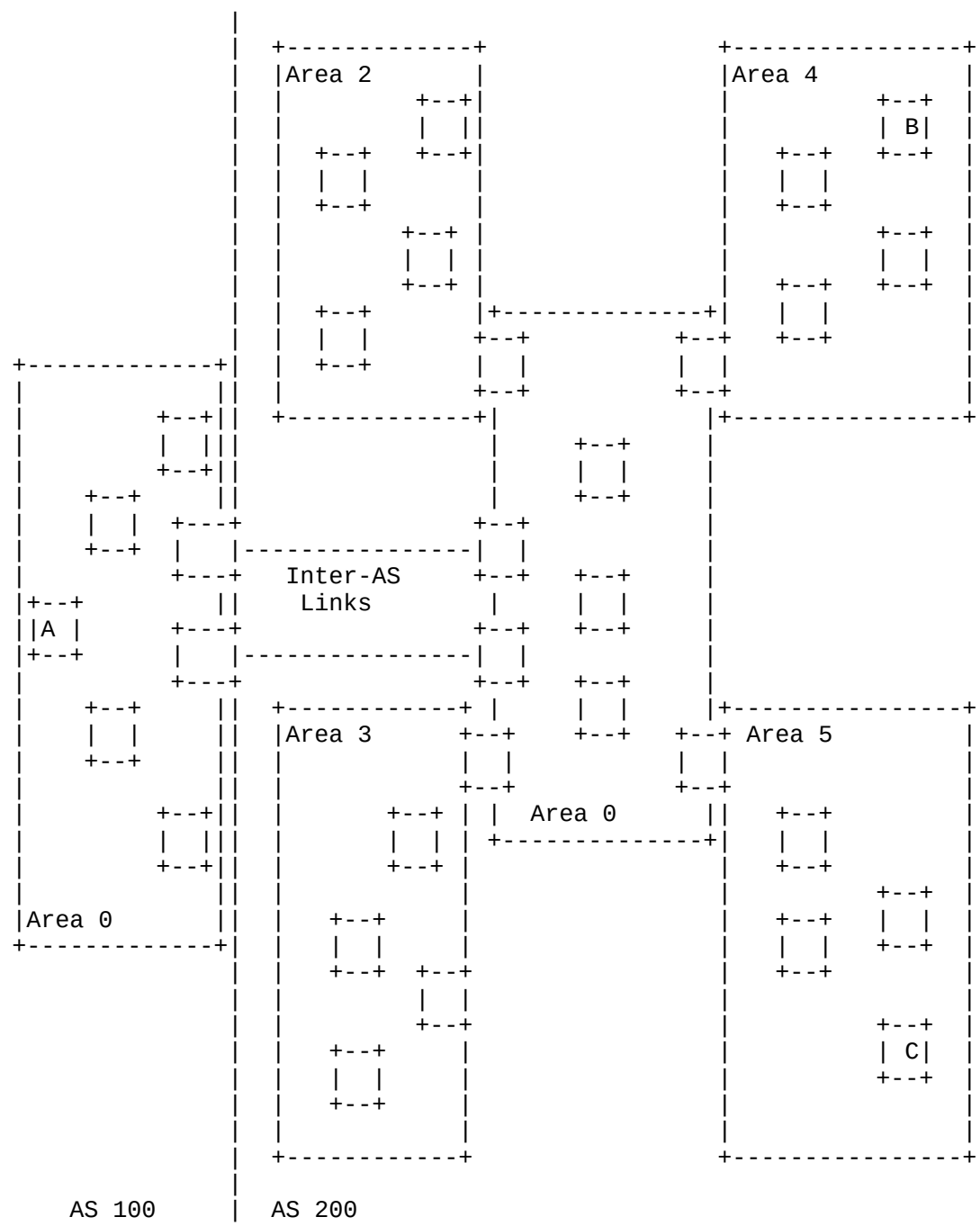
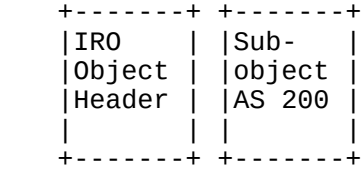
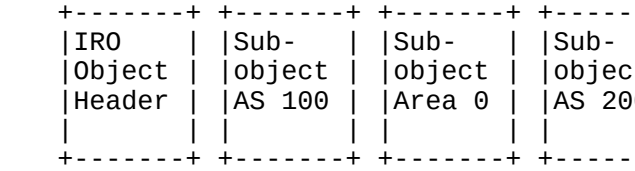


Figure 3: Inter-AS Path Computation

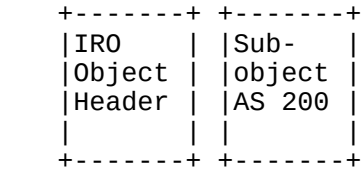
For LSP (A-B), where ingress A is in (AS 100, area 0), egress B is in (AS 200, area 4), and transit is through (AS 200, area 0), here are some possible ways a PCC can encode the IRO:



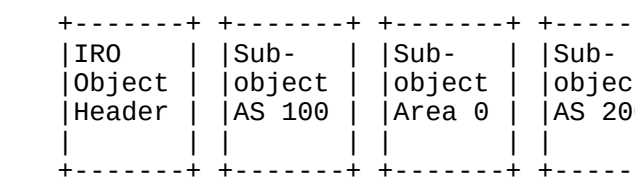
or



For LSP (A-C), where ingress A is in (AS 100, area 0), egress C is in (AS 200, area 5), and transit is through (AS 200, area 0), here are some possible ways a PCC can encode the IRO:



or

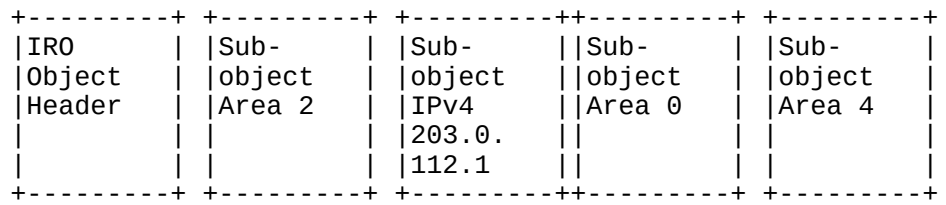


4.3. Boundary Node and Inter-AS Link

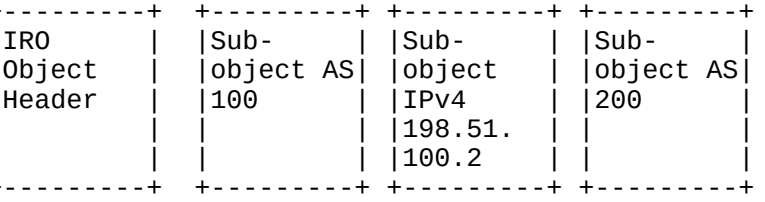
A PCC or PCE can include additional constraints covering which boundary nodes (ABR or ASBR) or border links (inter-AS link) to be traversed while defining a domain sequence. In which case, the boundary node or link can be encoded as a part of the domain sequence.

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

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



For Figure 3, an inter-AS link (say, 198.51.100.1 - 198.51.100.2) to be traversed can be specified as:



4.4. PCE Serving Multiple Domains

A single PCE can be responsible for multiple domains; for example, PCE function deployed on an ABR could be responsible for multiple areas. A PCE that can support adjacent domains can internally handle those domains in the domain sequence without any impact on the other domains in the domain sequence.

4.5. P2MP

[RFC7334] describes an experimental inter-domain P2MP path computation mechanism where the path domain tree is described as a series of domain sequences; an example is shown in the figure below:

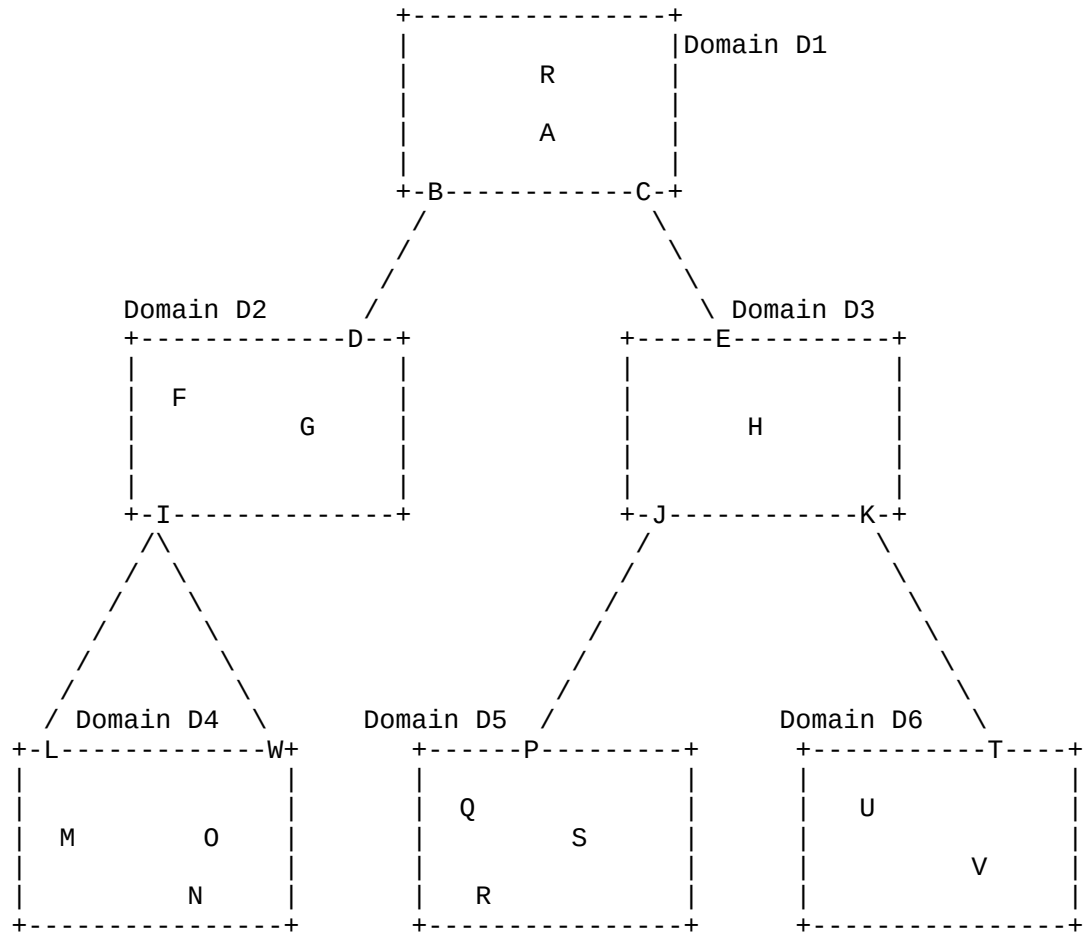


Figure 4: Domain Tree Example

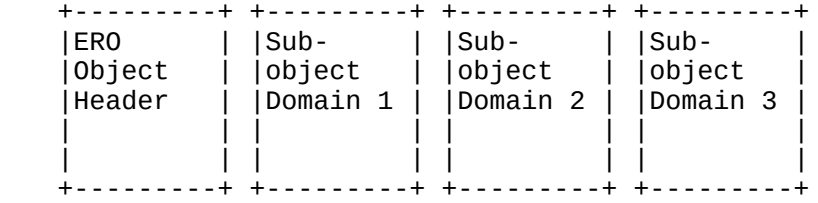
The domain tree can be represented as a series of domain sequences:

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

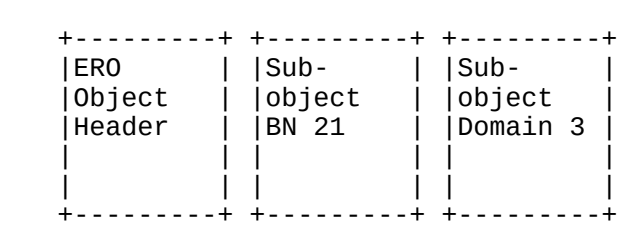
The domain sequence handling described in this document could be applied to the P2MP path domain tree.

4.6. Hierarchical PCE

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



or



5. Other Considerations

5.1. Relationship to PCE Sequence

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

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

5.2. Relationship to RSVP-TE

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

[RFC7898] extends the notion of abstract nodes by adding new subobjects for IGP areas and 4-byte AS numbers. These subobjects can be included in ERO, XRO, or EXRS in RSVP-TE.

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

6. IANA Considerations

6.1. New Subobjects

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

- IRO Subobjects
- XRO Subobjects

IANA has made identical additions to those registries as follows:

Value	Description	Reference
5	4-byte AS number	RFC 7897, [RFC7898]
6	OSPF Area ID	RFC 7897, [RFC7898]
7	IS-IS Area ID	RFC 7897, [RFC7898]

Further, IANA has added a reference to this document to the new RSVP numbers that are registered by [RFC7898], as shown on <http://www.iana.org/assignments/rsvp-parameters>.

7. Security Considerations

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

This document specifies a representation of the domain sequence and new subobjects, which could be used in inter-domain PCE scenarios as explained in [RFC5152], [RFC5441], [RFC6805], [RFC7334], etc. The security considerations set out in each of these mechanisms remain unchanged by the new subobjects and domain sequence representation in this document.

But the new subobjects do allow finer and more specific control of the path computed by a cooperating PCE(s). Such control increases the risk if a PCEP message is intercepted, modified, or spoofed because it allows the attacker to exert control over the path that the PCE will compute or to make the path computation impossible. Consequently, it is important that implementations conform to the relevant security requirements of [RFC5440]. These mechanisms include:

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

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

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

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

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

8. Manageability Considerations

8.1. Control of Function and Policy

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

8.2. Information and Data Models

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

8.3. Liveness Detection and Monitoring

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

8.4. Verify Correct Operations

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

8.5. Requirements on Other Protocols

In case of per-domain path computation [RFC5152], where the full path of an inter-domain TE LSP cannot be determined (or is not determined) at the ingress node, a signaling message can use the domain identifiers. The subobjects defined in this document SHOULD be supported by RSVP-TE. [RFC7898] 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 aside from those already listed in [RFC5440].

8.6. Impact on Network Operations

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

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

9. References

9.1. Normative References

[ISO10589] International Organization for Standardization, "Information technology -- Telecommunications and information exchange between systems -- Intermediate System to Intermediate System intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service (ISO 8473)", ISO/IEC 10589:2002, Second Edition, 2002.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <http://www.rfc-editor.org/info/rfc3209>.

[RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, DOI 10.17487/RFC3473, January 2003, <http://www.rfc-editor.org/info/rfc3473>.

[RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, DOI 10.17487/RFC3477, January 2003, <http://www.rfc-editor.org/info/rfc3477>.

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

[RFC5441] Vasseur, JP., Ed., 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, DOI 10.17487/RFC5441, April 2009, <http://www.rfc-editor.org/info/rfc5441>.

[RFC5521] Oki, E., Takeda, T., and A. Farrel, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Route Exclusions", RFC 5521, DOI 10.17487/RFC5521, April 2009, <http://www.rfc-editor.org/info/rfc5521>.

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

[RFC7896] Dhody, D., "Update to the Include Route Object (IRO) Specification in the Path Computation Element Communication Protocol (PCEP)", RFC 7896, DOI 10.17487/RFC7896, June 2016, <http://www.rfc-editor.org/info/rfc7896>.

[RFC7898] Dhody, D., Palle, U., Kondreddy, V., and R. Casellas, "Domain Subobjects for Resource Reservation Protocol - Traffic Engineering (RSVP-TE)", RFC 7898, DOI 10.17487/RFC7898, June 2016, <http://www.rfc-editor.org/info/rfc7898>.

9.2. Informative References

[PCEPS] Lopez, D., Dios, O., Wu, W., and D. Dhody, "Secure Transport for PCEP", Work in Progress, draft-ietf-pce-pceps-09, November 2015.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <http://www.rfc-editor.org/info/rfc4655>.

[RFC4726] Farrel, A., Vasseur, J., and A. Ayyangar, "A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering", RFC 4726, DOI 10.17487/RFC4726, November 2006, <http://www.rfc-editor.org/info/rfc4726>.

[RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, DOI 10.17487/RFC4873, May 2007, <http://www.rfc-editor.org/info/rfc4873>.

[RFC4874] Lee, CY., Farrel, A., and S. De Cnodder, "Exclude Routes - Extension to Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)", RFC 4874, DOI 10.17487/RFC4874, April 2007, <http://www.rfc-editor.org/info/rfc4874>.

[RFC5152] Vasseur, JP., Ed., Ayyangar, A., Ed., and R. Zhang, "A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)", RFC 5152, DOI 10.17487/RFC5152, February 2008, <http://www.rfc-editor.org/info/rfc5152>.

[RFC5520] Bradford, R., Ed., Vasseur, JP., and A. Farrel, "Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Path-Key-Based Mechanism", RFC 5520, DOI 10.17487/RFC5520, April 2009, <http://www.rfc-editor.org/info/rfc5520>.

[RFC5886] Vasseur, JP., Ed., Le Roux, JL., and Y. Ikejiri, "A Set of Monitoring Tools for Path Computation Element (PCE)-Based Architecture", RFC 5886, DOI 10.17487/RFC5886, June 2010, <http://www.rfc-editor.org/info/rfc5886>.

[RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <http://www.rfc-editor.org/info/rfc5925>.

[RFC6793] Vohra, Q. and E. Chen, "BGP Support for Four-Octet Autonomous System (AS) Number Space", RFC 6793, DOI 10.17487/RFC6793, December 2012, <http://www.rfc-editor.org/info/rfc6793>.

[RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", RFC 6952, DOI 10.17487/RFC6952, May 2013, <http://www.rfc-editor.org/info/rfc6952>.

[RFC7334] Zhao, Q., Dhody, D., King, D., Ali, Z., and R. Casellas, "PCE-Based Computation Procedure to Compute Shortest Constrained Point-to-Multipoint (P2MP) Inter-Domain Traffic Engineering Label Switched Paths", RFC 7334, DOI 10.17487/RFC7334, August 2014, <http://www.rfc-editor.org/info/rfc7334>.

[RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", RFC 7420, DOI 10.17487/RFC7420, December 2014, <http://www.rfc-editor.org/info/rfc7420>.

Acknowledgments

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

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

Thanks to Jonathan Hardwick for shepherding this document.

Thanks to Deborah Brungard for being the responsible AD.

Thanks to Amanda Baber for the IANA review.

Thanks to Joel Halpern for the Gen-ART review.

Thanks to Klaas Wierenga for the SecDir review.

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

Authors' Addresses

Dhruv Dhody
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

Email: dhruv.ietf@gmail.com

Udayasree Palle
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

Email: udayasree.palle@huawei.com

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

Email: ramon.casellas@cttc.es