

PCE Working Group
IETF Internet Draft
Document: [draft-boucadair-pce-interas-01.txt](#)
Proposed Status: Informational
Expires: November 2005

M. Boucadair (Ed.)
P. Morand (Ed.)
France Telecom R&D
May 2005

A Solution for providing inter-AS MPLS-based QoS tunnels
< [draft-boucadair-pce-interas-01.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/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 November 2005.

Abstract

This document describes a solution for providing inter-AS MPLS-based Quality of Service (QoS) tunnels. This solution makes use of Path Computation Elements (PCE) as a means to compute inter-domain constraint-based paths. Service considerations and agreements between two IP Network Providers (INP) implementing this solution are also described.

Copyright Notice

Internet Draft A Solution for
providing inter-AS MPLS-based QoS tunnels

Copyright (C) The Internet Society. (2005)

Table of Contents

1.	Contributors.....	2
2.	Changes since last version:.....	2
3.	Terminology.....	2
4.	Introduction.....	3
4.1.	General.....	3
4.2.	Assumptions.....	4
4.3.	Draft structure.....	5
5.	Conventions used in this document.....	5
6.	Inter-AS PCE model.....	5
7.	Inter-provider Service Considerations.....	7
8.	Path Computation Element functions.....	8
9.	PCE discovery.....	10
10.	PCE to PCE Communication.....	10
11.	Routing considerations.....	11
11.1.	Assumptions.....	11
11.2.	Finding inter-domain LSP paths.....	11
12.	Communication between PCE and LSR/LER for initiating LSP set-up.....	13
13.	Advanced features.....	13
13.1.	Exclusion of specific ASs from the path.....	13
13.2.	Feedback.....	13
14.	Security Considerations.....	14
15.	References.....	14
16.	Acknowledgments.....	14
17.	Author's Addresses.....	15

1. Contributors

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

2. Changes since last version:

The main changes occurred in this version are:

- o Add new contributor to the draft
- o Rewording of several sections of the draft

3. Terminology

This memo makes use of the following terms:

Boucadair (Ed.) Informational - Expires November 2005

[Page 2]

Internet Draft

A Solution for
providing inter-AS MPLS-based QoS tunnels

May 2005

- o Path Computation Element (PCE): an entity that is responsible for computing/finding inter/intra domain MPLS tunnels (LSPs). This entity can simultaneously act as client and a server. Several PCEs can be deployed in a given AS.
- 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: the service employing a PCE-based system as the underlying infrastructure for creating e.g., inter-domain QoS VPNs.
- o High-level service customer: the customer that subscribes to a High-level service.
- o pSLS(provider SLS): A provider level SLS, which is established for specific QoS class between two Internet Network Providers (INP) for exchanging traffic in the Internet with the purpose to expand the geographical span of their offered services. pSLSs are meant to support aggregate traffic and they are assumed to exist prior to any agreements with end customers.
- o SLS Management: a service level management system, which includes service ordering (i.e for establishing contracts between peer providers) and service invocation (i.e for 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 for its route selection process.

o Domain: within this document it denotes an Autonomous System.

[4.](#) Introduction

[4.1.](#) General

The level of Quality of Service (QoS) guarantees offered by INPs using a pure IP-based traffic engineering (TE) solution, other than overbooking, is not yet satisfactory for all corporate business services, for which strong guarantees must be provided. For this type of customers, hard QoS performance and bandwidth guarantees are considered as the major requirements.

Boucadair (Ed.) Informational - Expires November 2005

[Page 3]

Internet Draft

A Solution for
providing inter-AS MPLS-based QoS tunnels

May 2005

Currently, these requirements can be satisfied within a single domain or across several interconnected domains managed only by a single INP. However, it becomes very challenging when these domains are managed by different INPs. Each INP defines and deploys its own QoS policy within the scope of its domain(s), utilizes its proprietary TE functions, etc.

Providing QoS-based services within the scope of the Internet requires the collaboration among INPs in order to offer this type of services. This document aims at proposing a solution that will facilitate the introduction of such services in an Inter-provider environment. Service considerations are also taken into account.

[4.2.](#) Assumptions

In this document, we assume the following:

- o An AS can announce a given prefix in several QoS planes; each of these QoS planes being identified across inter-domain links by a unique DiffServ Code Point (DSCP);
- o Each announcement, except best effort ones, contains values of a set of QoS parameters that characterizes the likely end-to-end QoS to be experienced for reaching a given prefix (we call this

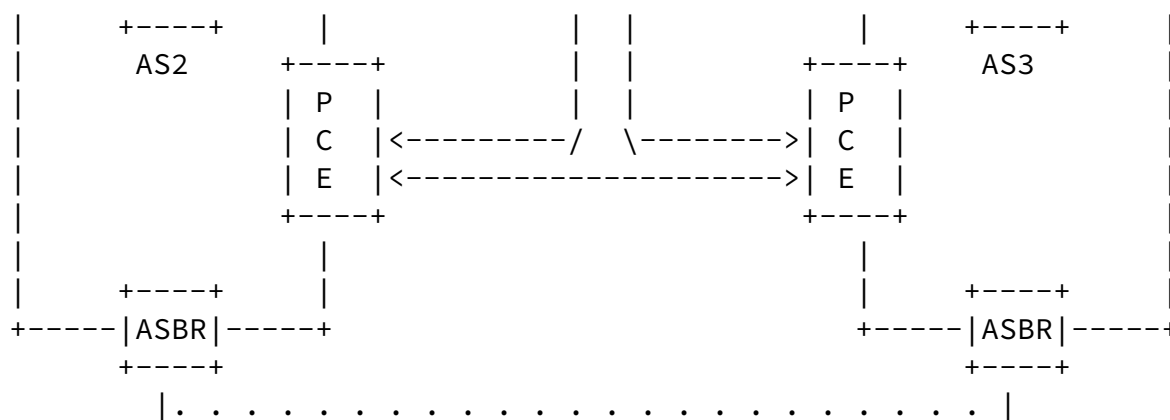
- end-to-end QoS as aggregated QoS);
- o The way aggregated QoS values are computed is out of the scope of this document;
 - o Adjacent ASs agree on the DSCP values to use in order to signal a given QoS class at inter-domain links (we call these DSCP values inter-domain DSCPs);
 - o Every AS has the freedom to bind an inter-domain DSCP to a local DSCP within its domain, which identifies its local QoS class for signalling a QoS planes in its domain;
 - o An AS supporting the inter-domain MPLS-based QoS tunnels service, owns a Path Computation Service Identifier(s) (PCSID), which is a routable IP address. This IP address is not necessarily the IP address(es) of the PCE(s) of the domain. These PCSIDs are announced per QoS plane basis, by an inter-domain routing protocol, together with the plane's aggregated QoS values;
 - o ASs can discover other ASs supporting the inter-domain MPLS-based QoS tunnels service by receiving inter-domain routing protocol announcements. These announcements provide an AS with

the end-to-end QoS characteristics of the path towards any prefix of the remote AS owning the PCSID.

[4.3.](#) Draft structure

The structure of this document is as follows:

- o [Section 5](#) describes the inter-AS PCE model.
- o [Section 6](#) discusses the service considerations.
- o [Section 7](#) highlights PCE functions.
- o [Section 8](#) explains the PCE discovery function.
- o [Section 9](#) gives an overview of the PCP protocol that is used for communication between PCEs.



...: Peering links
 ---: inter-PCE session

Figure 1: Inter-AS PCE scenario

In the above figure, we assume that each domain operates a set of QoS-classes. Each INP deploys its own local QoS-classes with specific QoS characteristics. QoS-classes in a domain have not necessarily the same QoS characteristics with QoS-classes in the other domains. We also assume that service level QoS-based contracts (pSLSs) have been established between adjacent domains. BGP is running between these adjacent ASs. Each AS learns, the set of destinations its peer domains can reach, together with end-to-end QoS performance characteristics.

When a pSLS is established, the domains exchange their respective PCE information (name, IP address, identifiers, authentication information, etc.).

In order to create an inter-domain constraint-based LSP, the domain which requests the establishment of the LSP asks its PCE to compute an inter-domain path satisfying a set of constraints, (for example bandwidth guarantee per QoS-Class, maximum end-to-end delay, etc.)

The first PCE selects one possible path among the set of alternatives and identifies the next-hop domain. It then verifies that appropriate resources are available in its own domain and sets up administrative pre-reservations in the management system of its domain. Then it

contacts the next hop PCE, requesting a path computation between the next hop AS Border Router (ASBR) and the termination address of the inter-domain LSP. This second PCE performs the same computation as

However, it is difficult to establish such a contract in advance especially when the LSP path is not known in advance. Thus, the sequence of operation for establishing an LSP should be:

- o Compute inter-domain path candidate(s);
- o Select and request an inter-domain path for this particular LSP using information returned by the PCE,
- o Establish the LSP once final negotiation terms have been agreed end-to-end between PCEs of adjacent domains.

The establishment of this cascade of negotiations can be difficult to achieve and can take some time. In particular, the risk is not negligible that the resources that were available when the PCE performed the path computation are no longer available along the path when the cascaded negotiation terms are agreed, because others LSPs have used the corresponding resources.

In order to solve this issue it is necessary that the PCE of each domain makes an administrative reservation of the corresponding resources and indicates the characteristics of the path. This information is registered by the management plane, which triggers in parallel the creation of a provisional contract referencing the technical characteristics of the future LSP. Subsequent path computation requests may be impacted because the management plane removes these resources from the available overall network resources. This provisional contract is valid for a limited time period, which is the minimum of time periods each specified and reported by a domain along the path. If time period expires, the provisional contract can be removed from the management systems, and related administrative network resources have to be informed.

It is the responsibility of the management plane of each domain to cooperate in agreeing the exact financial terms and additional clauses of this contract, including its duration. Each domain knows the entry and the exit point of the LSP within its own domain and consequently knows both the upstream and downstream ASs to deal with. This validation procedure SHOULD ideally be automated to speed up the process and could integrate pricing negotiation. The way that the other blocks of the management plane deal with this automation is out the scope of this document.

Thus, once the pre-contract is validated, the path computed by the PCE can be provided to the head-end LSP, which effectively sets up the LSP. Note that every ingress point of each domain SHOULD activate some outsourced policy functions that would allow RSVP TE to get an agreement from the management system.

8. Path Computation Element functions

Internet Draft

A Solution for
providing inter-AS MPLS-based QoS tunnels

May 2005

The main function provided by a PCE is to contribute to the overall path computation by computing part of the end-to-end inter-domain path satisfying a set of constraints. The management plane could call other elementary services such as requesting a path computation for informational purposes or canceling a request in progress for instance.

The deployment and the maintenance of the LSP connectivity require cooperation of several functional entities from different planes. Within this document, the PCE is only responsible for computing an inter-domain constraint-based path. The implementation of the service (whether it is automated or not) and the creation of inter-domain LSP results from the cooperation of functional blocks, control plane blocks and data plane blocks aren't described in this document.

The PCE does not itself trigger the establishment of any inter-domain LSP, but provides inter-domain paths, if they are available. Unlike the management plane, it is not aware of business considerations. A PCE entity provides an interface used by the service management block to request a path computation. It communicates with other remote PCE thanks to the PCP protocol and requests additional services from other functional blocks as illustrated in the figure below:

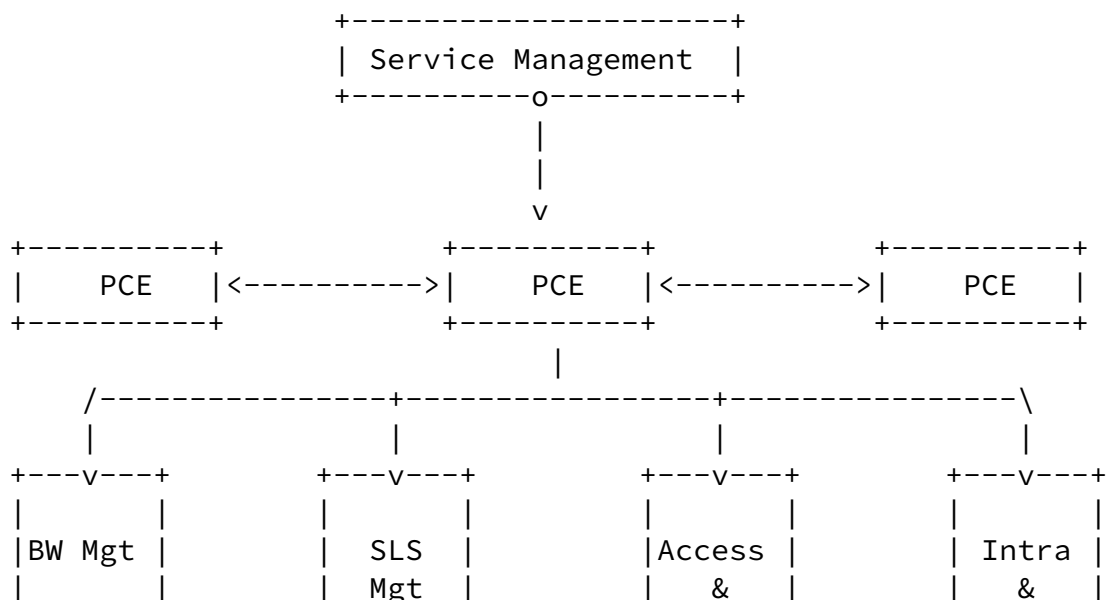




Figure 3: PCE Interactions

The PCE interacts with the dynamic routing processes to retrieve routing information that is used to compute an inter-domain path satisfying the expressed constraints. An interface **MUST** be made available to the PCE so that it can access routing information. Note that both intra and inter-domain routes **MUST** be made available to the PCE.

In addition, for access control and authorization purposes, the PCE **MUST** be provided with access to the list of other PCEs from which it will accept requests. This list is updated each time new pSLSs are negotiated by the INP.

9. PCE discovery

Within this document, we assume that during the service negotiation phase the peers exchange the IP addresses of their respective PCE(s). This information is stored in the SLS Management Systems of each INP.

As described in [[PCE-DISCOVERY](#)], instead of announcing all potential tail-end addresses in BGP, only an identifier is announced via BGP. It is called the Path Computation Service Identifier (PCSID). This particular BGP announcement is identified by a well-known community value (to be defined by IANA) and is represented by a routable IP address, which can be different from the real IP address of the PCE.

As a consequence, this particular route **SHOULD NOT** be installed in the Forwarding Information Base (FIB) since this PCSID is not necessarily the IP address of the PCE.

BGP announcements of PCSID will ease to discover the set of remote ASs supporting the inter-AS MPLS-based QoS tunnels service and associated end-to-end QoS-related information for reaching them. In order to compute a path towards a specific domain supporting this inter-AS MPLS-based QoS tunnels service, the local PCE chooses a route that serves the PCSID of that domain and extracts from the AS_PATH attribute the AS number of the next hop ASBR. Then, the local

PCE queries its SLS Management system and gets back the PCE's IP address of the next neighboring PCE to contact. Finally, the local PCE forms and forwards a path computation request to the next PCE. The process is iteratively repeated until the request reaches the PCE of the target AS identified by its PCSID.

10. PCE to PCE Communication

A PCE can act as a client (Path Computation Client, PCC) or a server (Path Computation Server, PCS). The PCC is responsible for issuing Path Computation requests. The PCS is responsible for handling requests received from PCCs.

The PCP (Path Computation Protocol) is a simple query and response protocol that is used for communication between PCE entities, i.e., PCC and PCS.

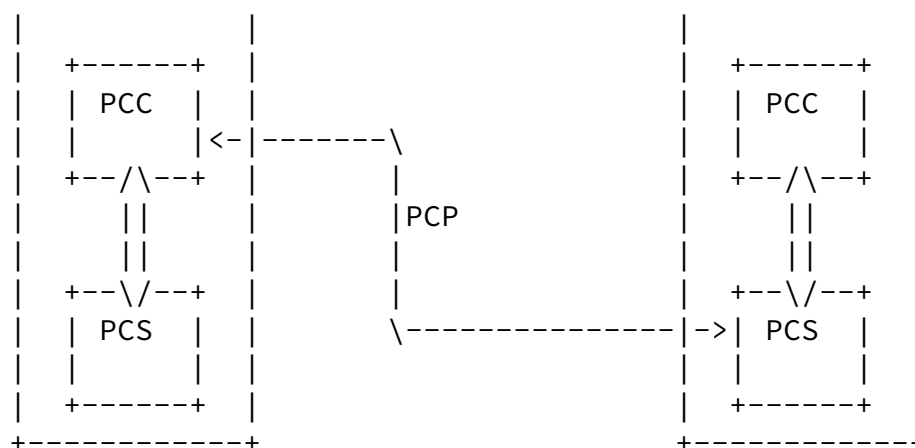


Figure 4: PCC to PCS communication

The main characteristics of the PCP protocol are:

- 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, providing the reliable exchange of messages between PCEs.

- o In its first version, PCP does not provide any message level security for authentication, message replay protection, and integrity. However, 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 current PCP protocol provides the service for supporting only a basic path computation function. In particular it does not support the service for additional path computation constraints, or provide enhanced reporting features in the case of path computation failure.

[11.](#) Routing considerations

[11.1.](#) Assumptions

We assume in this document that PCE addresses are only announced in a few QoS-Class planes. Addresses of LSR/LER interfaces could be announced in the best effort plane. This reduces the number of BGP announcements to one announcement per PCE per AS. By setting a well-known community value, we specify announcements that serve PCEs. These are not regarded as routes and are not stored in the FIBs.

[11.2.](#) Finding inter-domain LSP paths

In order to find an inter-domain path, the PCE MUST be provided with a set of attributes including the information describing head-end and tail-end of the LSP and the performance requirements for the LSP. The aforementioned information MUST include the loopback IP address of its LSR, and the IP Address of the PCE of its domain (notation is IPADDRESS@PCSID). This information MUST also include the performance guarantees required for the inter-domain constraint-based LSP. This information MAY encompass the requested QoS-class(es) so that the set of collaborating PCE can compute a path that will cross a set of domain satisfying the expressed constraints.

It can also contain, per QoS-class, additional QoS performance guarantees that the PCE must take into account. These performance guarantees include guaranteed end-to-end delay, jitter, loss rate and/or bandwidth. Note that these parameters can differ depending on

the type of requested QoS-class, and they MAY not all be present in the LSP set-up request. If included in the path computation request they MUST be taken into account by the PCE. If the PCE doesn't recognize a given QoS parameter, the PCE MUST stop its computation and MUST return an error (PCP Error Message).

When computing a path, a PCE interacts with other blocks of the management plane. In particular, it checks the availability of the resources within the boundaries of its domain. If the resources are available, and the sub-path (path between the ingress point of its domain and a potential ingress point of a given domain) conforms to the path constraints requested, it MUST inform the management plane of a pre-reservation concerning this path. This information can then be taken into account when processing other path computation requests. Once this operation succeeds, the request is propagated to the next domain PCE, which has been selected by the PCE of this domain.

Note that the performance guarantees requested MUST be updated before forwarding it to the next domain by taking into account the performance figures already observed along the computed sub-path plus the performance figures within its own domain. Therefore, PCE MUST be aware of the above performance figures of the QoS-classes.

The requesting PCE MUST use the QoS-class identifier they agreed during the pSLS negotiation phase in order to signal a given QoS-class.

If an end-to-end LSP has to be re-engineered because the associated constraints have changed in terms of QoS-class requested, bandwidth, delay, etc., a new end-to-end path needs to be computed. In order to improve its chances of finding a valid path, the requestor can specify that the path for which the request is issued will replace a previously established LSP. For doing so, the requestor can indicate the reference of the path corresponding to this LSP. A PCE can release, during the path computation process, the resources

corresponding to the former LSP, if the new path follows part of the former path. This reference is stored in the management plane of each domain and is generated by the initial requestor. This reference is globally unique.

The ability to address such additional constraints can be interesting in the case of backup LSPs, so that the PCE can compute a path along

a different route. These considerations are out of scope of this document.

12. Communication between PCE and LSR/LER for initiating LSP set-up

Communication between PCE and an LER/LSR could be achieved thanks to the use of Common Open Policy Service protocol (COPS, [[RFC2748](#)]). An RSVP client-type could be used in order to convey configuration data resulting from the computation operation executed by a PCE. Specification of RSVP configuration data is out of scope of this document.

13. Advanced features

13.1. Exclusion of specific ASs from the path

If a PCE in the chain wants to exclude particular AS(s) from the path, additional constraints (that can be expressed using the AS number of the excluded domain/s) MUST be added to the request message body and MUST be propagated downstream.

13.2. Feedback

When computing a path, the PCEs can fail to find a path for a number of reasons. These failures, in normal operations, will be mainly due to the lack of resources, or not meeting the requested QoS requirements. In such a situation, a path, which would have been the optimal path, would not be established. Identifying the domain/s where the path computation failed, together with the reasons, would be of a real added value for providers in order to improve their efficiency.

A proposal for achieving this is to rely on the Path Computation Protocol, which could be improved to return all the path alternatives which were tried but have failed. In doing so, the requesting provider would be aware of the reasons of the failure and possibly interact with that AS where the path computation failed and aborted.

The AS (where the path computation failed), faces with multiple requests, from external domains, could consider a possible

modification of some of its peering agreements based on business objectives.

14. Security Considerations

This document does not change the underlying security issues in the PCP and BGP protocols specifications. It is recommended that a security protocol like IPsec or TLS to be activated in order to protect PCP sessions.

15. 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
- [RFC2385] Heffernan, A., "Protection of BGP sessions via the TCP MD5 Signature Option", [RFC 2385](#), August 1998
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997
- [INTERAS-PCP] Boucadair M., Morand P., "Inter PCE Communication Protocol", [draft-boucadair-pcp-interas-01.txt](#), May 2005
- [PCE-DISCOVERY] Boucadair M., Morand P., "PCE discovery via Border Gateway Protocol", [draft-boucadair-pce-discovery-01.txt](#), Work in progress, May 2005
- [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
- [RFC2748] Boyle, J., Cohen, R., Durham, D., Herzog, S., Raja, R. and A. Sastry, "The COPS (Common Open Policy Service) Protocol", [RFC 2748](#), January 2000.

16. 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.

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 OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED,

Boucadair (Ed.) Informational - Expires November 2005

[Page 15]

Internet Draft

A Solution for
providing inter-AS MPLS-based QoS tunnels

May 2005

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 (2005). 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.

