Datapath Composition

James Won-Ki Hong, Seyeon Jeong, Jian Li

Dept. of Computer Science & Engineering

POSTECH

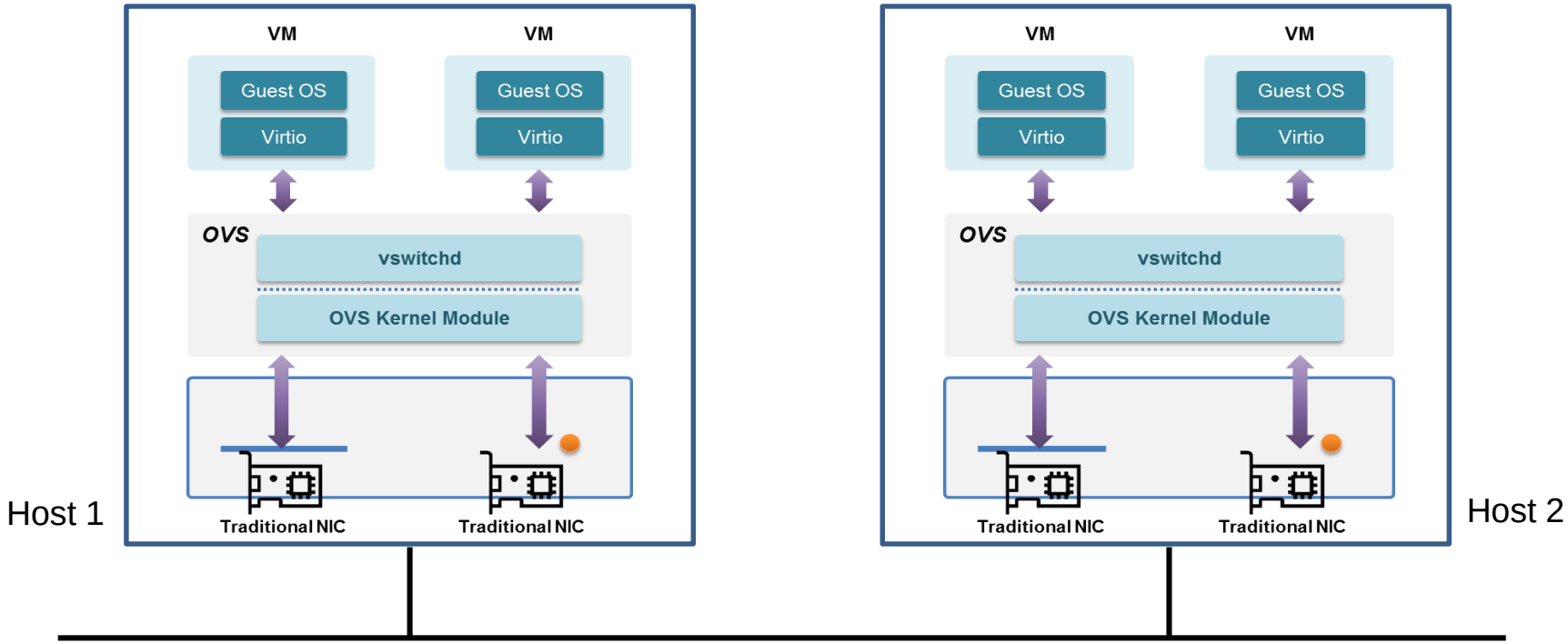
http://dpnm.postech.ac.kr/~jwkhong

jwkhong@postech.ac.kr

Datapath Composition

❖ (1) Original OVS in Host + Virtio in VM

- Uses OVS kernel datapath to deliver packets to VMs (Virtio-based) across all hosts servers



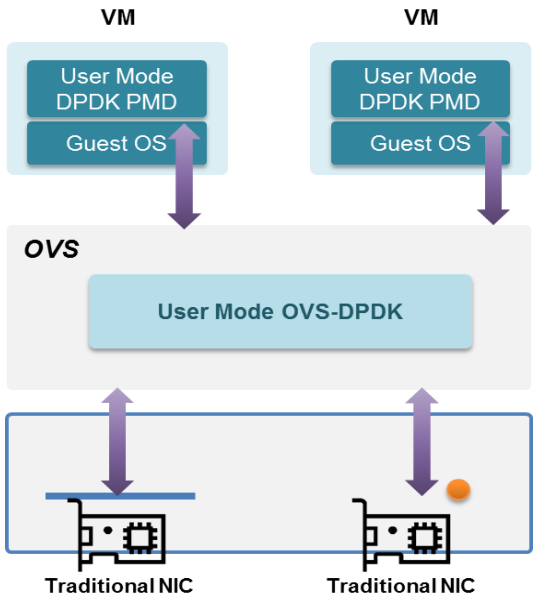
Datapath Composition

❖ (2) OVS-DPDK in Host + DPDK PMD in VM

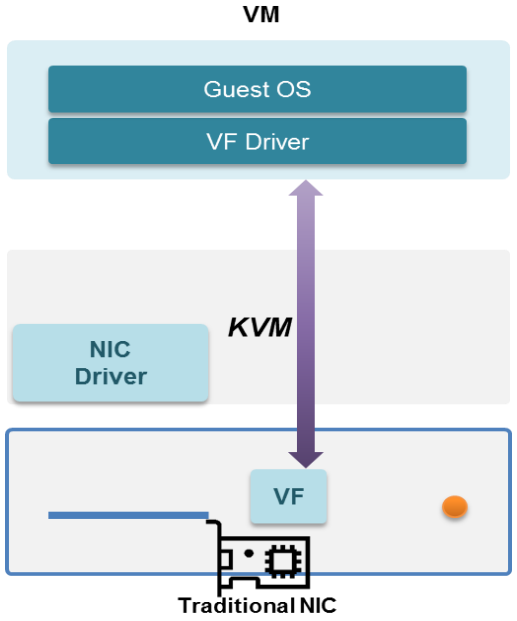
- Uses OVS-DPDK datapath in user space to deliver packets to VMs (DPDK PMD-based)

❖ (3) SR-IOV in Host + VF Driver in VM

- Uses datapath that bypasses the hypervisor to deliver packets from physical NIC to VMs (VF Driver-based)



(2)

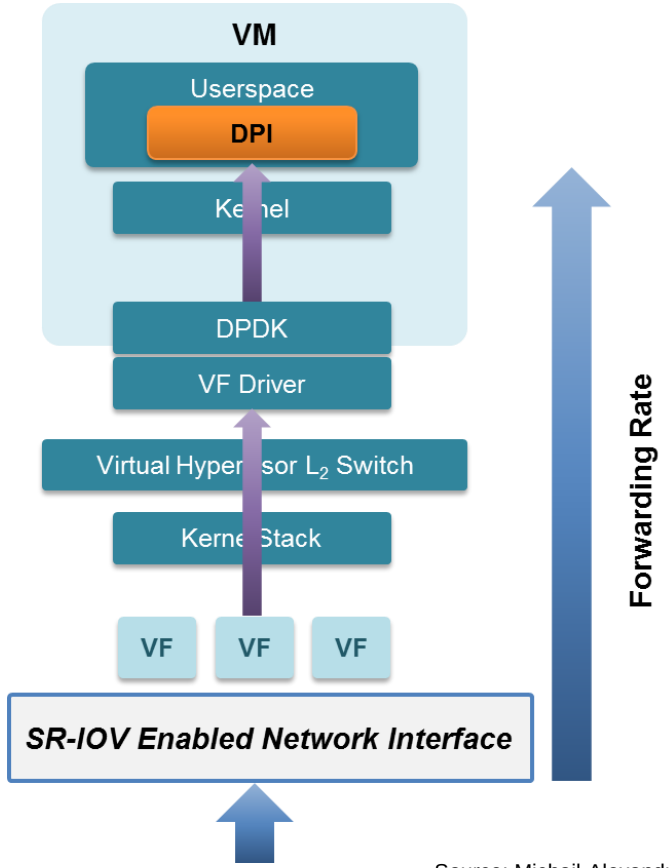


(3)

Datapath Composition

❖ (4) SR-IOV in Host + DPDK in VM

- Preserves NIC line rate in the whole stack
- Network Interface → VF → DPDK VF Driver → DPDK app.



Source: Michail-Alexandros Kourtis et al., Enhancing VNF performance by exploiting SR-IOV and DPDK packet processing acceleration, 2015 Network Function Virtualization and Software Defined Network (NFV-SDN)

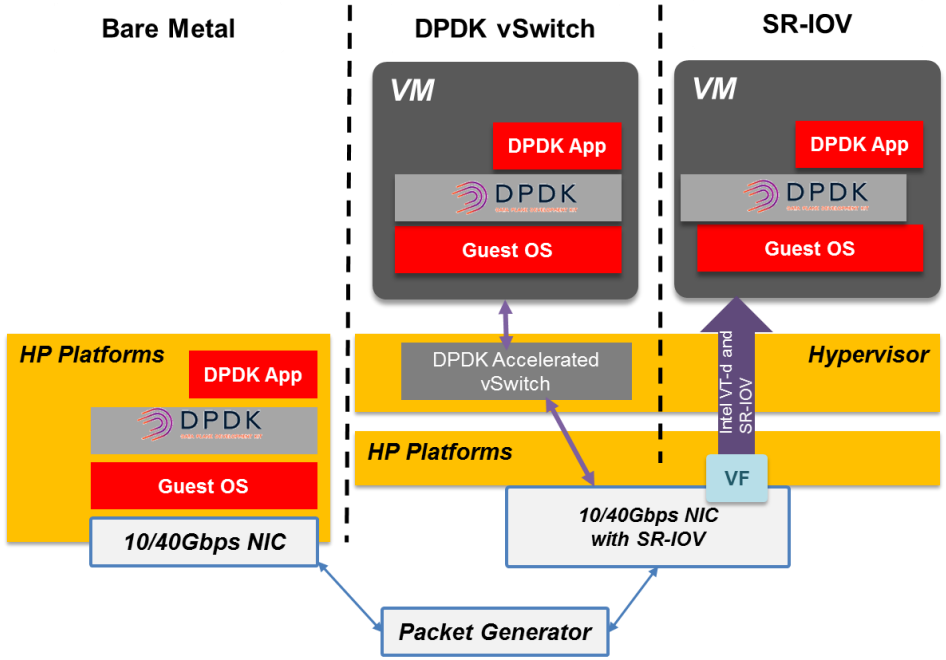
Datapath Composition

❖ Datapath Composition Comparison

	Pros.	Cons.
(1) Original OVS + Virtio	• Homogenous server deployment (reduce CAPEX/OPEX) → easy VM live migration	• Lower performance of VMs that require higher PPS (packet per sec)
(2) OVS-DPDK + PMD	• Homogenous server deployment (reduce CAPEX/OPEX) • Improved performance for VMs	• Applications need to be modified to leverage DPDK (increase OPEX) • Require to pin CPU cores to the datapath for predictable performance → Lack of host's CPU cores to spawn more VMs
(3) SR-IOV + VFD	• Performs better than case (2)	• Loss of all SDN-based network services that OVS provides → Users have to access the hardware to manage forwarding logics • VM live migration is not supported (increased OPEX)
(4) SR-IOV + DPDK	• Highest performance	• Includes all cons. of the case (3) • DPDK has to support the specific VF driver → http://dpdk.org/doc/nics

Performance Evaluation

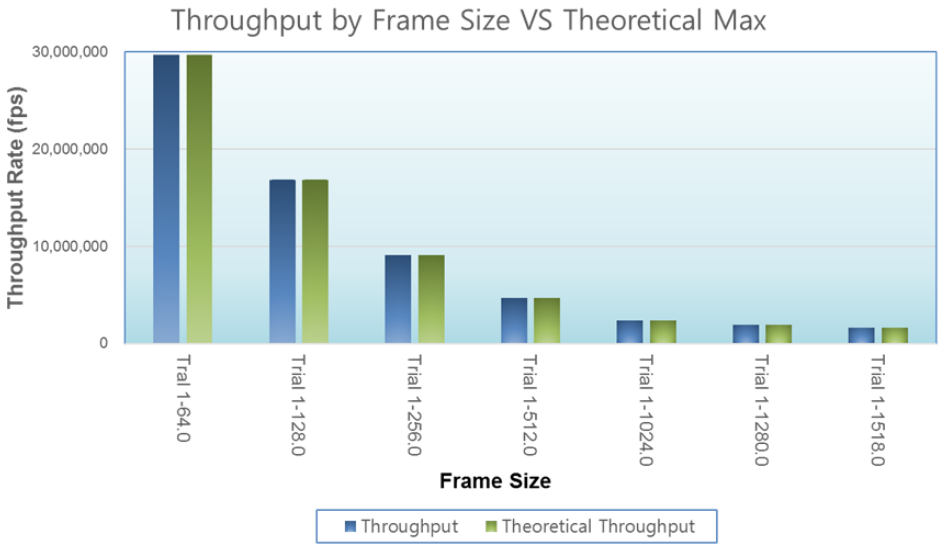
- **Evaluation and Characterization of NFV Infrastructure Solutions on Hewlett-Packard Server Platforms**
  - *Presented by Al Sanders (HP) in DPDK summit (Aug. 17, 2015)*
  - Performance comparison: Bare metal vs. SR-IOV vs. OVS-DPDK
- **Experiment Setup**
  - H/W packet generator(10/40Gbps)
    → HP Platform (Intel Xeon E5-2603 v4 6-core 1.7GHz, 16GB memory)
    → Measurement of L2 forwarding throughput in each DPDK app.

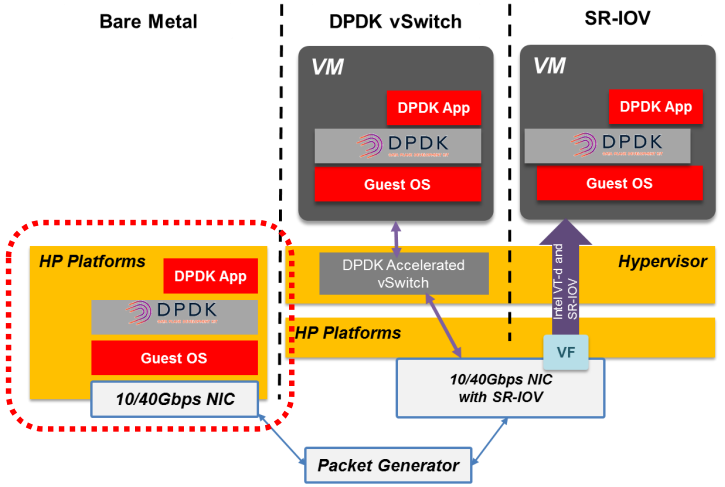


Performance Evaluation

❖ 1) Bare Metal

- Only measure RX throughput
  - No application packet processing time



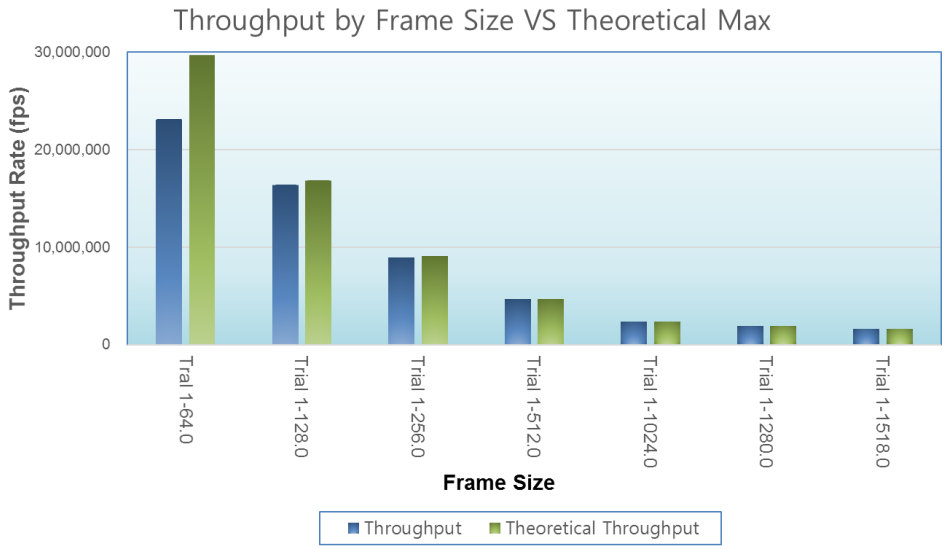


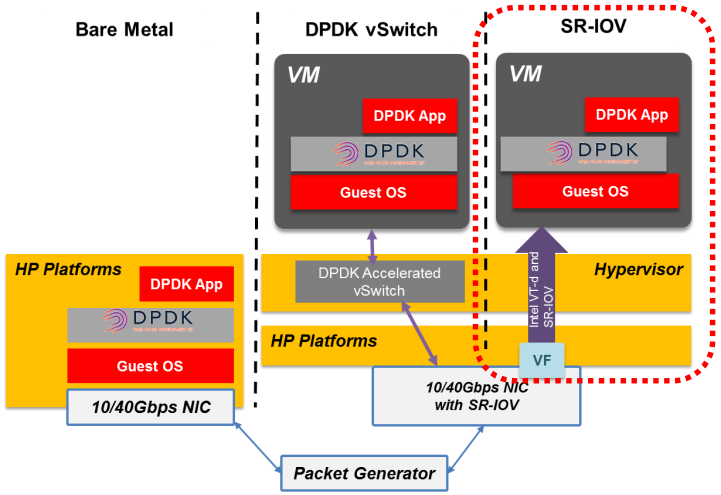
Total Trials	Number of Passed Trials	Frame size (bytes)	Intended Load(%)	Offered Load(%)	Throughput(%)	Aggregated Throughput(fps)	Aggregated Theoretical Max (fps)	Aggregated Throughput (Mbps)
1	1	64	100	100	100	29761904.767	29761904.762	20000
1	1	128	100	100	100	16891891.900	16891891.892	20000
1	1	256	100	100	100	9057971.017	9057971.014	20000
1	1	512	100	100	100	4699248.133	4699248.120	20000
1	1	1024	100	100	100	2394636.017	2394636.015	20000
1	1	1280	100	100	100	1923076.933	1923076.923	20000
1	1	1518	100	99.87	99.87	1623376.633	1625487.646	19974.026

Performance Evaluation

❖ 2) SR-IOV (with DPDK)

- Nearly identical to the bare metal performance (Line rate for > 64 bytes packets)
- Need to set huge pages (1GB) for VM to reduce loss of small size packets



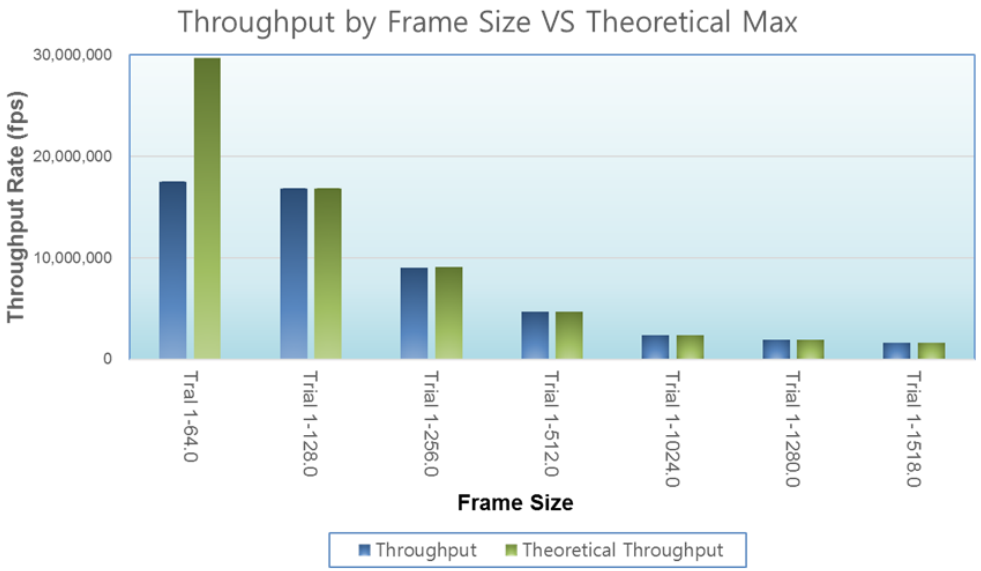


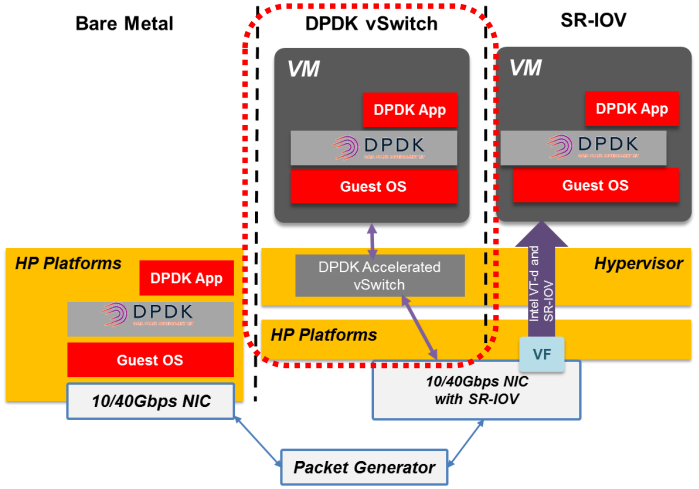
Total Trials	Number of Passed Trials	Frame size (bytes)	Intended Load(%)	Offered Load(%)	Throughput(%)	Aggregated Throughput(fps)	Aggregated Theoretical Max (fps)	Aggregated Throughput (Mbps)
1	1	64	79.688	77.778	77.778	23148148	29761904.762	15555.555
1	1	128	99.219	97.368	97.368	16447368	16891891.892	19473.684
1	1	256	99.219	98.571	98.571	8928570	9057971.014	19714.283
1	1	512	99.219	99.254	99.254	4664178	4699248.120	19850.742
1	1	1024	99.219	99.239	99.239	2376424	2394636.015	19847.893
1	1	1280	99.219	99.085	99.085	1905486	1923076.923	19817.054
1	1	1518	100	99.870	99.870	1623376	1625487.646	19974.018

Performance Evaluation

❖ 3) OVS-DPDK

- Performance on 64 bytes packets is significantly lower
- Near line rate for other packet sizes





Total Trials	Number of Passed Trials	Frame size (bytes)	Intended Load(%)	Offered Load(%)	Throughput(%)	Aggregated Throughput(fps)	Aggregated Theoretical Max (fps)	Aggregated Throughput (Mbps)
1	1	64	59.727	58.921	58.921	17536013.067	29761904.762	11784.201
1	1	128	99.922	99.688	99.688	16839228.173	16891891.892	19937.676
1	1	256	99.922	99.688	99.688	9029731.053	9057971.014	19937.676
1	1	512	99.922	99.688	99.688	4684597.313	4699248.120	19937.676
1	1	1024	99.922	99.688	99.688	2387170.280	2394363.015	19937.676
1	1	1280	99.922	99.693	99.693	1917177.920	1923076.923	19938.650
1	1	1518	100	99.870	99.870	1623376.627	1625487.646	19974.026

Performance Evaluation

❖ Conclusion

- OVS-DPDK achieved 90% of the performance seen in SR-IOV
  - But it can provide SDN-based flexibility
  - New releases keep showing performance improvement
    - OVS 2.7.x + DPDK 16.11.x
- Refer to Intel performance test reports on OVS-DPDK for more details
  - https://download.01.org/packet-processing/

All Tests	Bare Metal	SR-IOV	OVS-DPDK
Packet Size (Bytes)	Throughput (Gbps)	Throughput (Gbps)	Throughput (Gbps)
64	20	15.55	11.78
128	20	19.47	19.93
256	20	19.71	19.93
512	20	19.85	19.93
1024	20	19.84	19.93
1280	20	19.81	19.93
1518	19.97	19.97	19.97