

A Design of Virtual Machine aware Flow Switch in the Cloud based Network Function Virtualization System

Kang Il Choi, Bhum Cheol Lee

Network Research Division, ETRI, Daejeon, Korea

forerunner@etri.re.kr, bclee@etri.re.kr

Abstract— In this paper, we present a design of Virtual Machine aware flow switch in the Cloud based Network Function Virtualization System, which capable of providing a service at a stable speed even though the number of virtual machines connected to a hypervisor is increased by extracting flow information about each of a plurality of packets and providing network virtualization. In this paper, we also present an architecture of the VM aware flow switch, detailed operation method of it, and the POC of the VM aware flow switch.

Keywords— Flow Switch, Cloud, Network Function Virtualization

I. INTRODUCTION

In the server virtualization environment, one or more different virtual machines (several, several tens, or several hundreds of virtual machines) may exist in one computer server. There is a hypervisor (or a virtual machine monitor (VMM)) sharing and simultaneously executing hardware (a CPU, a memory, a storage, a network interface, and the like) resources of a computer server in which several (several tens or several hundreds of) different virtual machines are virtualized. The hypervisor performs functions, such as a generation of a virtual machine within a server, a removal of a virtual machine, and a management of resources of virtual machines. The hypervisor enables the virtual machines to share a network and a storage. In a case of the storage, when the hypervisor is set so that each virtual machine uses a storage of a logically or physically divided region, the virtual machines may share the entire storage without interfering each other.

However, in a case of the network, several virtual machines (several tens or several hundreds of virtual machines) installed in one computer server generally share one or more network devices. When one or more network devices are shared by one or more virtual machines, one or more virtual machines need to be provided so that the virtual machines share a network without interfering each other in a network level for each virtual machine. Researches on a network virtualization technology for solving the aforementioned problems have been recently conducted [1-7].

II. ARCHITECTURE OF THE VM AWARE FLOW SWITCH

A. System Diagram of the VM aware Flow Switch

Fig. 1 is a system diagram schematically illustrating a network system including a VM aware flow switch, several

servers and VM Flow Manager. Each of the servers may include several virtual machines and hypervisors[1].

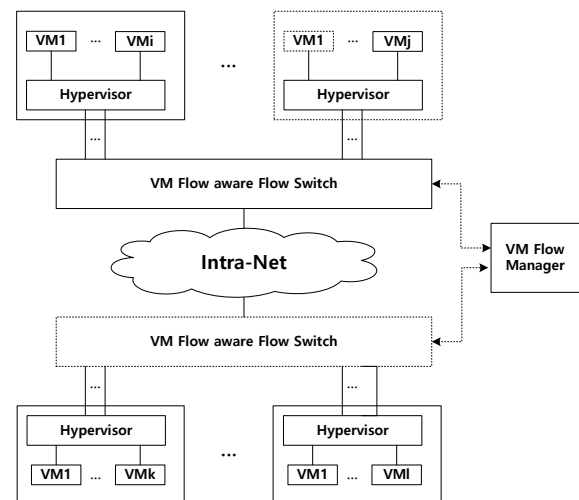


Figure 1 System Diagram of the VM aware Flow Switch

In Fig.1, the virtual machines may be operating systems (OS), operated in logical hardware provided by the hypervisor. The virtual machines generate packets according to a demanded service, for example, a web server, a mail server, a file server, a video server, and a cloud server, business finance, finance, and stocks, and packets include flow information representing the demands for different qualities of service (QoS) according to the demanded services.

In Fig.1, the VM aware flow switches may communicate with each other through the Internet or Intranet. The first VM aware flow switch receives packets transmitted through the network interface connected with servers. In this case, the first VM aware flow switch receives packets transmitted from the hypervisor, extracts the flow information from the first packet of the first flow, and extracts the second flow information from the first packet of the second flows. The first VM aware flow switch checks whether first and second previous flow information corresponding to flow information, respectively, is stored in a flow information table stored based on flow information. Then, when previous flow information corresponding to flow information, respectively, exists, the first VM aware flow switch may process packets according to first and second previous QoS policy information corresponding to previous flow information, respectively.

When previous flow information exists, the first VM aware flow switch may check whether previous flow information is included in QoS limitation information set for packets before previous QoS policy information is applied, and determine whether to apply previous QoS policy information. This will be described in detail with reference to Fig.2 below. When previous flow information corresponding to at least one of flow information does not exist in the flow information table, the first VM aware flow switch generates new QoS policy information according to a QoS generation policy set for at least one of flow information, and processes at least one of flow information according to the QoS policy information.

B. Control Block Diagram of the VM aware Flow Switch

Fig. 2 is a control block diagram illustrating a control configuration of the VM aware flow switch illustrated in Fig.1. Referring to Fig. 2, the VM aware flow switch includes a virtual flow recognition switch, a flow determination unit, and a packet processing control unit. The virtual flow recognition switch receives packets transmitted from the hypervisor included in the first server connected through the network interface, and transfers the received first and second packets to another virtual flow recognition switch through the Internet or the Intranet according to a control of the packet processing control unit to provide or receive a service. The virtual flow recognition switch may be a logically virtual switch or a physical switch. The flow determination unit extracts the first flow information from the first packet received from the virtual flow recognition switch, extracts the second flow information from the first packet of second flow, and checks whether previous flow information corresponding to flow information is stored in a stored previous flow information list.

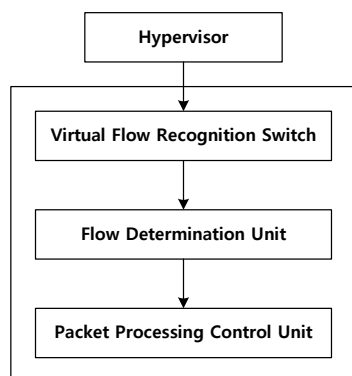


Figure 2 A Control Block Diagram of the VM aware Flow Switch

Here, each of flow information may include at least one of an IP source address, an IP destination address, a protocol number, source transport information, and destination transport information included in packets. The packet processing control unit controls the virtual flow recognition switch so that packets are processed according to a result of a determination of the flow determination unit. That is, when there is no first and second previous flow information matched with flow information, respectively, as a result of the determination of the flow determination unit, the packet processing control unit generates QoS policy information according to the QoS generation policy set based on each of flow information, transmits QoS policy information to the virtual flow recognition switch, and controls

the virtual flow recognition switch so that packets are processed. In this case, the QoS generation policy may include at least one of the type of services, the amount of bandwidth used, and predetermined user demands based on flow information.

When there is the first previous flow information matched with the first flow information, and there is no second previous flow information matched with the second flow information as the result of the determination of the flow determination unit, the packet processing controller generates the second QoS policy information based on the second flow information and transmits the generated second QoS policy information to the virtual flow recognition switch as described above. In this case, in a case where there is the first previous flow information matched with the first flow information, the packet processing control unit checks whether the first flow information is included in the predetermined QoS limitation information in order to apply first previous QoS policy information corresponding to the first previous flow information. That is, in a case where there is the first previous flow information, the packet processing control unit calculates the total amount of bandwidths used for packets, and the amount of first individual bandwidth used for the first packet. Then, the packet processing control unit compares the total amount of bandwidths used and a port bandwidth of the network interface included in the QoS limitation information, and when the total amount of bandwidths used is equal to or smaller than the port bandwidth, the packet processing control unit controls the virtual flow recognition switch unit so that the first packet is processed without applying the first previous QoS information.

When the total amount of bandwidths used is larger than the port bandwidth, the packet processing control unit re-checks whether the amount of total bandwidths used is included in the amount of first individual bandwidth used and an individual bandwidth included in the QoS limitation information, and when the amount of first individual bandwidth used is equal to or smaller than the individual bandwidth, the packet processing control unit controls the virtual flow recognition switch so that the first packet is processed without applying the first previous QoS information. When the amount of first individual bandwidth used is larger than the individual bandwidth, the packet processing control unit controls the virtual flow recognition switch so that the first packet having the amount of first individual bandwidth used is processed by applying the first previous QoS policy information.

In the diagram, flow information of packets contains information about virtual machines and generating packets, and the flow information table may include at least one of a virtual machine information list and a flow information list. Accordingly, in a case where the packet processing control unit stores flow information in the flow information table, the packet processing control unit may make a control so that flow information is divided and stored in at least one of the virtual machine information list and the flow information list.

C. Operation Flowchart of the VM aware flow switch

Fig. 3 is a flowchart illustrating an operating method of the VM aware flow switch. Referring to Fig. 3, when first and second packets and are transmitted from the hypervisors through the network interface, the VM aware flow switch extracts first

and second flow information corresponding to packets, respectively(S110).

That is, the virtual flow recognition switch receives packets transmitted from the hypervisor included in the first server connected through the network interface, and transfers the received first and second packets to another virtual flow recognition switch through the Internet or the Intranet according to a control of the packet processing control unit to provide or receive a service. The flow determination unit extracts the first flow information from the first packet received by the virtual flow recognition switch and the second flow information from the first packet of second flow.

After step S110, the VM aware flow switch determines whether first and second previous flow information matched with flow information, respectively (S120), and when there is not first and second previous flow information or the first previous flow information, the VM aware flow switch generates each of first and second QoS policy information according to QoS generation policy corresponding to each of flow information, or generates first QoS policy information according to a QoS generation policy corresponding to the first flow information (S130), and processes packets according to QoS policy information or processes the first packet according to the first QoS policy information (S140).

That is, when there is no stored first and second previous flow information matched with flow information, respectively, or there is no first previous flow information matched with the first flow information among flow information as a result of the determination of the flow determination unit, the packet processing control unit generates QoS policy information according to a QoS generation policy set based on each of flow information, or generates the first QoS policy information according to a QoS generation policy set based on the first flow information. Then, the packet processing control unit transmits QoS policy information or the first QoS policy information to the virtual flow recognition switch, and controls the virtual flow recognition switch so that packets or the first packet is processed.

When there is previous flow information or there is only the second previous flow information as the result of the determination of step S120, the VM aware flow switch calculates the total amount of bandwidths used for packets and the amounts of first and second individual bandwidths used for the respective first and second packets (S150).

That is, when there is the stored first and second previous flow information matched with flow information, respectively or there is only the second previous flow information matched with the second flow information among flow information as a result of the determination of the flow determination unit, the packet processing control unit calculates the total amount of bandwidths used for packets, and the amounts of first and second individual bandwidths used for the respective first and second packets. The VM aware flow switch compares whether the total amount of bandwidths used calculated in step S150 is larger than a port bandwidth included in predetermined QoS limitation information (S160), when the total amount of bandwidths used is smaller than the port bandwidth, the VM aware flow switch controls so as to process packets or the second packet without applying previous QoS policy information or the second

previous QoS policy information (S170). That is, when the total amount of bandwidths used is smaller than the port bandwidth included in the predetermined QoS limitation information, for example, a bandwidth of the network interface, the packet processing control unit controls so as to process packets or the second packet without applying previous QoS policy information or the second previous QoS policy information.

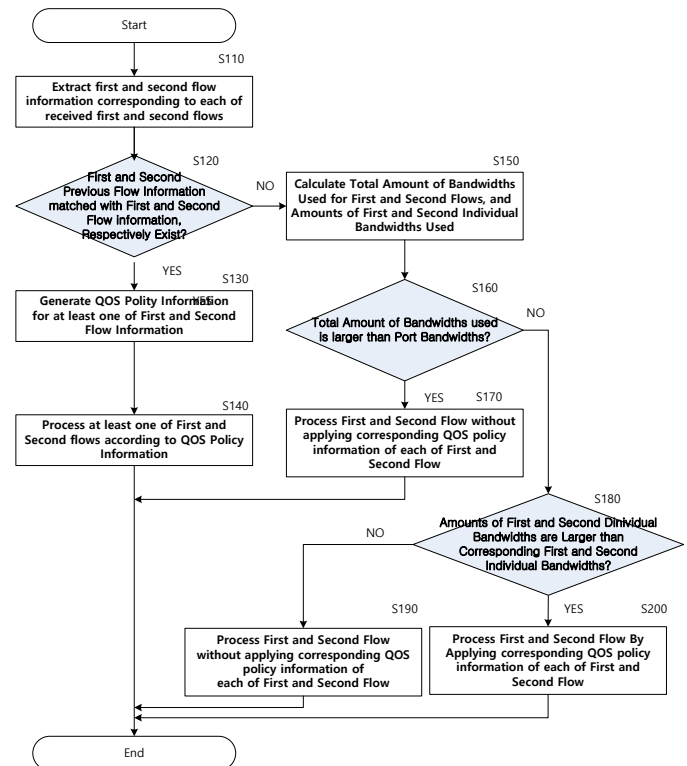


Figure 3 Operation of the VM aware Flow Switch

When the total amount of bandwidths used is larger than the port bandwidth in step S160, the VM aware flow switch compares the amounts of first and second individual bandwidths used and first and second individual bandwidths included in the QoS limitation information and corresponding to the amounts of respective first and second individual bandwidths used, and compares whether the amounts of first and second individual bandwidths used are larger than individual bandwidths, respectively, or whether the amount of second individual bandwidth used is larger than the second individual bandwidth (S180). When the amounts of first and second individual bandwidths used are smaller than individual bandwidths, respectively, or when the amount of second individual bandwidth used is smaller than the second individual bandwidth, the VM aware flow switch controls the virtual flow recognition switch so that packets are processed without applying previous QoS policy information (S190). That is, the packet processing control unit compares individual bandwidths matched with the amounts of first and second individual bandwidths used, respectively, and compares whether the amount of first individual bandwidth used is larger than the first individual bandwidth, or whether the amount of second individual bandwidth used is larger than the second individual bandwidth.

Here, when the amount of first individual bandwidth used is smaller than the first individual bandwidth, and the amount of second individual bandwidth used is smaller than the second individual bandwidth used, the packet processing control unit controls the virtual flow recognition switch so that packets are processed without applying previous QoS policy information.

When at least one of the amounts of first and second individual bandwidths used is larger than individual bandwidths in step S180, the VM aware flow switch controls the virtual flow recognition switch so that packets are processed by applying previous QoS policy information corresponding to at least one of the corresponding first and second packets (S200). That is, when both the amounts of first and second individual bandwidths used are larger than individual bandwidths, the packet processing control unit controls the virtual flow recognition switch so that packets are processed by applying previous QoS policy information, and when the amount of first individual bandwidth used is larger than the first individual bandwidth or the amount of second individual bandwidth used is larger than the second individual bandwidth, the packet processing control unit controls the virtual flow recognition switch so that packets are processed by applying the corresponding first and second previous QoS policy information.

D. Proof of Concept(POC) of the VM aware flow switch

Fig. 4 is a Cloud based Network Function Virtualization System, a POC platform of the VM aware flow switch, which integrates computing, storage and networking system.

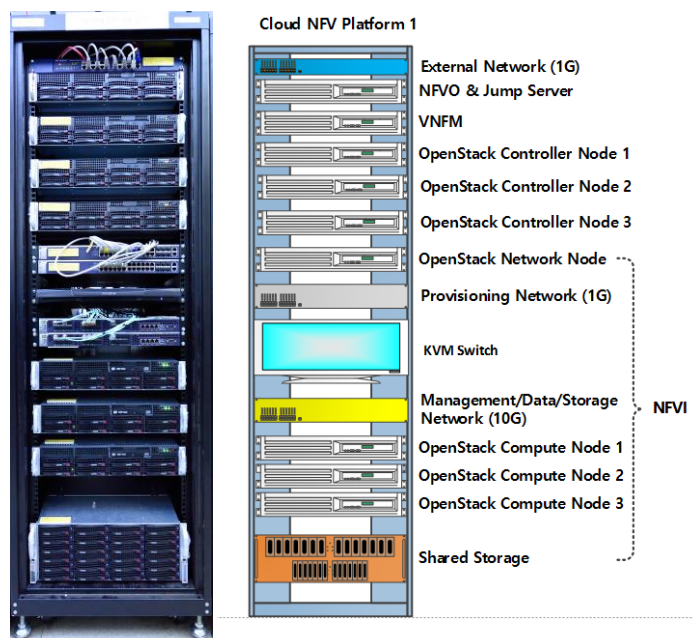


Figure 4 a POC platform of the VM aware Flow Switch

III. CONCLUSIONS

In this paper, we provide the design of VM aware flow switch in the Cloud based Network Function Virtualization System, which capable of providing network virtualization. In this paper, we also provide an architecture of the VM aware flow

switch and present an operation method of it. We are currently testing the scheme in the Cloud based NFV platform.

ACKNOWLEDGMENT

This work was supported by ICT R&D program of MSIP/IITP. [B0101-16-0233, Smart Networking Core Technology Development]

REFERENCES

- [1] Kang Il Choi, Bhum Cheol Lee, Seung Woo Lee, Young Ho Park, Jung Hee Lee and Sang Min Lee, "Virtual Machine Flow Aware Flow Switch and Method for Operation Thereof", US2014061862 A1, May 8, 2014
- [2] Kang Il Choi, Bhum Chul Lee, "Applying Data Plane Acceleration Technology into a Cloud Network Function Virtualization Platform," ITC-CSCC 2015, pp. 381-384, July 2015.
- [3] Kyeong-Hwan Doo, Bin-Yeong Yoon, Bhum-Cheol Lee, Soon-Seok Lee, Man-Soo Han, and Whan-Woo Kim, "Multicore Flow Processor with Wire-Speed Flow Admission Control", ETRI Journal, vol. 34, no. 6, Dec. 2012, pp. 827-837.
- [4] B.Y. Yoon, B.C. Lee, and S.S. Lee, "Scalable Flow-Based Network Processor for Premium Network Services", ICTC2011, Nov. 2011, pp. 436-440.
- [5] Jae Won Park, Won Young Chung, Hyun Pil Kim, Jung Hee Lee, Yong Surk Lee, "Parallel Processing Architecture for Network L2~4 and L7", ITC-CSCC 2010, pp.147-150.
- [6] Jong Geun Park, "Inter Cloud Networking between Distributed Micro Data Centers", ICGHIT2015, 2015.02.
- [7] HongSik Jeon, "CDN Service Scenarios in Distributed NFV Infrastructures using OpenStack Heat", ICGHIT2015, 2015.02.



Kang Il Choi received B.S. degree in Computer Science from KAIST, Korea and M.S. degree in Computer Science from Sogang University in 1992 and 1994, respectively. He is currently principal researcher of Electronics and Telecommunications Research Institute (ETRI), Korea. His research interests are Multicore Parallel Processing, Distributed Cloud Data Center, Data Plane Acceleration Technology (Intel DPDK, ODP etc) and Network Virtualization.



Dr. Bhum Cheol Lee received M.S. and Ph.D. degree in Electric Engineering from Yonsei University, Korea in 1983 and 1997, respectively. He is currently Technical Lead of Networking Research division in Electronics and Telecommunications Research Institute (ETRI), Korea. His research interests are Smart Network, Parallel Flow Processing and Network Virtualization.