Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: November 30, 2007

Y. Lee
Huawei
JL. Le Roux
France Telecom
D. King
Aria Networks
E. Oki
NTT

May 29, 2007

Path Computation Element Communication Protocol (PCECP) Requirements and Protocol Extensions In Support of Global Concurrent Optimization

draft-lee-pce-global-concurrent-optimization-04.txt

Status of this Memo

Copyright Notice

Abstract

The Path Computation Element (PCE) is a network component, application, or node that is capable of performing path computations at the request of Path Computation Clients (PCCs). The PCE is applied in Multiprotocol Label Switching Traffic Engineering (MPLS-TE) networks and in Generalized MPLS (GMPLS) networks to determine the routes of Label Switched Paths (LSPs) through the network. The Path Computation Element Communication Protocol (PCEP) is specified for communications between PCCs and PCEs, and between cooperating PCEs.

When computing or re-optimizing the routes of a set of LSPs through a network it may be advantageous to perform bulk path computations in order to avoid blocking problems and to achieve more optimal network-wide solutions. Such bulk optimization is termed Global Concurrent Optimization (GCO). A Global Concurrent Optimization is able to simultaneously consider the entire topology of the network and the complete set of existing LSPs, and their respective constraints, and look to optimize or re-optimize the entire network to satisfy all constraints for all LSPs. The Global Concurrent Optimization (GCO) application is primarily an NMS based solution.

This document provides application-specific requirements and the PCEP extensions in support of a global concurrent optimization (GCO) application.

Table of Contents

1. Terminology

The terminology explained herein complies with [RFC4655].

PCC: Path Computation Client: Any client application requesting a path computation to be performed by a Path Computation Element.

PCE: Path Computation Element: An entity (component, application or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints.

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

PCECP: The PCE Communication Protocol: PCECP is the generic abstract idea of a protocol that is used to communicate path computation requests a PCC to a PCE, and to return computed paths from the PCE to the PCC. The PCECP can also be used between cooperating PCEs.

PCEP: The PCE communication Protocol: PCEP is the actual protocol that implements the PCECP idea.

GCO: Global Concurrent Optimization: A concurrent path computation application where a set of TE paths are computed concurrently in order to efficiently utilize network resources. A GCO path computation is able to simultaneously consider the entire topology of the network and the complete set of existing LSPs, and their respective constraints, and look to optimize or re-optimize the entire network to satisfy all constraints for all LSPs. A GCO path computation can also provide an optimal way to migrate from an existing set of LSPs to a reoptimized set (Morphing Problem).

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119]. These terms are also used in the parts of this document that specify requirements for clarity of specification of those requirements.

2. Introduction

[RFC4655] defines the PCE based Architecture and explains how a Path Computation Element (PCE) may compute Label Switched Paths (LSP) in Multiprotocol Label Switching Traffic Engineering (MPLS-TE) and Generalized MPLS (GMPLS) networks at the request of PCCs. A PCC is shown to be any network component that makes such a request and may be for instance a Label Switching Router (LSR) or a Network Management System (NMS). The PCE, itself, is shown to be located anywhere within the network, and may be within an LSR, an NMS or Operational Support System (OSS), or may be an independent network server.

The PCECP is the communication protocol used between PCC and PCE, and may also be used between cooperating PCEs. [RFC4657] sets out the common protocol requirements for the PCECP. Additional application-specific requirements for PCECP are deferred to separate documents.

This document provides a set of PCECP extension requirements and solutions in support of concurrent path computation applications that may arise during network operations. A concurrent path computation is a path computation application where a set of TE paths are computed concurrently in order to efficiently utilize network resources. The computation method involved with a concurrent path computation is referred to as global concurrent optimization in this document. Appropriate computation algorithms to perform this type of optimization are out of the scope of this document.

The Global Concurrent Optimization (GCO) application is primarily an NMS based solution. Due to complex synchronization issues associated with GCO application, the management based PCE architecture defined in section 5.5 of [RFC4655] is considered as the most suitable usage to support GCO application. This does not automatically preclude other architectural alternatives to support GCO application, but they are not recommended. For instance, GCO may be enabled by distributed LSRs through complex synchronization. However, this approach may suffer from significant synchronization overhead between the PCE and each of the PCCs. It would likely affect the network stability and hence significantly diminish the benefits of deploying PCEs.

As new LSPs are added sequentially or removed from the network over time, the global network resources become fragmented and the network no longer provides the optimal use of the available capacity. A global concurrent path computation is able to simultaneously consider the entire topology of the network and the complete set of existing LSPs, and their respective constraints, and look to re-optimize the entire network to satisfy all constraints for all LSPs. Alternatively, the application may consider a subset of the LSPs

and/or a subset of the network topology.

The need for a global concurrent path computation may also arise when network operators need to establish a set of TE LSPs in their network planning process. It is also envisioned that network operators might require a global concurrent path computation in the event of catastrophic network failures, where a set of TE LSPs need to be optimally rerouted. The nature of this work does promote such systems for offline processing. Online application of this work should only be considered with proven empirical validation.

As the PCE is envisioned to provide solutions in all path computation matters, it is anticipated that the PCE would provide solutions for global concurrent path computation needs.

The main focus of this document is to highlight the PCC-PCE communication needs in support of a concurrent path computation application and to define protocol extensions to meet those needs.

The PCC-PCE requirements addressed herein are specific to the context where the PCE is a specialized PCE that is capable of solving global concurrent path computation applications. Discovery of such capabilities might be desirable and could be achieved through extensions to the PCE discovery mechanisms [RFC4674], but that is out of the scope of this document.

3. Applicability of Global Concurrent Optimization (GCO)

This section discusses the PCE architecture to which GCO is applied. It also discusses various application scenarios for which global concurrent path computation may be applied.

3.1. Application of the PCE Architecture

Figure 1 shows the PCE-based network architecture as defined in [RFC4655] to which GCO application is applied. It must be observed that the PCC is not necessarily an LSR [RFC4655]. The GCO application is primarily an NMS-based solution in which an NMS plays the function of the PCC. Although Figure 1 shows the PCE as remote from the NMS, it might be collocated with the NMS. Note that in the collocated case there is no need for a standard communication protocol, this can rely on internal APIs.

```
                             -----------
          Application       |   -----   |
            Request         |  | TED |  |
               |            |   -----   |
               v            |     |     |
         ------------- Request/ |     v     |
        |     PCC     | Response|   -----   |
        |    (NMS)    |<--------+> | PCE |  |
        |             |         |   -----   |
         -------------           -----------
     Service |
     Request |
             v
          ----------  Signaling   ----------
         | Head-End | Protocol   | Adjacent |
         |  Node    |<---------->|   Node   |
          ----------              ----------
```

Figure 1: PCE-Based Architecture for Global Concurrent Optimization

Upon receipt of an application request (e.g., a traffic demand matrix is provided to the NMS by the operator's network planning procedure), the NMS requests a global concurrent path computation from the PCE. The PCE then computes the requested paths concurrently applying some algorithms. When the requested path computation completes, the PCE sends the resulting paths back to the NMS. The NMS then supplies the

head-end LSRs with a fully computed explicit path for each TE LSP that needs to be established.

3.2. Re-optimization of Existing Networks

The need for global concurrent path computation may arise in existing networks. When an existing TE LSP network experiences sub-optimal use of its resources, the need for re-optimization or reconfiguration may arise. The scope of re-optimization and reconfiguration may vary depending on particular situations. The scope of re-optimization may be limited to bandwidth modification to an existing TE LSP. However, it could well be that a set of TE LSPs may need to be re-optimized concurrently. In an extreme case, the TE LSPs may need to be globally re-optimized.

In loaded networks, with large size LSPs, a sequential re-optimization may not produce substantial improvements in terms of overall network optimization. The potential for network-wide gains from reoptimization of LSPs sequentially is depependent upon the network usage and size of the LSPs being optimized. However, the key point remains: computing the reoptimized path of one LSP at a time with giving no consideration to the other LSPs in the network could result in sub-optimal use of network resources. This may be far more visible in an optical network with a low ratio of potential LSPs per link, and far less visible in packet networks with micro-flow LSPs.

With regards to applicability of GCO in the event of catastrophic failures, there may be a real benefit in computing the paths of the LSPs as a set rather than computing new paths from the head-end LSRs in a distributed manner. It is to be noted, however, that a centralized system will typically suffer from a slower response time than a distributed system.

3.2.1. Reconfiguration of the Virtual Network Topology (VNT)

Reconfiguration of the VNT [MLN-REQ] is a typical application scenario where global concurrent path computation may be applicable. Triggers for VNT reconfiguration, such as traffic demand changes, network failures, and topological configuration changes, may require a set of existing LSPs to be re-computed.

3.2.2. Traffic Migration

When migrating from one set of TE LSPs to a reoptimized set of TE LSPs it is important that the traffic be moved without causing disruption. Various techniques exist in MPLS and GMPLS, such as make-before-break [RFC3209], to establish the new LSPs before tearing down the old LSPs. When multiple LSP routes are changed according to

the computed results, some of the LSPs may be disrupted due to the resource constraints. In other words, it may prove to be impossible to perform a direct migration from the old LSPs to the new optimal LSPs without disrupting traffic because there are insufficient network resources to support both sets of LSPs when make-before-break is used. However, the PCE may be able to determine an order of LSP rerouting actions so that make-before-break can be performed within the limited resources.

It may be the case that the reoptimization is radical. This could mean that it is not possible to apply make-before-break in any order to migrate from the old LSPs to the new LSPs. In this case a migration strategy is required that may necessitate LSPs being rerouted using make-before-break onto temporary paths in order to make space for the full reoptimization. A PCE might indicate the order in which reoptimized LSPs must be established and take over from the old LSPs, and may indicate a series of different temporary paths that must be used. Alternatively, the PCE might perform the global reoptimization as a series of sub-reoptimizations by reoptimizing subsets of the total set of LSPs.

The benefit of this multi-step rerouting includes minimization of traffic discruption and optimization gain. However this approach may imply some transient packets desequencing, jitter as well as control plane stress.

Note also that during reoptimization, traffic disruption may be allowed for some LSPs carrying low priority services (e.g., Internet traffic) and not allowed for some LSPs carrying mission critical services (e.g., voice traffic).

3.3. Greenfield Optimization

Greenfield optimization is a special case of GCO application when there is no LSP setup. Once the LSPs are setup, it is not a greenfield. The need for greenfield arises when network planner will want to make use of computation servers to plan the LSPs that will be provisioned in the network.

When a new TE network needs to be provisioned from a green-field perspective, a set of TE LSPs need to be created based on traffic demand, network topology, service constraints, and network resources. Under this scenario, concurrent computation ability is highly desirable, or required, to utilize network resources in an optimal manner and avoid blocking risks.

3.3.1. Single-layer Traffic Engineering

Greenfield optimization can be applied when layer-specific TE LSPs need to be created from a green-field perspective. For example, MPLS-TE network can be established based on layer 3 specific traffic demand, network topology, and network resources. Greenfield optimization for single-layer traffic engineering can be applied to optical transport networks such as SDH/Sonet, Ethernet Transport, WDM, etc.

3.3.2. Multi-layer Traffic Engineering

Greenfield optimization is not limited to single-layer traffic engineering. It can also be applied to multi-layer traffic engineering. Both the client and the server layers network resources and topology can be considered simultaneously in setting up a set of TE LSPs that traverse the layer boundary.

4. PCECP Requirements

This section provides the PCECP requirements to support global concurrent path computation applications. The requirements specified here should be regarded as application-specific requirements and are justifiable based on the extensibility clause found in section 6.1.14 of [RFC4657]:

> The PCECP MUST support the requirements specified in the application-specific requirements documents. The PCECP MUST also allow extensions as more PCE applications will be introduced in the future.

It is also to be noted that some of the requirements discussed in this section have already been discussed in the PCECP requirement document [RFC4657]. For example, Section 5.1.16 in [RFC4657] provides a list of generic constraints while Section 5.1.17 in [RFC4657] provides a list of generic objective functions that MUST be supported by the PCECP. While using such generic requirements as the baseline, this section provides application-specific requirements in the context of global concurrent path computation and in a more detailed level than the generic requirements.

The PCEP SHOULD support the following capabilities either via creation of new objects and/or modification of existing objects where applicable.

o An indicator to convey that the request is for a global concurrent path computation. This indicator is necessary to ensure consistency in applying global objectives and global constraints in all path computations. Note: This requirement is covered by "synchronized path computation" in [RFC4655] and [RFC4657]. However, an explicit indicator to request a global concurrent optimization is a new requirement.

o A Global Objective Function (GOF) field in which to specify the global objective function. The global objective function is the overarching objective function to which all individual path computation requests are subjected in order to find a globally optimal solution. Note that this requirement is covered by "synchronized objective functions" in section 5.1.7 [RFC4657]. A list of available global objective functions SHOULD include the following objective functions at the minimum and SHOULD be expandable for future addition:

  * Minimize the sum of all TE LSP costs (min cost)

   * Evenly allocate the network load to achieve the most uniform link utilization across all links (this can be achieved by the following objective function: minimize max over all links {(C(i)-A(i))/C(i)} where C(i) is the link capacity for link i and A(i) is the total bandwidth allocated on link i.

- o A Global Constraints (GC) field in which to specify the list of global constraints to which all the requested path computations should be subjected. This list SHOULD include the following constraints at the minimum and SHOULD be expandable for future addition:

   * Maximum link utilization value -- This value indicates the highest possible link utilization percentage set for each link. (Note: to avoid floating point numbers, the values should be integer values.)

   * Minimum link utilization value -- This value indicates the lowest possible link utilization percentage set for each link. (Note: same as above)

   * Overbooking Factor -- The overbooking factor allows the reserved bandwidth to be overbooked on each link beyond its physical capacity limit.

   * Maximum number of hops for all the LSPs -- This is the largest number of hops that any LSP can have. Note that this constraint can also be provided on a per LSP basis (as requested in [RFC4657] and defined in [PCEP]).

   * Exclusion of links/nodes in all LSP path computation (i.e., all LSPs should not include the specified links/nodes in their paths). Note that this constraint can also be provided on a per LSP basis (as requested in [RFC4657] and defined in [PCEP]).

   * An indication should be available in a path computation response that further reoptimization may only become available once existing traffic has been moved to the new LSPs.

- o A Global Concurrent Vector (GCV) field in which to specify all the individual path computation requests that are subject to concurrent path computation and subject to the global objective function and all of the global constraints. Note that this requirement is entirely fulfilled by the SVEC object in the PCEP specification [PCEP]. Since the SVEC object as defined in [PCEP] allows identifying a set of concurrent path requests, the SVEC can be reused to specify all the individual concurrent path requests

for a global concurrent optimization.

o An indicator field in which to indicate the outcome of the request. When the PCE could not find a feasible solution with the initial request, the reason for failure SHOULD be indicated. This requirement is partially covered by [RFC4657], but not in this level of detail. The following indicators SHOULD be supported at the minimum:

  * no feasible solution found. Note that this is already covered in [PCEP].

  * memory overflow

  * PCE too busy. Note that this is already covered in [PCEP].

  * PCE not capable of concurrent reoptimization

  * no migration path available

  * administrative privileges do not allow global reoptimization

o In order to minimize disruption associated with bulk path provisioning, the following requirements MUST be supported:

  * The request message MUST allow requesting the PCE to provide the order in which LSPs should be reoptimized (i.e., the migration path) in order to minimize traffic disruption during the migration. That is the request message MUST allow indicating to the PCE that the set of paths that will be provided in the response message (PCRep) has to be ordered.

  * In response to the "ordering" request from the PCC, the PCE MUST be able to indicate in the response message (PCRep) the order in which LSPs should be reoptimized so as to minimize traffic disruption. It should indicate for each request the order in which the old LSP should be removed and the order in which the new LSP should be setup. If the removal order is lower than the setup order this means that make-before-break cannot be done for this request. It might also be desirable to have the PCE indicate whether ordering is in fact required or not.

  * As stated in RFC 4657, the request for a reoptimization MUST support the inclusion of the set of previously computed paths along with their bandwidth. This is to avoid double bandwidth accounting and also this allows running an algorithm that minimizes perturbation and that can compute a migration path

(LSP setup/removal orders). This is particularly required for stateless PCEs.

* During a migration it may not be possible to do a make-before-break for all existing LSPs. The request message must allow indicating for each request whether make-before-break is required (e.g. Voice traffic) or break-before-make is acceptable (e.g. Internet traffic). The response message must allow indicating LSPs for which make-before-break reoptimization is not possible (this will be deduced from the LSP setup and deletion orders).

* During a reoptimization it may be required to move a LSP several times so as to avoid traffic disruption. The response message must allow indicating the sequence of successive paths for each request.

5. Protocol extensions for support of global concurrent optimization

This section provides protocol extensions for support of global concurrent optimization. Protocol extensions discussed in this section are built on [PCEP].

The format of a PCReq message is currently as follows per [PCEP]:

```
<PCReq Message>::= <Common Header>
                        [<SVEC-list>]
                        <request-list>

           where:
                   <SVEC-list>::=<SVEC> [<SVEC-list>]
                   <request-list>::=<request> [<request-list>]
                   <request>::=<RP>
                           [<END-POINTS>]
                           [<LSPA>]
                           [<BANDWIDTH>]
                           [<METRIC>]
                           [<RRO>]
                           [<IRO>]
                           [<LOAD-BALANCING>]
```

The format of a PCReq message after incorporating new requirements for support of global concurrent optimization is as follows:

```
<PCReq Message>::=<Common Header>
                   [<SVEC-list>]
                   <request-list>
```

The <SVEC-list> is changed as follows:

```
<SVEC-list>:: =<SVEC>
               [<OF>]
               [<GC>]
               [<XRO>]
               [<SVEC-list>]
```

Note that three optional objects are added, following the SVEC object: the OF (Objective Function) object, which is defined in [PCE-OF], the GC (Global Constraints) object, which is defined in this document (section 5.5), as well as the eXclude Route Object (XRO) which is defined in [PCE-XRO]. Details of this change will be

discussed in the following sections

5.1. Global Objective Function (GOF) Specification

The global objective function can be specified in the PCEP Objective Function (OF) object, defined in [PCE-OF]. The OF object includes a 16 bit Objective Function identifier. As per discussed in [PCE-OF], objective function identifier code points are managed by IANA.

Two global objective functions are defined in this document and their identifier should be assigned by IANA (suggested value)

Function Code	Description
1	Minimize the sum of all TE LSP costs (min cost)
2	Evenly allocate the network load to achieve the most uniform link utilization across all links*

* Note: This can be achieved by the following objective function: minimize max over all links {(C(i)-A(i))/C(i)} where C(i) is the link capacity for link i and A(i) is the total bandwidth allocated on link i.)

5.2. Indication of Global Concurrent Optimization Requests

All the path requests in this application should be indicated so that the global objective function and all of the global constraints are applied to each of the requested path computation. In order to support this requirement, a new flag is defined in the SVEC object:

C flag (1 bit): This is a new flag in the SVEC object. When C flag is set, this indicates that all of the path requests listed in the body of the SVEC object should be computed applying the global constraints and the global objective function.

When the C Flag is set in the SVEC Object, the OF and the GC objects, if included, should directly follow the SVEC Object.

5.3. Request for the order of LSP

In order to minimize disruption associated with bulk path provisioning, the PCC may indicate to the PCE that the response MUST be ordered. That is, the PCE has to include the order in which LSPs MUST be moved so as to minimize traffic disruption. To support such

indication a new flag, the D flag, is defined in the RP object as follows:

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |       Reserved    |              Flags      |D|M|F|O|B|R| Pri |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                        Request-ID-number                      |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    //                      Optional TLV(s)                        //
    |                                                               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: RP object body format in the PCReq Message

D bit (orDer - 1 bit): when set, in a PCReq message, the requesting PCC requires the PCE to specify in the PCRep message the order in which this particular path request is to be provisioned relative to other requests.

M bit (Make-before-break - 1 bit): when set, this indicates that a make-before-break reoptimization is required for this request.

When M bit is not set, this implies that a break-before-make reoptimization is allowed for this request. Note that M bit can be set only if the R (Reoptimization) flag is set.

5.4. The Order Response

The PCE MUST specify the order number in response to the Order Request made by the PCC in the PCReq message if so requested by the setting of the D bit in the RP object in the PCReq message. To support such ordering indication a new optional TLV is defined in the RP object, the Order TLV, as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Reserved    |             Flags      |D|M|F|O|B|R| Pri |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Request-ID-number                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                        Order TLV (Optional TLV)             //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: RP object body format in the PCRep Message

The Order TLV is an optional TLV in the RP object, that indicates the order in which the old LSP must be removed and the new LSP must be setup during a reoptimization. It is carried in the PCRep message in response to a reoptimization request.

The Order TLV SHOULD be included in the RP object in the PCRep message if the D bit is set in the RP object in the PCReq message.

The format of the Order TLV is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |    Length     |                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Delete Order                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Setup Order                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

          Type     To be defined by IANA (suggested value = )
          Length   Variable
          Value    Orders in which the old path should be removed
                     and the new path should be setup
```

Figure 7: The Order TLV in the RP object in the PCRep Message

Delete Order: 32 bit integer that indicates the order in which the old LSP should be removed

Setup Order: 32 bit integer that indicates the order in which the new LSP should be setup

The delete order should not be equal to the setup order. If the delete order is higher than the setup order, this means that the reoptimization can be done in a make-before-break manner, else it cannot be done in a make-before-break manner.

To illustrate, consider a network with two established requests: R1 with path P1 and R2 with path P2. During a reoptimization the PCE may provide the following ordered reply:

R1, path P1', remove order 1, setup order 4
R2, path P2', remove order 3, setup order 2

This indicates that the NMS should do the following sequence of tasks:

1: Remove path P1
2: Setup path P2'
3: Remove path P2
4: Setup path P1'

That is, R1 is reoptimized in a break-before-make manner and R2 in a make-before-break manner.

5.5. Global Constraints (GC) Object

The Global Constraints (GC) Object is used in a PCReq message to specify the necessary global constraints that should be applied to all individual path computations for a global concurrent path optimization request.

The format of the GC object body that includes the global constraints is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       MU      |       mU      |       OB      |       MH      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                         Optional TLV(s)                     //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 10: GC body object format

MU (Max Utilization) (8 bits) : 8 bit integer that indicates the upper bound utilization percentage by which all link should be bound. Utilization = (Link Capacity - Allocated Bandwidth on the Link)/ Link Capacity

mU (minimum Utilization) (8 bits) : 8 bit integer that indicates the lower bound utilization percentage by which all link should be bound.

OB (Over Booking factor) (8 bits) : 8 bit integer that indicates the overbooking percentage that allows the reserved bandwidth to be overbooked on each link beyond its physical capacity limit. The value, for example, 10% means that 110 Mbps can be reserved on a 100Mbps link.

MH (Max Hop) (8 bits): 8 bit integer that indicates the maximum hop count for all the LSPs.

GC Object-Class is to be assigned by IANA.

GC Object-Type is to be assigned by IANA.

The exclusion of the list of nodes/links from a global path computation can be done by including the XRO object following the GC object in the new SVEC list definition.

5.6. Error Indicator

To indicate errors associated with the global concurrent path optimization request, a new Error-Type (14) and subsequent error-values are defined as follows for inclusion in the PCEP-ERROR object:

A new Error-Type (14) and subsequent error-values are defined as follows:

Error-Type=14 and Error-Value=1: if a PCE receives a global concurrent path optimization request and the PCE is not capable of the request due to insufficient memory, the PCE MUST send a PCErr message with a PCEP ERROR object (Error-Type=14) and an Error-Value (Error-Value=1). The corresponding global concurrent path optimization request MUST be cancelled.

Error-Type=14; Error-Value=2: if a PCE receives a global concurrent path optimization request and the PCE is not capable of global concurrent optimization, the PCE MUST send a PCErr message with a PCEP-ERROR Object (Error-Type=14) and an Error-Value (Error-Value=2). The corresponding global concurrent path optimization MUST be

cancelled.

To indicate an error associated with policy violation, a new error value "global concurrent optimization not allowed" should be added to an existing error code for policy violation (Error-Type=5) as defined in [PCEP].

Error-Type=5; Error-Value=3: if a PCE receives a global concurrent path optimization request which is not compliant with administrative privileges (i.e., the PCE policy does not support global concurrent optimization), the PCE sends a PCErr message with a PCEP-ERROR Object (Error-Type=5) and an Error-Value (Error-Value=3). The corresponding global concurrent path computation MUST be cancelled.

5.7. NO-PATH Indicator

To communicate the reason(s) for not being able to find global concurrent path computation, the NO-PATH object can be used in the PCRep message. The format of the NO-PATH object body is as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|C|         Flags               |          Reserved             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//                      Optional TLV(s)                        //
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 11: NO-PATH object format

Flags (16 bits). The C flag is defined in [PCEP].

Two new bit flags are defined in the NO-PATH-VECTOR TLV carried in the NO-PATH Object:

0x08: when set, the PCE indicates that no migration path was found.

0x10: when set, the PCE indicates no feasible solution was found that meets all the constraints associated with global concurrent path optimization in the PCRep message.

When either 0x08 or 0x10 flag is set in the NO-PATH-VECTOR TLV carried in the NO-PATH object in the PCRep Message, a subsequent multi-session feature may be triggered if the PCC's local policy

allows it. The multi-session feature allows the original global concurrent optimization to be split into a number of multiple sessions so that the PCE would compute a number of smaller-scale optimizations in a sequential manner. The trade-off is that a partial feasible solution may be obtained using this approach which is better than not having any solution at all, although such solution might not be a global optimal solution. How to divide up the original set of global concurrent optimization requests into multiple numbers of smaller-scale optimizations is out of the scope of this document.

6. Manageability Considerations

Manageability of Global Concurrent Path Computation with PCE must address the following considerations:

6.1. Control of Function and Policy

In addition to the parameters already listed in section 8.1 of [PCEP], a PCEP implementation SHOULD allow configuring the following PCEP session parameters on a PCC:

- o The ability to send a GCO request.

In addition to the parameters already listed in section 8.1 of [PCEP], a PCEP implementation SHOULD allow configuring the following PCEP session parameters on a PCE:

- o The support for Global Concurrent Optimization.

- o The maximum number of synchronized path requests per request message.

- o A set of GCO specific policies (authorized sender, request rate limiter, etc).

These parameters may be configured as default parameters for any PCEP session the PCEP speaker participates in, or may apply to a specific session with a given PCEP peer or a specific group of sessions with a specific group of PCEP peers.

6.2. Information and Data Models, e.g. MIB module

Extensions to the PCEP MIB module defined in [PCEP-MIB] should be defined, so as to cover the GCO information introduced in this document.

6.3. Liveness Detection and Monitoring

Mechanisms defined in this draft does not imply any new liveness detection and monitoring requirements in addition to those already listed in section 8.3 of [PCEP].

6.4. Verifying Correct Operation

Mechanisms defined in this draft does not imply any new verification requirements in addition to those already listed in section 8.4 of [PCEP]

6.5. Requirements on Other Protocols and Functional Components

The PCE Discovery mechanisms ([ISIS PCED] and [OSPF PCED]) may be used to advertise global concurrent path computation capabilities to PCCs.

6.6. Impact on Network Operation

Mechanisms defined in this draft does not imply any new network operation requirements in addition to those already listed in section 8.6 of [PCEP].

7. Security Considerations

When global re-optimization is applied to an active network, it could be extremely disruptive. Although the real security and policy issues apply at the NMS, if the wrong results are returned to the NMS, the wrong actions may be taken in the network. Therefore, it is very important that the operator issuing the commands has sufficient authority and is authenticated, and that the computation request is subject to appropriate policy.

The mechanism defined in [PCEP] to secure a PCEP session can be used to secure global concurrent path computation requests/responses.

8. Acknowledgements

We would like to thank Jerry Ash, Adrian Farrel, J-P Vasseur, Ning So and Lucy Yong for their useful comments and suggestions.

9. IANA Considerations

A future revision of this document will present requests to IANA for codepoint allocation.

10. References

10.1. Normative References

[PCE-OF] Le Roux, JL., Vasseur, JP., and Y. Lee, "Objective Function encoding in Path Computation Element communication and discovery protocols, draft-leroux-pce-of, work in progress".

[PCE-XRO] Oki, E. and A. Farrel, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Route Exclusions, draft-ietf-pce-pcep-xro-00.txt, work in progress".

[PCEP] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) communication Protocol (PCEP) - Version 1, draft-ietf-pce-pcep, work in progress".

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

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

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

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

[RFC4674] Le Roux, J., "Requirements for Path Computation Element (PCE) Discovery, RFC 4674, October 2006.".

10.2. Informative References

[ISIS-PCED] Le Roux, J. and JP. Vasseur, "IS-IS protocol extensions for Path Computation Element (PCE) Discovery, draft-ietf-pce-disco-proto-isis, work in progress.".

[MLN-REQ] Shiomoto, K., Ed., "Requirements for GMPLS-based multi-region and multi-layer networks (MRN/MLN), draft-ietf-ccamp-gmpls-mln-reqs, work in progress".

[OSPF-PCED] Le Roux, J. and JP. Vasseur, "OSPF protocol extensions for

Path Computation Element (PCE) Discovery,
draft-ietf-pce-disco-proto-ospf, work in progress.".

[PCE-MLN] Oki, E., Le Roux, J., and A. Farrel, "Framework for PCE-
based inter-layer MPLS and GMPL traffic engineering,
draft-ietf-pce-inter-layer-frwk, work in progress.".

[PCEP-MIB]
Stephen, E. and K. Koushik, "PCE communication
protocol(PCEP) Management Information Base,
draft-kkoushik-pce-pcep-mib, work in progress.".

Authors' Addresses

Young Lee
Huawei
1700 Alma Drive, Suite 100
Plano, TX 75075
US

Phone: +1 972 509 5599 x2240
Fax: +1 469 229 5397
Email: ylee@huawei.com

JL Le Roux
France Telecom
2, Avenue Pierre-Marzin
Lannion 22307
FRANCE

Email: jeanlouis.leroux@orange-ftgroup.com

Daniel King
Aria Networks
44/45 Market Place
Chippenham SN15 3HU
United Kingdom

Phone: +44 7790 775187
Fax: +44 1249 446530
Email: daniel.king@aria-networks.com

Eiji Oki
NTT
Midori 3-9-11
Musashino, Tokyo 180-8585
JAPAN

Email: oki.eiji@lab.ntt.co.jp

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).