A Host-based Performance Comparison of 40G NFV Environments Focusing on Packet Processing Architectures and Virtual Switches

R. Kawashima, S. Muramatsu, H. Nakayama
T. Hayashi, and H. Matsuo
Contents

- **Backgrounds**
  - NFV
  - Related work
- **Packet Processing Architectures**
- **Virtual Switches**
- **Evaluation**
- **Conclusion**
Network Functions Virtualization (NFV)

Traditional

- HW-based
- High cost
- Low flexibility
- Closed

NFV

- SW-based
- Low cost
- High flexibility
- Open
Which architecture or vswitch should be used for NFV hosts?

How much performance differs depending on its architecture and vswitch?
Related Work

- **Performance of baremetal servers [1][2]**
  - They focused on performance bottlenecks

- **Performance of KVM and Container-based virtualization [3]**


Our Evaluation

- Combination of PM and VM architectures and vswitches
- Throughput and Latency/Jitter
- Intel and Mellanox 40GbE NICs
- SR-IOV
Contents

- Backgrounds
  - NFV
  - Related work
- Packet Processing Architectures
- Virtual Switches
- Evaluation
- Conclusion
Packet Processing Architecture

- A way to forward packets (NIC <=> Applications)
  - Interruption or Polling
  - Single core or Multi cores
  - Kernel space or User space
  - Packet buffer structure
  - Packet buffer management

The architecture has a major effect on the performance!
Three Architectures

**NAPI**
- User
- vSW
- Network Stack
- Rx
- NIC
- Interruption or Polling

**Netmap**
- App
- Netmap API
- Network Stack
- VALE
- Tx
- Rx
- Interruption or Polling

**DPDK**
- vSW
- DPDK API
- TX
- RX
- PMD
- Polling

User -> Kernel

NIC

DMA
Contents

- Backgrounds
  - NFV
  - Related work
- Packet Processing Architectures
- Virtual Switches
- Evaluation
- Conclusion
Virtual Switch

- A virtual switch bridges the host and VMs

- The virtual switch has an impact on performance too!
## Six Virtual Switches

<table>
<thead>
<tr>
<th>Name</th>
<th>Running Space</th>
<th>Architecture</th>
<th>Virtual I/O Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linux Bridge</td>
<td>Kernel</td>
<td>NAPI</td>
<td>TAP/vhost-net</td>
</tr>
<tr>
<td>OVS</td>
<td>Kernel</td>
<td>NAPI</td>
<td>TAP/vhost-net</td>
</tr>
<tr>
<td>VALE</td>
<td>Kernel</td>
<td>Netmap</td>
<td>(QEMU)</td>
</tr>
<tr>
<td>L2FWD-DPDK</td>
<td>User</td>
<td>DPDK</td>
<td>-</td>
</tr>
<tr>
<td>OVS-DPDK</td>
<td>User</td>
<td>DPDK</td>
<td>vhost-user</td>
</tr>
<tr>
<td>Lagopus</td>
<td>User</td>
<td>DPDK</td>
<td>(to be supported)</td>
</tr>
</tbody>
</table>
Contents

❖ Backgrounds
  ● NFV
  ● Related work

❖ Packet Processing Architectures

❖ Virtual Switches

❖ Evaluation

❖ Conclusion
Goals

Clarify performance characteristics of existing systems

Propose appropriate NFV host environments

Find a proper direction for performance improvement
Experiment 1 (Baremetal)

<table>
<thead>
<tr>
<th>Physical</th>
<th>Server 1</th>
<th>Server 2 (DUT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>CentOS 7.2</td>
<td></td>
</tr>
<tr>
<td>CPU</td>
<td>Core i7-3770 3.40 GHz (4 cores with HT)</td>
<td>Core i7-6700K 4.00 GHz (4 cores with HT)</td>
</tr>
<tr>
<td>Memory</td>
<td>64 GB</td>
<td>32 GB</td>
</tr>
<tr>
<td>VMM</td>
<td>-</td>
<td>KVM</td>
</tr>
<tr>
<td>NIC</td>
<td>Intel Ethernet Converged Network Adapters XL710 Mellanox ConnectX-3 EN</td>
<td></td>
</tr>
</tbody>
</table>
Throughput

Throughputs differ depending on the NIC type

Throughputs with short packet sizes are far from wire rate
Latency

L2FWD-DPDK and Lagopus show worse latency

Jitter values are less than 10 µs
Experiment 2 (VM)

**Server 1**
- pktgen-dpdk (Throughput)
- OSTA (Latency)
- eth1
- eth2

**Server 2 (DUT)**
- vSwitch
- eth1
- eth2

**40GbE**

**UDP traffic**

<table>
<thead>
<tr>
<th>Physical</th>
<th>Server 1</th>
<th>Server 2 (DUT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>CentOS 7.2</td>
<td></td>
</tr>
<tr>
<td>CPU</td>
<td>Core i7-3770 3.40 GHz (4 cores with HT)</td>
<td>Core i7-6700K 4.00 GHz (4 cores with HT)</td>
</tr>
<tr>
<td>Memory</td>
<td>64 GB</td>
<td>32 GB</td>
</tr>
<tr>
<td>VMM</td>
<td>-</td>
<td>KVM</td>
</tr>
<tr>
<td>NIC</td>
<td>Intel Ethernet Converged Network Adapters XL710 Mellanox ConnectX-3 EN</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Virtual</th>
<th>VM (Experiment 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>CentOS 7.2</td>
</tr>
<tr>
<td>vCPU</td>
<td>2 cores</td>
</tr>
<tr>
<td>Memory</td>
<td>4 GB</td>
</tr>
<tr>
<td>vNIC</td>
<td>virtio-net</td>
</tr>
</tbody>
</table>
Throughput

NAPI/ vhost-net

The virtualization overhead is fairly large

DPDK/ vhost-user
Latency

NAPI/vhost-net

DPDK/vhost-user

The virtualization amplifies jitters
SR-IOV shows the best performance!

SR-IOV lacks flexibility of flow handling
Adequate NFV Host Environment

Throughput
Latency/Jitter
Usability

NIC
- Intel XL710
- Mellanox ConnectX-3

Architecture/Virtualization
- NAPI/vhost-net
- DPDK/vhost-user

Virtual Switch
- Linux Bridge
- OVS
- L2FWD-DPDK
- OVS-DPDK
Contents

- Backgrounds
  - NFV
  - Related work
- Packet Processing Architectures
- Virtual Switches
- Evaluation
- Conclusion
Conclusion

❖ **Summary**

- We have evaluated NFV host environments with 40GbE
  - A NIC device affects performance characteristics
  - DPDK should be used for both the host and the guest
  - We cannot reach the wire rate with short packet sizes
  - Virtualization worsens both throughput and latency
  - SR-IOV showed better throughput and latency

❖ **Future Work**

- Further evaluations
  - VALE/Netmap based virtualization
  - VALE and Lagopus on the VM
  - Bidirectional and lots of flows