DPDK VIRTIO PERFORMANCE - STATUS AND NEAR-TERM PLAN

Zhihong Wang, Software Engineering Manager, Intel

Dong Wang, Platform Application Engineer, Intel

Notices and Disclaimers

© Copyright 2019 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.

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].

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 modelling, and provided to you for informational purposes. Any differences in your system hardware, software or configuration may affect your actual performance.

Agenda

Virtio introduction

DPDK Virtio/Vhost performance status

Spec and code readiness and plan

Some Facts About Virtio

Open, evolving, backwards-compatible spec

Big ecosystem, de-facto virtualization interface

- Native support in KVM/QEMU ecosystem
- VM, Container, IPC, User & Kernel space
- Live Migration support

Good performance: Optimization + Acceleration

- SW and HW friendly

Good Cloud-Ready choice for VM and Container

Virtio Spec

A virtualization standard across Hypervisors and Operating Systems

- Network, storage, GPU…

Backwards compatible

- Feature negotiation

Milestone	Time	Main changes
Virtio	2008	Initial paper
Virtio 0.95 draft	2012	Non-normative spec
Virtio 1.0	2014	CSR PCI capacity, MMIO
Virtio 1.1	2019	Packed ring layout, IN_ORDER, HW SRIOV
…	…	…

DPDK, Kernel, QEMU support for v1.1 underway

Active and evolving

- https://github.com/oasis-tcs/virtio-spec/issues
- https://lists.oasis-open.org/archives/virtio-dev/

Who Is Using It?

Almost all open source Hypervisors and Guest OSes

Containers

vSwitches and SmartNICs

DPDK

DATA PLANE DEVELOPMENT KIT

Firecracker

katacontainers

rust-vmm

Benchmark Methodology

Define the DUT

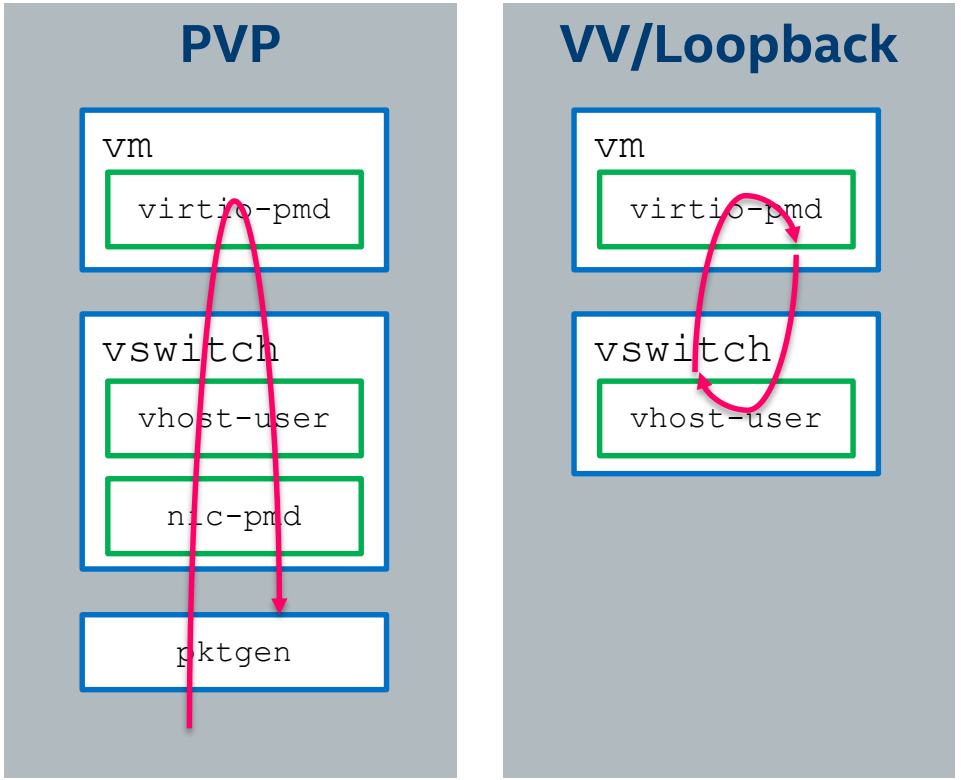
- Virtio, Vhost, IPC, vSwitch

Define the use case

- Header & data access
- Max throughput vs. RFC2544

Make DUT the BOTTLENECK

- Multi queue, port, core



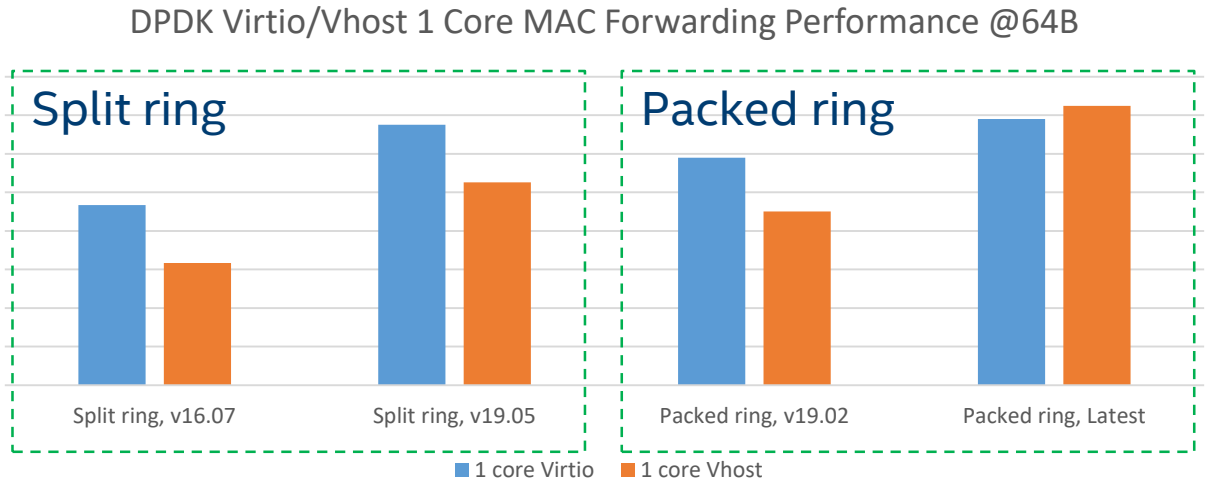
DPDK Virtio/Vhost Performance Status

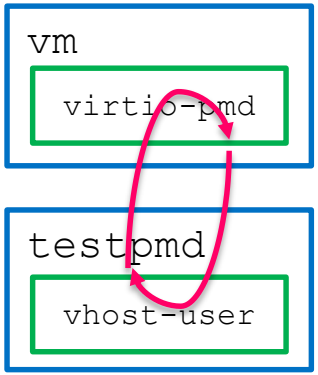
Test and System Configurations: Estimates are based on internal Intel analysis using Intel® Server Board S2600WFT, Intel(R) Xeon(R) CPU 8180 @ 2.50GHz, Intel® XL710-QDA2 Gigabit Ethernet Controller. Performance results are based on testing as of 08/28/2019 and may not reflect all publicly available security updates. No product or component can be absolutely secure. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit www.intel.com/benchmarks.

Huge improvement over releases

Packed ring brings ~35% Vhost improvement

Virtio driver optimization underway



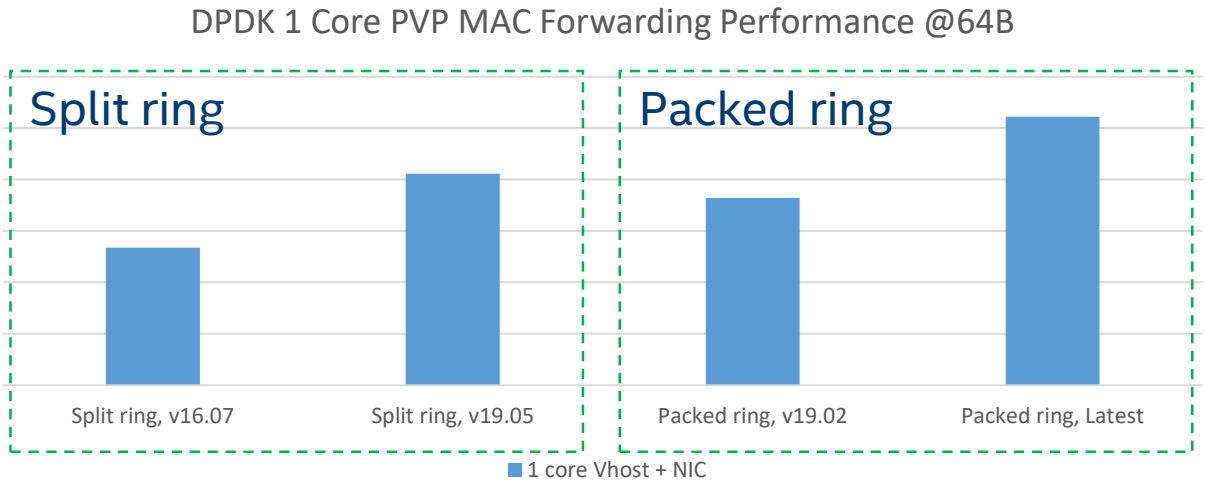


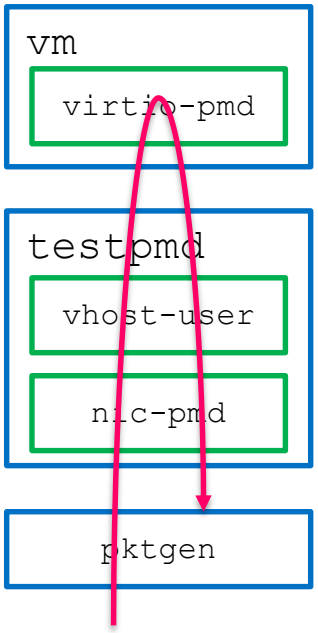
DPDK PVP Performance Status

Test and System Configurations: Estimates are based on internal Intel analysis using Intel® Server Board S2600WFT, Intel(R) Xeon(R) CPU 8180 @ 2.50GHz, Intel® XL710-QDA2 Gigabit Ethernet Controller.
Performance results are based on testing as of 08/28/2019 and may not reflect all publicly available security updates. No product or component can be absolutely secure. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit www.intel.com/benchmarks.

Huge improvement over releases

Packed ring brings ~25% PVP improvement

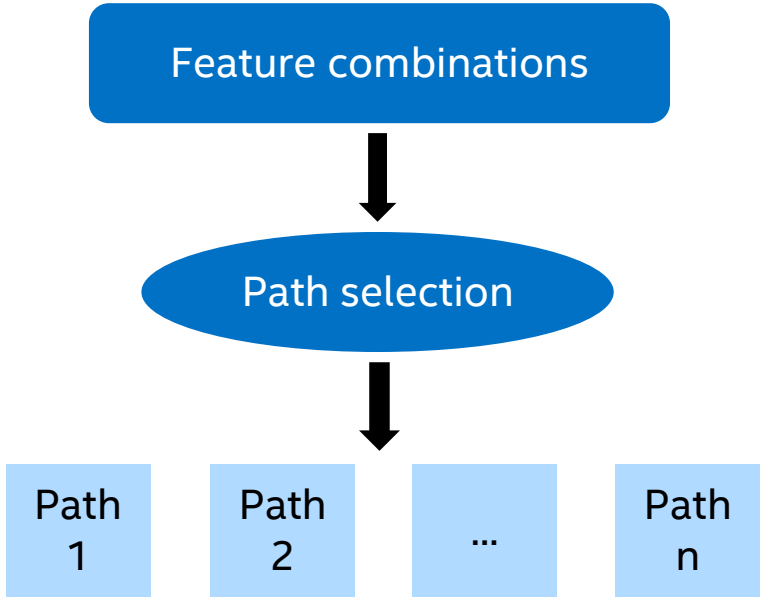




Under The Hood

Dedicated Virtio/Vhost paths for targeted optimization

- `RING_PACKED`
- `IN_ORDER`
- `MRG_RXBUF`



Virtio/Vhost Performance Internals

Critical factors for optimal & consistent Virtio performance

- Consistent cache latency (Mesh)
- Fast memcpy (IOAT, AVX-512) underway

Optimization techniques

- Separated fast path
- Loop unrolling
- Cache line aligning & reducing conflict
- Batch descriptors write back

Vhost Performance In Virtual Switching

Checkout enhancements in latest DPDK

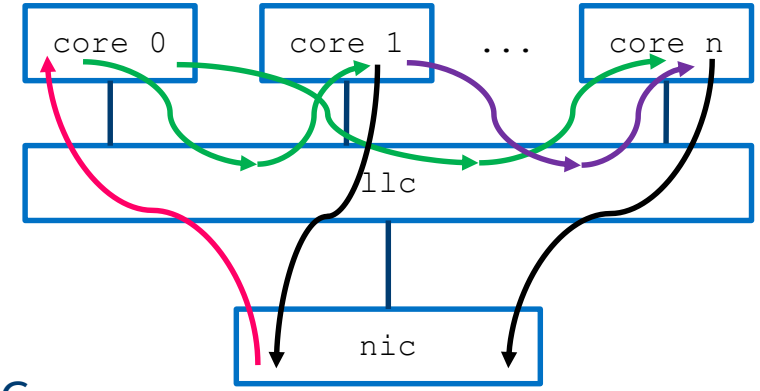
- RING_PACKED, IN_ORDER, optimizations

NIC RSS to reduce OVS PMD contentions

- Batching, locking

In Cross-NUMA cases

- Prefer Vhost stays with Virtio
- If not applicable, prefer Vhost stays with NIC

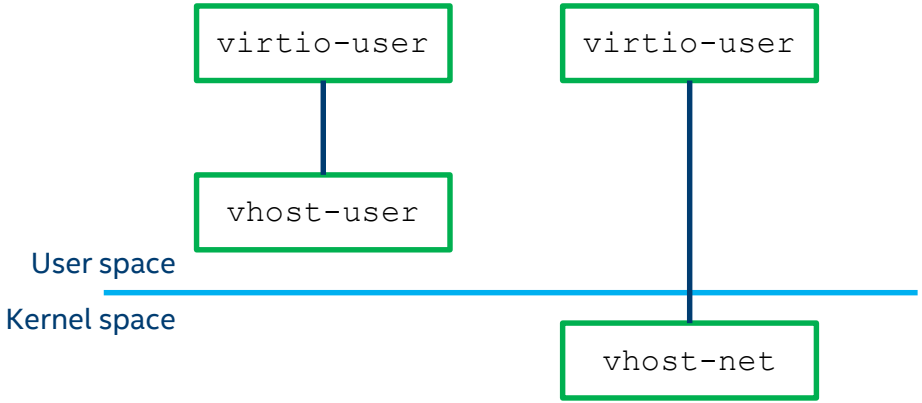


Virtio-user

Container interface for kernel-bypass virtual switching

User space kernel stack injection channel

Kernel stack injection optimization underway



ACM DL DIGITAL LIBRARY

VIRTIO-USER: A New Versatile Channel for Kernel-Bypass Networks

Published in:

· Proceeding

KBNets '17 Proceedings of the Workshop on Kernel-Bypass Networks

Pages 13-18

Los Angeles, CA, USA — August 21 - 21, 2017

ACM New York, NY, USA ©2017

table of contents ISBN: 978-1-4503-5053-2 doi>10.1145/3098583.3098586

https://dl.acm.org/citation.cfm?id=3098583.3098586

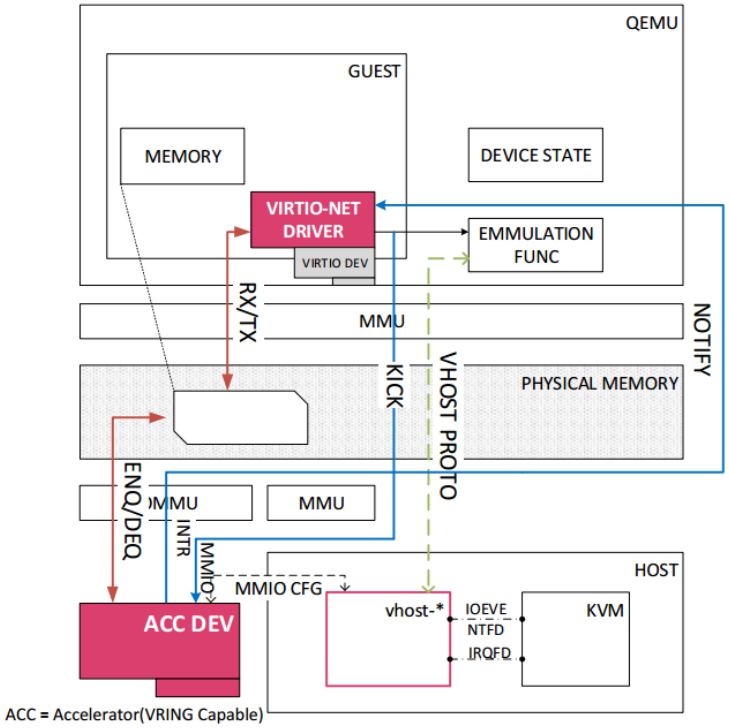
vDPA

Virtio data path acceleration framework

- Native performance for Cloud-Ready
- Relay within driver if not Virtio device
  - Zero-copy for better performance

Fully supported in DPDK since v18.08

Kernel vhost-mdev underway



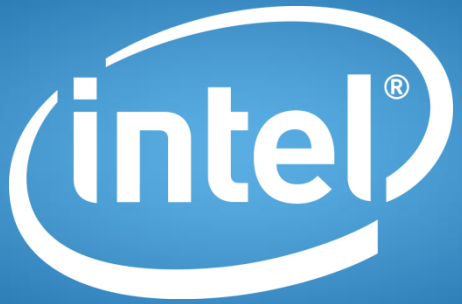
https://www.linux-kvm.org/images/8/87/KVM17vDPA-v4_0.pdf

Key Takeaways

Virtio is fast evolving in both features and performance

Innovation empowers new use scenarios: Container, Acceleration

Intel Platform provides optimal & consistent Virtio performance


intel®