



OPNFV SUMMIT

JUNE 2017 | BEIJING, CHINA

Dataplane Performance, Capacity, and Benchmarking in OPNFV

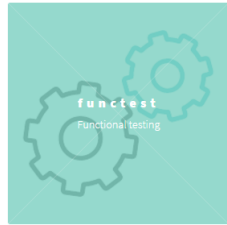
Trevor Cooper, Intel Corp.

Sridhar Rao, Spirent Communications

Al Morton, AT&T Labs

... with acknowledgement to VSPERF committers

Danube reporting



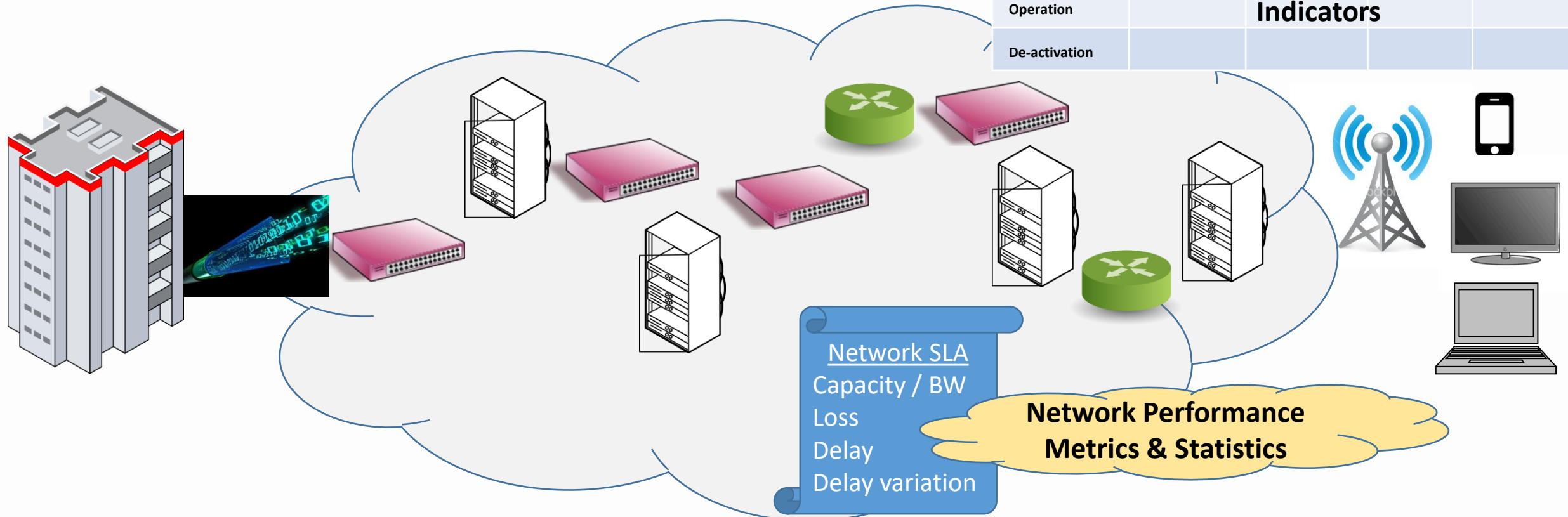
Agenda

1. Dataplane Performance Measurement with VSPERF
2. VSPERF Example Results and Analysis
3. Moving Ahead with VSPERF

E2E Dataplane Performance Measurement & Analysis ...

User-Application Quality-of-Service

COVERAGE	SPEED	ACCURACY	RELIABILITY	SCALABILITY
Activation	Service Performance Indicators			
Operation				
De-activation				

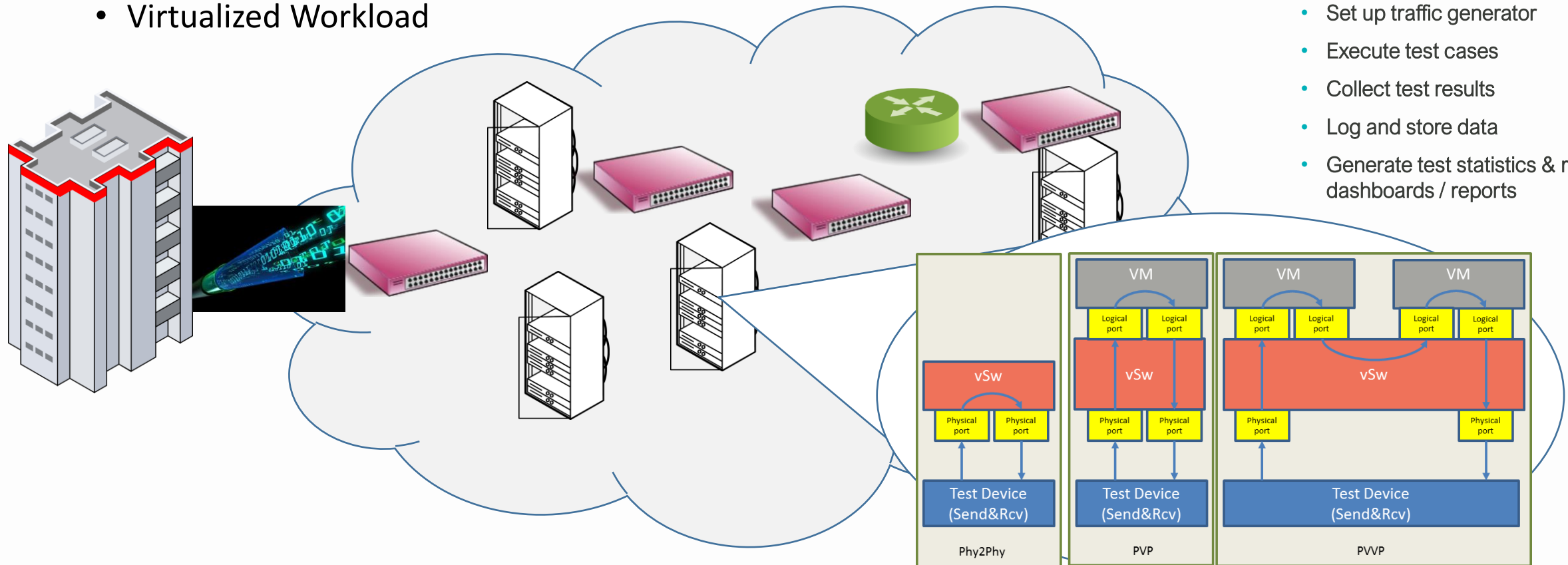


VSPERF DUT is an important part of the E2E Data Path

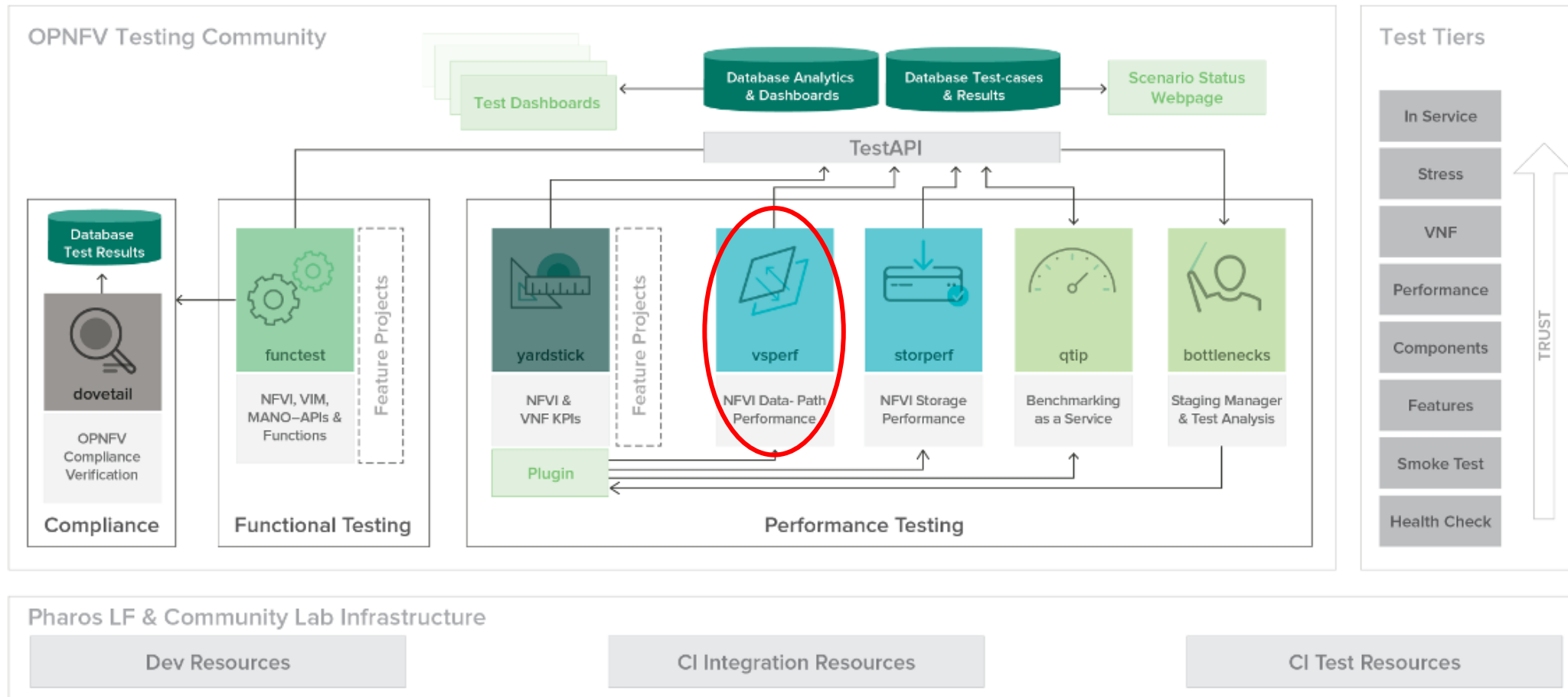
- Virtual Switching technology and NIC offloads
- Physical and virtual ports
- Virtualized Workload

VSPERF Test Automation

- Source/build SW components
- Set up vSwitch
- Set up workload
- Set up traffic generator
- Execute test cases
- Collect test results
- Log and store data
- Generate test statistics & result dashboards / reports



VSPERF and the OPNFV Testing Community















Dataplane Performance Testing Options

Solution Stack

Workload (DUT)	Traffic Generator (HW or SW)	Automation code (Test framework)
Sample VNFs <ul style="list-style-type: none"> SampleVNF (vACL, vFW, vCGNAT, ...) Open source VNF catalogue 	Hardware - commercial <ul style="list-style-type: none"> Ixia Spirent Xena 	Compliance <ul style="list-style-type: none"> Dovetail
Test VMs <ul style="list-style-type: none"> vloop-vnf (dpdk-testpmd, Linux bridge, L2fwd module) Spirent stress-VM Virtual Traffic classifier 	Virtual - commercial <ul style="list-style-type: none"> Ixia Spirent 	VIM and MANO <ul style="list-style-type: none"> NFVbench
Virtual switching <ul style="list-style-type: none"> OVS OVS-dpdk VPP 	Software - Open Source <ul style="list-style-type: none"> Pktgen Moongen TREX PROX 	VIM, no MANO <ul style="list-style-type: none"> Yardstick Qtip Bottleneck
Physical / virtual interfaces <ul style="list-style-type: none"> NIC (10GE, 40GE, ...) Vhost-user Pass-through, SR-IOV 		No VIM or MANO <ul style="list-style-type: none"> VSPERF Storperf
HW offload <ul style="list-style-type: none"> TSO encrypt/decrypt SmartNIC 		*Used in test examples presented

Daily tests on master and stable branch in OPNFV Lab
<https://build.opnfv.org/ci/view/vswitchperf/>

		vswitchperf-daily-danube
		vswitchperf-daily-master
		vswitchperf-merge-danube
		vswitchperf-merge-master
		vswitchperf-verify-danube
		vswitchperf-verify-master

Specifications

- IETF BMWG RFCs for Dataplane Performance
- ETSI NFV Test Specifications

Topologies

- vSwitch
- SR-IOV etc.
- Phy2Phy
- PVP
- PVVP (multi-VM)

```

/tmp/results_2016-01-19_17-04-26/result_phy2phy_tput_p2p.md
-----
- Test ID: phy2phy_tput
- Description: LTD.Throughput.RFC2544.PacketLossRatio
- Deployment: p2p
- Traffic type: rfc2544
- Bidirectional : True

-----
Metric                                     Result
-----
throughput_rx_fps                           22626032.812
throughput_rx_mbps                           11584.529
tx_rate_percent                              76.023
throughput_rx_percent                         76.023
min_latency_ns                               5260.000
max_latency_ns                               138000.000
avg_latency_ns                               7392.000
type                                          rfc2544
packet_size                                  64
traffic_type                                 udp
  
```

VSPERF Example Results and Analysis

Results and Analysis from Recent Tests

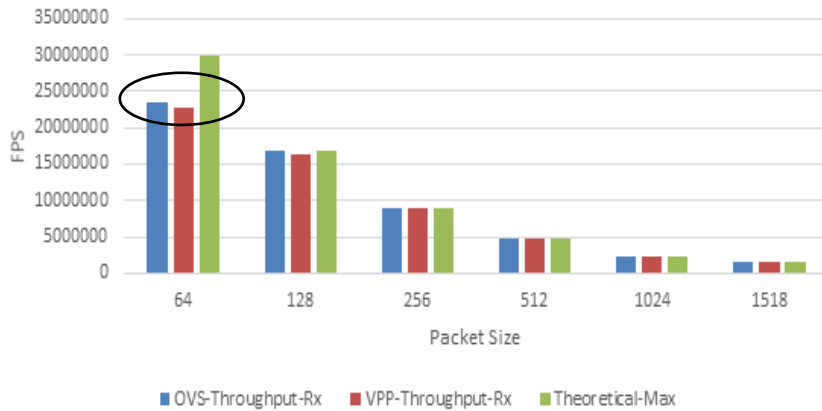
Using VSPERF to Analyse:

1. OVS and VPP
2. Traffic Generators
3. Impact of noisy neighbor
4. Back2Back frame testing with CI

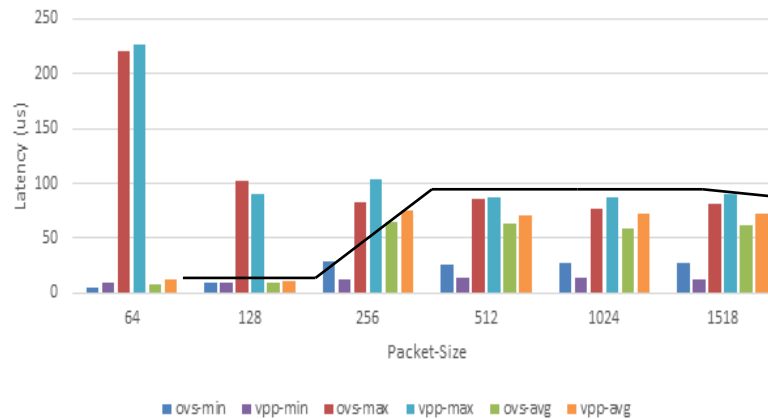
Virtual Switches in VSPERF: OVS and VPP

RFC2544, Phy2Phy OVS2.6.90,
VPP 17.01DPDK: 16.07.0

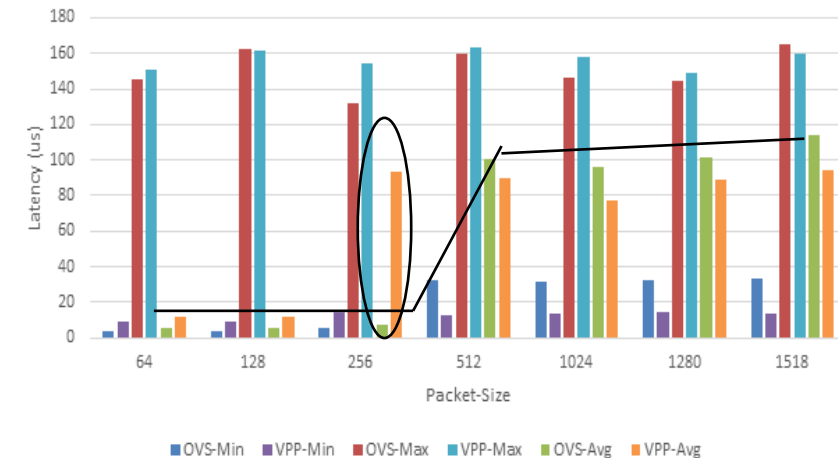
OVS and VPP-RFC2544: FPS vs Packet Size - Single Flow, Bidirectional



OVS and VPP- RFC2544: Latency vs Packet-Size - Single Flow - Bidirectional



OVS and VPP RFC2544-4096Flows- Latency Vs Packet-size



Both OVS and VPP (64 B, 1-Flow, bi-dir), the throughput is ~80% of line-rate
NIC has known processing limits that could be the bottleneck
For uni-directional traffic line-rate is achieved for 64B

Avg. latency for OVS and VPP varies from 10-90us with minimal (1-9%) difference between them
Average latency jumps significantly after 128 B

For multi-stream, latency variation are:

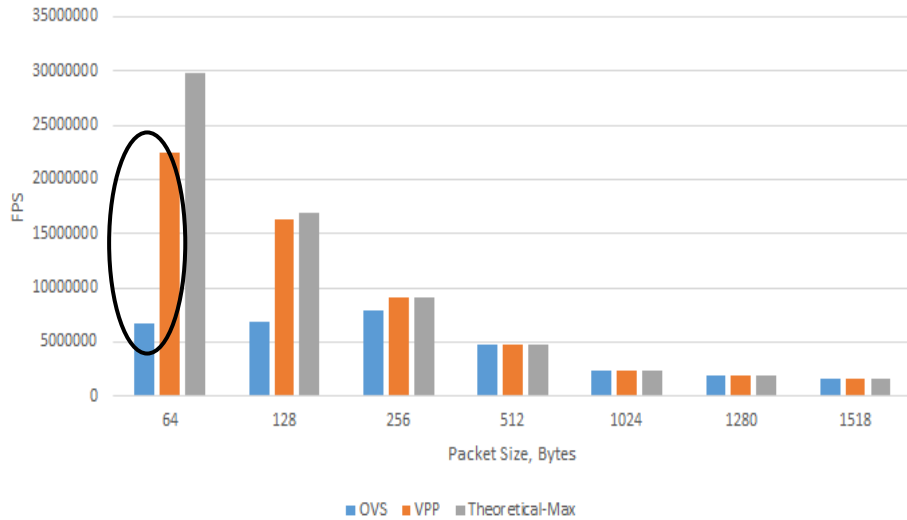
- Min: 2-30us
- Avg: 5-110us

Inconsistency for 256B with OVS vs VPP

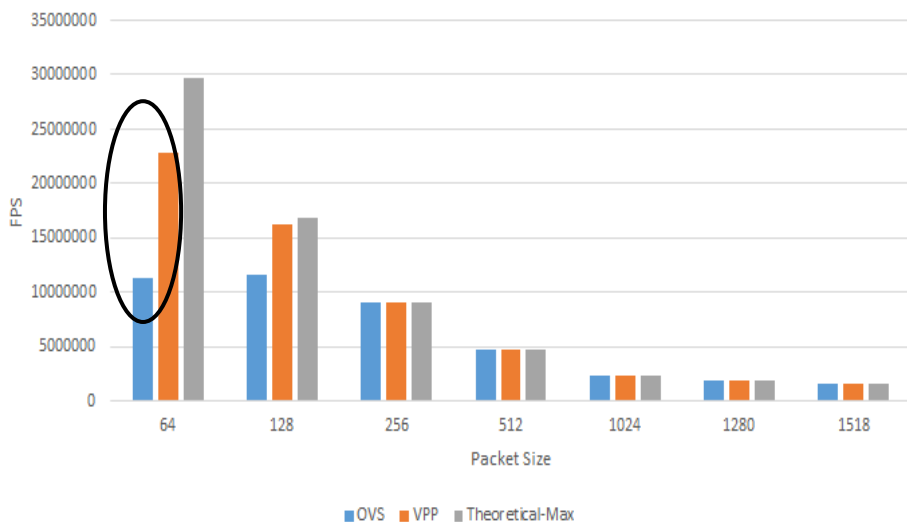
A jump in latency for higher packet-sizes, is seen in almost all cases



OVS and VPP-RFC2544-4096Flows:FPS vs Packet-Size



OVS and VPP-RFC2544-1Mflows: FPP vs Packet Size



ERF: OVS and VPP

For multi-stream, 64 and 128B – VPP throughput can go up to 70% higher than OVS. But ...

Inconsistencies

- OVS: 4K flows lower TPUT vs 1M
- Traffic generator results differ

Possible Reasons

- Packet-handling architectures
- Packet construction variation
- Test traffic is fixed size

RFC2544, Phy2Phy
OVS2.6.90, VPP 17.01
DPDK: 16.07.0

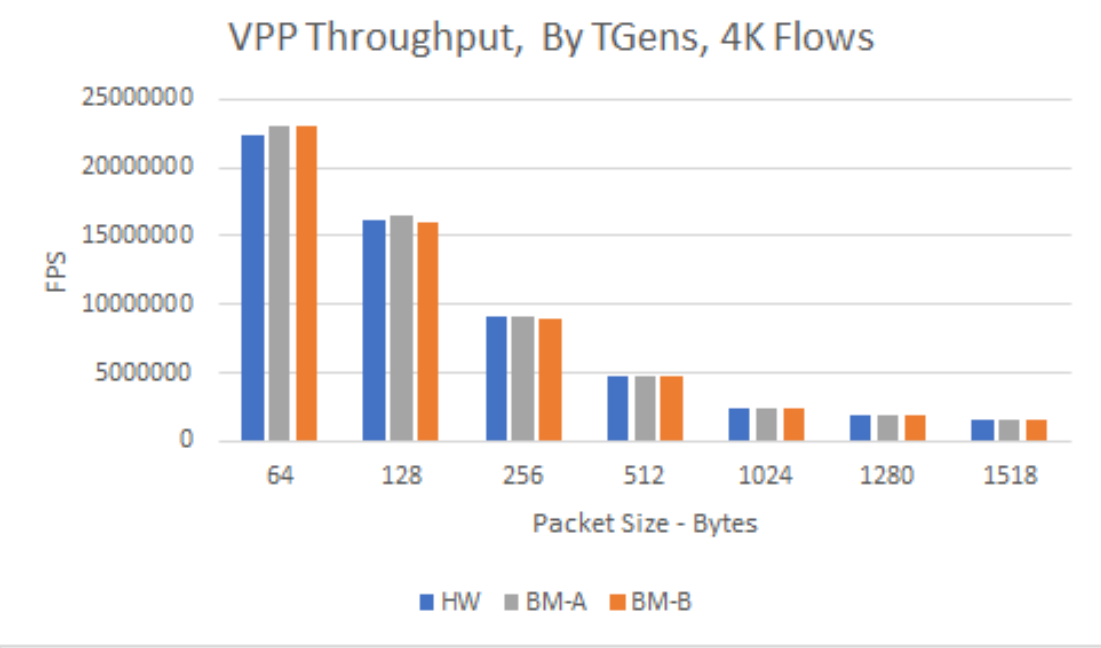
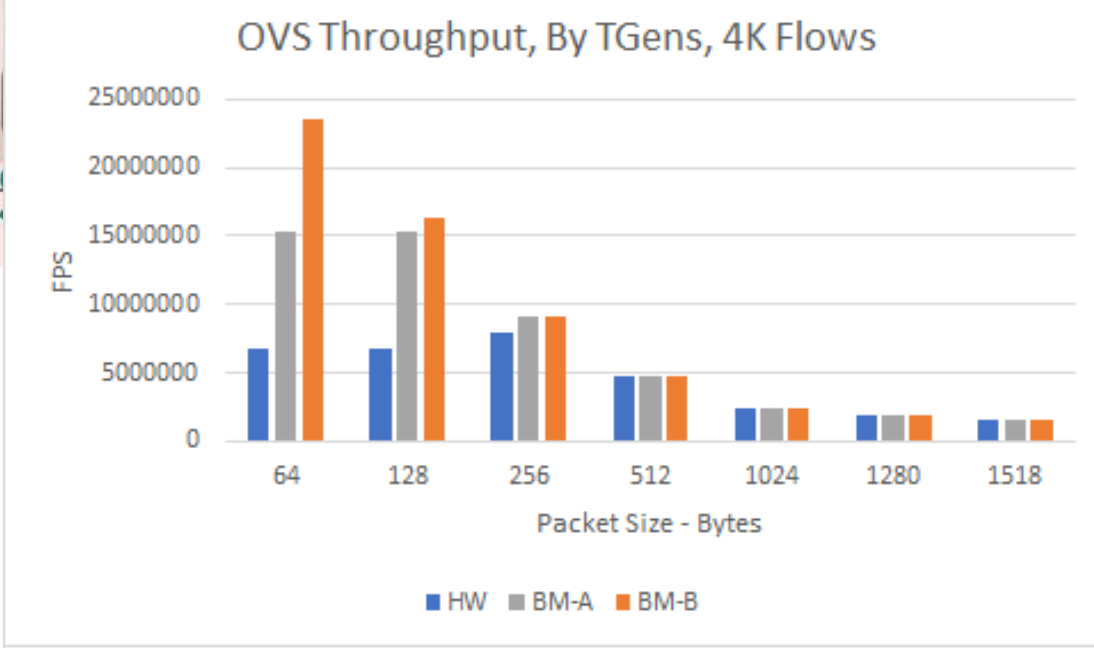
Analysis of Cache Miss: [SNAP monitoring tool]

The cache miss of VPP is 6% lower compared to the cache-misses for OVS.

Requires further analysis!

Lessons Learned – OVS and VPP

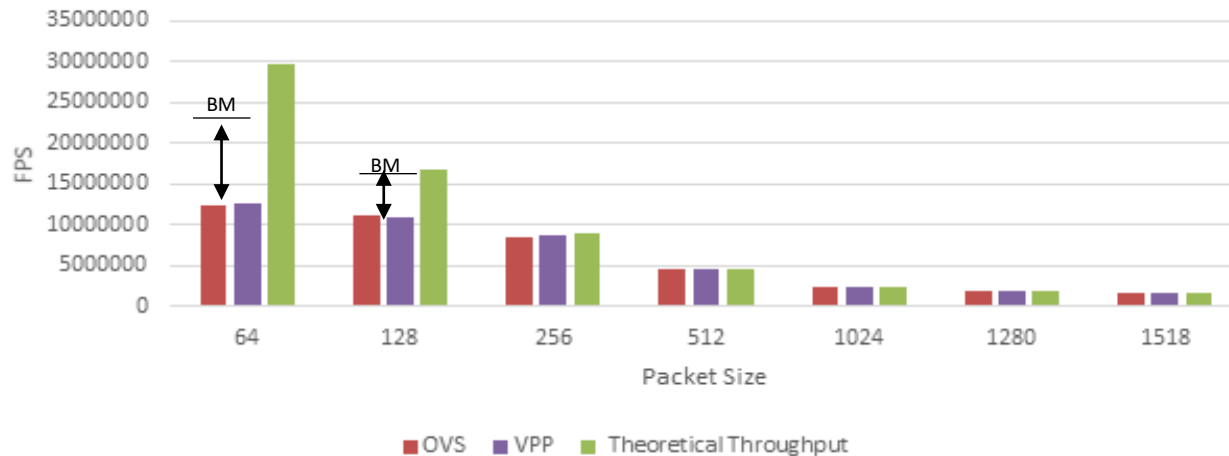
- Simple performance test-cases (#flows + pps) may not provide meaningful comparisons
 - [EANTC Validates Cisco's Network Functions Virtualization \(NFV\) Infrastructure](#) - (Oct 2015)
 - *Test case topology is VM to VM ... 0.001% packet loss accepted ... Pass-through connects physical interfaces to VNF ... VPP and OVS use a “single core” ... Software versions – OVS-dpdk 2.4.0, DPDK 2.0, QEMU 2.2.1 ... Processor E5-2698 v3 (Haswell – 16 physical cores), NW adaptor X520-DA2*
- Results are use-case dependent
 - Topology and encapsulation impact workloads under-the-hood
 - Realistic and more complex tests (beyond L2) may impact results significantly
 - Measurement methods (searching for max) may impact results
 - DUT always has multiple configuration dimensions
 - Hardware and/or software components can limit performance (but this may not be obvious)
 - Metrics / statistics can be deceiving – without proper considerations to above points!



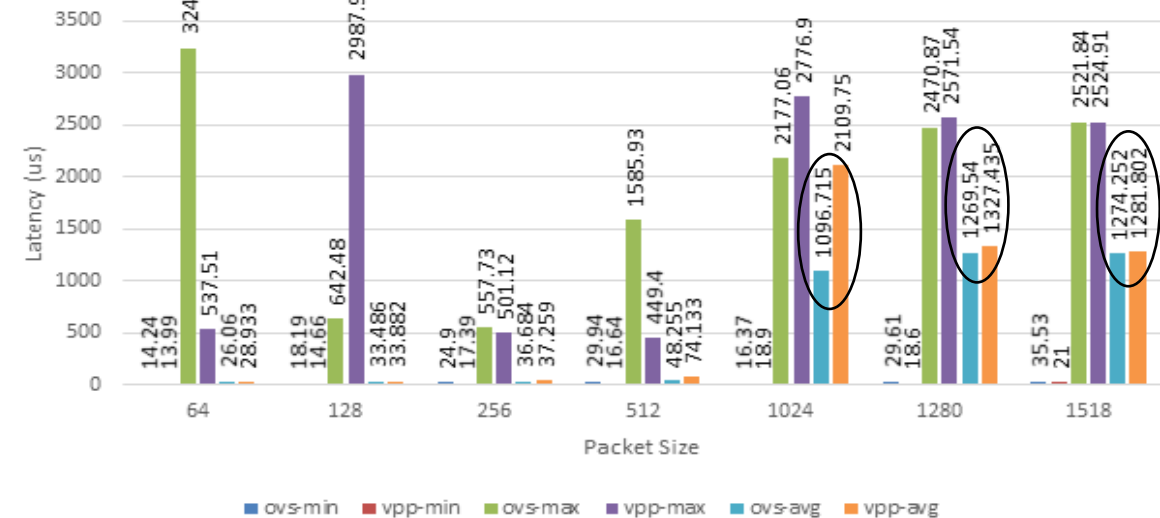
- Software Traffic Generators on bare-metal are comparable to HW reference for larger pkt sizes
- Small pkt sizes show inconsistent results
 - Across different generators
 - Between VPP and OVS
 - For both single and multi-stream scenarios
- For now, **in VSPERF**, existing baremetal software trafficgens, are unable to provide latency values*

*Running vsperf in "trafficgen-off" mode, it is possible to obtain latency values for some SW TGen.

OVS and VPP-RFC2544: Fps vs Packet-Size - Single Flow, Bidirectional



OVS and VPP RFC2544 - Latency Vs Packet-Size - Single Flow



With TGen-as-a-VM, the throughput is lower (upto 40%) in comparison with baremetal traffic generator.

Mostly restricted to lower packet size.

Reasons

Inherent baremetal vs VM differences.

Resource allocations.

Processes per packet.

In VSPERF, TGen-a-VM, can provide latency values.

* The latency values (min and avg) can be 10x times the values provided by the hardware traffic-generator *

[Configuration of NTP servers]

Software Traffic Generators – Lessons Learned

TG characteristics can impact measurements

- Inconsistent results seen for small packet sizes across TGs
- Packet stream characteristics may impact results ... bursty traffic is more realistic!
- Back2Back tests confirm sensitivity of DUT at small frame sizes
- Switching technology (DUT) are not equally sensitive to packet stream characteristics

Configuration of 'environment' for Software traffic-generators is critical

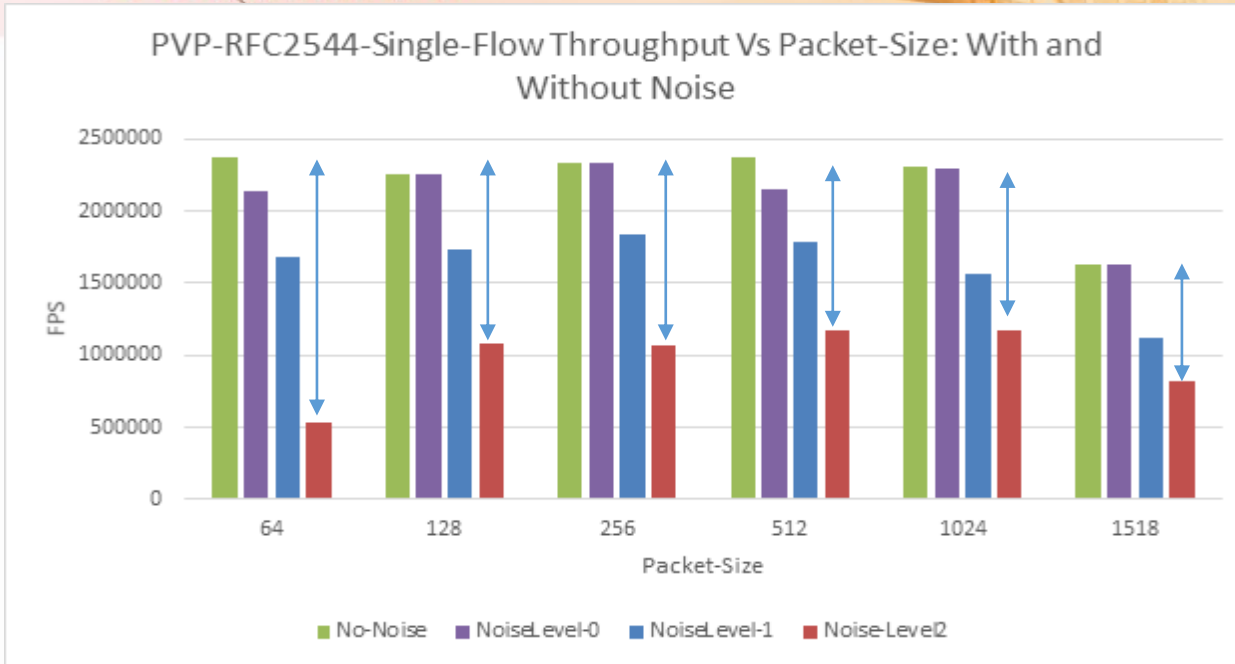
CPU: Count and
affinity definition

Memory: RAM,
Hugepages and
NUMA Configuration

DPDK Interfaces:
Tx/Rx Queues

PCI Passthrough or
SRIOV configurations

Software version



DUT: VSPERF with OVS and L2FWD VNF.

Traffic Generator: Hardware.

Noisy Neighbor: Stressor VM

Test: RFC2544 Throughput

Level	Last level cache consumption by the noisy neighbor VM
0	Minimal L3 cache consumption (<10%)
1	Average L3 cache consumption (50%)
2	High L3 cache consumption (100%)

CPU affinity configuration and NUMA configuration can protect from majority of Noise.

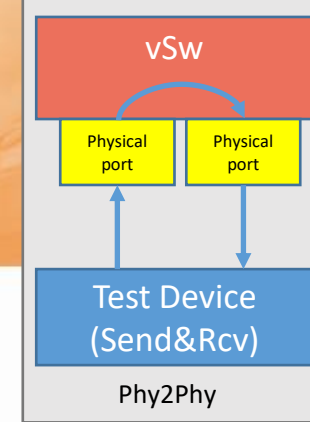
Consumption of Last-level cache (L3) is key to creating noise*

If the noisy neighbor can thrash the L3-Cache, it can lower the forwarding performance – throughput – upto 80%

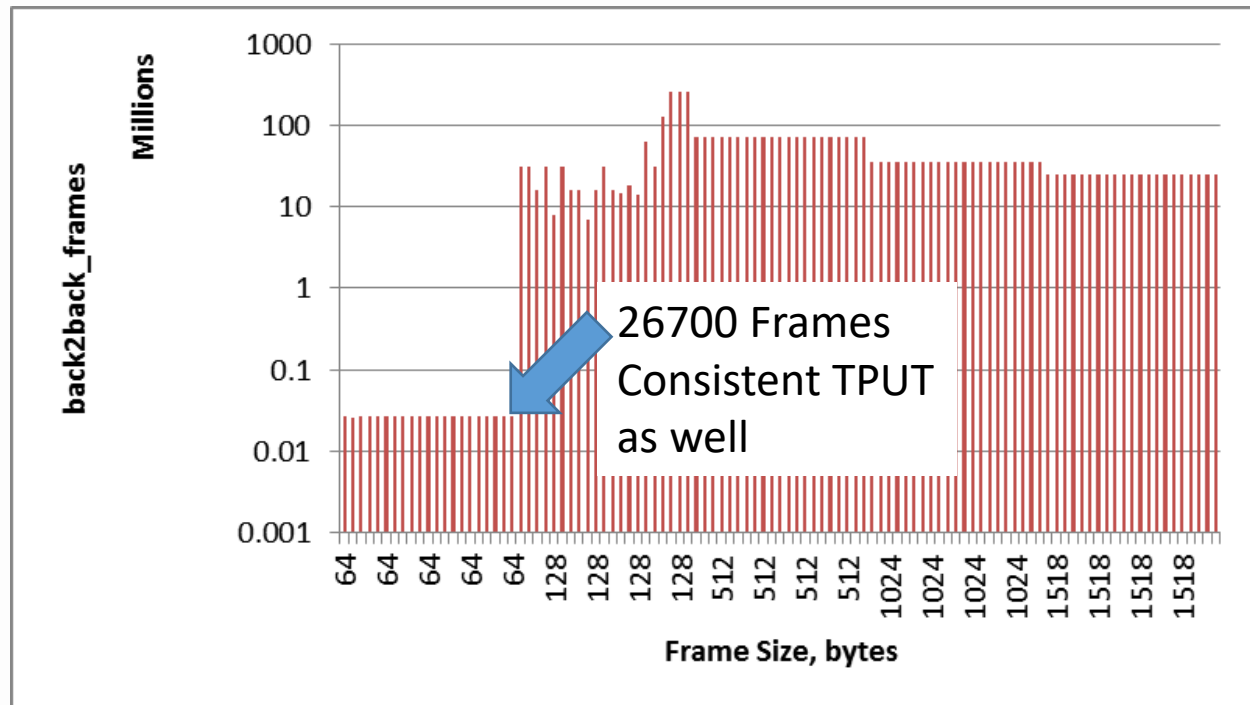
*It maybe be worth studying the use of tools such as cache-allocation-technology (Libpqos) to manage noisy-neighbors as shown here:

<https://www.openstack.org/assets/presentation-media/Collectd-and-Vitrage-integration-an-eventful-presentation2.final.pdf>

Back2Back Frame Testing Analysis

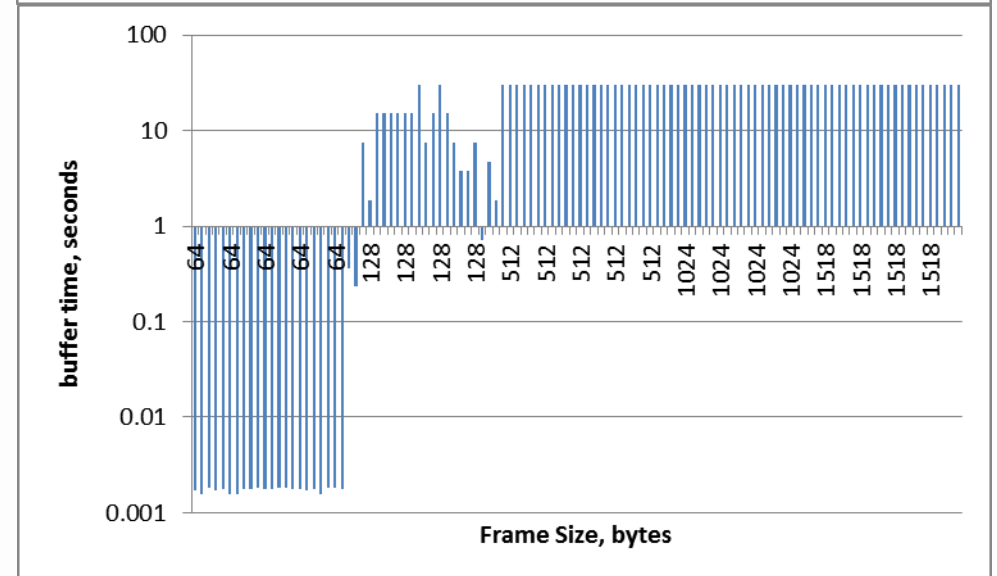
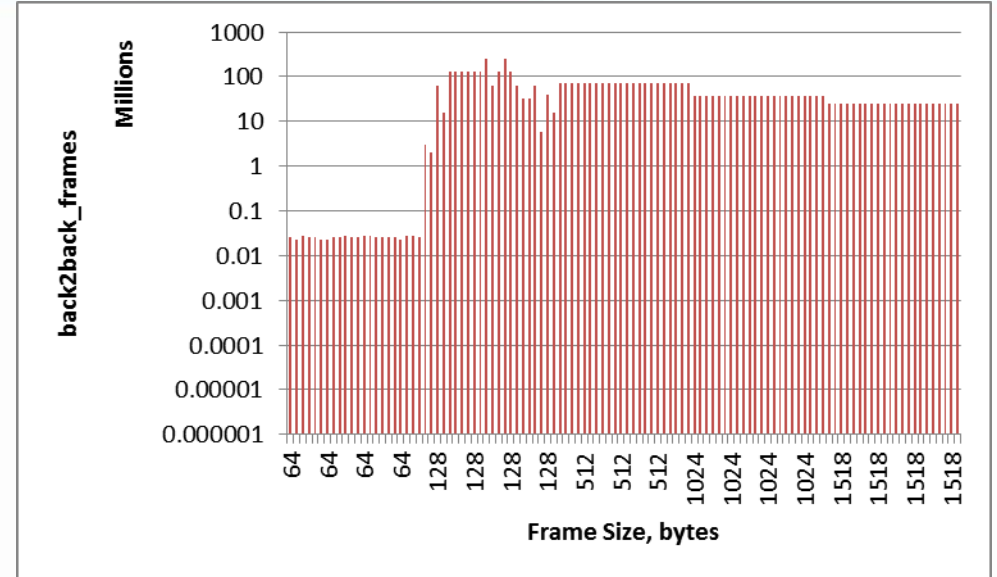
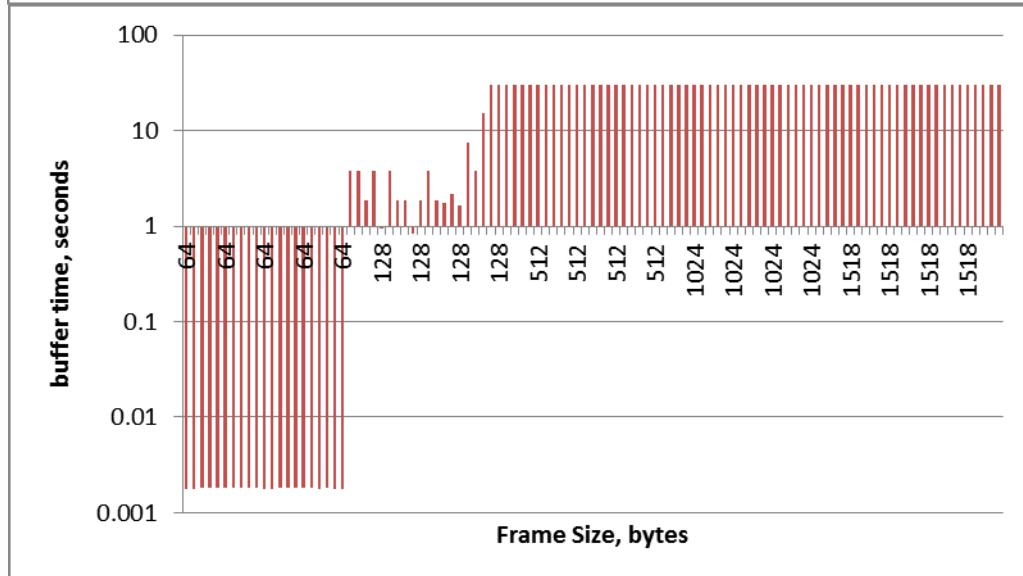
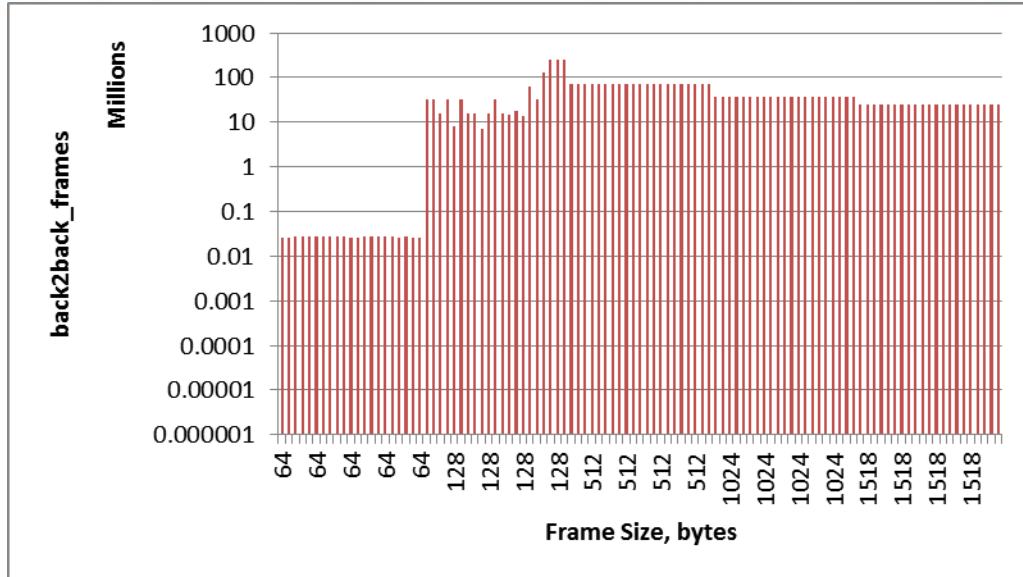


- Seek Maximum burst length (sent with min. spacing, or back-to-back) that can be transmitted through the DUT without loss (est.Buffer size)
- HW Tgen, Phy2Phy, OVS, CI tests on Intel Pod 12, Feb-May 2017



- Model: Tgen->Buff->HeaderProc->Rcv
- Only 64byte Frames are buffered!
- Ave Burst length = 26,700 Frames
- Source of Error: many Frames are processed before buffer overflow
- Corr_Buff=5713 frames, or 0.384ms
- Similar results for Intel Pod 3

• Pod 12



• Pod 3

Moving Ahead with VSPERF

STUDIES

Comparing virtual switching technologies and NFVI setups

More realistic traffic profiles
More complex topologies (e.g. full mesh)
Additional real-world use-cases (e.g. overlays)

Custom VNFs (dpdk workloads)
Stress tests (e.g. noisy neighbor)
Additional test cases (e.g. TCP)

FEATURES

Visualization and Interpretation of test results

New NFVI test specs & metrics (IETF, ETSI NFV)
Display of latency measurements
Test environment and DUT configurations
Traffic generator capabilities

Dashboards and analytics
Correlation of statistics
Simplification of results

INTEGRATION

Tool support and integration with other OPNFV frameworks

Metrics agents & monitoring systems
Additional traffic generators (e.g. 40GE)
CI unit tests for developers

OPNFV scenario support
Installer integration
Yardstick integration