

Network Working Group

M. Boucadair (Ed.)

P. Morand (Ed.)

Internet Draft

France Telecom R&D

Document: [draft-boucadair-pcp-interas-00.txt](#)

October 2004

Category: Standards Track

Inter-AS PCE Communication protocol  
draft-boucadair-pcp-interas-00.txt

## Status of this Memo

This document is an Internet-Draft and is subject to all provisions of [section 3 of RFC 3667](#) [RFC3667]. By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she become aware will be disclosed, in accordance with [RFC 3668](#) [RFC3668].

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/lid-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 April 2005.

## Abstract

This draft describes a new protocol allowing communication between two Path Computation Elements (PCEs) located in different domains in order to compute inter-domain paths satisfying a set of QoS constraints. This protocol could also be used for intra-domain purposes.

Internet Draft

PCE Communication protocol

October 2004

## Table of Contents

<a href="#">1.</a>	Contributors.....	<a href="#">2</a>
<a href="#">2.</a>	Terminology.....	<a href="#">2</a>
<a href="#">3.</a>	Introduction.....	<a href="#">3</a>
<a href="#">4.</a>	Conventions used in this document.....	<a href="#">4</a>
<a href="#">5.</a>	Overview of overall service approach.....	<a href="#">4</a>
<a href="#">6.</a>	PCE to PCE communication.....	<a href="#">5</a>
<a href="#">7.</a>	PCP messages.....	<a href="#">6</a>
<a href="#">7.1.</a>	Common header.....	<a href="#">6</a>
<a href="#">7.2.</a>	OPEN message.....	<a href="#">7</a>
<a href="#">7.3.</a>	ACCEPT message.....	<a href="#">7</a>
<a href="#">7.4.</a>	CLOSE message.....	<a href="#">7</a>
<a href="#">7.5.</a>	REQUEST message.....	<a href="#">8</a>
<a href="#">7.6.</a>	RESPONSE-PATH message.....	<a href="#">11</a>
<a href="#">7.7.</a>	PATH-ERROR message.....	<a href="#">12</a>
<a href="#">7.8.</a>	CANCEL message.....	<a href="#">13</a>
<a href="#">7.9.</a>	ACKNOWLEDGE message.....	<a href="#">14</a>
<a href="#">7.10.</a>	KEEPALIVE message (KA).....	<a href="#">14</a>
<a href="#">8.</a>	Exchange of PCP messages.....	<a href="#">14</a>
<a href="#">8.1.</a>	Communication.....	<a href="#">14</a>
<a href="#">8.2.</a>	OPEN (OPN).....	<a href="#">14</a>
<a href="#">8.3.</a>	ACCEPT (ACP).....	<a href="#">15</a>
<a href="#">8.4.</a>	CLOSE (CLO).....	<a href="#">15</a>
<a href="#">8.5.</a>	REQUEST (REQ).....	<a href="#">15</a>
<a href="#">8.6.</a>	RESPONSE (RSP).....	<a href="#">18</a>
<a href="#">8.7.</a>	ACKNOWLEDGE (ACK).....	<a href="#">18</a>
<a href="#">8.8.</a>	CANCEL (CCL).....	<a href="#">19</a>
<a href="#">9.</a>	State diagram.....	<a href="#">19</a>
<a href="#">10.</a>	Security Considerations.....	<a href="#">19</a>
<a href="#">11.</a>	References.....	<a href="#">20</a>
<a href="#">12.</a>	Acknowledgments.....	<a href="#">20</a>
<a href="#">13.</a>	Author's Addresses.....	<a href="#">21</a>

[1.](#) Contributors

- o Hamid Asgari (Thales Research and Technology)
- o Panagiotis Georgatsos (Algonet)
- o David Griffin (University College London)
- o Micheal Howarth (University of Surrey)

## [2.](#) Terminology

This memo makes use of the following terms:

- o Path Computation Element (PCE): an entity that is responsible for computing/finding inter/intra domain LSPs. This entity can simultaneously act as client and a server. Several PCEs can be deployed in a given AS.

Boucadair (Ed.)      Standards Track - Expires April 2005

[Page 2]

---

Internet Draft

PCE Communication protocol

October 2004

- o Path Computation Client (PCC): a PCE acting as a client. This entity is responsible for issuing path computation requests that fulfill the Service Management constraints for the establishment of inter/intra domain LSPs.
- o Path Computation Server (PCS): a PCE acting as a server. This entity is responsible for handling path computation requests including neighboring PCC constraints.
- o High-level service: is the service using a PCE-based system as an underlying infrastructure (an inter-domain QoS VPNs service for instance)
- o High-level service customer: is a customer that subscribes to a High-level service.
- o pSLS: A provider SLS is an SLS established between two Internet Network Providers (INP) with the purpose of extending the geographical span of their service offers.
- o SLS Management: this includes service ordering (i.e establishing contracts between peers) and invocation (i.e committing resources before traffic can be admitted)
- o Q-BGP: QoS-inferred BGP. A modified BGP protocol that takes into account QoS information as input to for its route selection process.
- o Domain: within this draft it denotes an Autonomous system.

## [3.](#) Introduction

Nowadays, services are deployed on a same basic infrastructure (best-

effort shared IP network) on which more elaborate functionalities (MPLS for instance) rely for providing enhanced network services. Especially those intended for specific corporate customers or providers needs. These extra functionalities were introduced because the basic IP approach failed to support those added-value services or was not considered to be efficient enough.

MPLS is a technical solution that has been successfully deployed by a large number of providers for supporting connection-oriented services such as IP VPN services for which traffic isolation is an important criterion. Then, the solution evolved to encompass QoS issues, and Traffic engineering functions were then progressively introduced. Up to now, some providers have deployed MPLS TE but only within their own domains.

Extending the scope of offered intra-domain services (like QoS-based services), using MPLS as infrastructure, to the Internet scale is conditioned by the cooperation between service providers. Several proposals have been proposed within the IETF in order to deal with this issue but only from inter-AS point of view (see for example [\[INTERAREA-REQ\]](#), [\[INTERAS-REQ\]](#), [\[PCE-ARCH\]](#) and [\[PCE-FWK\]](#)).

Inter-provider issues need to be studied further in order to build a complete end-to-end solution.

Draft [\[INTERAS-PCE\]](#) describes a solution that could be implemented in order to offer end-to-end services. This solution requires a close cooperation between distinct Path Computation Elements (PCE) that are located in distinct domains.

This draft describes a protocol to use for communication between two Path computation Elements.

The structure of this draft is as follows:

- o [Section 5](#) presents an overview of the overall service approach;.
- o [Section 6](#) lists characteristics of the PCP protocol;.
- o Sections [7](#) and [8](#) detail the PCP messages and operations.

#### [4](#). Conventions used in this document

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](#)].

## 5. Overview of overall service approach

Neighboring domains establish pSLs between themselves. An inter-domain routing protocol runs between the domains. This inter-domain routing protocol is used to announce PCE unique identifiers [PCE-DISCOVERY] across the Internet in order for other PCEs to be able to discover possible paths towards every AS having a PCE. Therefore, when an AS wants to establish an LSP between 2 addresses, its PCE forms a path computation request containing the HEAD-END-ADDRESS and the TAIL-END-ADDRESS defining the future LSP. In addition to the IP address of the head and the tail of the LSP, each X-END-ADDRESS contains also the PCE unique identifier of the AS these IP addresses belong to. Using information reported by BGP the PCE identifies possible paths that reach the target AS identified by its PCE unique identifier. It then selects one of these paths and forms a new request, which is sent to the neighboring PCE it selected along that path.

The path computation request is propagated downstream to the appropriate PCEs and is repeated until the request reaches the

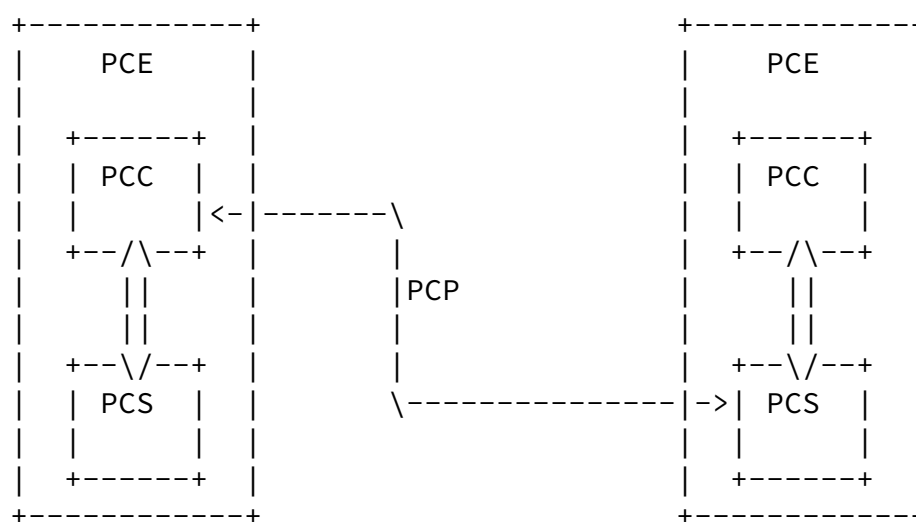
destination PCE. Each PCE along the path ensures that the constraints expressed by the request are satisfied. Each PCE is responsible for computing both the intra- and inter-domain sub-path and to ensure that resources are available and will remain available until the LSP is effectively created. If for some reasons the path computation aborts, all resources must be relaxed.

After authenticating the identity of LSP requester (originating) PCE, the destination PCE sends a reply message back to the downstream domain's PCE accepting the request. The LSP sub-path (from the ingress ASBR and the final destination) is inserted in the message. The next downstream domain's PCE does the same adding its own relevant sub-path to the overall loose or strict path. At the end of the chain, the originating PCE does also the same. An end-to-end path has thus been computed. The originating PCE is now in a position to provide the service request handler with appropriate information (end-to-end inter-domain path) allowing an RSVP reservation to be issued for the establishment of the LSP.

At the service/application level, when an originating AS wants to establish an LSP towards a destination ASs, there MUST exist a preliminary agreement between the two ASs (Service providers owning these PCEs). This agreement specifies both the tail-end and head-end address of the LSP, together with the PCE unique identifier of the originating and destination AS. This allows only agreed LSP to be established.

## 6. PCE to PCE communication

A PCE can act as a client (PCC) or a server (PCS). A PCC is responsible for issuing requests. PCS is responsible for handling requests received from PCCs.



PCP protocol is used for communication between a PCC and a PCS.

PCP is a simple query and response protocol that can be used between PCE entities to collaborate for computing an inter-domain QoS constrained path.

The main characteristics of the PCP protocol include:

- o The protocol employs a client/server model in which a PCE can both act as a client and/or a server at the same time. A PCE Client (PCC) sends requests, cancellation and receives responses.

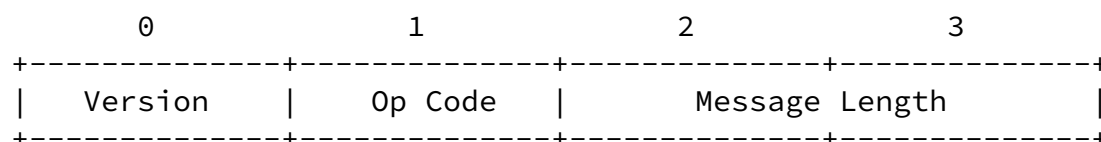
- o The protocol uses TCP as its transport protocol for reliable exchange of messages between PCE. Therefore, no additional mechanisms are necessary for reliable communication between two PCE.
- o In this first version, PCP does not provide any encryption mechanism, replay protection, and message integrity. But PCP can reuse existing protocols for security such as IPSEC [[RFC2401](#)] or TLS [[RFC2246](#)] to authenticate and secure the channel between two PCE.
- o The PCP protocol described below supports only a basic path computation service. In particular it doesn't support additional path computation constraints, nor enhanced reporting features in case of path computation failure.

## [7.](#) PCP messages

This section discusses the PCP message formats and objects exchanged between PCE entities.

### [7.1.](#) Common header

Each PCP message consists of the PCP header followed by a number of arguments depending on the nature of the operation.



Global note: //// implies field is reserved, set to 0.

The fields in the header are:

Version: 8 bits. PCP version number. Current version is 1.

Op Code: 8 bits. The PCP operations are:

1 = OPEN (OPN)

2 = ACCEPT (ACP)

3 = CLOSE (CLO)

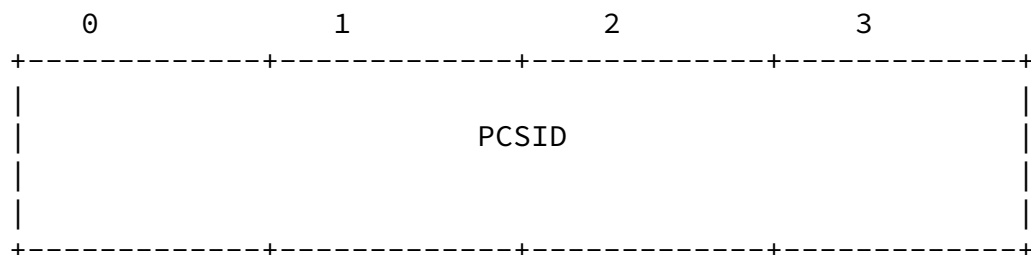
4 = REQUEST (REQ)

5 = RESPONSE	(RSP)
6 = PATH-ERROR	(ERR)
7 = CANCEL	(CCL)
8 = ACKNOWLEDGE	(ACK)
9 = KEEP-ALIVE	(KA)

Message Length: 16 bits

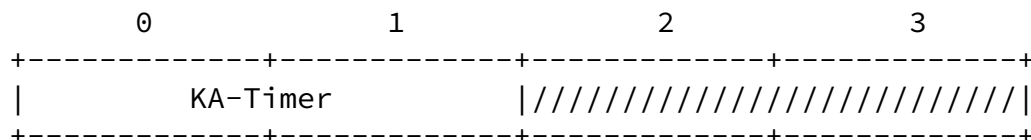
This is the size of the message in octets, which includes the standard PCP header and all encapsulated objects. Messages **MUST** be aligned on 4 octet intervals.

## 7.2. OPEN message



The message contains only one argument. This PCSID is propagated by BGP between the domains. This is a routable IPv4 or IPv6 address identifying a PCS of a domain. This PCSID must be inserted by the PCE opening a PCP session. The size of the PCSID is 4 or 16 bytes.

## 7.3. ACCEPT message

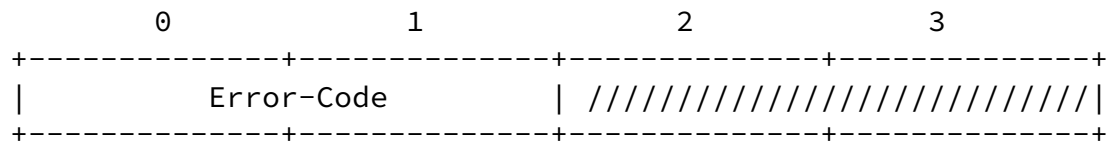


- o KA-Timer (Keep-Alive Timer): The argument of the accept message is a 2 octets integer value which represents a timer value expressed in units of seconds. This timer value is treated as a delta. KA-Timer is used to specify the maximum time interval over which a PCP message **MUST** be sent by the two communication entities. The range of finite timeouts is 1 to 65535 seconds represented as an unsigned two-octet integer. The value of zero implies infinity.

## 7.4. CLOSE message

The close message contains an error code indicating the reason of the close of the session.



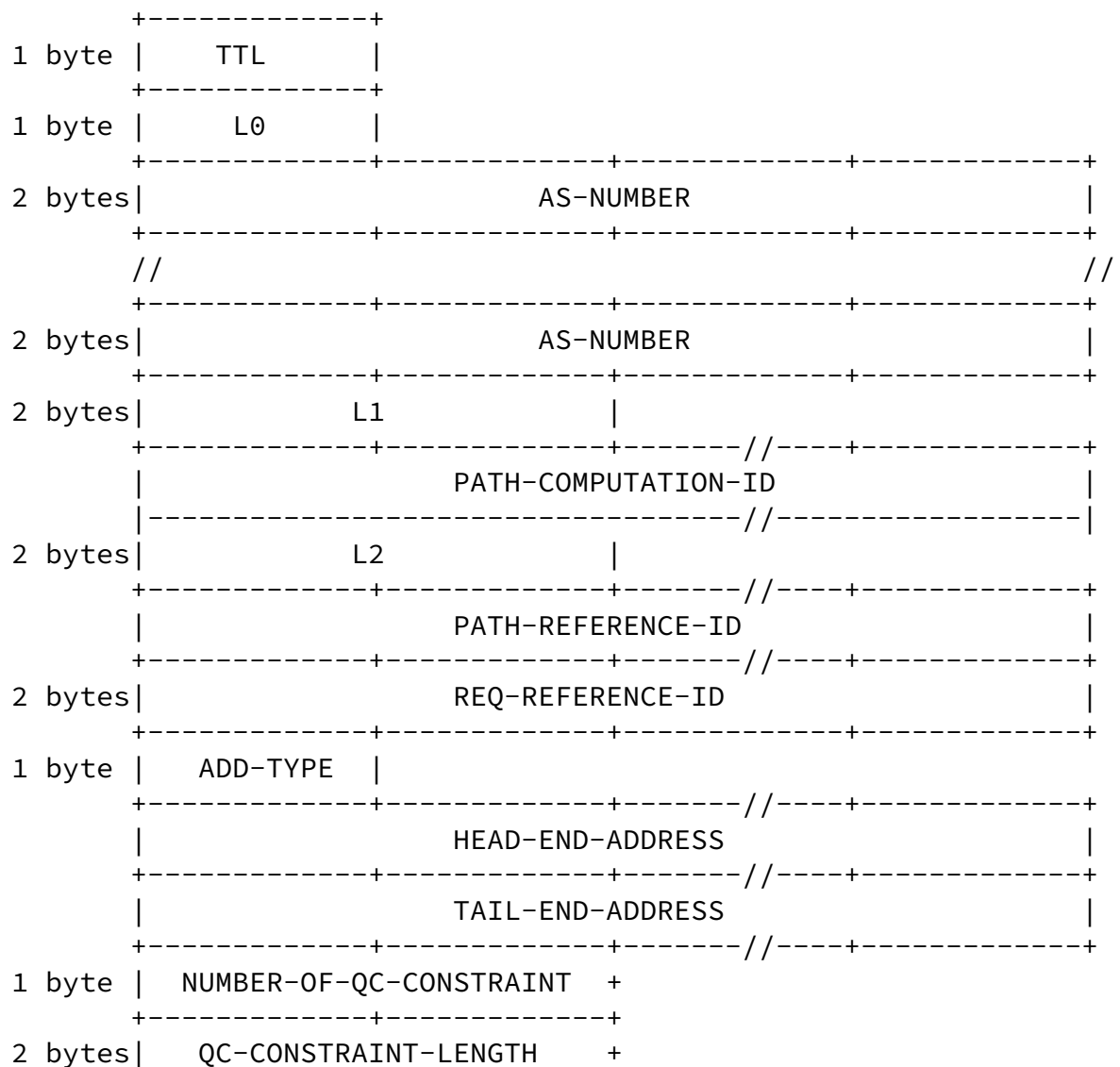


#### Error-Code:

- 1 = Shutting Down
- 2 = Bad Message Format
- 3 = Incorrect identifier
- 4 = Unable to process
- 5 = Protocol error

### [7.5.](#) REQUEST message

The Request message is sent by the PCC for computing and inter-domain path.



1 byte		QOS-CLASS-IDENTIFIER	+	
1 byte		QOS-INFO-CODE	+	QOS-INFO-SUB-CODE

2 bytes		QOS-INFO-VALUE	
		QOS-INFO-CODE	+
		QOS-INFO-SUB-CODE	
		QOS-INFO-VALUE	
		QOS-INFO-CODE	+
		QOS-INFO-SUB-CODE	
		QOS-INFO-VALUE	

- o TTL: is the maximum number of ASs that can be crossed by the path. This field is decremented by one each time a PCS issues a request.
- o L0: is a 1-byte length field. It represents the number of ASs that have already been crossed.
- o AS-NUMBER: is a 2 bytes length field representing an AS number. The first AS-NUMBER value of the list is the AS-NUMBER of the PCC that initialized a path computation.
- o L1: is the length in bytes of the PATH-COMPUTATION-ID. Size of this field is 2 bytes.
- o PATH-COMPUTATION-ID: is a globally unique value that identifies a path computation occurrence. It is a variable-length field. It is suggested, at least in this first specification, that this identifier is computed using the PCSID of the domain, concatenated with the date and an identifier that will be computed by the first requesting PCC each time a request will have to be issued. Across PCC reboots, this identifier must be unique. This PATH-COMPUTATION-ID will be replicated in all subsequent request initiated by the PCEs along the path.
- o L2: is the length in bytes of the PATH-REFERENCE-ID. Size of this field is 2 bytes.

- o PATH-REFERENCE-ID: is a variable-length field. It is an identifier that represents a pre-agreement between the head and the tail-end domain that allows the PCS from the terminating domain to accept or reject the path computation request.
- o REQ-REFERENCE-ID: is a 2 bytes length field representing an unsigned integer. This field is used to identify the REQUEST. It allows making the difference between several REQ issued for different path computation (but same PATH-COMPUTATION-ID) between two neighbor ASs interconnected via multiple links.

- o ADD-TYPE: indicates the nature of the IP addresses of the tail-end and head-end termination:
  - o 1 = IPv4
  - o 2 = IPv6
- o HEAD-END-ADDRESS: is the head-end address of the future LSP represented in the form HEAD-END@PCSID. This is a couple of IPv4 or IPv6 address. The first address of the couple identifies a loopback or an interface address of a network element, the second element is the PCSID of the domain owning the previous address.
- o TAIL-END-ADDRESS: is the tail-end address of the LSP represented in the form TAIL-END@PCSID. This is a couple of IPv4 or IPv6 address. The first address of the couple identifies a loopback or an interface address of a network element, the second element is the PCSID of the domain owning the previous address.

These above parameters MUST be present in each REQUEST and in the same order.

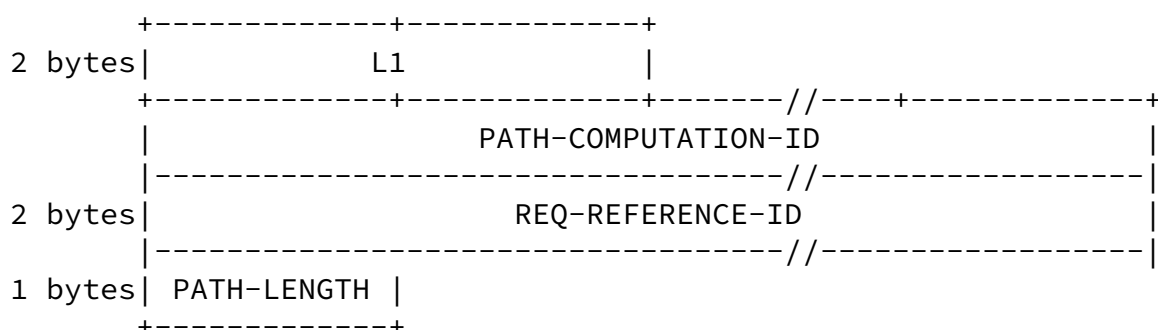
- o NUMBER-OF-QC-CONSTRAINT: represents the number of QoS class constraints the PCS must take into account when computing a path. A QoS class constraint contains a QoS-Class-Identifier (like a DSCP value) followed by additional constraints. The size of this field is 1 byte. This field is not really necessary in this first version of the specification but it could become useful if additional path constraints were included in the request.

- o QC-CONSTRAINT-LENGTH: is the number in byte of the QoS-Class-Constraint that follows. The size of this field is 2 bytes.
- o QOS-CLASS-IDENTIFIER: identifies a particular QoS-class. The size of the field is 1 byte.
- o QOS-INFO-CODE: this field identifies the type of QoS information. The size of this field is 4 bits. This code could be:
  - o (0) Reserved
  - o (1) Packet rate
  - o (2) One-way delay metric
  - o (3) Inter-packet delay variation
- o QOS-INFO-SUB-CODE: this field carries the sub-type of the QoS information. The following sub-types have been identified. The size of this field is 4 bits.
  - o (0) None
  - o (1) Reserved rate
  - o (2) Available rate

- o (3) Loss rate
- o (4) Minimum one-way delay
- o (5) Maximum one-way delay
- o (6) Average one-way delay
- o QOS-INFO-VALUE: this field indicates the value of the QoS information. These are the constraints that the PCE should respect. The corresponding units depend on the instantiation of the QoS information code.

#### [7.6.](#) RESPONSE-PATH message

This message is sent back when a path has been successfully computed.



```

1 byte |  ADD-TYPE  |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     NEXT-HOP                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
//                                                                                      //
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     NEXT-HOP                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
8 bytes|  VALIDITY-DATE  +
+-----+-----+-----+-----+-----+-----+-----+-----+
1 byte |  NUMBER-OF-QC-CONSTRAINT  +
+-----+-----+-----+-----+-----+-----+-----+-----+
2 bytes|  QC-CONSTRAINT-LENGTH  +
+-----+-----+-----+-----+-----+-----+-----+-----+
1 byte |  QOS-CLASS-IDENTIFIER  +
+-----+-----+-----+-----+-----+-----+-----+-----+
1 byte |  QOS-INFO-CODE          +  QOS-INFO-SUB-CODE          |
+-----+-----+-----+-----+-----+-----+-----+-----+
2 bytes|                                     QOS-INFO-VALUE          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|  QOS-INFO-CODE          +  QOS-INFO-SUB-CODE          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     QOS-INFO-VALUE          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|  QOS-INFO-CODE          +  QOS-INFO-SUB-CODE          |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     QOS-INFO-VALUE          |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

- o L1: is the length in bytes of the PATH-COMPUTATION-ID. Size of this field is 2 bytes.
- o PATH-COMPUTATION-ID: is a globally unique value that identifies a path computation occurrence. It is a variable-length field. This value of this identifier MUST be the same as the one provided by the REQUEST.
- o REQ-REFERENCE-ID: is a 2 bytes length field representing an unsigned integer. This field is used to reference the initial REQUEST.
- o PATH-LENGTH: indicates the number of next hops that form the path. The size of this field is 1 byte.

- o ADD-TYPE: indicates the nature of the IP addresses in the PATH.  
The size of this field is 1 byte.
  - o 1 = IPv4
  - o 2 = IPv6
- o NEXT-HOP: IP address of a next hop that is part of the computed path. Size of this field depends on the nature of the IP address.
- o VALIDITY-DATE: represents the GMT date after which the computed path returned will not be valid. The size of this field is 8 bytes.

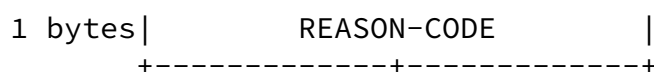
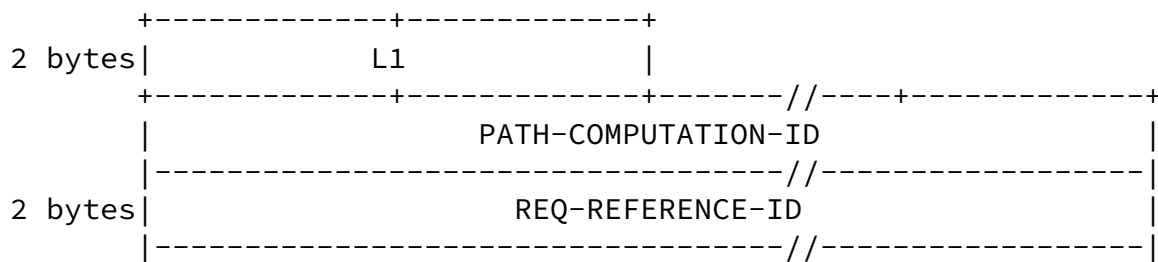
These above parameters MUST be present in each RESPONSE and in the same order.

The other parameters have the same meaning than for the REQUEST except:

- o QOS-INFO-VALUE: represents the QoS guarantees of the path, for this particular QOS-INFO-CODE parameter (delay, jitter,à) between the ingress ASBR of the responding PCS domain and the tail-end of the path.

### 7.7. PATH-ERROR message

This message is sent back when a path could not be computed.



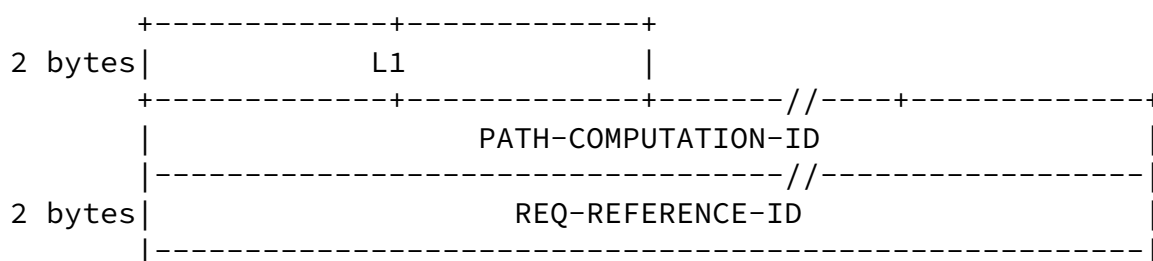
- o L1: is the length in bytes of the PATH-COMPUTATION-ID. Size of this field is 2 bytes.
- o PATH-COMPUTATION-ID: is a globally unique value that identifies

a path computation occurrence. It is a variable-length field. This identifier MUST be the same as the one provided by the REQUEST.

- o REQ-REFERENCE-ID: is a 2 bytes length field representing an unsigned integer. This field is used to reference the initial REQUEST.
- o REASON-CODE: indicate the reason of the failure. Identified failure are:
  - 1 = No resource available
  - 2 = Path reference error
  - 3 = Abnormal termination
  - 4 = PATH-COMPUTATION-ID already used
  - 5 = TTL expired
  - 6 = Loop detected
  - 7 = Request already handled

#### 7.8. CANCEL message

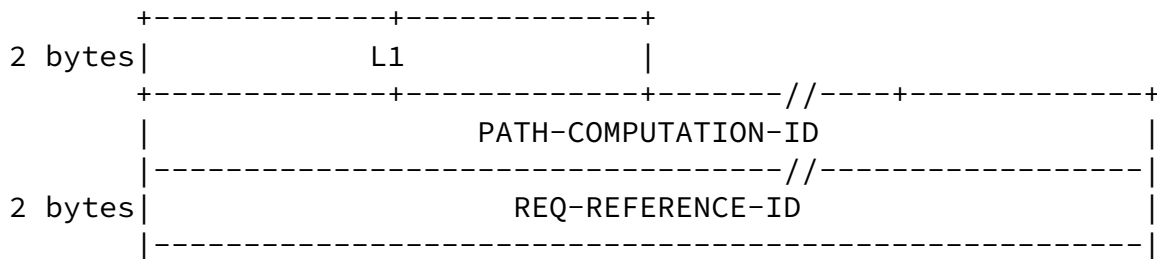
This message is sent by a PCC or a PCS when a path computation must be cancelled.



- o L1: is the length in bytes of the PATH-COMPUTATION-ID. Size of this field is 2 bytes.
- o PATH-COMPUTATION-ID: is a globally unique value that identifies a path computation occurrence. It is a variable-length field. This identifier MUST be the same as the one provided by the REQUEST.
- o REQ-REFERENCE-ID: is a 2 bytes length field representing an unsigned integer. This field is used to reference the initial REQUEST.

### [7.9.](#) ACKNOWLEDGE message

This message is sent by a PCC to a PCS to confirm the reservation of the path. This feature is particularly used when a PCC launches multiple REQUESTs during its path computation phase.



- o L1: is the length in bytes of the PATH-COMPUTATION-ID. Size of this field is 2 bytes.
- o PATH-COMPUTATION-ID: is globally unique value that identifies a path computation occurrence. It is a variable-length field. This identifier MUST be the same as the one provided by the REQUEST.
- o REQ-REFERENCE-ID: is a 2 bytes length field representing an unsigned integer. This field is used to reference the initial REQUEST.

### [7.10.](#) KEEPALIVE message (KA)

Message exchanged between two PCEs to maintain TCP session when no other messages are exchanged.

This message has no argument.

## [8.](#) Exchange of PCP messages

### [8.1.](#) Communication

The PCP protocol uses a single persistent TCP connection between a PCC and a remote PCS. One PCE server implementation per server MUST listen on a well-known TCP port number (to be defined). The PCC is responsible for initiating the TCP connection to the PCS. The location of the remote PCS is deduced and retrieved from the management plane blocks during the path computation process or at PCS boot via the SLS management block. PCE can have crossed communication; some are acting as a client role, others as a server role.

### [8.2.](#) OPEN (OPN)

An OPN message MUST be sent before any other message exchange. As part of the open message, the PCC provide its PCSID, which allows the



Internet Draft                      PCE Communication protocol                      October 2004

retrieve the client context near its management plane. Only one OPN message can be issued at a time.

If the PCS receives malformed message it MUST close the session using the appropriate error code.

### [8.3.](#) ACCEPT (ACP)

The ACP message is used to positively respond to the OPN message from the PCC. This message will return to the PCC a timer value object indicating the maximum time interval between keep-alive messages. If the PCS refuses the PCC open message, it will instead issue a CLOSE message.

The KA-Timer corresponds to maximum acceptable intermediate time between the generation of messages by the PCEs. The timer value is determined by the PCS and is specified in seconds.

### [8.4.](#) CLOSE (CLO)

The CLOSE message can be issued by either the PCC or the PCS to notify the other that it is no longer available.

The Error code is included to describe the reason for the close.

When issuing a CLOSE both the PCC and the PCS MUST delete all the internal states related to this PCP session. Additionally, all pending requests MUST be explicitly cancelled using a CCL message in order to free as much as possible all pending resources reservations and/or pre-contracts that could have been established.

### [8.5.](#) REQUEST (REQ)

A request is issued by a PCC when it has found a potential path toward the target final destination. This request can be issued as a consequence of a request received from another domain it has agreement with or from its own service management plane.

When the service request comes from a remote PCC, the server achieves the following tasks:

- (0) If the receiving TTL is zero the PCS MUST discard the request.

The receiving PCS, decrements by one the received TTL value. If the TTL is equal to zero, the request is rejected if the PCS is not the last PCS in the chain. In addition the PCS examines the AS-PATH included in the received REQ and reject it if it finds its own AS number in the list. This mechanism allows avoiding possible loops when a limited set of QoS constraints are provided in the request.

- (1) It checks if the PATH-COMPUTATION-ID of the received REQ is already associated to a pre-contract or contract. If this is the case, it returns a PATH-ERROR message with a reason-code = 4. It checks if the PATH-COMPUTATION-ID and the REQ-REFERENCE-ID of the received REQ are already associated to a pre-reservation record. If a pre-reservation is found, it returns a PATH-ERROR message with a reason-code = 4.
  - (2) It considers the HEAD-END-ADDRESS and the TAIL-END-ADDRESS parameters present in the request. The HEAD-END-ADDRESS MUST indicate a valid entry point in its domain. If not, the PCS returns a PATH-ERROR with an appropriate reason value.
  - (3) Then it extracts the PCSID from the TAIL-END-ADDRESS and parses the QoS constraints provided at part of the request message. It has thus identified all QoS-class required together with their associated QoS constraints.
  - (4) The PCS achieves some policing and verifies that the request constraints will not exceed the resources negotiated in the pSLS. If resources are exceeded, the PCS returns a PATH-ERROR message.
  - (5) If the PCS recognizes its own PCSID in the TAIL-END-ADDRESS, it considers the PATH-REFERENCE-ID otherwise it jumps to step (6). If this identifier is known from its management plane, the request is accepted and processing continues on (51). Otherwise the PCS returns a PATH-ERROR message with a reason-code = 2.
- (51) The PCS computes an intra-domain path and verifies the availability of the resources along this internal path. If available, the PCS interacts with its management plane and create a contract, which triggers the administrative reservation of the resources. When interacting with the management blocks,

the PCS MUST provide all information necessary to identify the sub-path it selected. In particular it MUST provide the PATH-COMPUTATION-ID, the ingress point ASBR address used in its domain and the termination point in its domain. The PCS sends a RESPONSE-PATH message back to the requesting PCC. If resources are not available a PATH-ERROR message is generated.

- (6) It then queries the dynamic inter-domain traffic-engineering block with the retrieved PCSID and the list of requested QoS-classes. The dynamic inter-domain TE block returns the available BGP announcements. The PCS then verifies whether it can find a next-hop ASBR, which announces the PCSID within the requested QoS-class. If cannot find it the procedure stops and a PATH-ERROR message is returned back to the requesting entity with an appropriate reason-code value.

- (7) If one or several next-hops are found, the PCS examines the QoS performance guarantees of the announcements and compare the values with those requested in the request. If it doesn't understand one of the requested QoS constraints, PATH-ERROR message is sent back. Otherwise, QoS constraints are successively compared to those received from q-BGP. All next-hops propagating the set of announcements satisfying the required QoS constraints are kept. The others are left on side.
- (8) For each possible next hop ASBR the PCS checks is there are enough available resources available at the domain boundaries. In particular if some bandwidth guarantees are required the PCS checks if the administrative maximum bandwidth agreed during the pSLS negotiation phase will not be exceeded. If resources are not available the ASBR is left on side and the next ASBR in the list is considered. If resources are available, the PCS pre-reserves the corresponding resources near the management plane. At this stage, the management plane doesn't create any contract since we are not sure that an end-to-end path exists. This pre-reservation can be taken into account by the PCS for subsequent requests. It can use it as a lock and delay the incoming requests or introduce the pre-reservations in its resource availability computation according to the local policy enforced. When interacting with the management blocks, the PCS must

provide all information necessary to identify the sub-path it selected. In particular it must provide the PATH-COMPUTATION-ID, the ingress point address of its domain and the ingress point address of the next domain. This latter information can be used by the management plane to identify the upstream and downstream involved domains.

- o (81) The PCS computes an intra-domain path and verifies the availability of the resources along this internal path. If resources are available, the sub-path is valid and the PCE forms a new REQUEST message which is sent to the PCS of the remote domain owning the next-hop ASBR. It adds its own AS number to the existing list. If internal resources are not available, the PCS discards the pre-reservation and considers the next hop ASBR in the list. When building the request the PCC keeps the PATH-COMPUTATION-ID, the PATH-REFERENCE-ID, the TAIL-END-ADDRESS unchanged. The initial HEAD-END-ADDRESS is replaced by the address of the ingress next-hop ASBR identified during the path computation. The QoS constraints characteristics are modified in order to take into account the QoS performance guarantees provided by the domain.
- (9) If QoS constraints cannot be satisfied for any of the ASBR, the PCS returns a PATH-ERROR message.

Note that it is quite possible that several next hops ASBR can satisfy the requested constraints. In such a case the PCS can process one next-hop ASBR at a time or several in parallel. For one incoming request, there can be multiple simultaneous outgoing requests towards different PCS. If several requests are sent toward the same neighbor, for a same PATH-COMPUTATION-ID, the REQ-REFERENCE-ID must be different. Nevertheless, this feature can lead to scalability issues and needs further investigations.

#### [8.6](#). RESPONSE (RSP)

A RESPONSE message is sent by a PCS in response to a request issued by a PCC. RSP messages are sent back when a valid end-to-end path has been computed. The RSP message MUST be initiated by the tail-end domain.

When a valid end-to-end path has been computed, the PCS of the last domain on the path, forms a RSP message. It first inserts the original PATH-COMPUTATION-ID. Then it forms a path argument that

MUST contains the IP address of the tail-end LSP and the IP address interface of the ingress ASBR supporting that path. It MAY insert between these two extremities, the IP address of additional hops. It MAY also indicates the date after which the path will not be valid anymore because administratively reserved resources will have been relaxed. Then, it MUST indicate QoS guarantees it provides between the ingress ASBR and the tail-end address of the LSP. The RSP message is then sent to the requesting PCC.

On receipt, the PCC adds its own intra-domain sub-path to the list. It does not indicate the next-hop ASBR since this latter has already been inserted by the downstream PCS. This sub-path can be a strict or loose description. It also modifies the QoS guarantee parameters so that they reflect the QoS guarantees it can provide for its part of the path. This is achieved in the same way than for the request, but it is an "addition" operation if we consider the delay, for example. The VALIDITY-DATE MUST modified so that the value indicates now the smaller date between the date received in the RSP message and the date reported by the management plane.

If the PCC sent multiple REQUEST messages in parallel, it MAY wait for a RSP or ERR message for all the requests it sent. If the PCC got multiple RSP messages it MUST select only one and inform the un-selected PCS that they can cancel their reservation. It forms CANCEL messages, sends them to the appropriate PCS and cancels its own pre-reservation for the corresponding requests. If the PCC doesn't wish to wait for a reply, it can send a CANCEL message at any time.

The PCS can send the consolidated RES message to the requesting PCC after sending ACK message to the PCS it decided to keep in the path.

#### [8.7.](#) ACKNOWLEDGE (ACK)

The ACK message is used by PCS to confirm to its management plane that the resources needed for the path referenced by PATH-COMPUTATION-ID present in the message need to be reserved. It allows the management plane to create a contract based on information previously stores by the PCS during the computation phase. If no ACK is received, no contract is created and the negotiation at the management level will fail. If for some reasons, no ACK were received, the VALIDITY-DATE would be used and the administrative pre-reservation automatically removed for that path. ACK messages are only accepted if they arrive after the server has issued a RSP

otherwise they are ignored.

#### [8.8. CANCEL \(CCL\)](#)

A CANCEL message can be sent by PCC and PCS. CCL messages can be generated during the normal path computation cycle but also in case of an abnormal termination of a PCE to PCE communication.

If a PCE, acting as a server for the PCP session, received a CCL message from the PCC, it MUST form new CCL messages and forward a CCL message to each PCS to which it sent a REQ for which it did not received any positive or negative reply. Once this has been achieved it MUST delete all its internal states referencing the PATH-COMPUTATION-ID indicated in the message. If the PCE has no pending request concerning this PATH-COMPUTATION-ID, it can optionally query its management plane to retrieve a possible existing contract referenced by this PATH-COMPUTATION-ID and delete it. Just before deleting this contract, it can form a new CCL message and forward it to the next PCS in the path. If it does not, the VALIDITY-DATE will be applied.

The same procedure applies if the PCE server detects a communication problem with one of its PCC. In that case, the PCS issues CCL messages for all pending request received from this PCC.

When a PCE, acting as a client for the PCP session, received a CCL message from a PCE server, this indicates that a PCS along the path towards the target destination has experienced communication problems leading to close a PCP communication. In such a case, each PCC cancels all the internal states referencing this PATH-COMPUTATION-ID and forward this indication to the upstream client PCS up to the initial requestor.

#### [9. State diagram](#)

TBD.

#### [10. Security Considerations](#)

PCP is a communication protocol that is used between two PCEs. No security mechanisms are defined in this PCP specification. It is

recommended that a security protocol like IPSec or TLS MUST be activated in order to protect PCP sessions.

## 11. References

- [RFC3667] Bradner, S., "IETF Rights in Contributions", [RFC 3667](#), February 2004
- [RFC3668] Bradner, S., "Intellectual Property Rights in IETF Technology", [RFC 3668](#), February 2004
- [INTERAREA-REQ] Le Roux, J., Vasseur, JP, Boyle, J., "Requirements for Support of Inter-Area and Inter-AS MPLS Traffic Engineering", [draft-ietf-tewg-interarea-mpls-te-req-00.txt](#), March 2004 (work in progress)
- [INTERAS-REQ] Zhang, R., Vasseur, JP., et. al., "MPLS Inter-AS Traffic Engineering requirements", [draft-ietf-tewg-interas-mpls-te-req-06.txt](#), January 2004 (work in progress).
- [PCE-ARCH] Ash, J., Farrel, A., Vasseur, JP., " Path Computation Element (PCE) Architecture", [draft-ash-pce-architecture-00.txt](#), September 2004
- [PCE-FWK] Farrel, A., Vasseur, JP., Ayyangar, A., "A Framework for Inter-Domain MPLS Traffic Engineering", [draft-ietf-ccamp-inter-domain-framework-00.txt](#), August 2004
- [INTERAS-PCE] Boucadair, M., Morand, P., "A Solution for providing inter-AS QoS tunnels", [draft-mescal-pce-interas-00.txt](#), October 2004
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997
- [PCE-DISCOVERY] Boucadair, M., Morand, P., "PCE Discovery via Border Gateway Protocol", [draft-mescal-pce-discovery-00.txt](#), October 2004
- [RFC2401] Atkinson R., "Security Architecture for the Internet Protocol", [RFC 2401](#), August 1998.
- [RFC2246] Dierks T., Allen C., " The TLS Protocol", [RFC 2246](#), January 1999

## 12. Acknowledgments

The authors would also like to thank all the partners of the Mescal (Management of End-to-End Quality of Service Across the Internet At Large, <http://www.mescal.org>) project for the fruitful discussions.

### 13. Author's Addresses

Mohamed Boucadair  
France Telecom R & D  
42, rue des Coutures  
BP 6243  
14066 Caen Cedex 4  
France  
Phone: +33 2 31 75 92 31  
Email: mohamed.boucadair@francetelecom.com

Pierrick Morand  
France Telecom R & D  
42, rue des Coutures  
BP 6243  
14066 Caen Cedex 4  
France  
Email: pierick.morand@francetelecom.com

### Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org).



## Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS

Boucadair (Ed.) Standards Track - Expires April 2005

[Page 21]

---

Internet Draft

PCE Communication protocol

October 2004

OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY 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.

## Copyright Statement

Copyright (C) The Internet Society (2004). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

