High Performance NFV Infrastructure (NFVI): DPDK Host Applications with Neutron/OpenStack and VNF Acceleration

Vincent JARDIN
6WIND – CTO
@VincentJardin
Telecom - The technology behind
Telecom - The technology behind
Telecom - The technology behind

Proprietary hardware
Telecom - The technology behind

Full re-deployment for each new generation.
Network Architecture Transformation Towards NFV

Application
Proprietary Hardware Platform

Application
Proprietary Hardware Platform

Application
Proprietary Hardware Platform

Application
Generic Hardware Platforms

Virtualization

Application
Generic Hardware Platform

ironic

openstack

openstack
1. SDN: freedom of ability to create any networks and their overlays

2. Troll: is an Openflow HW/ASIC switch a SDN equipment?
   - How to support any new “what if”? => wait and buy a new switch

- SDN != NFV, but SDN can be made of VNF & NFVI
… but SysOps/DevOps virtualizing their networks…

“...It's time we face reality, my friends. ... We're not exactly rocket scientists.”

Where is my network performance?
Performance first
High Performance East-West Communications

Virtual Machine
Application
Linux

Virtual Machine
Application
Windows

Virtual Machine
Application
Any OS

Virtual Machine
Application
Any OS

Hypervisor
Virtual Switch
Throughput
Hardware Independence
What if SRIOV?

or what if XYZ PCI passthru technologies?

High Performance East-West Communications

Virtual Machine
Application
Linux

Virtual Machine
Application
Windows

Virtual Machine
Application
Any OS

Virtual Machine
Application
Any OS

Hypervisor

Throughput
SR-IOV

Hardware independence
Number of VMs is limited by virtual switching / networking capacity in both compute and network nodes.
Virtual Network Functions (VNFs)

Virtualization Layer

Hardware resources

NFV / ETSI Simplified Architecture
Network Function Virtualization Use Case

- VNF (North-South) and service chain (East-West) throughput is limited by Linux virtual switching / networking capacity
Appliance Virtualization Use Case

- Appliance is based on specialized architectures
  - Rigid
  - High development costs
  - Long TTM
Typical NFV Performance Bottlenecks

1. Driver Level Bottleneck
2. Virtual Switch Bottleneck
3. Communication Bottleneck - Host vs Guest OS
4. Virtual Machine Bottleneck
6WINDGate for Industry-Leading Processor Platforms

Architecture-independent “Fast Path Modules”
- Generic, processor-independent source code
- Cycle-level and pipeline-level optimizations

Architecture-specific "Fast Path Networking SDK"
- Zero-overhead API for fast path modules
- Support for processor-specific features and resources
- Leverages processor suppliers' SDKs
Linux Compatibility is Critical
Linux Acceleration via 6WINDGate

- Standard Linux functions are accelerated by 6WINDGate

Linux Networking Stack
Linux Kernel

iptables  iproute2

Protocol Tables
Shared Memory
Statistics

Fast Path Modules

Fast Path

Fast Path Configuration
Fast Path Statistics

OpenStack
Quagga
OpenFlow
HAProxy
NGINX
libvirt

©6WIND 2014
Neutron’s protocols – strong needs for a fast path

Say no to proprietary plugins
Say no to SRIOV to be SDN ready.
Accelerate Neutron

Openstack

6WINDGate fast path

VLAN
IPv4/IPv6 Forwarding
IPsec IPv4/IPv6
OVS Acceleration
VXLAN

Link Aggregation
IPv4/IPv6 Multi-cast
IPv4/IPv6 Tunneling (IPinIP)
GRE
MPLS/VPLS Encapsulation
Ethernet Bridging

NAT
Filtering IPv4/IPv6
QoS
Flow Inspection

Neutron’s protocols

Nova
Neutron

6WINDGate FPN-SDK

6WINDGate DPDK

Mellanox ConnectX®-3 EN Series PMD
Intel® QuickAssist Crypto
Cavium NITROX SDK 5.X Crypto
VIRTIO Guest XEN-KVM PMD

Emulex OCE14000 Series PMD
Intel® Multi-Buffer Crypto
VMXNET3 Guest VMware PMD
Fast vNIC PMD

©6WIND 2014

v9.3 | 19
6WINDGate Extensions to Virtualization

Drivers for Virtual Appliance
- 6WIND drivers for high performance communications
- Standard drivers for existing Virtual Appliances
- Extensible for all OSs

Virtual Acceleration
- 6WIND drivers for high performance communications
- Accelerated virtual switch and bridging
- Extended network services
- Dpdk.org with multi-vendor NIC support
6WIND Virtual Accelerator in OpenStack Compute Node

- 240Gbps 6WIND Virtual Accelerator throughput on 12 cores of Xeon E5-2697 v2 @ 2.70GHz

- 1 core provides a 20Gbps Virtual Accelerator bandwidth

- Examples on a dual socket / 24 cores server
  - 120Gbps North-South traffic delivered to standard VMs or VNFs with 12 cores remaining for VMs
  - 40Gbps North-South traffic with 20 cores remaining for VMs
  - 40Gbps North-South traffic and 160 Gbps East-West traffic for service chaining with 3 Turbo Boosted VNFs
6WIND Virtual Accelerator in OpenStack Network Node

- 255Gbps 6WIND Virtual Accelerator throughput on 8 cores of Xeon E5-2697 v2 @ 2.70GHz
- 1 core provides a 30 Gbps Virtual Accelerator bandwidth
- Examples
  - 40Gbps North-South traffic on a dual socket / 24 cores server hosting both Network and Compute Node on 6 cores, with 18 cores remaining for VMs
  - 240Gbps North-South traffic on a single socket / 8 cores server feeding six 40Gbps Compute Nodes, each hosting a 3 Turbo Boosted VNFs service chain
Data Center Virtualization Use Case

- 6WIND Virtual Accelerator removes Linux performance bottlenecks and enables high VM density without any change to the environment
- Number of VMs is limited by virtual switching / networking capacity in both compute and network nodes
Network Function Virtualization Use Case

- VNF (North-South) and service chain (East-West) throughput is limited by Linux virtual switching / networking capacity

- 6WIND Virtual Accelerator removes Linux performance bottlenecks and maximizes North-South and East-West throughput with higher VNF density without any change to the environment
Appliance Virtualization Use Case

- Appliance is based on specialized architectures
  - Rigid
  - High development costs
  - Long TTM

- 6WIND Virtual Accelerator removes Linux networking performance bottlenecks on standard servers and enables flexibility brought by virtualization
6WINDGate NFVI + VM Performance Comparison

- **Fast Path**
  - IPv4/IPv6 Forwarding
  - DPDK
  - Fast vNIC PMD

- **Virtual Machine**
  - Fast vNIC PMD
  - Virtual Networking
  - DPDK

- **Hypervisor**
  - Linux kernel

- **R720**
  - 12 x 10G Ports

- **Traffic Generator**
  - ixia
Test 1:
Standard Open vSwitch + Virtio

L2 Throughput

7.2 Gbps

Limited Bandwidth To Linux Based Virtual Machines

©SWIND 2014
Test 2: 6WINDGate OVS Acceleration + Fast vNIC Linux

L2 Throughput

- 7.2 Gbps
- 59.2 Gbps

9X Throughput Performance Increase

12 x 10G Ports

Fast Path

- IPv4/IPv6 Forwarding
- Fast vNIC Linux
- Fast vNIC PMD
- Virtual Accelerator
- DPDK

Bottleneck

Linux Based Virtual Machine

Hypervisor

Linux kernel

ixia Traffic Generator

DELL R720

12 x 10G Ports

©6WIND 2014
Test 3:
6WINDGate OVS Acceleration + Fast vNIC PMD

L2 Throughput

7.2 Gbps
59.2 Gbps
118.4 Gbps

Wire Speed Performance

12 x 10G Ports
Compute nodes with 6WINDGate, Openstack-horizon

Compute node/host:
# yum install 6windgate*.rpm
# systemctl enable 6windgate.service
6WIND Virtual Accelerator Performance Summary

For Data Center Virtualization

- Servers are dual socket Xeon E5-2697 v2 @ 2.70GHz (12 cores per socket)
- Independently of HW constraints (number of ports per socket for example)

For Network Function Virtualization

For Appliance Virtualization
SPEED MATTERS
Turbo Boost Linux
The OEM Advantage

Unlock Hidden Performance
Reduce Time-To-Market
Enable Transition To SDN / NFV

L2-L4 Acceleration
IPsec VPN Gateways
TCP / UDP Termination
Virtual Switching
DPDK
And More...

增加数据平面性能
无变更Linux环境
可移植于所有主要平台
支持广泛的协议集

增加数据平面性能
无变更Linux环境
可移植于所有主要平台
支持广泛的协议集

Packet Processing Software
Up To 10X Network Performance
backup
Virtual Accelerator
Lowest Latency and Flexible Chaining

6WINDGate Virtual Accelerator
- Hardware independent virtual switching (NIC driver)
- Aggregate 500 Gbps bandwidth with low latency
- No external limit to number of chained VNFs
- DPDK ready

Physical Switching Limitations
- Hardware dependent switching (SR-IOV, RDMA, NIC embedded switching)
- Throughput is limited by PCI Express (50 Gbps) and faces PCI Express and DMA additional latencies
- Available PCI slots limit the number of chained VNFs
- At 30 Gbps a single VNF is supported per node!
Introduction to NFV & and why Openstack – 1 slide

- **NFV = Network Function Virtualization**
  - How do I spawn a VM that is an IPsec router, that is a L2TP/DSL (LNS) server, that is a firewall, etc…? Instead of using a physical HW?

- **You need network cables …**
  - Network cables => **NFVI** – Network Function Virtualization Infrastructure
    - It provides interconnect between the physical cables and the virtual network equipments that are spawned
    - L2 switch, L3 switch, ACL/Switch

- **… and network equipments**
  - Network Function (equipments) => Virtualized => **VNF** – Virtual Network Function (vNF)
    - vIPsec, vBRAS, vRouter, vFirewall, vXXX

- **Openstack provides the framework to spawn Virtual Machines**
  - Virtual Machines can have multiple virtual network interfaces (vmxnet3, virtio) per VM
    - vm1# ip link show | grep virtio | wc –l
  - Any VMs with multiple network interfaces can be any network nodes
    - IPsec, Firewall, Proxy
  - To some extends, the VMs are the containers of the **VNF**
    - run vIPsec, vFirewall, vProxy

- **Openstack/Neutron and Openstack/Nova VIF drivers provides the automations of the provisioning for interconnecting the VMs**
  - vhost/tuntap, brctl, OVS, L3/routes
  - To some extends, this software interconnect of VMs is the **NFVI**