NFV with OpenStack

Vincent Jardin  
CTO  
6Wind  
@VincentJardin

Nick Barcet  
VP Products  
eNovance  
@nijaba
Hello? The technology behind
Hello? The technology behind
Hello? The technology behind

Proprietary hardware
Hello? The technology behind Full re-deployment for each new generation network
What if this went went virtual?
What if this went went virtual?
What if this went scalable?
What if this went virtual?
What if this went scalable?
What if this was just like cloud?
Creation of ETSI NFV

• In November 2012 seven of the world's leading telecoms network operators selected ETSI to be the home of the Industry Specification Group for NFV.

• 2 years later: 230 individual companies including 37 of the world's major service providers as well as representatives from both telecoms and IT vendors.
NFV business opportunities

- Latest major disruption in the telco marketplace since IP introduction
- Create opportunities for service providers:
  - To accelerate development of new services
  - To implement Network and IT convergence
- Force NEP
  - To change their business model (transform themselves as Software Providers)
  - To redesign their software
OpenStack NFV subteam

- NFV support for OpenStack aims to provide the best possible infrastructure for such workloads to be deployed in, while respecting the design principles of a IaaS cloud.

- In order for VNF to perform correctly in a cloud world, the underlying infrastructure needs to provide a certain number of functionalities which range from scheduling to networking and from orchestration to monitoring capacities.

- This means that to correctly support NFV use cases in OpenStack, implementations may be required across most, if not all, main OpenStack projects, starting with Neutron and Nova.

https://wiki.openstack.org/wiki/Teams/NFV
3 main specific problems for NFVi

- North/South connectivity requirements
3 main specific problems for NFVi

• North <-> South connectivity requirements

• High Service Level Agreement requirements
3 main specific problems for NFVi

- North <-> South connectivity requirements
- High Service Level Agreement requirements
- East <-> West connectivity requirements
NFV Simplified Architecture

Virtual Network Functions (VNFs)

- VNF
- VNF
- VNF
- VNF
- VNF

Network Functions Virtualization Infrastructure (NFVi)

- Virtual Compute
- Virtual Storage
- Virtual Network

Virtualization Layer

Hardware resources

NFV Management and Orchestration
## Use Cases

<table>
<thead>
<tr>
<th>Workload Type</th>
<th>Description</th>
<th>Characteristics</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data plane</td>
<td>Tasks related to packet handing in an end-to-end communication between edge applications.</td>
<td>Intensive I/O requirements - potentially millions of small VoIP packets per second per core Intensive memory R/W requirements</td>
<td>CDN cache node Router IPSec tunneller Session Border Controller - media relay function</td>
</tr>
<tr>
<td>Control plane</td>
<td>Any other communication between network functions that is not directly related to the end-to-end data communication between edge applications.</td>
<td>Less intensive I/O and R/W requirements than data plane, due to lower packets per second More complicated transactions resulting in (potentially) higher CPU load per packet.</td>
<td>PPP session management Border Gateway Protocol (BGP) routing Remote Authentication Dial In User Service (RADIUS) authentication in a Broadband Remote Access Server (BRAS) network function Session Border Controller - SIP signaling function IMS core functions (S-CSCF / I-CSCF / BGCF)</td>
</tr>
<tr>
<td>Signal processing</td>
<td>All network function tasks related to digital processing</td>
<td>Very sensitive to CPU processing capacity. Delay sensitive.</td>
<td>Fast Fourier Transform (FFT) decoding Encoding in a Cloud-Radio Access Network (C-RAN) Base Band Unit (BBU) Audio transcoding in a Session Border Controller</td>
</tr>
<tr>
<td>Storage</td>
<td>All tasks related to disk storage.</td>
<td>Varying disk, SAN, or NAS, I/O requirements based on applications, ranging from low to extremely high intensity.</td>
<td>Logger Network probe</td>
</tr>
</tbody>
</table>
## Blueprints implemented in Juno

<table>
<thead>
<tr>
<th>Description</th>
<th>Project(s)</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support two interfaces from one VM attached to the same network</td>
<td>Nova</td>
<td>Design Approved / Implemented</td>
</tr>
<tr>
<td>SR-IOV Networking Support</td>
<td>Nova</td>
<td>Design Approved / Needs Code Review</td>
</tr>
<tr>
<td>Virt driver guest vCPU topology configuration</td>
<td>Nova</td>
<td>Design Approved / Implemented</td>
</tr>
<tr>
<td>Evacuate instance to scheduled host</td>
<td>Nova</td>
<td>Approved / Implemented (juno-2)</td>
</tr>
</tbody>
</table>
## What's brewing for Kilo? (Highest Priority)

<table>
<thead>
<tr>
<th>Description</th>
<th>Project(s)</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>- VLAN trunking networks for NFV</td>
<td>Neutron</td>
<td>New</td>
</tr>
<tr>
<td>- VLAN tagged traffic transmissible over a tenant network</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- decomposition of VLAN trunks to virtual networks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- VLAN tagged traffic to a physical appliance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- management of VLANs on ports as subports</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permit unaddressed interfaces for NFV use cases</td>
<td>Neutron</td>
<td>New</td>
</tr>
</tbody>
</table>


## What's brewing for Kilo (the rest)

<table>
<thead>
<tr>
<th>Description</th>
<th>Project(s)</th>
<th>Status</th>
<th>Description</th>
<th>Project(s)</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discless VM</td>
<td>Nova</td>
<td>Under discussion</td>
<td>Port mirroring</td>
<td>Neutron</td>
<td>Under discussion</td>
</tr>
<tr>
<td>Framework for Advanced Services in Virtual Machines</td>
<td>Neutron</td>
<td>Under Discussion</td>
<td>Schedule vms per flavour cpu overcommit</td>
<td>Nova</td>
<td>New</td>
</tr>
<tr>
<td>I/O (PCIe) Based NUMA Scheduling</td>
<td>Nova</td>
<td>Design Approved / Needs Code Review</td>
<td>Snabb NFV mechanism driver</td>
<td>Neutron</td>
<td>Approved</td>
</tr>
<tr>
<td>Network QoS API</td>
<td>Neutron</td>
<td>Under discussion</td>
<td>Solver Scheduler - complex constraints scheduler with NFV use cases</td>
<td>Nova</td>
<td>Design review in progress</td>
</tr>
<tr>
<td>Neutron Services Insertion, Chaining, and Steering</td>
<td>Neutron</td>
<td>Design Approved / Needs Code Review</td>
<td>Support userspace vhost in ovs vif bindings</td>
<td>Nova</td>
<td>Design review in progress</td>
</tr>
<tr>
<td>NIC state aware scheduling</td>
<td>Nova</td>
<td>Rejected</td>
<td>Traffic Steering Abstraction</td>
<td>Neutron</td>
<td>Design review in progress</td>
</tr>
<tr>
<td>Open vSwitch to use patch ports in place of veth pairs for vlan n/w</td>
<td>Neutron</td>
<td>Superseded / Unknown</td>
<td>VIF_VHOSTUSER (qemu vhost-user) support</td>
<td>Nova</td>
<td>Approved</td>
</tr>
<tr>
<td>Open vSwitch-based Security Groups: Open vSwitch Implementation of FirewallDriver</td>
<td>Neutron</td>
<td>Design review in progress</td>
<td>Virt driver guest NUMA node placement &amp; topology</td>
<td>Nova</td>
<td>Design Approved / Needs Code Review</td>
</tr>
<tr>
<td>OVF Meta-Data Import via Glance</td>
<td>Glance</td>
<td>New</td>
<td>Virt driver large page allocation for guest RAM *</td>
<td>Nova</td>
<td>Design Approved / Needs Code Review</td>
</tr>
<tr>
<td>Persist scheduler hints</td>
<td>Nova</td>
<td>Design review in progress</td>
<td>Virt driver pinning guest vCPUs to host pCPUs</td>
<td>Nova</td>
<td>Design Approved / Needs Code Review</td>
</tr>
</tbody>
</table>
Openstack related challenges

- Cultural Changes:
  - Introduction of devops, Continuous Integration, ...
  - Work with opensource communities
  - No more standards...
  - 80% is good enough...

- Technical, R&D:
  - Openstack is evolving very fast (major release every 6 months)
  - Limited availability of skills
  - Redesign of application to become cloud application

- Operation:
  - Need to educate/train people
North ↔ South
East ↔ West

Virtual Network Performances
With 6WIND
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

Hardware Independence

Throughput
SR-IOV
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
- Fast Path Configuration
- Fast Path Statistics

Linux Kernel
- Fast Path Statistics

Protocol Tables
- Shared Memory
- Statistics

Fast Path Modules

Fast Path

Software Tools:
- openstack
- Quagga
- iptables
- iproute2
- HAPROXY
- NGINX
- libvirt
- Debian
- Fedora
- Red Hat Linux
- Ubuntu
Neutron’s protocols – strong needs from a fast path

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

NICs
6WINDGate NFVI + VM Performance Comparison

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

Virtual Machine
- Fast vNIC PMD
- Virtual Networking
- Linux kernel

Hypervisor

Fast Path
- Fast vNIC PMD
- Virtual Networking
- DPDK

Dell R720
- 12 x 10G Ports

ixia Traffic Generator
Test 1:
Standard Open vSwitch + Virtio

L2 Throughput

7.2 Gbps

Limited Bandwidth To Linux Based Virtual Machines

Bottleneck

IPv4/IPv6 Forwarding
Virtio Guest Linux

Bottleneck

Virtio Host Linux
Linux kernel
Kernel Drivers

Bottleneck

Hypervisor

12 x 10G Ports

Dell R720

ixia
Traffic Generator

Linux Based Virtual Machine
Test 2:
6WINDGATE OVS Acceleration + Fast vNIC Linux

L2 Throughput
7,2 Gbps
59 Gbps

9X Throughput Performance Increase

12 x 10G Ports

IPv4/IPv6 Forwarding
Fast vNIC Linux

Linux Based Virtual Machine

Fast vNIC PMD
Virtual Networking
DPDK

Fast Path

Bottleneck

Hypervisor

Linux kernel

ixia Traffic Generator

DELL R720

12 x 10G Ports
Test 3:
6WINDGate OVS Acceleration + Fast vNIC PMD

L2 Throughput

7,2 Gbps
59 Gbps
118 Gbps

Wire Speed Performance

DPDK Based Virtual Machine

Hypervisor

Linux kernel

12 x 10G Ports

Dell R720

Traffic Generator
Compute nodes with 6WINDGate, Openstack-horizon

Compute node/host:
# yum install 6windgate*.rpm
# systemctl enable 6windgate.service
Virtual Switch-Based NFVI
Lowest Latency and Flexible Chaining

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

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!
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...

Packet Processing Software
Up To 10X Network Performance

Increase Data Plane Performance
No Change To Linux Environments
Portable Across All Major Platforms
Support Extensive Set Of Protocols
DPDK.org meetup – Jeudi prochain – Santa Clara @ 6WIND
THANK YOU!