PCE Implementation and Testing by Using Virtual Machines

Byeongsik Kim and Bhum-Cheol Lee

Smart Node Platform Research Sector
Electronics and Telecommunications Research Institute (ETRI)
Daejeon, KOREA
bskim25@etri.re.kr, bclee@etri.re.kr

Abstract— **In this paper, we describe the way to implement and test a PCE on the virtualized environment. The PCE is a special computational entity that will corporate with similar entities to compute the best possible path through single or multiple domains. Generally, in real network, the testing and implementation of network-related works are very difficult and complex but those works are much easier on the virtualized environment using virtual machines and networks. In this paper, we introduce the result of PCE implementation and test environment by using virtual machines and networks of VMware hypervisor, VMware ESX Server.**

Keywords— **Virtualization, Hypervisor, Virtual Machine, Virtual Network, Path Computation Element**

I. INTRODUCTION

A Virtual Machine (VM) is a software implemented abstraction of the underlying hardware, which is presented to the application layer of the system. VMs may be based on specifications of a hypothetical computer or emulate the computer architecture and functions of a real world computer.

In computing, a hypervisor is a piece of computer software, firmware or hardware that creates and runs VMs. A computer on which a hypervisor is running one or more VMs is defined as a host machine. Each VM is called a guest machine. The hypervisor presents the guest operating systems with a virtual operating platform and manages the execution of the guest operating systems. Multiple instances of a variety of operating systems may share the virtualized hardware resources.

VMware ESX is an enterprise-level computer virtualization product offered by VMware, Inc. ESX is a component of VMware's larger offering, VMware Infrastructure, which adds management and reliability services to the core server product. VMware is replacing the original ESX with ESXi. VMware ESX and VMware ESXi are Type 1 hypervisors that are VMware's enterprise software hypervisors for guest virtual servers that run directly on host server hardware without requiring an additional underlying operating system.

A Path Computation Element (PCE) is an entity that is capable of computing a network path or route based on a network graph, and of applying computational constraints during the computation. The PCE entity is an application that can be located within a network node or component, on an out-of-network server, etc. For example, a PCE would be able to compute the path of a Traffic Engineering (TE) Label Switched Path (LSP) by operating on the Traffic Engineering Data Base (TED) and considering bandwidth and other constraints applicable to the TE LSP service request [1].

In this paper, we describe the way to implement and test a PCE on the virtualized environment. The PCE is a special computational entity that will corporate with similar entities to compute the best possible path through single or multiple domains. Generally, in real network, the testing and implementation of network-related works are very difficult and complex but those works are much easier on the virtualized environment using virtual machines and networks. In this paper, we introduce the result of PCE implementation and test environment by using virtual machines and networks of VMware hypervisor, VMware ESX Server

The remainder of this paper is organized as follows. We describe the details of the hypervisor (VMware ESX Server) and PCE. We present implementation results of the PCE, and discuss the test environment based on the VMs.

II. RELATED WORKS

A. *VMware ESX Server*

VMware ESX Server [2], an enterprise-level product, can deliver greater performance than the freeware VMware Server, due to lower system overhead. VMware ESX is a "bare-metal" product, running directly on the server hardware, allowing virtual servers to also use hardware more or less directly as shown in Figure 1.

VMware ESX is VMware's enterprise server virtualization platform. The platform is available in two versions -- ESX Server and ESXi Server. VMware ESX and ESXi can be deployed as part of the VMware infrastructure (vSphere or VMware View) to enable centralized management for data center applications and enterprise desktops. One main feature of the enterprise server virtualization platform is that it allows users to contain server sprawl by running software applications in virtual machines on fewer physical servers.

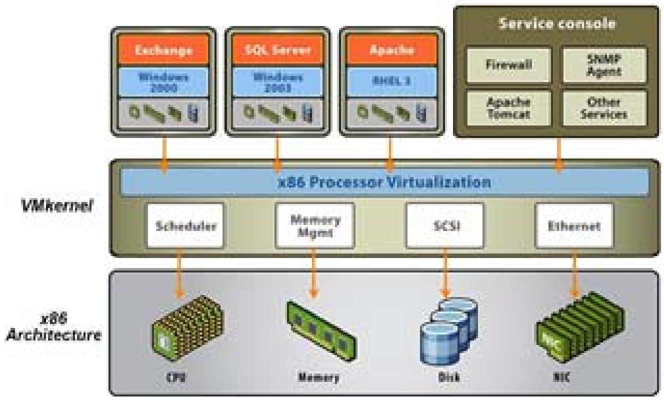

Figure 1. VMware ESX Server Architecture

B. *Path Computation Element (PCE)*

The PCE architecture is shown in Figure 2. The PCE is a server that can perform constrained path computation using topology and traffic engineering information stored in TED. The use of a TED allows the PCE to compute optimal constrained paths, which is especially useful in provisioning services in transport networks, such as WDM, which have strict QoS requirements.

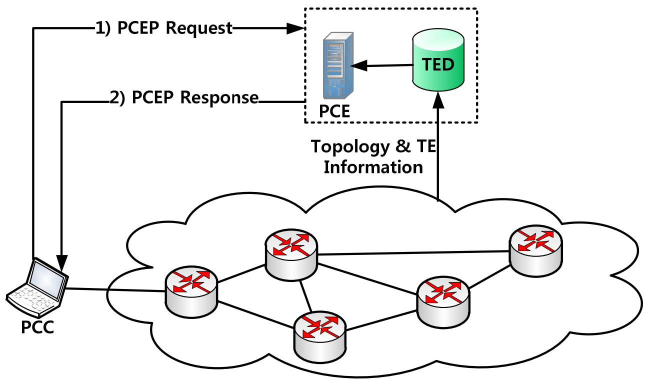

Figure 2. PCE Architecture

In this architecture, a Path Computation Client (PCC) can communicate with the PCE using PCE communication Protocol (PCEP) [3]. Using the PCEP, the PCC requests the PCE to compute a path from the source node to the destination node under necessary constraints.

III. PCE IMPLEMENTATION

We implemented the PCE according to the IETF standard. Figure 3 shows the architecture of the implementation. It consists of five major blocks, (a) the management and processing of PCEP messages from PCC and PCE peers, (b) the management and update processing of the TED by Routing Listener, (c) the path computation itself, (d) the management and update processing of the PCE Database (PCED) by PCE Discovery Manager, and (e) the management of the inter-domain path computation by the Backward Recursive PCE-based Computation (BRPC) Manager [4].

The PCE architecture requires that a PCC be aware of the location of one or more PCEs in its domain, and, potentially, of PCEs in other domains, e.g., in the case of inter-domain TE LSP computation.

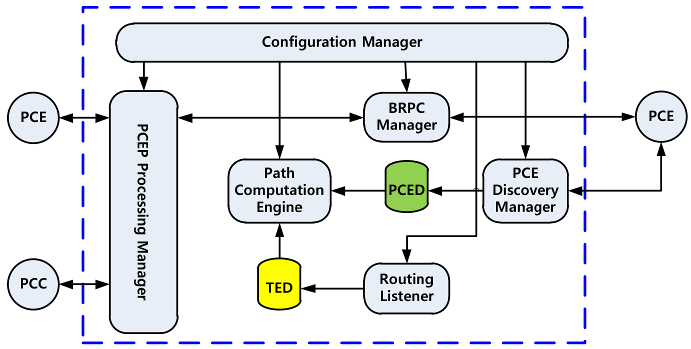

Figure 3. Functional Architecture of Our PCE

A network may contain a large number of PCEs, each with potentially distinct capabilities. In such a context, it is highly desirable to have a mechanism for automatic and dynamic PCE discovery that allows PCCs to automatically discover a set of PCEs, along with additional information about each PCE that may be used by a PCC to perform PCE selection. Additionally, it is valuable for a PCC to dynamically detect new PCEs, failed PCEs, or any modification to the PCE information. Detailed requirements for such a PCE discovery mechanism are provided in [5][6]. When PCCs are LSRs participating in the IGP (OSPF or IS-IS), and PCEs are either LSRs or servers also participating in the IGP, an effective mechanism for PCE discovery within an IGP routing domain consists of utilizing IGP advertisements.

A routing domain may, in practice, contain multiple PCEs:

- The path computation load may be balanced among a set of PCEs to improve scalability.
- For the purpose of redundancy, primary and backup PCEs may be used.
- PCEs may have distinct path computation capabilities (multi-constrained path computation, backup path computation, etc.).
- In an inter-domain context, there can be several PCEs with distinct inter-domain functions (inter-area, inter-AS, inter-layer), each PCE being responsible for path computation in one or more domains.

In order to allow for effective PCE selection by PCCs, that is, to select the appropriate PCE based on its capabilities and perform efficient load balancing of requests, a PCC needs to know the location of PCEs in its domain, along with some information relevant to PCE selection, and also potentially needs to know the location of some PCEs in other domains, for inter-domain path computation purpose (Figure 4).

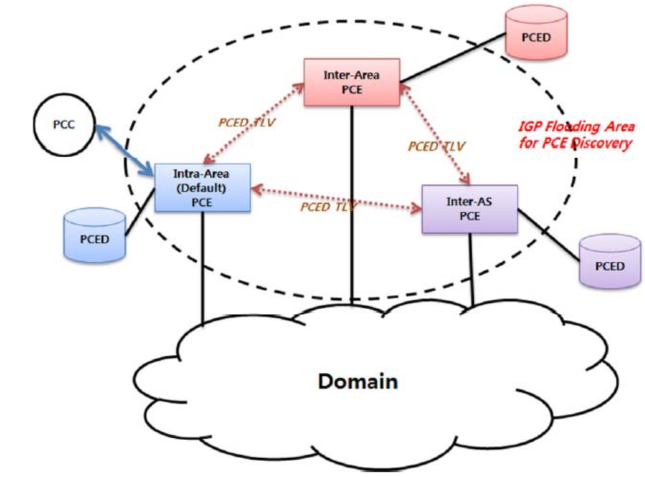

Figure 4. PCE Discovery

Figure 5 shows three connected domains with one inter-domain PCE per domain (area). In domain 1, PCC sends an inter-domain path computation request to default PCE, which forwards it requests to the inter-domain PCE 1. Using BRPC, PCE 1 sends a request to PCE 0, which forward it to PCE 2. The PCE 0 replies with the distance from its boundary nodes with the Domain 2. The PCE 0 carries out the same operation sending a tree with the possible combinations. In our implementation, BRPC manager in Figure 3 performs the function to compute the inter-domain paths.

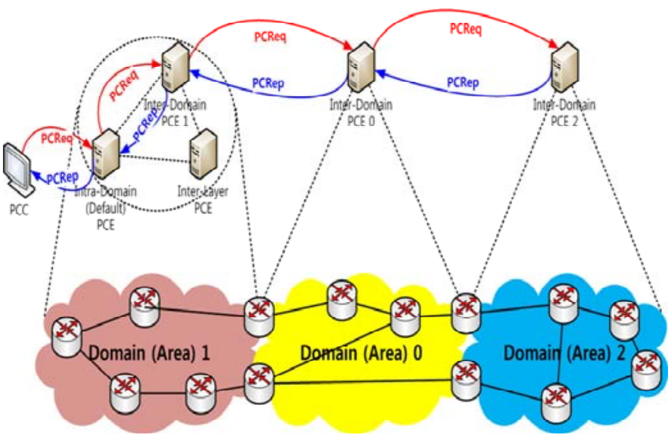

Figure 5. Implementation of BRPC

IV. Test Environment and Results

In order to evaluate our implementations, we constructed a test bed consisting of three domains that correspond to OSPF-TE routing areas as shown in Figure 3. OSPF-TE routers in the test bed are based on Virtual Label Switching Router (VLSR) software package that is one of key functional elements in the DRAGON control plane architecture [5]. Basically, we implemented all PCEs and routers on the virtual machines and constructed the virtual networks using the routers. To construct the network in Figure 5, we created the each network using the Virtual Switch in the VMware ESX as shown in Figure 6. This configuration is for the Area 1 network in Figure 5. Using this method, we created all test network including all PCEs and all routers in Area 0 and Area 2.

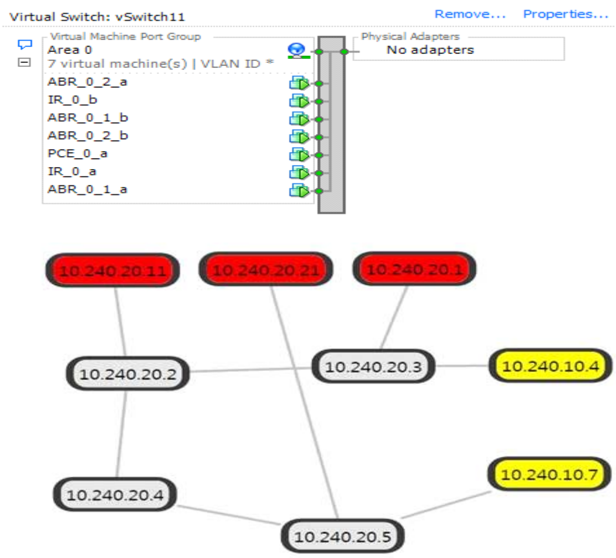

Figure 6. Area1 Network Creation using Virtual Machines

Figure 7 shows the results that compute the path satisfied the constraints (bandwidth) from the source router in Area 1 to the destination router in Area 2.

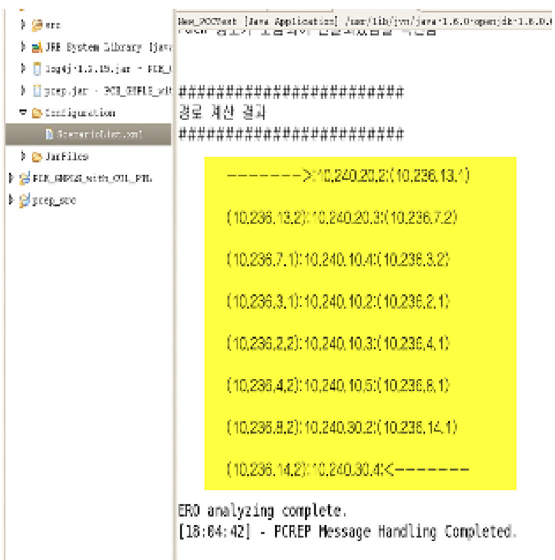

Figure 7. Path Computation Results

V. Conclusions

This paper describes the results of PCE implementation using virtual machines and also construction of test networks using the virtual networks based on the VMware ESX's virtual switches. Virtual machines and virtual network can be flexible and scalable tools for implementations and tests of the communication protocols and networks.

Acknowledgment

This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013.

References

[1] A. Farrel, J.P. Vasseur, and J. Ash, "A Path Computation Element (PCE)-Based Architecture", *RFC 4655*, Aug. 2006

[2] A. Muller, S. Wilson, D. Happe, and G.J. Humphrey, "Virtualization with ESX Server", O'Reilly Media, 2005

[3] JP. Vasseur and J.L. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, Mar. 2009

[4] J. P. Vasseur, et. al. "A Backward Recursive PCE-based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-domain Traffic Engineering Label Switched Paths", RFC5441, Apr. 2009.

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

[6] J.L. Le Roux, JP. Vasseur, Y. Ikejiri, and R. Zhang, "OSPF Protocol Extensions for Path Path Computation Element (PCE) Discovery", RFC5088, Jan. 2008.

[7] DRAGON-Project, http://dragon.east.isi.edu/twiki/bin/view/DRAGON/WebHome

Byeongsik Kim is a principle member of engineering staff in Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. He received his Ph.D., M.S., and B.S. degrees in computer engineering from Chungnam National University in 2001, 1997, and 1995, respectively. He was a visiting researcher at National Institute of Standards and Technology (NIST), USA in 1997. His recent research areas include network resource management, network QoS, cloud operating systems, and network virtualization.

Bhum-Cheol Lee received the B.S. degree from Kyung-Hee University, Seoul, Korea, in 1982, and the M.S. and Ph.D. degrees from Yonsei University, Seoul, in 1983 and 1997, respectively. From 1983 to 1995, he was with the Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea, as an Interface, Switching, Link, and Network Synchronization Engineer. Currently he is the Project Leader of the Smart Node Platform Research Section, ETRI. His research interests are synchronization, line coding, analog and digital circuit design, and network virtualization.