SDN software switch “Lagopus” and NFV enabled software node

Kazuaki OBANA
NTT Network Innovation Laboratories
SDN software switch “Lagopus”
Motivation

• **Agile and flexible networking**
  – Full automation in provisioning, operation and management
  – Seamless networking for customers

• **Server virtualization and NFV needs a high-performance software switch**
  – Small latency
  – Wire-rate with short packet (64B)

• **NO high-performance OpenFlow 1.3 software switch for carrier network**
  – 10M flow rules
  – 10Gbps-wire-rate
  – Management protocol
Lagopus: SDN/OpenFlow Software Switch

- The world’s highest performance SDN software switch with 1M flow entries
- To be released as OSS soon

- Best OpenFlow 1.3 compliant software-based switch
  - Ryu certification score: 920/928 (http://osrg.github.io/ryu/certification.html)
  - Multi tables, Group tables support
  - MPLS, PBB, QinQ, support

- ONF standard specification support
  - OpenFlow Switch Specification 1.3.3
  - OF-CONFIG 1.1

- Multiple data-plane configuration
  - High performance software data-plane on Intel x86 bare-metal server
    - Intel DPDK, Raw socket
  - Bare metal switch

- Various management/configuration interfaces
  - OF-CONFIG, OVSDB, CLI
  - SNMP, Ethernet-OAM functionality

- Modular architecture
- To be released as OSS soon
  - Apache v2 license
OpenFlow 1.3 Conformance Status

- Conformance test results by Ryu Certification
WAN-DC Gateway Demonstration

- Gateway performance between WAN (IP VPN/MPLS) and data center (VLAN)

**Diagram Description**

- **SDN / OpenFlow controller**: Ryu
  - OpenFlow version 1.3
- **Agent (OpenFlow protocol processing)**
- **Data-plane (Packet processing)**: I/O TX, I/O RX, flow lookup & processing
- **OpenFlow port #1**, **OpenFlow port #2**
- **General purpose PC server**
- **Payload (size)**
  - IP
  - VLAN
  - Ethernet
- **Wide Area Network**
  - Payload (size)
  - IP
  - MPLS
  - Ethernet
- **Data Center Network**
  - Payload (size)
  - IP
  - VLAN
  - Ethernet

**Traffic tester generates 10Gbps pseudo traffic**

**DUT Hardware Specifications**
- CPU: Dual Intel Xeon Processor E5-2660 (Sandy bridge)
  - 16 thread/8 core, 20M Cache, 2.2GHz, 8.00GT/s OPI
- Memory: DDR3-1600 ECC 64GByte (Quad-channel: 8x8GByte構成)
- Chipset: Intel C602
- NIC: Intel Ethernet Converged Network Adapter X520-DA2 (Intel 82599ES, PCIe v2.0)

**DUT Software Specifications**
- OS: Ubuntu 12.04 LTS x86_64
- Linux kernel 3.8.0-29-generic
- GCC version 4.6.3-1ubuntu5
glibc version 2.15-ubuntu10.5
- Intel DPDK version 1.5.0
Publicity Activities

- **Event demonstrations: 4**
  - Okinawa Open Days, R&D Forum, ONS, O₃ Symposium

- **Press releases: 3**
  - O₃ Project launched for achieving the world’s first wide area SDN
  - Prototype High Performance SDN Software Switch Launched
  - Technologies developed for providing flexible wide area networks with SDN, a global first

- **Research presentations: 10**
  - SDN Japan 2013, TEIN4, Okinawa Open Days, OCS, TELESAA, JICA, ONS, ONF Workday, O₃ Symposium
NFV enabled software node
Motivation

• **Rapid development for new network service demands**
  – Hardware-based new network equipment takes much more time to develop

• **Reducing costs for develop and maintain “the long tail” network services**
  – Reducing CAPEX/OPEX is needed

• **NFV and SDN are suitable answers for above challenges but there is no software node both enabled for carrier networks**
Requirements of Network Nodes for NFV

- High-performance and low-latency packet processing
- Easy development
- Multi-tenancy and isolation
- Fine-grained control of network functions
- Control interface for external systems
- Monitoring
- Fault tolerance
3 options of NFV deployment

Orchestrator / Network Management System / VM Manager / OpenFlow Controller

- Bare-metal server
  - NW Func.

- Hypervisor-enabled server
  - NFV VM
    - NW Func.
  - NFV VM
    - NW Func.
  - Hypervisor

- Hypervisor-enabled server
  - NFV VM
    - NW Func.
  - NFV VM
    - NW Func.
  - NFV VM
    - NW Func.
  - Hypervisor
  - Software-based OpenFlow switch

Hardware-based OpenFlow switch (ASIC based)
2 options of NFV deployment using NSOD*

* NSOD stands for Network Service-Oriented Distributor

Hypervisor-enabled server

- NFV VM
  - NW Func.
- Hypervisor
- Software-based OpenFlow switch

Direct attached model

Hardware-based OpenFlow switch (ASIC based)

Hypervisor-enabled server

- NFV VM
  - NW Func.
- Hypervisor
- Software-based OpenFlow switch

Internal OpenFlow model
How to make it faster

- Bypassing OS networking
  - Nowadays polling model is more suitable for high-speed packet processing rather than interrupt-driven model
  - Memory copies are dominant reason for low-performance

- Para-virtualized I/O acceleration
  - Virtio poll mode driver is the key feature for accelerating Linux kvm hypervisor for high speed networking
Our proposal of NFV design approach

Conventional design approach

Bare-metal server

- Socket Based Network Application, etc.
- new function
  - Network Stack (basic or advanced)
- Packet handling Layer as network device driver
- OS Kernel (Linux kernel or etc.)

Our NFV design approach

- Network Application Logic
- Userspace Standardized Network Stack
- Abstraction
- Packet handling API
- Virtual Machine (NFV VM)
  - software-based OpenFlow Switch
  - Bypass
    - Packet handling API
      - OS Kernel
- VIRTUALIZED Environment

- Bare-metal

- Network Application Logic
- Standardized Network Stack
- Abstraction
- Packet handling API
- Bypass
  - OS Kernel

Copyright©2014 NTT corp. All Rights Reserved.
Application example: virtual BRAS server

* : PMD stands for Poll-Mode Driver
Key takeaways

■ There was NO high-performance OpenFlow 1.3 software switch for carrier network before “Lagopus”
■ “Lagopus” passed almost all of OpenFlow 1.3 conformance tests
■ “Lagopus” will be available as OSS soon
■ Rapid development and reducing costs are needed
■ Making it faster by bypassing OS networking and para-virtualized I/O acceleration
■ NFV design approach with user land application and abstraction

These researches have been executed under the Commissioned Research “O3 project” of Ministry of Internal Affairs and Communications