OpenFlow Ethernet Fabric for Large-Scale Data Centers

TAKASHIMA Masanori, TORIGOE Keisuke, OKUMIYA Keiji, SUZUKI Yoji

Abstract

As cloud services, big data, virtualized data centers, and IoT become ever more critical to the operation of today's businesses, there is a growing need for dynamic, elastic network services capable of keeping pace with the growth of data centers. This paper introduces the OpenFlow Ethernet Fabric (OEF), an innovative network architecture used in our UNIVERGE PF Series, which offers large-scale data centers the extensibility and flexibility they need in today's rapidly evolving computing environment. OEF is a leaf-spine network fabric that boasts excellent cost performance. To get around the need to use 4,904 VLANs, the OEF uses the QinQ method. This paper examines network virtualization with OEF, as well as reviewing the issues faced by data center operators and providing a general overview of OpenFlow Spec Version 1.3 and OCP White Box, which are used in the UNIVERGE PF5340.

Keywords

OpenFlow, UNIVERGE PF Series, data center, QinQ, Ethernet fabric, SDN

1. Introduction

To keep pace with the rapid and dramatic changes that are disrupting today's business and market environments, more and more firms are leveraging the power of cloud computing, big data, and IoT to add value to the services and products they provide. The heart of this new information-oriented world is the data center. But responding to the needs of users is not easy when their requirements cannot be reliably predicted. This means that data centers need to install expandable systems that allow them to quickly offer new services and add capacity as necessary. At the same time, the new technologies proposed to meet these challenges require data centers to employ personnel able to deal with a rapidly changing software and hardware environments.

The UNIVERGE PF Series features a newly developed architecture designed to power the growth of data center businesses by enabling them to quickly scale up to meet demand, while solving the issues involved in adapting to new technology. Our new architecture makes it possible to scale up a data center that started small scale using network virtualization to build a multi-tenant network that can be configured in order for each tenant to get their own virtual network so that users are isolated form one another, while orchestration software manages and integrates all the servers and networks. This new architecture uses the reliable QinQ (IEEE 802.1ad) method, which has been found to effectively minimize the difficulties associated with the introduction of new technology, allowing telecom operators to innovate and expand while taking advantage of the human resources and technological expertise they already possess.

2. Social Requirements for the Advancement of ICT

2.1 Issues of ICT Complexity

As ICT evolves, its importance to the advancement of social and public infrastructure - such as urban infrastructure, agriculture, disaster prevention, traffic, healthcare, and energy - is increasing demand for technology specifically designed to respond to these needs. To address these requirements, NEC is supporting the advancement of ICT by focusing on four key areas - Software-Defined Networking (SDN), cloud computing, big data, and safety. We will contribute to establishing a better social infrastructure under the slogan "A richer life for people".

At the same time, while ICT for social infrastructure provides significant benefits, it also adds new layers of complexity to management and operation. So, for example, when a network system needs to be built or modified quickly, bottlenecks in management or operation can frustrate attempts to move forward. This added complexity also makes it more difficult to recruit and retain personnel with the expertise and flexibility to handle highly complex systems.

This makes the development of a reliable, efficient system that can alleviate the complexity and human resource issues of conventional ICT systems that's not just good business, but a public good as well.

2.2 The Growing Importance of Data Centers

As they seek to maintain growth of their businesses, corporations are expected to strive to, which in turn will propel growth of big data and cloud computing. Similarly, incorporation of IoT in public and social infrastructure will further accelerate this trend. But all of this depends on data centers where the cloud networks that process the big data and make possible a wide range of services are housed.

However, it is difficult for data centers to expect and build capacity for customer demands. Ongoing market volatility means that today's companies only invest as and when required, making it almost impossible to predict future needs. What this means is that data centers need to flexible enough to respond to changing demand. In other words, they need a system that's scalable, cost-effective, and that can be added to incrementally, while maintaining maximum performance and without interrupting service.

3. OCP-Based White Box Switches

The Open Compute Project (OCP)[1] was established in order to standardize (convert to open source) all hardware for data centers - such as servers, networks, storage devices, and racks. As of today, the10 G and 40 G White Box switches with Ethernet interface have been set as the standard for the network device specifications.

The UNIVERGE PF5340 (**Photo**) is a newly developed switch that incorporates NEC's SDN technology in its hardware platform (based on the industry-standard OCP), and is capable of providing better network service functions at higher speed and lower cost. The advantage of adopting the OCP is its ability to concentrate value

Photo UNIVERGE PF5340
(Top: 10 G switch, Bottom: 40 G switch).

as a network system. In other words, prompt adoption of standardized hardware that meets the needs of data center operators is now possible.

4. OpenFlow Spec. Version 1.3

OpenFlow Spec. Version 1.3 is designated as a stable version by the Open Networking Foundation (ONF)[2], a user-driven standardization organization dedicated to development and promotion of SDN through open standards development. The biggest change from OpenFlow Spec. Version 1.0 - the previous de facto standard - is that pipeline control of data paths using multiple table technology is now possible. The adoption of multiple table technology has made it possible to use the OpenFlow protocol to control various functions provided with ordinary-use commercial switching chips - functions that were difficult to use with the original OpenFlow.

The UNIVERGE PF5340 incorporates OpenFlow Datapath Abstraction (OF-DPA)[3], a pipeline for ordinary-use commercial switching chips using multiple table technology proposed by the ONF as a de facto pipeline standard. For example, OF-DPA makes possible MAC address-based path control in layer 2 and control of access lists as far as layer 4.

However, flexible control of QinQ with OpenFlow Spec. Version 1.3 or OF-DPA is currently not possible. The problem is that rules with stacked virtual LAN (VLAN) tags cannot be expressed as a flow. To deal with this

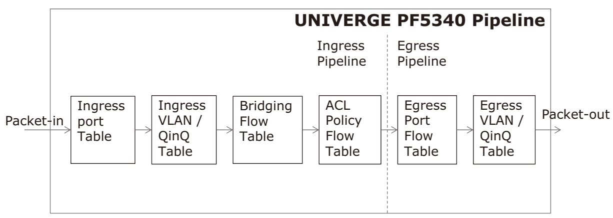

Fig.1 UNIVERGE PF5340 Pipeline.

problem, the UNIVERGE PF5340 employs an extension that makes it possible to use expressions with stacked VLAN tags as NEC-original flow rules (**Fig. 1**).

5. QinQ Network Virtualization Using the UNIVERGE PF5340 (OpenFlow Ethernet Fabric)

5.1 Scalable Network Infrastructure at Data Centers

One potential obstacle to data center scalability is the limitation and operability of VLANs (**Fig. 2**).

(1) Limit on the number of VLANs

At data centers, VLANs are used primarily to separate user networks. However, because it is only possible to configure up to 4,094 VLANs, expansion can become difficult in an environment where a large number of VLANs are used.

(2) Limited flexibility in VLAN ID assignment

When a cloud service - especially a tenant of Infrastructure as a Service (IaaS) - is to be used, it is important to interconnect between housing spaces inside data centers and on-premises. Thus, it is necessary that the tenant networks be connected to these existing networks. At this time, flexibility is required in the assignment of VLAN IDs.

(3) Limitations on the growth of services and the number of servers

As the use of cloud services expands, data centers install more servers to meet demands. However, as the number of racks that accommodate servers increases, the number of switches also increases proportionally, eventually reaching a point where it is no longer physically possible to expand the network any further. What is needed right from the start is an expandable network that makes it possible to easily increase the numbers of servers and racks with the same architecture.

(4) Solving these issues with the UNIVERGE PF Series

The UNIVERGE PF Series new network virtualization architecture built on the UNIVERGE PF5340 - which uses a QinQ method that performs VLAN tag stacking - makes it possible to expand the number of usable VLANs well beyond the 4,094 limitation, flexibly allocate VLAN IDs, and provides seamless network expansion, effectively solving the problem of expandability at conventional data centers. This architecture is called OpenFlow Ethernet Fabric (OEF).

5.2 OEF Architecture

The OEF (**Fig. 3**) is a layer 2-based network fabric comprised of two types of physical network domains - leaf and spine. Within the OEF network, the number of VLANs can be expanded using the QinQ packet format. The spine domain is used only for connection between leaf domains. In the leaf domains, the user VLAN IDs are mapped and OEF routing control is performed, facilitating connection to the racks and external networks. Moreover, VLAN IDs can be assigned arbitrarily according to ports in the leaf domain, enabling flexible ID assignment when the system is linked to existing networks.

The OEF virtual tenant network (OEF-VTN) achieves unique mapping of virtual networks and physical networks using the QinQ outer VLAN tag. Allocation of the physical networks is performed with the UNIVERGE PF6800, which controls the networks, and the UNIVERGE PF6800 Network Coordinator (UNC), which bundles them. This ensures smooth utilization and operation without forcing users and operators to deal with the complexities of network virtualization.

The OEF controls the routes in layer 2 from the controller via the OpenFlow interface. These routes are called a flooding tree. For this reason, switches inde-

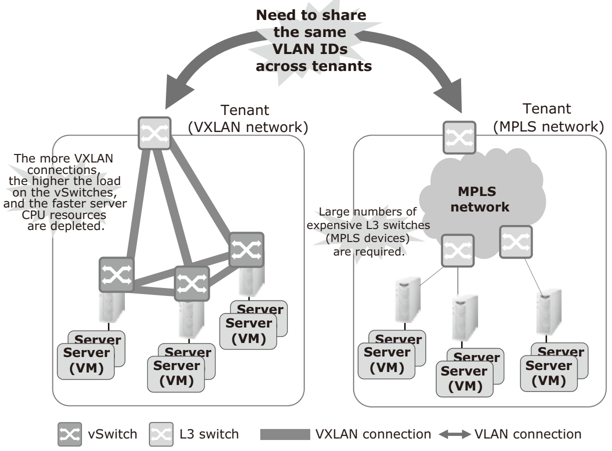

Fig.2 Problems with VLANs at data centers.

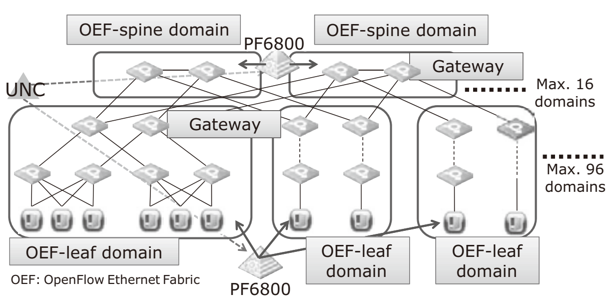

Fig.3 OEF architecture.

pendently learn MAC addresses according to the tree. Minimizing the data shared between the switches and the controller enables changes to be easily made to switches and controllers - that is, changes can be made on physical network scale, in other words. The OEF is also provided with a fail-safe function that prevents learning problems in cases where MAC addresses in different QinQ inner VLAN tags are redundant, enabling continuous operation of the network even when MAC addresses are being updated.

As many as 40 UNIVERGE PF6800 units can be managed from the UNC, making it possible to configure an OEF network with up to16 spine domains, 96 leaf domains, and about 10,000 switches.

6. Conclusion

This paper has introduced the OEF, a new network architecture featured in the UNIVERGE PF Series. This advanced new architecture provides a simple and effective solution to data center scalability, making it much easier and less costly for data centers to expand server capacity and add new services to meet the demands of cloud computing and big data. Because the new architecture is based on existing technology, it can be easily incorporated in existing systems, relieving staff from the stress of having to introduce and learn all-new technology. Not only does the OEF solve the problems of data center operators in terms of system expansion, it will also support the growth of their businesses, while helping reduce their exposure the pressure to reduce capital expenditures (CAPEX) and operating expenses (OPEX) generated by a volatile business environment and the intensification of competition. We are now planning to provide the UNIVERGE PF Series with a centralized configuration function that will help reduce introduction costs, as well as a flow-tracking function to minimize the cost of troubleshooting once the system is up and running.

We at NEC are committed to co-creating, with our customers, values that will be a driving force for further growth by leveraging the power of UNIVERGE Series networks to solve the problems of data center operators.

Reference

1) OCP (Open Compute Project):
http://www.opencompute.org/
2) ONF (Open Networking Foundation):
https://www.opennetworking.org/
3) OF-DPA (OpenFlow-DataPlane Abstruction):
https://www.broadcom.com/products/Switching/Software-Defined-Networking-Solutions/OF-DPASoftware

Authors' Profiles

TAKASHIMA Masanori
Manager
Converged Network Division

TORIGOE Keisuke
Assistant Manager
Converged Network Division

OKUMIYA Keiji
Manager
Converged Network Division

SUZUKI Yoji
Manager
Converged Network Division

Information about the NEC Technical Journal

Thank you for reading the paper.
If you are interested in the NEC Technical Journal, you can also read other papers on our website.

Link to NEC Technical Journal website

Japanese | English

Vol.10 No.3 Special Issue on Telecom Carrier Solutions for New Value Creation

Vol.10 No.3
July 2016

Special Issue TOP

General Paper

NEC Information