



OVS, DPDK AND SOFTWARE DATAPLANE ACCELERATION

FOSDEM 2016

Kevin Traynor <kevin.traynor@intel.com>

Legal Disclaimer

General Disclaimer:

© Copyright 2016 Intel Corporation. All rights reserved. Intel, the Intel logo, Intel Inside, the Intel Inside logo, Intel. Experience What's Inside are trademarks of Intel. Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others.

Technology Disclaimer:

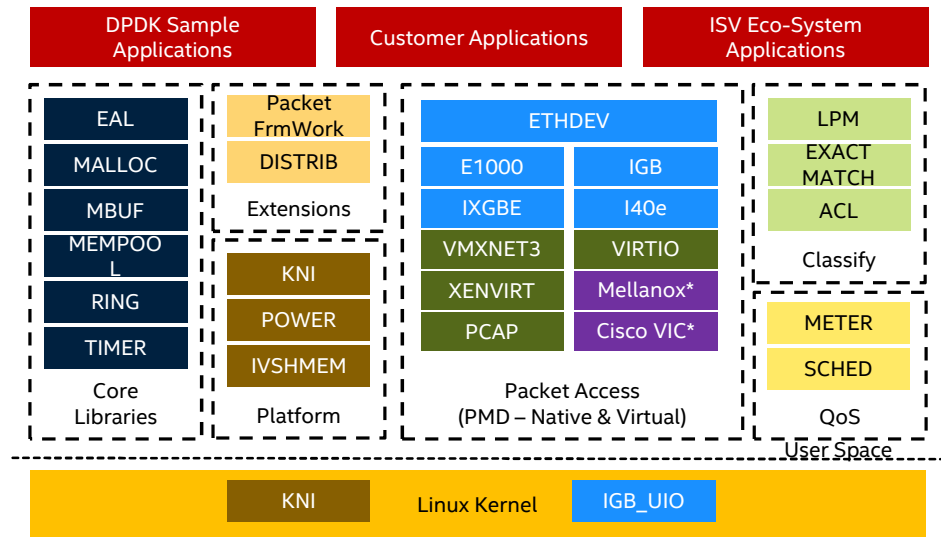
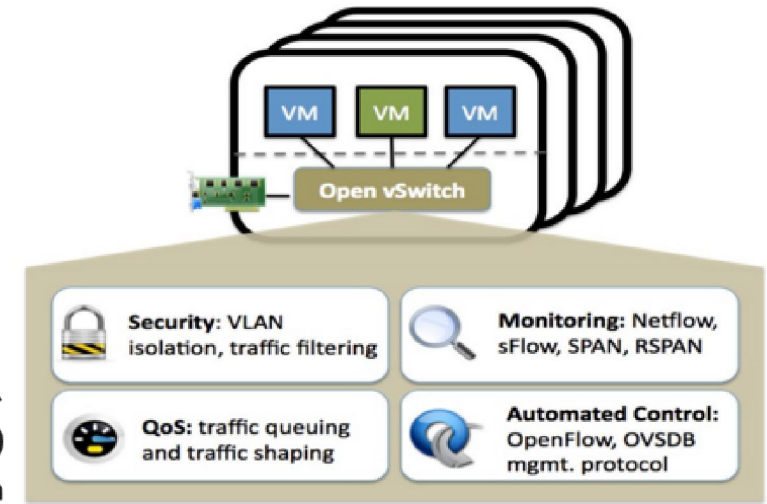
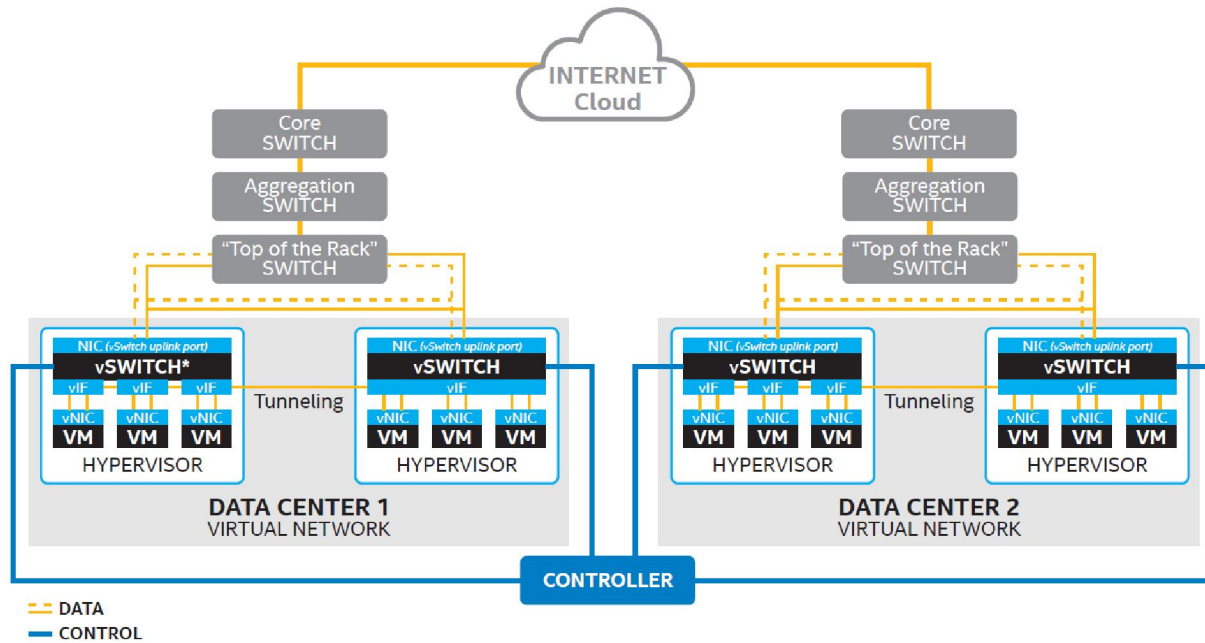
Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at [intel.com].

Performance Disclaimers (include only the relevant ones):

Cost reduction scenarios described are intended as examples of how a given Intel- based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

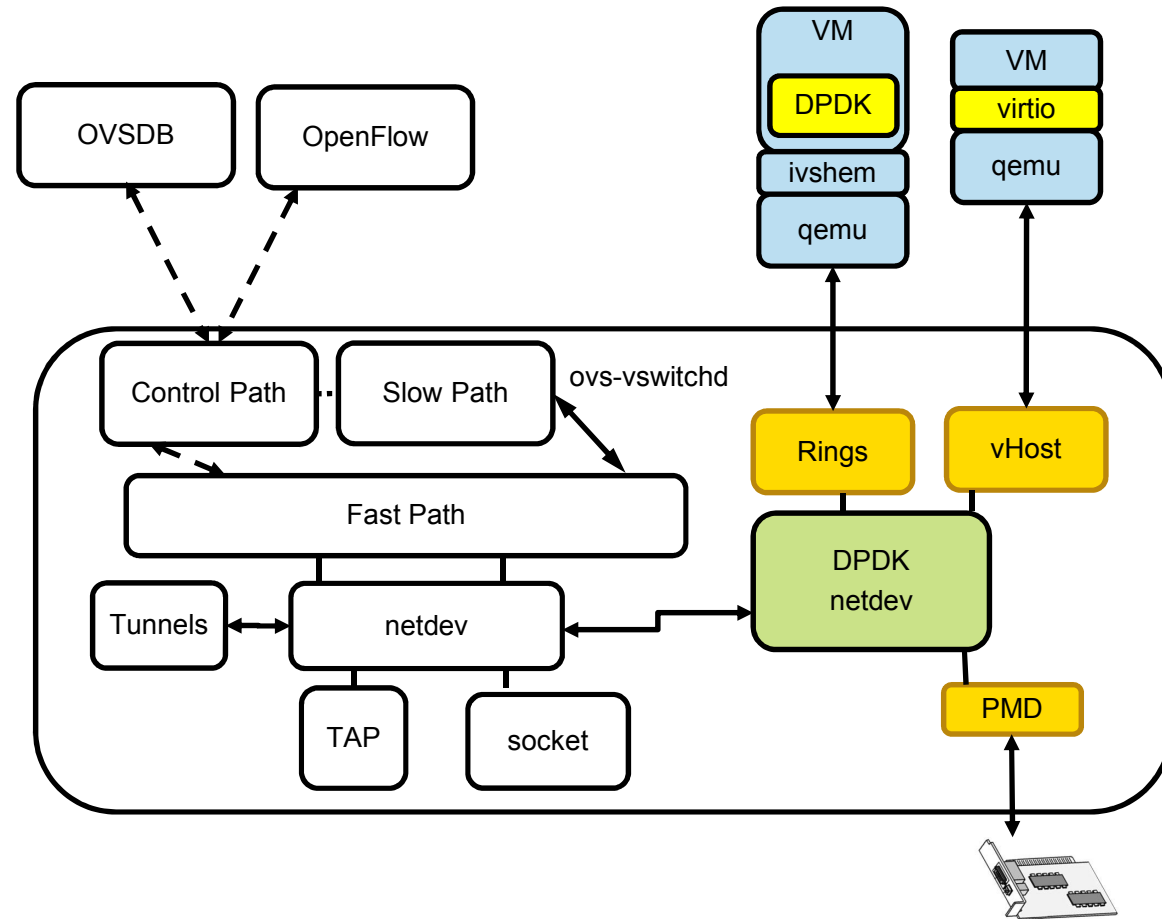
Results have been estimated or simulated using internal Intel analysis or architecture simulation or modeling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.

Background

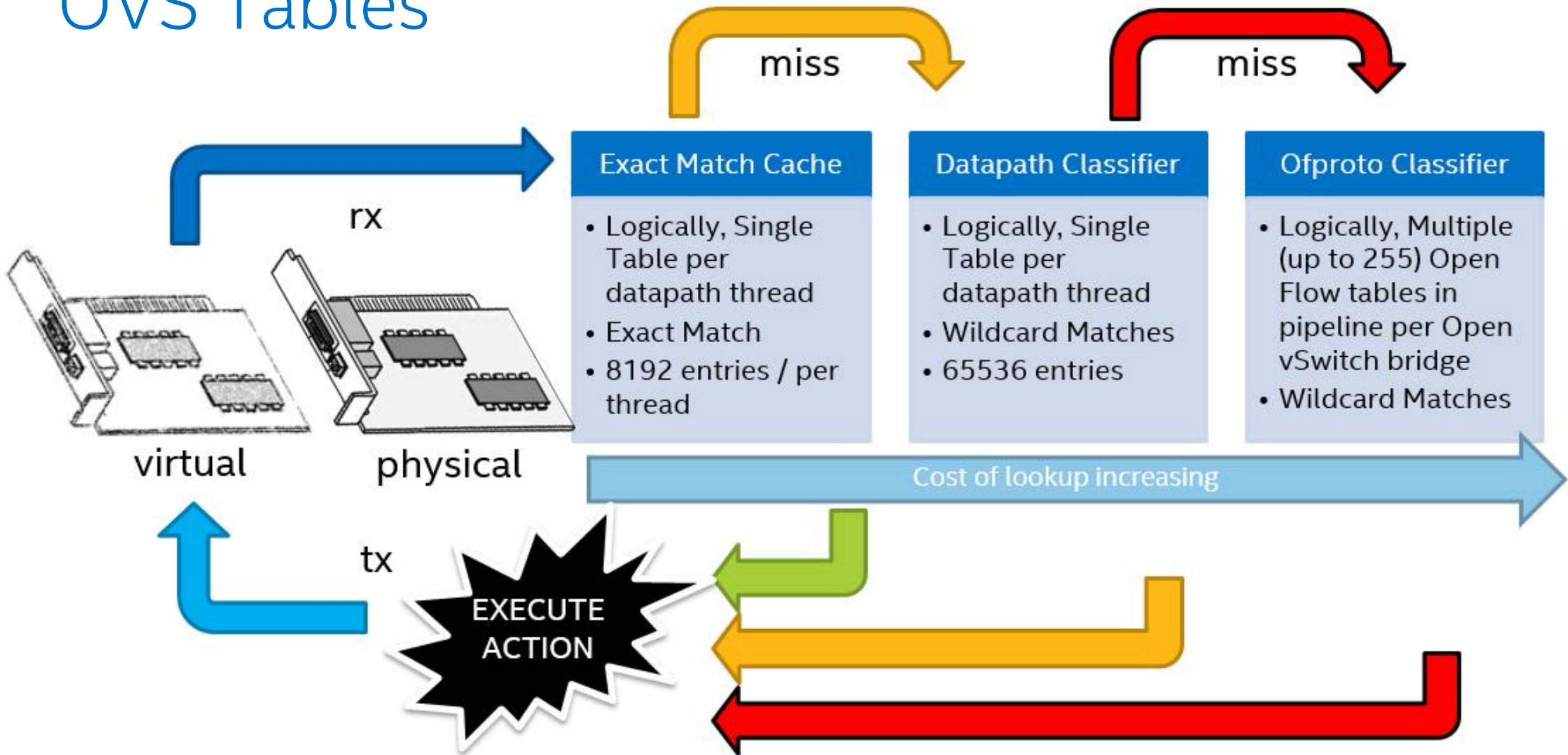


Architecture

OpenvSwitch with DPDK



OVS Tables



OVS Tables

rx cost

lookup cost

tx cost

action cost

Performance

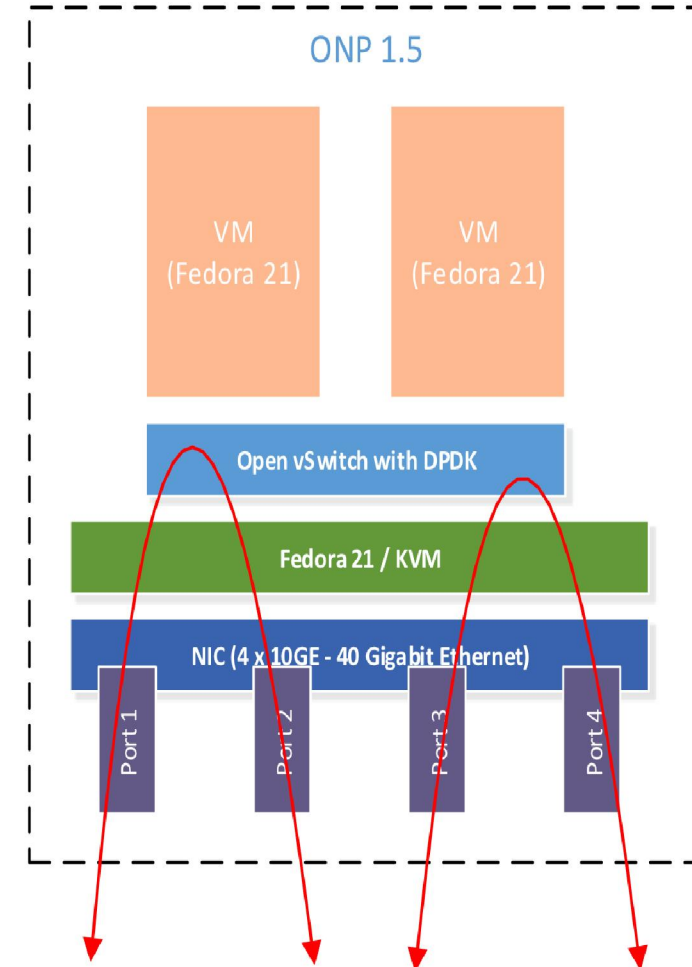
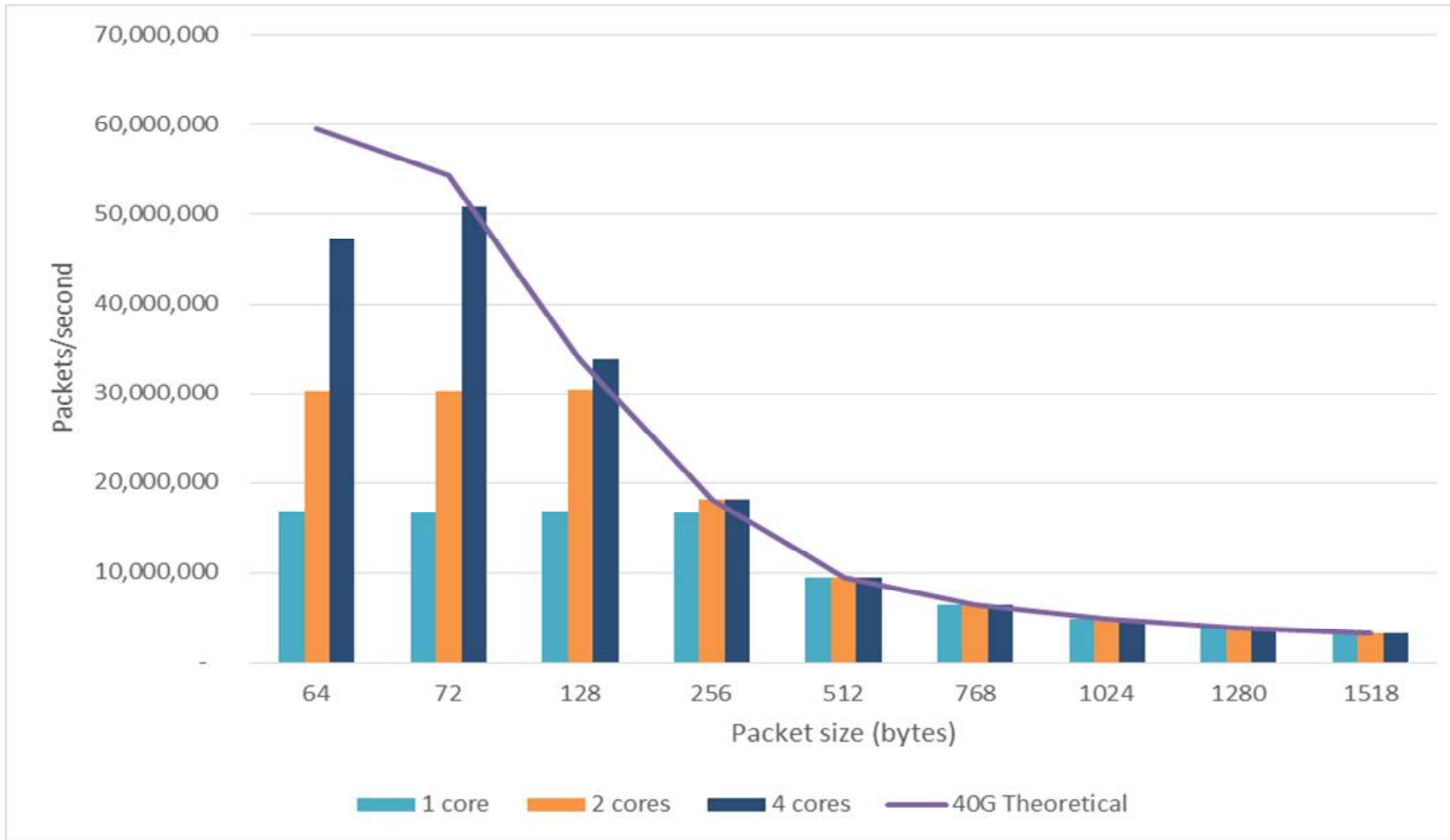
OpenvSwitch 2.4

Platform Performance Configuration

Item	Description
Server Platform	Intel® Server Board S2600WT2 DP (Formerly Wildcat Pass) 2 x 1GbE integrated LAN ports Two processors per platform
Chipset	Intel® C610 series chipset (Formerly Wellsburg)
Processor	Intel® Xeon® Processor E5-2697 v3 (Formerly Haswell) Speed and power: 2.60 GHz, 145 W Cache: 35 MB per processor Cores: 14 cores, 28 hyper-threaded cores per processor for 56 total hyper-threaded cores QPI: 9.6 GT/s Memory types: DDR4-1600/1866/2133, Reference: http://ark.intel.com/products/81059/Intel-Xeon-Processor-E5-2697-v3-35M-Cache-2_60-GHz
Memory	Micron 16 GB 1Rx4 PC4-2133MHz, 16 GB per channel, 8 Channels, 128 GB Total
Local Storage	500 GB HDD Seagate SATA Barracuda 7200.12 (SN:9VMKQZMT)
PCIe	Port 3a and Port 3c x8
NICs	2 x Intel® Ethernet CAN X710-DA2 Adapter (Total: 4 x 10GbE ports) (Formerly Fortville)
BIOS	Version: SE5C610.86B.01.01.0008.021120151325 Date: 02/11/2015

OpenvSwitch 2.4

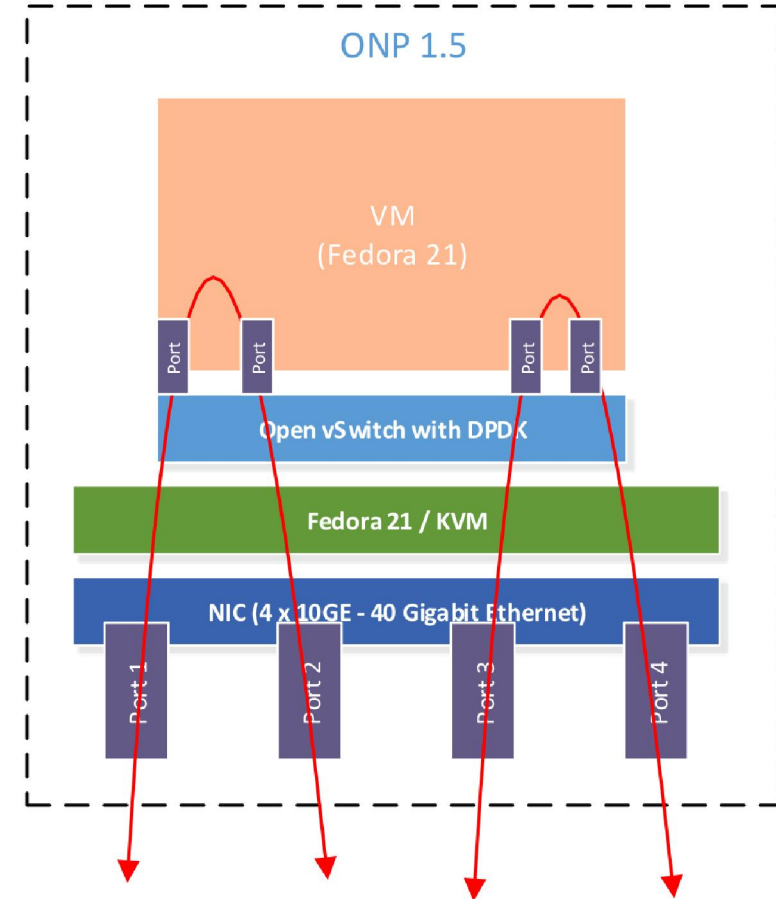
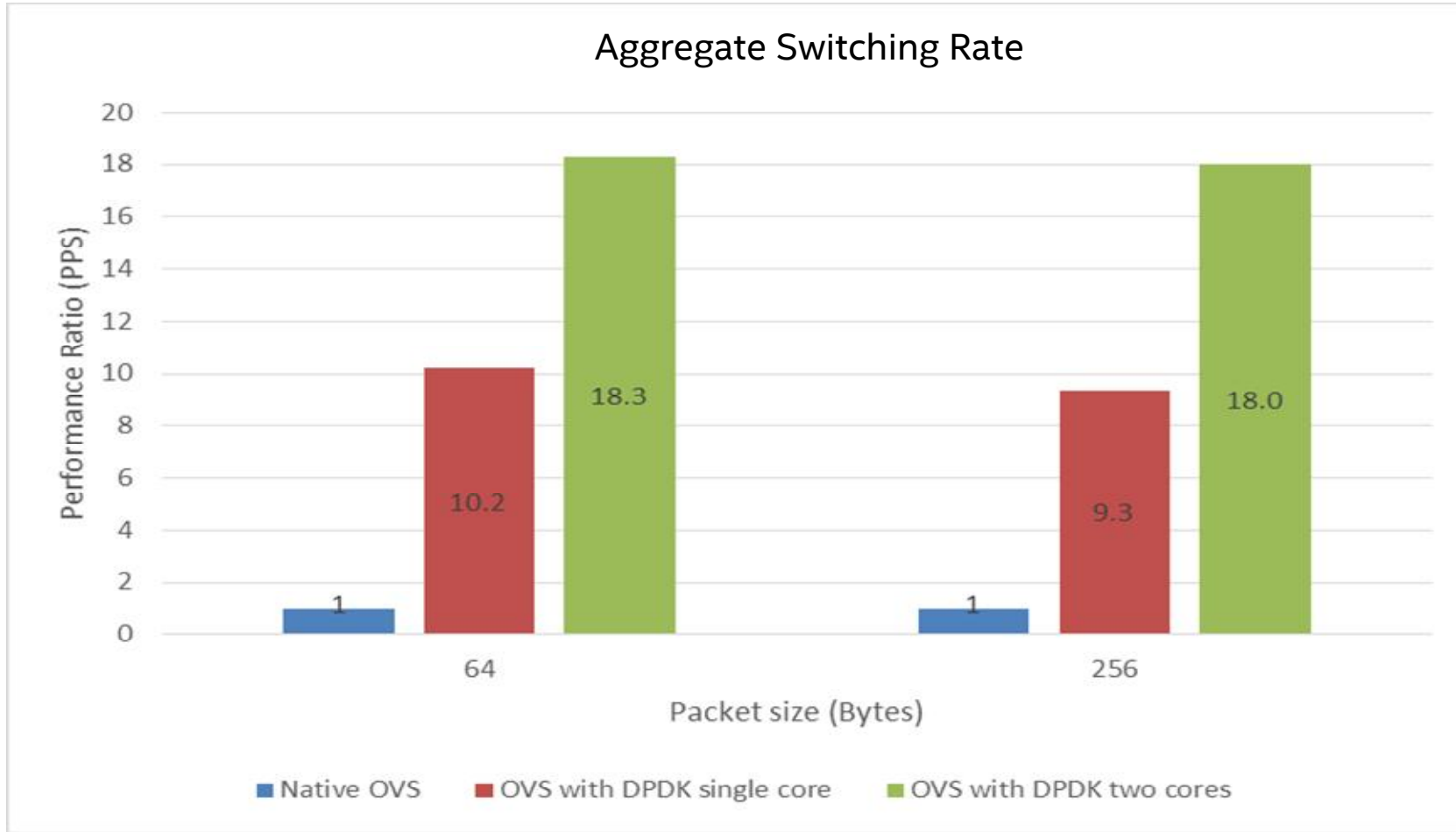
Phy-OVS-Phy Performance



Disclaimer: For more complete information about performance and benchmark results, visit www.intel.com/benchmarks and https://download.01.org/packet-processing/ONPS1.5/Intel_ONP_Server_Release_1.5_Performance_Test_Report_Rev1.2.pdf

OpenvSwitch 2.4

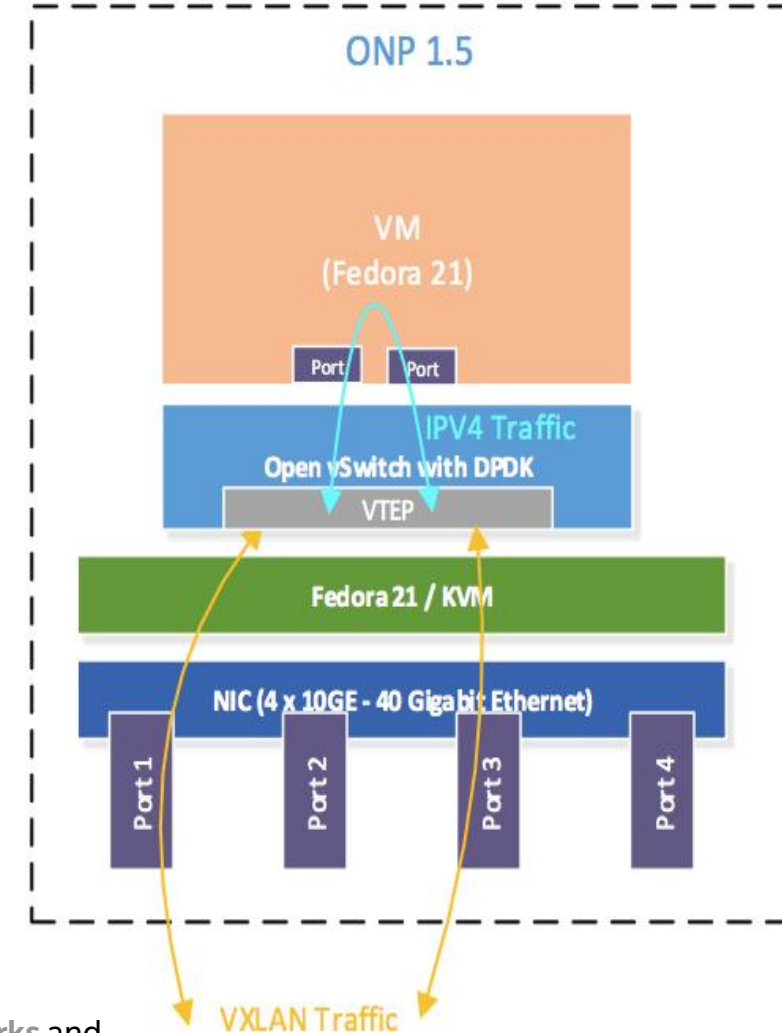
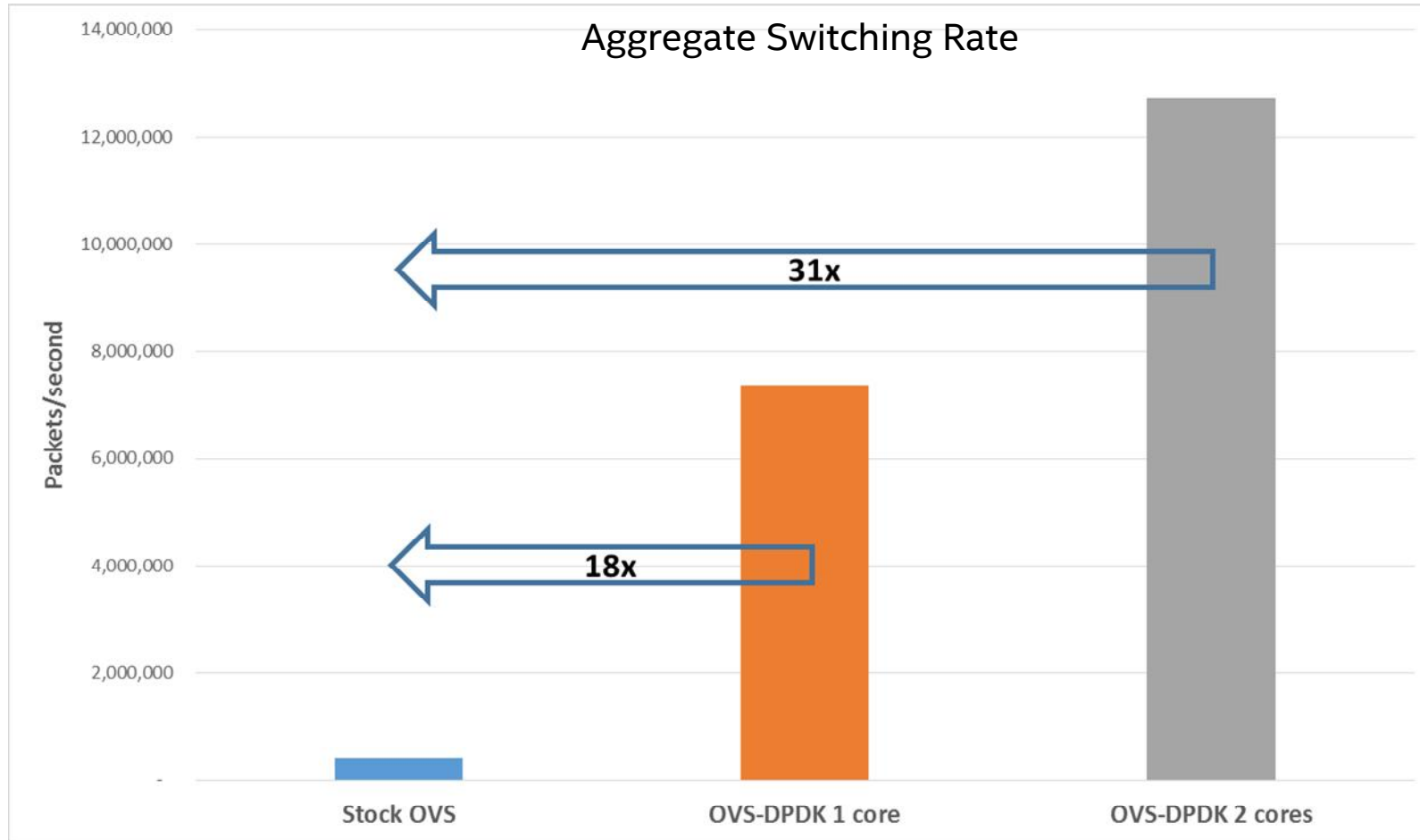
Phy-VM-Phy Performance



Disclaimer: For more complete information about performance and benchmark results, visit www.intel.com/benchmarks and https://download.01.org/packet-processing/ONPS1.5/Intel_ONP_Server_Release_1.5_Performance_Test_Report_Rev1.2.pdf

OpenvSwitch 2.4

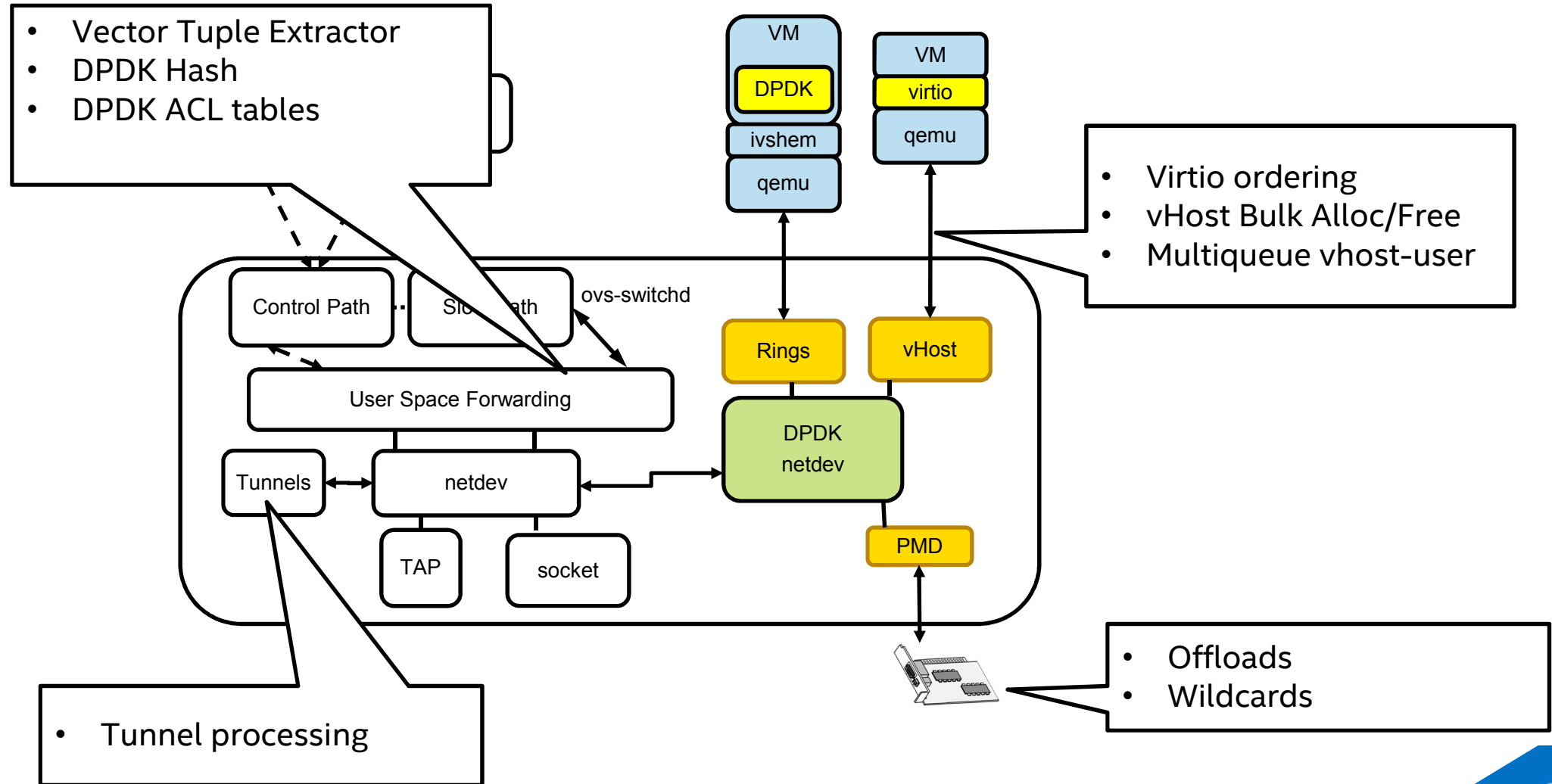
Phy-OVS Tunnel-Phy Performance



Disclaimer: For more complete information about performance and benchmark results, visit www.intel.com/benchmarks and https://download.01.org/packet-processing/ONPS1.5/Intel_ONP_Server_Release_1.5_Performance_Test_Report_Rev1.2.pdf

OpenvSwitch 2.x DPDK 2.x

netdev-DPDK Performance Enhancements



usable

adjective | us·able | \ˈyü-zə-bəl\


Simple Definition of USABLE


Popularity: Bottom 50% of words

: capable of being used : in good enough condition to be used

Full Definition of USABLE

- 1 : capable of being used
- 2 : convenient and practicable for use

—us·abil·i·ty  \ˈyü-zə-ˈbi-lə-tē\ *noun*

—us·able·ness  \ˈyü-zə-bəl-nəs\ *noun*

—us·ably  \-blē\ *adverb*

Usability



Usability Examples

Cmd Line args - `ovs-vswitchd --dpdk -c 0x40 -n 4 --socket-mem 1024,0`

- <http://openvswitch.org/pipermail/dev/2016-January/063959.html>

Testing - VSPERF – OPNFV project

- https://wiki.opnfv.org/characterize_vswitch_performance_for_telco_nfv_use_cases

DPDK Device Management - Driverctl

- <http://dpdk.org/ml/archives/dev/2015-December/029500.html>

DPDK Debug/tcpdump – Several ideas proposed

- <http://dpdk.org/ml/archives/dev/2015-November/029085.html>

Documentation

- <https://github.com/openvswitch/ovs/blob/master/INSTALL.DPDK.md>

Out of the box Performance

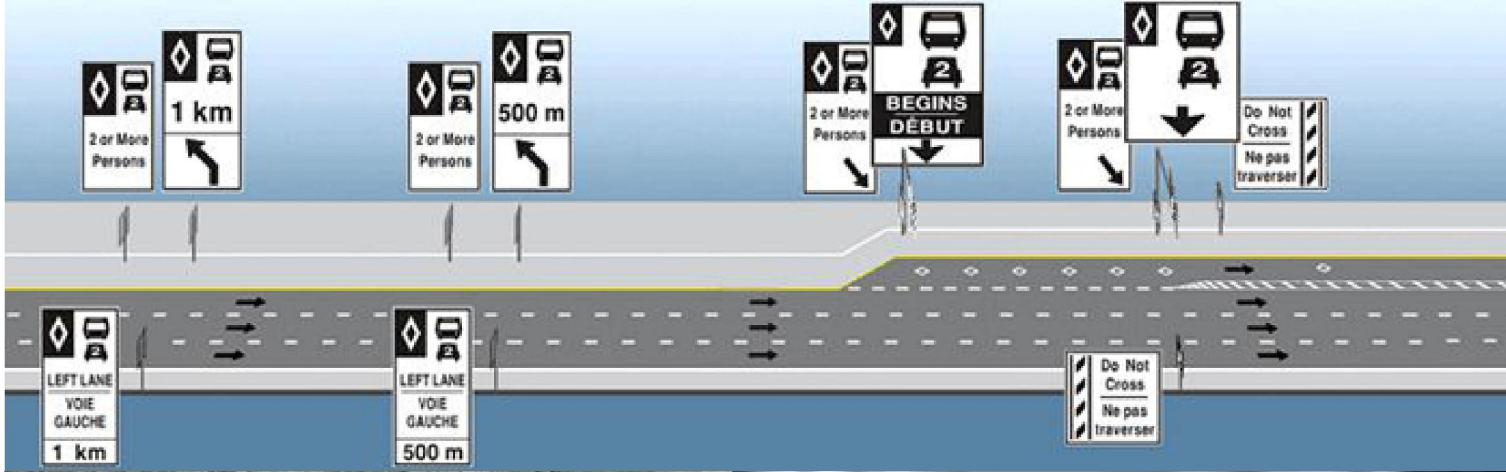
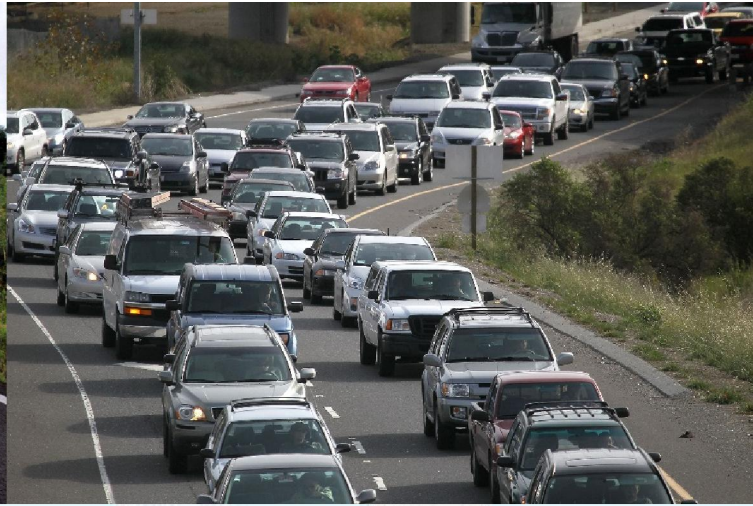
- `ovs-vsctl --no-wait set Open_vSwitch . other_config:pmd-cpu-mask=f`

Availability: Distro Packages and Git

- ISV and OSV recognizing the progress of OVS with DPDK
- Centos7: 7.4: DPDK 2.1; 7.2: OVS 2.4
- Fedora: F23;F22 updates DPDK 2.0; F24: DPDK 2.1
- Fedora Copr repo for latest:<https://copr.fedoraproject.org/coprs/pmatilai/dpdk/>
- Red Hat OSP8:
 - OVS 2.4/DPDK 2.0 Integrated with Neutron
- Ubuntu: 15.10: OVS with DPDK package
- OVSNFV OPNFV Project planning deployment in future OPNFV releases
- <https://01.org/packet-processing/intel%C2%AE-onp-servers>
- `git clone` <http://dpdk.org/git/dpdk>
- `git clone` <https://github.com/openvswitch/ovs.git>

Get yours now!





Wrap-up

- Feel free to join us in the OVS-DPDK lane...
 - Performance
 - Usability
 - Testing
 - <http://openvswitch.org/mlists/>
 - <http://dpdk.org/ml>
- But not like this...



I don't always violate the HOV lane law...but when I do, I get a \$124 ticket! We'll give him an A for creativity! 🚔👍



Backup

OVS-DPDK Setup #1

```
# Build DPDK
```

```
export RTE_SDK=/home/ktraynor/vswitch/ovs/code/dpdk_210
```

```
cd /home/ktraynor/vswitch/ovs/code/dpdk_210
```

```
make install T=x86_64-native-linuxapp-gcc CONFIG_RTE_BUILD_COMBINE_LIBS=y
```

```
# Build OVS
```

```
./boot.sh
```

```
./configure --with-dpdk=/home/ktraynor/vswitch/ovs/code/dpdk_210/x86_64-native-linuxapp-gcc
```

```
make 'CFLAGS=-g -Ofast -march=native' ; make install
```

```
# Mount Hugepages
```

```
mkdir -p /mnt/huge
```

```
mount -t hugetlbfs nodev /mnt/huge
```

```
echo 64 > /sys/devices/system/node/node0/hugepages/hugepages-2048kB/nr_hugepages
```

```
# Bind ports to DPDK
```

```
modprobe uio
```

```
insmod $dpdk_dir/x86_64-native-linuxapp-gcc/kmod/igb_uio.ko
```

```
$dpdk_dir/tools/dpdk_nic_bind.py -b igb_uio 05:00.0 05:00.1
```

OVS-DPDK Setup #2

```
# Run vswitchd
```

```
ovs-vswitchd --dpdk -c 0x40 -n 4 --socket-mem 1024,0 -- unix:/usr/local/var/run/openvswitch/db.sock
```

```
# Set Forwarding cores
```

```
ovs-vsctl --no-wait set Open_vSwitch . other_config:pmd-cpu-mask=f
```

```
# Add Bridge, Ports and Rule
```

```
ovs-vsctl add-br br0 -- set bridge br0 datapath_type=netdev
```

```
ovs-vsctl add-port br0 dpdk0 -- set Interface dpdk0 type=dpdk
```

```
ovs-vsctl add-port br0 dpdk1 -- set Interface dpdk1 type=dpdk
```

```
ovs-ofctl add-flow br0 in_port=1,action=output:2
```

```
# Debug
```

```
ovs-ofctl add-flow br0 in_port=1,action=output:2, LOCAL
```

```
tcpdump -i <ip of LOCAL port>
```

OVS Setup

```
# Build OVS
./boot.sh
./configure
make ; make install
# Add OVS kernel module
modprobe openvswitch
# Run vswitchd
ovs-vswitchd unix:/usr/local/var/run/openvswitch/db.sock
# Add Bridge, Ports and Rule
ovs-vsctl add-br br0
ovs-vsctl add-port br0 p3p1
ovs-vsctl add-port br0 p3p2
ovs-ofctl add-flow br0 in_port=1,action=output:2
# Debug
tcpdump -i p3p2
```

Paths to the guest

