ï»¿

        Applicability of the Path Computation Element to Inter-area and Inter-AS
                    MPLS and GMPLS Traffic Engineering

Abstract

   The Path Computation Element (PCE) may be used for computing services
   that traverse multi-area and multi-Autonomous System (multi-AS)
   Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS)
   Traffic-Engineered (TE) networks.

   This document examines the applicability of the PCE architecture,
   protocols, and protocol extensions for computing multi-area and
   multi-AS paths in MPLS and GMPLS networks.

Status of This Memo

   This document is not an Internet Standards Track specification; it is
   published for informational purposes.

   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 candidates 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
   https://www.rfc-editor.org/info/rfc8694.

Table of Contents

1. Introduction

   Computing paths across large multi-domain environments may require
   special computational components and cooperation between entities in
   different domains capable of complex path computation.

   Issues that may exist when routing in multi-domain networks include
   the following:

   * There is often a lack of full topology and TE information across
      domains.

   * No single node has the full visibility to determine an optimal or
      even feasible end-to-end path across domains.

   * Knowing how to evaluate and select the exit point and next domain
      boundary from a domain.

   * Understanding how the ingress node determines which domains should
      be used for the end-to-end path.

   An information exchange across multiple domains is often limited due
   to the lack of trust relationship, security issues, or scalability
   issues, even if there is a trust relationship between domains.

   The Path Computation Element (PCE) [RFC4655] provides an architecture

and a set of functional components to address the problem space and
the issues highlighted above.

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 PCE-based computation (BRPC)
mechanism [RFC5441] defines a path computation procedure to compute
inter-domain constrained Multiprotocol Label Switching (MPLS) and
Generalized MPLS (GMPLS) Traffic-Engineered (TE) networks. 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.

In more advanced deployments (including multi-area and multi-
Autonomous System (multi-AS) environments), the sequence of domains
may not be known in advance, and the choice of domains in the end-to-
end domain sequence might be critical to the determination of an
optimal end-to-end path. In this case, the use of the hierarchical
PCE [RFC6805] architecture and mechanisms may be used to discover the
intra-area path and select the optimal end-to-end domain sequence.

This document describes the processes and procedures available when
using the PCE architecture and protocols for computing inter-area and
inter-AS MPLS and GMPLS Traffic-Engineered paths.

The scope of this document does not include discussions of deployment
scenarios for stateful PCE, active PCE, remotely initiated PCE, or
PCE as a central controller (PCECC).

1.1. Domains

Generally, a domain can be defined as a separate administrative,
geographic, or switching environment within the network. A domain
may be further defined as a zone of routing or computational ability.
Under these definitions, a domain might be categorized as an
Autonomous System (AS) or an Interior Gateway Protocol (IGP) area (as
per [RFC4726] and [RFC4655]).

For the purposes of this document, a domain is considered to be a
collection of network elements within an area or AS that has a common
sphere of address management or path computational responsibility.
Wholly or partially overlapping domains are not within the scope of
this document.

In the context of GMPLS, a particularly important example of a domain
is the Automatically Switched Optical Network (ASON) subnetwork
[G-8080]. In this case, computation of an end-to-end path requires
the selection of nodes and links within a parent domain where some
nodes may, in fact, be subnetworks. Furthermore, a domain might be
an ASON routing area [G-7715]. A PCE may perform the path
computation function of an ASON Routing Controller as described in
[G-7715-2].

It is assumed that the PCE architecture is not applied to a large
group of domains, such as the Internet.

1.2. Path Computation

For the purpose of this document, it is assumed that path computation
is the sole responsibility of the PCE as per the architecture defined
in [RFC4655]. When a path is required, the Path Computation Client
(PCC) will send a request to the PCE. The PCE will apply the
required constraints, compute a path, and return a response to the
PCC. In the context of this document, it may be necessary for the
PCE to cooperate with other PCEs in adjacent domains (as per BRPC
[RFC5441]) or with a parent PCE (as per [RFC6805]).

It is entirely feasible that an operator could compute a path across
multiple domains without the use of a PCE if the relevant domain
information is available to the network planner or network management
platform. The definition of what relevant information is required to

perform this network planning operation and how that information is
discovered and applied is outside the scope of this document.

1.2.1. PCE-Based Path Computation Procedure

As highlighted, the PCE is an entity capable of computing an inter-
domain TE path upon receiving a request from a PCC. There could be a
single PCE per domain or a single PCE responsible for all domains. A
PCE may or may not reside on the same node as the requesting PCC. A
path may be computed by either a single PCE node or a set of
distributed PCE nodes that collaborate during path computation.

According to [RFC4655], a PCC should send a path computation request
to a particular PCE using [RFC5440] (PCC-to-PCE communication). This
negates the need to broadcast a request to all the PCEs. Each PCC
can maintain information about the computation capabilities of the
PCEs it is aware of. The PCC-PCE capability awareness can be
configured using static configurations or by automatic and dynamic
PCE discovery procedures.

If a network path is required, the PCC will send a path computation
request to the PCE. A PCE may then compute the end-to-end path if it
is aware of the topology and TE information required to compute the
entire path. If the PCE is unable to compute the entire path, the
PCE architecture provides cooperative PCE mechanisms for the
resolution of path computation requests when an individual PCE does
not have sufficient TE visibility.

End-to-end path segments may be kept confidential through the
application of Path-Keys to protect partial or full path information.
A Path-Key is a token that replaces a path segment in an explicit
route. The Path-Key mechanism is described in [RFC5520].

1.3. Traffic Engineering Aggregation and Abstraction

Networks are often constructed from multiple areas or ASes that are
interconnected via multiple interconnect points. To maintain network
confidentiality and scalability, the TE properties of each area and
AS are not generally advertised outside each specific area or AS.

TE aggregation or abstraction provide a mechanism to hide information
but may cause failed path setups or the selection of suboptimal end-
to-end paths [RFC4726]. The aggregation process may also have
significant scaling issues for networks with many possible routes and
multiple TE metrics. Flooding TE information breaks confidentiality
and does not scale in the routing protocol.

The PCE architecture and associated mechanisms provide a solution to
avoid the use of TE aggregation and abstraction.

1.4. Traffic-Engineered Label Switched Paths

This document highlights the PCE techniques and mechanisms that exist
for establishing TE packet and optical Label Switched Paths (LSPs)
across multiple areas (inter-area TE LSP) and ASes (inter-AS TE LSP).
In this context and within the remainder of this document, we
consider all LSPs to be constraint based and traffic engineered.

Three signaling options are defined for setting up an inter-area or
inter-AS LSP [RFC4726]:

* Contiguous LSP

* Stitched LSP

* Nested LSP

All three signaling methods are applicable to the architectures and
procedures discussed in this document.

1.5. Inter-area and Inter-AS-capable PCE Discovery

When using a PCE-based approach for inter-area and inter-AS path
computation, a PCE in one area or AS may need to learn information
related to inter-AS-capable PCEs located in other ASes. The PCE
discovery mechanism defined in [RFC5088] and [RFC5089] facilitates
the discovery of PCEs and disclosure of information related to inter-
area and inter-AS-capable PCEs.

1.6. Objective Functions

An Objective Function (OF) [RFC5541] or a set of OFs specifies the
intentions of the path computation and so defines the "optimality" in
the context of the computation request.

An OF specifies the desired outcome of a computation. It does not
describe or specify the algorithm to use. Also, an implementation
may apply any algorithm or set of algorithms to achieve the result
indicated by the OF. A number of general OFs are specified in
[RFC5541].

Various OFs may be included in the PCE computation request to satisfy
the policies encoded or configured at the PCC, and a PCE may be
subject to policy in determining whether it meets the OFs included in
the computation request or whether it applies its own OFs.

During inter-domain path computation, the selection of a domain
sequence, the computation of each (per-domain) path fragment, and the
determination of the end-to-end path may each be subject to different
OFs and policies.

2. Terminology

This document also uses the terminology defined in [RFC4655] and
[RFC5440]. Additional terminology is defined below:

ABR: IGP Area Border Router -- a router that is attached to more
      than one IGP area.

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

Inter-area TE LSP: A TE LSP whose path transits through two or more
      IGP areas.

Inter-AS MPLS TE LSP: A TE LSP whose path transits through two or
      more ASes or sub-ASes (BGP confederations)

SRLG: Shared Risk Link Group.

TED: Traffic Engineering Database, which contains the topology and
      resource information of the domain. The TED may be fed by
      Interior Gateway Protocol (IGP) extensions or potentially by
      other means.

3. Issues and Considerations

3.1. Multihoming

Networks constructed from multi-areas or multi-AS environments may
have multiple interconnect points (multihoming). End-to-end path
computations may need to use different interconnect points to avoid a
single-point failure disrupting both the primary and backup services.

3.2. Destination Location

A PCC asking for an inter-domain path computation is typically aware
of the identity of the destination node. If the PCC is aware of the
destination domain, it may supply the destination domain information
as part of the path computation request. However, if the PCC does
not know the destination domain, this information must be determined

by another method.

3.3. Domain Confidentiality

When the end-to-end path crosses multiple domains, it may be possible that each domain (AS or area) is administered by separate Service Providers. Thus, if a PCE supplies a path segment to a PCE in another domain, it may break confidentiality rules and could disclose AS-internal topology information.

If confidentiality is required between domains (ASes and areas) belonging to different Service Providers, then cooperating PCEs cannot exchange path segments; otherwise, the receiving PCE or PCC will be able to see the individual hops through another domain.

This topic is discussed further in Section 8 of this document.

4. Domain Topologies

Constraint-based inter-domain path computation is a fundamental requirement for operating traffic-engineered MPLS [RFC3209] and GMPLS [RFC3473] networks in inter-area and inter-AS (multi-domain) environments. Path computation across multi-domain networks is complex and requires computational cooperational entities like the PCE.

4.1. Selecting Domain Paths

Where the sequence of domains is known a priori, various techniques can be employed to derive an optimal multi-domain path. If the domains are connected to a simple path with no branches and single links between all domains or if the preferred points of interconnection are also known, the per-domain path computation [RFC5152] technique may be used. Where there are multiple connections between domains and there is no preference for the choice of points of interconnection, BRPC [RFC5441] can be used to derive an optimal path.

When the sequence of domains is not known in advance or the end-to-end path will have to navigate a mesh of small domains (especially typical in optical networks), the optimum path may be derived through the application of a hierarchical PCE [RFC6805].

4.2. Domain Sizes

Very frequently, network domains are composed of dozens or hundreds of network elements. These network elements are usually interconnected in a partial-mesh fashion to provide survivability against dual failures and to benefit from the traffic-engineering capabilities of MPLS and GMPLS protocols. Network operator feedback in the development of the document highlighted that the node degree (the number of neighbors per node) typically ranges from 3 to 10 (4-5 is quite common).

4.3. Domain Diversity

Domain and path diversity may also be required when computing end-to-end paths. Domain diversity should facilitate the selection of paths that share ingress and egress domains but do not share transit domains. Therefore, there must be a method allowing the inclusion or exclusion of specific domains when computing end-to-end paths.

4.4. Synchronized Path Computations

In some scenarios, it would be beneficial for the operator to rely on the capability of the PCE to perform synchronized path computation.

Synchronized path computations, known as Synchronization VECtors (SVECs), are used for dependent path computations. SVECs are defined in [RFC5440], and [RFC6007] provides an overview of the use of the PCE SVEC list for synchronized path computations when computing

dependent requests.

In hierarchical PCE (H-PCE) deployments, a child PCE will be able to
request both dependent and synchronized domain-diverse end-to-end
paths from its parent PCE.

4.5. Domain Inclusion or Exclusion

A domain sequence is an ordered sequence of domains traversed to
reach the destination domain. A domain sequence may be supplied
during path computation to guide the PCEs or are derived via the use
of hierarchical PCE (H-PCE).

During multi-domain path computation, a PCC may request specific
domains to be included or excluded in the domain sequence using the
Include Route Object (IRO) [RFC5440] and Exclude Route Object (XRO)
[RFC5521]. The use of Autonomous Number (AS) as an abstract node
representing a domain is defined in [RFC3209]. [RFC7897] specifies
new subobjects to include or exclude domains such as an IGP area or a
4-byte AS number.

An operator may also need to avoid a path that uses specified nodes
for administrative reasons. If a specific connectivity service is
required to have a 1+1 protection capability, two separate disjoint
paths must be established. A mechanism known as Shared Risk Link
Group (SRLG) information may be used to ensure path diversity.

5. Applicability of the PCE to Inter-area Traffic Engineering

As networks increase in size and complexity, it may be required to
introduce scaling methods to reduce the amount of information flooded
within the network and make the network more manageable. An IGP
hierarchy is designed to improve IGP scalability by dividing the IGP
domain into areas and limiting the flooding scope of topology
information to within area boundaries. This restricts visibility of
the area to routers in a single area. If a router needs to compute
the route to a destination located in another area, a method would be
required to compute a path across area boundaries.

In order to support multiple vendors in a network in cases where data
or control-plane technologies cannot interoperate, it is useful to
divide the network into vendor domains. Each vendor domain is an IGP
area, and the flooding scope of the topology (as well as any other
relevant information) is limited to the area boundaries.

Per-domain path computation [RFC5152] exists to provide a method of
inter-area path computation. The per-domain solution is based on
loose hop routing with an Explicit Route Object (ERO) expansion on
each Area Border Router (ABR). This allows an LSP to be established
using a constrained path. However, at least two issues exist:

* This method does not guarantee an optimal constrained path.

* The method may require several crankback signaling messages, as
   per [RFC4920], increasing signaling traffic and delaying the LSP
   setup.

PCE-based architecture [RFC4655] is designed to solve inter-area path
computation problems. The issue of limited topology visibility is
resolved by introducing path computation entities that are able to
cooperate in order to establish LSPs with the source and destinations
located in different areas.

5.1. Inter-area Routing

An inter-area TE-LSP is an LSP that transits through at least two IGP
areas. In a multi-area network, topology visibility remains local to
a given area for scaling and privacy purposes. A node in one area
will not be able to compute an end-to-end path across multiple areas
without the use of a PCE.

5.1.1. Area Inclusion and Exclusion

The BRPC method [RFC5441] of path computation provides a more optimal method to specify inclusion or exclusion of an ABR. Using the BRPC procedure, an end-to-end path is recursively computed in reverse from the destination domain towards the source domain. Using this method, an operator might decide if an area must be included or excluded from the inter-area path computation.

5.1.2. Strict Explicit Path and Loose Path

A strict explicit path is defined as a set of strict hops, while a loose path is defined as a set of at least one loose hop and zero or more strict hops. It may be useful to indicate whether a strict explicit path is required during the path computation request. An inter-area path may be strictly explicit or loose (e.g., a list of ABRs as loose hops).

A PCC request to a PCE does allow indication of whether a strict explicit path across specific areas ([RFC7897]) is required or desired or whether the path request is loose.

5.1.3. Inter-Area Diverse Path Computation

It may be necessary to compute a path that is partially or entirely diverse from a previously computed path to avoid fate sharing of a primary service with a corresponding backup service. There are various levels of diversity in the context of an inter-area network:

*   Per-area diversity (the intra-area path segments are a link, node, or SRLG disjoint).

*   Inter-area diversity (the end-to-end inter-area paths are a link, node, or SRLG disjoint).

Note that two paths may be disjointed in the backbone area but non-disjointed in peripheral areas. Also, two paths may be node disjointed within areas but may share ABRs, in which case path segments within an area are node disjointed but end-to-end paths are not node disjointed. Per-domain [RFC5152], BRPC [RFC5441], and H-PCE [RFC6805] mechanisms all support the capability to compute diverse paths across multi-area topologies.

6. Applicability of the PCE to Inter-AS Traffic Engineering

As discussed in Section 5 (Applicability of the PCE to Inter-area Traffic Engineering), it is necessary to divide the network into smaller administrative domains, or ASes. If an LSR within an AS needs to compute a path across an AS boundary, it must also use an inter-AS computation technique. [RFC5152] defines mechanisms for the computation of inter-domain TE LSPs using network elements along the signaling paths to compute per-domain constrained path segments.

The PCE was designed to be capable of computing MPLS and GMPLS paths across AS boundaries. This section outlines the features of a PCE-enabled solution for computing inter-AS paths.

6.1. Inter-AS Routing

6.1.1. AS Inclusion and Exclusion

[RFC5441] allows the specification of AS or ASBR inclusion or exclusion. Using this method, an operator might decide whether an AS must be included or excluded from the inter-AS path computation. Exclusion and/or inclusion could also be specified at any step in the LSP path computation process by a PCE (within the BRPC algorithm), but the best practice would be to specify them at the edge. In opposition to the strict and loose path, AS inclusion or exclusion doesn't impose topology disclosure as ASes and their interconnection are public entities.

6.2. Inter-AS Bandwidth Guarantees

Many operators with multi-AS domains will have deployed the MPLS-TE
Diffserv either across their entire network or at the domain edges on
CE-PE links. In situations where strict QoS bounds are required,
admission control inside the network may also be required.

When the propagation delay can be bounded, the performance targets,
such as maximum one-way transit delay, may be guaranteed by providing
bandwidth guarantees along the Diffserv-enabled path. These
requirements are described in [RFC4216].

One typical example of the requirements in [RFC4216] is to provide
bandwidth guarantees over an end-to-end path for VoIP traffic
classified as an EF (Expedited Forwarding) class in a Diffserv-
enabled network. In cases where the EF path is extended across
multiple ASes, an inter-AS bandwidth guarantee would be required.

Another case for an inter-AS bandwidth guarantee is the requirement
to guarantee a certain amount of transit bandwidth across one or
multiple ASes.

6.3. Inter-AS Recovery

During a path computation process, a PCC request may contain the
requirement to compute a backup LSP for protecting the primary LSP,
such as 1+1 protection. A single LSP or multiple backup LSPs may
also be used for a group of primary LSPs; this is typically known as
m:n protection.

Other inter-AS recovery mechanisms include [RFC4090], which adds Fast
Reroute (FRR) protection to an LSP. So, the PCE could be used to
trigger computation of backup tunnels in order to protect inter-AS
connectivity.

Inter-AS recovery clearly requires backup LSPs for service
protection, but it would also be advisable to have multiple PCEs
deployed for path computation redundancy, especially for service
restoration in the event of catastrophic network failure.

6.4. Inter-AS PCE Peering Policies

Like BGP peering policies, inter-AS PCE peering policies are required
for an operator. In an inter-AS BRPC process, the PCE must cooperate
in order to compute the end-to-end LSP. Therefore, the AS path must
not only follow technical constraints, e.g., bandwidth availability,
but also the policies defined by the operator.

Typically, PCE interconnections at an AS level must follow the agreed
contract obligations, also known as peering agreements. The PCE
peering policies are the result of the contract negotiation and
govern the relation between the different PCEs.

7. Multi-domain PCE Deployment Options

7.1. Traffic Engineering Database and Synchronization

An optimal path computation requires knowledge of the available
network resources, including nodes and links, constraints, link
connectivity, available bandwidth, and link costs. The PCE operates
on a view of the network topology as presented by a TED. As
discussed in [RFC4655], the TED used by a PCE may be learned by the
relevant IGP extensions.

Thus, the PCE may operate its TED by participating in the IGP running
in the network. In an MPLS-TE network, this would require OSPF-TE
[RFC3630] or ISIS-TE [RFC5305]. In a GMPLS network, it would utilize
the GMPLS extensions to OSPF and IS-IS defined in [RFC4203] and
[RFC5307]. Inter-AS connectivity information may be populated via
[RFC5316] and [RFC5392].

An alternative method to providing network topology and resource
   information is offered by [RFC7752], which is described in the
   following section.

7.1.1. Applicability of BGP-LS to PCE

   The concept of the exchange of TE information between Autonomous
   Systems (ASes) is discussed in [RFC7752]. The information exchanged
   in this way could be the full TE information from the AS, an
   aggregation of that information, or a representation of the potential
   connectivity across the AS. Furthermore, that information could be
   updated frequently (for example, for every new LSP that is set up
   across the AS) or only at threshold-crossing events.

   In an H-PCE deployment, the parent PCE will require the inter-domain
   topology and link status between child domains. This information may
   be learned by a BGP-LS speaker and provided to the parent PCE.
   Furthermore, link-state performance, including delay, available
   bandwidth, and utilized bandwidth, may also be provided to the parent
   PCE for optimal path link selection.

7.2. Pre-planning and Management-Based Solutions

   Offline path computation is performed ahead of time before the LSP
   setup is requested. That means that it is requested by or performed
   as part of an Operation Support System (OSS) management application.
   This model can be seen in Section 5.5 of [RFC4655].

   The offline model is particularly appropriate for long-lived LSPs
   (such as those present in a transport network) or for planned
   responses to network failures. In these scenarios, more planning is
   normally a feature of LSP provisioning.

   The management system may also use a PCE and BRPC to pre-plan an AS
   sequence, and the source domain PCE and per-domain path computation
   to be used when the actual end-to-end path is required. This model
   may also be used where the operator wishes to retain full manual
   control of the placement of LSPs, using the PCE only as a computation
   tool to assist the operator and not as part of an automated network.

   In environments where operators peer with each other to provide end-
   to-end paths, the operator responsible for each domain must agree on
   the extent to which paths must be pre-planned or manually controlled.

8. Domain Confidentiality

   This section discusses the techniques that cooperating PCEs can use
   to compute inter-domain paths without each domain disclosing
   sensitive internal topology information (such as explicit nodes or
   links within the domain) to the other domains.

   Confidentiality typically applies to inter-provider (inter-AS) PCE
   communication. Where the TE LSP crosses multiple domains (ASes or
   areas), the path may be computed by multiple PCEs that cooperate
   together, with each local PCE responsible for computing a segment of
   the path. With each local PCE responsible for computing a segment of
   the path.

   In situations where ASes are administered by separate Service
   Providers, it would break confidentiality rules for a PCE to supply
   path segment details to a PCE responsible for another domain, thus
   disclosing AS-internal or area topology information.

8.1. Loose Hops

   A method for preserving the confidentiality of the path segment is
   for the PCE to return a path containing a loose hop in place of the
   segment that must be kept confidential. The concept of loose and
   strict hops for the route of a TE LSP is described in [RFC3209].

   [RFC5440] supports the use of paths with loose hops; whether it

returns a full explicit path with strict hops or uses loose hops is a local policy decision at a PCE. A path computation request may require an explicit path with strict hops or may allow loose hops, as detailed in [RFC5440].

8.2. Confidential Path Segments and Path-Keys

   [RFC5520] defines the concept and mechanism of a Path-Key. A Path-Key is a token that replaces the path segment information in an explicit route. The Path-Key allows the explicit route information to be encoded and is contained in the Path Computation Element Communication Protocol (PCEP) ([RFC5440]) messages exchanged between the PCE and PCC.

   This Path-Key technique allows explicit route information to be used for end-to-end path computation without disclosing internal topology information between domains.

9. Point to Multipoint

   For inter-domain point-to-multipoint application scenarios using MPLS-TE LSPs, the complexity of domain sequences, domain policies, and the choice and number of domain interconnects is magnified compared to point-to-point path computations. As the size of the network grows, the number of leaves and branches increases, further increasing the complexity of the overall path computation problem. A solution for managing point-to-multipoint path computations may be achieved using the PCE inter-domain point-to-multipoint path computation [RFC7334] procedure.

10. Optical Domains

   The International Telecommunication Union (ITU) defines the ASON architecture in [G-8080]. [G-7715] defines the routing architecture for ASON and introduces a hierarchical architecture. In this architecture, the Routing Areas (RAs) have a hierarchical relationship between different routing levels, which means a parent (or higher level) RA can contain multiple child RAs. The interconnectivity of the lower RAs is visible to the higher-level RA.

   In the ASON framework, a path computation request is termed a route query. This query is executed before signaling is used to establish an LSP, which is termed a Switched Connection (SC) or a Soft Permanent Connection (SPC). [G-7715-2] defines the requirements and architecture for the functions performed by Routing Controllers (RC) during the operation of remote route queries. An RC is synonymous with a PCE.

   In the ASON routing environment, an RC responsible for an RA may communicate with its neighbor RC to request the computation of an end-to-end path across several RAs. The path computation components and sequences are defined as follows:

   * Remote route query. An operation where a Routing Controller communicates with another Routing Controller, which does not have the same set of layer resources, in order to compute a routing path in a collaborative manner.

   * Route query requester. The connection controller or RC that sends a route query message to a Routing Controller that requests one or more routing paths satisfying a set of routing constraints.

   * Route query responder. An RC that performs the path computation upon reception of a route query message from a Routing Controller or connection controller, and sends a response back at the end of the computation.

   When computing an end-to-end connection, the route may be computed by a single RC or multiple RCs in a collaborative manner, and the two scenarios can be considered a centralized remote route query model and a distributed remote route query model. RCs in an ASON

environment can also use the hierarchical PCE [RFC6805] model to
fully match the ASON hierarchical routing model.

10.1. Abstraction and Control of TE Networks (ACTN)

Where a single operator operates multiple TE domains (including
optical environments), an Abstraction and Control of TE Networks
(ACTN) framework [RFC8453] may be used to create an abstracted
(virtualized network) view of underlay-interconnected domains. This
underlay connectivity is then exposed to higher-layer control
entities and applications.

ACTN describes the method and procedure for coordinating the underlay
per-domain Provisioning Network Controllers (PNCs), which may be
PCEs, via a hierarchical model to facilitate setup of end-to-end
connections across interconnected TE domains.

11. Policy

Policy is important in the deployment of new services and the
operation of the network. [RFC5394] provides a framework for PCE-
based policy-enabled path computation. This framework is based on
the Policy Core Information Model (PCIM) as defined in [RFC3060] and
further extended by [RFC3460].

When using a PCE to compute inter-domain paths, policy may be invoked
by specifying the following:

* Each PCC must select which computations it will request from a
   PCE.

* Each PCC must select which PCEs it will use.

* Each PCE must determine which PCCs are allowed to use its services
   and for what computations.

* The PCE must determine how to collect the information in its TED,
   whom to trust for that information, and how to refresh/update the
   information.

* Each PCE must determine which objective functions and algorithms
   to apply.

12. Manageability Considerations

General PCE management considerations are discussed in [RFC4655]. In
the case of multi-domains within a single service provider network,
the management responsibility for each PCE would most likely be
handled by the same service provider. In the case of multiple ASes
within different service provider networks, it will likely be
necessary for each PCE to be configured and managed separately by
each participating service provider, with policy being implemented
based on a previously agreed set of principles.

12.1. Control of Function and Policy

As per [RFC5440], PCEP implementation allows the user to configure a
number of PCEP session parameters. These are detailed in Section 8.1
of [RFC5440].

In H-PCE deployments, the administrative entity responsible for the
management of the parent PCEs for multi-areas would typically be a
single service provider. In multiple ASes (managed by different
service providers), it may be necessary for a third party to manage
the parent PCE.

12.2. Information and Data Models

A PCEP MIB module is defined in [RFC7420], which describes managed
objects for modeling PCEP communication, including:

* PCEP client configuration and status.

* PCEP peer configuration and information.

* PCEP session configuration and information.

* Notifications to indicate PCEP session changes.

A YANG module for PCEP has also been proposed [PCEP-YANG].

An H-PCE MIB module or YANG data model will be required to report parent PCE and child PCE information, including:

* Parent PCE configuration and status.

* Child PCE configuration and information.

* Notifications to indicate session changes between parent PCEs and child PCEs.

* Notification of parent PCE TED updates and changes.

12.3. Liveness Detection and Monitoring

PCEP includes a keepalive mechanism to check the liveliness of a PCEP peer and a notification procedure allowing a PCE to advertise its overloaded state to a PCC. In a multi-domain environment, [RFC5886] provides the procedures necessary to monitor the liveliness and performance of a given PCE chain.

12.4. Verifying Correct Operation

It is important to verify the correct operation of PCEP. [RFC5440] specifies the monitoring of key parameters. These parameters are detailed in [RFC5520].

12.5. Impact on Network Operation

[RFC5440] states that in order to avoid any unacceptable impact on network operations, a PCEP implementation should allow a limit to be placed on the number of sessions that can be set up on a PCEP speaker and that it may also be practical to place a limit on the rate of messages sent by a PCC and received by the PCE.

13. Security Considerations

PCEP security considerations are discussed in [RFC5440] and [RFC6952]. Potential vulnerabilities include spoofing, snooping, falsification, and using PCEP as a mechanism for denial of service attacks.

As PCEP operates over TCP, it may make use of TCP security encryption mechanisms, such as Transport Layer Security (TLS) and TCP Authentication Option (TCP-AO). Usage of these security mechanisms for PCEP is described in [RFC8253], and recommendations and best current practices are described in [RFC7525].

13.1. Multi-domain Security

Any multi-domain operation necessarily involves the exchange of information across domain boundaries. This represents a significant security and confidentiality risk.

It is expected that PCEP is used between PCCs and PCEs that belong to the same administrative authority while also using one of the aforementioned encryption mechanisms. Furthermore, PCEP allows individual PCEs to maintain the confidentiality of their domain path information using path-keys.

14. IANA Considerations

This document has no IANA actions.

15. References

15.1. Normative 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, DOI 10.17487/RFC3209, December 2001,
              <https://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,
              <https://www.rfc-editor.org/info/rfc3473>.

   [RFC4216] Zhang, R., Ed. and J.-P. Vasseur, Ed., "MPLS Inter-
              Autonomous System (AS) Traffic Engineering (TE)
              Requirements", RFC 4216, DOI 10.17487/RFC4216, November
              2005, <https://www.rfc-editor.org/info/rfc4216>.

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

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

   [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,
              <https://www.rfc-editor.org/info/rfc5152>.

   [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation
              Element (PCE) Communication Protocol (PCEP)", RFC 5440,
              DOI 10.17487/RFC5440, March 2009,
              <https://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,
              <https://www.rfc-editor.org/info/rfc5441>.

   [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,
              <https://www.rfc-editor.org/info/rfc5520>.

   [RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of
              Objective Functions in the Path Computation Element
              Communication Protocol (PCEP)", RFC 5541,
              DOI 10.17487/RFC5541, June 2009,
              <https://www.rfc-editor.org/info/rfc5541>.

   [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,
              <https://www.rfc-editor.org/info/rfc6805>.

15.2. Informative References

   [RFC3060] Moore, B., Ellesson, E., Strassner, J., and A. Westerinen,

                "Policy Core Information Model -- Version 1
                Specification", RFC 3060, DOI 10.17487/RFC3060, February
                2001, <https://www.rfc-editor.org/info/rfc3060>.

[RFC3460]   Moore, B., Ed., "Policy Core Information Model (PCIM)
                Extensions", RFC 3460, DOI 10.17487/RFC3460, January 2003,
                <https://www.rfc-editor.org/info/rfc3460>.

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

[RFC4090]   Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast
                Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090,
                DOI 10.17487/RFC4090, May 2005,
                <https://www.rfc-editor.org/info/rfc4090>.

[RFC4203]   Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in
                Support of Generalized Multi-Protocol Label Switching
                (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005,
                <https://www.rfc-editor.org/info/rfc4203>.

[RFC4920]   Farrel, A., Ed., Satyanarayana, A., Iwata, A., Fujita, N.,
                and G. Ash, "Crankback Signaling Extensions for MPLS and
                GMPLS RSVP-TE", RFC 4920, DOI 10.17487/RFC4920, July 2007,
                <https://www.rfc-editor.org/info/rfc4920>.

[RFC5088]   Le Roux, JL., Ed., Vasseur, JP., Ed., Ikejiri, Y., and R.
                Zhang, "OSPF Protocol Extensions for Path Computation
                Element (PCE) Discovery", RFC 5088, DOI 10.17487/RFC5088,
                January 2008, <https://www.rfc-editor.org/info/rfc5088>.

[RFC5089]   Le Roux, JL., Ed., Vasseur, JP., Ed., Ikejiri, Y., and R.
                Zhang, "IS-IS Protocol Extensions for Path Computation
                Element (PCE) Discovery", RFC 5089, DOI 10.17487/RFC5089,
                January 2008, <https://www.rfc-editor.org/info/rfc5089>.

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

[RFC5307]   Kompella, K., Ed. and Y. Rekhter, Ed., "IS-IS Extensions
                in Support of Generalized Multi-Protocol Label Switching
                (GMPLS)", RFC 5307, DOI 10.17487/RFC5307, October 2008,
                <https://www.rfc-editor.org/info/rfc5307>.

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

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

[RFC5394]   Bryskin, I., Papadimitriou, D., Berger, L., and J. Ash,
                "Policy-Enabled Path Computation Framework", RFC 5394,
                DOI 10.17487/RFC5394, December 2008,
                <https://www.rfc-editor.org/info/rfc5394>.

[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, <https://www.rfc-editor.org/info/rfc5521>.

[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,
                <https://www.rfc-editor.org/info/rfc5886>.

[RFC6007]   Nishioka, I. and D. King, "Use of the Synchronization
            VECtor (SVEC) List for Synchronized Dependent Path
            Computations", RFC 6007, DOI 10.17487/RFC6007, September
            2010, <https://www.rfc-editor.org/info/rfc6007>.

[G-8080]   ITU-T, "Architecture for the automatically switched
            optical network", ITU-T Recommendation G.8080/Y.1304,
            February 2012.

[G-7715]   ITU-T, "Architecture and requirements for routing in the
            automatically switched optical networks", ITU-T
            Recommendation G.7715/Y.1706, June 2002.

[G-7715-2] ITU-T, "ASON routing architecture and requirements for
            remote route query", ITU-T
            Recommendation G.7715.2/Y.1706.2, February 2007.

[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,
            <https://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,
            <https://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,
            <https://www.rfc-editor.org/info/rfc7420>.

[RFC7525]   Sheffer, Y., Holz, R., and P. Saint-Andre,
            "Recommendations for Secure Use of Transport Layer
            Security (TLS) and Datagram Transport Layer Security
            (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May
            2015, <https://www.rfc-editor.org/info/rfc7525>.

[RFC7752]   Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and
            S. Ray, "North-Bound Distribution of Link-State and
            Traffic Engineering (TE) Information Using BGP", RFC 7752,
            DOI 10.17487/RFC7752, March 2016,
            <https://www.rfc-editor.org/info/rfc7752>.

[RFC7897]   Dhody, D., Palle, U., and R. Casellas, "Domain Subobjects
            for the Path Computation Element Communication Protocol
            (PCEP)", RFC 7897, DOI 10.17487/RFC7897, June 2016,
            <https://www.rfc-editor.org/info/rfc7897>.

[RFC8253]   Lopez, D., Gonzalez de Dios, O., Wu, Q., and D. Dhody,
            "PCEPS: Usage of TLS to Provide a Secure Transport for the
            Path Computation Element Communication Protocol (PCEP)",
            RFC 8253, DOI 10.17487/RFC8253, October 2017,
            <https://www.rfc-editor.org/info/rfc8253>.

[RFC8453]   Ceccarelli, D., Ed. and Y. Lee, Ed., "Framework for
            Abstraction and Control of TE Networks (ACTN)", RFC 8453,
            DOI 10.17487/RFC8453, August 2018,
            <https://www.rfc-editor.org/info/rfc8453>.

[PCEP-YANG]
            Dhody, D., Hardwick, J., Beeram, V., and J. Tantsura, "A
            YANG Data Model for Path Computation Element
            Communications Protocol (PCEP)", Work in Progress,
            Internet-Draft, draft-ietf-pce-pcep-yang-13, 31 October
            2019,

<https://tools.ietf.org/html/draft-ietf-pce-pcep-yang-13>.

Contributors

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

   Email: dhruv.ietf@gmail.com

   Quintin Zhao
   Huawei Technologies
   125 Nagog Technology Park
   Acton, MA 01719
   United States of America

   Email: qzhao@huawei.com

   Julien Meuric
   France Telecom
   2, avenue Pierre-Marzin
   22307 Lannion Cedex
   France

   Email: julien.meuric@orange.com

   Olivier Dugeon
   France Telecom
   2, avenue Pierre-Marzin
   22307 Lannion Cedex
   France

   Email: olivier.dugeon@orange.com

   Jon Hardwick
   Metaswitch Networks
   100 Church Street
   Enfield
   EN2 6BQ
   United Kingdom

   Email: jonathan.hardwick@metaswitch.com

   Ã\223scar GonzÃ¡lez de Dios
   Telefonica I+D
   Emilio Vargas 6
   Madrid
   Spain

   Email: oscar.gonzalezdedios@telefonica.com

Authors' Addresses

   Daniel King
   Old Dog Consulting

   Email: daniel@olddog.co.uk

Haomian Zheng
Huawei Technologies
H1, Huawei Xiliu Beipo Village, Songshan Lake
Dongguan
Guangdong, 523808
China

Email: zhenghaomian@huawei.com

Additional contact information:

é\203\221å¥½æ£\211
ä¸å\233½
523808
å¹¿ä¸\234 ä¸\234è\216\236
æ\235¾å±±æ¹\226å\215\216ä¸ºæºªæµ\201è\203\214å\235¡æ\235\221H1
å\215\216ä¸ºæ\212\200æ\234¯æ\234\211é\231\220å\205¬å\217¸