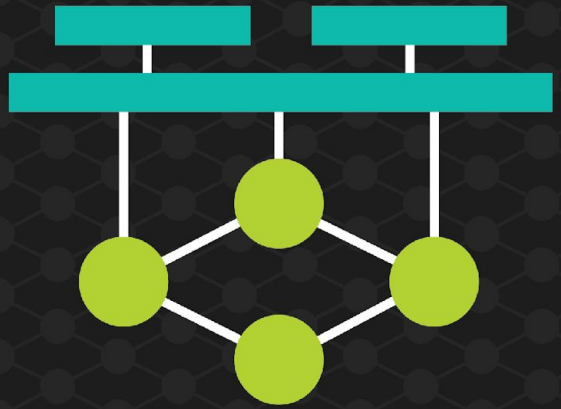
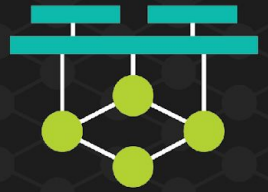


ons

EUROPE

OPEN NETWORKING //
Enabling Collaborative
Development & Innovation

Hosted By
THE LINUX FOUNDATION | LF NETWORKING



ons
EUROPE
OPEN NETWORKING //
Enabling Collaborative
Development & Innovation

Plumbing Hardware Accelerated Interfaces into Containerized Network Functions

Thomas F Herbert Red Hat

Hosted By
THE LINUX FOUNDATION | LF NETWORKING

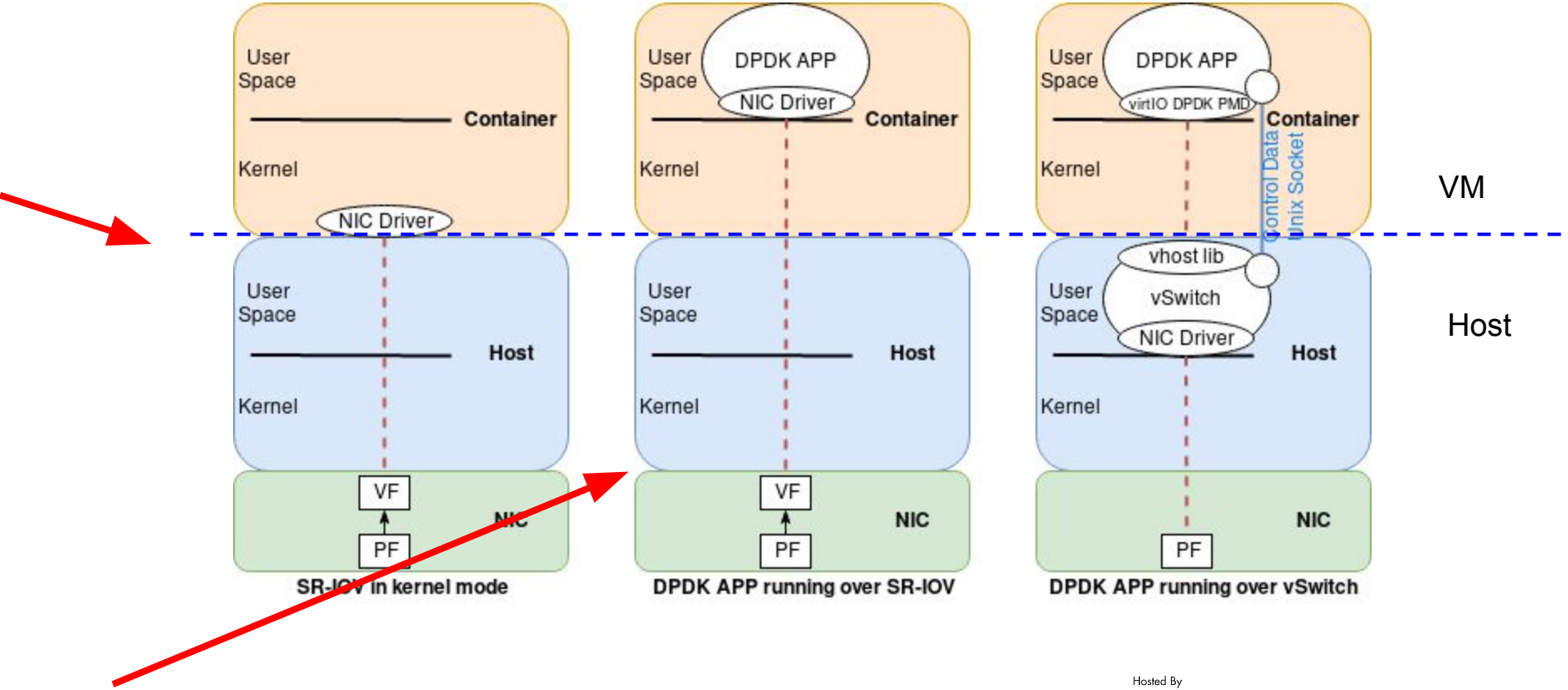
Plumbing HW Accelerated Interfaces into CNFs

- Legacy Virtualization
- HW Acceleration into Containers
  - SR/IOV
  - TCFlower - OVS OffLoad
  - vDPA vHost Data Path Acceleration

Legacy Virtualization

- Legacy Virtualization
  - Device DMA into Port on vSwitch
    - Data Forwarded to vHost-User Port
    - Data Copied via vHost-User to
    - User Space or Kernel Driver in VM

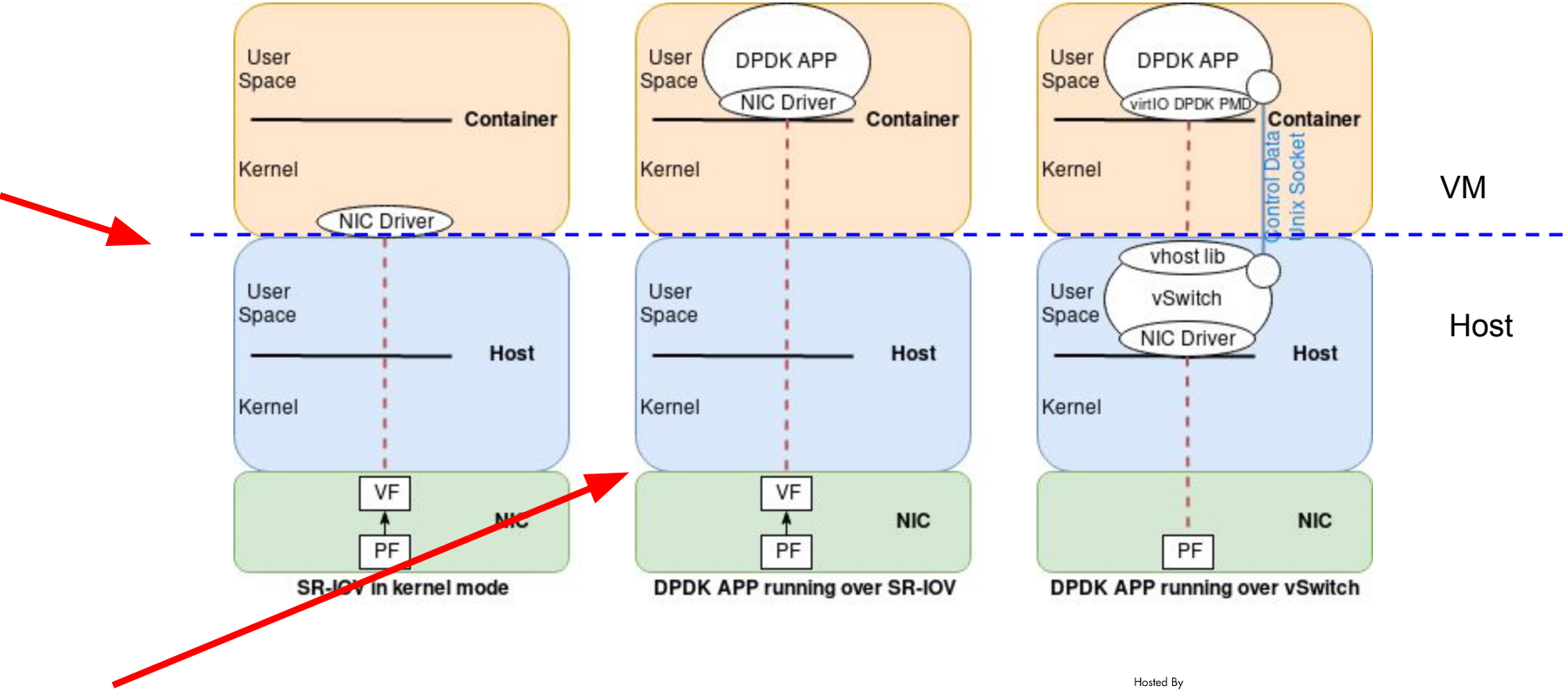
Legacy Virtualization



SR/IOV

- Single Root Input/Output Virtualization
  - Device Bandwidth split into Multiple VFs
    - Supports Zero copy
      - Directly to VMs or Containers
      - With Either User Space or Kernel Driver in VM
    - Vendor Specific

SR/IOV


User Space
Container
Kernel
NIC Driver
User Space
Host
Kernel
VF
PF
NIC
SR-IOV in kernel mode
User Space
DPDK APP
NIC Driver
Container
Kernel
User Space
Host
Kernel
VF
PF
NIC
DPDK APP running over SR-IOV
User Space
DPDK APP
virtIO DPDK PMD
Container
Kernel
control Data
Unix Socket
vhost lib
vSwitch
NIC Driver
User Space
Host
Kernel
PF
NIC
DPDK APP running over vSwitch
VM
Host

TCFlower - OVS OffLoad

- OVS OffLoad
  - OVS Virtual Switch on Host
  - Fast Packet Forward Directly on NIC
  - Flow Management by OVS on Host
  - Forwarding Directly on NIC at Wire speeds
- Currently No Zero Copy Support for User Space
- Solves Different Problem
  - But Maybe Future vDPA Compatible
  - Data and Control Path

vDPA -- vHost Data Path Acceleration

- vDPA - What is it?
- How Does it Work
  - Data and Control Path
- Orchestration and Deployment

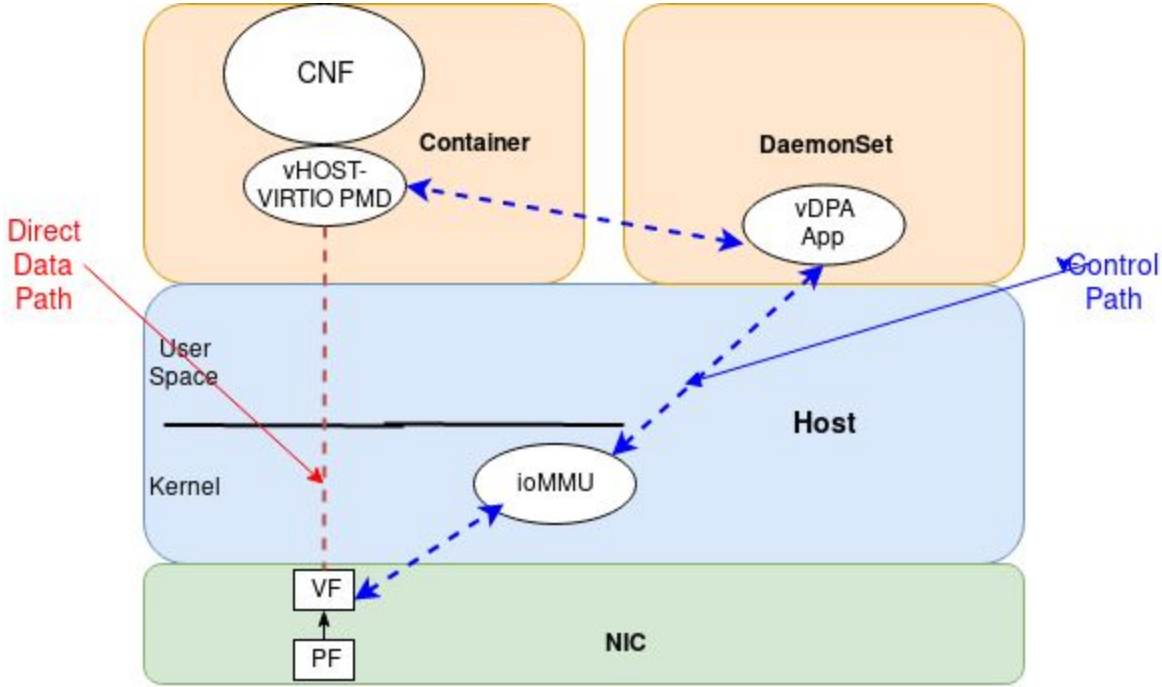
vDPA -- What is It?

- vDPA Features
  - For Fast layer 2 and layer 3 access
- Enable CNF Acceleration with
  - Vendor Agnostic
  - Standard HW IF and User CNF API
- Similar to SR/IOV but Vendor Independent
  - NIC “Knows” Virtio rings
  - Maps Data Direct to User Space CNF

vDPA -- How Does it Work?

- User Space DPDK App Sees vHost User
  - ReUse of DPDK App
  - Except Maps Direct to HW
- vDPA App Sets Up Control Plane
- Kernel IOMMU Maps Buffer Descriptors
  - Zero Copy
  - Device DMA Direct To/From CNF User Space
- To User Space App: Standard NIC IF
  - DPDK CNF Similar to vHost User
    - Same Backend Abstraction

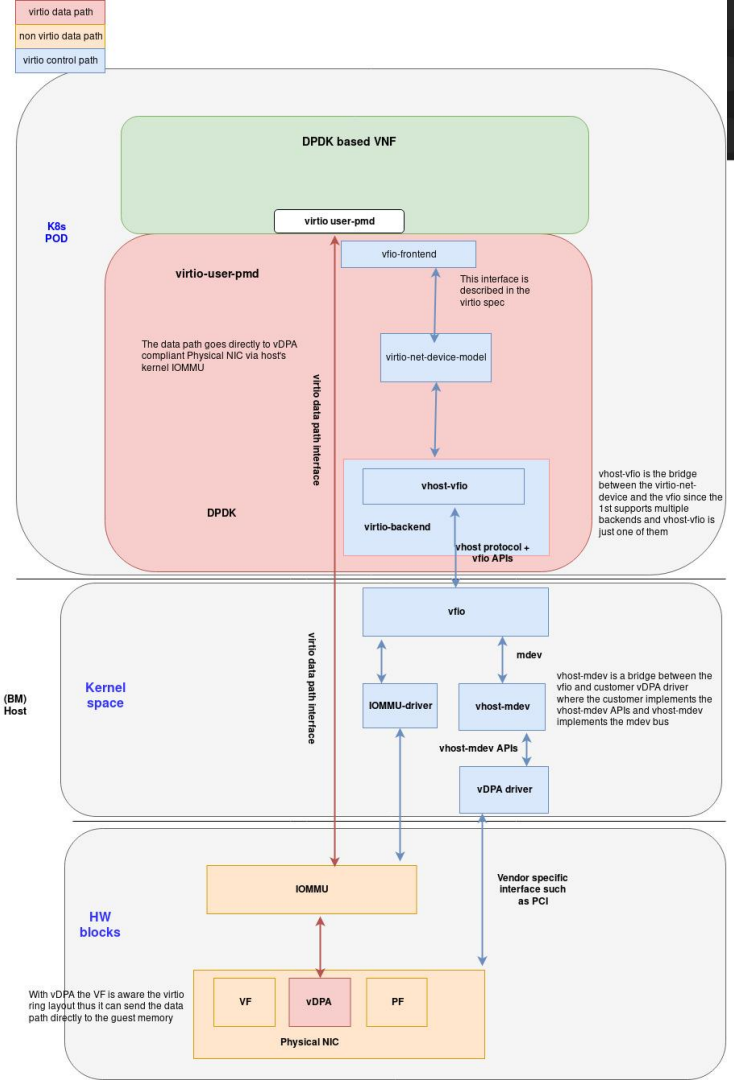
vDPA with DPDK - Simplified view



DPDK Container vDPA

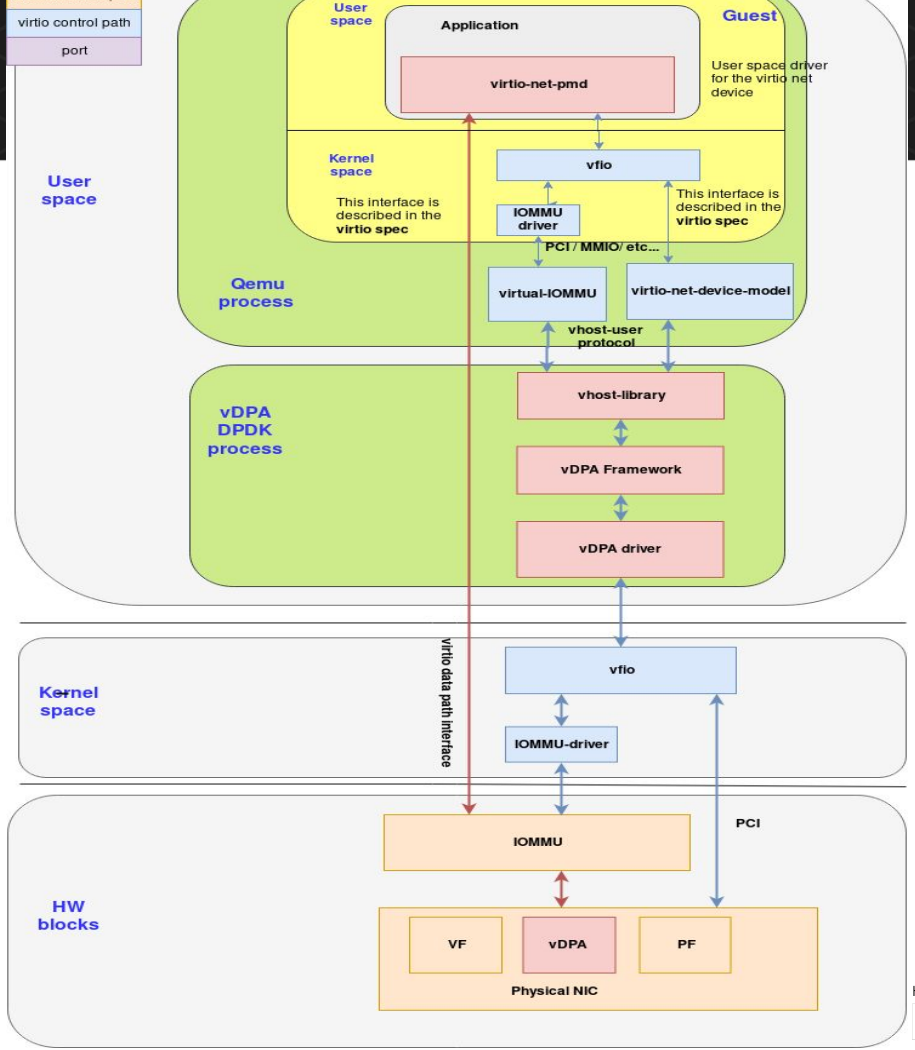
vDPA with DPDK

Detailed View



vDPA Stack

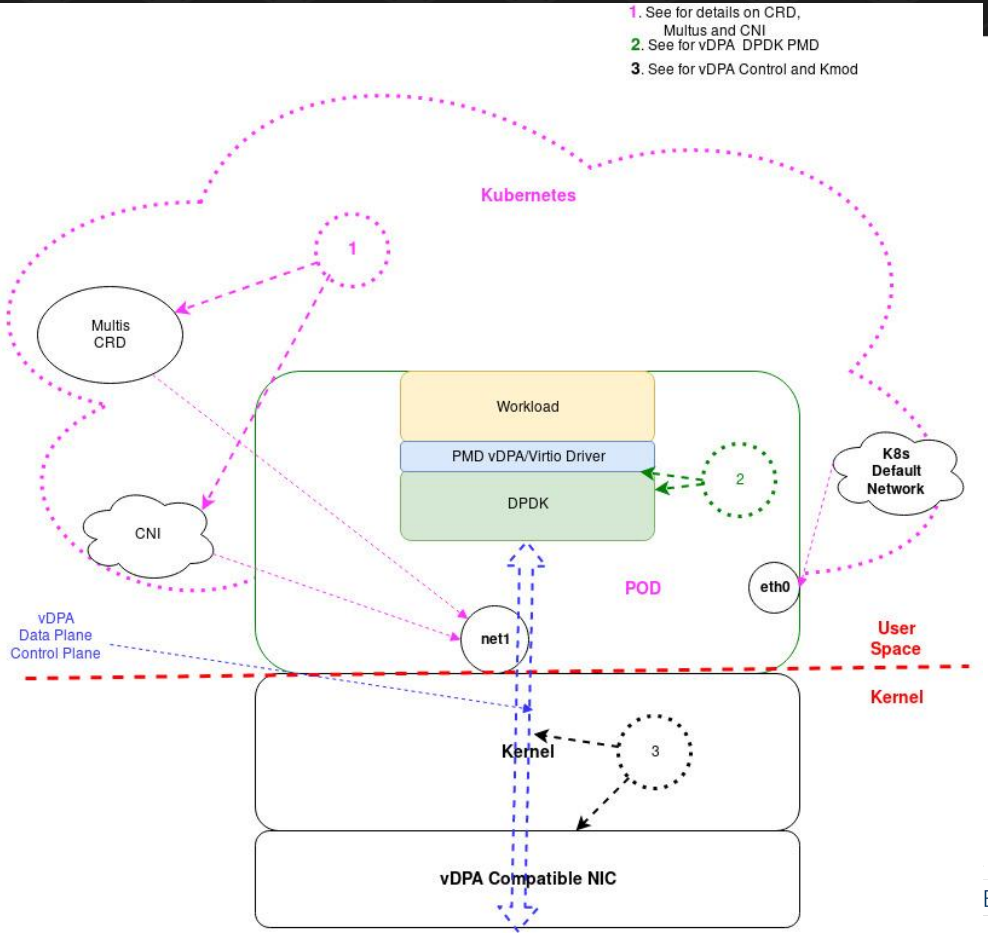
Legacy VM



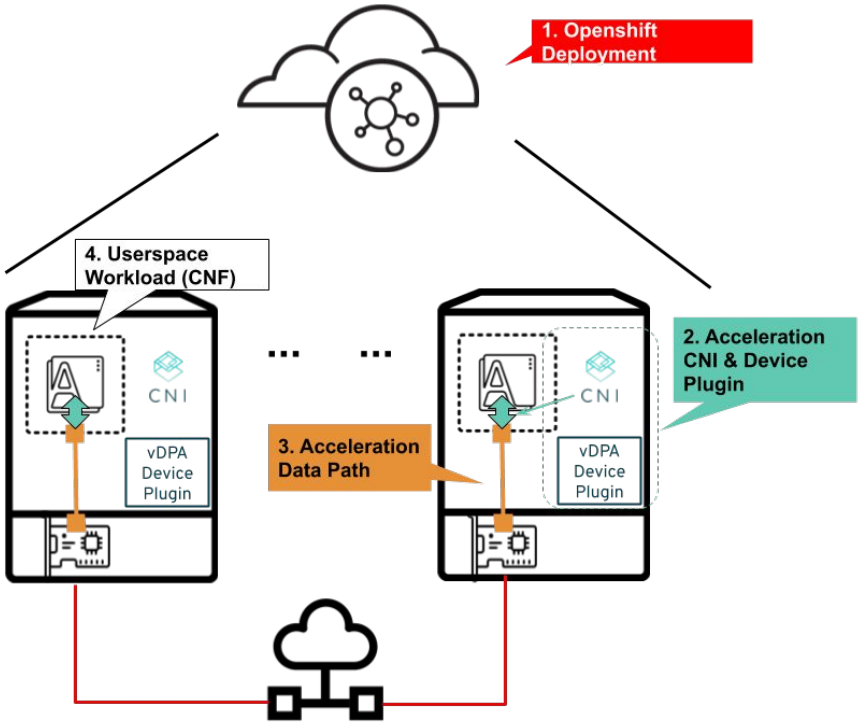
vDPA -- Orchestration and Deployment

- Similar to SR/IOV
  - CNF for DPDK
    - VF PCI-ID or
    - Control Protocol Unix Domain Socket
  - Device Plugin- DaemonSet with vDPA “App”
    - Started Before Container

Kubernetes Deployment of vDPA



Kubernetes Deployment



vDPA -- Current Status

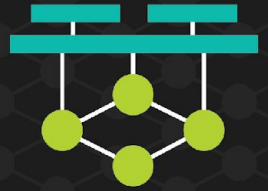
- POCs Will be Ready at Kubecon NA 2019
- HW Status
  - Using Intel Cascade Glacier
  - Mellanox bluefield
- SW Status
  - Virtio 1.1 But Backward Compatible
  - DPDK vHost-User PMD for “Backend”

vDPA -- Near Term and Long Term

- Near Term
  - Using Intel Cascade Glacier
  - Mellanox Bluefield
  - SW Status
    - Vertio 1.1
- Long Term
  - “Kernel Mode” Support
    - With Unix Domain Socket for Control Plane
    - Unified SW Stack with AF_XDP

References

- Combining Networking and Virtualization https://www.redhat.com/en/blog/introducing-virtio-networking-combining-virtualization-and-networking-modern-it?source=bloglisting
- OVS and TC Offload: https://github.com/Mellanox/mlxsw/wiki/OVS
- SR/IOV: https://en.wikipedia.org/wiki/Single-root_input/output_virtualization
- Scylla DB and DPDK: https://docs.scylladb.com/kb/dpdk-hardware/
- Vhost mdev: https://lwn.net/Articles/750770/
- Virtio-user and dpdk: https://doc.dpdk.org/guides/howto/virtio_user_as_exceptional_path.html
- vDPA for Live Migration: https://www.youtube.com/watch?v=WTawv0GwWSU
- vDPA Sample application: https://doc.dpdk.org/guides/sample_app_ug/vdpa.html
- vHost for Data Path Acceleration: https://www.youtube.com/watch?v=v50-rXGMr-Y



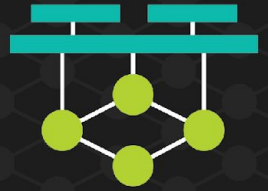
ons

EUROPE

OPEN NETWORKING //
Enabling Collaborative
Development & Innovation

Thank You

Thomas F Herbert Red Hat



ons

EUROPE

OPEN NETWORKING //
Enabling Collaborative
Development & Innovation

Backup

Thomas F Herbert Red Hat

Hosted By
THE LINUX FOUNDATION | LF NETWORKING

vDPA Stack

Legacy VM