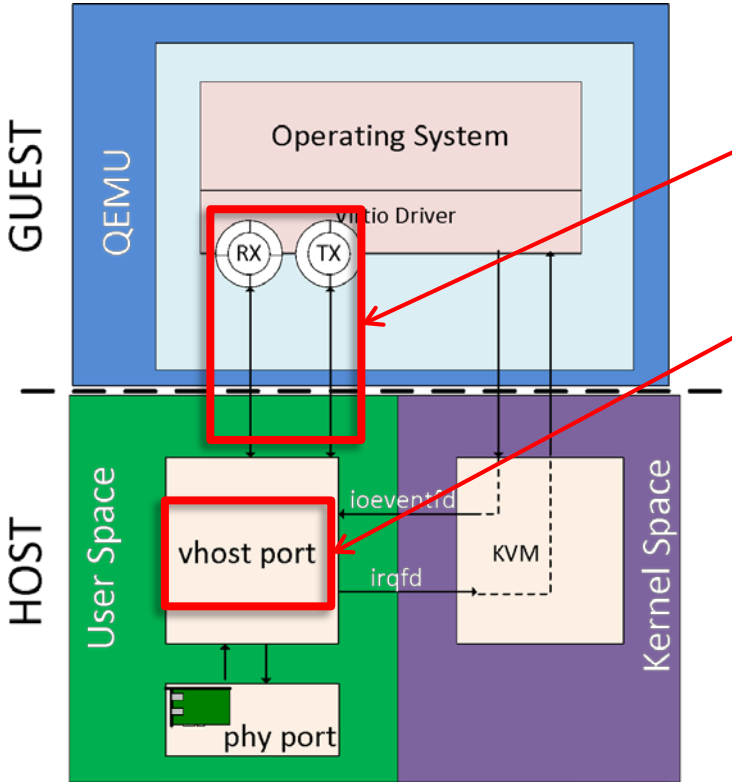
Topic: Accelerate virtio/vhost using DPDK in NFV/Cloud Environment

Company: Intel
Title: Software Engineer
Name: Xie, huawei; Tan, Jianfeng

Agenda

- virtio/vhost background
- virtio in NFV/Cloud (challenges, solutions)
  1. virtio PMD optimization
  2. vhost TSO
  3. vhost reconnect
  4. VM2VM fastpath
- virtio in container

virtio/vhost background



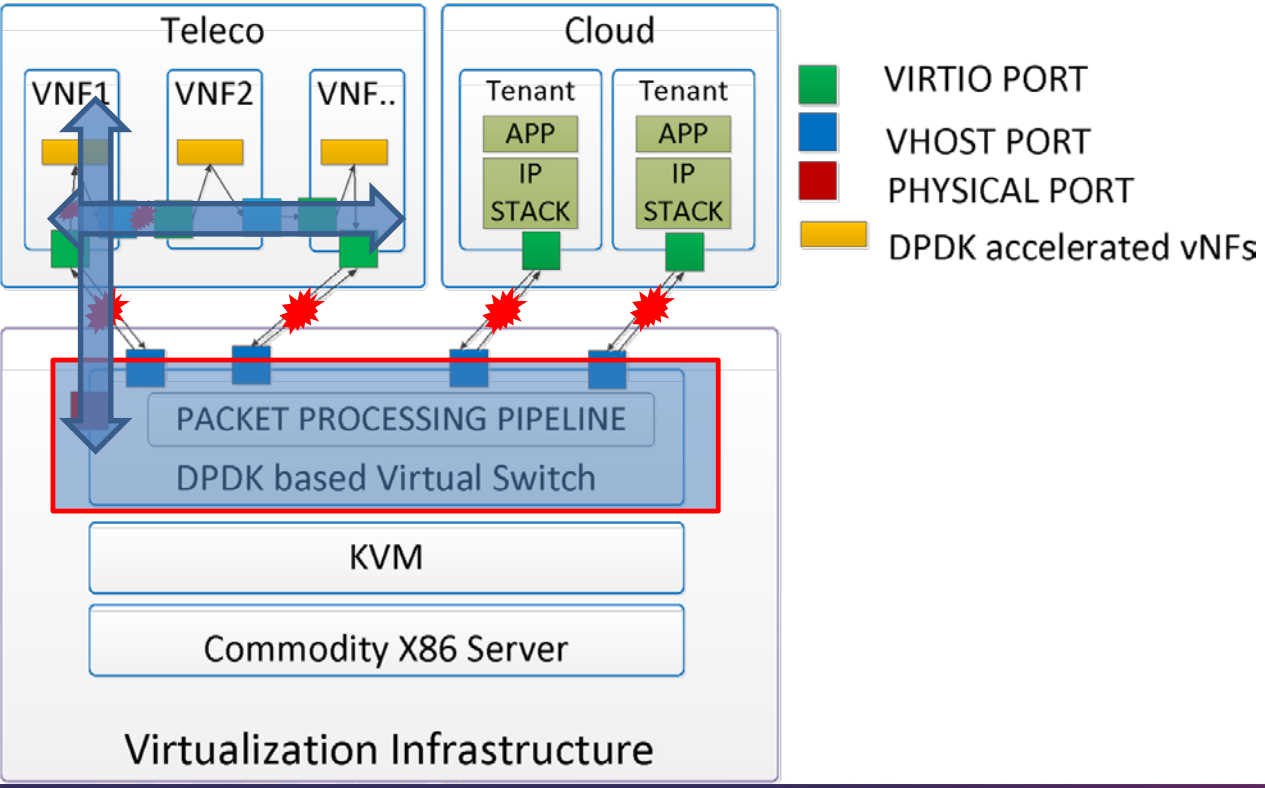
Virtio is the defacto para-virtualization standard for communicating with Virtual Machines (VM) efficiently.

Vhost is the KVM backend for Virtio, supplying packets to a Virtio Frontend.

Packet Flow

A bridge/virtual switch, switches packets to the backend (vhost) and these are forwarded to the frontend (virtio) in the Guest.

virtio in NFV/Cloud



North2South Perf

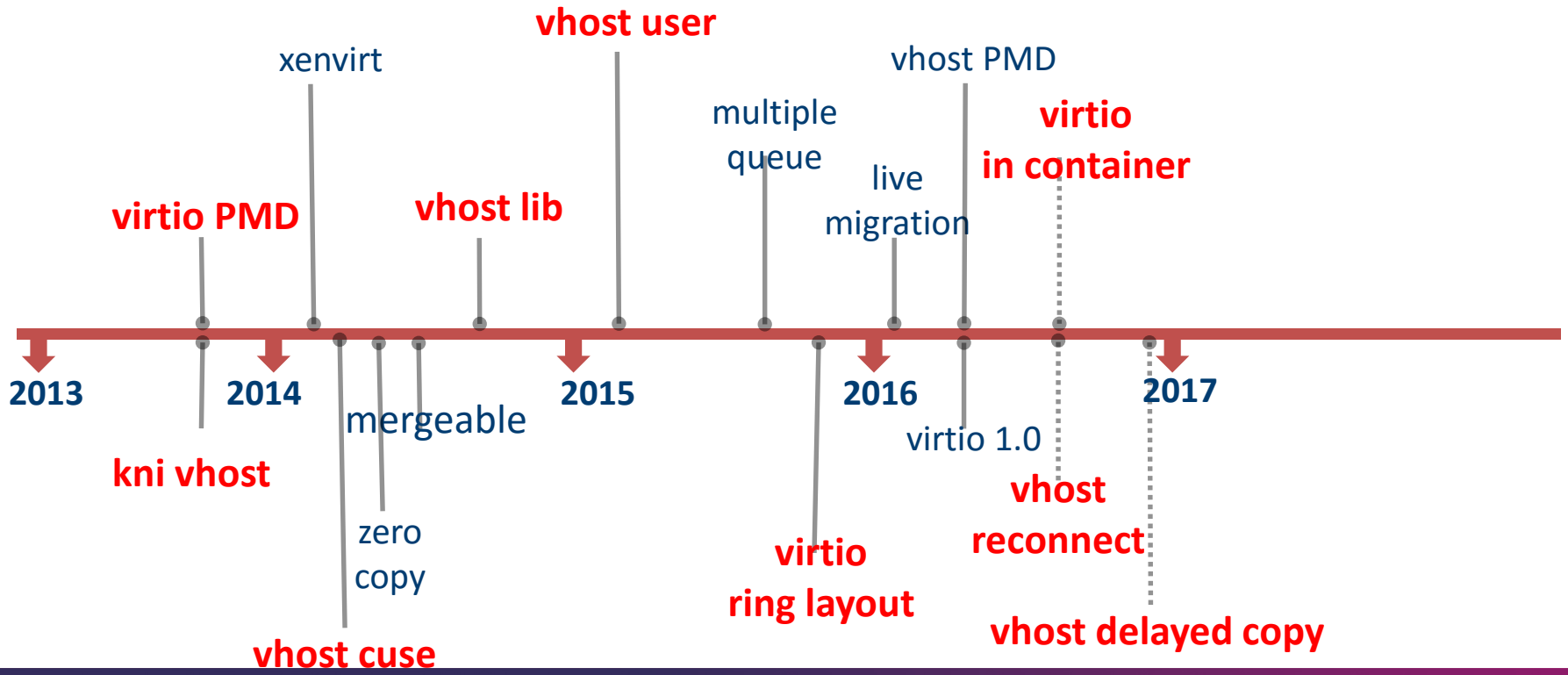
- virtio PMD optimization
- vhost AVX, delayed copy
- vhost TSO

East2West Perf :VM2VM

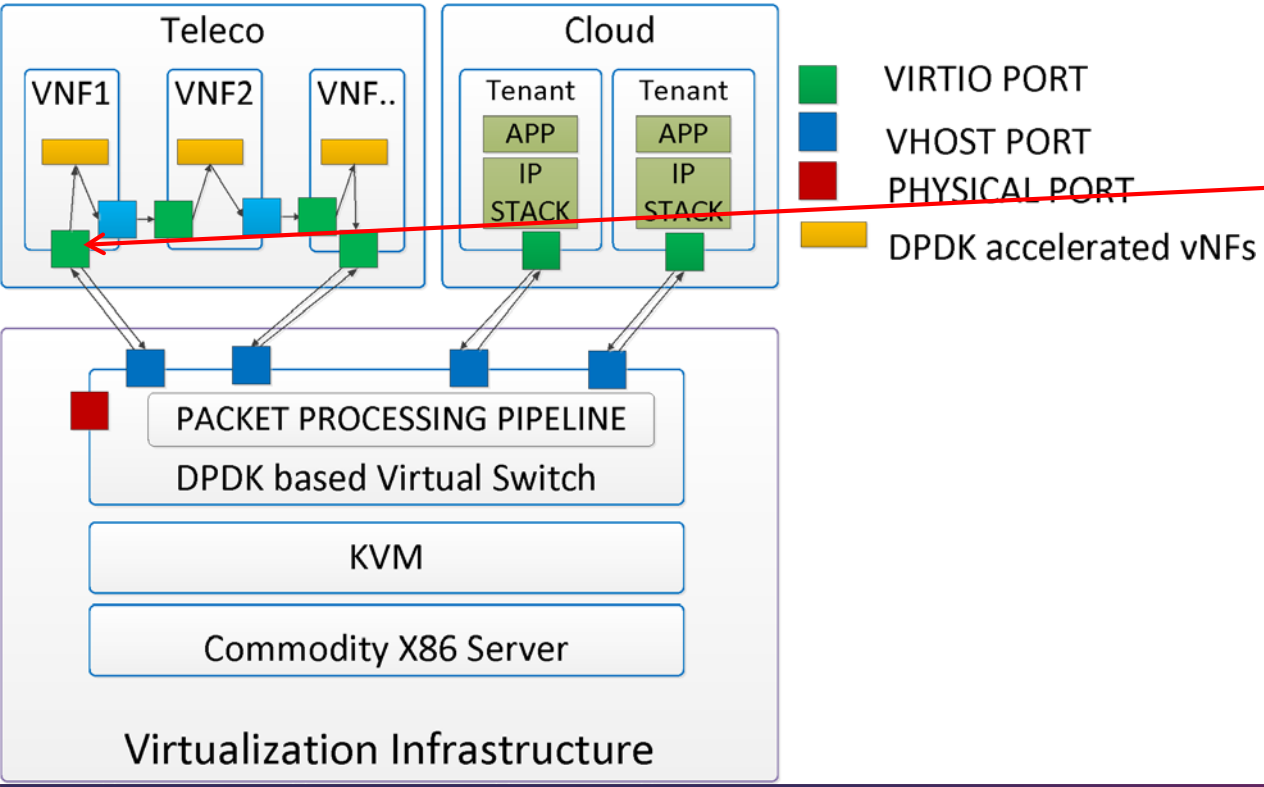
Stability

- vhost reconnect

DPDK virtio development journey



virtio in NFV/Cloud



North2South Perf

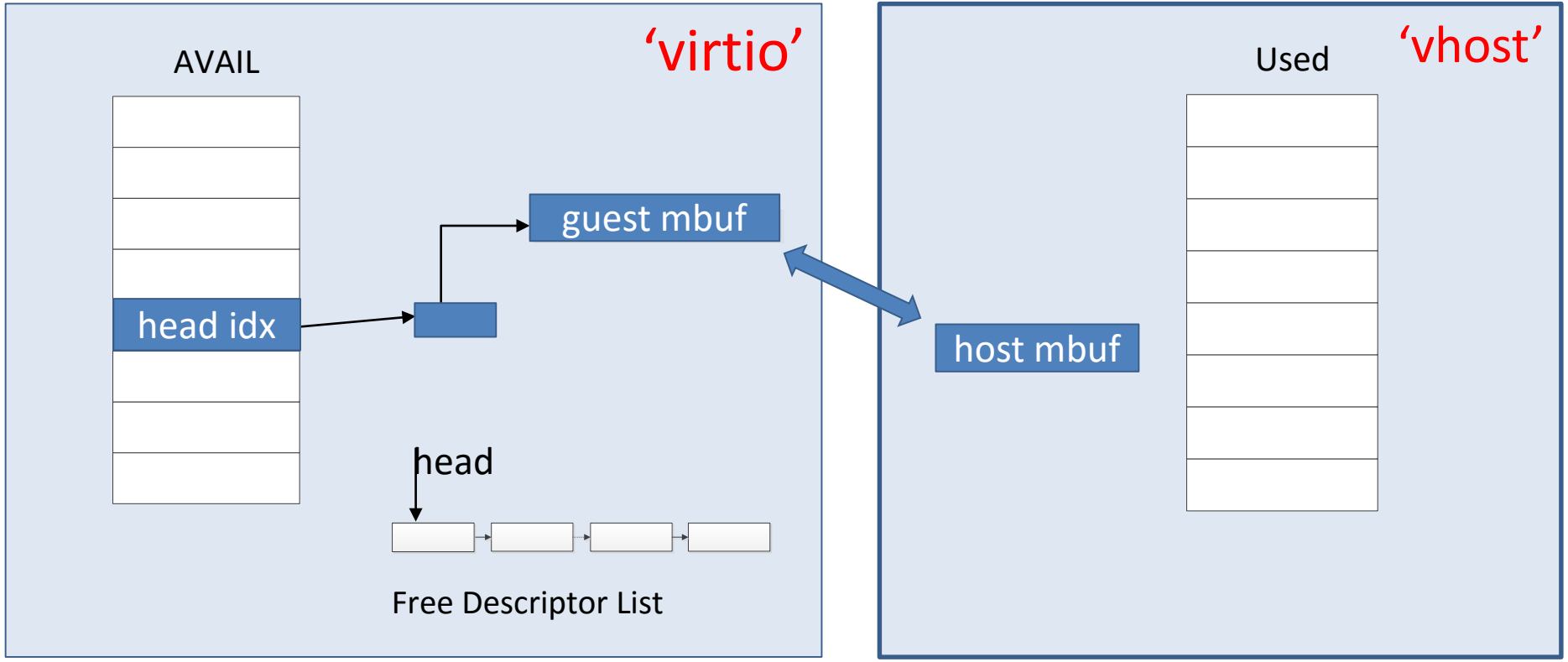
- virtio PMD optimization
- vhost AVX, delayed copy
- vhost TSO

East2West Perf :VM2VM

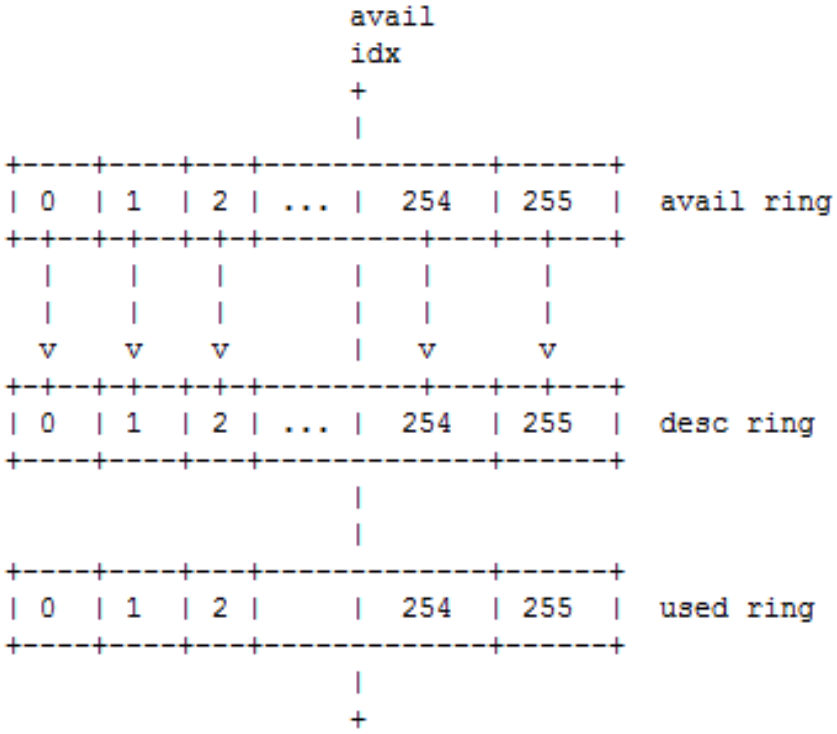
Stability

- vhost reconnect

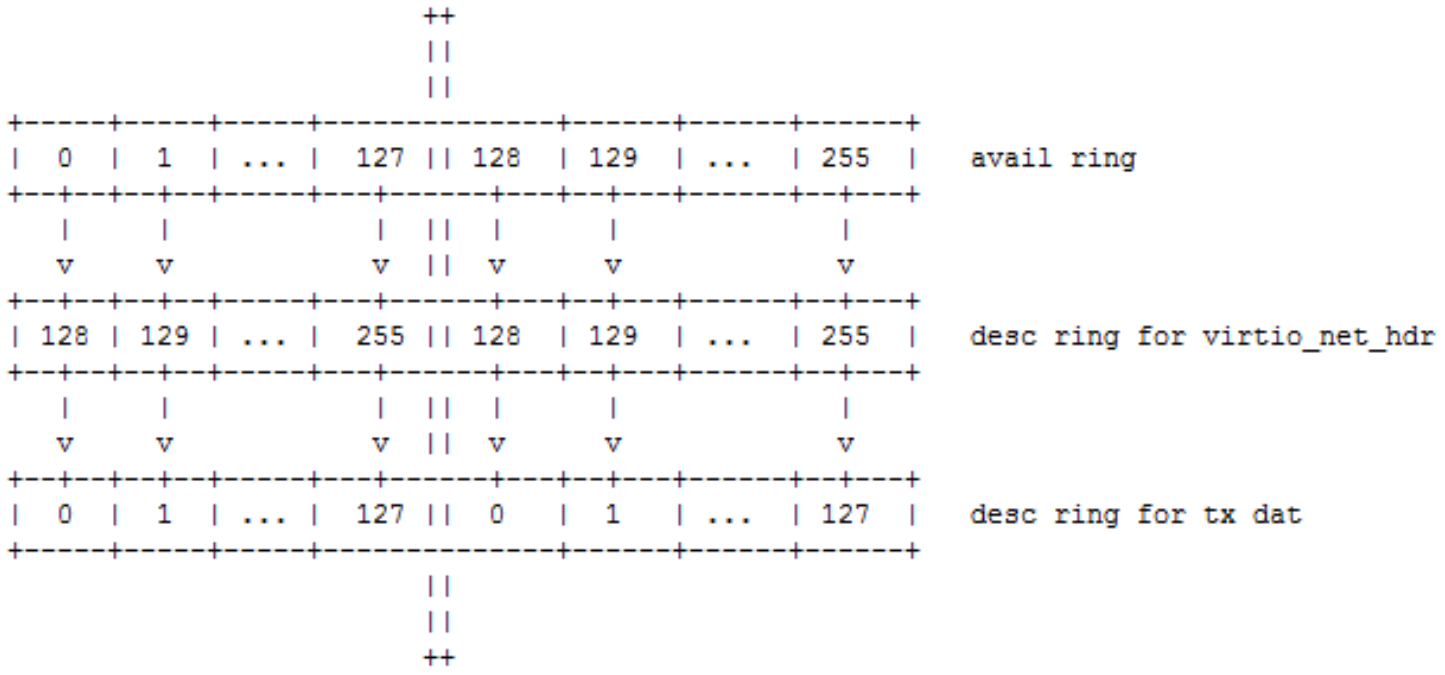
normal virtio process



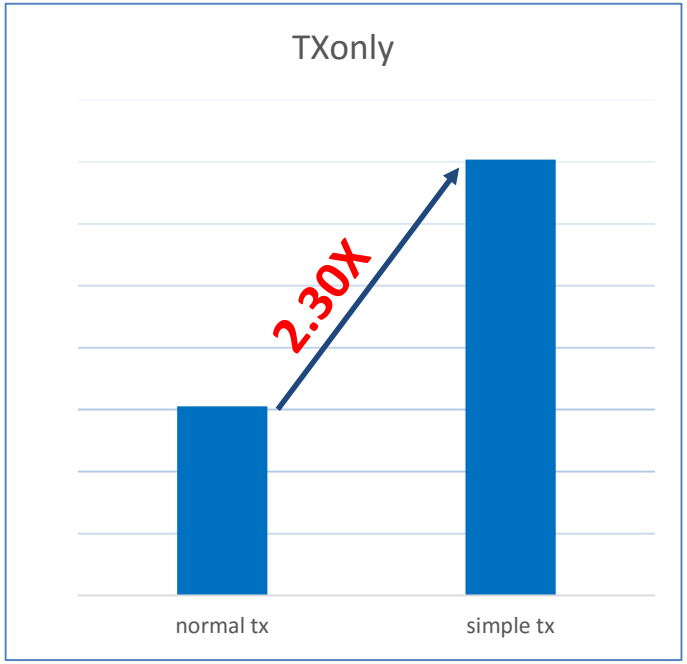
RX ring layout optmization

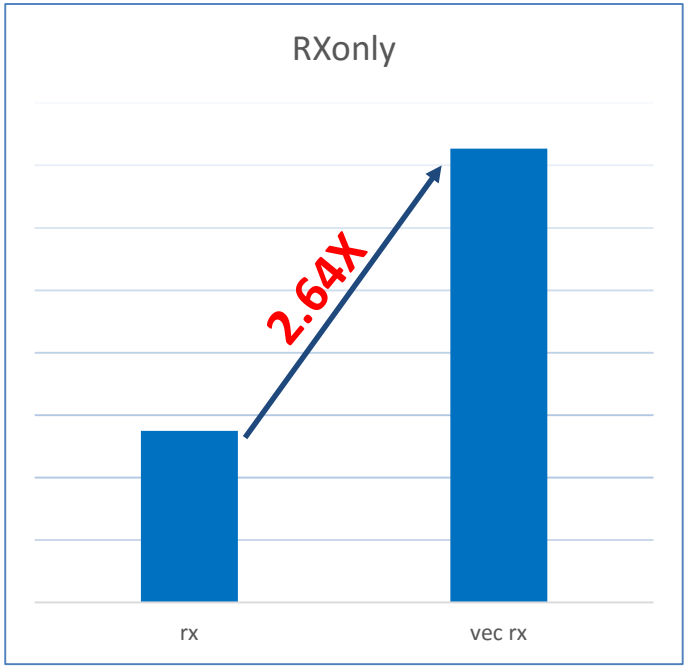


TX ring layout optmization



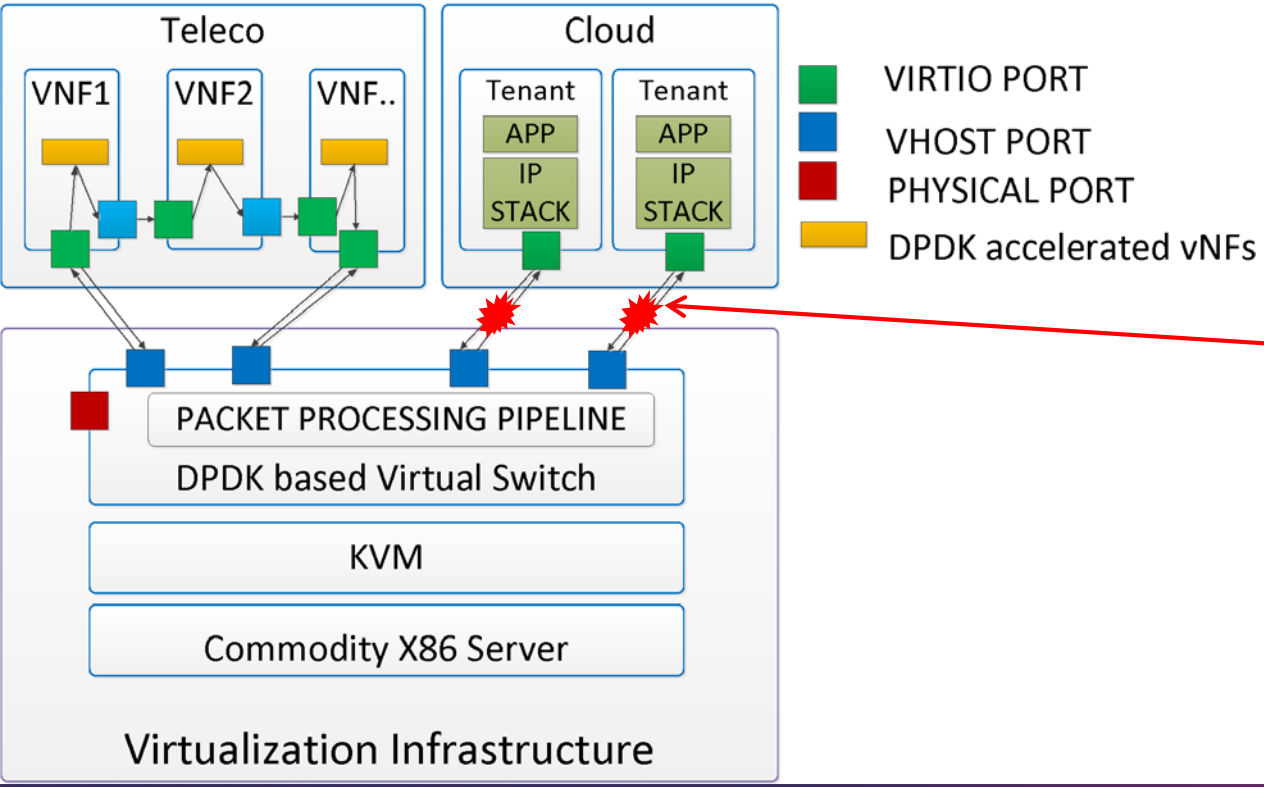
ring layout opt. and vectorization





new ring layout ?

virtio in NFV/Cloud



North2South Perf

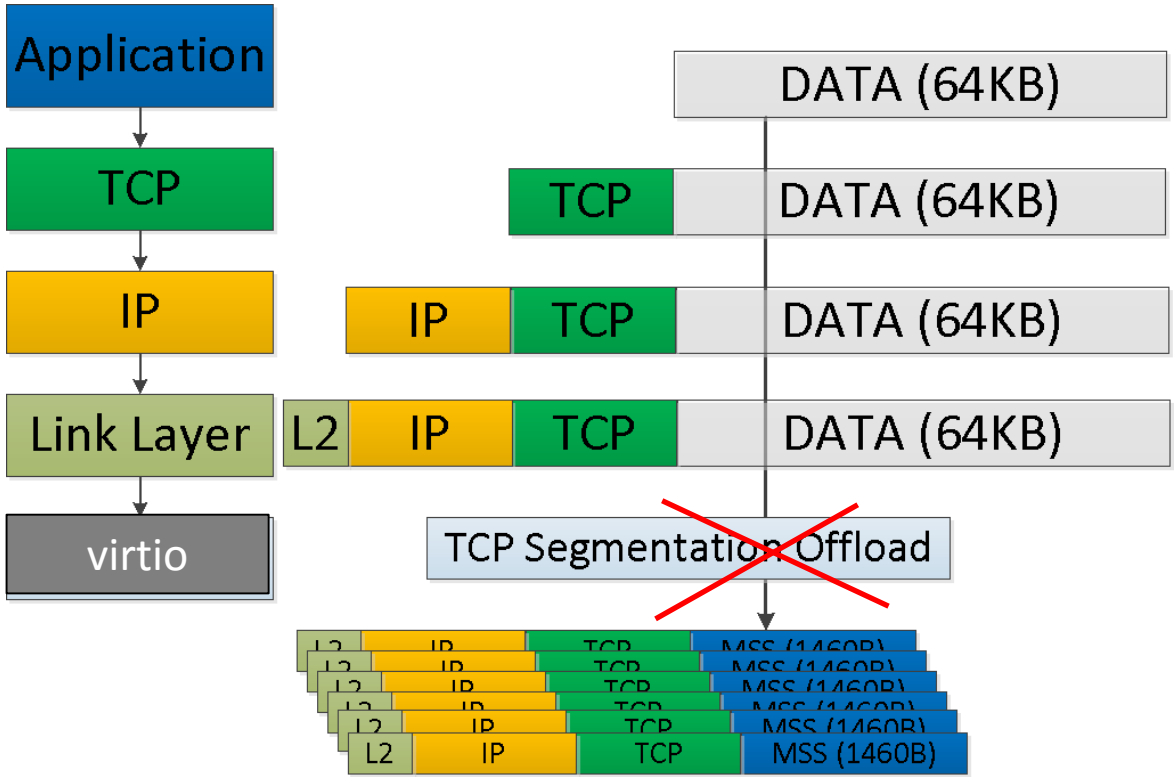
- virtio PMD and optimization
- vhost AVX, delayed copy
- vhost TSO

East2West Perf :VM2VM

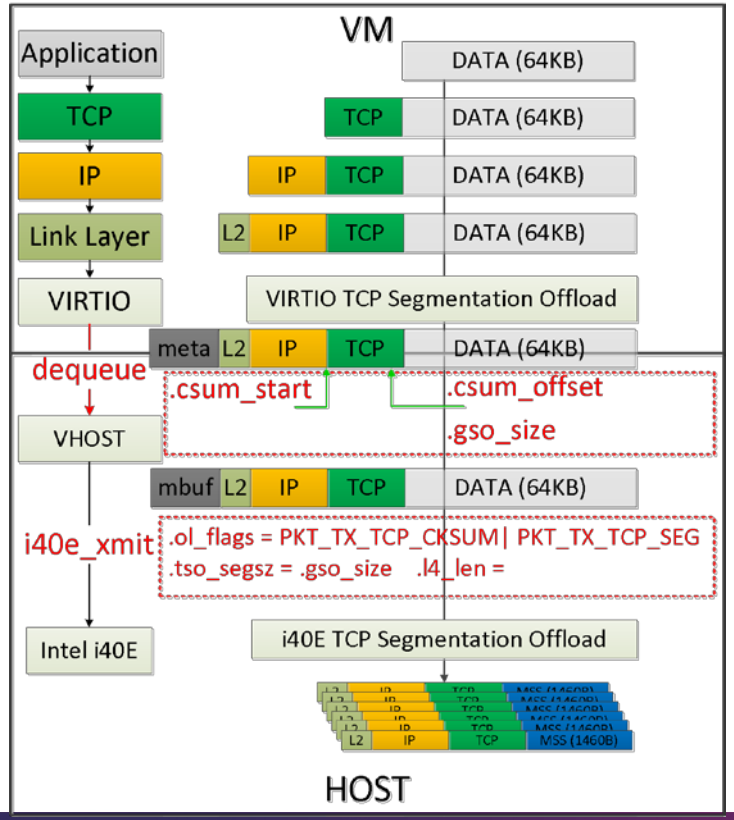
Stability

- vhost reconnect

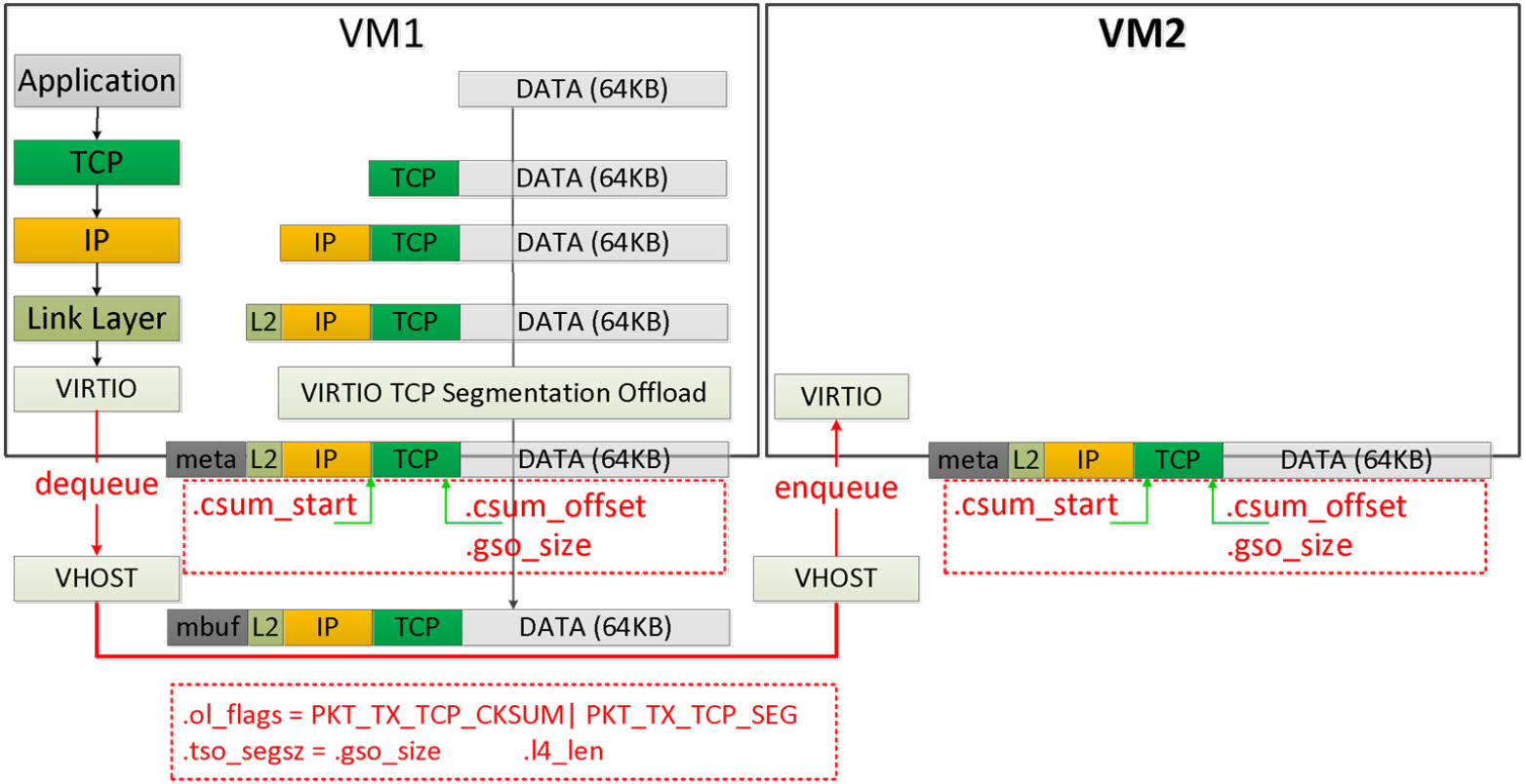
VHOST TSO


Application
TCP
IP
Link Layer
virtio
DATA (64KB)
TCP
DATA (64KB)
IP
TCP
DATA (64KB)
L2
IP
TCP
DATA (64KB)
TCP Segmentation Offload
L2
IP
TCP
MSS (1460B)

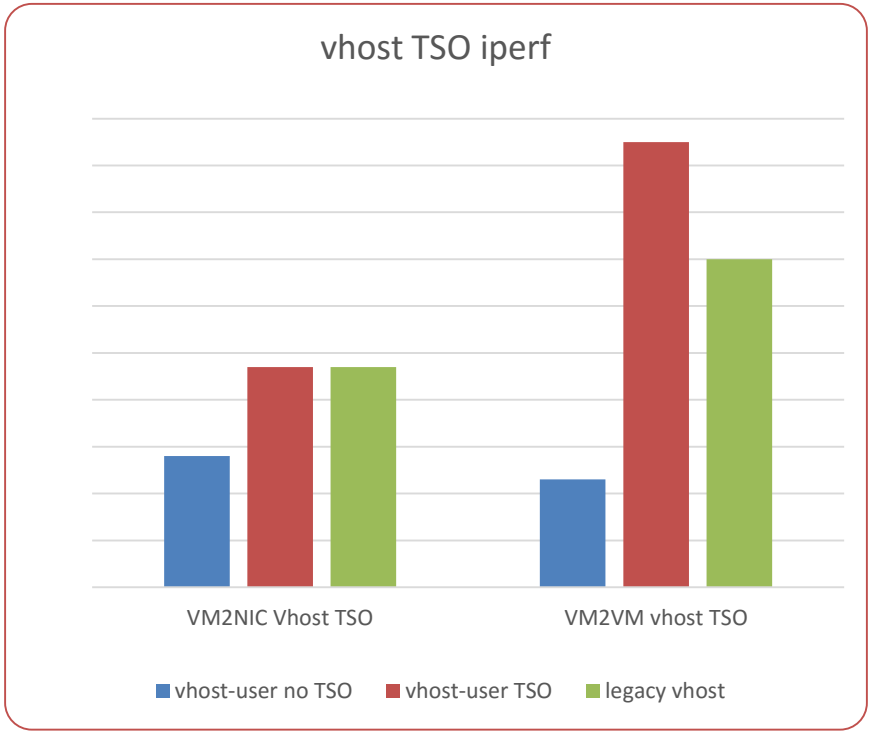
VHOST TSO To NIC



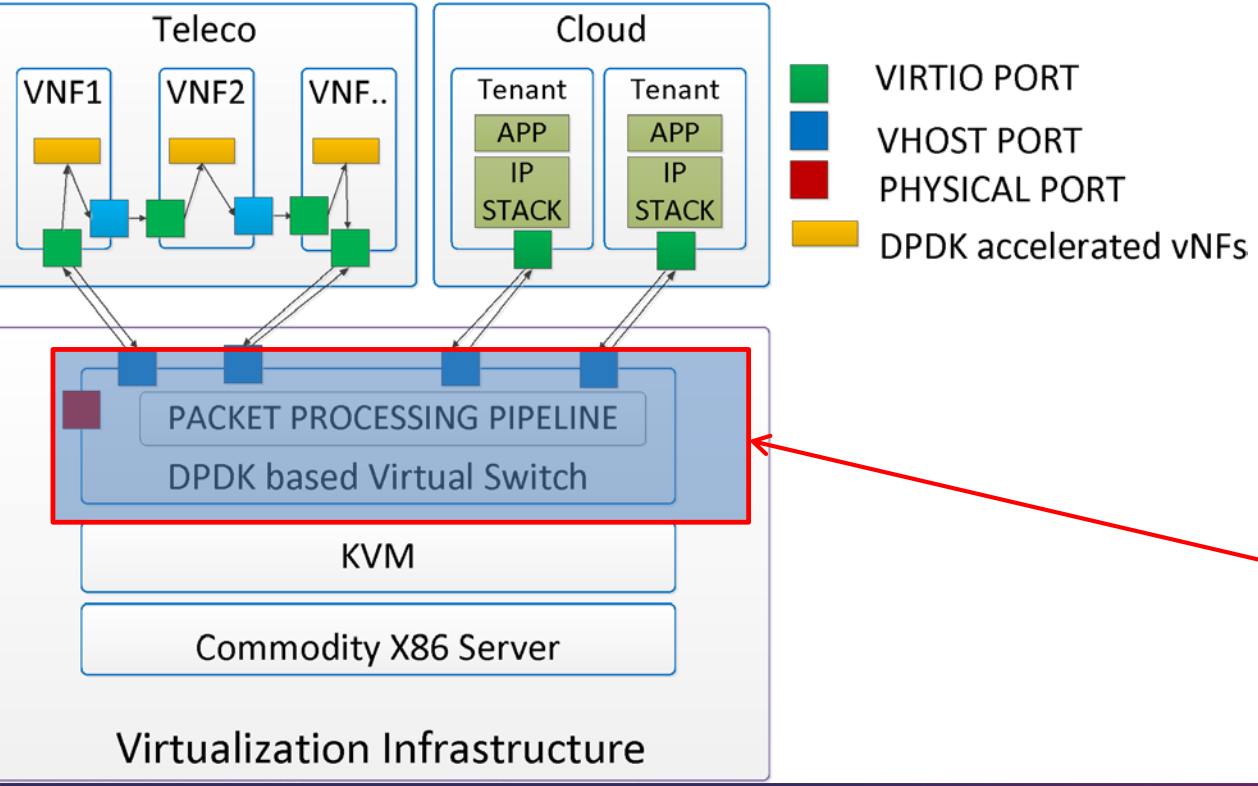
vhost TSO in VM2VM



vhost TSO performance



virtio in NFV/Cloud



North2South Perf

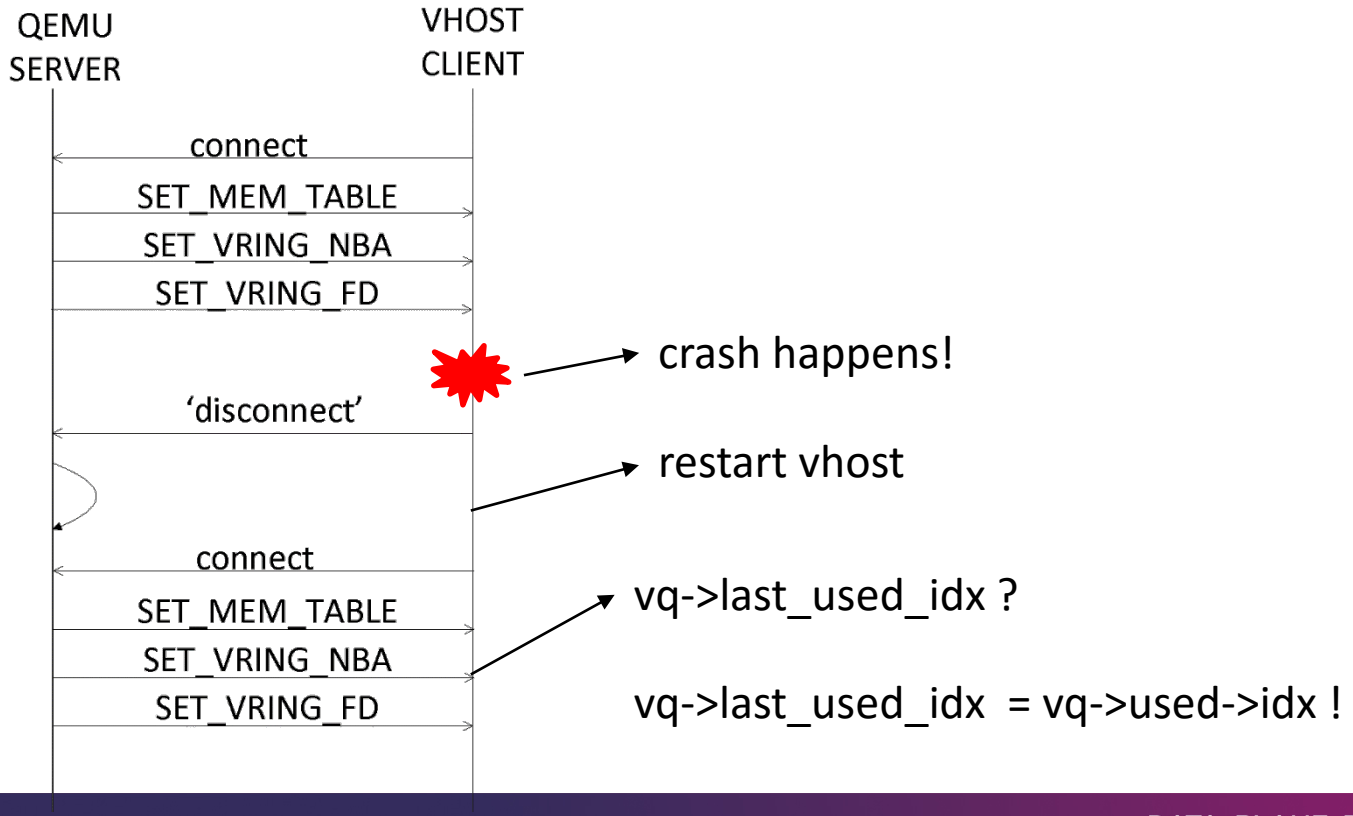
- virtio PMD and optimization
- vhost AVX, delayed copy
- vhost TSO

East2West Perf :VM2VM

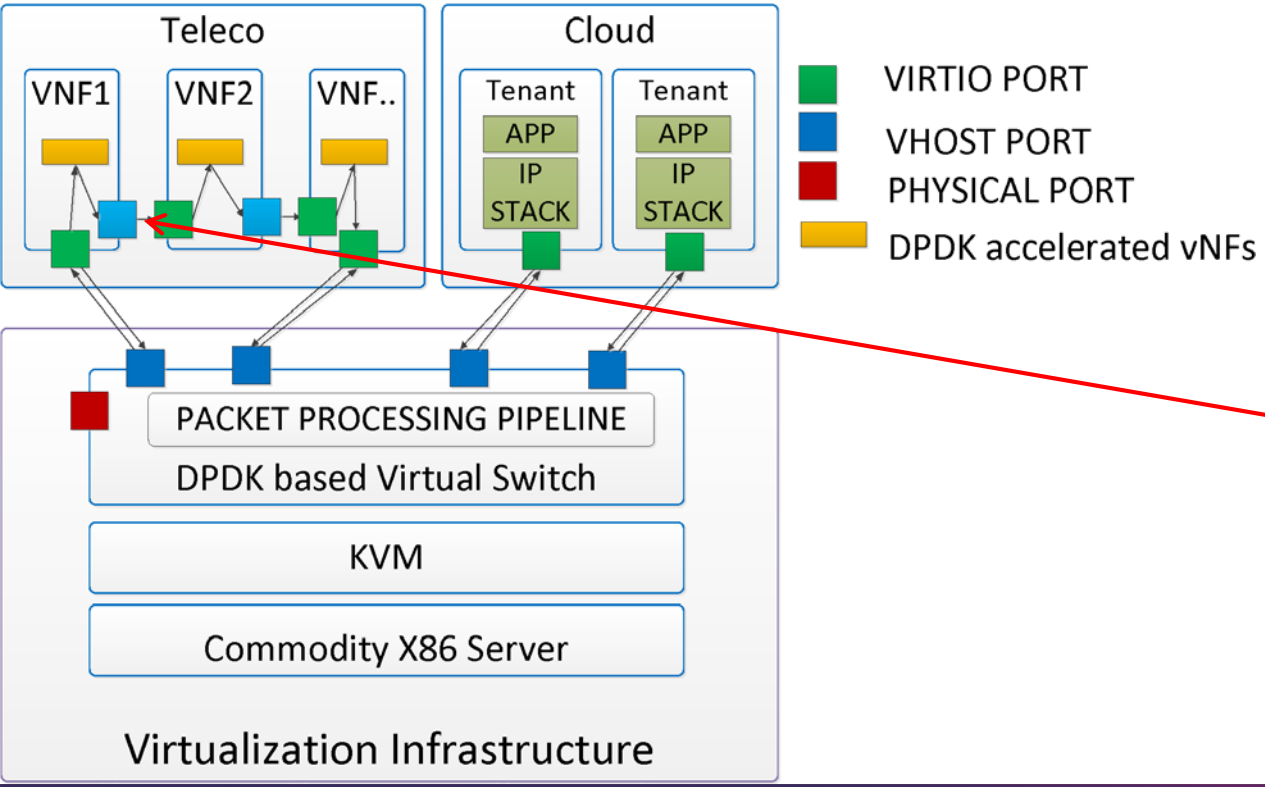
Stability

- vhost reconnect

vhost reconnect



virtio in NFV/Cloud



North2South Perf

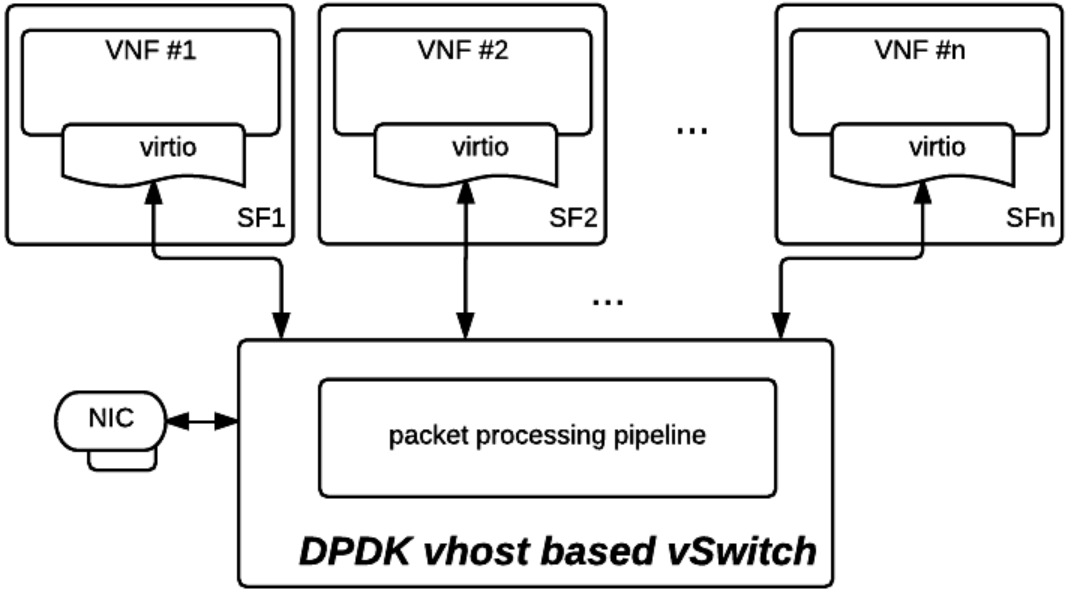
- virtio PMD and optimization
- vhost AVX, delayed copy
- vhost TSO

East2West Perf :VM2VM

Stability

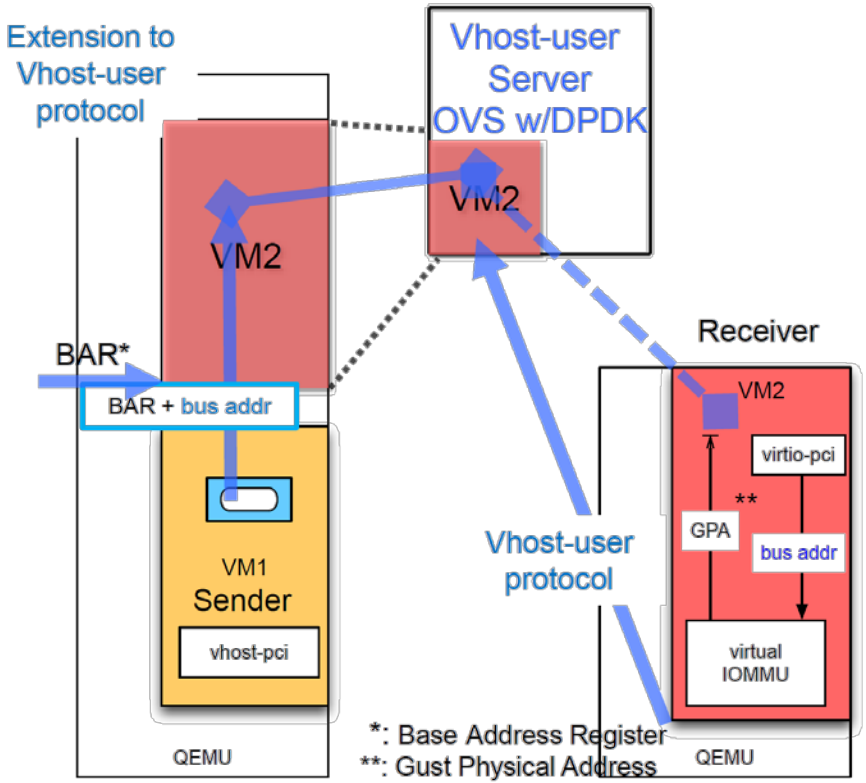
- vhost reconnect

VM2VM fastpath(WIP)

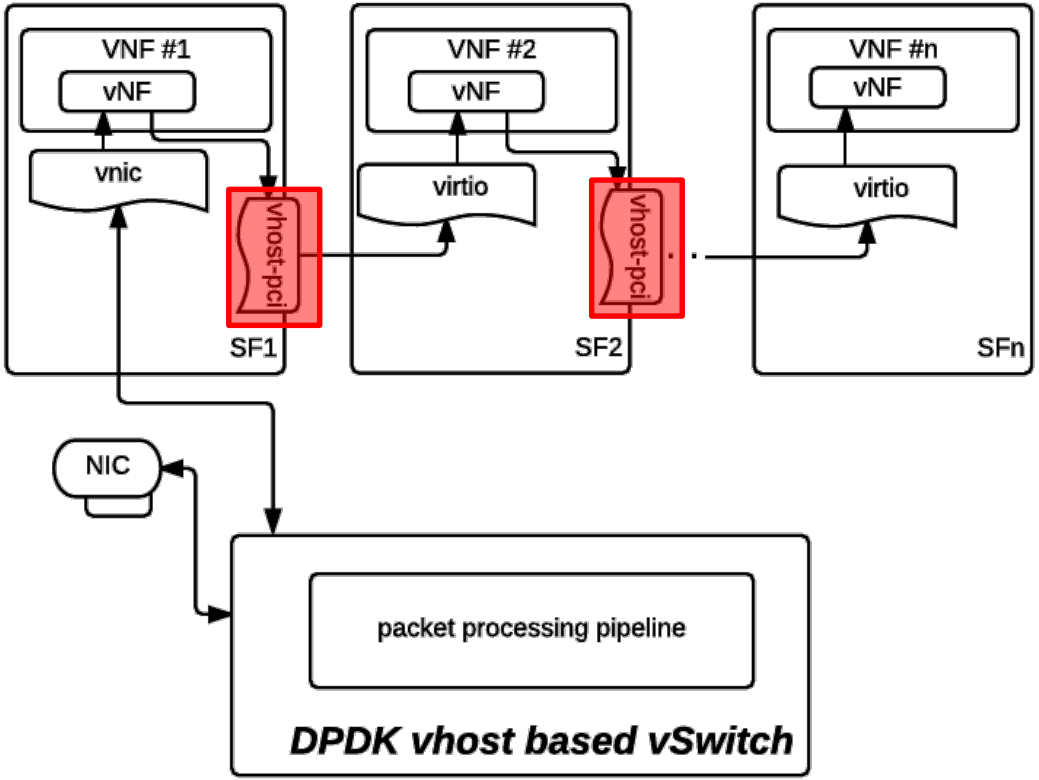


http://www.linux-kvm.org/images/8/87/02x09-Aspen-Jun_Nakajima-KVM_as_the_NFV_Hypervisor.pdf
http://schd.ws/hosted_files/ons2016/36/Nakajima_and_Ergin_PreSwitch_final.pdf

VM2VM fastpath(WIP)



VM2VM fastpath(WIP)



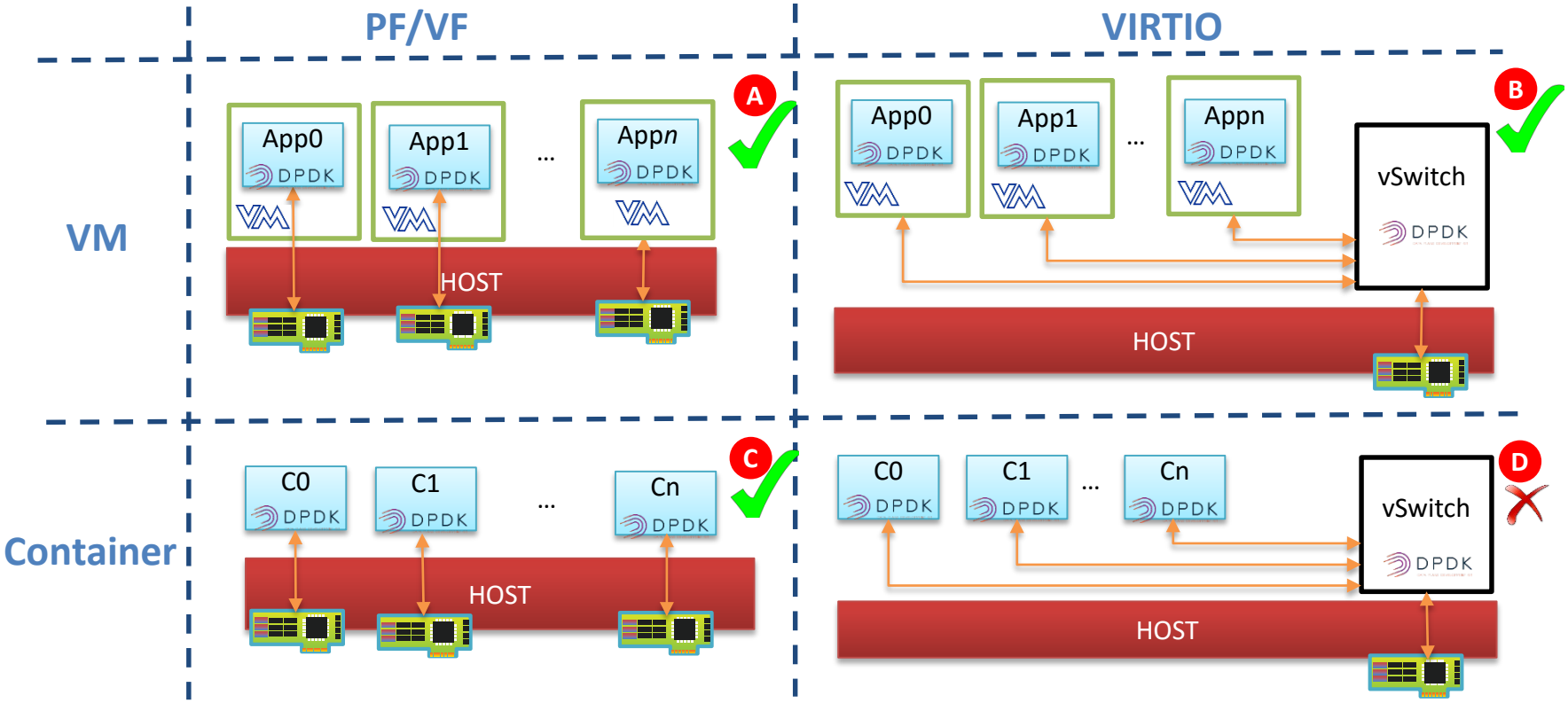
Future work

- new ISA
- vhost delayed copy
- vhost AVX
- vhost FPGA

virtio for container - Motivations

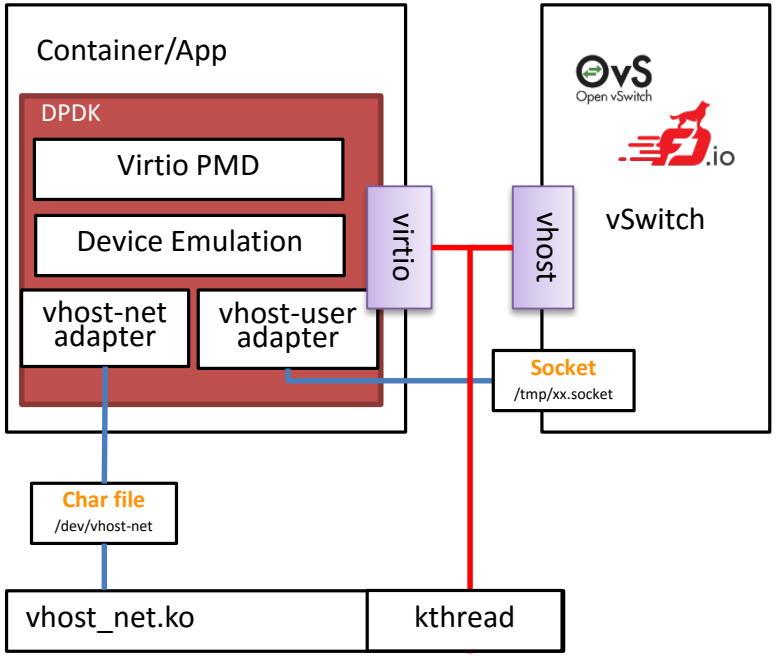
- Requirements for Container-based NFV
  - high throughput
  - low latency

virtio for container - Status quo

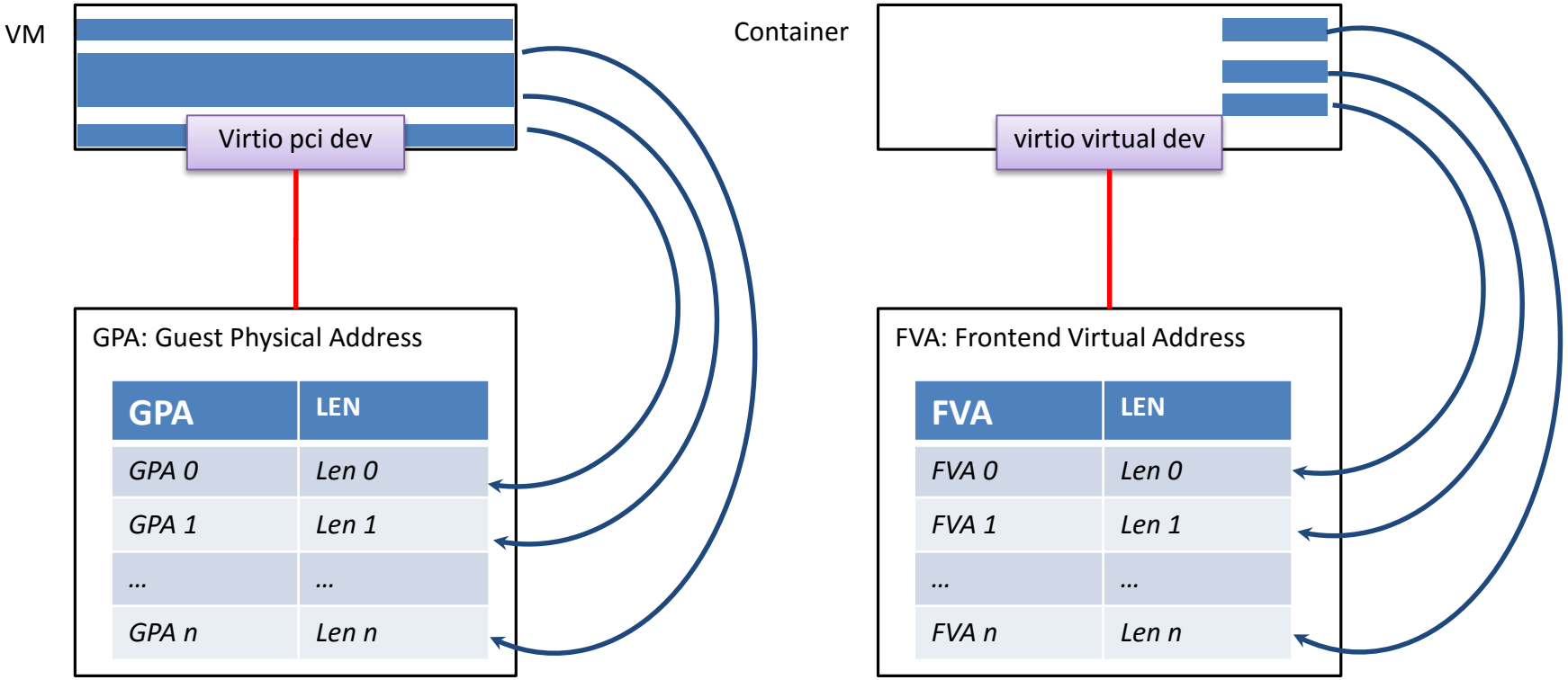


virtio for container - Architecture

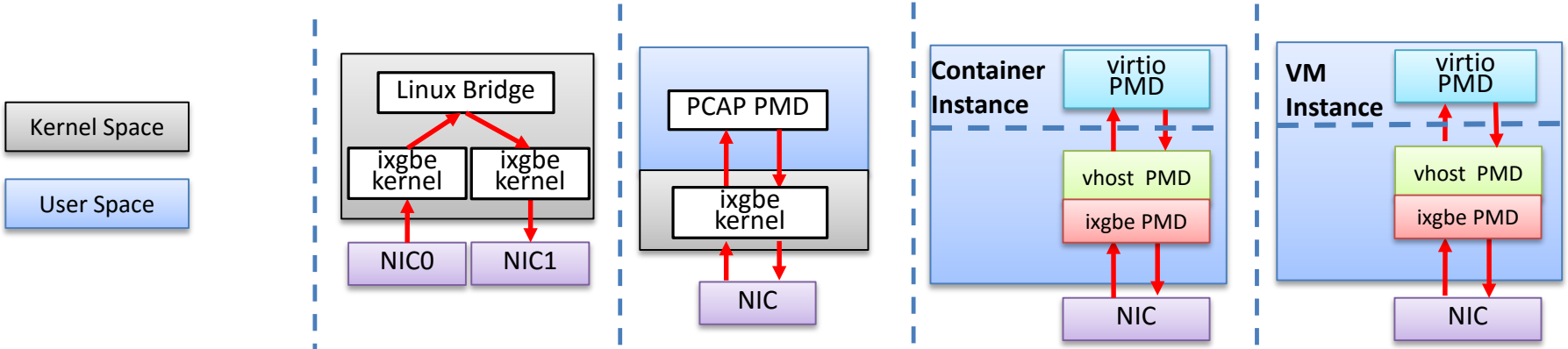
- A new IPC in essence
  - Kernel-bypass
  - Well defined msg format
  - Cache friendly
- Virtio in Container *vs* VM
  - Device emulation
  - Address translation



Virtio for container – Addr trans



Virtio/Container - Performance



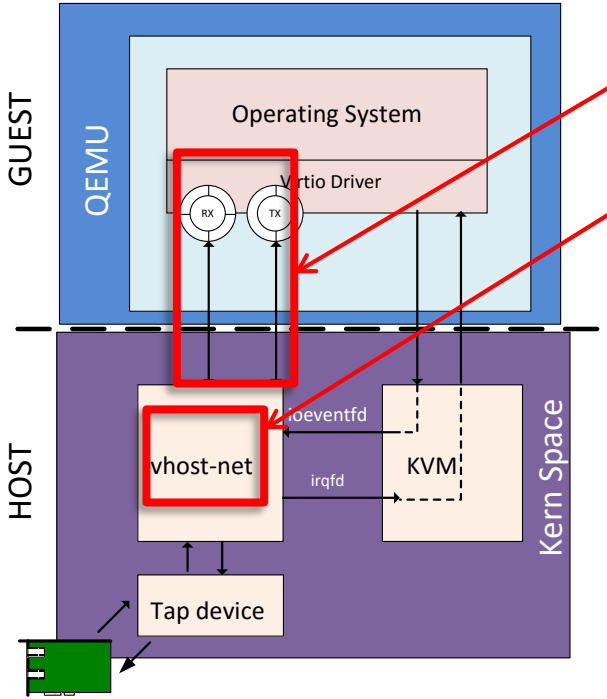
Test with various packet size	Kernel forwarding (Mpps)	DPDK Pcap Test (Mpps)	Container with virtio (Mpps)	VM with virtio (Mpps)
64 Bytes	0.933	0.699	**9.8**	**9.5**
256 Bytes	0.909	0.673	4.5	4.5
512 Bytes	0.821	0.655	2.3	2.3
1024 Bytes	0.819	0.623	1.2	1.2

CPU: Intel(R) Xeon(R) CPU E5-2699 v3 @ 2.30GHz with HT disabled

Disclaimer: prototyping result, subject to change with different system configurations


DPDK
DATA PLANE DEVELOPMENT KIT

virtio/vhost background



Virtio is the KVM standard for communicating with Virtual Machines (VM) efficiently

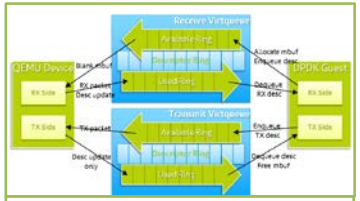
Vhost is the KVM backend for Virtio, supplying packets to a Virtio Frontend.

Packet Flow

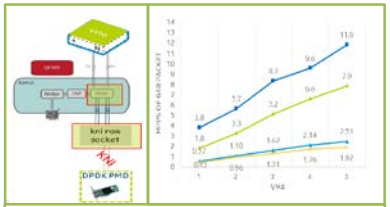
A virtual switch, switches packets to the backend (vhost) and these are forwarded to the frontend (virtio) in the Guest.

Existing kernel space components.

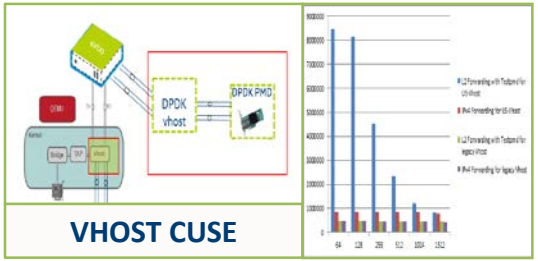
DPDK virtio development journey



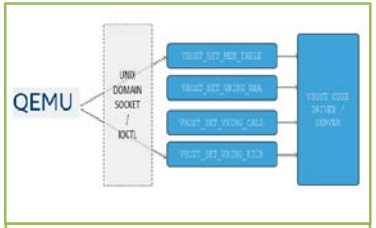
VIRTIO PMD



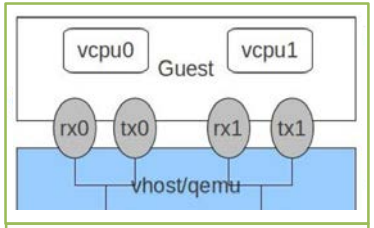
KNI VHOST



VHOST CUSE



VHOST USER and lib

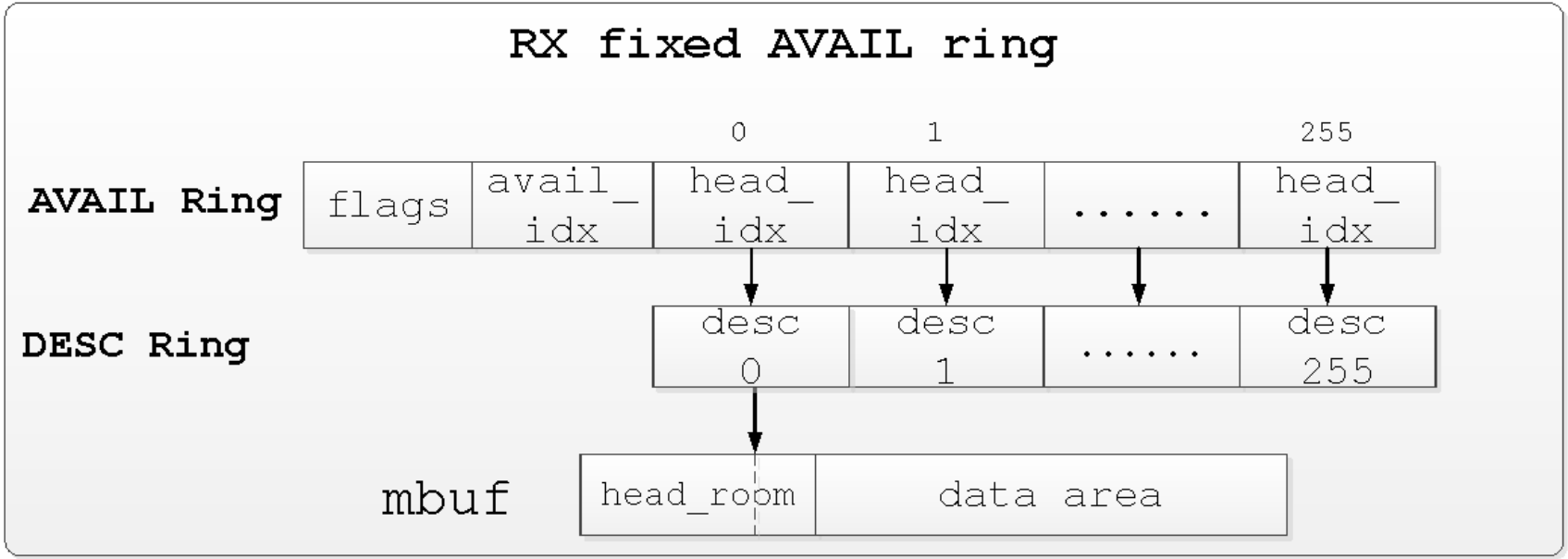


virtio multiple QUEUE

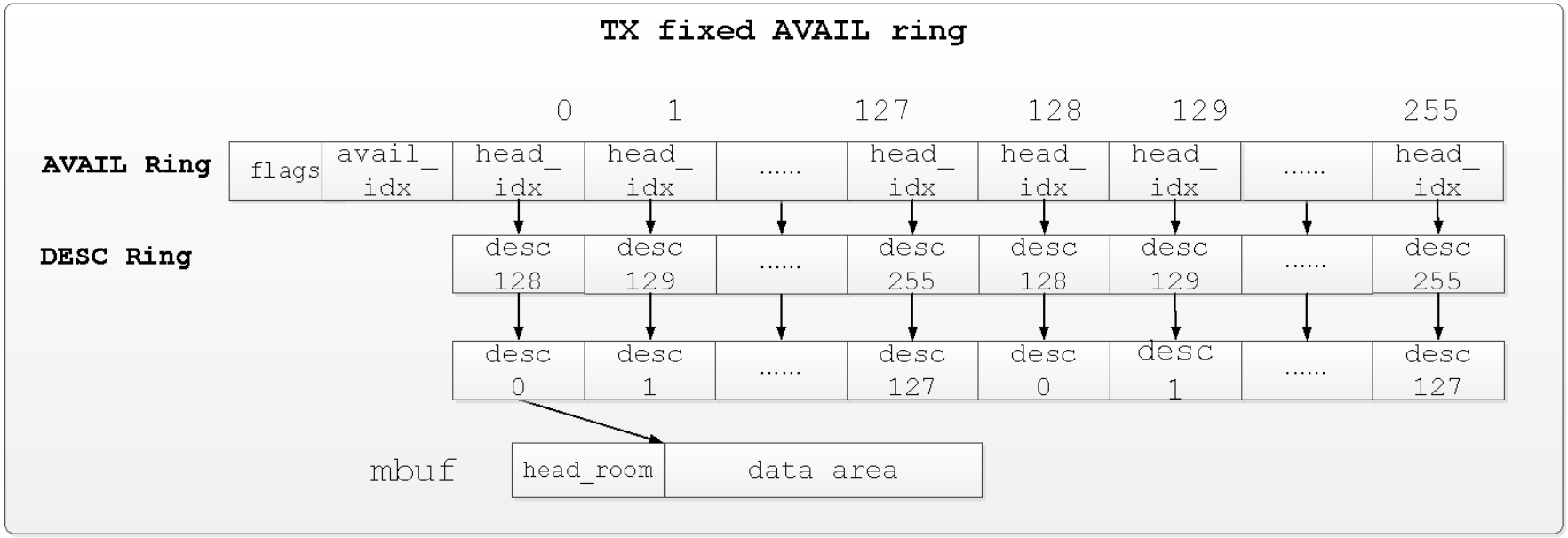
VIRTIO 1.0

VHOST live migration

virtio optimization: ring layout and vectorization



virtio optimization: ring layout and vectorization



VM2VM fastpath(WIP)