

PCE Working Group
Internet Draft
Document: draft-grampin-pce-mgmt-if-00.txt
Expires: January 2006

Eduardo Grampin
Udelar
July 2005

PCE Management Interface

Status of this Memo

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 becomes aware will be disclosed, in accordance with Section 6 of 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/lid-abstracts.txt>

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

Abstract

The Path Computation Element (PCE) Architecture provide a framework to support the Constraint-Based Routing (CBR) functionality for traffic engineering systems in Multiprotocol Label Switching (MPLS) and Generalized Multiprotocol Label Switching (GMPLS) networks. The model define the PCC and PCE entities at the Control Plane, which communicate using a Request/Reply protocol. The PCE architecture enable network operators a tight control of the resource assignment by means of the administrative constraints that are included in the Traffic Engineering Database (TED), and used in the path computation process. Moreover, appropriate objective funtions assist operators in the fulfillment of the overall network objectives.

This document present a system architecture for the PCE, namely Routing and Management Agent (RMA), with a standard management interface, which permit established management frameworks to rely on

the PCE for the CBR functionality and inject network-wide policies into the PCE path computation process. Some reliability issues of the PCE Architecture are addressed in the document.

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 [i].

Table of Contents

1. Introduction.....	2
2. RMA Functional Components.....	3
2.1 Signalling Interface.....	4
2.2 IGP Interface.....	4
2.3 CBR Computation Component.....	4
2.4 Traffic Engineering DataBase Component.....	4
2.5 Management Plane Interface.....	4
3. Two Tier RMA Architecture.....	5
3.1 RMA availability and fault-tolerance.....	5
3.2 Traffic Engineering Database (TE-DB).....	6
3.3 The CBR process.....	7
4. Autodiscovery of RMAs.....	7
5. Security Considerations.....	7
6. Acknowledgments.....	7
7. References.....	7
8. Author's Addresses.....	8
9. Full Copyright Statement.....	8

1. Introduction

The Routing and Management Agent (RMA), interacts with both the Control Plane and the Management Plane, assuming the PCE role in the PCE Architecture [PCE-ARCH], as shown in Figure 1.

This entity, while enabling a Control Plane-based provisioning, can be used as a traffic engineering tool by management applications, using its interface towards the Management Plane. This interface can be used to establish cooperative relationship with other RMAs in the "multi-PCE path computation with PCE-intercommunication model" , as specified in [PCE-ARCH].

The RMA management interface is reachable in the management DCN, therefore integrating the RMA into a distributed management framework.

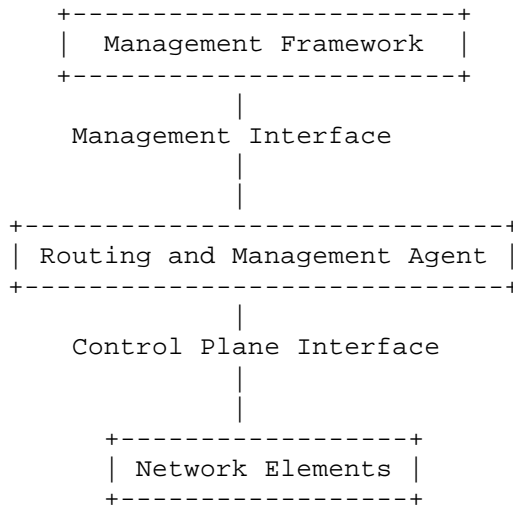


Figure 1. Routing and Management Agent (RMA)

2. RMA Functional Components

The RMA is built assuming a component-based framework, with basic scheduling and other supporting modules needed to build the described functionality. The interfaces and "core" components shown in Figure 2 are described below:

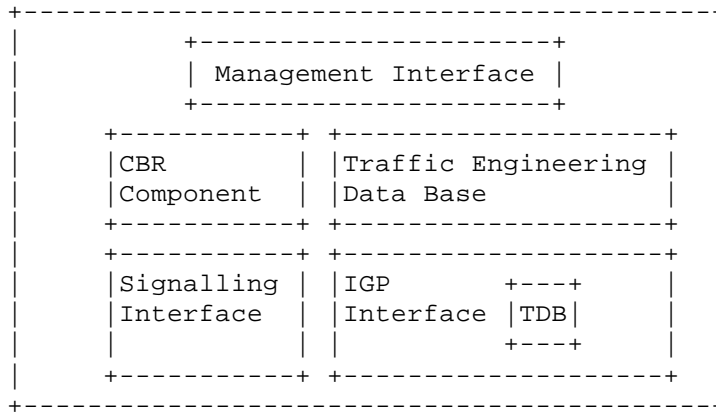


Figure 2. RMA Functional Components

2.1 Signalling Interface

This component interfaces with PCCs via a suitable Request/Reply protocol, as required by [PCE-COMM-REQ], and is out of the scope of this document.

2.2 IGP Interface

The RMA gathers network topology running an instance of the network TE IGP (OSPF-TE [RFC3630] or ISIS-TE [RFC3784]), communicating through this interface. The RMA is like any node in the routing graph, usually without forwarding capabilities, even though the PCE functionality is orthogonal to packet forwarding, as defined in [PCE-ARCH]. The goal of this component is to maintain the Topology DataBase (TDB), which in turn is used to feed the Traffic Engineering DataBase (TE-DB), the basic information source for CBR computation.

2.3 CBR Computation Component

This is the core of the RMA, which provides the intended functionality: a computation engine for Constraint-Based Routing. The component implements the needed algorithms to solve the Path Computation problem with multiple restrictions. Well-known algorithms and heuristics can be used to accomplish the intended goal, making sure that route computation time is limited (i.e. by the usage of polynomial-time CBR algorithms).

The two tier architecture with a computation cluster is presented in Section 3 is able to fulfill this objective.

2.4 Traffic Engineering DataBase Component

The TE-DB contains the up-to-date information regarding link states in the network, gathered by the IGP Interface. Additional information, like constraints and administrative policies can also be made persistent in the TE-DB. This information, which defines the TE objectives of the network, will typically come from Policy-Based Management applications.

2.5 Management Plane Interface

This component implements the interaction with management applications, which enables the RMA to be used as a traffic engineering component for high-level applications. Other possibilities of this interface include the usage of the RMA as a network-wide monitoring agent.

This interface may be based in well-known distributed component frameworks like CORBA, widely used by telecommunication operators, and/or J2EE, .NET or other usual packages in use in the enterprise and Internet environments.

The information model used in this interfaces, as well as object access granularity issues (coarse or fine grain) are out of the scope of this document, and may be further standardized.

3. Two Tier RMA Architecture

The RMA is designed as a component-based system, as detailed in the previous section. As described in other documents of the PCE WG, several issues arise regarding the design and implementation of a PCE Architecture; in particular availability and fault tolerance are of major impact in network operation. This section review some of these problems and propose some implementation guidelines.

3.1 RMA availability and fault-tolerance

A PCE centralized solution may suffer potential bottleneck problems and is a single point of failure is not careful designed.

A Two Tier Architecture comprises a reliable communication front-end with a computation back-end cluster, where the well-known High Performance Computing paradigm may be of use. This architecture is shown in Figure 3. This is a proven architecture used by Service and Content Providers for high-availability services such as web-server farms, VoD headends, E-Mail distributed servers. In the figure there are two front-end sets, one to handle Control Plane communication, and the other for the Management Plane. This separation, while not mandatory, is advisable given that very different kind of protocols need to be supported.

High availavility is given by two factors:

- a) arbitrary large set of front-end (i.e. signaling) processors and,
- b) arbitrary large set of computation nodes in the back-end cluster.

The remaining point of failure is network connectivity (both internal and external). Internal connectivity (i.e. between front and back-ends) can be protected by redundant LAN switches, while different options exist to overcome potential external connectivity failure. A straightforward (and expensive) solution is to place disjoint RMA clusters in the network, while an acceptable solution would be to have a multi-homed approach, i.e. use multiple load-sharing links. Other useful techniques include VRRP [RFC3768] and DNS Round-Robin, among others.

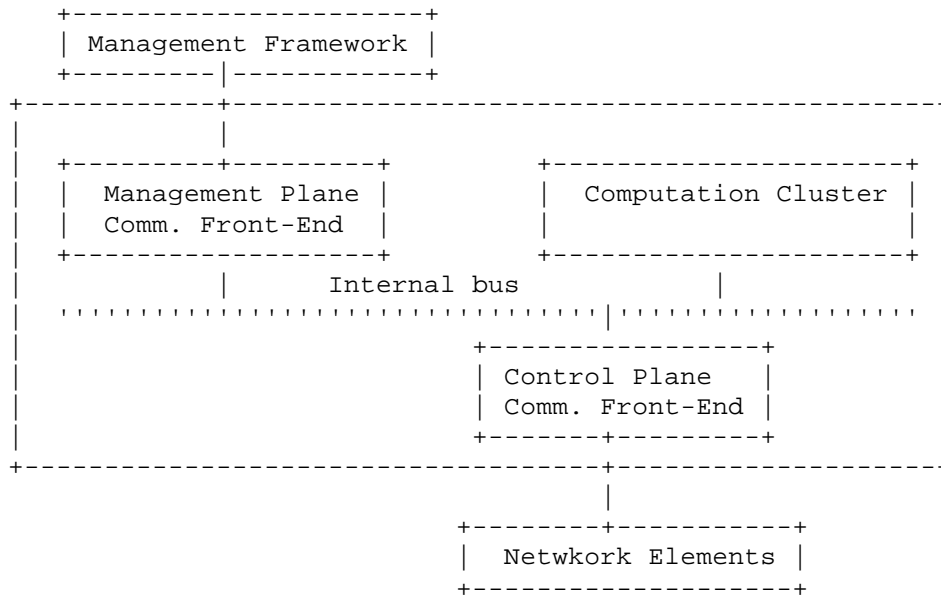


Figure 3 - Two Tier RMA Architecture

3.2 Traffic Engineering Database (TE-DB)

The construction of the TE-DB involves two asynchronous process:

- a) update of Topology Database (TDB) by the IGP
- b) policy and administrative information insertion from management applications.

The topology database (TDB) suffer intrinsic inaccuracies, due to the update mechanisms of the IGPs and the hierarchical routing approach with information summarization. Some form of inaccuracy reduction shall be implemented to reduced the gap between the gathered TDB and the actual network state. This, consequently, will reduce the blocking probability and the need for cranchback procedures in provisioning time.

Policy and administrative information is inserted in the TE-DB by management processes. Policy-Based Network Management enable network administrators to manage network behaviour using high-level definitions, or policies, which shall be expressed unambiguously. These policies, associated with an abstract model of the managed objects and accurate network status information permit to trigger or refrain certain actions on network elements, resulting in an enforcement of the high level policies. The policy language and information model is out of the scope of this proposal.

Typical administrative information comprises link resource class (colour), SRLGs, etc, while a simple routing policy is to deny the establishment of LSPs with bandwidth greater than a certain value, to tune the load sharing of traffic demands and minimize blocking probability.

3.3 The CBR process

There are a number of heuristics and a few exact algorithms to solve the CBR process in near real time. The implementation shall evaluate the applicable approaches to the RMA, taking into account the objective functions, the system of demands, network and administrative constraints that need to be satisfied.

4. Autodiscovery of RMAs

Autodiscovery requirements are defined in [PCE-DISC-REQ]. The RMA architecture could implement a very basic static strategy, relying on LSRs' local configuration; since the RMA architecture is redundant and fault-tolerant, this may be considered a minor drawback.

5. Security Considerations

A potential security issue is raised by the fact that the proposed architecture has connections to the network elements via the Control Plane interface, and to the management applications via the Management Plane interface. Specially crafted packets in the Control Plane could eventually be used to gain access to the RMA with potential incidence in network management applications. This is for further study.

6. Acknowledgments

The author would like to express his warmest thanks to Joan Serrat, Javier Baliosian, and the networking team at UdelaR for their support, review and valuable suggestions.

7. References

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

[PCE-ARCH] Farrel, A., Vasseur, JP, Ash, J., "Path Computation Element (PCE) Architecture", work in progress.

[PCE-COMM-REQ] Ash, J., Le Roux JL, "Path Communication Protocol Generic Requirements", work in progress.

[RFC3630] Katz D., Kompella K., Yeung D., "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC3784] Smit H., Li T., "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, June 2004.

[RFC3768] Hinden R. et al., "Virtual Router Redundancy Protocol (VRRP)", RFC 3768, April 2004.

[PCE-DISC-REQ] Le Roux, JL, et. al., "Requirements for Path Computation Element (PCE) Discovery," work in progress.

8. Author's Addresses

Eduardo Grampin
Universidad de la Republica - UdelaR
J.H. Reissig 565
11300 Montevideo - Uruguay
Email: grampin@fing.edu.uy

9. Full 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.

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, 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.