

Network Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: Aug 2009

Y. Lee  
Huawei  
JL. Le Roux  
France Telecom  
D. King  
Old Dog Consulting  
E. Oki  
Univeristy of Electro Communications  
March 29, 2009

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

[draft-ietf-pce-global-concurrent-optimization-10.txt](#)

### Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on August 29, 2009.



## Abstract

The Path Computation Element Communication Protocol (PCEP) allows Path Computation Clients (PCCs) to request path computations from Path Computation Elements (PCEs), and lets the PCEs return responses. When computing or re-optimizing the routes of a set of TE 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 GCO is able to simultaneously consider the entire topology of the network and the complete set of existing TE LSPs, and their respective constraints, and look to optimize or re-optimize the entire network to satisfy all constraints for all TE LSPs. A GCO may also be applied to some subset of the TE LSPs in a network. The GCO application is primarily a Network Management System (NMS) solution.

This document provides application-specific requirements and the PCEP extensions in support of GCO applications.

## Table of Contents

|                        |                                                                                             |                    |
|------------------------|---------------------------------------------------------------------------------------------|--------------------|
| <a href="#">1.</a>     | <a href="#">Introduction . . . . .</a>                                                      | <a href="#">4</a>  |
| <a href="#">2.</a>     | <a href="#">Terminology . . . . .</a>                                                       | <a href="#">6</a>  |
| <a href="#">3.</a>     | <a href="#">Applicability of Global Concurrent Optimization (GCO) . . . . .</a>             | <a href="#">7</a>  |
| <a href="#">3.1.</a>   | <a href="#">Application of the PCE Architecture . . . . .</a>                               | <a href="#">7</a>  |
| <a href="#">3.2.</a>   | <a href="#">Greenfield Optimization . . . . .</a>                                           | <a href="#">8</a>  |
| <a href="#">3.2.1.</a> | <a href="#">Single-layer Traffic Engineering . . . . .</a>                                  | <a href="#">8</a>  |
| <a href="#">3.2.2.</a> | <a href="#">Multi-layer Traffic Engineering . . . . .</a>                                   | <a href="#">8</a>  |
| <a href="#">3.3.</a>   | <a href="#">Re-optimization of Existing Networks . . . . .</a>                              | <a href="#">8</a>  |
| 3.3.1.                 | <a href="#">Reconfiguration of the Virtual Network Topology (VNT) . . . . .</a>             | <a href="#">9</a>  |
| <a href="#">3.3.2.</a> | <a href="#">Traffic Migration . . . . .</a>                                                 | <a href="#">9</a>  |
| <a href="#">4.</a>     | <a href="#">PCECP Requirements . . . . .</a>                                                | <a href="#">10</a> |
| <a href="#">5.</a>     | <a href="#">Protocol Extensions for Support of Global Concurrent Optimization . . . . .</a> | <a href="#">14</a> |
| <a href="#">5.1.</a>   | <a href="#">Global Objective Function (GOF) Specification . . . . .</a>                     | <a href="#">14</a> |
| <a href="#">5.2.</a>   | <a href="#">Indication of Global Concurrent Optimization Requests . . . . .</a>             | <a href="#">15</a> |
| <a href="#">5.3.</a>   | <a href="#">Request for The Order of TE LSP . . . . .</a>                                   | <a href="#">15</a> |
| <a href="#">5.4.</a>   | <a href="#">The Order Response . . . . .</a>                                                | <a href="#">16</a> |
| <a href="#">5.5.</a>   | <a href="#">GLOBAL CONSTRAINTS (GC) Object . . . . .</a>                                    | <a href="#">17</a> |
| <a href="#">5.6.</a>   | <a href="#">Error Indicator . . . . .</a>                                                   | <a href="#">18</a> |
| <a href="#">5.7.</a>   | <a href="#">NO-PATH Indicator . . . . .</a>                                                 | <a href="#">19</a> |
| <a href="#">6.</a>     | <a href="#">Manageability Considerations . . . . .</a>                                      | <a href="#">20</a> |
| <a href="#">6.1.</a>   | <a href="#">Control of Function and Policy . . . . .</a>                                    | <a href="#">20</a> |
| <a href="#">6.2.</a>   | <a href="#">Information and Data Models, e.g. MIB module . . . . .</a>                      | <a href="#">20</a> |
| <a href="#">6.3.</a>   | <a href="#">Liveness Detection and Monitoring . . . . .</a>                                 | <a href="#">20</a> |
| <a href="#">6.4.</a>   | <a href="#">Verifying Correct Operation . . . . .</a>                                       | <a href="#">20</a> |
| 6.5.                   | <a href="#">Requirements on Other Protocols and Functional Components . . . . .</a>         | <a href="#">21</a> |
| <a href="#">6.6.</a>   | <a href="#">Impact on Network Operation . . . . .</a>                                       | <a href="#">21</a> |
| <a href="#">7.</a>     | <a href="#">Security Considerations . . . . .</a>                                           | <a href="#">21</a> |
| <a href="#">8.</a>     | <a href="#">Acknowledgements . . . . .</a>                                                  | <a href="#">21</a> |
| <a href="#">9.</a>     | <a href="#">IANA Considerations . . . . .</a>                                               | <a href="#">22</a> |
| <a href="#">9.1.</a>   | <a href="#">Request Parameter Bit Flags . . . . .</a>                                       | <a href="#">22</a> |
| <a href="#">9.2.</a>   | <a href="#">New PCEP TLV . . . . .</a>                                                      | <a href="#">22</a> |
| <a href="#">9.3.</a>   | <a href="#">New Flag in PCE-CAP-FLAGS Sub-TLV in PCED . . . . .</a>                         | <a href="#">22</a> |
| <a href="#">9.4.</a>   | <a href="#">New PCEP Object . . . . .</a>                                                   | <a href="#">23</a> |
| <a href="#">9.5.</a>   | <a href="#">New PCEP Error Codes . . . . .</a>                                              | <a href="#">23</a> |
| <a href="#">9.5.1.</a> | <a href="#">New Error-Values for Existing Error-Types . . . . .</a>                         | <a href="#">23</a> |
| <a href="#">9.5.2.</a> | <a href="#">New Error-Types and Error-Values . . . . .</a>                                  | <a href="#">23</a> |
| <a href="#">9.6.</a>   | <a href="#">New No-Path Reasons . . . . .</a>                                               | <a href="#">24</a> |
| <a href="#">10.</a>    | <a href="#">References . . . . .</a>                                                        | <a href="#">24</a> |
| <a href="#">10.1.</a>  | <a href="#">Normative References . . . . .</a>                                              | <a href="#">24</a> |
| <a href="#">10.2.</a>  | <a href="#">Informative References . . . . .</a>                                            | <a href="#">25</a> |
|                        | <a href="#">Authors' Addresses . . . . .</a>                                                | <a href="#">25</a> |
|                        | <a href="#">Intellectual Property and Copyright Statements . . . . .</a>                    | <a href="#">26</a> |



## 1. Introduction

[RFC4655] defines the Path Computation Element (PCE) based Architecture and explains how a PCE may compute Label Switched Paths (LSPs) in Multiprotocol Label Switching Traffic Engineering (MPLS-TE) and Generalized MPLS (GMPLS) networks at the request of Path Computation Clients (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 PCE Communication Protocol (PCEP) is the communication protocol used between PCC and PCE, and may also be used between cooperating PCEs. [RFC4657] sets out generic protocol requirements for PCEP. Additional application-specific requirements for PCEP are defined in separate documents.

This document provides a set of requirements and PCEP extensions in support of concurrent path computation applications. 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 or a PCE Server based solution. Owing to complex synchronization issues associated with GCO applications, 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 preclude other architectural alternatives to support GCO application, but they are NOT RECOMMENDED. For instance, GCO might be enabled by distributed LSRs through complex synchronization mechanisms. However, this approach might 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.

The need for 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 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 promote the use of such systems for offline processing. Online application of this work



should only be considered with proven empirical validation.

As new TE LSPs are added or removed from the network over time, the global network resources become fragmented and the existing placement of TE LSPs within network no longer provides 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 TE LSPs and their respective constraints, and look to re-optimize the entire network to satisfy all constraints for all TE LSPs. Alternatively, the application may consider a subset of the TE LSPs and/or a subset of the network topology. Note that other preemption can also help reducing the fragmentation issues.

While GCO is applicable to any simultaneous request for multiple TE LSPs (for example, a request for end-to-end protection), it is NOT RECOMMENDED that global concurrent reoptimization would be applied in a network (such as an MPLS-TE network) that contains a very large number of very low bandwidth or zero bandwidth TE LSPs since the large scope of the problem and the small benefit of concurrent reoptimization relative to single TE LSP reoptimization is unlikely to make the process worthwhile. Further, applying global concurrent reoptimization in a network with a high rate of change of TE LSPs (churn) is NOT RECOMMENDED because of the likelihood that TE LSPs would change before they could be globally reoptimized. Global reoptimization is more applicable to stable networks such as transport networks or those with long-term TE LSP tunnels.

The main focus of this document is to highlight the PCC-PCE communication needs in support of a concurrent path computation applications 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 performing computations in support of GCO. Discovery of such capabilities might be desirable and could be achieved through extensions to the PCE discovery mechanisms [[RFC4674](#)], [[RFC5088](#)], [[RFC5089](#)], but that is out of the scope of this document.

It is to be noted that Backward Recursive Path Computation (BRPC) [[BRPC](#)] is a multi-PCE path computation technique used to compute a shortest constrained inter-domain path whereas this ID specifies a technique where a set of path computation requests are bundled and send to a PCE with the objective of "optimizing" the set of computed paths.



## 2. Terminology

Most of the terminology used in this document is explained in [\[RFC4655\]](#). A few key terms are repeated here for clarity.

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 from 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 optimize network resources. A GCO path computation is able to simultaneously consider the entire topology of the network and the complete set of existing TE LSPs, and their respective constraints, and look to optimize or re-optimize the entire network to satisfy all constraints for all TE LSPs. A GCO path computation can also provide an optimal way to migrate from an existing set of TE 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.



### 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.

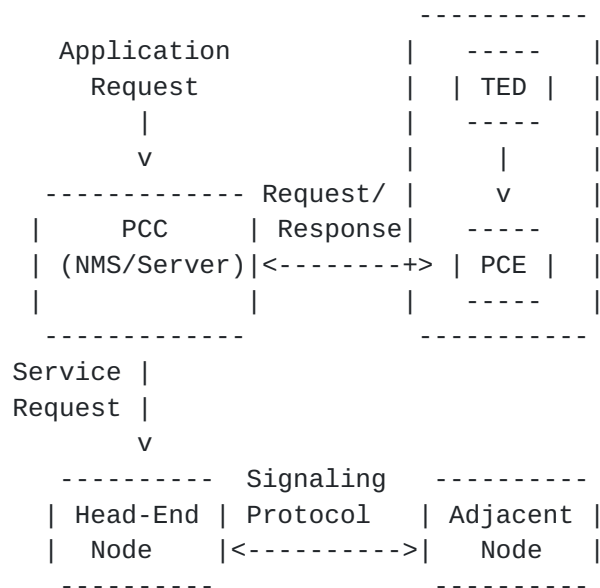


Figure 1: PCE-Based Architecture for Global Concurrent Optimization

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. Various algorithms and computation techniques have been proposed to perform this function. Specification of such algorithms or techniques is outside the scope of this document.



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. Greenfield Optimization**

Greenfield optimization is a special case of GCO application when there are no TE LSPs already set up in the network. The need for greenfield optimization arises when network planner wants to make use of a computation server to plan the TE LSPs that will be provisioned in the network. Note that once greenfield operation is one-time optimization. When network conditions change due to failure or other changes, then re-optimization mode of operation will kick in.

When a new TE network needs to be provisioned from a greenfield perspective, a set of TE LSPs needs to be created based on traffic demand, network topology, service constraints, and network resources. In this scenario, the ability to perform concurrent computation is desirable, or required, to utilize network resources in an optimal manner and avoid blocking.

#### **3.2.1. Single-layer Traffic Engineering**

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

#### **3.2.2. Multi-layer Traffic Engineering**

Greenfield optimization is not limited to single-layer traffic engineering. It can also be applied to multi-layer traffic engineering [[PCE-MLN](#)]. 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.

### **3.3. 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 TE LSPs, a sequential re-optimization may not produce substantial improvements in terms of overall network optimization. Sequential re-optimization refers to a path computation method that computes the re-optimized path of one TE LSP at a time without giving any consideration to the other TE LSPs that need to be re-optimized in the network. The potential for network-wide gains from reoptimization of TE LSPs sequentially is dependent upon the network usage and size of the TE LSPs being optimized. However, the key point remains: computing the reoptimized path of one TE LSP at a time without giving any consideration to the other TE 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 TE LSPs per link, and far less visible in packet networks with micro-flow TE 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 TE LSPs as a set rather than computing new paths from the head-end LSRs in a distributed manner. Distributed jittering is a technique that could prevent race condition (i.e., competing for the same resource from different head-end LSRs) with a distributed computation. GCO provides an alternative way that could also prevent race condition in a centralized manner. However, a centralized system will typically suffer from a slower response time than a distributed system.

#### **3.3.1. Reconfiguration of the Virtual Network Topology (VNT)**

Reconfiguration of the VNT [[RFC5212](#)] [[PCE-MLN](#)] 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 TE LSPs to be re-computed.

#### **3.3.2. 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 TE LSPs before tearing down the old TE LSPs. When multiple TE LSP routes are changed according to the computed results, some of the TE 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 TE



LSPs to the new optimal TE LSPs without disrupting traffic because there are insufficient network resources to support both sets of TE LSPs when make-before-break is used. However, a PCE may be able to determine a sequence of make-before-break replacement of individual TE LSPs or small sets of TE LSPs so that the full set of TE LSPs can be migrated without any disruption. This scenario assumes that the bandwidth of existing TE LSP is kept during the migration which is required in optical networks. In packet networks, this assumption can be relaxed as the bandwidth of temporary TE LSPs during migration can be zeroed.

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 TE LSPs to the new TE LSPs. In this case a migration strategy is required that may necessitate TE 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 TE LSPs must be established and take over from the old TE 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 TE LSPs.

The benefit of this multi-step rerouting includes minimization of traffic disruption 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 TE LSPs carrying low priority services (e.g., Internet traffic) and not allowed for some TE LSPs carrying mission critical services (e.g., voice traffic).

#### **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 \[RFC4657\]](#) and that [\[PCE-OF\]](#) defined three global objective functions as follows. A list of available global objective functions SHOULD include the following objective functions at the minimum and SHOULD be expandable for future addition:
  - \* Minimize aggregate Bandwidth Consumption (MBC)
  - \* Minimize the load of the Most Loaded Link (MLL)
  - \* Minimize Cumulative Cost of a set of paths (MCC)
- 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 TE LSPs -- This is the largest number of hops that any TE LSP can have. Note that this constraint can also be provided on a per TE LSP basis (as requested in [[RFC4657](#)] and defined in [PCEP]).
  - \* Exclusion of links/nodes in all TE LSP path computation (i.e., all TE LSPs should not include the specified links/nodes in their paths). Note that this constraint can also be provided on a per TE 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 TE 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 cannot 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 TE 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 TE LSPs should be reoptimized so as to minimize traffic disruption. It should indicate for each request the order in which the old TE LSP should be removed and the order in which the new TE 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 MAY also be desirable to have the PCE indicate whether ordering is in fact required or not.
  - \* During a migration it may not be possible to do a make-before-break for all existing TE 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 TE LSPs for which make-before-break reoptimization is not possible (this will be deduced from the TE LSP setup and deletion orders).



## 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 after incorporating new requirements for support of global concurrent optimization is as follows. The message format uses Reduced Backus-Naur Format as defined in [RBNF].

```
<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]. The placement of the OF object (in which the global objective function is specified) in the SVEC-list is defined in [PCE-OF]. Details of this change will be discussed in the following sections.

Note also that when the XRO is global to a SVEC, and F bit is set, it SHOULD be allowed to specify multiple Record Route Objects in the PCReq message.

### 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.

Three global objective functions defined in [PCE-OF] are used in the context of GCO.



| Function Code | Description                                      |
|---------------|--------------------------------------------------|
| 4             | Minimize aggregate Bandwidth Consumption (MBC)   |
| 5             | Minimize the load of the Most Loaded Link (MLL)* |
| 6             | Minimize Cumulative Cost of a set of paths (MCC) |

\* 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. This can be indicated implicitly by placing the GCO related objects (GOF, GC or XRO) after the SVEC object. That is, if any of these objects follows the SVEC object in the PCReq message, all of the requested path computations specified in the SVEC object are subject to GOF, GC or XRO.

### **5.3. Request for The Order of TE 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 TE 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:

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.

To support the determination of whether make-before-break optimization is required, a new flag, the M flag, is defined in the RP object as follows.

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



Setup Order: 32 bit integer that indicates the order in which the new TE 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.

For a new TE LSP the delete order is not applicable. The value 0 is designated to specify this case. When the value of the delete order is 0, it implies that the resulting TE LSP is a new TE LSP.

To illustrate this, consider a network with two established TE LSPs: 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.

GLOBAL CONSTRAINTS Object-Class is to be assigned by IANA (recommended value=24)

GLOBAL CONSTRAINTS Object-Type is to be assigned by IANA (recommended value=1)

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



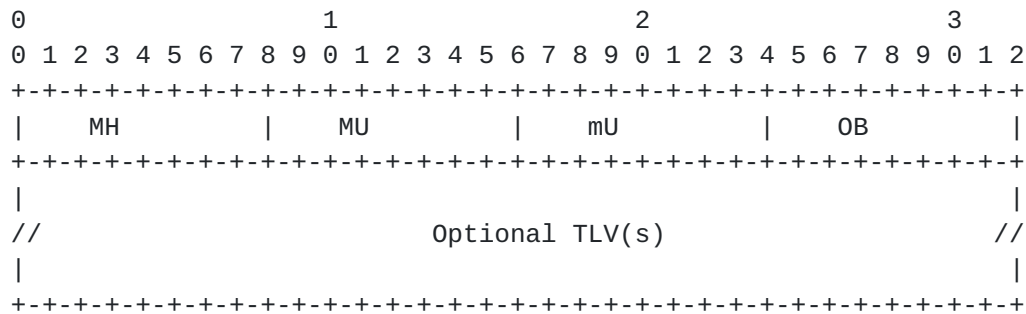


Figure 3: GC body object format

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

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

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

OB (Over Booking factor Percentage: 8 bits) : 8 bits 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.

Reserved bits (24 bits) of the GLOBAL CONSTRAINTS Object SHOULD be transmitted as zero and SHOULD be ignored upon receipt.

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.

Optional TLVs may be included within the GC object body to specify additional global constraints. The TLV format and processing is consistent with [Section 7.1 of RFC5440](#). Any TLVs will be allocated from the "PCEP TLV Type Indicators" registry. Note that no TLVs are defined in this document.

## 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 (15) and subsequent error-values are defined as follows:

Error-Type=15 and Error-Value=1: if a PCE receives a global concurrent path optimization request and the PCE is not capable of processing the request due to insufficient memory, the PCE MUST send a PCErr message with a PCEP ERROR object (Error-Type=15) and an Error-Value (Error-Value=1). The PCE stops processing the request. The corresponding global concurrent path optimization request MUST be cancelled at the PCC.

Error-Type=15; 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=15) and an Error-Value (Error-Value=2). The PCE stops processing the request. The corresponding global concurrent path optimization MUST be cancelled at the PCC.

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 [[RFC5440](#)].

Error-Type=5; Error-Value=5: 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=5). The PCE stops the processing the request. The corresponding global concurrent path computation MUST be cancelled at the PCC.

### **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 defined in [[RFC5440](#)]. The object may contain a NO-PATH-VECTOR TLV to provide additional information about why a path computation has failed.

Two new bit flags are defined to be carried in the Flags field in the NO-PATH-VECTOR TLV carried in the NO-PATH Object.

Bit 6: When set, the PCE indicates that no migration path was found.

Bit 7: 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.



## **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 \[RFC5440\]](#), 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 \[RFC5440\]](#), 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 document do not imply any new liveness detection and monitoring requirements in addition to those already listed in [Section 8.3 of \[RFC5440\]](#).

### **6.4. Verifying Correct Operation**

Mechanisms defined in this document do not imply any new verification requirements in addition to those already listed in [Section 8.4 of \[RFC5440\]](#)



### **6.5. Requirements on Other Protocols and Functional Components**

The PCE Discovery mechanisms ([[RFC5088](#)] and [[RFC5089](#)]) may be used to advertise global concurrent path computation capabilities to PCCs. A New Flag (value=9) in PCE-CAP-FLAGS Sub-TLV should be assigned to be able to indicate GCO capability.

### **6.6. Impact on Network Operation**

Mechanisms defined in this document do not imply any new network operation requirements in addition to those already listed in [Section 8.6 of \[RFC5440\]](#).

### **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 [[RFC5440](#)] 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, Lucy Yong and Fabien Verhaeghe for their useful comments and suggestions.



## 9. IANA Considerations

IANA maintains a registry of PCEP parameters. IANA is requested to make allocations from the sub-registries as described in the following sections.

### 9.1. Request Parameter Bit Flags

As described in [Section 5.3](#), two new bit flags are defined for inclusion in the Flags field of the RP object. IANA is requested to make the following allocations from the "Request Parameter Bit Flags" sub-registry.

| Bit | Name  | Description              | Reference  |
|-----|-------|--------------------------|------------|
| 22  | D-bit | Report the request order | [This.I-D] |
| 21  | M-bit | Make-before-break        | [This.I-D] |

### 9.2. New PCEP TLV

As described in [Section 5.4](#), a new PCEP TLV is defined to indicate the setup and delete order of TE LSPs in a GCO. IANA is requested to make the following allocation from the "PCEP TLV Types" sub-registry.

| TLV Type | Meaning   | Reference  |
|----------|-----------|------------|
| 5        | Order TLV | [This.I-D] |

### 9.3. New Flag in PCE-CAP-FLAGS Sub-TLV in PCED

As described in [Section 6.5](#), a new PCE-CAP-FLAGS Sub-TLV is defined to indicate a GCO capability. IANA is requested to make the following allocation from the "PCE-CAP-FLAGS TLV Types" sub-registry. The "PCE Capability Flags Registry" is created by [section 7.2 of RFC 5088](#). It is an OSPF registry.

| FLAG | Meaning                              | Reference  |
|------|--------------------------------------|------------|
| 9    | Global Concurrent Optimization (GCO) | [This.I-D] |



#### **9.4. New PCEP Object**

As described in [Section 5.5](#), a new PCEP object is defined to carry global constraints. IANA is requested to make the following allocation from the "PCEP Objects" sub-registry.

| Object Class | Name                  | Reference  |
|--------------|-----------------------|------------|
| 24           | GLOBAL-CONSTRAINTS    | [This.I-D] |
|              | Object-Type           |            |
|              | 1: Global Constraints | [This.I-D] |

#### **9.5. New PCEP Error Codes**

As described in [Section 5.6](#), new PCEP error codes are defined for GCO errors. IANA is requested to make allocations from the "PCEP Error Types and Values" sub-registry as set out in the following sections.

##### **9.5.1. New Error-Values for Existing Error-Types**

| Error Type | Meaning                                    | Reference  |
|------------|--------------------------------------------|------------|
| 5          | Policy violation                           |            |
|            | Error-value=5:                             | [This.I-D] |
|            | Global concurrent optimization not allowed |            |

##### **9.5.2. New Error-Types and Error-Values**

| Error Type | Meaning                                      | Reference  |
|------------|----------------------------------------------|------------|
| 15         | Global Concurrent Optimization Error         | [This.I-D] |
|            | Error-value=1:                               |            |
|            | Insufficient memory                          | [This.I-D] |
|            | Error-value=2:                               |            |
|            | Global concurrent optimization not supported | [This.I-D] |



## 9.6. New No-Path Reasons

IANA is requested to make the following allocations from the "No-Path Reasons" sub-registry for bit flags carried in the NO-PATH-VECTOR TLV in the PCEP NO-PATH object as described in [Section 5.7](#).

| Bit Number | Name                        | Reference  |
|------------|-----------------------------|------------|
| 26         | No GCO migration path found | [This.I-D] |
| 25         | No GCO solution found       | [This.I-D] |

## 10. References

### 10.1. Normative References

- [BRPC] Vasseur, JP., Ed., "A Backward Recursive PCE-based Computation (BRPC) procedure to compute shortest inter-domain Traffic Engineering Label Switched Paths, [draft-ietf-pce-brpc](#), work in progress".
- [PCE-OF] Le Roux, JL., Vasseur, JP., and Y. Lee, "Objective Function encoding in Path Computation Element communication and discovery protocols, [draft-ietf-pce-of](#), work in progress".
- [PCE-XR0] Oki, E. and A. Farrel, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Route Exclusions, [draft-ietf-pce-pcep-xro](#), work in progress".
- [RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) communication Protocol (PCEP)", [RFC 5440](#), March 2009.
- [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.
- [RFC5088] Le Roux, J., Vasseur, J., Ikejiri, Y., and R. Zhang, "OSPF Protocol Extensions for Path Computation Element (PCE) Discovery", [RFC 5088](#), January 2008.
- [RFC5089] Le Roux, J., Vasseur, J., Ikejiri, Y., and R. Zhang, "IS-IS Protocol Extensions for Path Computation Element (PCE) Discovery", [RFC 5089](#), January 2008.



## **10.2. Informative References**

- [PCE-MLN] Oki, E., Le Roux, J., and A. Farrel, "Framework for PCE-based inter-layer MPLS and GMPLS 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.
- [RBNF] A. Farrel, "Reduced Backus-Naur Form (RBNF) - A Syntax Used in Various Protocol Specifications", [draft-farrel-rtg-common-bnf](#), work in progress.
- [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.
- [RFC5212] Shiomoto, K., Ed., "Requirements for GMPLS-based multi-region and multi-layer networks (MRN/MLN)", [RFC 5212](#), July 2008.

### 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  
Old Dog Consulting  
United Kingdom

Phone:  
Fax:  
Email: daniel@olddog.co.uk

Eiji Oki  
University of Electro-Communications  
1-5-1 Chofugaoka  
Chofu, Tokyo 182-8585  
JAPAN

Email: oki@ice.uec.ac.jp

#### Full Copyright Statement

Copyright (c) 2009 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

ALL IETF Documents and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

#### Intellectual Property

The definitive version of an IETF Document is that published by, or under the auspices of, the IETF. Versions of IETF Documents that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of IETF Documents. The definitive version of these Legal Provisions is that published by, or under the auspices of, the IETF. Versions of these Legal Provisions that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of these Legal Provisions.



For the avoidance of doubt, each Contributor to the IETF Standards Process licenses each Contribution that he or she makes as part of the IETF Standards Process to the IETF Trust pursuant to the provisions of [RFC 5378](#). No language to the contrary, or terms, conditions or rights that differ from or are inconsistent with the rights and licenses granted under [RFC 5378](#), shall have any effect and shall be null and void, whether published or posted by such Contributor, or included with or in such Contribution.