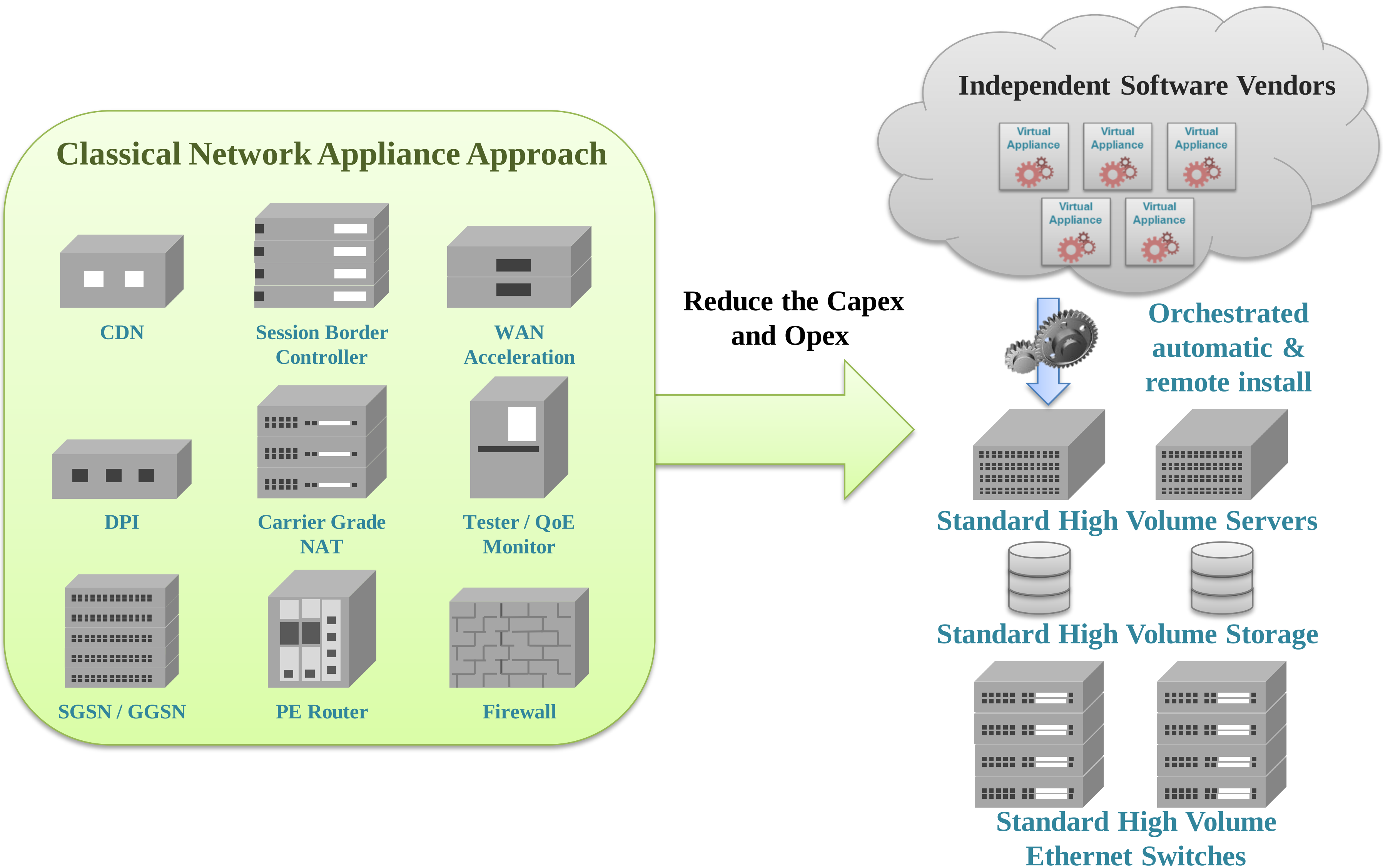
OPNFV & NFV Performance Lab

Industrial Technology Research Institute

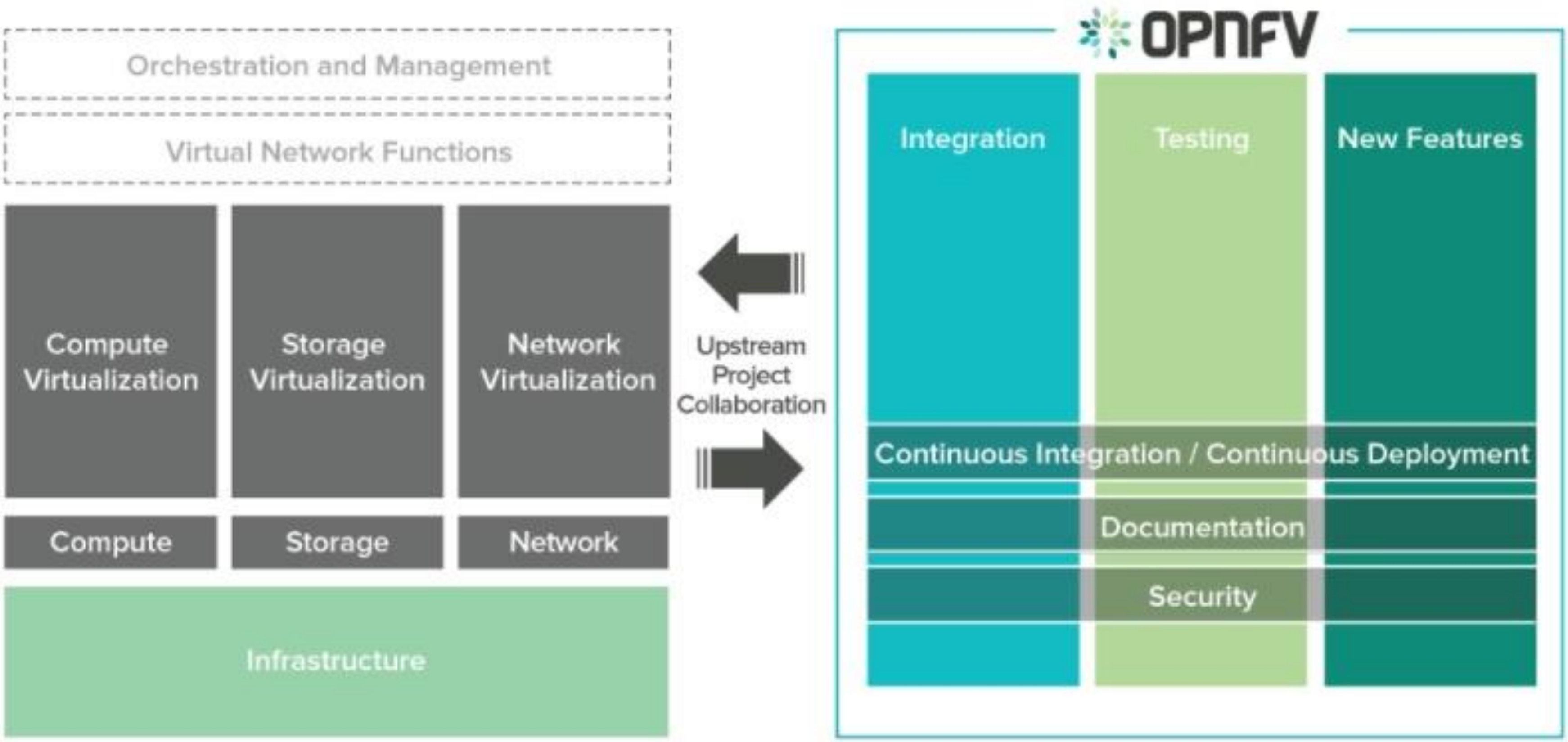
ICL-X Yu-Wei, Lee (Ray)

Network Function Virtualization (NFV)

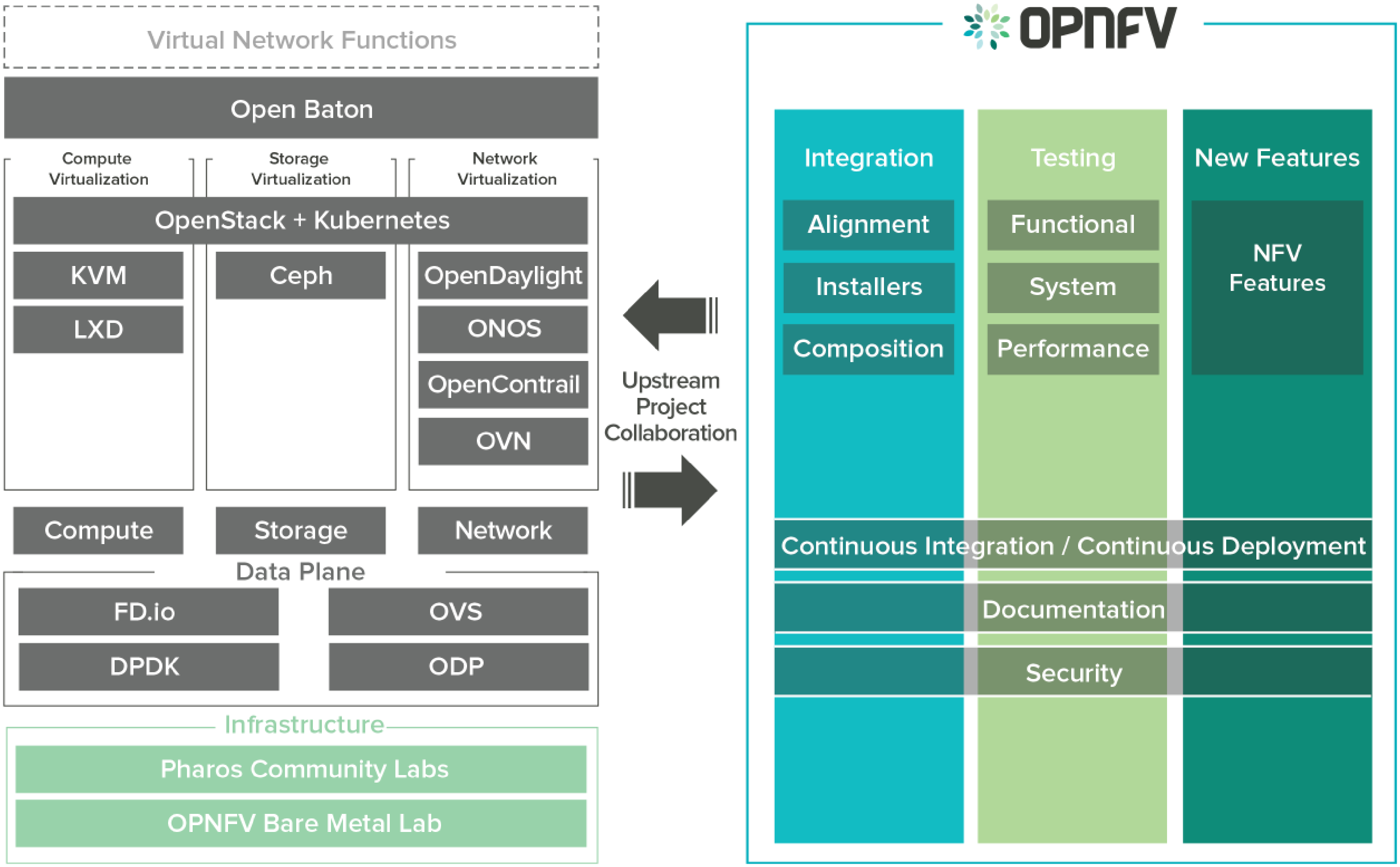


Open Platform for NFV (OPNFV)

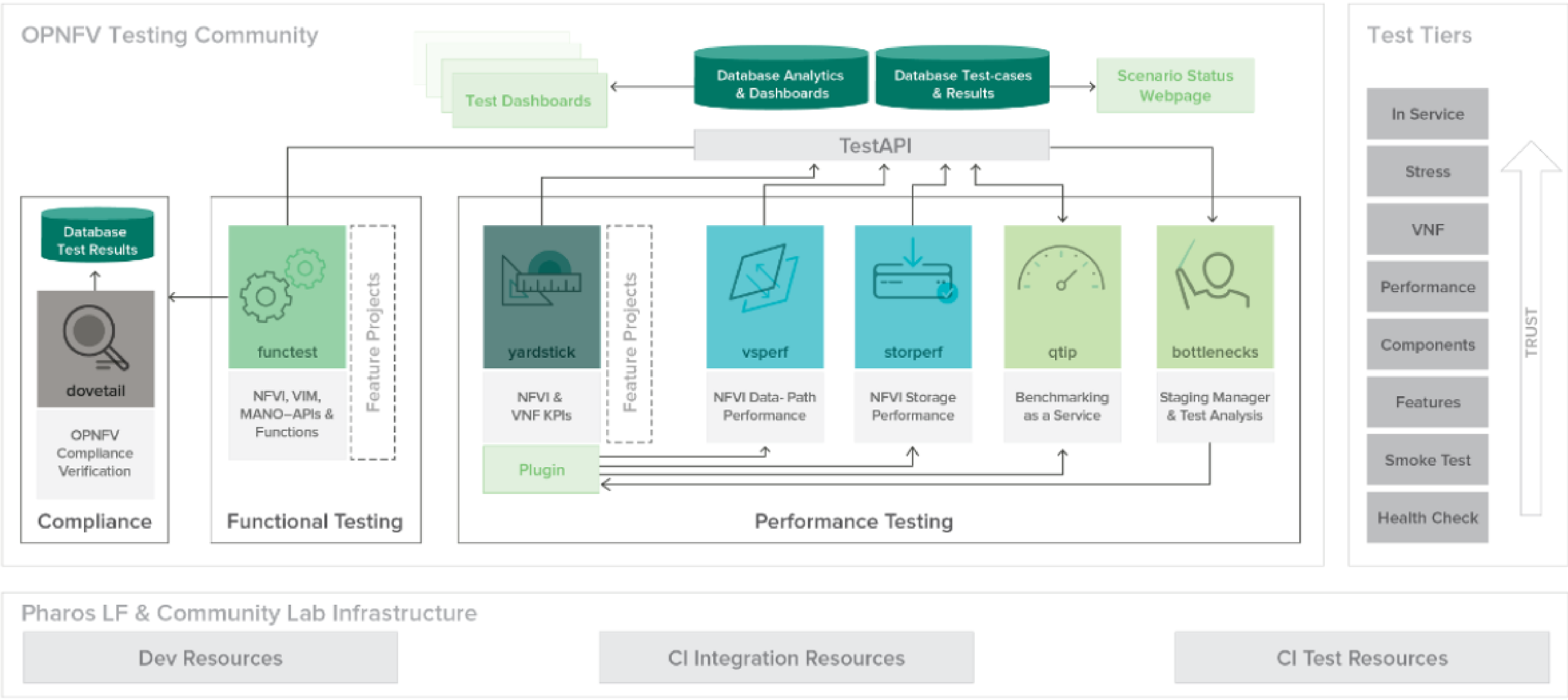
- **Open Platform for NFV (OPNFV) facilitates the development and evolution of NFV components across various open source ecosystems. Through system level integration, deployment and testing, OPNFV creates a reference NFV platform to accelerate the transformation of enterprise and service provider networks.**



OPNFV Euphrates Release



OPNFV Testing Ecosystem

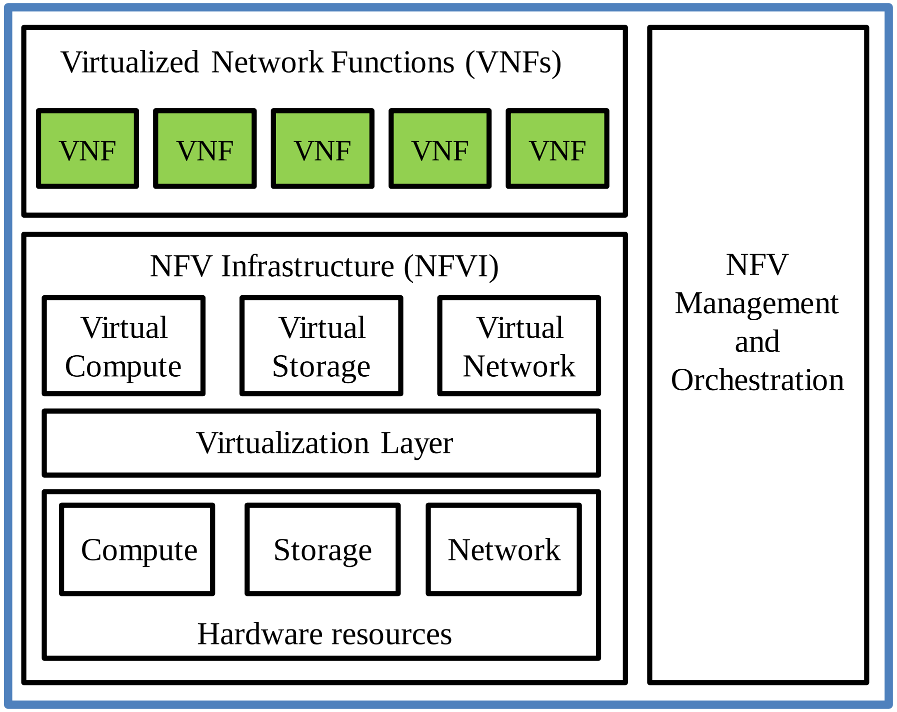


Functest: VIM and NFVI functional testing.
Yardstick: Verification of the infrastructure compliance when running VNF applications.
VSperf: Data-plane performance testing
Storperf: Storage Perfomance testing
QTIP: Platform Performance Benchmarking

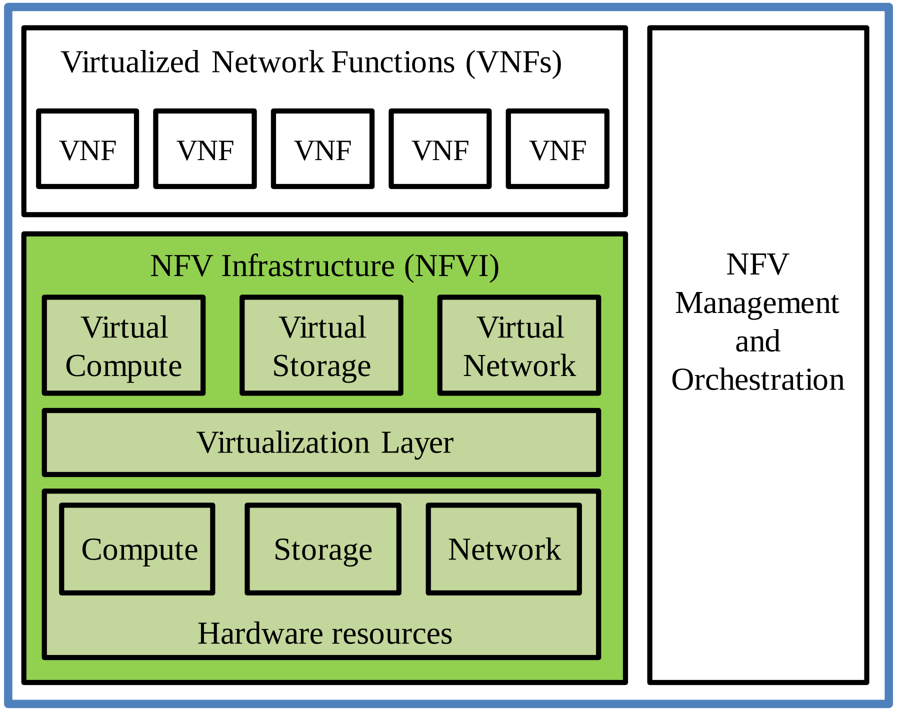
NFV Performance Lab

- Intel and ITRI build the NFV performance lab cooperatively
- Goal: NFV performance characterization
- NFVI characterization requires some “VNF”
- VNF characterization requires some NFV infrastructure

VNF Characterization focus



NFVI Characterization focus



Why Performance Lab

Why Performance Lab

- General lack of telco grade conformance / benchmarks
- Unclear network workload dimensions and stress vectors
- Missing system level capacity requirements
- Network workload scalability/agility implications on NFV
- Operators lack comprehensive information for TCO models to plan, procure and deploy NFV

OPNFV Yardstick – NSB Project (Network Service Benchmarking)

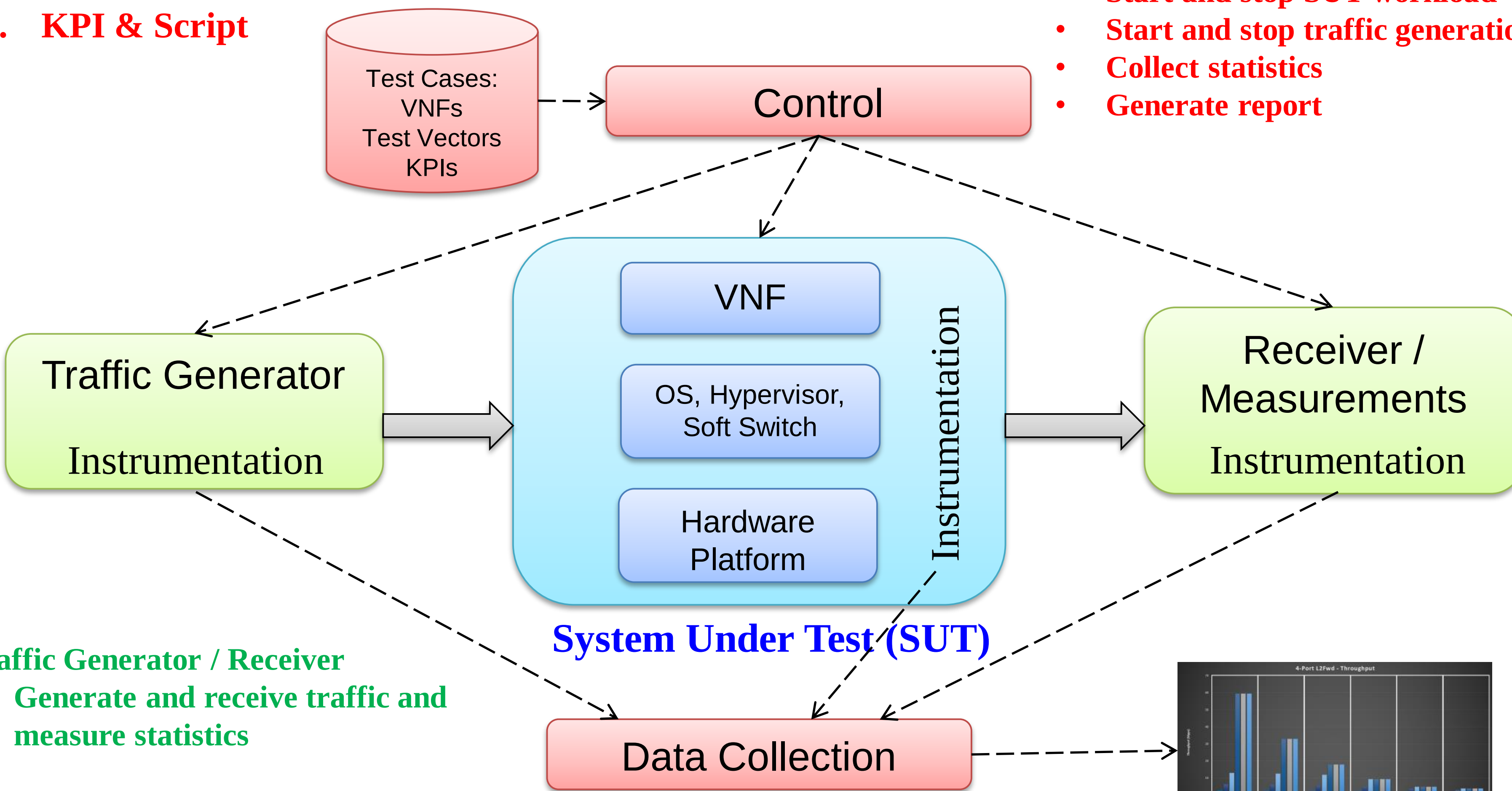
- NFVI/VNF characterization and benchmarking
- Different execution environments:
  - Native Linux environment
  - Standalone virtual environment
  - Managed virtualized environment (e.g. OpenStack etc.)
- Contributed by Intel

VNF and NFV Infrastructure Characterization

1. **System Under Test (SUT)**
2. **Traffic Generator**
3. **KPI & Script**

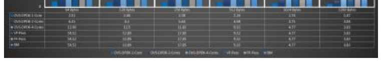
Control and Data Collection

- **Test cases measurement**
- **Start and stop SUT workload**
- **Start and stop traffic generation**
- **Collect statistics**
- **Generate report**



Traffic Generator / Receiver

- **Generate and receive traffic and measure statistics**



Generators

- IXIA
- SPIRENT

Commercial HW tools & VNF – full stack support

- PKTGEN
- MOONGEN
- PROX
- TREX

DPDK based tools

- Ostinato
- IPerf

Non DPDK based tools

- NG4T

Commercial tool for EPC characterization

- IMS Bench SIPP

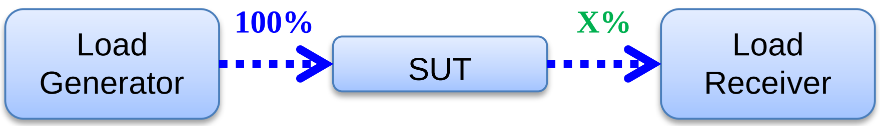
Open Source tool for IMS characterization

- And many others

Throughput

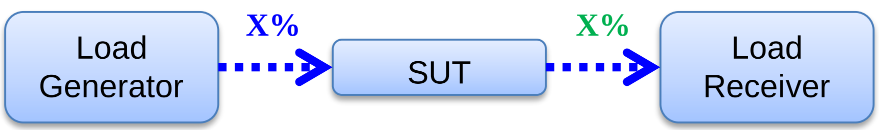
Forwarding Rate at Maximum Offered Load (FROML, RFC 2889)

- Generate at line rate, and measure forwarded packets
- Easy to run, fast
- Measure high load behavior (overload?)
- Might represent a completely different number than maximum forwarding rate



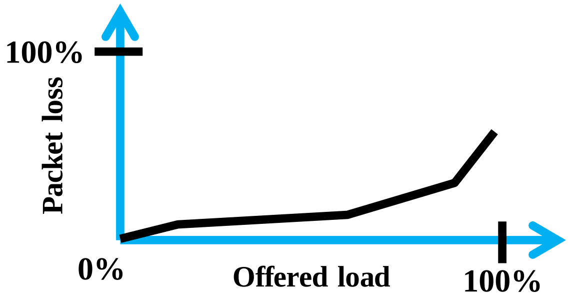
Maximum Forwarding Rate (MFR, RFC 2889)

- Start generating at 100% line rate and binary search for higher rate without packet loss
- Quite fast
- Might be very sensitive to spurious packet loss



Measure packet loss for any rate, starting from 0.1% to 100% of line rate

- Slow
- Better picture of the performance
- Will highlight spurious packet loss



ITRI
Industrial Technology
Research Institute

Intel Xeon Processor E5-2695 v4
NFVI Performance Report

Produced by ITRI Performance Lab

NFVI Characterization using NSB

VNF:
L2-Fwd
L3-Fwd
ACL
MPLS

Packet Size:
64 Bytes
128 Bytes
256 Bytes
512 Bytes
1024 Bytes
1280 Bytes

Test Cases:
VNFs
Test Vectors
KPIs

Control

Control and Data Collection
- **Test cases measurement**
- **Start and stop SUT workload**
- **Start and stop traffic generation**
- **Collect statistics**
- **Generate report**

Traffic Generator
Instrumentation

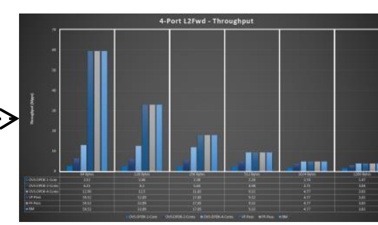
VNF
OS, Hypervisor, Soft Switch
Hardware Platform
Instrumentation

System Under Test (SUT)

Receiver / Measurements
Instrumentation

Traffic Generator: PROX

Data Collection



Traffic Generator / Receiver
- **Generate and receive traffic and measure statistics**

Platform Configuration

Item	Description
Server Platform	Supermicro X10DRH-C/I Dual Integrated 1GbE ports via Intel® i350-AM2 Gigabit Ethernet
Chipset	Intel® C612 chipset (formerly Lynx-H Chipset)
Processor	2x Intel® Xeon® Processor E5-2695 v4 (formerly Broadwell-EP) 2.10 GHz; 120 W; 45 MB cache per processor 18 cores(HT disabled)
Memory	64GB Total; Samsung 8GB 2Rx8 PC4-2400MHz, 8GB per channel, 4 Channels
PCIe	6x PCI-E 3.0 x8 slot 1x PCI-E 3.0 x16 slot
Local Storage	2 x Seagate Nytro® XF1230 SATA SSD 240G
NICs	2 x Intel® Ethernet Converged Network Adapter X710-DA4 Total: 8 Ports; 4 ports are used in tests.
BIOS	AMIBIOS Version: 2.0a Build Date : 06/30/2016

Software Configurations (1/2)

Software Component	Version
Host Operating System	Ubuntu 16.04 Kernel version : 4.4.0-101-generic
VM Operating System	Ubuntu 16.04 Kernel version : 4.4.0-98-generic
QEMU-KVM	QEMU emulator version 2.10.1 (v2.10.1-dirty)
Open vSwitch (native OvS and OvS with DPDK)	Open vSwitch 2.8.1 DB Schema 7.15.0
Intel® Ethernet Drivers (Physical Function)	driver: i40e version: 2.2.4 firmware-version: 5.05 0x80002a9b 1.1313.0 expansion-rom-version: bus-info: 0000:04:00.2 supports-statistics: yes supports-test: yes supports-eeprom-access: yes supports-register-dump: yes

Software Configurations (2/2)

System Compilation	Configuration
Intel® Ethernet Drivers (Virtual Function)	driver: i40evf version: 3.1.4 firmware-version: 5.05 0x80002a9b 1.1313.0 supports-statistics: yes supports-test: yes supports-eeprom-access: yes supports-register-dump: yes
DPDK	DPDK version : 17.05.2 (http://fast.dpdk.org/rel/dpdk-17.05.2.tar.xz) Compiled with : EXTRA_CFLAGS="–Ofast –march=native" make –j10
Packetgen	PROX (HEAD detached at 4a31f76)
OVS	OVS compiled as follows: # ./configure --with-dpdk=$DPDK_DIR/x86_64-native-linuxapp-gcc CFLAGS="-Ofast –march=native" # make "CFLAGS=-Ofast –march=native" -j10

BIOS Tuning Configurations (1/2)

Menu (Advanced)	BIOS Setting	Required Settings for Performance	BIOS Default
CPU Configuration -> Advanced Power Management Configuration			
	Hyper-Threading (ALL)	Disable	Enable
	Power Technology	Disable	Custom
	Energy Performance Tuning	Disable	Enable
	Energy Performance BIAS Setting	Performance	Enable
	Energy Efficient Turbo	Disable	Enable
-> CPU P State Control	EIST (P-States)	Disable	Enable
-> CPU P State Control	Turbo Mode	Disable	Enable
-> CPU P State Control	P-State Coordination	HW_ALL	HW_ALL
-> CPU C State Control	Package C State Limit	[C0/C1 State]	[C6 (Retention)]
-> CPU C State Control	CPU C3 Report	Disable	Enable
-> CPU C State Control	CPU C6 Report	Disable	Enable
-> CPU C State Control	Enhanced Halt State (C1E)	Disable	Enable

Ref :
Intel® Xeon® Processor E5-2695 v4 Performance report: Open vswitchwith dpdk(OVS-DPDK) VS. VT-d PCI passthrough P.10

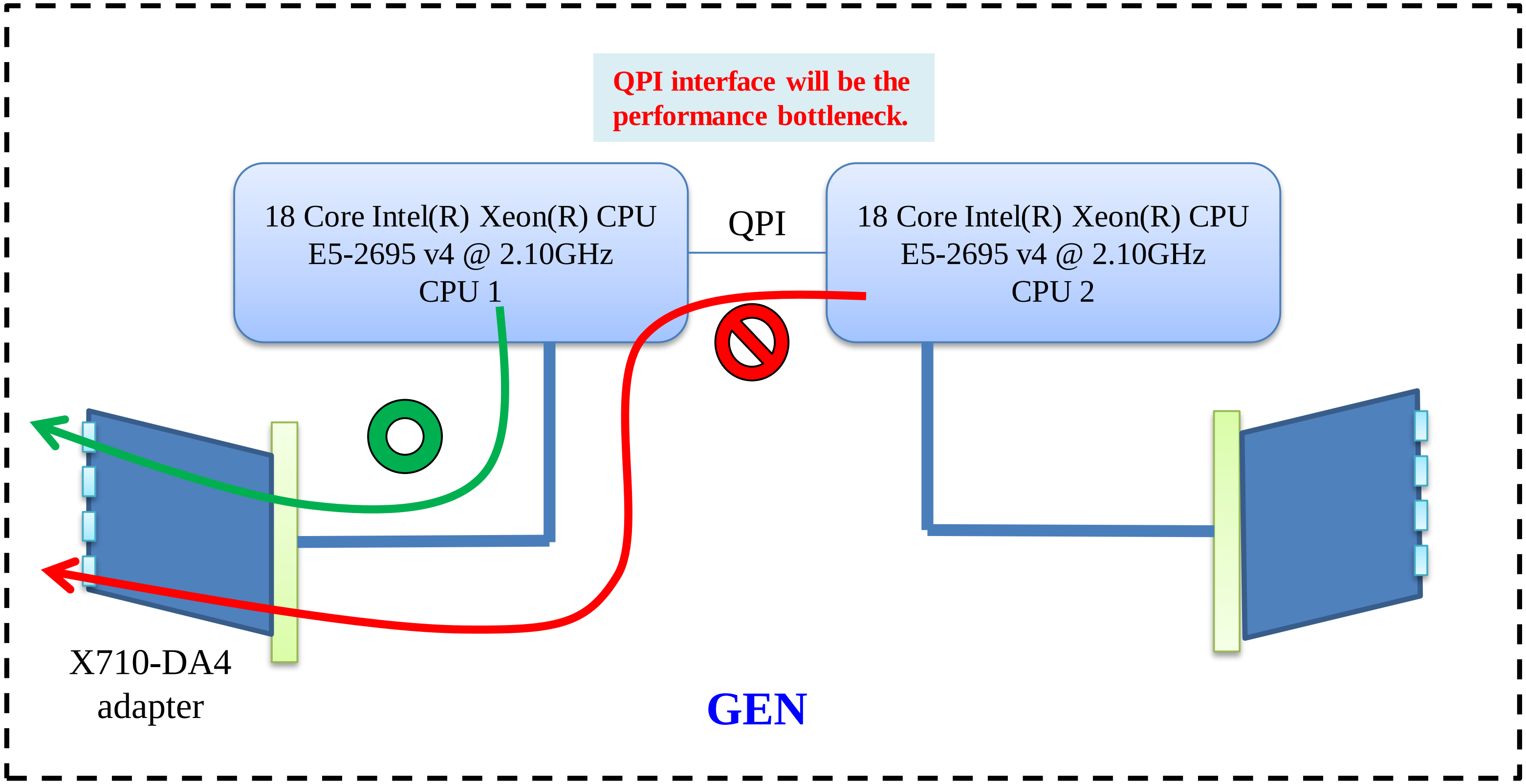
BIOS Tuning Configurations (2/2)

*In passthrough test cases, we will turn on VT-d

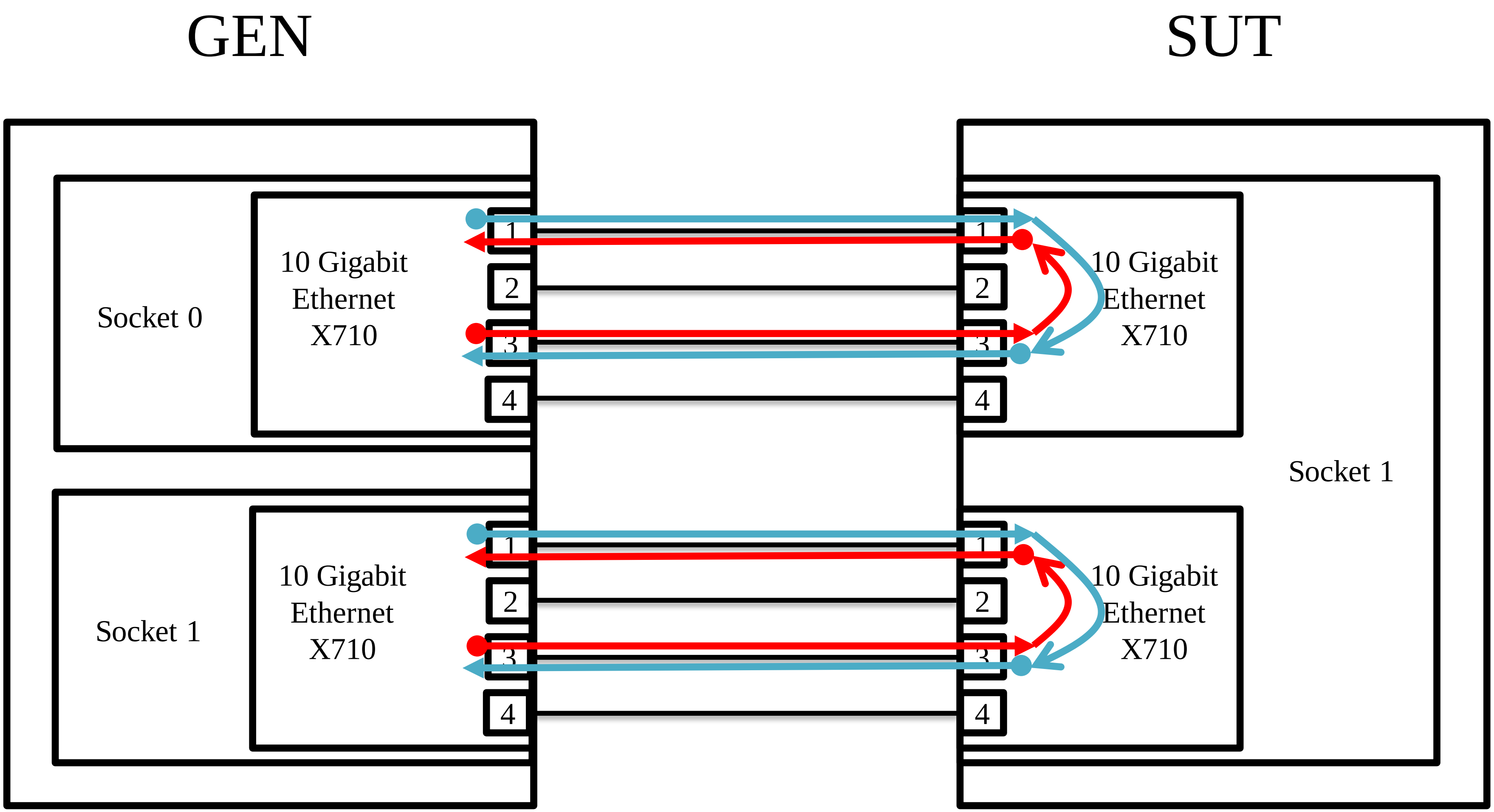
Chipset Configuration			
-> North Bridge -> IIO Configuration	EV DFX Features	Disable	Disable
	Intel VT for Directed I/O (VT-d)	Disable	Enable
->North Bridge -> IOAT Configuration	Enable IOAT	Enable	Enable
	No Snoop	Disable	Disable
	Relaxed Ordering	Disable	Disable
-> North Bridge -> QPI Configuration			
	Link L0 P	Disable	Enable
	Link L1	Disable	Enable
	COD Enable	Disable	Auto
	Early Snoop	Disable	Auto
	Isoc Mode	Disable	Disable
-> North Bridge ->Memory Configuration	Enforce POR	Disable	Auto
	Memory Frequency	2400	Auto
	DRAM RAPL Baseline	Disable	Auto
	A7 Mode	Enable	Enable
-> South Bridge	EHCI Hand-off	Disable	Auto
	USB3.0 Support	Disable	Enable
PCIe/PCI/PnP Configuration	ASPM	Disable	Enable
	Onboard LAN 1 OPROM	Disable	PXE

The Mapping Between CPU and NIC

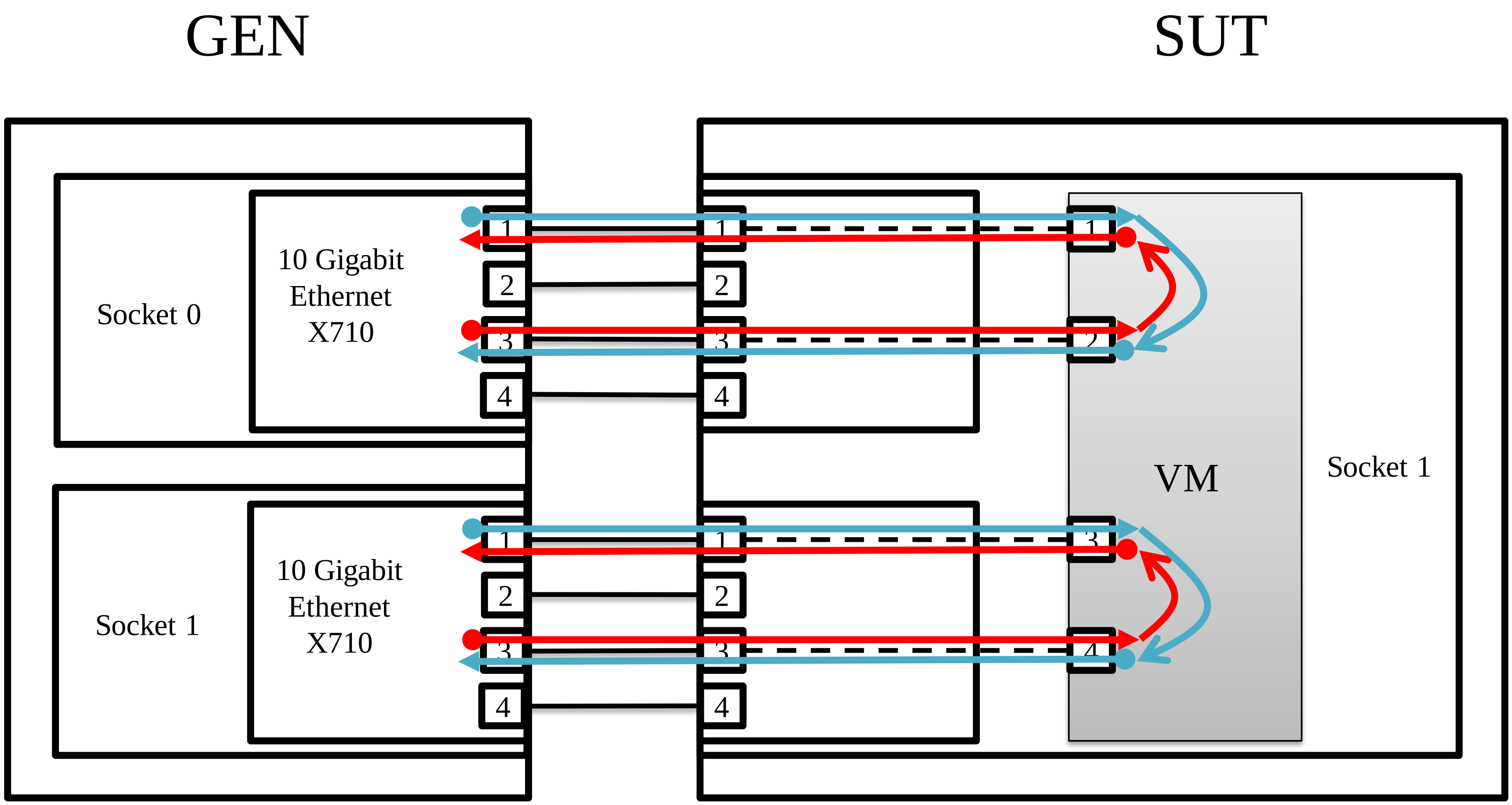
- The CPU core and NIC for packet generator should be in the same CPU socket, or the QPI interface will be the performance bottleneck.
  - 10GbE line rate (64bytes packet) = 10.00e9 bits / (8 bits * (64 + 20)bytes) = 14.88e6 packets.
  - CPU and NIC in same socket, 10G port could generate 14.88 Mpps.
  - CPU and NIC in different socket, 10G port only could generate 10 Mpps.



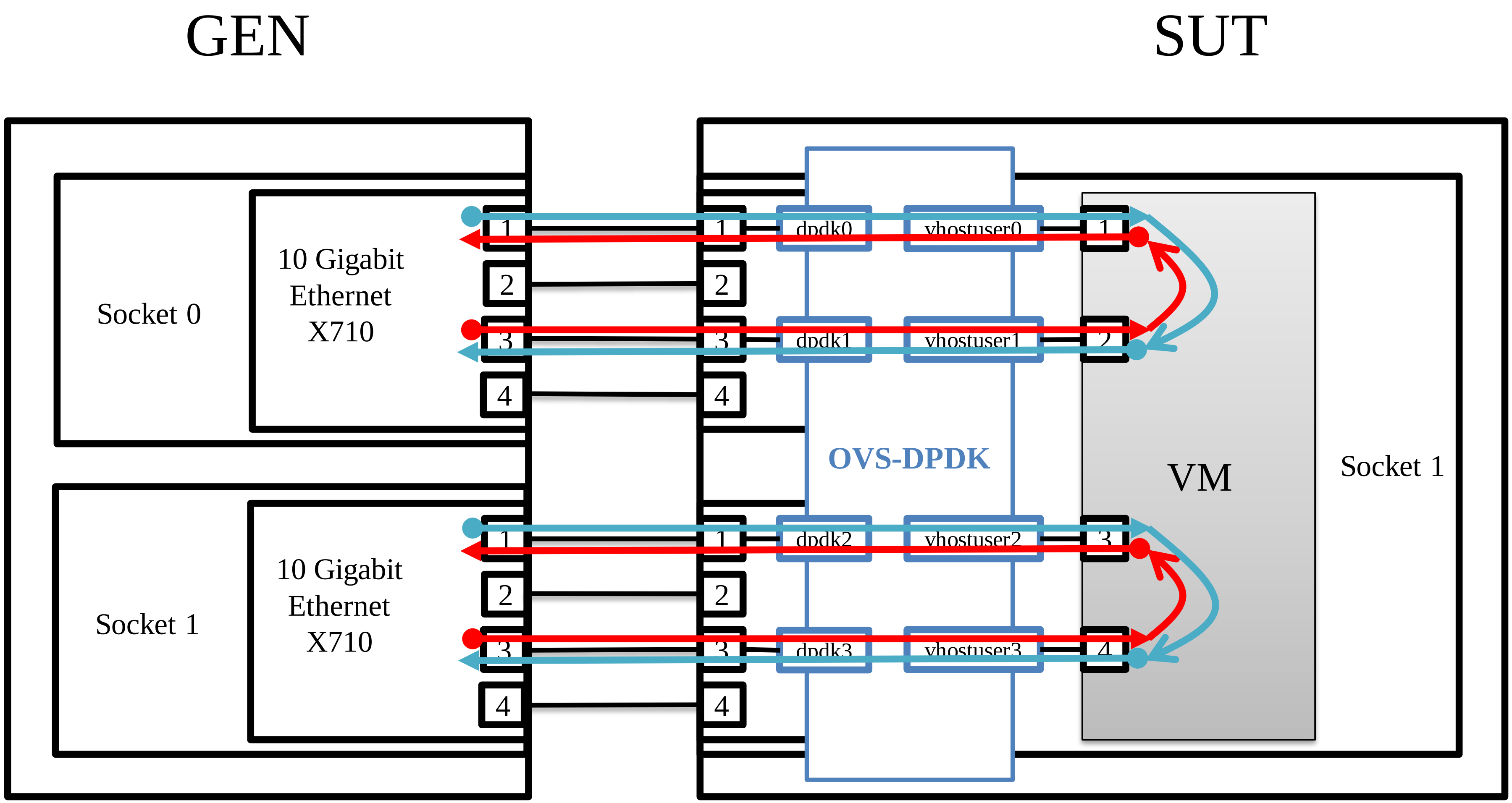
Baremetal Architecture



SRIOV Passthrough Architecture



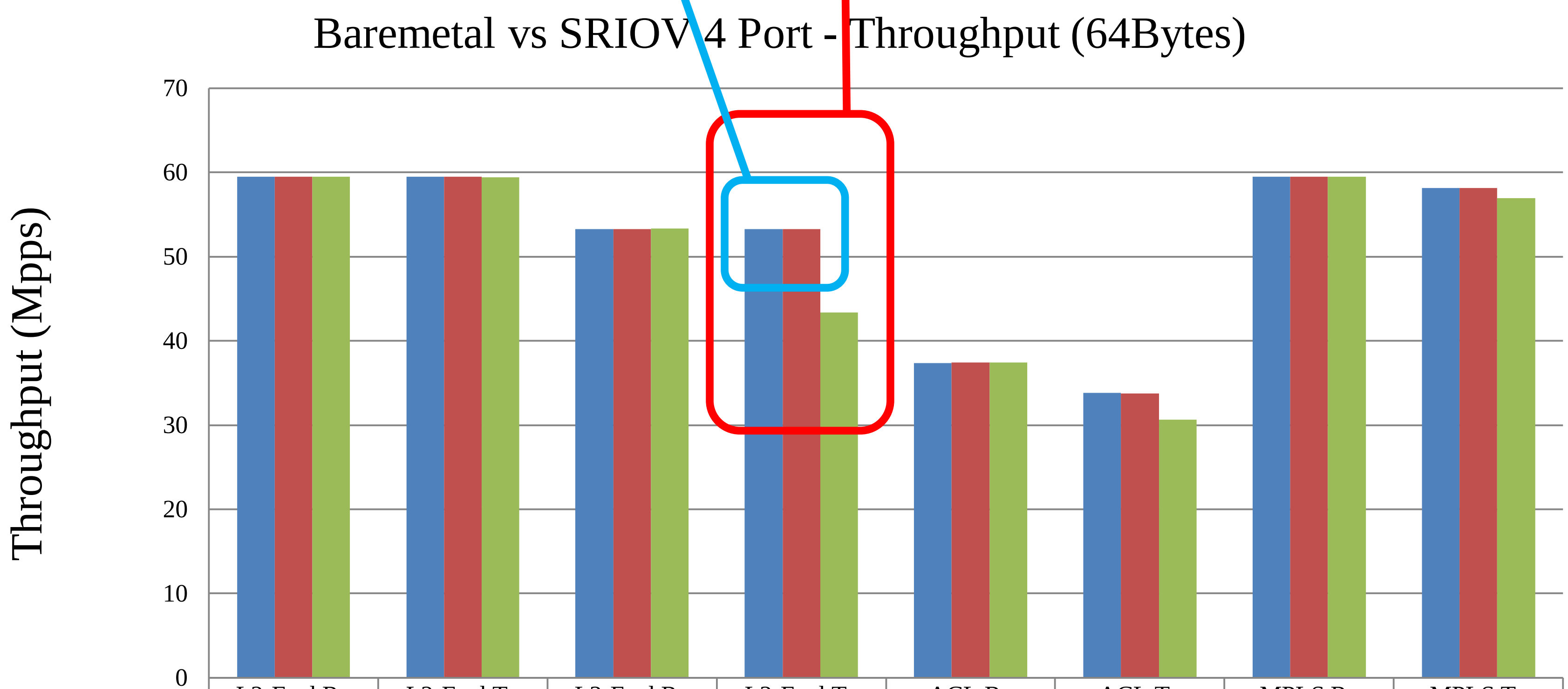
OVS-DPDK Architecture



SRIOV Performance

VM performance with NUMA-aware configuration is similar to Baremetal performance.

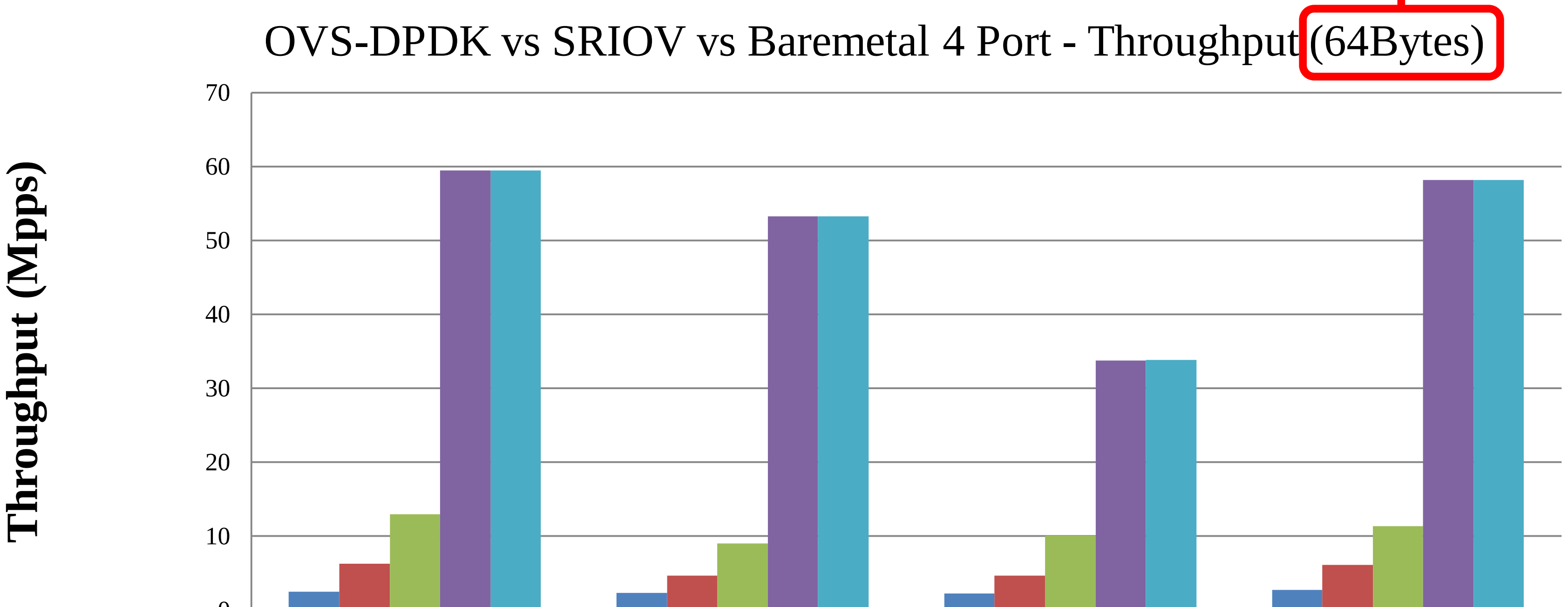
VM without NUMA-aware configuration may decrease 10 – 20% performance.



	L2-Fwd Rx	L2-Fwd Tx	L3-Fwd Rx	L3-Fwd Tx	ACL Rx	ACL Tx	MPLS Rx	MPLS Tx
BM	59.52	59.52	53.29	53.29	37.38	33.82	59.52	58.17
VM w/ NUMA	59.52	59.52	53.29	53.29	37.41	33.78	59.52	58.17
VM w/o NUMA	59.52	59.46	53.36	43.4	37.46	30.66	59.52	56.96

OVS-DPDK Performance

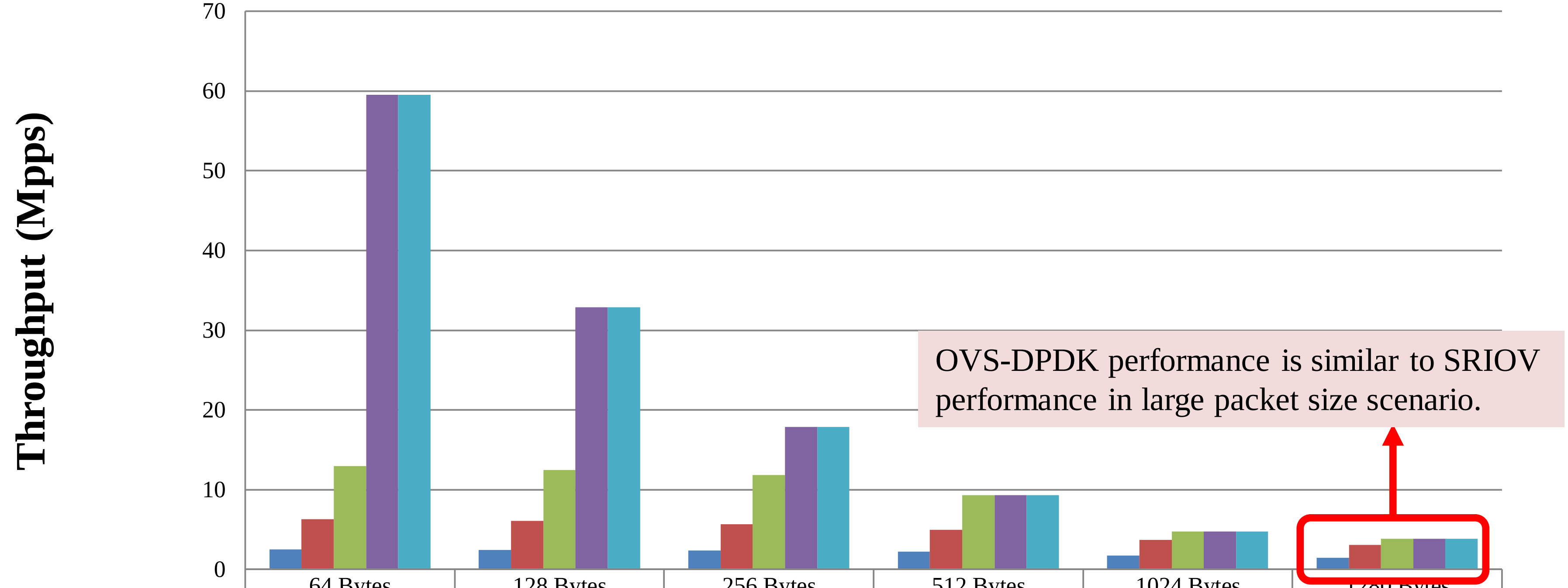
SRIOV could get better performance than OVS-DPDK in small packet size scenario.



	L2-Fwd Tx	L3-Fwd Tx	ACL Tx	MPLS Tx
OVS-DPDK-1-Core	2.51	2.33	2.24	2.71
OVS-DPDK-2-Cores	6.31	4.71	4.64	6.11
OVS-DPDK-4-Cores	12.96	9.06	10.1	11.36
SRIOV	59.52	53.29	33.78	58.17
BM	59.52	53.29	33.82	58.17

OVS-DPDK Performance

OVS-DPDK vs SRIOV vs Baremetal 4 Port - Throughput (L2-Fwd)

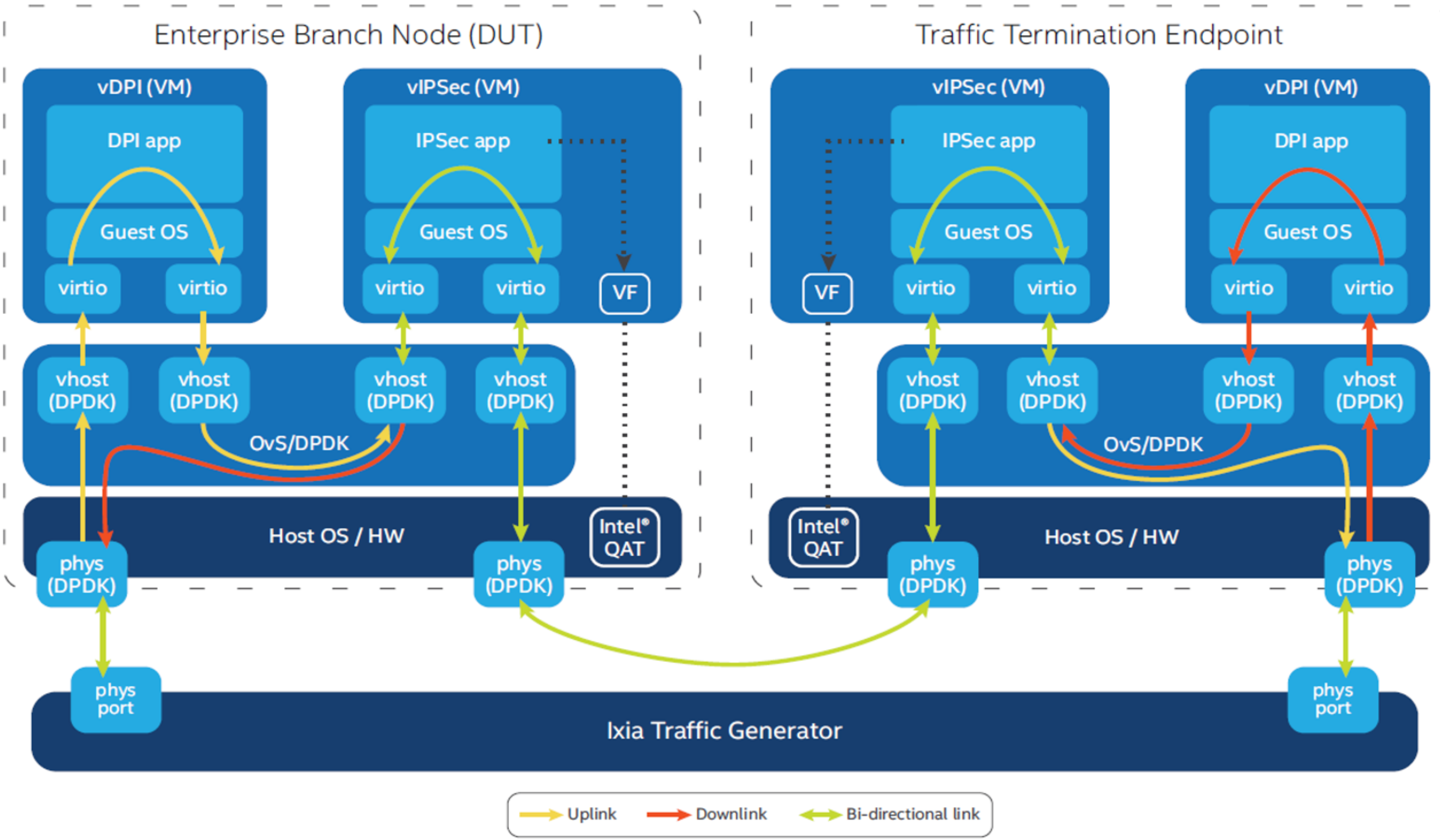


	64 Bytes	128 Bytes	256 Bytes	512 Bytes	1024 Bytes	1280 Bytes
OVS-DPDK-1-Core	2.51	2.46	2.38	2.26	1.74	1.47
OVS-DPDK-2-Cores	6.31	6.1	5.64	4.98	3.71	3.04
OVS-DPDK-4-Cores	12.96	12.5	11.82	9.32	4.77	3.83
SRIOV	59.52	32.89	17.85	9.32	4.77	3.83
BM	59.52	32.89	17.85	9.32	4.77	3.83

Intel Atom Processor C3758 SDWAN Performance Report

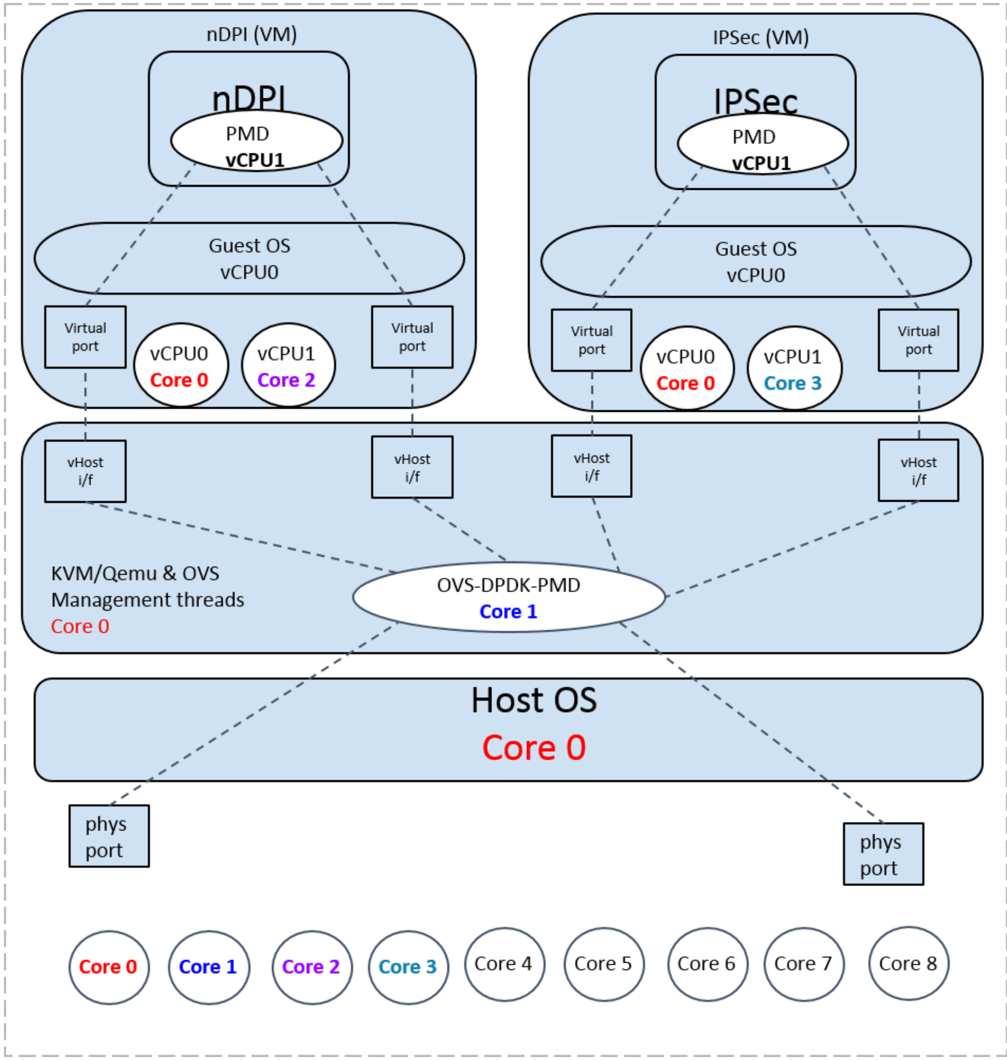
Produced by ITRI Performance Lab

SDWAN Topology



CPU Core Assignment – 4 Cores, 1 PMD Thread

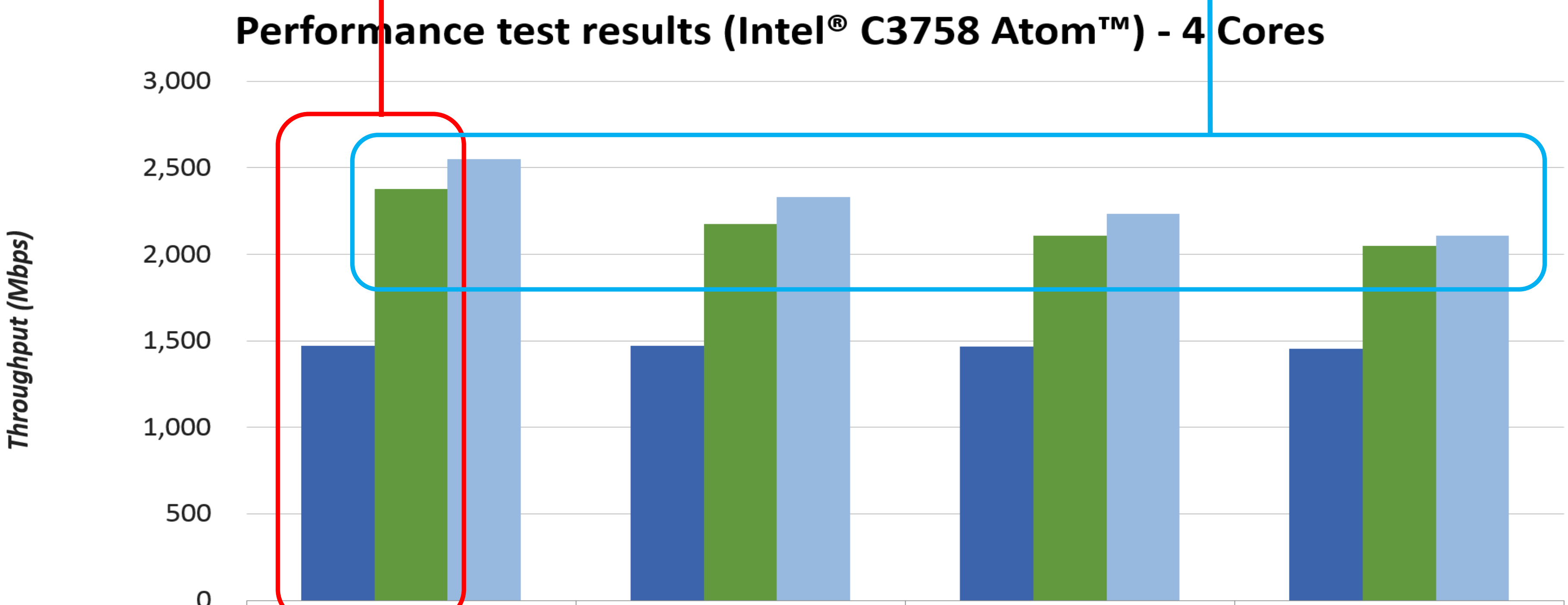
4-Core Configuration	
Host OS	Core 0
Hypervisor (KVM/Qemu)	Core 0
OVS Mgmt Threads	Core 0
OVS DPDK PMD	Core 1
DPI VNF (vCPU threads)	Core 0 (VM vCPU 0), Core 2 (VM vCPU 1)
IPsec VNF (vCPU threads)	Core 0 (VM vCPU 0), Core 3 (VM vCPU 1)



SDWAN Performance with 4 Cores

Use QAT accelerator can improve IPSec performance

QAT performance will be decrease when flow number is increase, OVS is the performance bottleneck.



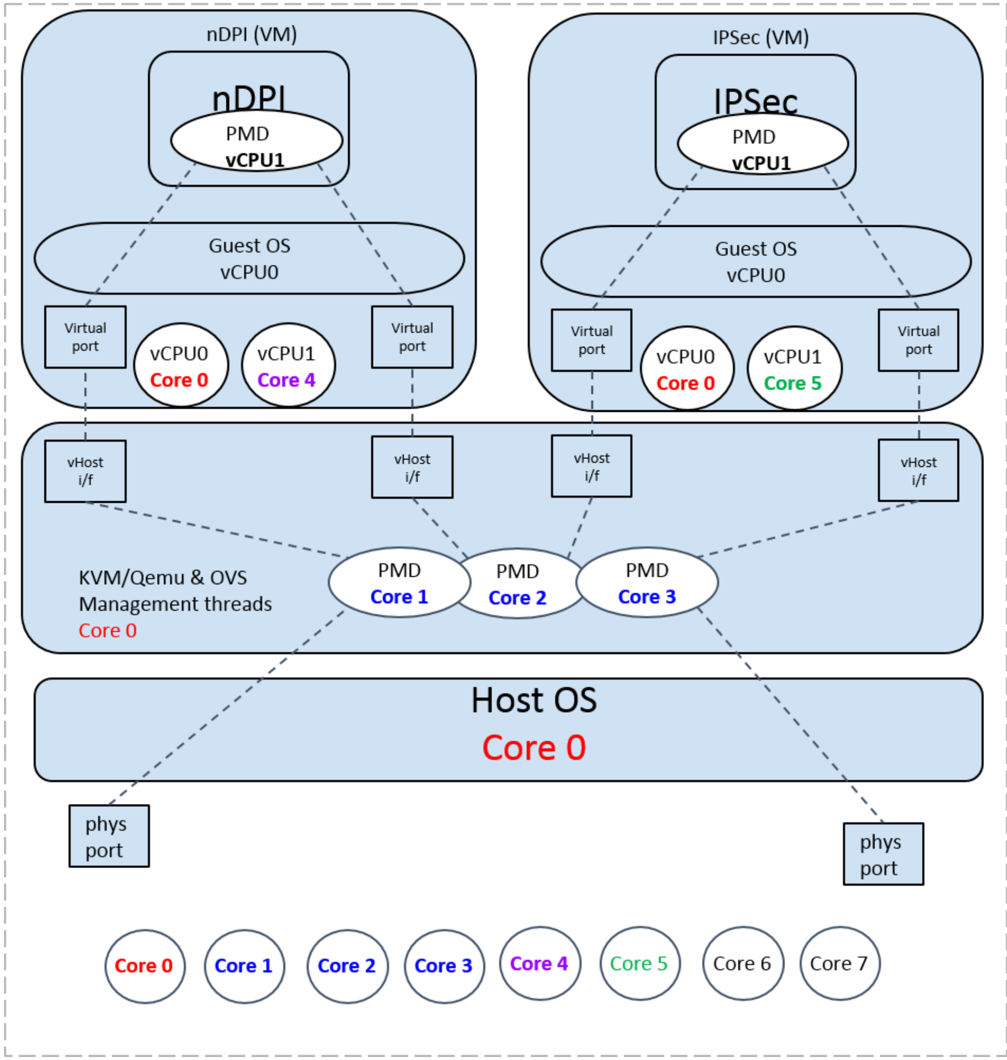
	100	500	1,000	2,000
ITRI Lab SHA1 - 10G	1,469.20	1,469.20	1,466.45	1,454.09
ITRI Lab SHA1 (QAT) - 10G	2,378.46	2,176.69	2,107.14	2,049.09
ITRI Lab GCM - 10G	2,550.45	2,330.27	2,234.31	2,107.14

Flows

ITRI Lab SHA1 - 10G　ITRI Lab SHA1 (QAT) - 10G　ITRI Lab GCM - 10G

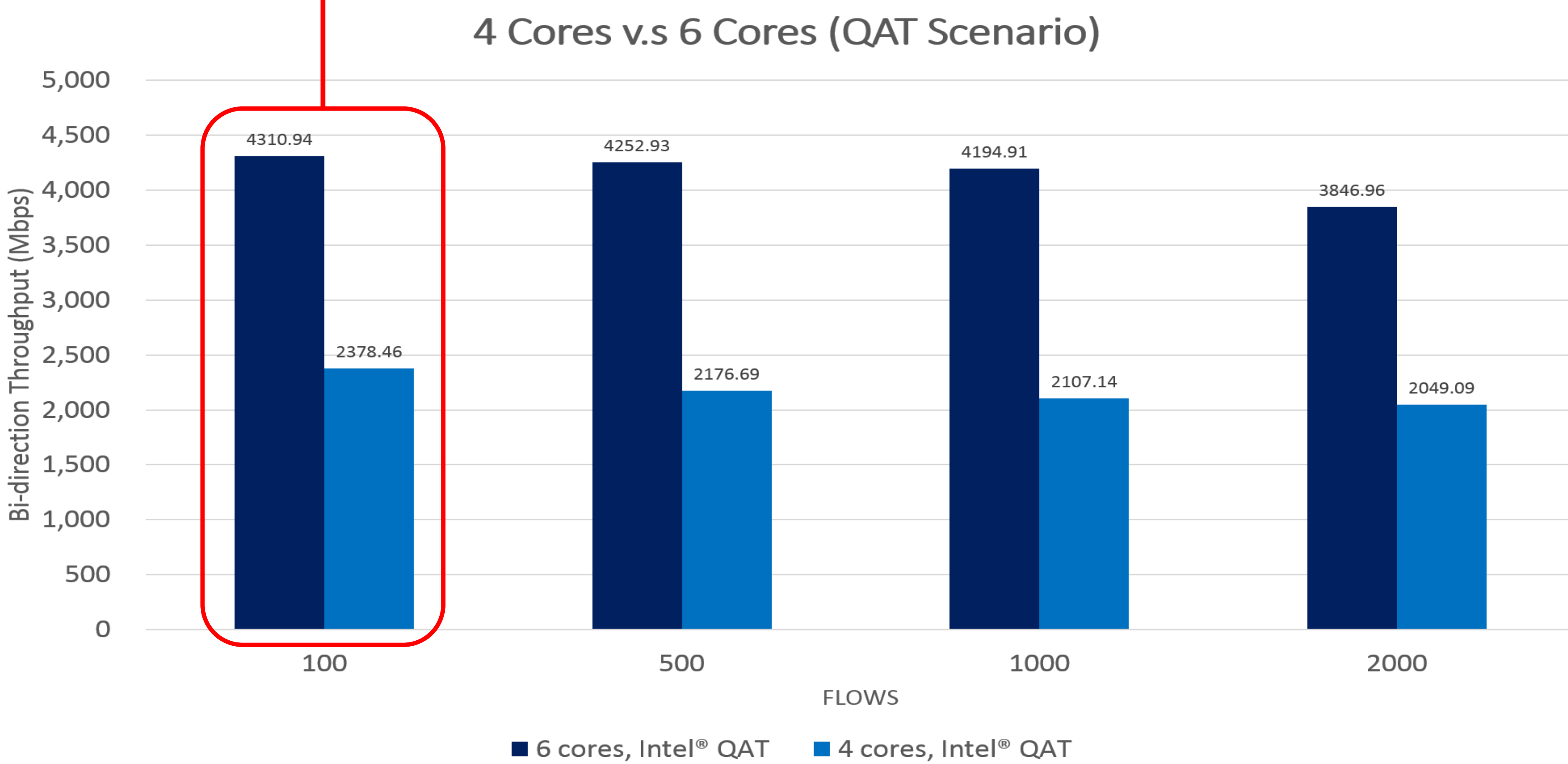
CPU Core Assignment – 6 Cores, 3 PMD Thread

6-Core Configuration	
Host OS	Core 0
Hypervisor (KVM/Qemu)	Core 0
OVS Mgmt Threads	Core 0
OVS DPDK PMD	Core 1 Core 2 Core 3
DPI VNF (vCPU threads)	Core 0 (VM vCPU 0), Core 4 (VM vCPU 1)
IPsec VNF (vCPU threads)	Core 0 (VM vCPU 0), Core 5 (VM vCPU 1)



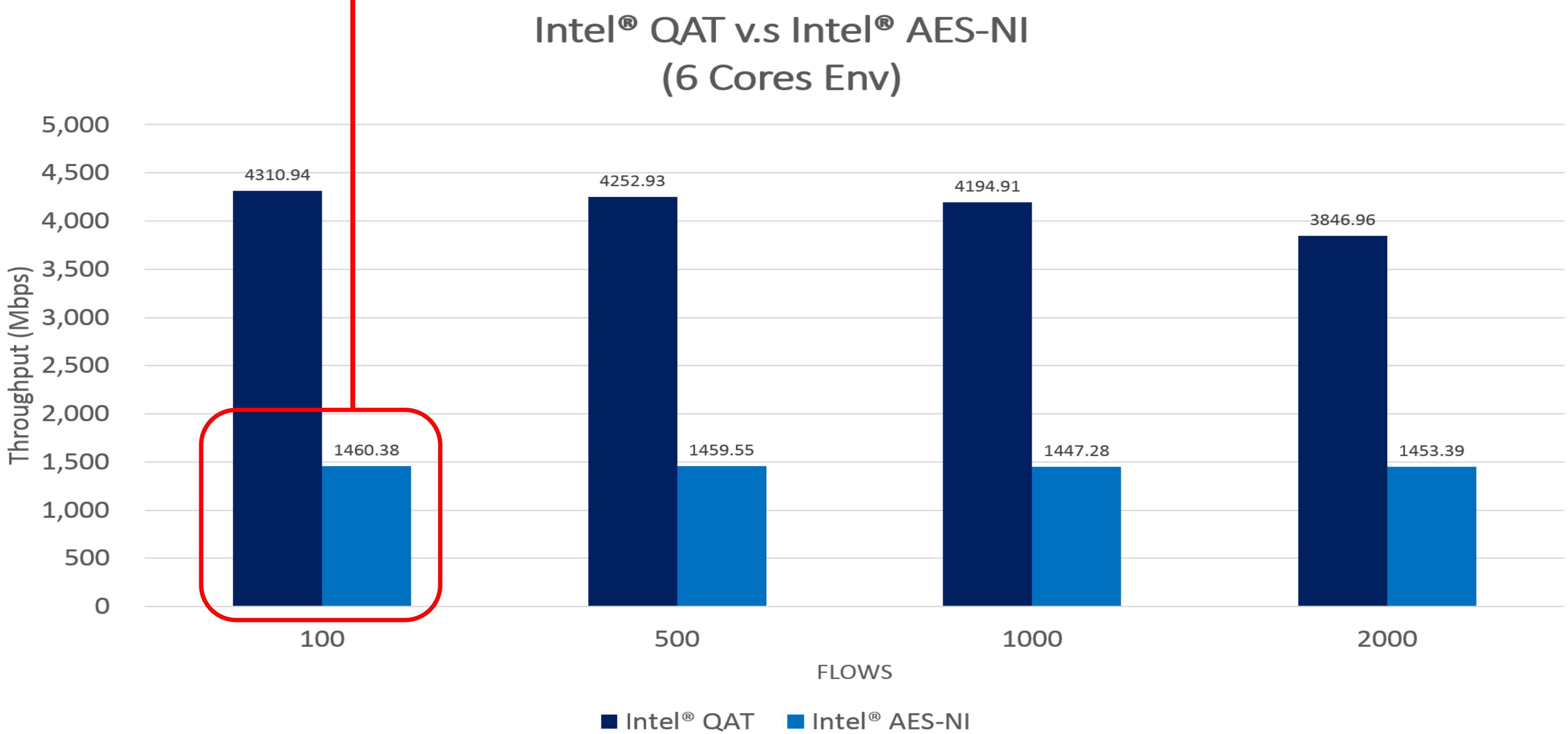
SDWAN Performance of QAT Scenario

QAT performance will be increase when we allocate more CPU core to OVS.



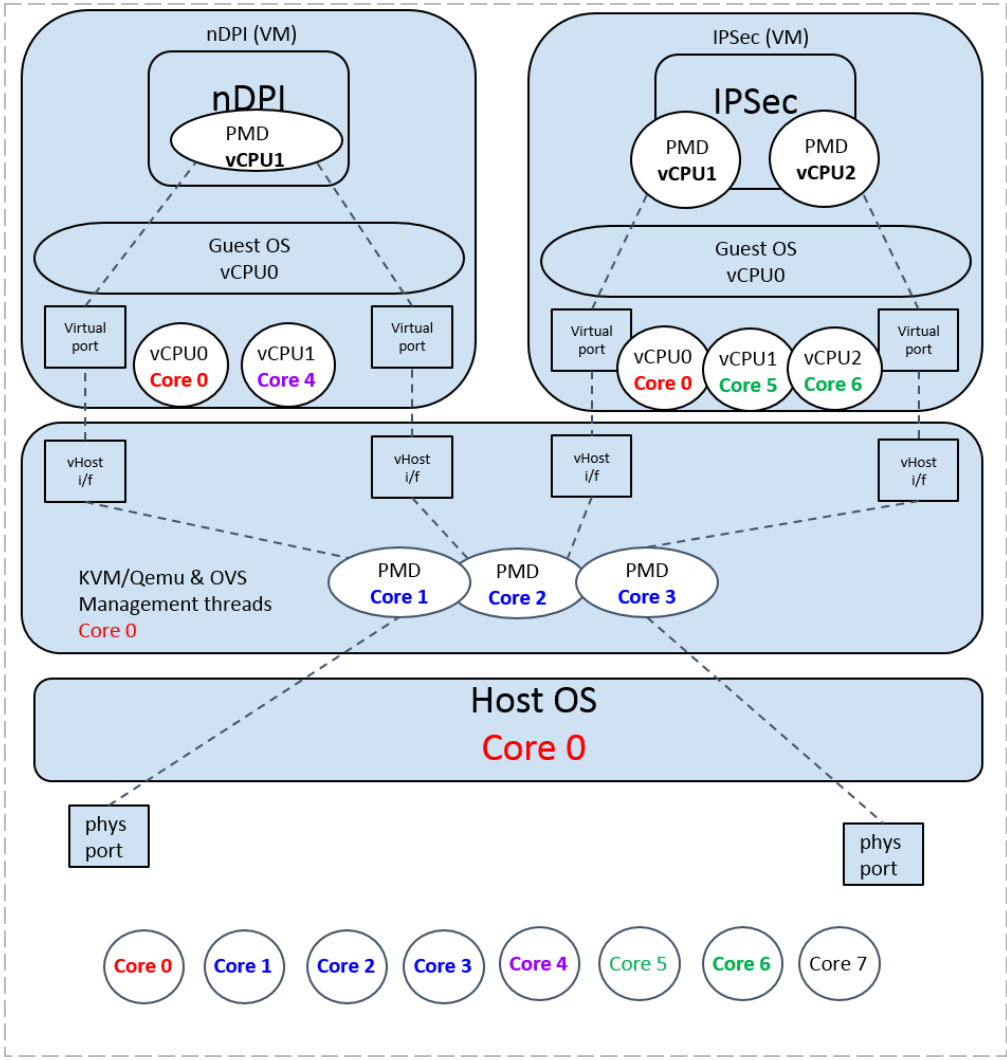
SDWAN Performance on 6 Core Environment

The performance of IPSec SHA without QAT is almost the same with 4 core scenario, bottleneck is IPSec itself.



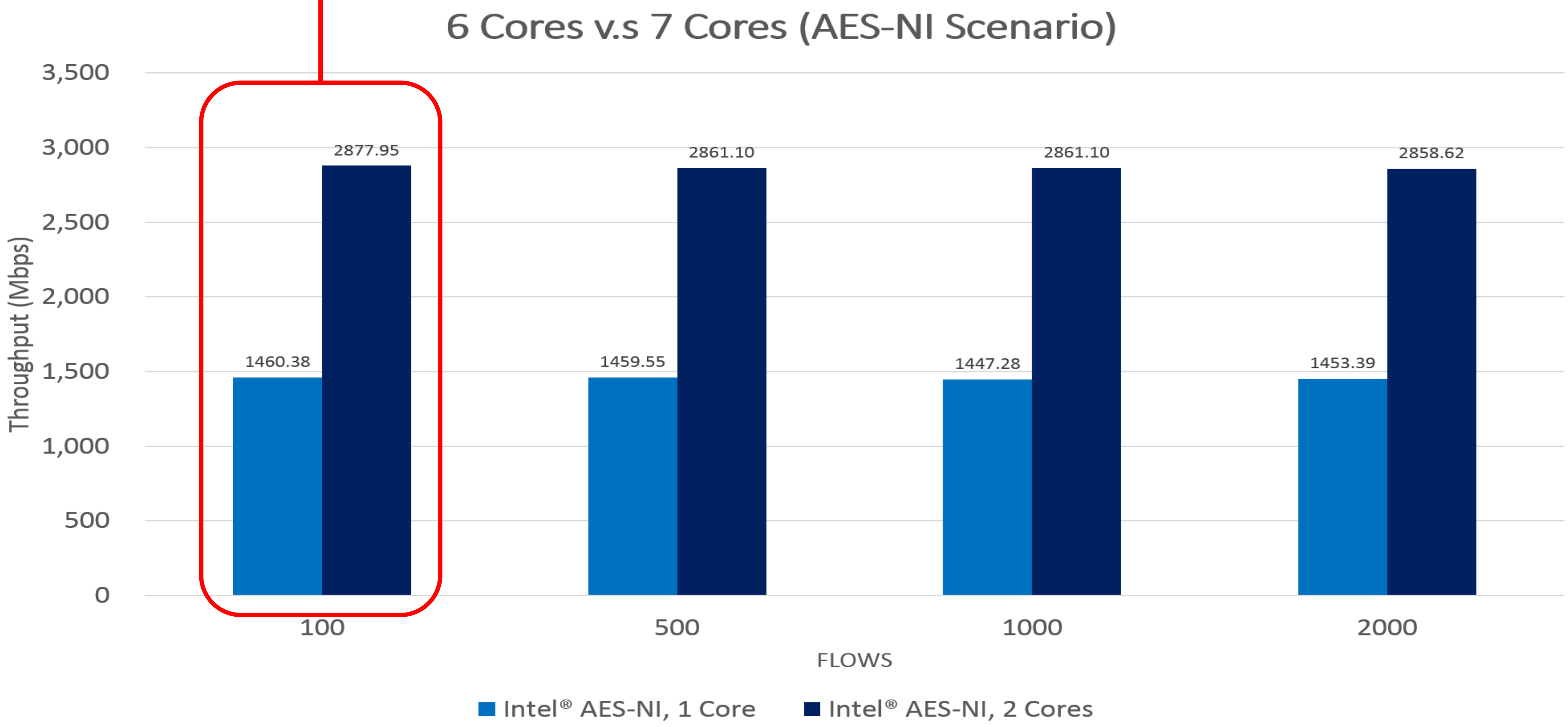
CPU Core Assignment – 7 Cores, 3 PMD Thread

7-Core Configuration	
Host OS	Core 0
Hypervisor (KVM/Qemu)	Core 0
OVS Mgmt Threads	Core 0
OVS DPDK PMD	Core 1 Core 2 Core 3
DPI VNF (vCPU threads)	Core 0 (VM vCPU 0), Core 4 (VM vCPU 1)
IPsec VNF (vCPU threads)	Core 0 (VM vCPU 0), Core 5 (VM vCPU 1), Core 6 (VM vCPU 2)



SDWAN Performance of AES-NI Scenario

The performance of IPSec AES-NI will be increase when we allocate more CPU core to IPSec VNF.



Intel Xeon Processor E5-2695 v4 Performance Report: 2-8 VMs Service Chain with SR-IOV, OVS-DPDK, VPP and SPP

Produced by ITRI Performance Lab

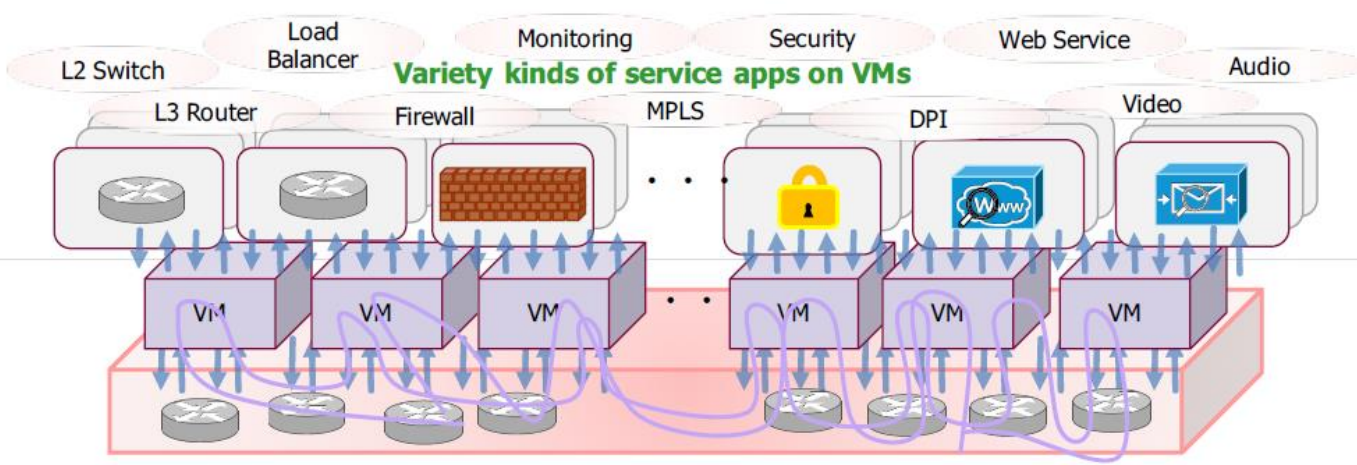
NTT's Presentation on DPDK Summit

VM Chaining Scenario

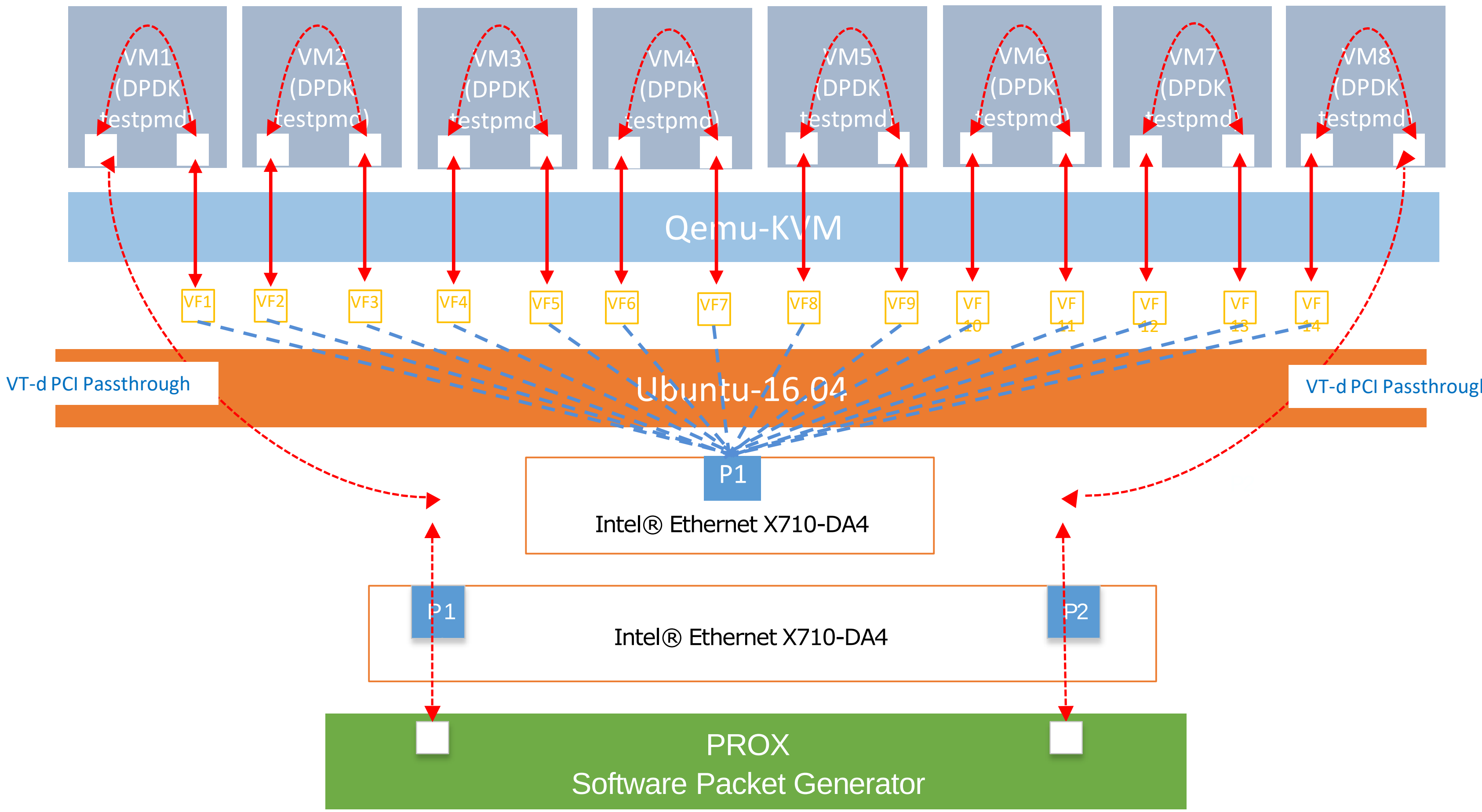
Motivation

DPDK

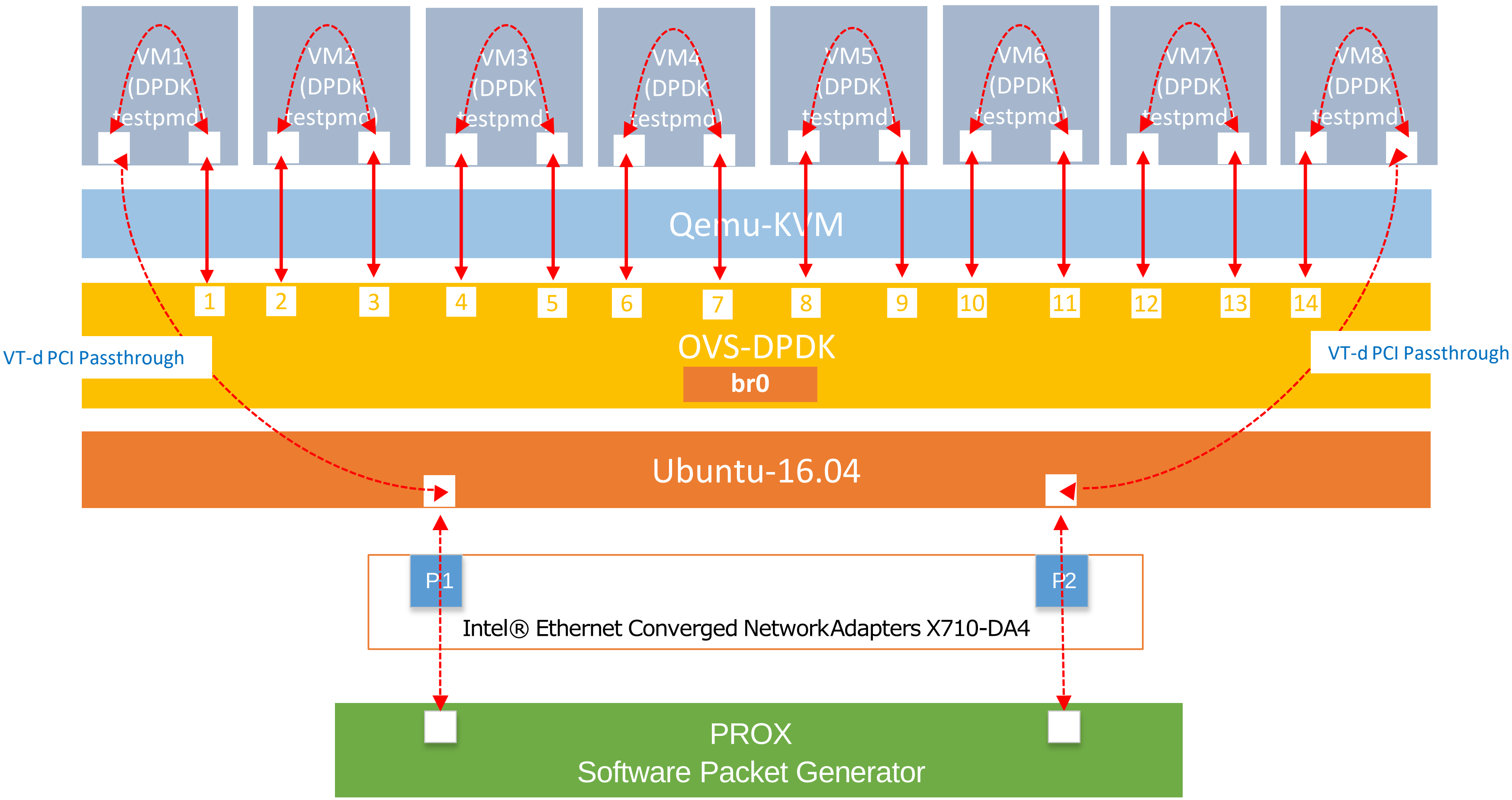
- Large-scale cloud for telecom services
- Service Function Chaining for virtual network appliances
- Flexibility, Maintainability and High-Performance



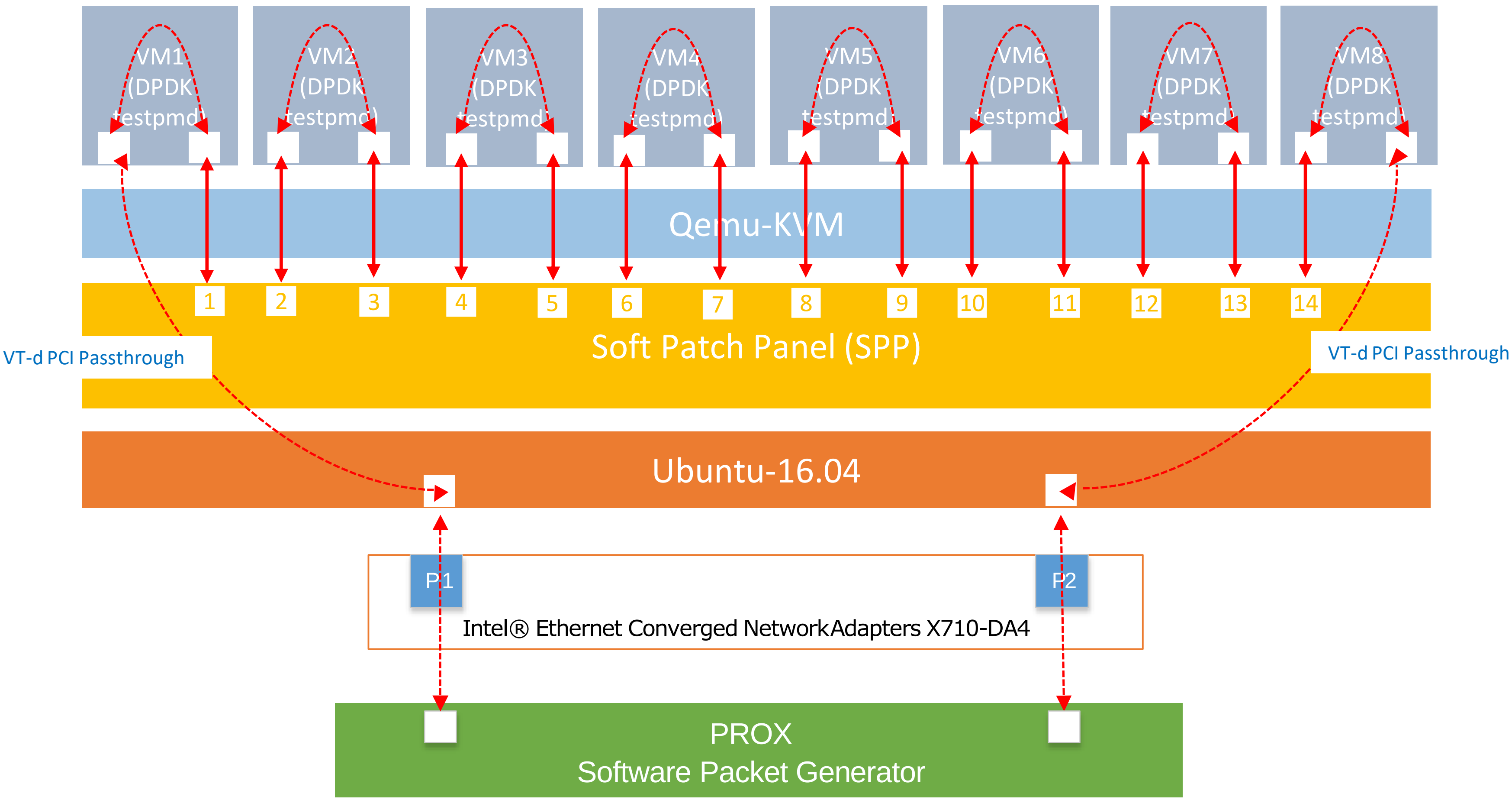
8 VMs Service Chain Test Setup Diagram (SRIOV)



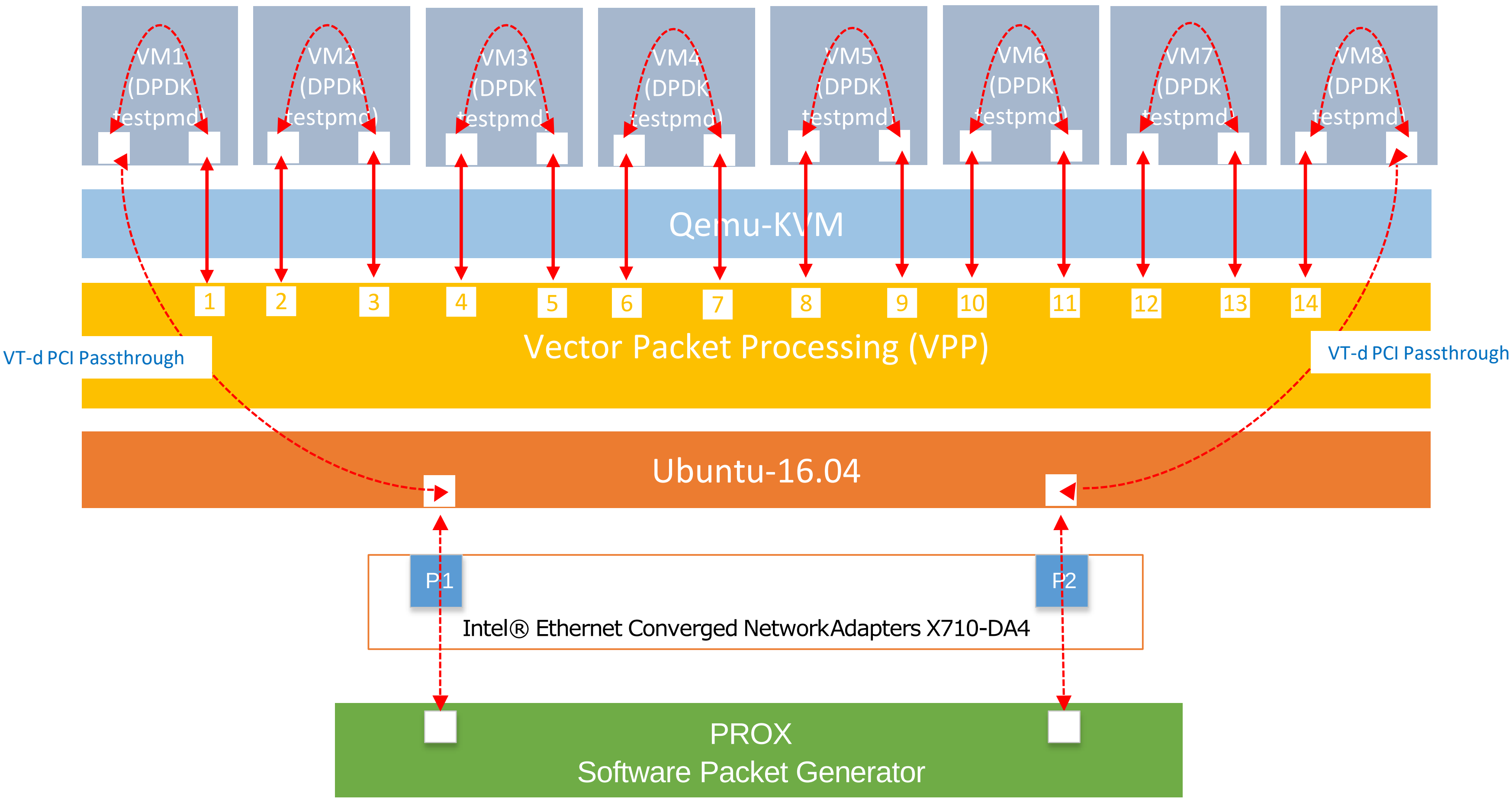
8 VMs Service Chain Test Setup Diagram (OVS-DPDK)



8 VMs Service Chain Test Setup Diagram (SPP)



8 VMs Service Chain Test Setup Diagram (VPP)



Performance Result : Summary

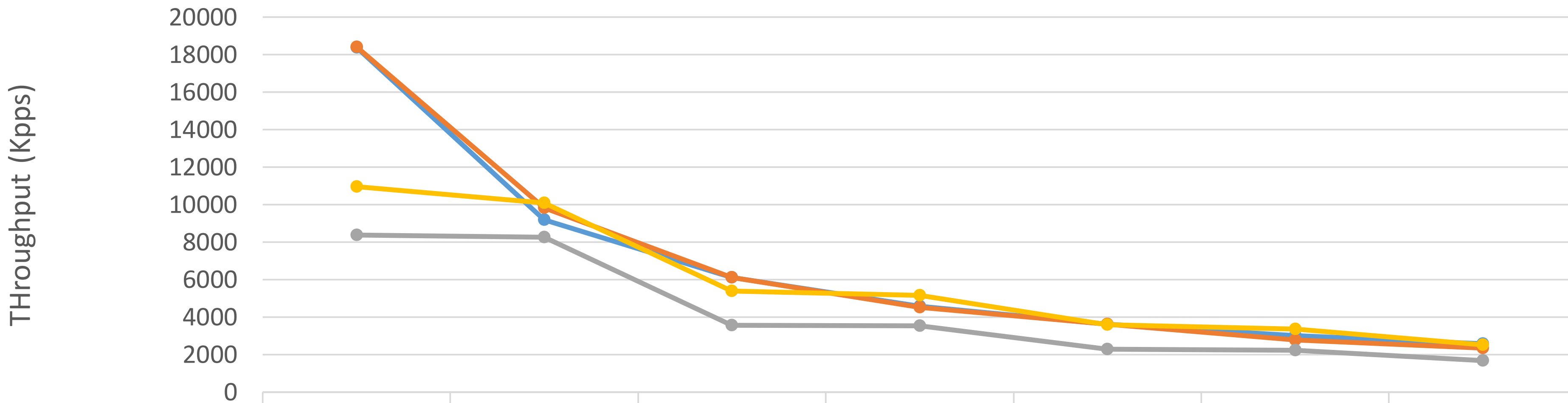
2 to 8 VMs Service Chain Performance
64 Byte, 10K Flow
(Bi-Direction, total 20G)

Binary-Search Pktgen

2 to 8 VMs Service Chain Performance:
64 Bytes, 10K Flow, Bi-Direction, Packet per second

SRIOV's performance is not good in VM chaining scenario, the bottleneck is NIC's limitation.

Phy-VM-Phy Core Scalability Performance running One Flows on
Intel® Xeon® Processor E5-2695 v4
Multiple VM Service Chain (64Byte, 10K Flow, Bi-direction)



	2 VMs	3 VMs	4 VMs	5 VMs	6 VMs	7 VMs	8 VMs
SRIOV	18370.39	9176	6106.92	4582.13	3610.25	3004.99	2580.51
SPP-vhost-4PMD	18405.19	9829.09	6117.86	4523.16	3624.33	2786.65	2343.46
VPP-vhost-4PMD	10959.9	10083.83	5393	5143.81	3581.41	3363.37	2522.53
OVS-DPDK-4PMD	8367.34	8263.44	3566.33	3537.34	2290.57	2228.31	1669.36

Number of Chaining VM

ITRI
Industrial Technology
Research Institute

Performance Result : OVS-DPDK

2 to 8 VMs Service Chain Performance
64 Byte, OVS-DPDK, 4 PMD
with Different number of Flow
(Single Direction 10G / Bi-Direction 20G)

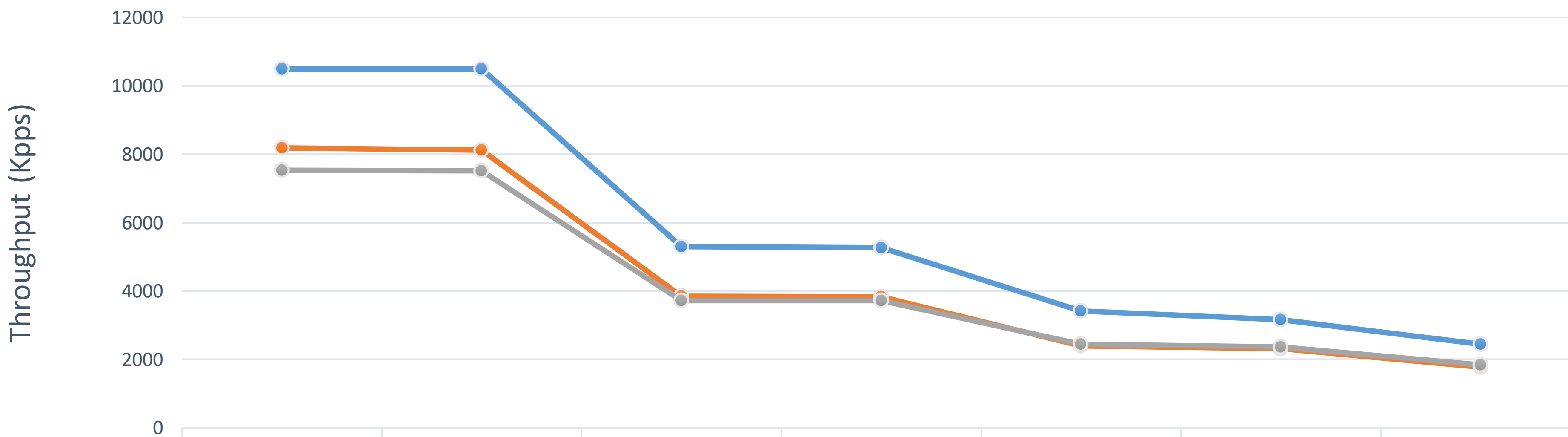
Full-Speed Pktgen

工業技術研究院
Industrial Technology
Research Institute

2 to 8 VMs Service Chain Performance: 64 Bytes, OVS-DPDK, 4 PMD, Single Direction

OVS performance will be effected by flow number, it may cause 20% performance decrease.

Phy-VM-Phy Core Scalability Performance running One Flows on
Intel® Xeon® Processor E5-2695 v4
Multiple VM Service Chain (64 Bytes, OVS-DPDK, 4 PMD, Bi-direction)



	2 VMs	3 VMs	4 VMs	5 VMs	6 VMs	7 VMs	8 VMs
1 Flow	10491	10491	5292	5264	3416	3168	2438
10K Flows	8184	8127	3844	3824	2404	2316	1782
1M flows	7523	7513	3716	3716	2448	2368	1838

Packet Size (Bytes)

NFV Performance Lab

VM/Container Performance Tuning

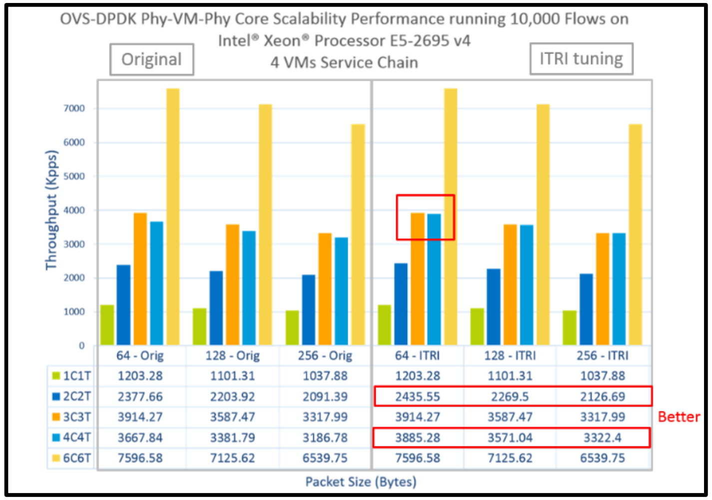
- CPU pinning
- NUMA configuration
- BIOS configuration

Data Plane Acceleration

- SRIOV passthrough (include SmartNIC)
- Enable DPDK (OVS-DPDK, VPP, SPP…etc)
- QAT, Intel AES-NI

Scenario / Use Cases

- NFVI / VNF performance characterization
- SDWAN scenario (uCPE, DPI, IPSEC)
- VM chaining with different data plane
- More use cases will be added into lab.



	64 - Orig	128 - Orig	256 - Orig	64 - ITRI	128 - ITRI	256 - ITRI
1C1T	1203.28	1101.31	1037.88	1203.28	1101.31	1037.88
2C2T	2377.66	2203.92	2091.39	2435.55	2269.5	2126.69
3C3T	3914.27	3587.47	3317.99	3914.27	3587.47	3317.99
4C4T	3667.84	3381.79	3186.78	3885.28	3571.04	3322.4
6C6T	7596.58	7125.62	6539.75	7596.58	7125.62	6539.75

ITRI
Industrial Technology Research Institute

Thanks for your attention