    Framework for PCE-Based Inter-Layer MPLS and GMPLS Traffic Engineering

Abstract

   A network may comprise multiple layers. It is important to globally
   optimize network resource utilization, taking into account all layers
   rather than optimizing resource utilization at each layer
   independently. This allows better network efficiency to be achieved
   through a process that we call inter-layer traffic engineering. The
   Path Computation Element (PCE) can be a powerful tool to achieve
   inter-layer traffic engineering.

   This document describes a framework for applying the PCE-based
   architecture to inter-layer Multiprotocol Label Switching (MPLS) and
   Generalized MPLS (GMPLS) traffic engineering. It provides
   suggestions for the deployment of PCE in support of multi-layer
   networks. This document also describes network models where PCE
   performs inter-layer traffic engineering, and the relationship
   between PCE and a functional component called the Virtual Network
   Topology Manager (VNTM).

Status of This Memo

   This memo provides information for the Internet community. It does
   not specify an Internet standard of any kind. Distribution of this
   memo is unlimited.

to this document. Code Components extracted from this document must
include Simplified BSD License text as described in Section 4.e of
the Trust Legal Provisions and are provided without warranty as
described in the BSD License.

Table of Contents

1. Introduction

   A network may comprise multiple layers. These layers may represent
   separations of technologies (e.g., packet switch capable (PSC), time
   division multiplex (TDM), or lambda switch capable (LSC)) [RFC3945],
   separation of data plane switching granularity levels (e.g., PSC-1,
   PSC-2, VC4, or VC12) [RFC5212], or a distinction between client and
   server networking roles. In this multi-layer network, Label Switched
   Paths (LSPs) in a lower layer are used to carry higher-layer LSPs
   across the lower-layer network. The network topology formed by
   lower-layer LSPs and advertised as traffic engineering links (TE
   links) in the higher-layer network is called the Virtual Network
   Topology (VNT) [RFC5212].

   It may be effective to optimize network resource utilization
   globally, i.e., taking into account all layers rather than optimizing
   resource utilization at each layer independently. This allows better
   network efficiency to be achieved and is what we call inter-layer
   traffic engineering. Inter-layer traffic engineering includes using
   mechanisms that allow the computation of end-to-end paths across
   layers (known as inter-layer path computation) and mechanisms that
   control and manage the Virtual Network Topology (VNT) by setting up
   and releasing LSPs in the lower layers [RFC5212].

   Inter-layer traffic engineering is included in the scope of the Path
   Computation Element (PCE)-based architecture [RFC4655], and PCE can
   provide a suitable mechanism for resolving inter-layer path
   computation issues.

   PCE Communication Protocol requirements for inter-layer traffic
   engineering are set out in [PCC-PCE].

   This document describes a framework for applying the PCE-based
   architecture to inter-layer traffic engineering. It provides
   suggestions for the deployment of PCE in support of multi-layer
   networks. This document also describes network models where PCE
   performs inter-layer traffic engineering as well as describing the
   relationship between PCE and a functional component in charge of the
   control and management of the VNT, called the Virtual Network
   Topology Manager (VNTM).

1.1. Terminology

   This document uses terminology from the PCE-based path computation
   architecture [RFC4655] and also common terminology from Multi-
   Protocol Label Switching (MPLS) [RFC3031], Generalized MPLS (GMPLS)
   [RFC3945], and Multi-Layer Networks [RFC5212].

2. Inter-Layer Path Computation

   This section describes key topics of inter-layer path computation in
   MPLS and GMPLS networks.

   [RFC4206] defines a way to signal a higher-layer LSP that has an
   explicit route and includes hops traversed by LSPs in lower layers.
   The computation of end-to-end paths across layers is called inter-
   layer path computation.

   A Label Switching Router (LSR) in the higher layer might not have
   information on the topology of the lower layer, particularly in an
   overlay or augmented model deployment, and hence may not be able to
   compute an end-to-end path across layers.

   PCE-based inter-layer path computation consists of using one or more
   PCEs to compute an end-to-end path across layers. This could be
   achieved by a single PCE path computation, where the PCE has topology
   information about multiple layers and can directly compute an end-
   to-end path across layers, considering the topology of all of the
   layers. Alternatively, the inter-layer path computation could be
   performed as a multiple PCE computation, where each member of a set
   of PCEs has information about the topology of one or more layers (but
   not all layers) and the PCEs collaborate to compute an end-to-end
   path.

```
       -----     -----                   -----    -----
      | LSR |--| LSR |................| LSR |--| LSR |
      | H1  |  | H2  |                | H3  |  | H4  |
       -----    -----\                /-----    -----
                      \-----    -----/
                      | LSR |--| LSR |
                      | L1  |  | L2  |
                       -----    -----
```
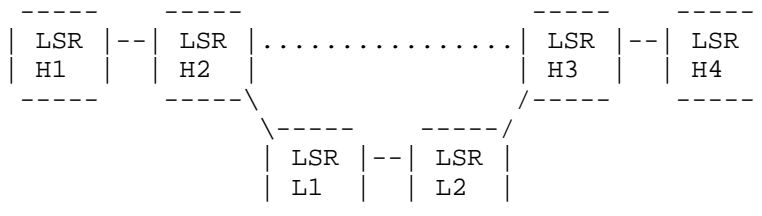
            Figure 1: A Simple Example of a Multi-Layer Network

   Consider, for instance, the two-layer network shown in Figure 1,
   where the higher-layer network (LSRs H1, H2, H3, and H4) is a
   packet-based IP/MPLS or GMPLS network, and the lower-layer network
   (LSRs, H2, L1, L2, and H3) is a GMPLS optical network. An ingress
   LSR in the higher-layer network (H1) tries to set up an LSP to an
   egress LSR (H4) also in the higher-layer network across the lower-
   layer network, and needs a path in the higher-layer network.
   However, suppose that there is no TE link in the higher-layer network
   between the border LSRs located on the boundary between the higher-
   layer and lower-layer networks (H2 and H3). Suppose also that the
   ingress LSR does not have topology visibility into the lower layer.

If a single-layer path computation is applied in the higher-layer,
the path computation fails because of the missing TE link. On the
other hand, inter-layer path computation is able to provide a route
in the higher-layer (H1-H2-H3-H4) and to suggest that a lower-layer
LSP be set up between the border LSRs (H2-L1-L2-H3).

Lower-layer LSPs that are advertised as TE links into the higher-
layer network form a Virtual Network Topology (VNT) that can be used
for routing higher-layer LSPs. Inter-layer path computation for end-
to-end LSPs in the higher-layer network that span the lower-layer
network may utilize the VNT, and PCE is a candidate for computing the
paths of such higher-layer LSPs within the higher-layer network.
Alternatively, the PCE-based path computation model can:

- Perform a single computation on behalf of the ingress LSR using
  information gathered from more than one layer. This mode is
  referred to as single PCE computation in [RFC4655].

- Compute a path on behalf of the ingress LSR through cooperation
  with PCEs responsible for each layer. This mode is referred to as
  multiple PCE computation with inter-PCE communication in [RFC4655].

- Perform separate path computations on behalf of the TE-LSP head-
  end and each transit border LSR that is the entry point to a new
  layer. This mode is referred to as multiple PCE computation
  (without inter-PCE communication) in [RFC4655]. This option
  utilizes per-layer path computation, which is performed
  independently by successive PCEs.

Note that when a network consists of more than two layers (e.g., MPLS
over SONET over Optical Transport Network (OTN)) and a path
traversing more than two layers needs to be computed, it is possible
to combine multiple PCE-based path computation models. For example,
the single PCE computation model could be used for computing a path
across the SONET layer and the OTN layer, and the multiple PCE
computation with inter-PCE communication model could be used for
computing a path across the MPLS layer (computed by higher-layer PCE)
and the SONET layer (computed by lower-layer PCE).

The PCE invoked by the head-end LSR computes a path that the LSR can
use to signal an MPLS-TE or GMPLS LSP once the path information has
been converted to an Explicit Route Object (ERO) for use in RSVP-TE
signaling. There are two options.

   - Option 1: Mono-Layer Path

     The PCE computes a "mono-layer" path, i.e., a path that includes
     only TE links from the same layer. There are two cases for this
     option. In the first case, the PCE computes a path that includes
     already established lower-layer LSPs or lower-layer LSPs to be
     established on demand. That is, the resulting ERO includes
     subobject(s) corresponding to lower-layer hierarchical LSPs
     expressed as the TE link identifiers of the hierarchical LSPs when
     advertised as TE links in the higher-layer network. The TE link
     may be a regular TE link that is actually established or a virtual
     TE link that is not established yet (see [RFC5212]). If it is a
     virtual TE link, this triggers a setup attempt for a new lower-
     layer LSP when signaling reaches the head-end of the lower-layer
     LSP. Note that the path of a virtual TE link is not necessarily
     known in advance, and this may require a further (lower-layer) path
     computation.

     The second case is that the PCE computes a path that includes a
     loose hop that spans the lower-layer network. The higher-layer
     path computation selects which lower-layer network to use and the
     entry and exit points of that lower-layer network, but does not
     select the path across the lower-layer network. A transit LSR that
     is the entry point to the lower-layer network is expected to expand
     the loose hop (either itself or relying on the services of a PCE).
     The path expansion process on the border LSR may result either in
     the selection of an existing lower-layer LSP or in the computation
     and setup of a new lower-layer LSP.

     Note that even if a PCE computes a path with a loose hop expecting
     that the loose hop will be expanded across the lower-layer network,
     the LSR (that is an entry point to the lower-layer network) may
     simply expand the loose hop in the same layer. If more strict
     control of how the LSR establishes the path is required, mechanisms
     such as Path Key [RFC5520] could be applied.

   - Option 2: Multi-Layer Path

     The PCE computes a "multi-layer" path, i.e., a path that includes
     TE links from distinct layers [RFC4206]. Such a path can include
     the complete path of one or more lower-layer LSPs that already
     exist or that are not yet established. In the latter case, the
     signaling of the higher-layer LSP will trigger the establishment of
     the lower-layer LSPs.

3. Inter-Layer Path Computation Models

   In Section 2, three models are defined to perform PCE-based inter-
   layer path computation -- namely, single PCE computation, multiple
   PCE computation with inter-PCE communication, and multiple PCE
   computation without inter-PCE communication. Single PCE computation
   is discussed in Section 3.1 below, and multiple PCE computation (with
   and without inter-PCE communication) is discussed in Section 3.2
   below.

3.1. Single PCE Inter-Layer Path Computation

   In this model, inter-layer path computation is performed by a single
   PCE that has topology visibility into all layers. Such a PCE is
   called a multi-layer PCE.

   In Figure 2, the network is comprised of two layers. LSRs H1, H2,
   H3, and H4 belong to the higher layer, and LSRs H2, H3, L1, and L2
   belong to the lower layer. The PCE is a multi-layer PCE that has
   visibility into both layers. It can perform end-to-end path
   computation across layers (single PCE path computation). For
   instance, it can compute an optimal path H1-H2-L1-L2-H3-H4 for a
   higher-layer LSP from H1 to H4. This path includes the path of a
   lower-layer LSP from H2 to H3 that is already in existence or not yet
   established.

```
                            -----
                           | PCE |
                            -----
     -----    -----                    -----    -----
    | LSR |--| LSR |................| LSR |--| LSR |
    | H1  |  | H2  |                | H3  |  | H4  |
     -----    -----\              /-----    -----
              \-----    -----/
             | LSR |--| LSR |
             | L1  |  | L2  |
              -----    -----
```
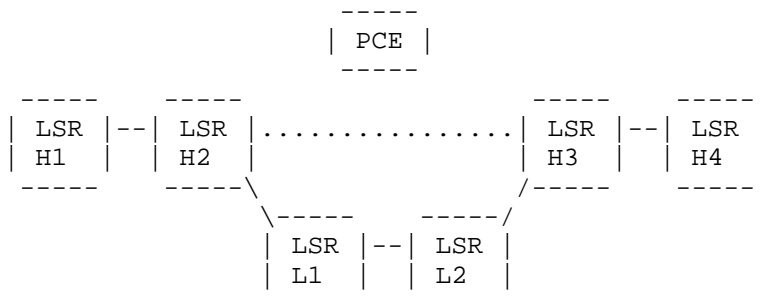
         Figure 2: Single PCE Inter-Layer Path Computation

3.2. Multiple PCE Inter-Layer Path Computation

   In this model, there is at least one PCE per layer, and each PCE has
   topology visibility restricted to its own layer. Some providers may
   want to keep the layer boundaries due to factors such as
   organizational and/or service management issues. The choice for
   multiple PCE computation instead of single PCE computation may also

be driven by scalability considerations, as in this mode a PCE only
needs to maintain topology information for one layer (resulting in a
size reduction for the Traffic Engineering Database (TED)).

These PCEs are called mono-layer PCEs. Mono-layer PCEs collaborate
to compute an end-to-end optimal path across layers.

Figure 3 shows multiple PCE inter-layer computation with inter-PCE
communication. There is one PCE in each layer. The PCEs from each
layer collaborate to compute an end-to-end path across layers. PCE
Hi is responsible for computations in the higher layer and may
"consult" with PCE Lo to compute paths across the lower layer. PCE
Lo is responsible for path computation in the lower layer. A simple
example of cooperation between the PCEs could be as follows:

- LSR H1 sends a request to PCE Hi for a path H1-H4.

- PCE Hi selects H2 as the entry point to the lower layer and H3 as
  the exit point.

- PCE Hi requests a path H2-H3 from PCE Lo.

- PCE Lo returns H2-L1-L2-H3 to PCE Hi.

- PCE Hi is now able to compute the full path (H1-H2-L1-L2-H3-H4) and
  return it to H1.

Of course, more complex cooperation may be required if an optimal
end-to-end path is desired.

```
                             -----
                            | PCE |
                            | Hi  |
                            --+--
                               |
   -----   -----               |               -----   -----
  | LSR |--| LSR |.............|...........| LSR |--| LSR |
  | H1  |  | H2  |             |             | H3  |  | H4  |
   -----   -----\            --+--          /-----   -----
                 \          | PCE |        /
                  \         | Lo  |       /
                   \         -----       /
                    \                   /
                     \-----     -----/
                      | LSR |--| LSR |
                      | L1  |  | L2  |
                       -----   -----
```
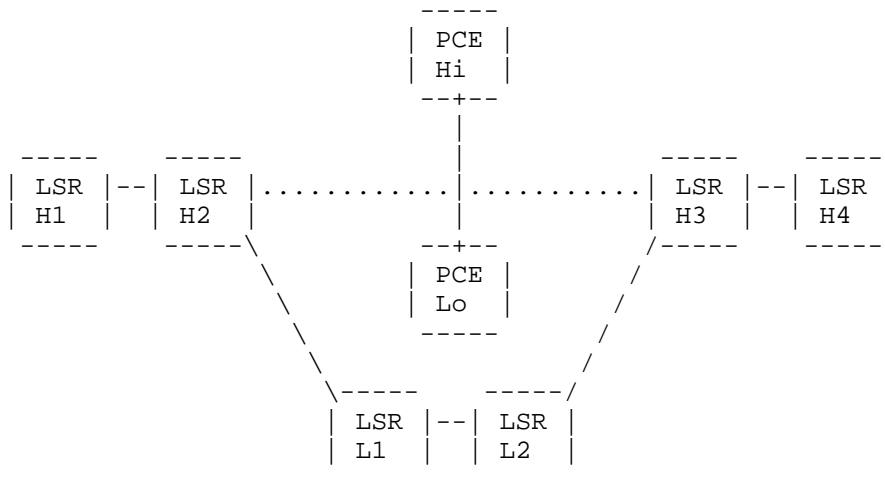
                 Figure 3: Multiple PCE Inter-Layer Path Computation
                         with Inter-PCE Communication

   Figure 4 shows multiple PCE inter-layer path computation without
   inter-PCE communication. As described in Section 2, separate path
   computations are performed on behalf of the TE-LSP head-end and each
   transit border LSR that is the entry point to a new layer.

```
                             -----
                            | PCE |
                            | Hi  |
                             -----
   -----   -----                                 -----   -----
  | LSR |--| LSR |.........................| LSR |--| LSR |
  | H1  |  | H2  |                             | H3  |  | H4  |
   -----   -----\            -----          /-----   -----
                 \          | PCE |        /
                  \         | Lo  |       /
                   \         -----       /
                    \                   /
                     \-----     -----/
                      | LSR |--| LSR |
                      | L1  |  | L2  |
                       -----   -----
```
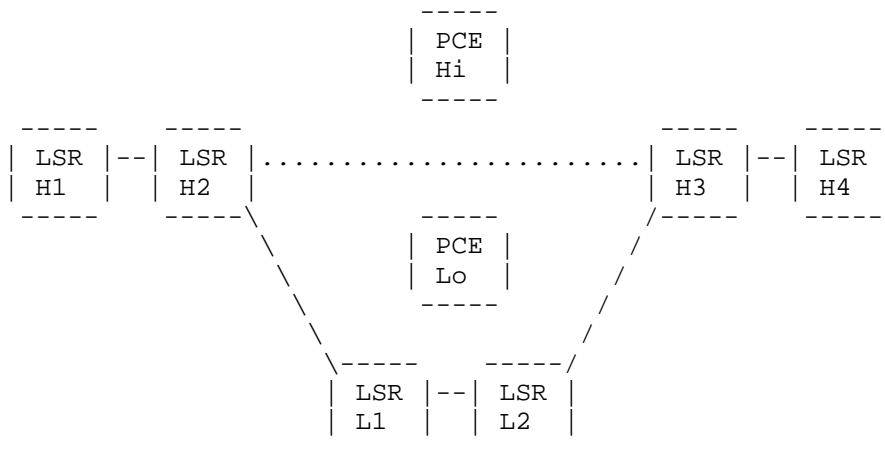
                 Figure 4: Multiple PCE Inter-Layer Path Computation
                        without Inter-PCE Communication

3.3. General Observations

   - Depending on implementation details, the time to perform inter-
     layer path computation in the single PCE inter-layer path
     computation model may be less than that of the multiple PCE model
     with cooperating mono-layer PCEs, because there is no requirement
     to exchange messages between cooperating PCEs.

   - When TE topology for all layer networks is visible within one
     routing domain, the single PCE inter-layer path computation model
     may be adopted because a PCE is able to collect all layers' TE
     topologies by participating in only one routing domain.

   - As the single PCE inter-layer path computation model uses more TE
     topology information in one computation than is used by PCEs in the
     multiple PCE path computation model, it requires more computation
     power and memory.

   When there are multiple candidate layer border nodes (we may say that
   the higher layer is multi-homed), optimal path computation requires
   that all the possible paths transiting different layer border nodes
   or links be examined. This is relatively simple in the single PCE
   inter-layer path computation model because the PCE has full
   visibility -- the computation is similar to the computation within a
   single domain of a single layer. In the multiple PCE inter-layer
   path computation model, backward-recursive techniques described in
   [RFC5441] could be used by considering layers as separate domains.

4. Inter-Layer Path Control

4.1. VNT Management

   As a result of mono-layer path computation, a PCE may determine that
   there is insufficient bandwidth available in the higher-layer network
   to support this or future higher-layer LSPs. The problem might be
   resolved if new LSPs are provisioned across the lower-layer network.
   Furthermore, the modification, re-organization, and new provisioning
   of lower-layer LSPs may enable better utilization of lower-layer
   network resources, given the demands of the higher-layer network. In
   other words, the VNT needs to be controlled or managed in cooperation
   with inter-layer path computation.

   A VNT Manager (VNTM) is defined as a functional element that manages
   and controls the VNT. The PCE and VNT Manager are distinct
   functional elements that may or may not be collocated.

4.2. Inter-Layer Path Control Models

4.2.1. PCE-VNTM Cooperation Model

```
        -----        ------
       | PCE |--->| VNTM |
        -----        ------
         ^             :
         :             :
         :             :
         v             V
        -----        -----                   -----        -----
       | LSR |----| LSR |................| LSR |----| LSR |
       | H1  |    | H2  |                | H3  |    | H4  |
        -----      -----\                /-----      -----
                         \-----    -----/
                        | LSR |--| LSR |
                        | L1  |  | L2  |
                         -----    -----
```
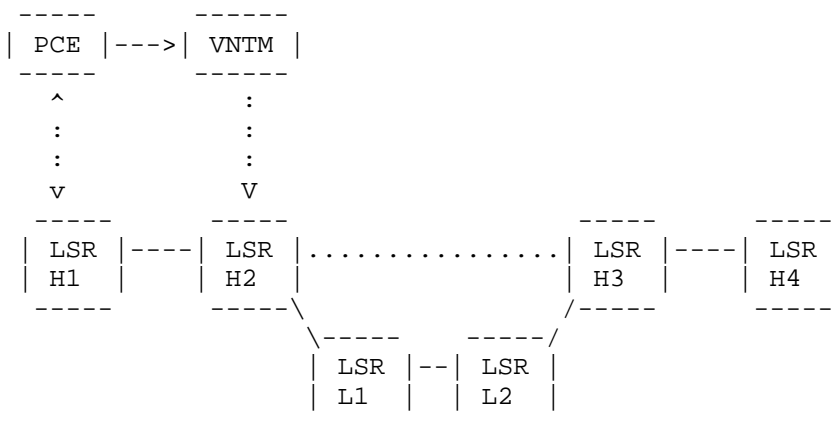
                   Figure 5: PCE-VNTM Cooperation Model

   A multi-layer network consists of higher-layer and lower-layer
   networks. LSRs H1, H2, H3, and H4 belong to the higher-layer
   network, and LSRs H2, L1, L2, and H3 belong to the lower-layer
   network, as shown in Figure 5. The case of single PCE inter-layer
   path computation is considered here to explain the cooperation model
   between PCE and VNTM, but multiple PCE path computation with or
   without inter-PCE communication can also be applied to this model.

   Consider that H1 requests the PCE to compute an inter-layer path
   between H1 and H4. There is no TE link in the higher layer between
   H2 and H3 before the path computation request, so the request fails.
   But the PCE may provide information to the VNT Manager responsible
   for the lower-layer network that may help resolve the situation for
   future higher-layer LSP setup.

   The roles of PCE and VNTM are as follows. PCE performs inter-layer
   path computation and is unable to supply a path because there is no
   TE link between H2 and H3. The computation fails, but PCE suggests
   to VNTM that a lower-layer LSP (H2-H3) could be established to
   support future LSP requests. Messages from PCE to VNTM contain
   information about the higher-layer demand (from H2 to H3), and may
   include a suggested path in the lower layer (if the PCE has
   visibility into the lower-layer network). VNTM uses local policy and
   possibly management/configuration input to determine how to process
   the suggestion from PCE, and may request an ingress LSR (e.g., H2) to

establish a lower-layer LSP. VNTM or the ingress LSR (H2) may
themselves use a PCE with visibility into the lower layer to compute
the path of this new LSP.

When the higher-layer PCE fails to compute a path and notifies VNTM,
it may wait for the lower-layer LSP to be set up and advertised as a
TE link. PCE may have a timer. After TED is updated within a
specified duration, PCE will know a new TE link. It could then
compute the complete end-to-end path for the higher-layer LSP and
return the result to the PCC. In this case, the PCC may be kept
waiting for some time, and it is important that the PCC understands
this. It is also important that the PCE and VNTM have an agreement
that the lower-layer LSP will be set up in a timely manner, or that
the PCE will be notified by the VNTM that no new LSP will become
available. In any case, if the PCE decides to wait, it must operate
a timeout. An example of such a cooperative procedure between PCE
and VNTM is as follows, using the example network in Figure 4.

   Step 1: H1 (PCC) requests PCE to compute a path between H1 and H4.

   Step 2: The path computation fails because there is no TE link
            across the lower-layer network.

   Step 3: PCE suggests to VNTM that a new TE link connecting H2 and
            H3 would be useful. The PCE notifies VNTM that it will be
            waiting for the TE link to be created. VNTM considers
            whether lower-layer LSPs should be established, if
            necessary and acceptable within VNTM's policy constraints.

   Step 4: VNTM requests an ingress LSR in the lower-layer network
            (e.g., H2) to establish a lower-layer LSP. The request
            message may include a lower-layer LSP route obtained from
            the PCE responsible for the lower-layer network.

   Step 5: The ingress LSR signals to establish the lower-layer LSP.

   Step 6: If the lower-layer LSP setup is successful, the ingress
            LSR notifies VNTM that the LSP is complete and supplies
            the tunnel information.

   Step 7: The ingress LSR (H2) advertises the new LSP as a TE link
            in the higher-layer network routing instance.

   Step 8: PCE notices the new TE link advertisement and recomputes
            the requested path.

      Step 9: PCE replies to H1 (PCC) with a computed higher-layer LSP
               route. The computed path is categorized as a mono-layer
               path that includes the already-established lower-layer LSP
               as a single hop in the higher layer. The higher-layer
               route is specified as H1-H2-H3-H4, where all hops are
               strict.

      Step 10: H1 initiates signaling with the computed path H2-H3-H4 to
               establish the higher-layer LSP.

4.2.2. Higher-Layer Signaling Trigger Model

```
        -----
       | PCE |
        -----
         ^
         :
         :
         v
        -----       -----                         -----     -----
       | LSR |----| LSR |................| LSR |--| LSR |
       | H1  |    | H2  |                | H3  |  | H4  |
        -----      -----\               /-----    -----
                        \-----    -----/
                        | LSR |--| LSR |
                        | L1  |  | L2  |
                         -----    -----
```
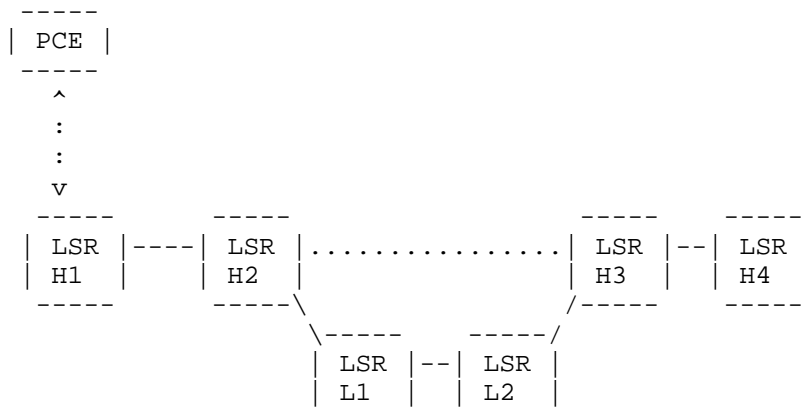
              Figure 6: Higher-Layer Signaling Trigger Model

   Figure 6 shows the higher-layer signaling trigger model. The case of
   single PCE path computation is considered to explain the higher-
   layer signaling trigger model here, but multiple PCE path computation
   with/without inter-PCE communication can also be applied to this
   model.

   As in the case described in Section 4.2.1, consider that H1 requests
   PCE to compute a path between H1 and H4. There is no TE link in the
   higher layer between H2 and H3 before the path computation request.

   PCE is unable to compute a mono-layer path, but may judge that the
   establishment of a lower-layer LSP between H2 and H3 would provide
   adequate connectivity. If the PCE has inter-layer visibility, it may
   return a path that includes hops in the lower layer (H1-H2-L1-L2-H3-
   H4), but if it has no visibility into the lower layer, it may return
   a path with a loose hop from H2 to H3 (H1-H2-H3(loose)-H4). The
   former is a multi-layer path, and the latter a mono-layer path that
   includes loose hops.

In the higher-layer signaling trigger model with a multi-layer path,
the LSP route supplied by the PCE includes the route of a lower-
layer LSP that is not yet established. A border LSR that is located
at the boundary between the higher-layer and lower-layer networks (H2
in this example) receives a higher-layer signaling message, notices
that the next hop is in the lower-layer network, and starts to set up
the lower-layer LSP as described in [RFC4206]. Note that these
actions depend on a policy being applied at the border LSR. An
example procedure of the signaling trigger model with a multi-layer
path is as follows.

   Step 1: H1 (PCC) requests PCE to compute a path between H1 and H4.
            The request indicates that inter-layer path computation is
            allowed.

   Step 2: As a result of the inter-layer path computation, PCE
            judges that a new lower-layer LSP needs to be established.

   Step 3: PCE replies to H1 (PCC) with a computed multi-layer route
            including higher-layer and lower-layer LSP routes. The
            route may be specified as H1-H2-L1-L2-H3-H4, where all
            hops are strict.

   Step 4: H1 initiates higher-layer signaling using the computed
            explicit router of H2-L1-L2-H3-H4.

   Step 5: The border LSR (H2) that receives the higher-layer
            signaling message starts lower-layer signaling to
            establish a lower-layer LSP along the specified lower-
            layer route of H2-L1-L2-H3. That is, the border LSR
            recognizes the hops within the explicit route that apply
            to the lower-layer network, verifies with local policy
            that a new LSP is acceptable, and establishes the required
            lower-layer LSP. Note that it is possible that a suitable
            lower-layer LSP has already been established (or become
            available) between the time that the computation was
            performed and the moment when the higher-layer signaling
            message reached the border LSR. In this case, the border
            LSR may select such a lower-layer LSP without the need to
            signal a new LSP, provided that the lower-layer LSP
            satisfies the explicit route in the higher-layer signaling
            request.

   Step 6: After the lower-layer LSP is established, the higher-layer
            signaling continues along the specified higher-layer route
            of H2-H3-H4 using hierarchical signaling [RFC4206].

On the other hand, in the signaling trigger model with a mono-layer
path, a higher-layer LSP route includes a loose hop to traverse the
lower-layer network between the two border LSRs. A border LSR that
receives a higher-layer signaling message needs to determine a path
for a new lower-layer LSP. It applies local policy to verify that a
new LSP is acceptable and then either consults a PCE with
responsibility for the lower-layer network or computes the path by
itself, and initiates signaling to establish the lower-layer LSP.
Again, it is possible that a suitable lower-layer LSP has already
been established (or become available). In this case, the border LSR
may select such a lower-layer LSP without the need to signal a new
LSP, provided that the existing lower-layer LSP satisfies the
explicit route in the higher-layer signaling request. Since the
higher-layer signaling request used a loose hop without specifying
any specifics of the path within the lower-layer network, the border
LSR has greater freedom to choose a lower-layer LSP than in the
previous example.

The difference between procedures of the signaling trigger model with
a multi-layer path and a mono-layer path is Step 5. Step 5 of the
signaling trigger model with a mono-layer path is as follows:

   Step 5': The border LSR (H2) that receives the higher-layer
            signaling message applies local policy to verify that a
            new LSP is acceptable and then initiates establishment of
            a lower-layer LSP. It either consults a PCE with
            responsibility for the lower-layer network or computes the
            route by itself to expand the loose hop route in the
            higher-layer path.

Finally, note that a virtual TE link may have been advertised into
the higher-layer network. This causes the PCE to return a path H1-
H2-H3-H4, where all the hops are strict. But when the higher-layer
signaling message reaches the layer border node H2 (that was
responsible for advertising the virtual TE link), it realizes that
the TE link does not exist yet, and signals the necessary LSP across
the lower-layer network using its own path determination (just as for
a loose hop in the higher layer) before continuing with the higher-
layer signaling.

```
    PCE
     ^
     :
     :
     V
    H1--H2                  H3--H4
       \                   /
         L1==L2==L3--L4--L5
                    |
                    |
                 L6--L7
                      \
                       H5--H6
```
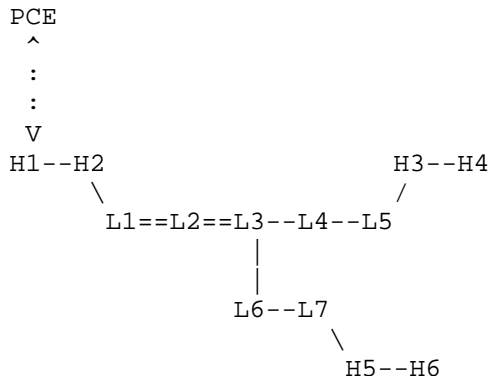
                Figure 7: Example of a Multi-Layer Network

Examples of multi-layer EROs are explained using Figure 7, which
shows how lower-layer LSP setup is performed in the higher-layer
signaling trigger model using an ERO that can include subobjects in
both the higher and lower layers. The higher-layer signaling trigger
model provides several options for the ERO when it reaches the last
LSR in the higher layer higher-layer network (H2).

1. The next subobject is a loose hop to H3 (mono-layer ERO).

2. The next subobject is a strict hop to L1, followed by a loose hop
   to H3.

3. The next subobjects are a series of hops (strict or loose) in the
   lower-layer network, followed by H3. For example, {L1(strict),
   L3(loose), L5(loose), H3(strict)}.

In the first example, the lower layer can utilize any LSP tunnel that
will deliver the end-to-end LSP to H3. In the third case, the lower
layer must select an LSP tunnel that traverses L3 and L5. However,
this does not mean that the lower layer can or should use an LSP from
L1 to L3 and another from L3 to L5.

4.2.3. NMS-VNTM Cooperation Model

In this model, NMS and VNTM cooperate to establish a lower-layer LSP.
There are two flavors in this model. One is where interaction
between layers in path computation is performed at the PCE level.
This is called "integrated flavor". The other is where interaction
between layers in path computation is achieved through NMS and VNTM
cooperation, which could be a point of application of administrative,
billing, and security policy. This is called "separated flavor".

   o NMS-VNTM Cooperation Model (integrated flavor)

```
     ------       -----
    | NMS   |<-->| PCE |
    |       |     -----
    | ----  |
    ||VNTM| |
    | ----  |
     ------
      :  :
      :   ---------
      :            :
      V            V
     -----       -----                     -----       -----
    | LSR  |----| LSR  |................. | LSR  |----| LSR  |
    | H1   |    | H2   |                  | H3   |    | H4   |
     -----       -----\                  /-----       -----
                   \-----    -----/
                   | LSR  |--| LSR  |
                   | L1   |  | L2   |
                    -----    -----
```
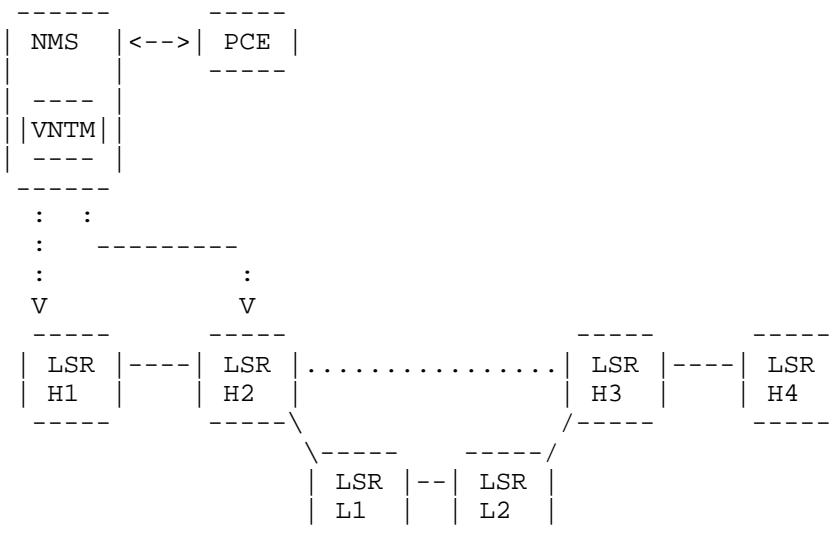
            Figure 8: NMS-VNTM Cooperation Model (integrated flavor)

   Figure 8 shows the NMS-VNTM cooperation model (integrated flavor).
   The case of single PCE path computation is considered to explain the
   NMS-VNTM cooperation model (integrated flavor) here, but multiple PCE
   path computation with inter-PCE communication can also be applied to
   this model. Note that multiple PCE path computation without inter-
   PCE communication does not fit in with this model. For this model to
   have meaning, the VNTM and NMS are closely coupled.

   The NMS sends the path computation request to the PCE. The PCE
   returns the inter-layer path computation result. When the NMS
   receives the path computation result, the NMS works with the VNTM and
   sends the request to LSR H2 to set up the lower-layer LSP. VNTM uses
   local policy and possibly management/configuration input to determine
   how to process the computation result from PCE.

   An example procedure of the NMS-VNTM cooperation model (integrated
   flavor) is as follows.

      Step 1: NMS requests PCE to compute a path between H1 and H4. The
               request indicates that inter-layer path computation is
               allowed.

      Step 2: PCE computes a path. The result (H1-H2-L1-L2-H3-H4) is
               sent back to the NMS.

Step 3: NMS discovers that a lower-layer LSP is needed. NMS works
         with VNTM to determine whether the new TE LSP H2-L1-L2-H3
         is permitted according to policy, etc.

Step 4: VNTM requests the ingress LSR in the lower-layer network
         (H2) to establish a lower-layer LSP. The request message
         includes the lower-layer LSP route obtained from PCE.

Step 5: H2 signals to establish the lower-layer LSP.

Step 6: If the lower-layer LSP setup is successful, H2 notifies
         VNTM that the LSP is complete and supplies the tunnel
         information.

Step 7: H2 advertises the new LSP as a TE link in the higher-layer
         network routing instance.

Step 8: VNTM notifies NMS that the underlying lower-layer LSP has
         been set up, and NMS notices the new TE link
         advertisement.

Step 9: NMS requests H1 to set up a higher-layer LSP between H1
         and H4 with the path computed in Step 2. The lower-layer
         links are replaced by the corresponding higher-layer TE
         link. Hence, the NMS sends the path H1-H2-H3-H4 to H1.

Step 10: H1 initiates signaling with the path H2-H3-H4 to establish
         the higher-layer LSP.

   o NMS-VNTM Cooperation Model (separate flavor)

```
         -----
        | NMS |
        |     |    -----
         -----    | PCE |
         ^   ^    | Hi  |
         :   :     -----
         :   :      ^
         :   :      :
         :   :      :
         :   v      v
         :  ------   -----                              -----   ------
         : | LSR  |--| LSR |........................| LSR  |--| LSR  |
         : | H1   |  | H2  |                        | H3   |  | H4   |
         :  ------   -----\                          /-----   ------
         :            ^    \                        /
         :            :     \                      /
         :   --------        \                    /
         v   :     -----      \-----   -----/
        ------   -----       \-----   -----/
       | VNTM |<-->| PCE |       | LSR |--| LSR |
       |      |    | Lo  |       | L1  |  | L2  |
        ------      -----         -----   -----
```
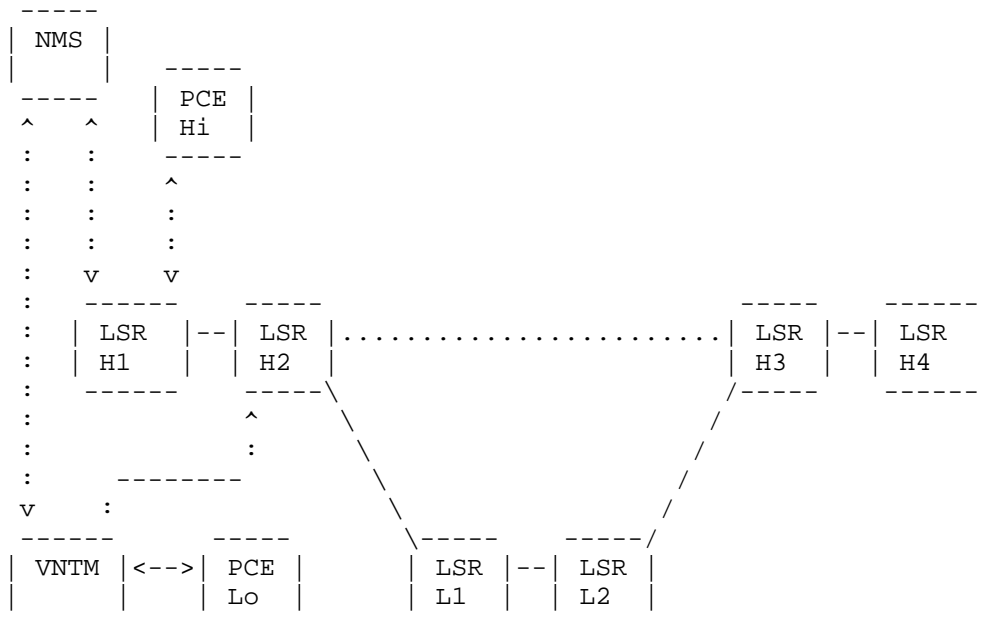
          Figure 9: NMS-VNTM Cooperation Model (separate flavor)

   Figure 9 shows the NMS-VNTM cooperation model (separate flavor). The
   NMS manages the higher layer. The case of multiple PCE computation
   without inter-PCE communication is used to explain the NMS-VNTM
   cooperation model here, but single PCE path computation could also be
   applied to this model. Note that multiple PCE path computation with
   inter-PCE communication does not fit in with this model.

   The NMS requests a head-end LSR (H1 in this example) to set up a
   higher-layer LSP between head-end and tail-end LSRs without
   specifying any route. The head-end LSR, which is a PCC, requests the
   higher-layer PCE to compute a path between head-end and tail-end
   LSRs. There is no TE link in the higher-layer between border LSRs
   (H2 and H3 in this example). When the PCE fails to compute a path,
   it informs the PCC (i.e., head-end LSR), which notifies the NMS. The
   notification may include information about the reason for failure
   (such as that there is no TE link between the border LSRs or that
   computation constraints cannot be met).

Note that it is equally valid for the higher-layer PCE to be
consulted by the NMS rather than by the head-end LSR. In this case,
the result is the same -- the NMS discovers that an end-to-end LSP
cannot be provisioned owing to the lack of a TE link between H2 and
H3.

The NMS may now suggest (or request) to the VNTM that a lower-layer
LSP between the border LSRs be established and be advertised as a TE
link in the higher layer to support future higher-layer LSP requests.
The communication between the NMS and the VNTM may be performed in an
automatic manner or in a manual manner, and is a key interaction
between layers that may also be separate administrative domains.
Thus, this communication is potentially a point of application of
administrative, billing, and security policy. The NMS may wait for
the lower-layer LSP to be set up and advertised as a TE link, or it
may reject the operator's request for the service that requires the
higher-layer LSP with a suggestion that the operator try again later.

The VNTM requests the lower-layer PCE to compute a path, and then
requests H2 to establish a lower-layer LSP. Alternatively, the VNTM
may make a direct request to H2 for the LSP, and H2 may consult the
lower-layer PCE. After the NMS is informed or notices that the
lower-layer LSP has been established, it can request the head-end LSR
(H1) to set up the higher-layer end-to-end LSP between H1 and H4.

Thus, cooperation between the higher layer and lower layer is
performed though communication between NMS and VNTM. An example of
such a procedure of the NSM-VNTM cooperation model is as follows,
using the example network in Figure 6.

   Step 1: NMS requests a head-end LSR (H1) to set up a higher-layer
            LSP between H1 and H4 without specifying any route.

   Step 2: H1 (PCC) requests PCE to compute a path between H2 and H3.

   Step 3: The path computation fails because there is no TE link
            across the lower-layer network.

   Step 4: H1 (PCC) notifies NMS. The notification may include an
            indication that there is no TE link between H2 and H4.

   Step 5: NMS suggests (or requests) to VNTM that a new TE link
            connecting H2 and H3 would be useful. The NMS notifies
            VNTM that it will be waiting for the TE link to be
            created. VNTM considers whether lower-layer LSPs should
            be established, if necessary and acceptable within VNTM's
            policy constraints.

      Step 6: VNTM requests the lower-layer PCE for path computation.

      Step 7: VNTM requests the ingress LSR in the lower-layer network
               (H2) to establish a lower-layer LSP. The request message
               includes a lower-layer LSP route obtained from the lower-
               layer PCE responsible for the lower-layer network.

      Step 8: H2 signals the lower-layer LSP.

      Step 9: If the lower-layer LSP setup is successful, H2 notifies
               VNTM that the LSP is complete and supplies the tunnel
               information.

      Step 10: H2 advertises the new LSP as a TE link in the higher-layer
               network routing instance.

      Step 11: VNTM notifies NMS that the underlying lower-layer LSP has
               been set up, and NMS notices the new TE link
               advertisement.

      Step 12: NMS again requests H1 to set up a higher-layer LSP between
               H1 and H4.

      Step 13: H1 requests the higher-layer PCE to compute a path and
               obtains a successful result that includes the higher-layer
               route that is specified as H1-H2-H3-H4, where all hops are
               strict.

      Step 14: H1 initiates signaling with the computed path H2-H3-H4 to
               establish the higher-layer LSP.

4.2.4. Possible Combinations of Inter-Layer Path Computation and
        Inter-Layer Path Control Models

   Table 1 summarizes the possible combinations of inter-layer path
   computation and inter-layer path control models. There are three
   inter-layer path computation models: the single PCE path computation
   model, the multiple PCE path computation with inter-PCE communication
   model, and the multiple PCE path computation without inter-PCE
   communication model. There are also four inter-layer path control
   models: the PCE-VNTM cooperation model, the higher-layer signaling
   trigger model, the NMS-VNTM cooperation model (integrated flavor),
   and the NMS-VNTM cooperation model (separate flavor). All the
   combinations between inter-layer path computation and path control
   models, except for the combination of the multiple PCE path
   computation with inter-layer PCE communication model and the NMS-
   VNTM cooperation model, are possible.

Table 1: Possible Combinations of Inter-Layer Path Computation
and Inter-Layer Path Control Models

Path computation \ Path control	Single PCE	Multiple PCE with inter-PCE	Multiple PCE w/o inter-PCE
PCE-VNTM cooperation	Yes	Yes	Yes
Higher-layer signaling trigger	Yes	Yes	Yes
NMS-VNTM cooperation (integrated flavor)	Yes	Yes	No
NMS-VNTM cooperation (separate flavor)	No*	No	Yes

   * Note that, in case of NSM-VNTM cooperation (separate flavor) and
     single PCE inter-layer path computation, the PCE function used by
     NMS and VNTM may be collocated, but it will operate on separate
     TEDs.

5. Choosing between Inter-Layer Path Control Models

   This section compares the PCE-VNTM cooperation model, the higher-
   layer signaling trigger model, and the NMS-VNTM cooperation model in
   terms of VNTM functions, border LSR functions, higher-layer signaling
   time, and complexity (in terms of number of states and messages). An
   appropriate model may be chosen by a network operator in different
   deployment scenarios, taking all these considerations into account.

5.1. VNTM Functions

   VNTM functions are required in both the PCE-VNTM cooperation model
   and the NMS-VNTM model. In the PCE-VNTM cooperation model,
   communications are required between PCE and VNTM and between VNTM and
   a border LSR. Communications between a higher-layer PCE and the VNTM
   are event notifications and may use Simple Network Management
   Protocol (SNMP) notifications from the PCE MIB modules [PCE-MIB].
   Note that communications from the PCE to the VNTM do not have any
   acknowledgements. VNTM-LSR communication can use existing GMPLS-TE
   MIB modules [RFC4802].

In the NMS-VNTM cooperation model, communications are required
between NMS and VNTM, between VNTM and a lower-layer PCE, and between
VNTM and a border LSR. NMS-VNTM communications, which are out of
scope of this document, may use proprietary or standard interfaces,
some of which, for example, are standardized in TM Forum.
Communications between VNTM and a lower-layer PCE use the Path
Computation Element Communication Protocol (PCEP) [RFC5440]. VNTM-
LSR communications are the same as in the PCE-VNTM cooperation model.

In the higher-layer signaling trigger model, no VNTM functions are
required, and no such communications are required.

If VNTM functions are not supported in a multi-layer network, the
higher-layer signaling trigger model has to be chosen.

The inclusion of VNTM functionality allows better coordination of
cross-network LSP tunnels and application of network-wide policy that
is far harder to apply in the trigger model since it requires the
coordination of policy between multiple border LSRs.

Also, VNTM functions could be applied to establish LSPs (or
connections) in non-MPLS/GMPLS networks, which do not have signaling
capabilities, by configuring each node along the path from the VNTM.

5.2. Border LSR Functions

In the higher-layer signaling trigger model, a border LSR must have
some additional functions. It needs to trigger lower-layer signaling
when a higher-layer Path message suggests that lower-layer LSP setup
is necessary. Note that, if virtual TE links are used, the border
LSRs must be capable of triggered signaling.

If the ERO in the higher-layer Path message uses a mono-layer path or
specifies a loose hop, the border LSR receiving the Path message must
obtain a lower-layer route either by consulting a PCE or by using its
own computation engine. If the ERO in the higher-layer Path message
uses a multi-layer path, the border LSR must judge whether lower-
layer signaling is needed.

In the PCE-VNTM and NMS-VNTM cooperation models, no additional
function for triggered signaling is required in border LSRs except
when virtual TE links are used. Therefore, if these additional
functions are not supported in border LSRs, where a border LSR is
controlled by VNTM to set up a lower-layer LSP, the cooperation model
has to be chosen.

5.3. Complete Inter-Layer LSP Setup Time

   The complete inter-layer LSP setup time includes inter-layer path
   computation, signaling, and the communication time between PCC and
   PCE, PCE and VNTM, NMS and VNTM, and VNTM and LSR. In the PCE-VNTM
   and the NMS-VNTM cooperation models, the additional communication
   steps are required compared with the higher-layer signaling trigger
   model. On the other hand, the cooperation model provides better
   control at the cost of a longer service setup time.

   Note that, in terms of higher-layer signaling time, in the higher-
   layer signaling trigger model, the required time from when higher-
   layer signaling starts to when it is completed is more than that of
   the cooperation model except when a virtual TE link is included.
   This is because the former model requires lower-layer signaling to
   take place during the higher-layer signaling. A higher-layer ingress
   LSR has to wait for more time until the higher-layer signaling is
   completed. A higher-layer ingress LSR is required to be tolerant of
   longer path setup times.

5.4. Network Complexity

   If the higher- and lower-layer networks have multiple interconnects,
   then optimal path computation for end-to-end LSPs that cross the
   layer boundaries is non-trivial. The higher-layer LSP must be routed
   to the correct layer border nodes to achieve optimality in both
   layers.

   Where the lower-layer LSPs are advertised into the higher-layer
   network as TE links, the computation can be resolved in the higher-
   layer network. Care needs to be taken in the allocation of TE
   metrics (i.e., costs) to the lower-layer LSPs as they are advertised
   as TE links into the higher-layer network, and this might be a
   function for a VNT Manager component. Similarly, attention should be
   given to the fact that the LSPs crossing the lower-layer network
   might share points of common failure (e.g., they might traverse the
   same link in the lower-layer network) and the shared risk link groups
   (SRLGs) for the TE links advertised in the higher-layer must be set
   accordingly.

   In the single PCE model, an end-to-end path can be found in a single
   computation because there is full visibility into both layers and all
   possible paths through all layer interconnects can be considered.

   Where PCEs cooperate to determine a path, an iterative computation
   model such as [RFC5441] can be used to select an optimal path across
   layers.

When non-cooperating mono-layer PCEs, each of which is in a separate
layer, are used with the triggered LSP model, it is not possible to
determine the best border LSRs, and connectivity cannot even be
guaranteed. In this case, crankback signaling techniques [RFC4920]
can be used to eventually achieve connectivity, but optimality is far
harder to achieve. In this model, a PCE that is requested by an
ingress LSR to compute a path expects a border LSR to set up a
lower-layer path triggered by high-layer signaling when there is no
TE link between border LSRs.

5.5. Separation of Layer Management

Many network operators may want to provide a clear separation between
the management of the different layer networks. In some cases, the
lower-layer network may come from a separate commercial arm of an
organization or from a different corporate body entirely. In these
cases, the policy applied to the establishment of LSPs in the lower-
layer network and to the advertisement of these LSPs as TE links in
the higher-layer network will reflect commercial agreements and
security concerns (see Section 8). Since the capacity of the LSPs in
the lower-layer network are likely to be significantly larger than
those in the client higher-layer network (multiplex-server model),
the administrator of the lower-layer network may want to exercise
caution before allowing a single small demand in the higher layer to
tie up valuable resources in the lower layer.

The necessary policy points for this separation of administration and
management are more easily achieved through the VNTM approach than by
using triggered signaling. In effect, the VNTM is the coordination
point for all lower-layer LSPs and can be closely tied to a human
operator as well as to policy and billing. Such a model can also be
achieved using triggered signaling.

6. Stability Considerations

Inter-layer traffic engineering needs to be managed and operated
correctly to avoid introducing instability problems.

Lower-layer LSPs are likely, by the nature of the technologies used
in layered networks, to be of considerably higher capacity than the
higher-layer LSPs. This has the benefit of allowing multiple higher-
layer LSPs to be carried across the lower-layer network in a single
lower-layer LSP. However, when a new lower-layer LSP is set up to
support a request for a higher-layer LSP because there is no suitable
route in the higher-layer network, it may be the case that a very
large LSP is established in support of a very small traffic demand.
Further, if the higher-layer LSP is short-lived, the requirement for
the lower-layer LSP will go away, either leaving it in place but

   unused or requiring it to be torn down. This may cause excessive
   tie-up of unused lower-layer network resources, or may introduce
   instability into the lower-layer network. It is important that
   appropriate policy controls or configuration features are available
   so that demand-led establishment of lower-layer LSPs (the so-called
   "bandwidth on demand") is filtered according to the requirements of
   the lower-layer network.

   When a higher-layer LSP is requested to be set up, a new lower-layer
   LSP may be established if there is no route with the requested
   bandwidth for the higher-layer LSP. After the lower-layer LSP is
   established, existing high-layer LSPs could be re-routed to use the
   newly established lower-layer LSP, if using the lower-layer LSP
   provides a better route than that taken by the existing LSPs. This
   re-routing may result in lower utilization of other lower-layer LSPs
   that used to carry the existing higher-layer LSPs. When the
   utilization of a lower-layer LSP drops below a threshold (or drops to
   zero), the LSP is deleted according to lower-layer network policy.

   But consider that some other new higher-layer LSP may be requested at
   once, requiring the establishment or re-establishment of a lower-
   layer LSP. This, in turn, may cause higher-layer re-routing, making
   other lower-layer LSPs under-utilized in a cyclic manner. This
   behavior makes the higher-layer network unstable.

   Inter-layer traffic engineering needs to avoid network instability
   problems. To solve the problem, network operators may have some
   constraints achieved through configuration or policy, where inter-
   layer path control actions such as re-routing and deletion of lower-
   layer LSPs are not easily allowed. For example, threshold parameters
   for the actions are determined so that hysteresis control behavior
   can be performed.

7. Manageability Considerations

   Inter-layer MPLS or GMPLS traffic engineering must be considered in
   the light of administrative and management boundaries that are likely
   to coincide with the technology layer boundaries. That is, each
   layer network may possibly be under separate management control with
   different policies applied to the networks, and specific policy rules
   applied at the boundaries between the layers.

   Management mechanisms are required to make sure that inter-layer
   traffic engineering can be applied without violating the policy and
   administrative operational procedures used by the network operators.

7.1. Control of Function and Policy

7.1.1. Control of Inter-Layer Computation Function

   PCE implementations that are capable of supporting inter-layer
   computations should provide a configuration switch to allow support
   of inter-layer path computations to be enabled or disabled.

   When a PCE is capable of, and configured for, inter-layer path
   computation, it should advertise this capability as described in
   [PCC-PCE], but this advertisement may be suppressed through a
   secondary configuration option.

7.1.2. Control of Per-Layer Policy

   Where each layer is operated as a separate network, the operators
   must have control over the policies applicable to each network, and
   that control should be independent of the control of policies for
   other networks.

   Where multiple layers are operated as part of the same network, the
   operator may have a single point of control for an integrated policy
   across all layers, or may have control of separate policies for each
   layer.

7.1.3. Control of Inter-Layer Policy

   Probably the most important issue for inter-layer traffic engineering
   is inter-layer policy. This may cover issues such as under what
   circumstances a lower-layer LSP may be established to provide
   connectivity in the higher-layer network. Inter-layer policy may
   exist to protect the lower-layer (high capacity) network from very
   dynamic changes in micro-demand in the higher-layer network (see
   Section 6). It may also be used to ensure appropriate billing for
   the lower-layer LSPs.

   Inter-layer policy should include the definition of the points of
   connectivity between the network layers, the inter-layer TE model to
   be applied (for example, the selection between the models described
   in this document), and the rules for path computation and LSP setup.
   Where inter-layer policy is defined, it must be used consistently
   throughout the network, and should be made available to the PCEs that
   perform inter-layer computation so that appropriate paths are
   computed. Mechanisms for providing policy information to PCEs are
   discussed in [RFC5394].

VNTM may provide a suitable functional component for the
implementation of inter-layer policy. Use of VNTM allows the
administrator of the lower-layer network to apply inter-layer policy
without making that policy public to the operator of the higher-layer
network. Similarly, a cooperative PCE model (with or without inter-
PCE communication) allows separate application of policy during the
selection of paths.

7.2. Information and Data Models

Any protocol extensions to support inter-layer computations must be
accompanied by the definition of MIB objects for the control and
monitoring of the protocol extensions. These MIB object definitions
will conventionally be placed in a separate document from that which
defines the protocol extensions. The MIB objects may be provided in
the same MIB module as used for the management of the base protocol
that is being extended.

Note that inter-layer PCE functions should, themselves, be manageable
through MIB modules. In general, this means that the MIB modules for
managing PCEs should include objects that can be used to select and
report on the inter-layer behavior of each PCE. It may also be
appropriate to provide statistical information that reports on the
inter-layer PCE interactions.

Where there are communications between a PCE and VNTM, additional MIB
modules may be necessary to manage and model these communications.
On the other hand, if these communications are provided through MIB
notifications, then those notifications must form part of a MIB
module definition.

Policy Information Base (PIB) modules may also be appropriate to meet
the requirements as described in Section 7.1 and [RFC5394].

7.3. Liveness Detection and Monitoring

Liveness detection and monitoring is required between PCEs and PCCs,
and between cooperating PCEs as described in [RFC4657]. Inter-layer
traffic engineering does not change this requirement.

Where there are communications between a PCE and VNTM, additional
liveness detection and monitoring may be required to allow the PCE to
know whether the VNTM has received its information about failed path
computations and desired TE links.

When a lower-layer LSP fails (perhaps because of the failure of a
lower-layer network resource) or is torn down as a result of lower-
layer network policy, the consequent change should be reported to the

higher layer as a change in the VNT, although inter-layer policy may
dictate that such a change is hidden from the higher layer. The
higher-layer network may additionally operate data plane failure
techniques over the virtual TE links in the VNT in order to monitor
the liveness of the connections, but it should be noted that if the
virtual TE link is advertised but not yet established as an LSP in
the lower layer, such higher-layer Operations, Administration, and
Management (OAM) techniques will report a failure.

7.4. Verifying Correct Operation

   The correct operation of the PCE computations and interactions are
   described in [RFC4657], [RFC5440], etc., and does not need further
   discussion here.

   The correct operation of inter-layer traffic engineering may be
   measured in several ways. First, the failure rate of higher-layer
   path computations owing to an absence of connectivity across the
   lower layer may be observed as a measure of the effectiveness of the
   VNT and may be reported as part of the data model described in
   Section 7.2. Second, the rate of change of the VNT (i.e., the rate
   of establishment and removal of higher-layer TE links based on
   lower-layer LSPs) may be seen as a measure of the correct planning of
   the VNT and may also form part of the data model described in Section
   7.2. Third, network resource utilization in the lower layer (both in
   terms of resource congestion and in consideration of under-
   utilization of LSPs set up to support virtual TE links) can indicate
   whether effective inter-layer traffic engineering is being applied.

   Management tools in the higher-layer network should provide a view of
   which TE links are provided using planned lower-layer capacity (that
   is, physical connectivity or permanent connections) and which TE
   links are dynamic and achieved through inter-layer traffic
   engineering. Management tools in the lower layer should provide a
   view of the use to which lower-layer LSPs are put, including whether
   they have been set up to support TE links in a VNT and, if so, for
   which client network.

7.5. Requirements on Other Protocols and Functional Components

   There are no protocols or protocol extensions defined in this
   document, and so it is not appropriate to consider specific
   interactions with other protocols. It should be noted, however, that
   the objective of this document is to enable inter-layer traffic
   engineering for MPLS-TE and GMPLS networks, and so it is assumed that
   the necessary features for inter-layer operation of routing and
   signaling protocols are in existence or will be developed.

This document introduces roles for various network components (PCE,
LSR, NMS, and VNTM). Those components are all required to play their
part in order that inter-layer TE can be effective. That is, an
inter-layer TE model that assumes the presence and operation of any
of these functional components obviously depends on those components
to fulfill their roles as described in this document.

7.6. Impact on Network Operation

The use of a PCE to compute inter-layer paths is expected to have a
significant and beneficial impact on network operations. Inter-layer
traffic engineering of itself may provide additional flexibility to
the higher-layer network while allowing the lower-layer network to
support more and varied client networks in a more efficient way.
Traffic engineering across network layers allows optimal use to be
made of network resources in all layers.

The use of PCE as described in this document may also have a
beneficial effect on the loading of PCEs responsible for performing
inter-layer path computation while facilitating a more independent
operation model for the network layers.

8. Security Considerations

Inter-layer traffic engineering with PCE raises new security issues
in all three inter-layer path control models.

In the cooperation model between PCE and VNTM, when the PCE
determines that a new lower-layer LSP is desirable, communications
are needed between the PCE and VNTM and between the VNTM and a border
LSR. In this case, these communications should have security
mechanisms to ensure authenticity, privacy, and integrity of the
information exchanged. In particular, it is important to protect
against false triggers for LSP setup in the lower-layer network,
since such falsification could tie up lower-layer network resources
(achieving a denial-of-service attack on the lower-layer network and
on the higher-layer network that is attempting to use it) and could
result in incorrect billing for services provided by the lower-layer
network. Where the PCE MIB modules are used to provide the
notification exchanges between the higher-layer PCE and the VNTM,
SNMPv3 should be used to ensure adequate security. Additionally, the
VNTM should provide configurable or dynamic policy functions so that
the VNTM behavior upon receiving notification from a higher-layer PCE
can be controlled.

The main security concern in the higher-layer signaling trigger model
is related to confidentiality. The PCE may inform a higher-layer PCC
about a multi-layer path that includes an ERO in the lower-layer

network, but the PCC may not have TE topology visibility into the
lower-layer network and might not be trusted with this information.
A loose hop across the lower-layer network could be used, but this
decreases the benefit of multi-layer traffic engineering. A better
alternative may be to mask the lower-layer path using a path key
[RFC5520] that can be expanded within the lower-layer network.
Consideration must also be given to filtering the recorded path
information from the lower-layer -- see [RFC4208], for example.

Additionally, in the higher-layer signaling trigger model,
consideration must be given to the security of signaling at the
inter-layer interface, since the layers may belong to different
administrative or trust domains.

The NMS-VNTM cooperation model introduces communication between the
NMS and the VNTM. Both of these components belong to the management
plane, and such communication is out of scope for this PCE document.
Note that the NMS-VNTM cooperation model may be considered to address
many security and policy concerns because the control and decision-
making is placed within the sphere of influence of the operator in
contrast to the more dynamic mechanisms of the other models.
However, the security issues have simply moved, and will require
authentication of operators and of policy.

Security issues may also exist when a single PCE is granted full
visibility of TE information that applies to multiple layers. Any
access to the single PCE will immediately gain access to the topology
information for all network layers -- effectively, a single security
breach can expose information that requires multiple breaches in
other models.

Note that, as described in Section 6, inter-layer TE can cause
network stability issues, and this could be leveraged to attack
either the higher- or lower-layer network. Precautionary measures,
such as those described in Section 7.1.3, can be applied through
policy or configuration to dampen any network oscillations.

9. Acknowledgments

We would like to thank Kohei Shiomoto, Ichiro Inoue, Julien Meuric,
Jean-Francois Peltier, Young Lee, Ina Minei, Jean-Philippe Vasseur,
and Franz Rambach for their useful comments.

10. References

10.1. Normative Reference

   [RFC3031]   Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol
               Label Switching Architecture", RFC 3031, January 2001.

   [RFC3945]   Mannie, E., Ed., "Generalized Multi-Protocol Label
               Switching (GMPLS) Architecture", RFC 3945, October 2004.

   [RFC4206]   Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP)
               Hierarchy with Generalized Multi-Protocol Label Switching
               (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.

10.2. Informative Reference

   [PCE-MIB]   Stephan, E., "Definitions of Textual Conventions for Path
               Computation Element", Work in Progress, March 2009.

   [PCC-PCE]   Oki, E., Le Roux, JL., Kumaki, K., Farrel, A., and T.
               Takeda, "PCC-PCE Communication and PCE Discovery
               Requirements for Inter-Layer Traffic Engineering", Work in
               Progress, January 2009.

   [RFC4208]   Swallow, G., Drake, J., Ishimatsu, H., and Y. Rekhter,
               "Generalized Multiprotocol Label Switching (GMPLS) User-
               Network Interface (UNI): Resource ReserVation Protocol-
               Traffic Engineering (RSVP-TE) Support for the Overlay
               Model", RFC 4208, October 2005.

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

   [RFC4657]   Ash, J., Ed., and J. Le Roux, Ed., "Path Computation
               Element (PCE) Communication Protocol Generic
               Requirements", RFC 4657, September 2006.

   [RFC4802]   Nadeau, T., Ed., and A. Farrel, Ed., "Generalized
               Multiprotocol Label Switching (GMPLS) Traffic Engineering
               Management Information Base", RFC 4802, February 2007.

   [RFC4920]   Farrel, A., Ed., Satyanarayana, A., Iwata, A., Fujita, N.,
               and G. Ash, "Crankback Signaling Extensions for MPLS and
               GMPLS RSVP-TE", RFC 4920, July 2007.

   [RFC5212]   Shiomoto, K., Papadimitriou, D., Le Roux, JL., Vigoureux,
               M., and D. Brungard, "Requirements for GMPLS-Based Multi-
               Region and Multi-Layer Networks (MRN/MLN)", RFC 5212, July
               2008.

   [RFC5394]   Bryskin, I., Papadimitriou, D., Berger, L., and J. Ash,
               "Policy-Enabled Path Computation Framework", RFC 5394,
               December 2008.

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

   [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, April
               2009.

   [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,
               April 2009.

Authors' Addresses

    Eiji Oki
    University of Electro-Communications
    Tokyo
    Japan
    EMail: oki@ice.uec.ac.jp

    Tomonori Takeda
    NTT
    3-9-11 Midori-cho,
    Musashino-shi, Tokyo 180-8585, Japan
    EMail: takeda.tomonori@lab.ntt.co.jp

    Jean-Louis Le Roux
    France Telecom R&D,
    Av Pierre Marzin,
    22300 Lannion, France
    EMail: jeanlouis.leroux@orange-ftgroup.com

    Adrian Farrel
    Old Dog Consulting
    EMail: adrian@olddog.co.uk