Network Virtualization Technology for Cloud Services

Hideo Kitazume[†]*, Takaaki Koyama, Toshiharu Kishi, and Tomoko Inoue*

Abstract

The network virtualization technology needed to effectively and efficiently construct and operate a cloud is already in place. In this article, we introduce the trends in the latest network virtualization technology for the cloud environment and its fields of application.

1. Introduction

In recent years, the development of server virtualization technology has led to changes in the cloud services environment, such as more efficient usage of physical servers (high aggregation), sharing of hardware resources such as server and network (multitenacy), and the need for practical migration. For networks within datacenters, attention has been drawn to problems such as the explosive increase in the number of medial access control (MAC) addresses and virtual local area networks (VLANs), the construction of layer 2 (L2) networks across offices, and network migration.

Technology for solving these problems has taken two major directions: architectures that extend existing technology and network equipment commoditization through virtualization. The former uses fast, high-capacity L2 switches ranging from 40GbE (40-Gbit/s Ethernet) ones to 100GbE (100-Gbit/s Ethernet) ones to make flat connections over multiple switches in one hop; these switches are operated and managed as one very large logical L2 switch. Although this approach is expected to be widely used in large, next-generation datacenter networks, there are problems such as strong dependence on the switch vendor, insufficient interworking with network equipment other than switches (e.g., firewalls and load balancers), and the same high construction cost as in the past. The latter approach, on the other hand, establishes a logical network independently of the physical network by logically integrating functions such as those of network equipment other than switches into a standard switch called an OpenFlow switch. As a result, a carrier can expect lower construction costs for datacenter networks using commodity network equipment. OpenFlow was initiated by the Open Network Foundation (ONF), a promotional organization that has been active in moving the technology toward a practical stage. ONF has focused on the flexibility of network programmability (reconfiguration of the network in connection with applications), which is considered to be a powerful network virtualization technology for the cloud environment.

In this article, we explain the need for network virtualization as well as the technology itself and its application areas.

2. Need for network virtualization

As described above, multitenancy requires the ability to provide an isolated network for each cloud services user, together with network flexibility and agile construction and configuration changes for high aggregation and migration needs. There are, however, difficult problems in satisfying those requirements with existing VLANs and products.

(1) Difficulty of network design and management in datacenters

In datacenters, the tagged VLAN is generally used to isolate each user's network, but there are two

[†] NTT Information Sharing Platform Laboratories
Musashino-shi, 180-8585 Japan

problems with this approach. One is the limited capacity of the tagged VLAN. This technique attaches one of 4094 VLAN identifiers (VLAN-IDs) to packets, so the maximum number of isolatable networks is 4094. A VLAN-ID must be unique over the entire datacenter network, so it is impossible to accommodate more than 4094 users at the same time. The other problem is the use of proprietary specifications by the vendors of the products used in datacenters. For the products currently used in datacenters, most vendors use proprietary specifications in both network design and setup. Network design, for example, may be premised on a virtual chassis* being used for switch products, in which multiple switches are seen as a single switch, and on the tagged VLAN being used for communication within server products. For setup methods, individual vendors have proprietary control protocol specifications etc. It is thus necessary to assign the VLAN-ID to match the vendor specifications, and the control protocol used must also match the vendor specifications. Furthermore, as the numbers of servers and switches in datacenters increase to more than hundreds of units, the use of products from multiple vendors to avoid vendor lock-in and reduce costs increases the difficulty of datacenter network design and management.

(2) Agile automatic changes in network configuration in cooperation with migration

For cloud services, virtual machine migration among multiple datacenters in times measured in hours is necessary in order to take advantage of nighttime electricity rates and to respond to disasters. Specifically, services can be migrated without interruption on a virtual machine that uses TCP (transmission control protocol), UDP (user datagram protocol), or other protocols, and the network configuration can be changed in cooperation with the migration within a very short time.

As a step toward solving the above two problems, a new standard network virtualization technology that does not use VLANs and that eliminates vendor dependence is needed for cloud services.

3. OpenFlow and ONF

OpenFlow originated as technology for an academic network at Stanford University in 2008, but it is now being studied by the OpenFlow Switch Consortium as a network control technology.

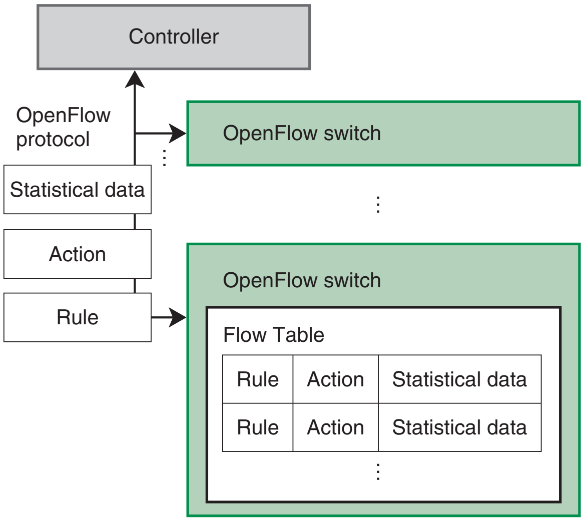


Fig. 1. Switch control in OpenFlow.

3.1 Controller-data separation

The basic concept of OpenFlow is that a controller performs central control by distributing programs to switches that conform to the OpenFlow specifications. Each switch then operates according to the program (**Fig. 1**). Unlike the conventional Internet, in which the various types of network equipment exchange path information and select paths autonomously, the controller performs all control centrally, and each switch operates according to the instructions given to it. This scheme is referred to as controller-data separation.

3.2 Control mechanism

OpenFlow control is specified as combinations of rules and actions. Rules identify the packets to be processed. It is thus possible to specify the evaluation conditions for L1–L4 header contents for packets whose TCP port number is 80, for example. Actions specify operations to be performed on the packets that match the rules. Specifically, it is possible to rewrite the header so that the packets are transferred to different ports or to specify that they are to be discarded, etc. For example, an action can specify that packets that arrive at a particular port number are to be discarded. Another way to regard this is that the controller can change a switch into a router, firewall, or load balancer as needed by sending a simple program to it. The flexibility of being able to control anything by programs is one reason that OpenFlow has been attracting attention.

* Chassis: In this context, chassis covers case, cabinet, and chassis.

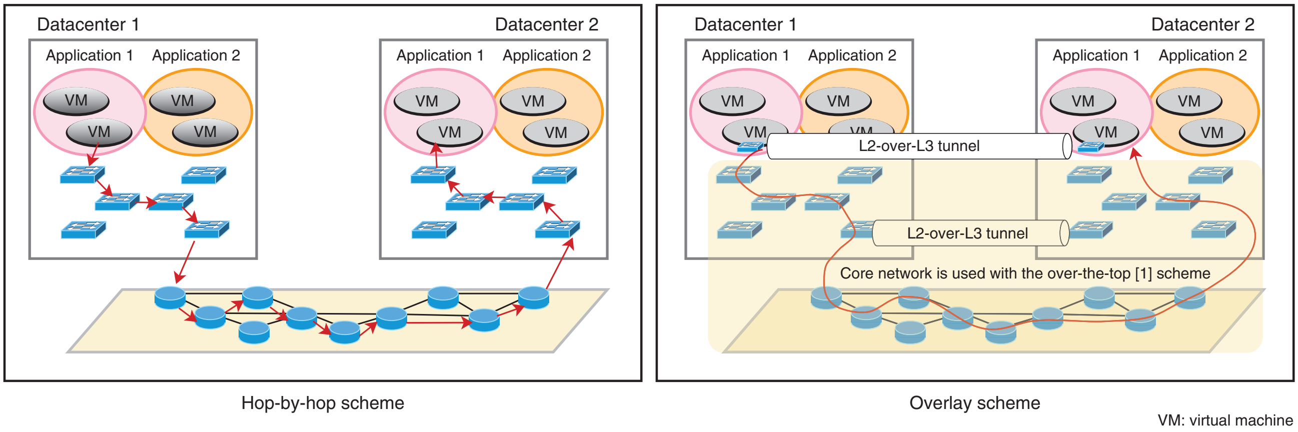


Fig. 2. Hop-by-hop and overlay schemes.

3.3 ONF

ONF is an organization of member companies such as Google, Facebook, and Yahoo that began promoting OpenFlow in March, 2011. NTT is also a participant. An interesting feature of ONF is that the board members are representatives of companies that operate large-scale datacenters, which is to say they are network user companies rather than network vendors. With the appearance of ONF, OpenFlow has taken on a commercial aspect in addition to its previous image of an academic system and it is now attracting much attention.

4. Use of OpenFlow

Considering routing, there are two schemes for using OpenFlow: hop-by-hop and overlay (**Fig. 2**).

(1) Hop-by-hop

In the hop-by-hop routing scheme, the controller knows all of the switches and designs paths service-by-service. Each switch operates according to instructions so as to repeatedly forward packets in a relay scheme that delivers the packets to their final destination. Although this scheme makes free use of the advantages of OpenFlow, each switch must hold all of the path information, so scalability may be a problem. It is suitable for the construction of small-scale networks, but application to large-scale networks requires measures against path congestion etc.

(2) Overlay scheme

In the overlay scheme, the controller does not control all of the paths, but uses the tunneling technique described later to control the communicating end points, a practice that is referred to as edge networking. With this scheme, the controller and the various switches need to know only the source and destination of the communication; the path is handled by the conventional routing mechanism.

This approach enables the amount of routing data to be managed to be kept down to a realistic level, even for a large-scale network. Early introduction to actual services is expected to be more feasible for the overlay scheme than for the hop-by-hop scheme. Next, we explain the implementation of an overlay virtual network with L2-over-L3 tunneling, which is one kind of overlay scheme.

5. Overlay virtual network

The overlay virtual network is implemented by encapsulating users' L2 frames inside L3 packets to achieve L2-over-L3 tunneling (**Fig. 3**). The three main points are explained below.

(1) L2-over-L3 tunnel

In this technique, an OpenFlow switch within a hypervisor is equipped with an L2-over-L3 tunnel endpoint function and a tunnel is established between two hypervisors. The connection between the on-premises environment (locally operated) and a hypervisor is also established by setting up an OpenFlow switch-based virtual gateway that has a tunnel endpoint function. The use of VLANs is abandoned to avoid VLAN-ID exhaustion and to escape from vendor dependence in VLAN-ID design.

(2) User isolation

User isolation is implemented with a function that assigns a user ID to each user and encapsulates the

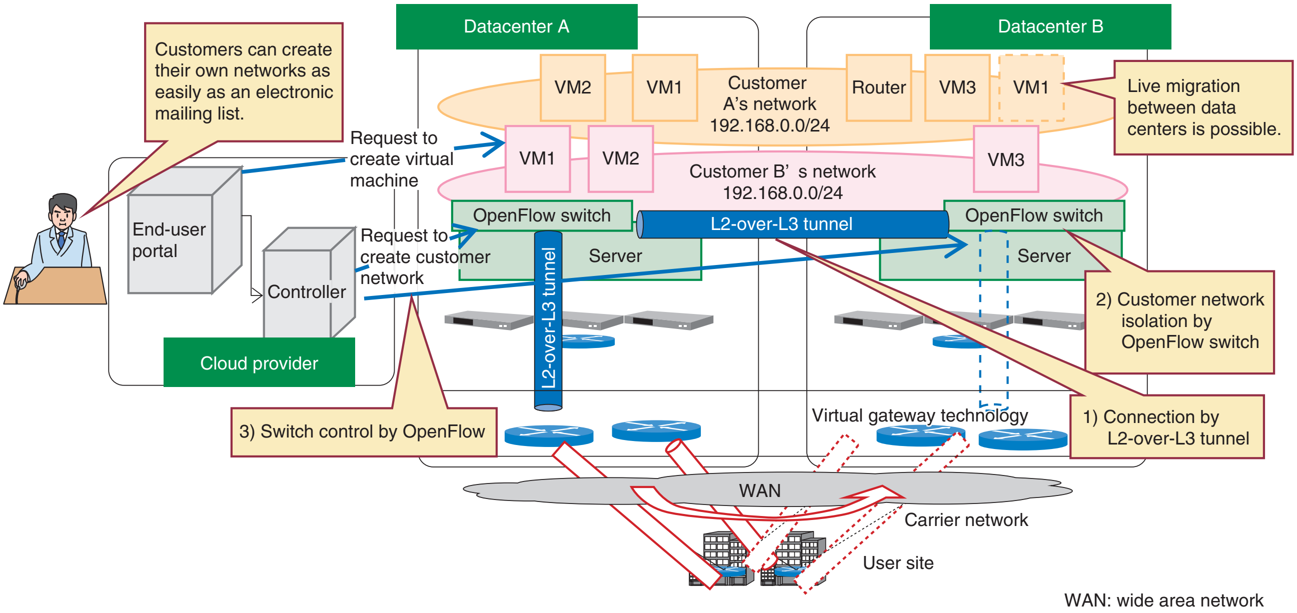

Customers can create their own networks as easily as an electronic mailing list.
Datacenter A
Datacenter B
VM2
VM1
Customer A's network 192.168.0.0/24
Router
VM3
VM1
Live migration between data centers is possible.
Request to create virtual machine
VM1
VM2
Customer B' s network 192.168.0.0/24
VM3
End-user portal
OpenFlow switch
L2-over-L3 tunnel
OpenFlow switch
Request to create customer network
Server
Server
Controller
L2-over-L3 tunnel
2) Customer network isolation by OpenFlow switch
Cloud provider
Virtual gateway technology
3) Switch control by OpenFlow
1) Connection by L2-over-L3 tunnel
WAN
Carrier network
User site
WAN: wide area network

Fig. 3. Overlay virtual network.

data within the tunnel endpoint function of the OpenFlow switch. The switching is then performed by the OpenFlow switch's switching function using both the virtual and physical interfaces and the user ID.

(3) Standard switch control protocol

The use of OpenFlow for the switch control protocol makes it possible to develop a (hardware) controller that can control the products of multiple vendors. That allows a reduction in equipment costs for servers and switches through multivendor sourcing. The development of this hardware also enables reductions in maintenance and operation costs.

6. Application areas

Next, we describe a few fields of application for overlay virtual networks as virtual network technology.

6.1 Disaster recovery and business continuity plans

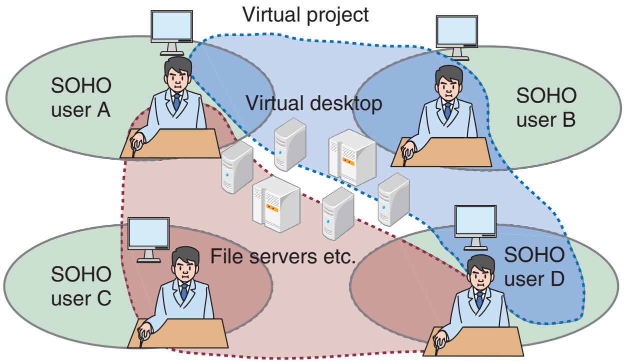
After the Great East Japan Earthquake on March 11, 2011, disaster recovery and business continuity plans that involve the backup of data used in offices have gained attention. Disaster recovery countermeasures require remote copying of data among distant offices and migration between virtual machines at different locations. The conventional movement of virtual machines involves the connection of special-purpose machines between offices and requires full network setup to be performed at both locations. That took months to accomplish in the past. By contrast, virtual network functions enable end users themselves to perform a live migration of a virtual machine to a remote location in minutes, which enables smooth disaster recovery.

NTT Information Sharing Platform Laboratories has constructed a logical network and remote live migration function by using virtual network control technology software to establish L2-over-L3 tunneling between cloud environments at the NTT Musashino Research and Development Center and the NTT Atsugi Research and Development Center. Evaluation tests have confirmed that smooth disaster recovery measures can be implemented in this manner.

6.2 Power consumption reduction

After the Great East Japan Earthquake in March 2011, which disrupted the electricity supply, attention turned to ways of reducing the power consumption of cloud services. The number of physical servers, and thus power consumption, can be reduced by concentrating the processing achieved by server virtualization using cloud services.

Nevertheless, the situation surrounding virtual machine use is changing over time, so having virtual



SOHO: small office home office

Fig. 4. Virtual desktop service.

servers running on a single physical server is not necessarily the optimum arrangement of physical servers and virtual servers.

Partial movement of a virtual machine among physical servers according to the virtual machine operation state enables concentrated processing that is always optimal to be achieved and power can be conserved by powering down empty physical servers. Furthermore, using the virtual network for migration between remote locations allows flexible operation, such as partial movement of virtual machines to areas that have a large surplus of power.

6.3 Desktop as a service

Desktop as a service (DaaS) puts the user desktop environment in the cloud so that inexpensive personal computers or smart phones can be used for the user environment while maintaining the same high degree of operability provided by a local desktop environment. Furthermore, the provision of new services by using DaaS in combination with virtual networks is being studied.

For example, it would be possible to place the desktop environments of corporate employees of affiliated companies in the cloud and also build logical networks between arbitrary employee desktops on demand. By setting up shared servers and chat servers, etc. on such logical networks, one could easily set up a shared space for projects that involve the employees of multiple organizations or related companies (**Fig. 4**).

In the past, it has been necessary to set up virtual desktops and a VPN for each project, and the end users had to access the virtual desktop of each particular project. With the combination of DaaS and virtual networks, on the other hand, the virtual desktops can be collected together for each user, and the end users only need to switch among the virtual desktops of the projects in which they are participating on demand.

7. Concluding remarks

We have introduced the network virtualization technology needed for the cloud environment. NTT Information Sharing Platform Laboratories is working toward the early implementation of overlay virtual network technology and its incorporation into CBoC (Common IT Bases over Cloud Computing) Type 1. In future work, we plan to investigate interworking between network virtualization technology within datacenters and VPN services in broadband networks, as well as gateway technology for maintaining quality and service level agreement guarantees.

Reference

[1] Over-the-top.
http://en.wikipedia.org/wiki/Over-the-top_content

Hideo Kitazume

Senior Research Engineer, Supervisor, Network Security Project, NTT Information Sharing Platform Laboratories.

He received the B.E. and M.E. degrees in computer science from Gunma University in 1987 and 1989, respectively. He joined NTT in 1989 and engaged in R&D of an ATM-LAN system, ATM traffic control studies, and the development of a global networking service platform. From 1998 to 2010, he worked in NTT EAST and engaged in the development, design, and operation of IP-VPN services. He is currently working on R&D of virtual networking technologies for cloud systems. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and the Operations Research Society of Japan.

Takaaki Koyama

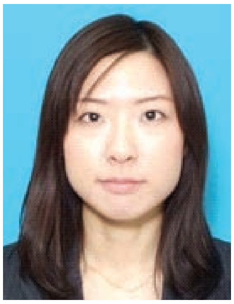
Senior Research Engineer, Network Security Project, NTT Information Sharing Platform Laboratories.

He received the B.E. and M.E. degrees in media and governance from Keio University, Kanagawa, in 1994 and 1996, respectively. He joined NTT Software Laboratory in 1996 and has been studying software CALS. Since 1999, he has been studying GMN-CL, which is a kind of IP-VPN technology and developing some network equipment. Recently, he has been interested in enterprise cloud network systems. He is a member of the Information Processing Society of Japan.

Toshiharu Kishi

Researcher, Secure Networking System Group, Network Security Project, NTT Information Sharing Platform Laboratories.

He received the B.E. and M.E. degrees in medical electronics from Chiba University in 2007 and 2009, respectively. He joined NTT Information Sharing Platform Laboratories in 2009 and worked on threat analysis of web applications. Since 2011, he has been interested in enterprise cloud network system and studying the architecture and construction of virtual networks in a cloud environment. He is a member of IEICE.



Tomoko Inoue

Researcher, Network Security Project, NTT Information Sharing Platform Laboratories.

She received the B.A. degree in literature from Ritsumeikan University, Kyoto, in 2003 and the M.A. degree in informatics from Kyoto University in 2005. She joined NTT WEST in 2005 and moved to NTT Information Sharing Platform Laboratories in 2011. Since 2011, she has been interested in enterprise cloud network systems and is studying the architecture and construction of virtual networks in a cloud environment.