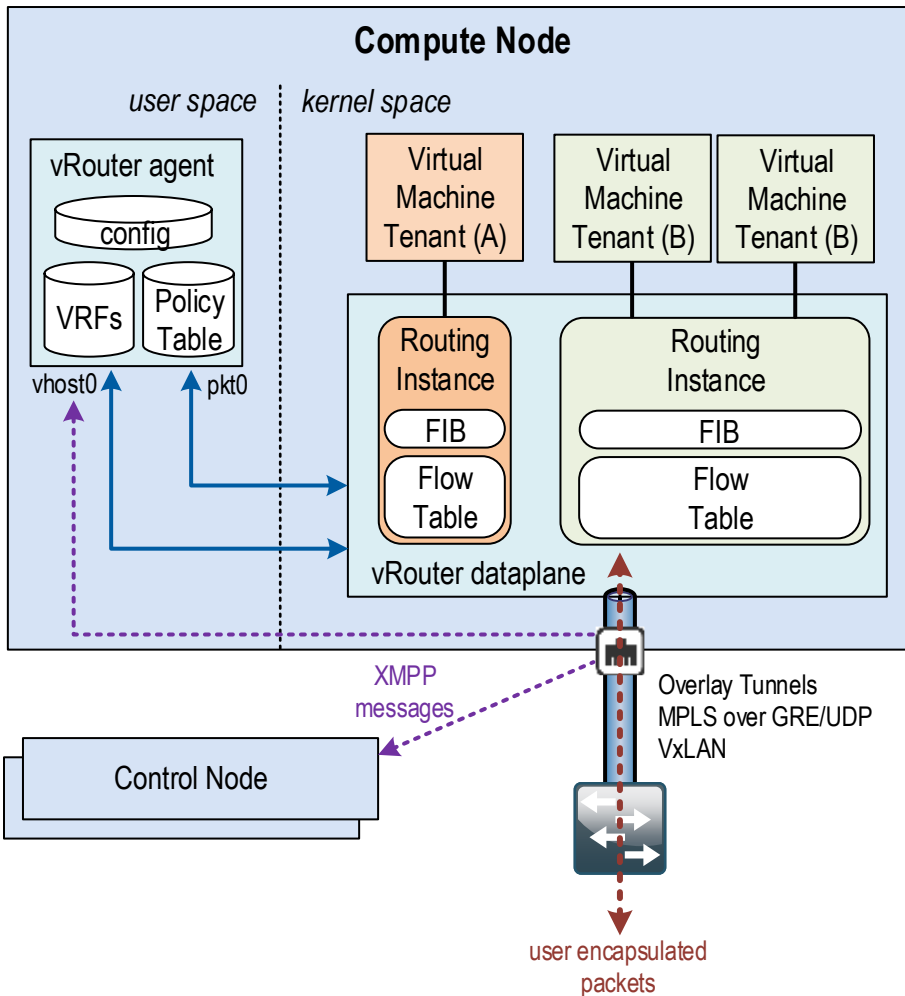


Title: vrouter overview

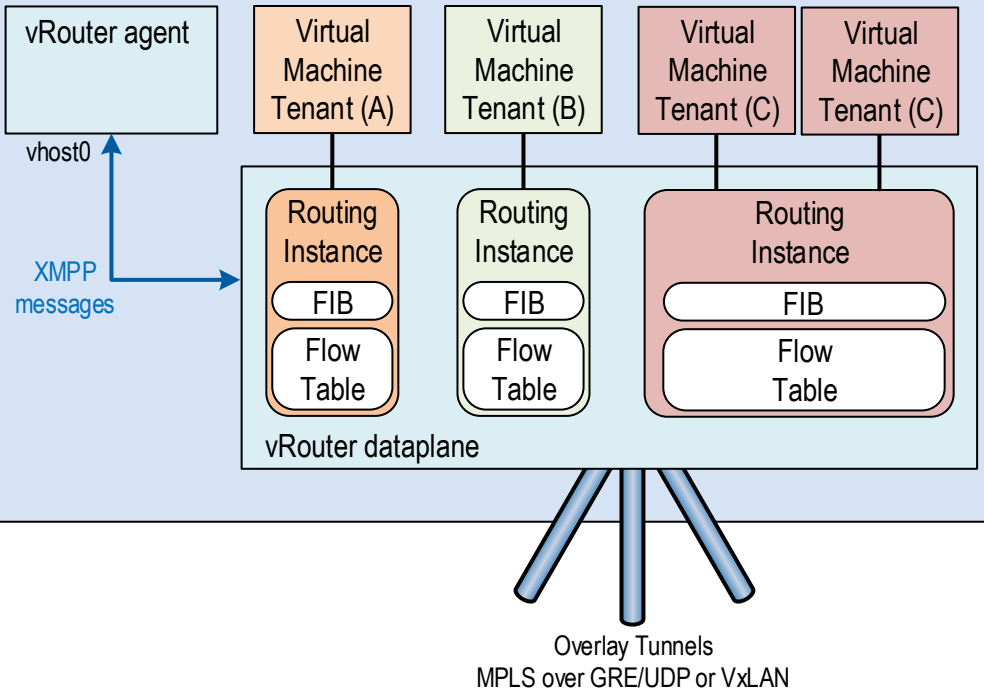
version: 1.0

date: 25/03/2020



Title: vRouter XMPP path	
version: 1.0	date: 01/10/2020

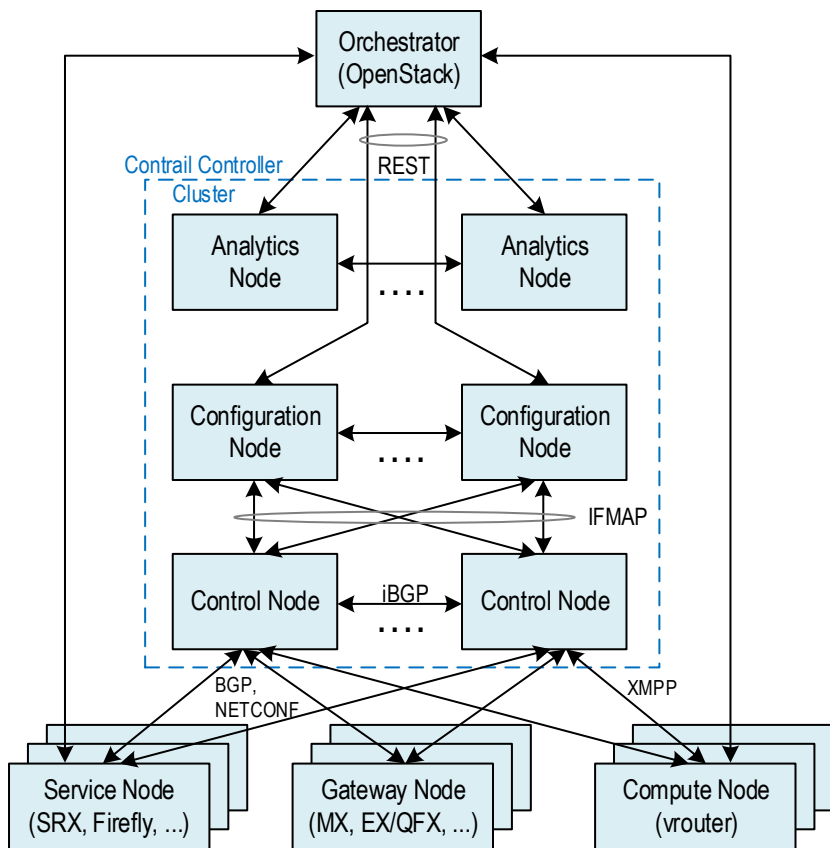
Compute Node



Title: vRouter internal paths

version: 1.0

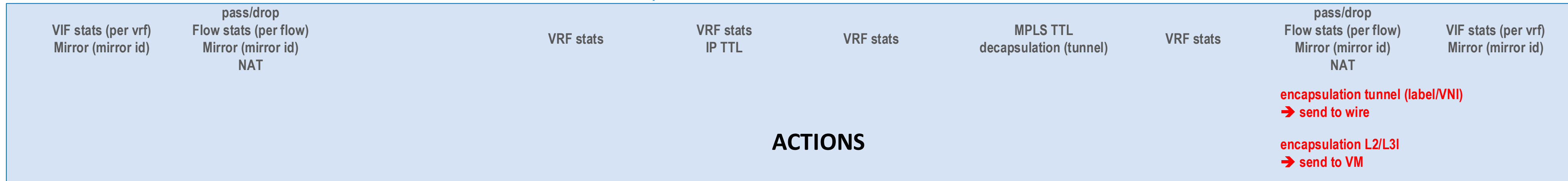
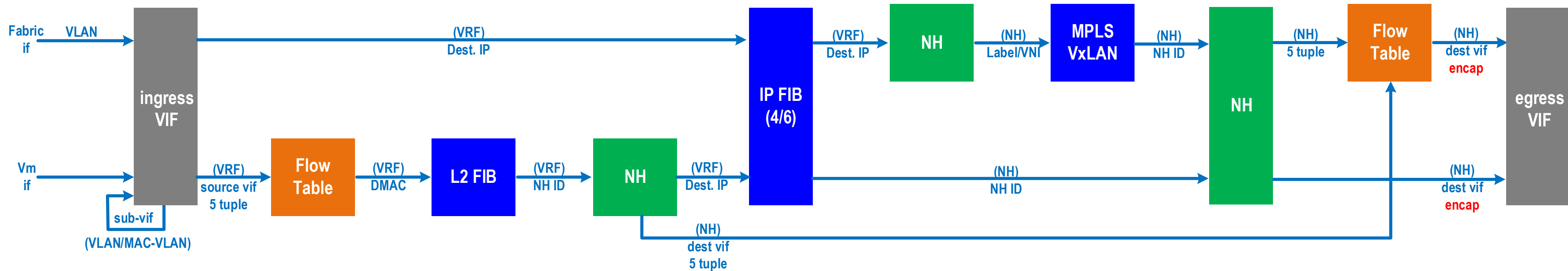
date: 01/10/2020



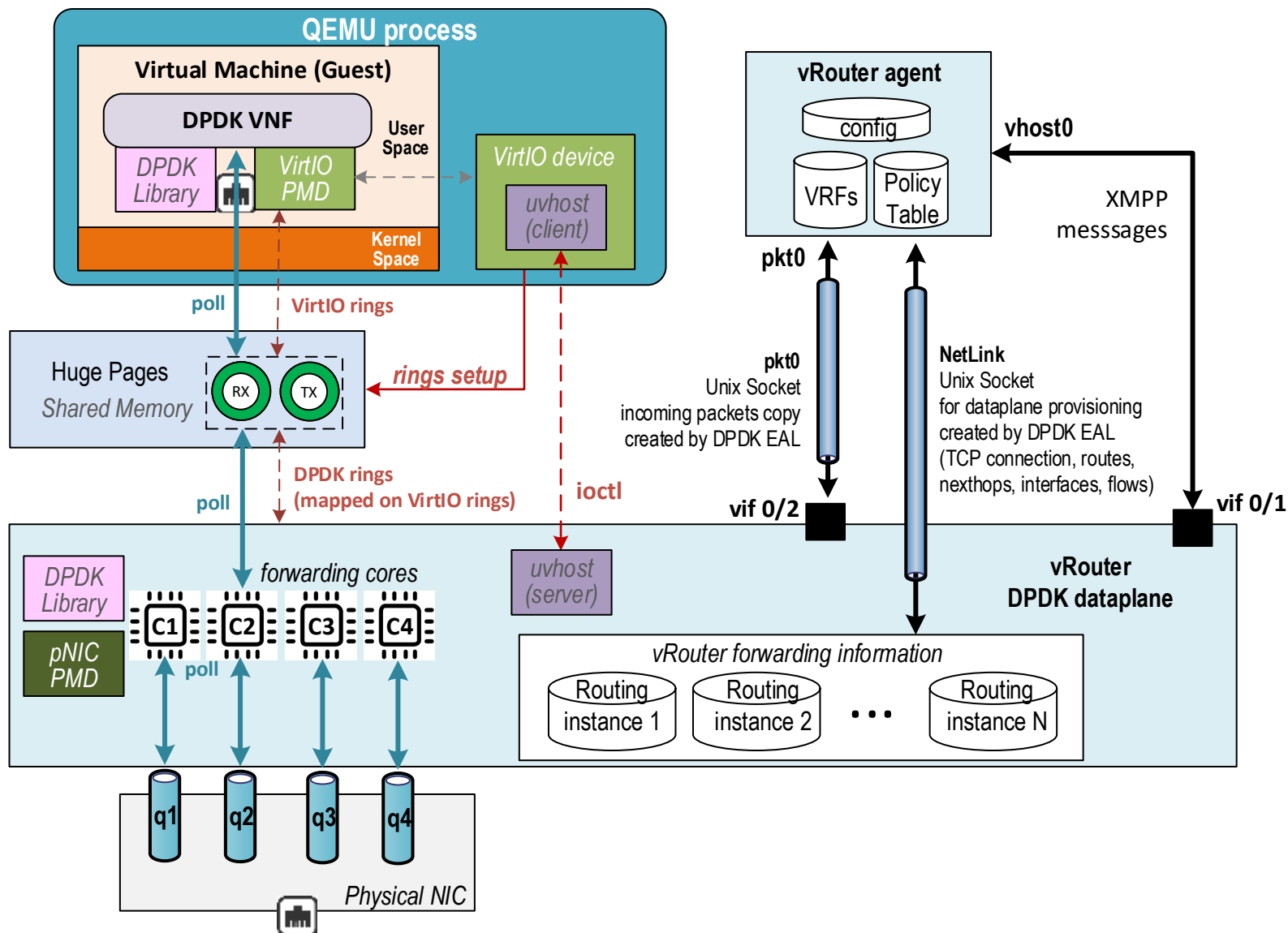
Title: Contrail Architecture

version: 1.0

date: 01/10/2020



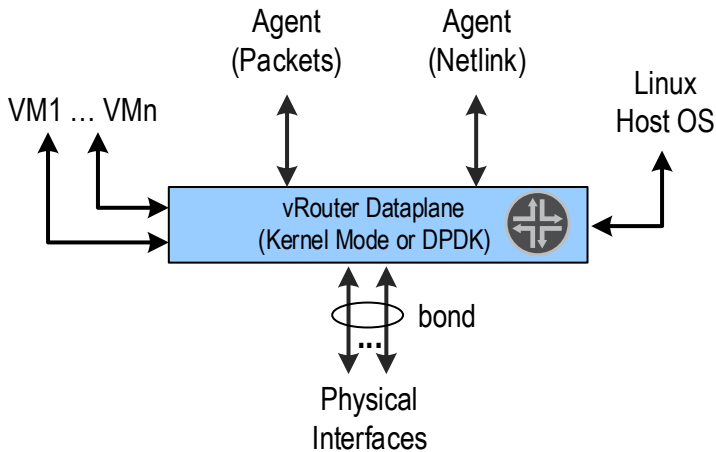
Title: vRouter Packet processing	
version: 1.0	date: 01/10/2020



Title: DPDK vRouter internal logic

version: 1.0

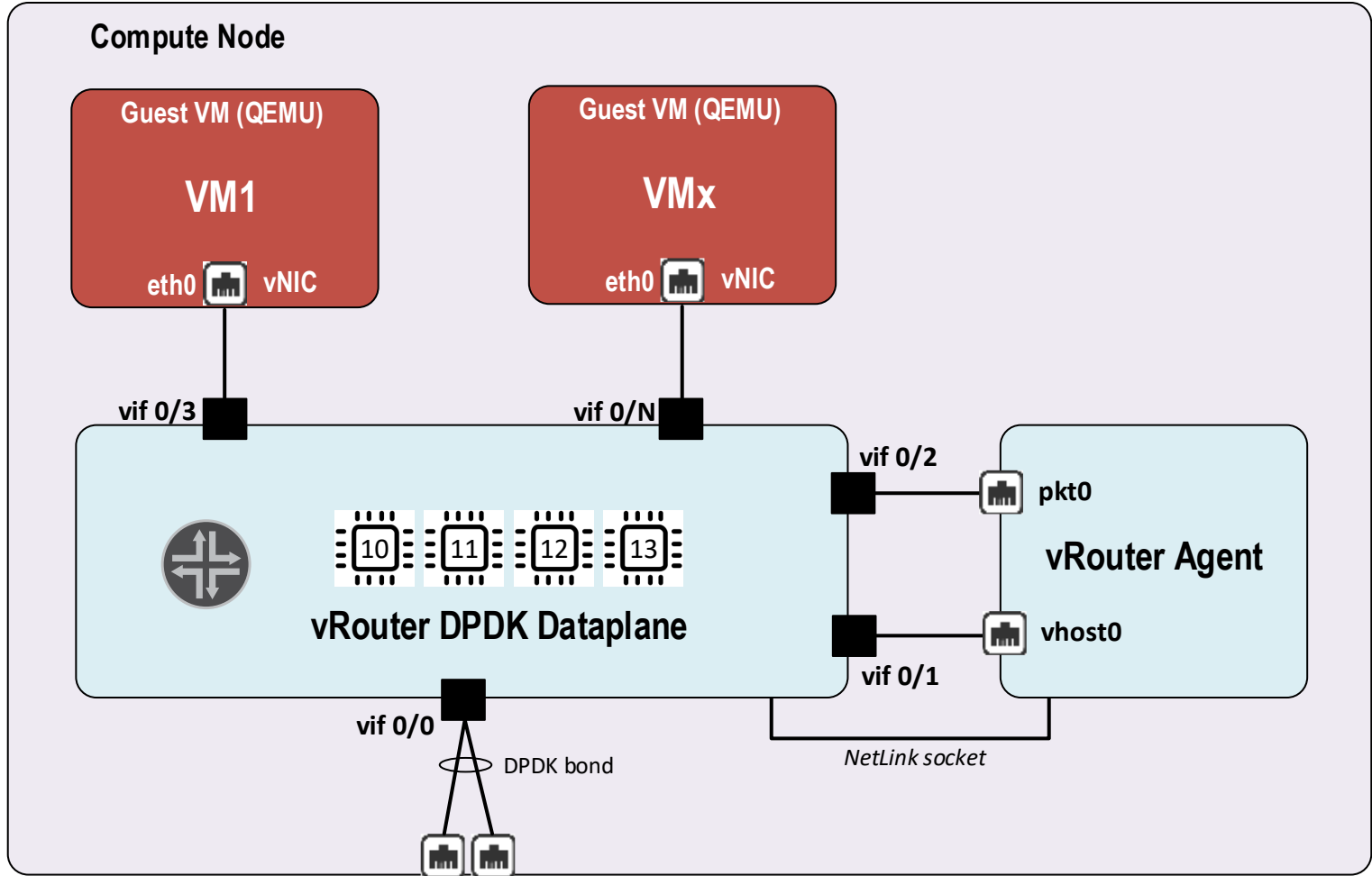
date: 01/10/2020



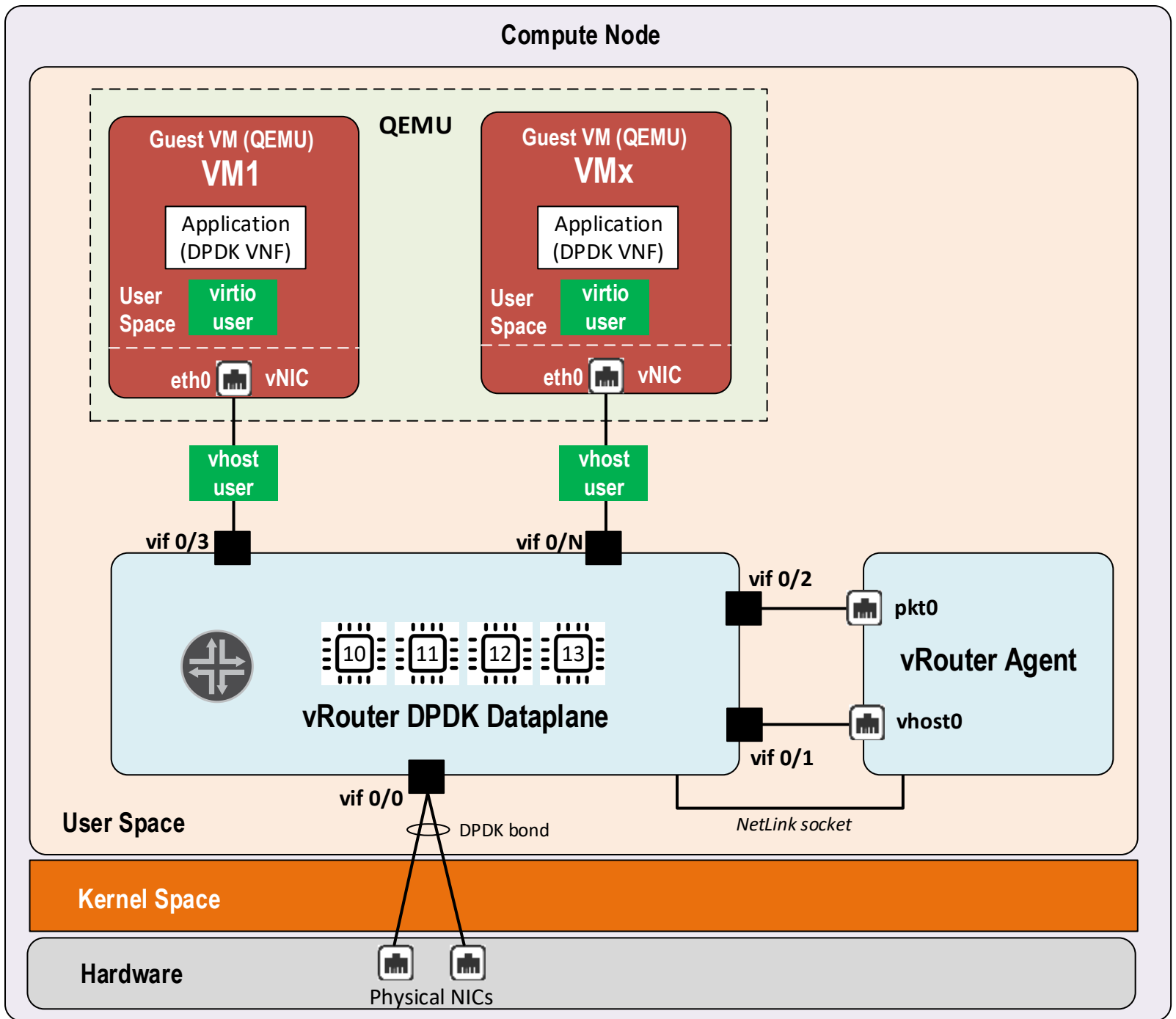
Title: vRouter interfaces

version: 1.0

date: 01/10/2020



Title: DPDK vRouter interfaces overview	
version: 1.0	date: 27/03/2020



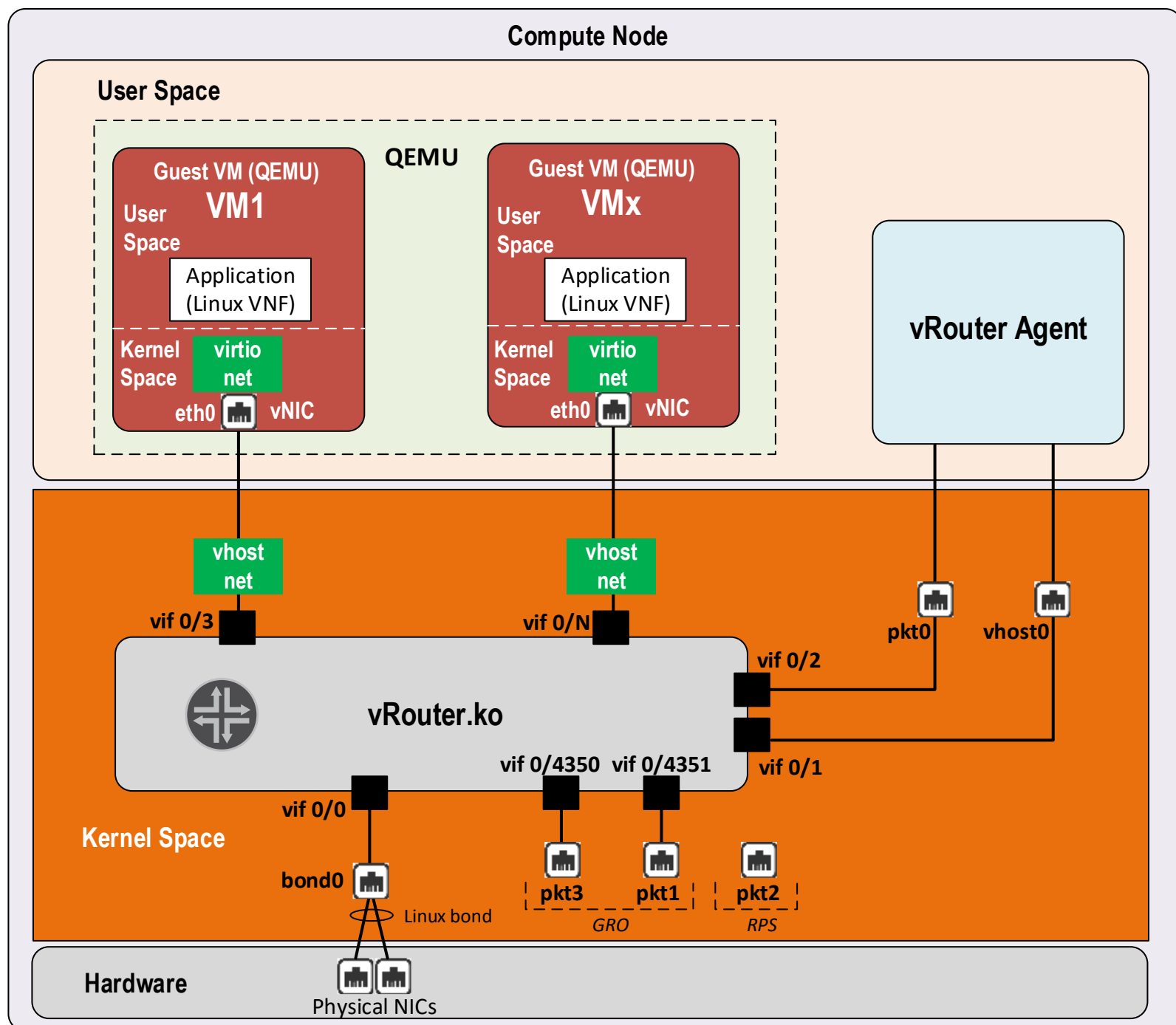
DPDK model fast:

- NIC DMA directly into user-space.
- DPDK memory allocation with hugepages reduces TLB misses.
- DPDK in host & guest copies directly from host user-space.

Title: DPDK vRouter interfaces overview v2

version: 1.0

date: 27/09/2020



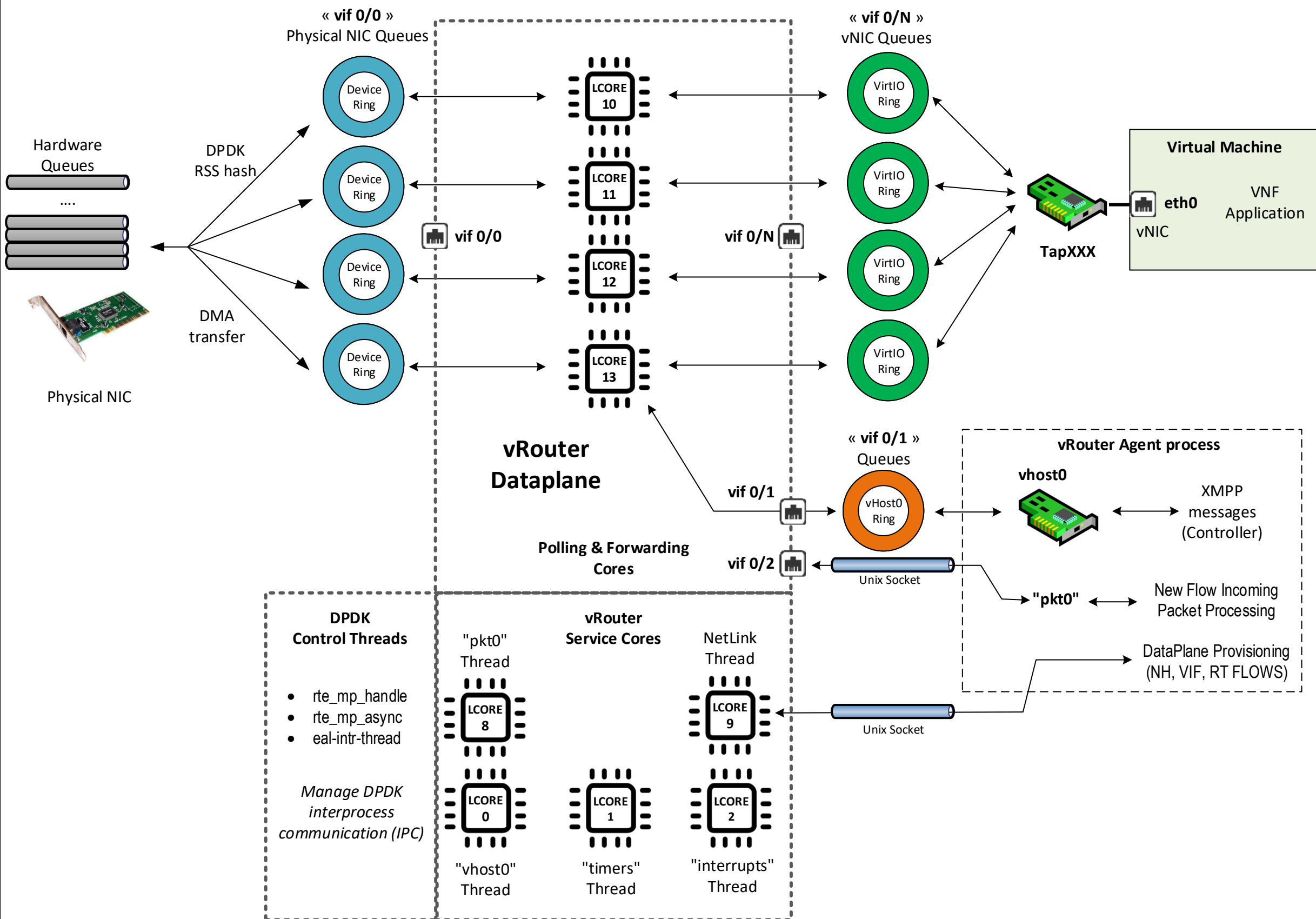
Title: Kernel Mode vRouter interfaces overview

version: 1.0

date: 27/09/2020

Kernel model slow:

- Packet needs to be copied across kernel/user-space boundary in both host and guest
- Kernel code responsible for packet I/O, tunneling, etc.

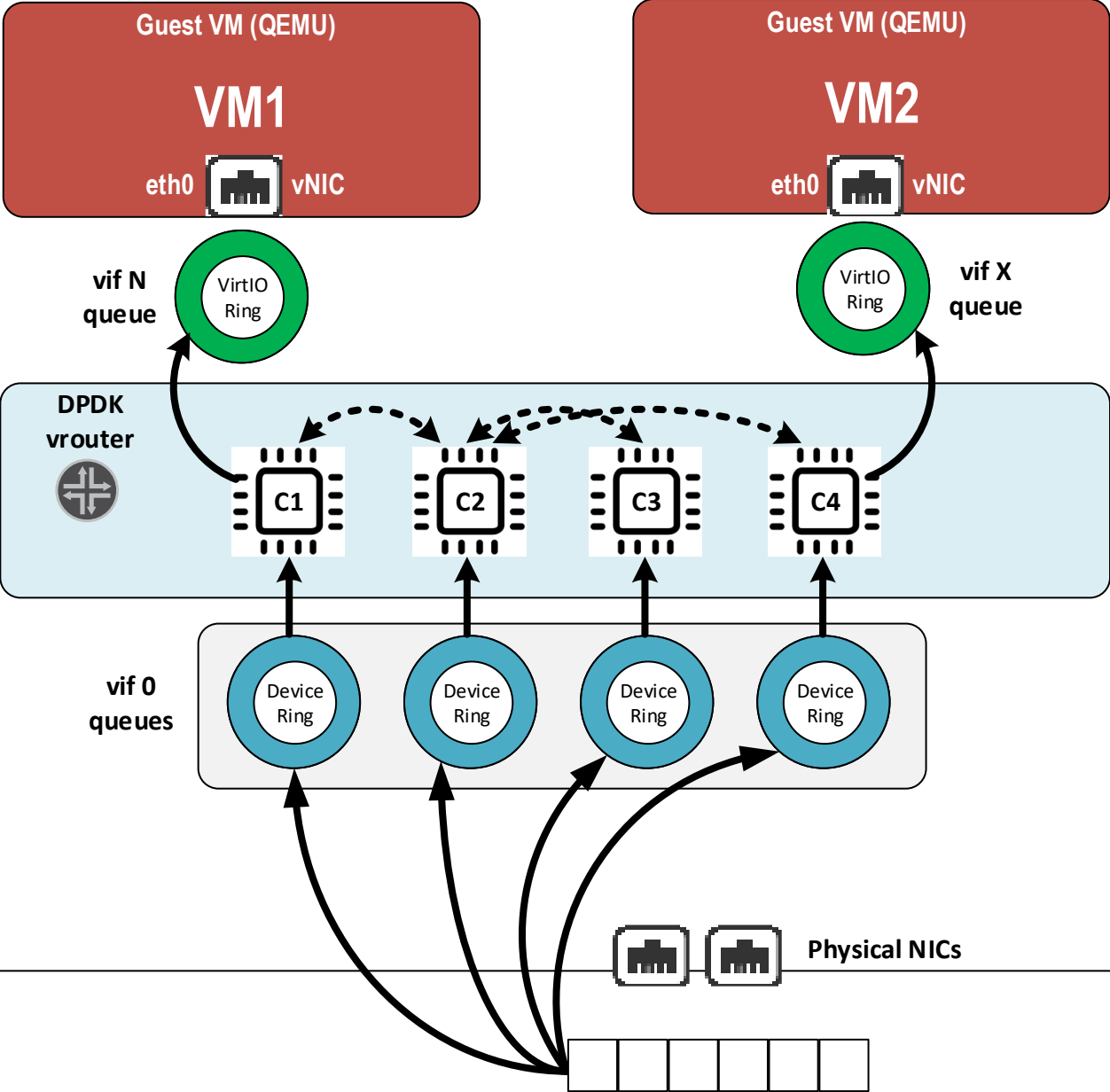


Title: vRouter Internal Architecture

version: 1.1

date: 22/05/2020

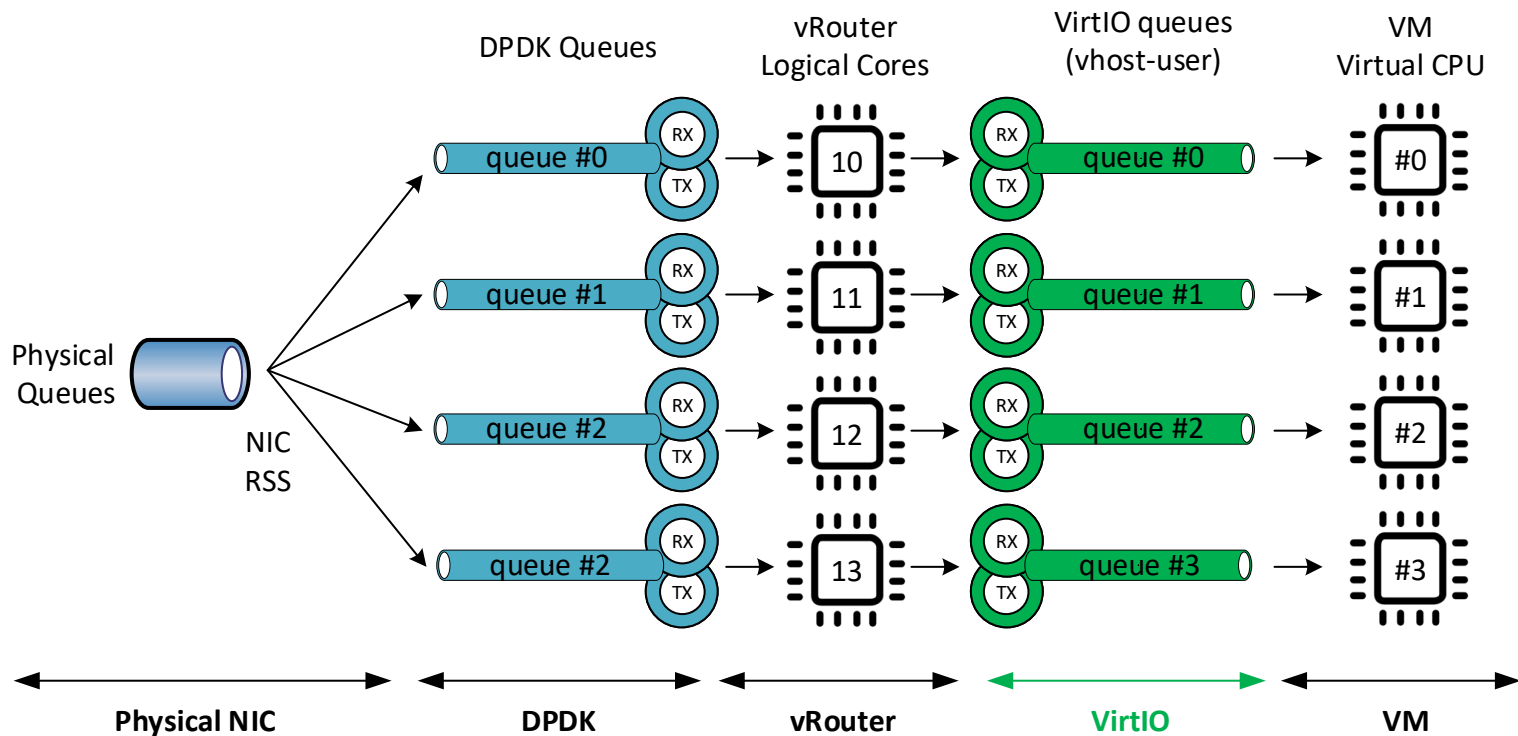
Compute Node



Title: DPDK Packet Processing Overview

version: 1.0

date: 25/03/2020



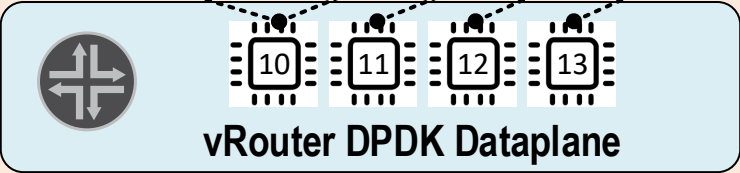
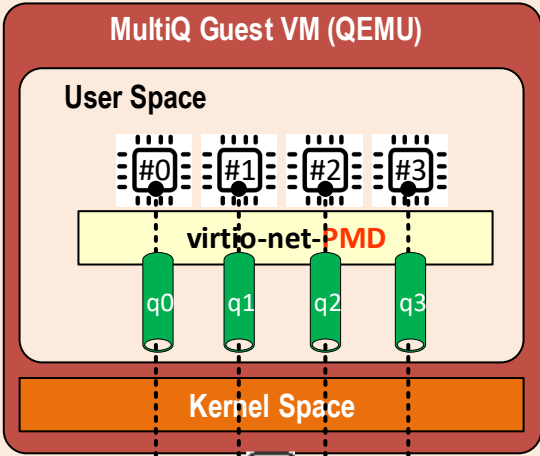
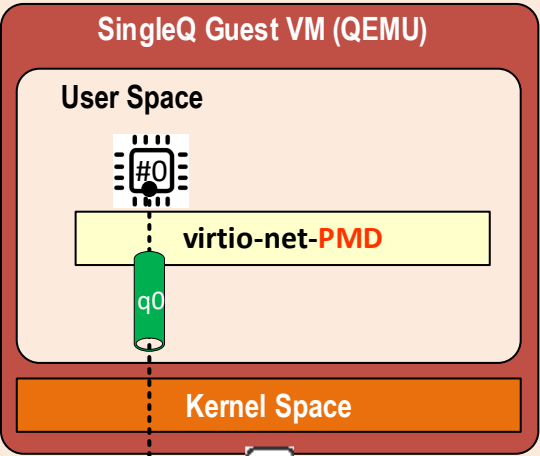
Title: DPDK End to End Packet Processing

version: 1.0

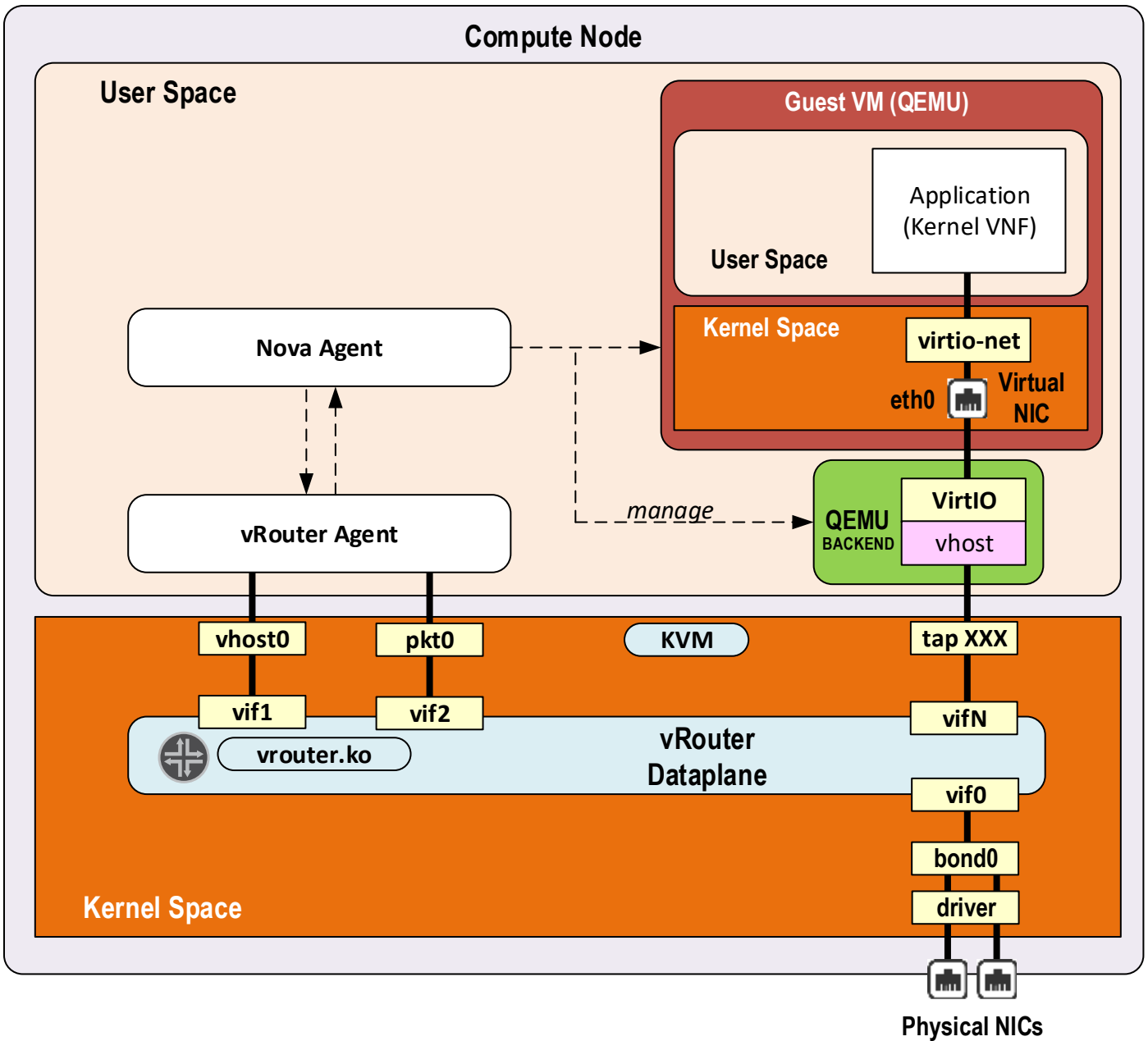
date: 27/03/2020

Compute Node

User Space



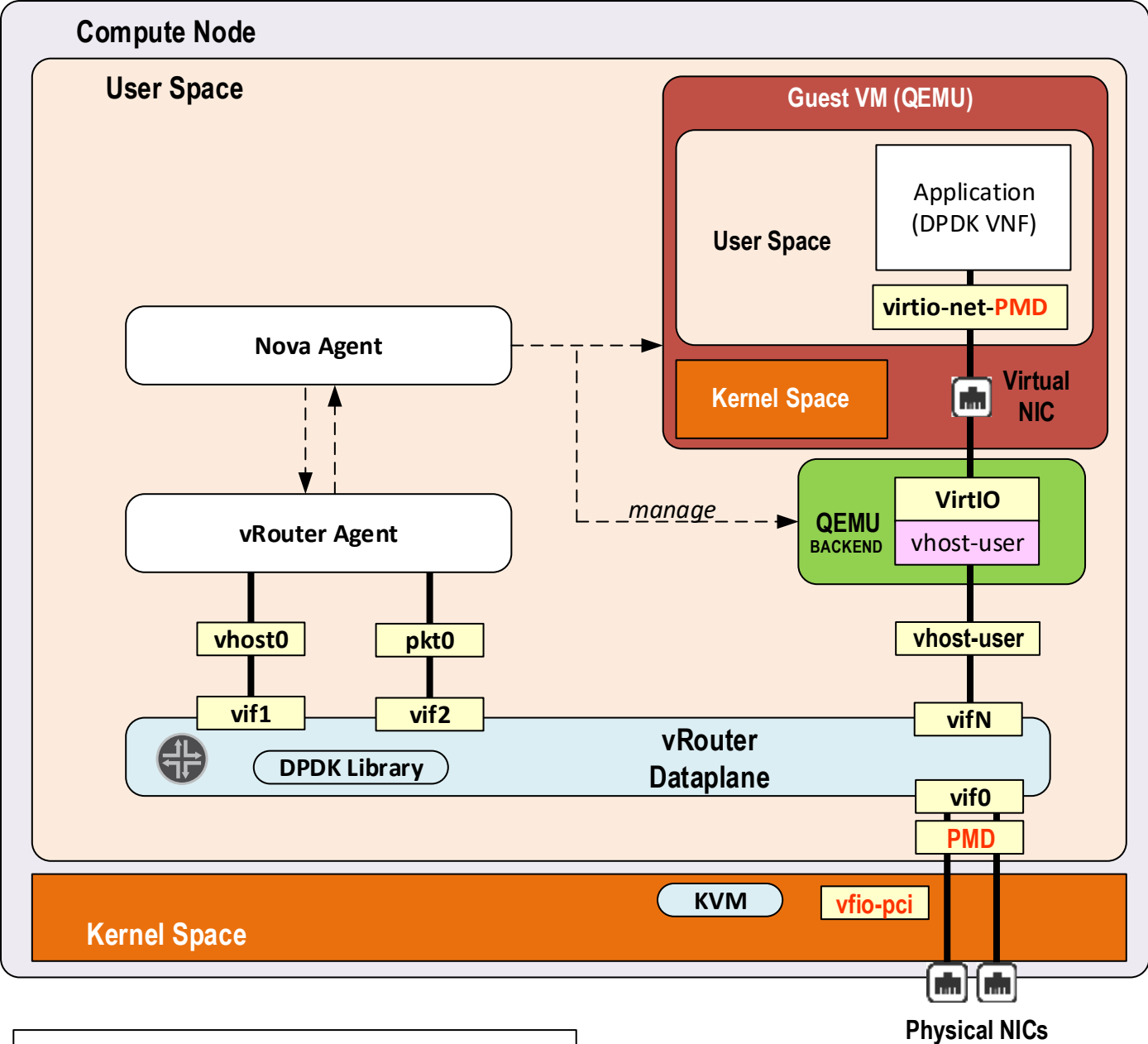
Title: SingleQ versus MultiQ VNF	
version: 1.0	date: 27/03/2020



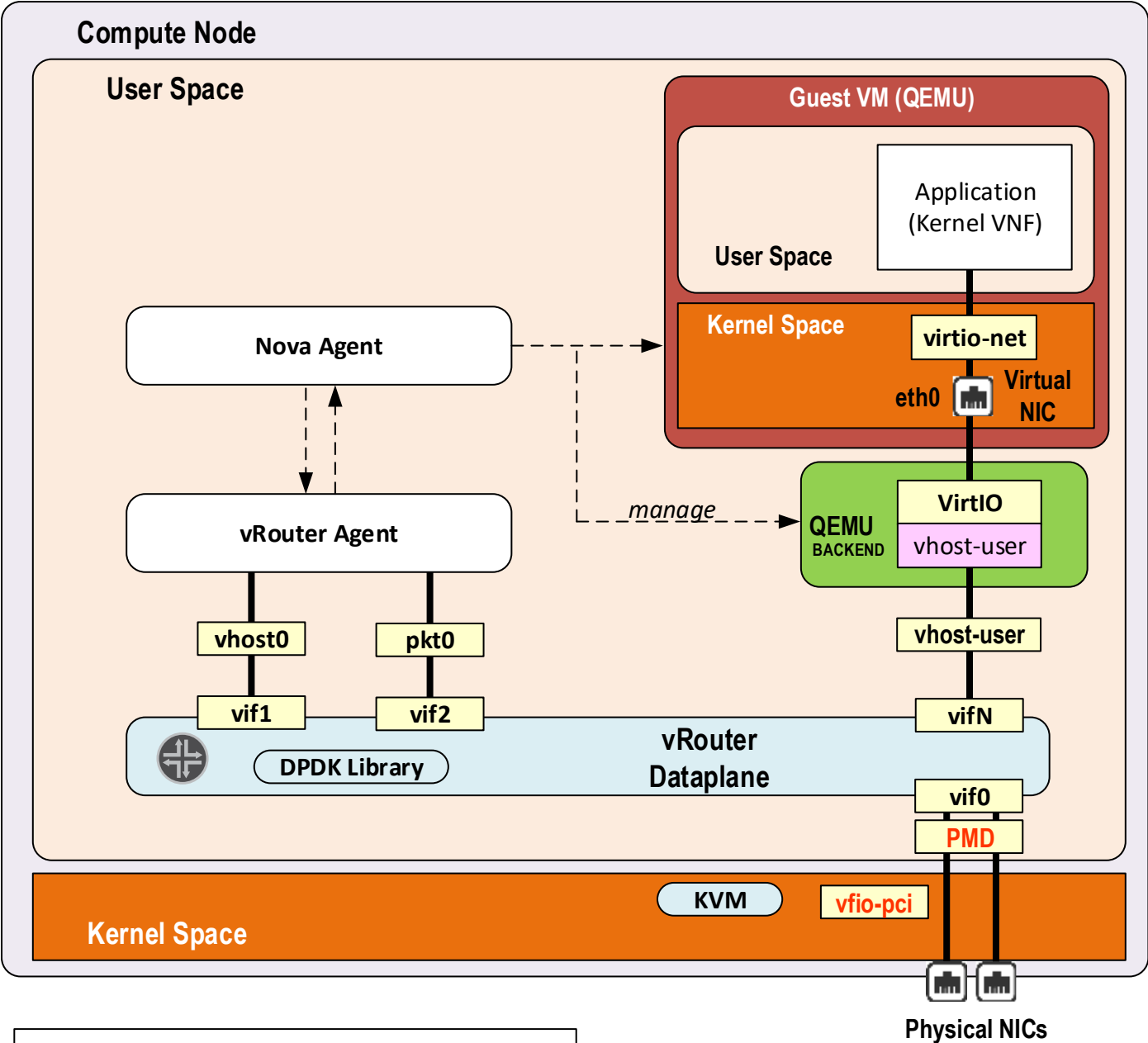
Title: Kernel vRouter details

version: 1.0

date: 25/03/2020



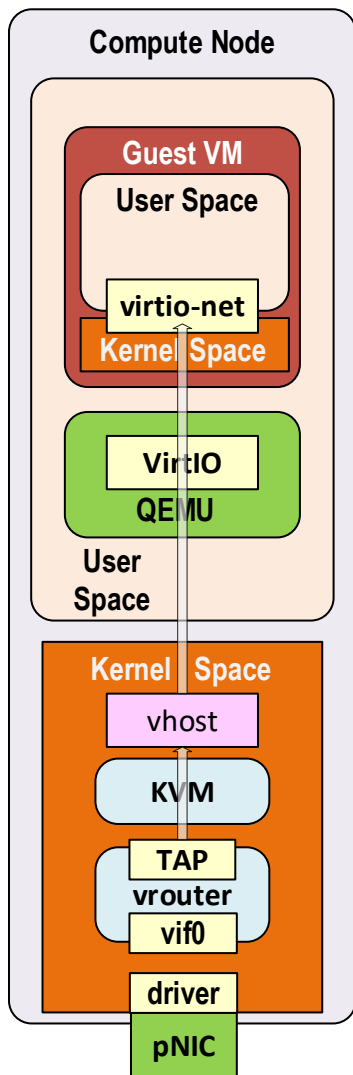
Title: DPDK vRouter details (DPDK App)	
version: 1.0	date: 25/03/2020



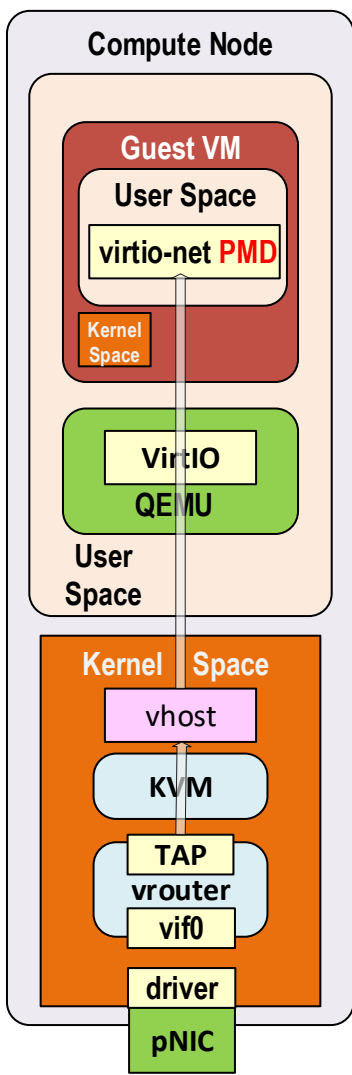
Title: DPDK vRouter details (Kernel App)	
version: 1.0	date: 25/03/2020

Kernel Mode vRouter

Virtio-net (Kernel)

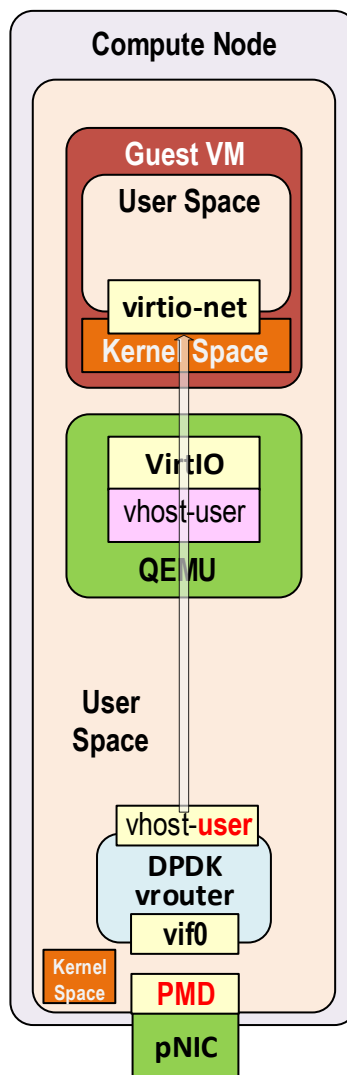


DPDK (user)

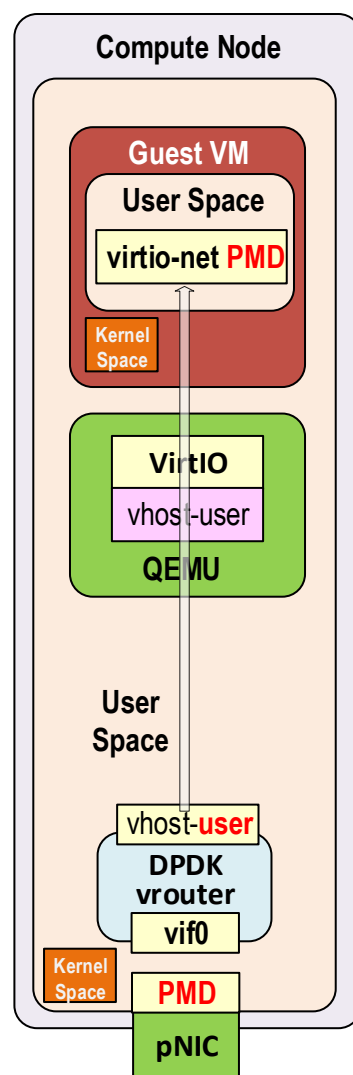


DPDK vRouter

Virtio-net (Kernel)



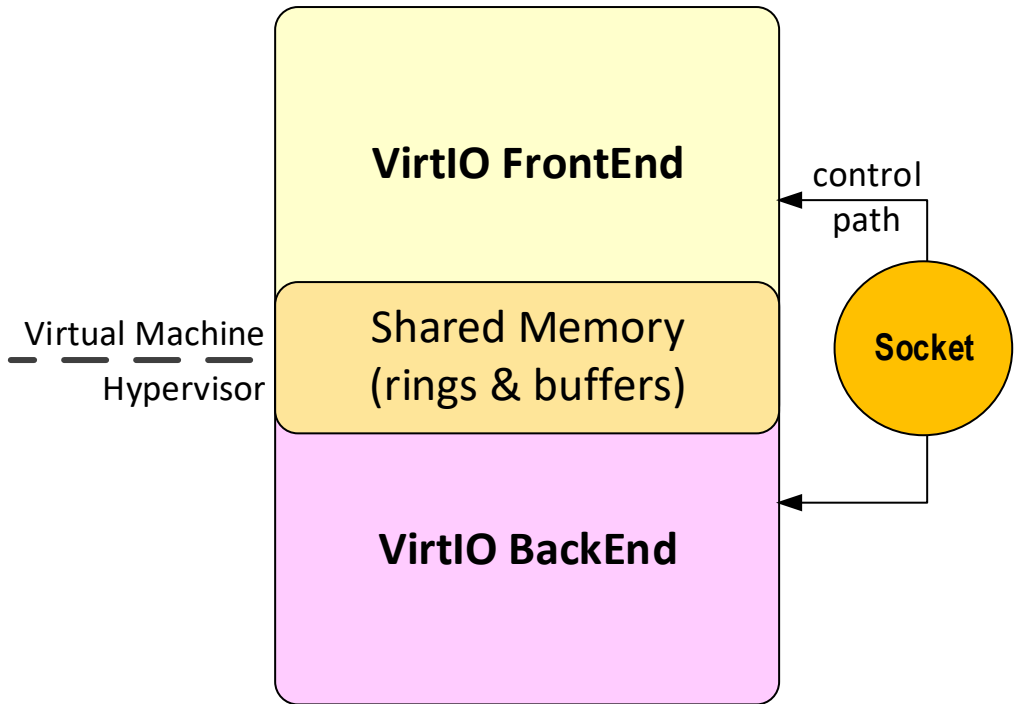
DPDK (user)



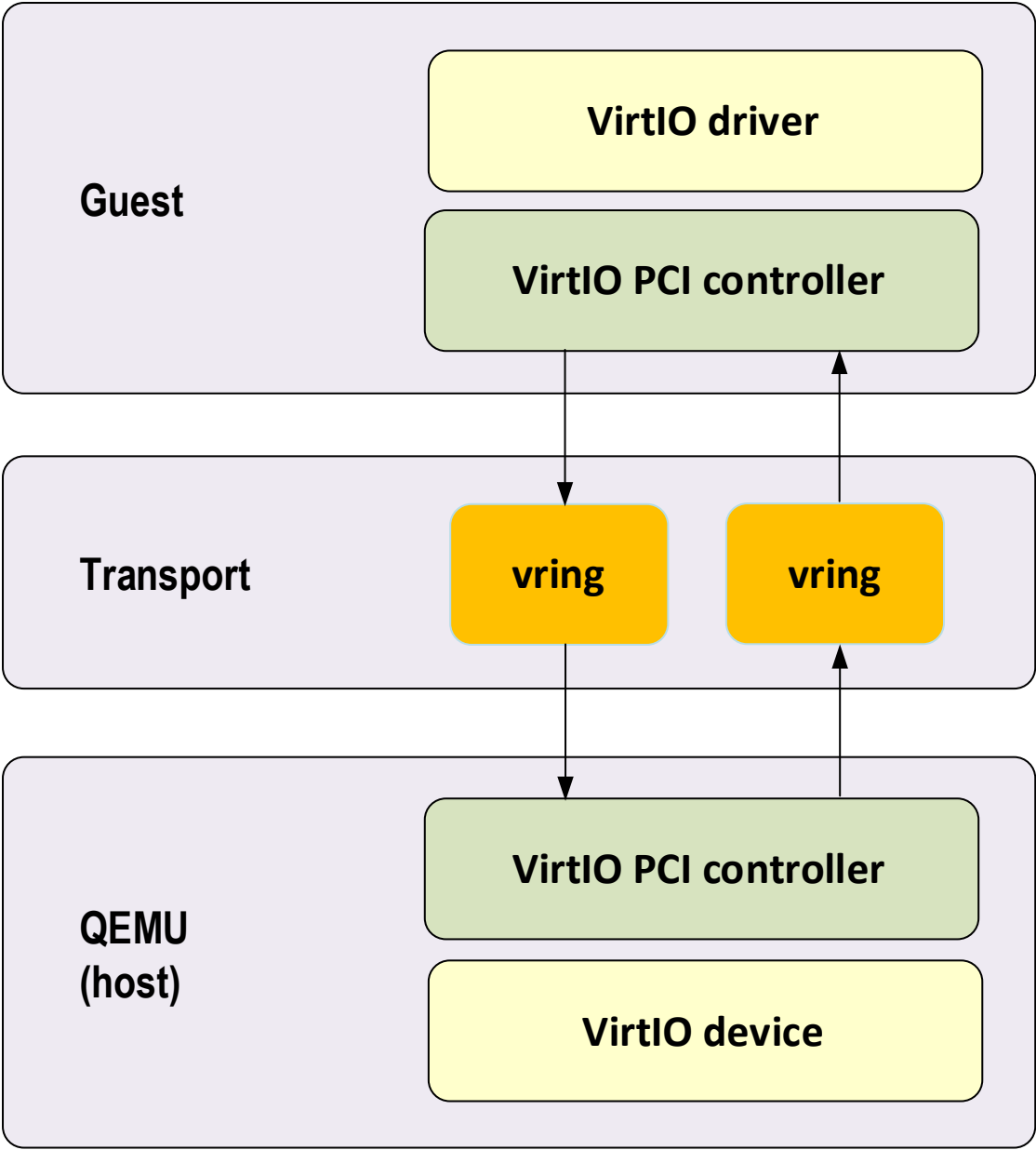
Title: Kernel vRouter All Cases

version: 1.1

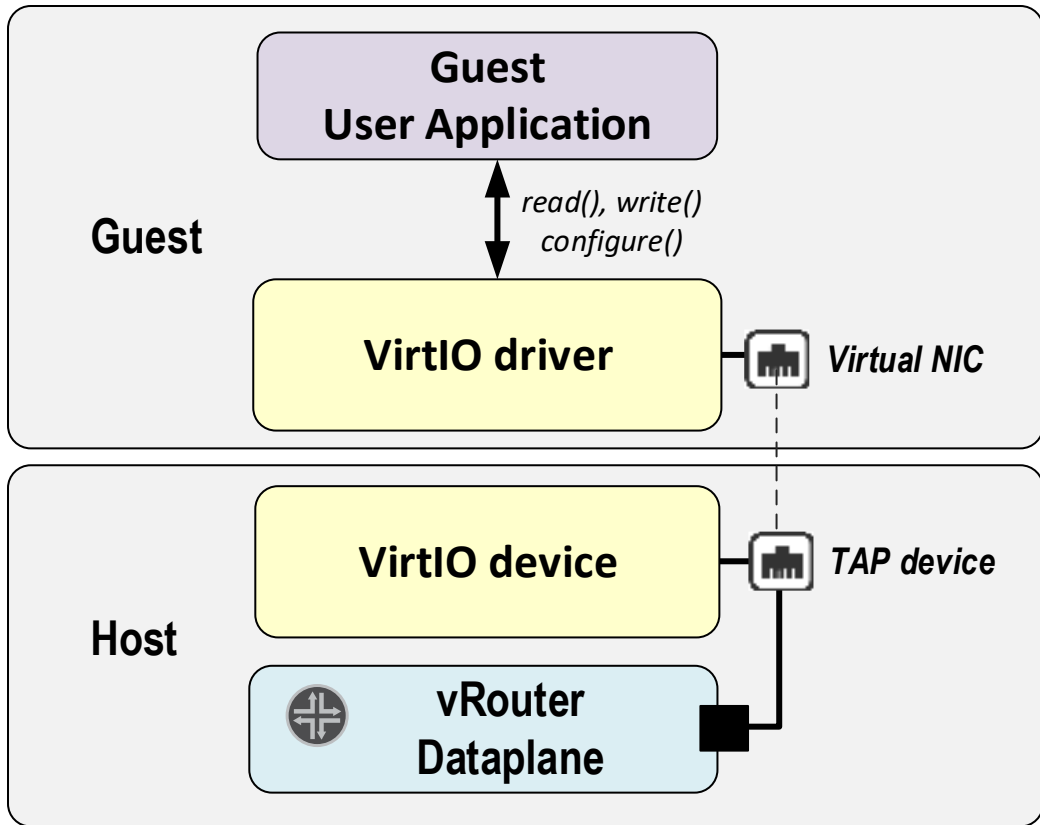
date: 22/05/2020



Title: VirtIO overview	
version: 1.0	date: 25/03/2020



Title: VirtIO driver and device	
version: 1.0	date: 02/06/2020



Title: Virtio & virtual infrastructure

version: 1.0

date: 02/06/2020

Compute Node

User Space

QEMU process

Guest VM

User Space

Kernel Space

**virtio-net driver
(virtio frontend)**

Hardware

eth0 Virtual
NIC

Shared Memory

virtqueues

Buffers

TX
Ring

RX
Ring

**virtio-net-device
(virtio backend)**

virtio datapath interface

*QEMU readmsg()
writemsg()*

irq

vmexit

*Notification
Messages*

*PCI
BAR*

Kernel Space

TAP

kvm.ko

Title: VirtIO Net

version: 1.0

date: 04/06/2020

Compute Node

User Space

QEMU process

Guest VM

User Space

Kernel Space

virtio-net driver
(virtio frontend)

Hardware

eth0 Virtual
NIC

QEMU Shared Memory

virtqueues

Buffers

TX
Ring

RX
Ring

PCI
BAR

virtio-net-device
(virtio backend)

virtio datapath interface

IRQ FD

ioevent FD

Notification
Messages

Kernel Space

QEMU
readmsg()
writemsg()

kvm.ko

TAP

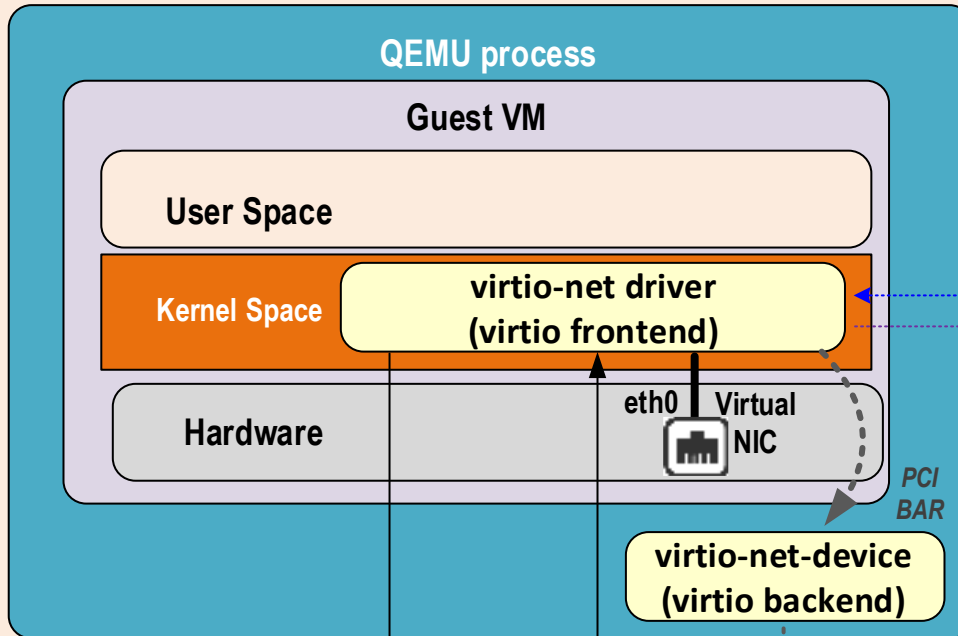
Title: Virtio-Net

version: 1.0

date: 27/09/2020

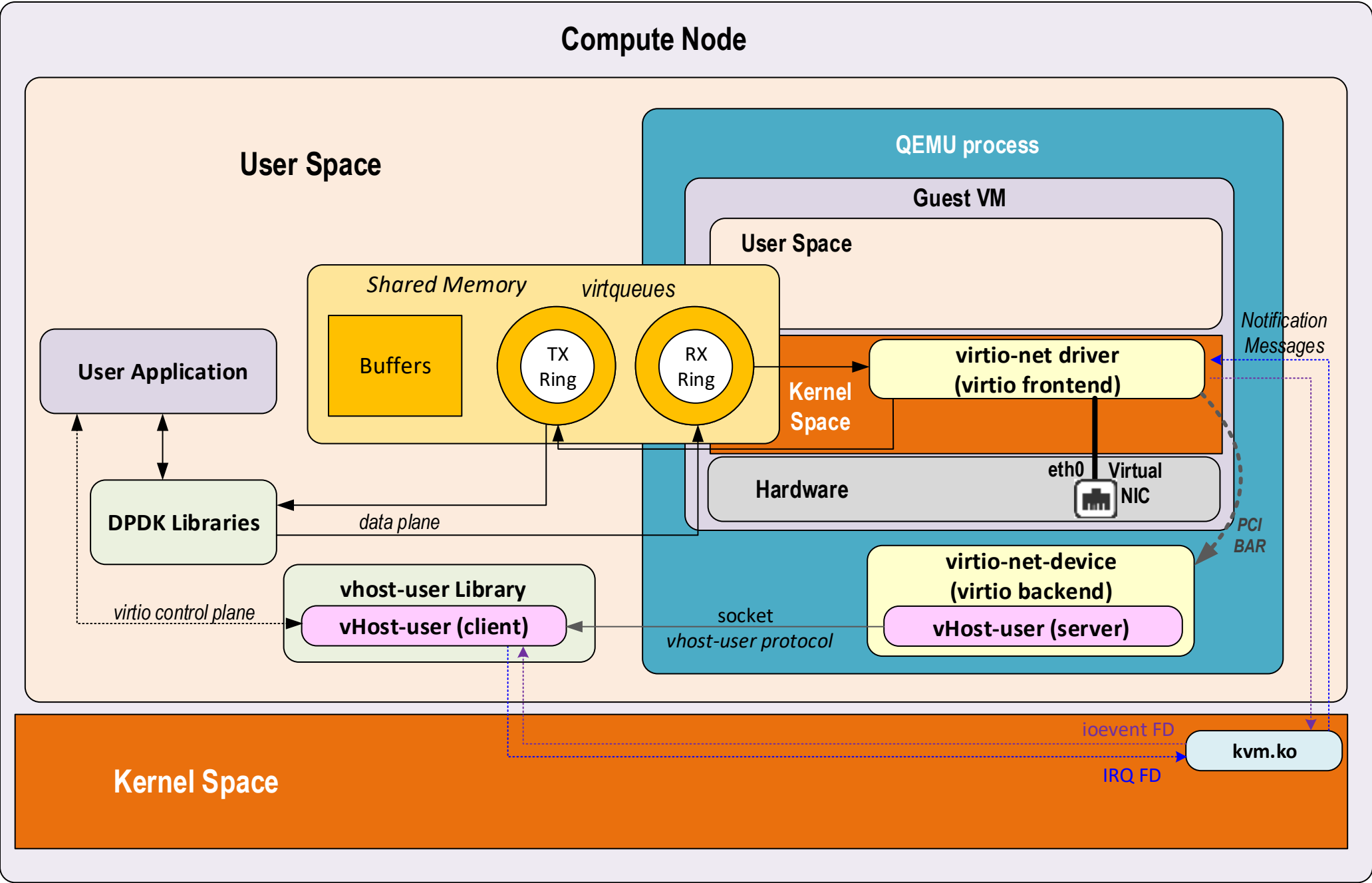
Compute Node

User Space

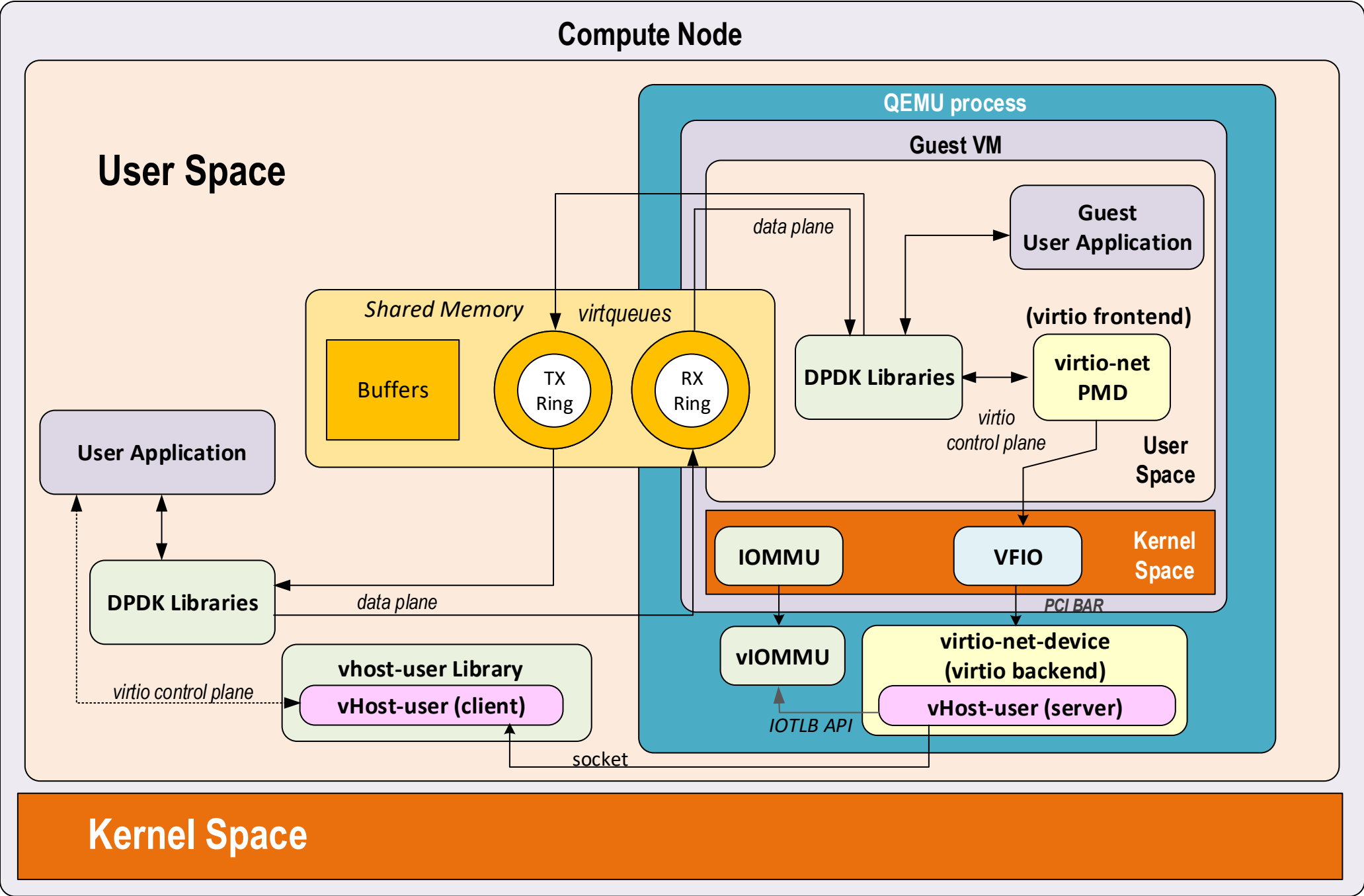


Kernel Space

Title: vHost Net	
version: 1.0	date: 04/06/2020



Title: vHost User	
version: 1.0	date: 04/06/2020



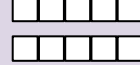
Title: VirtIO PMD (full DPDK)	
version: 1.0	date: 04/06/2020

Compute Node

User Space

User Application

Application Data



System Call Interface

Socket API

TCP/IP Stack

NIC Driver

RX

TX

DMA
transfer

Hard
Interrupts

Ring
buffers

Soft
IRQ

Kernel
SoftIRQ
Thread

Copy

Copy

Hardware

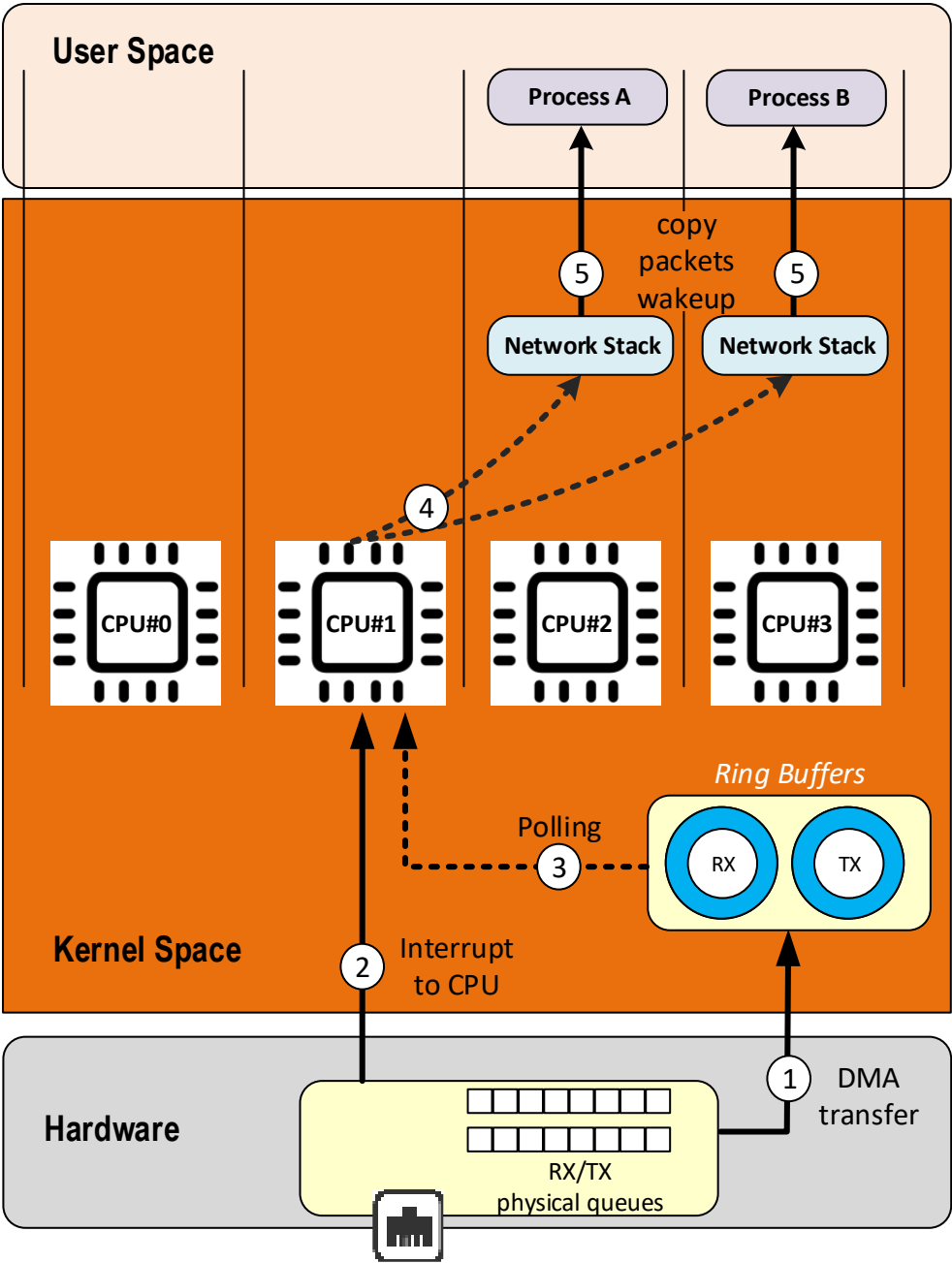
RX/TX
physical queues

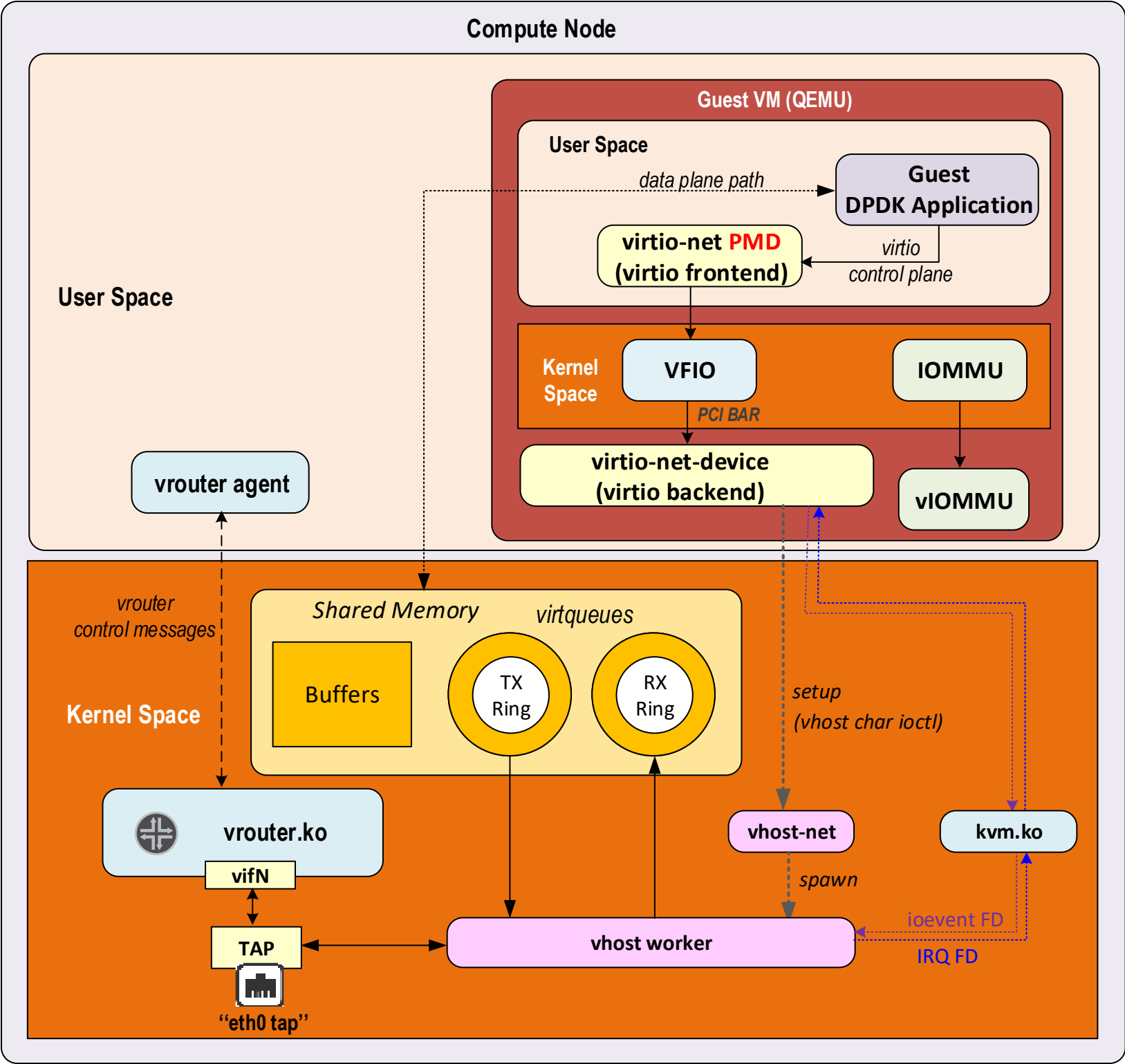


Title: Linux TCP/IP Stack

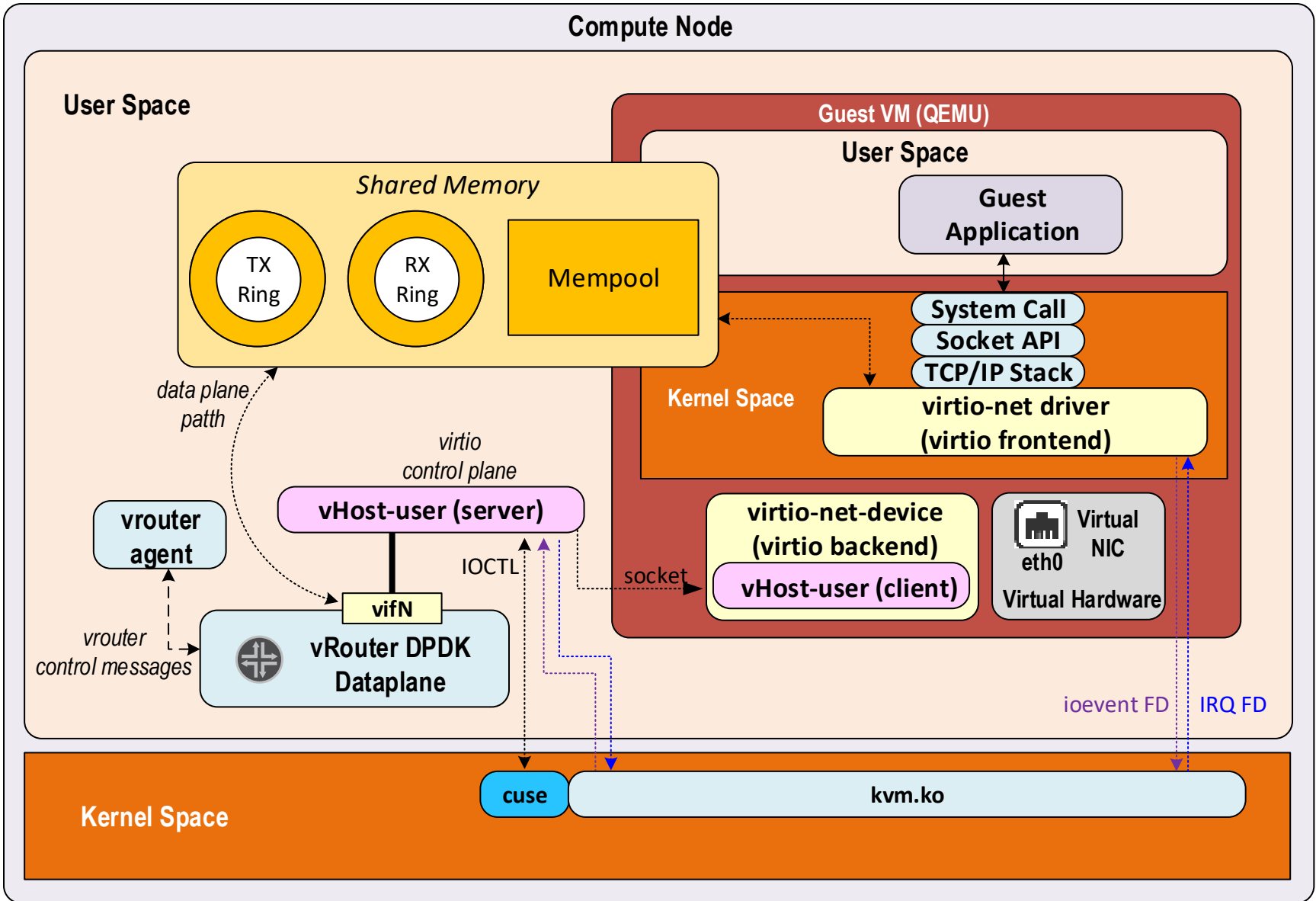
version: 1.0

date: 23/09/2020



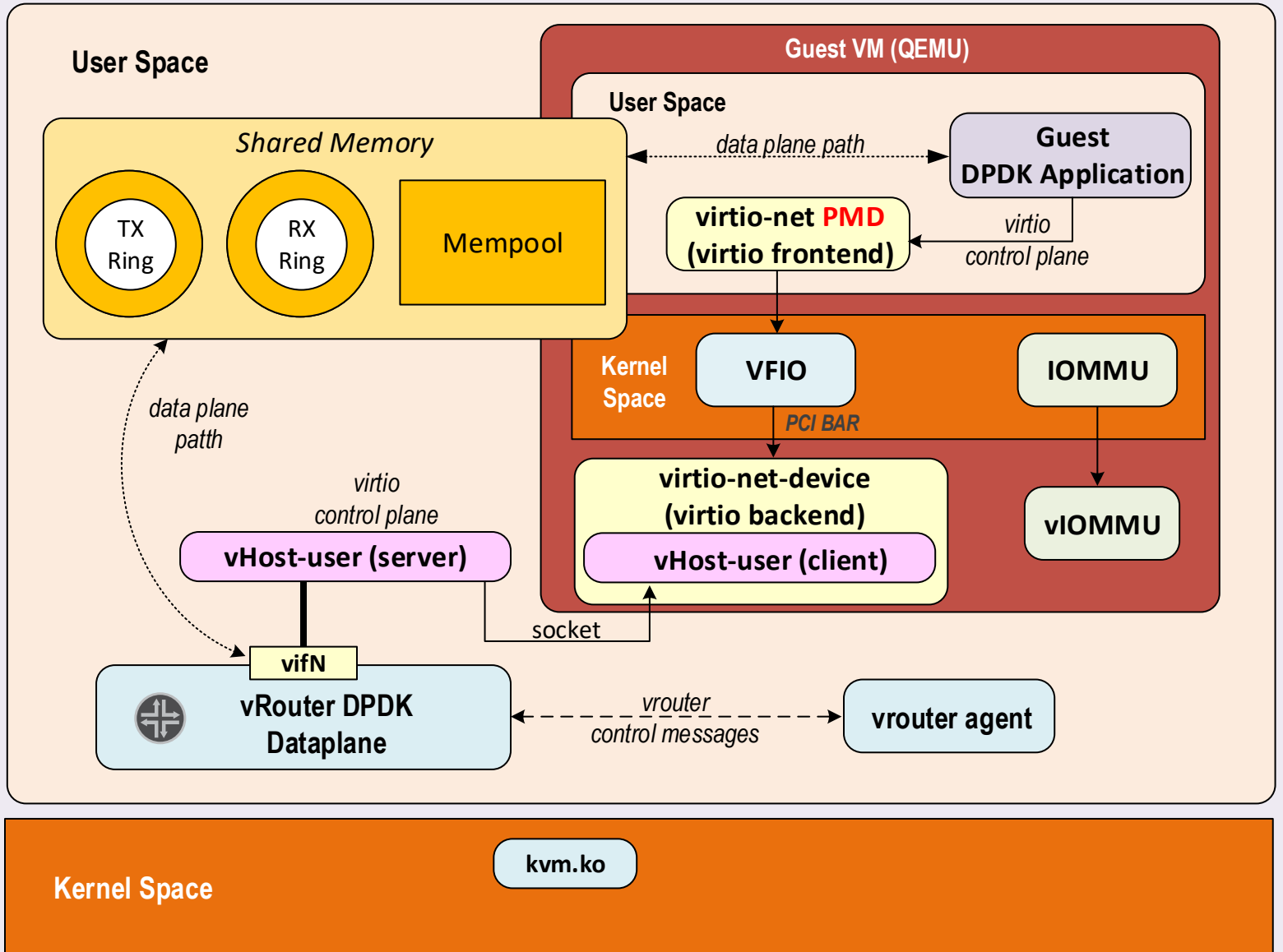


DPDK vrouter (Kernel mode App)	
version: 1.0	date: 27/09/2020



DPDK vrouter (Kernel mode App)	
version: 1.0	date: 04/06/2020

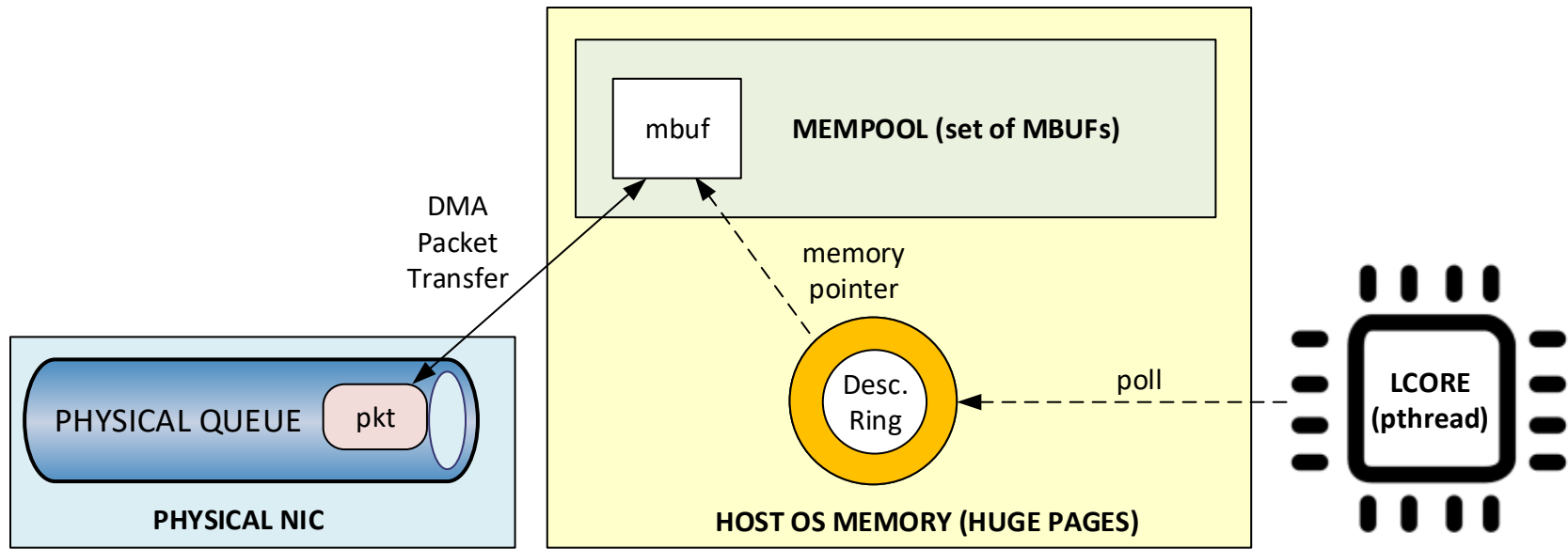
Compute Node



Title: DPDK vrouter (DPDK mode App)

version: 1.0

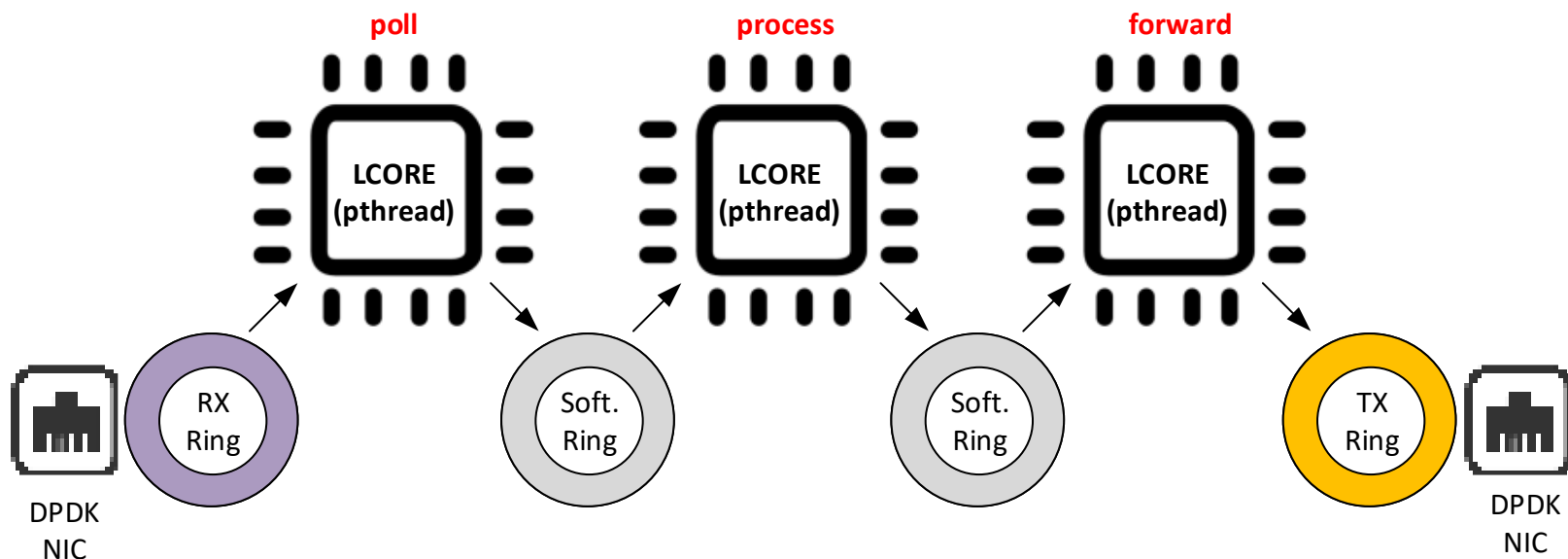
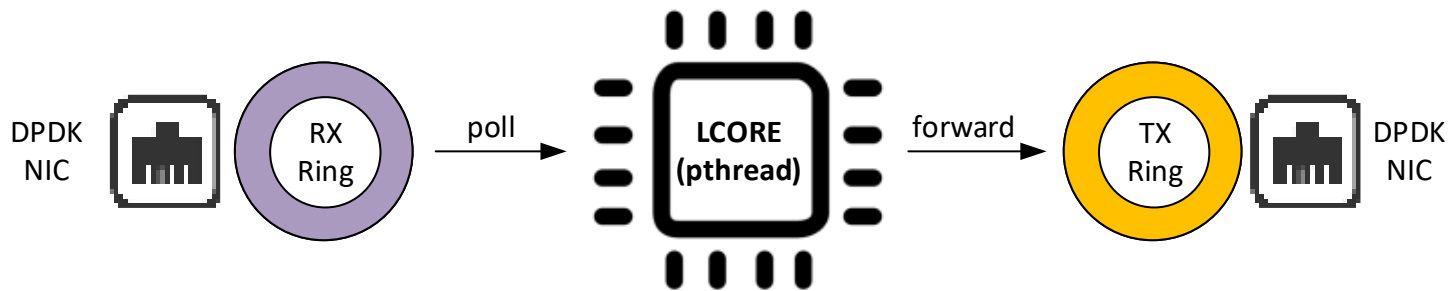
date: 04/06/2020



Title: DPDK High Level Components

version: 1.0

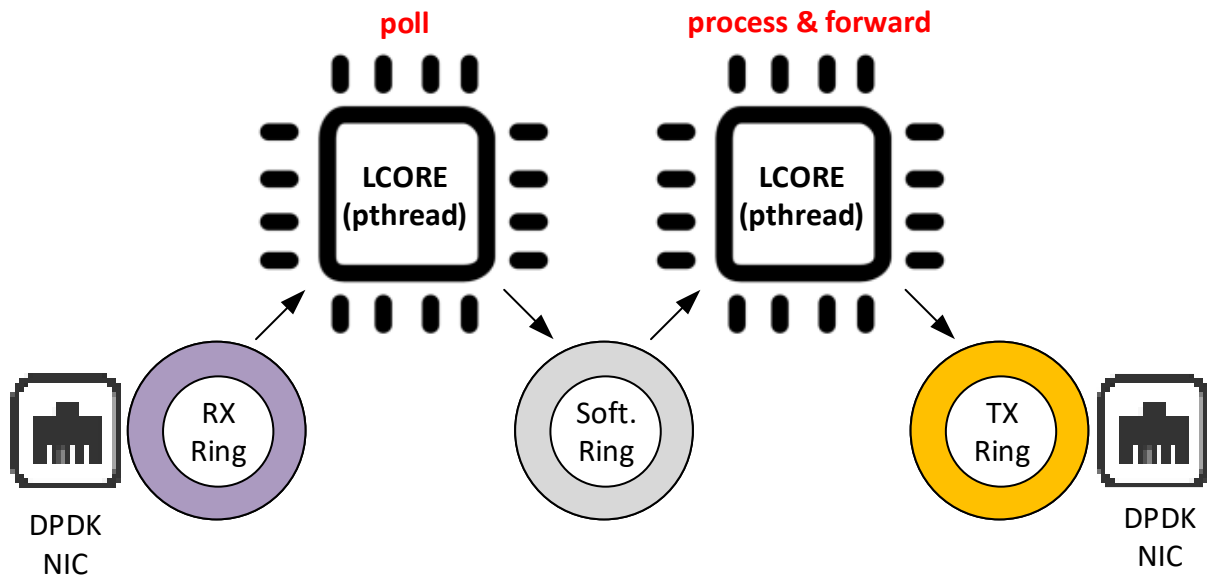
date: 02/06/2020



Title: DPK Packets processing models

version: 1.0

date: 02/06/2020

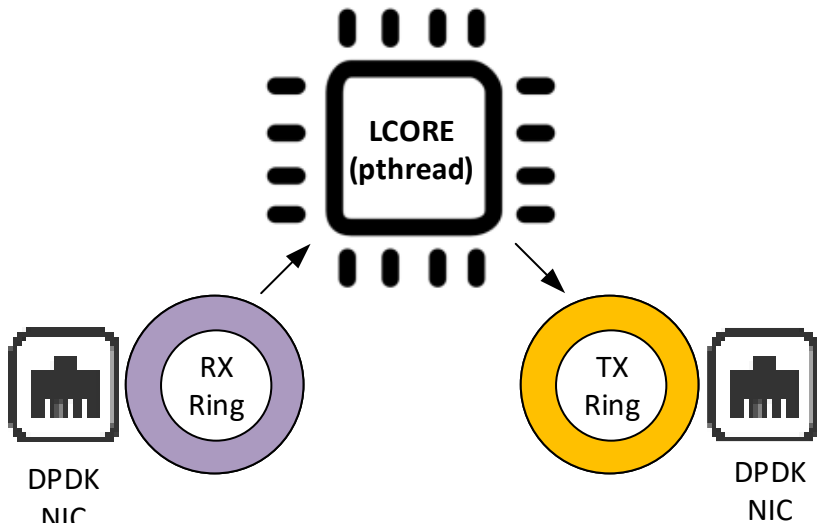


Title: Contrail vrouter DPKD processing model

version: 1.0

date: 15/06/2020

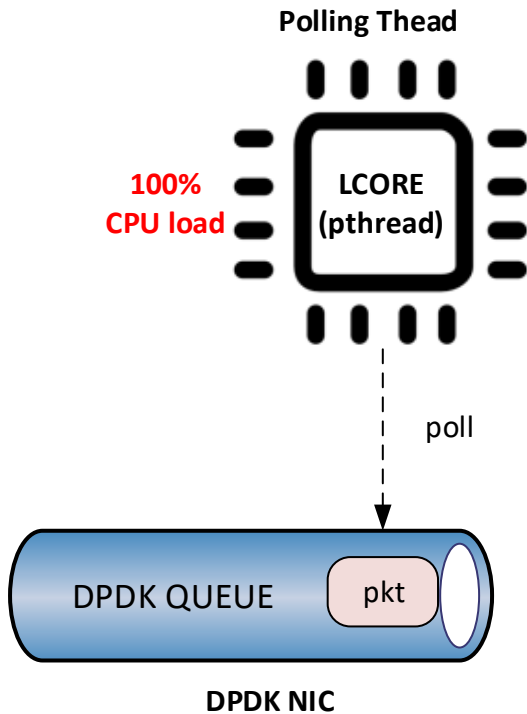
poll, process & forward



Title: Contrail vrouter DDPK new processing model

version: 1.0

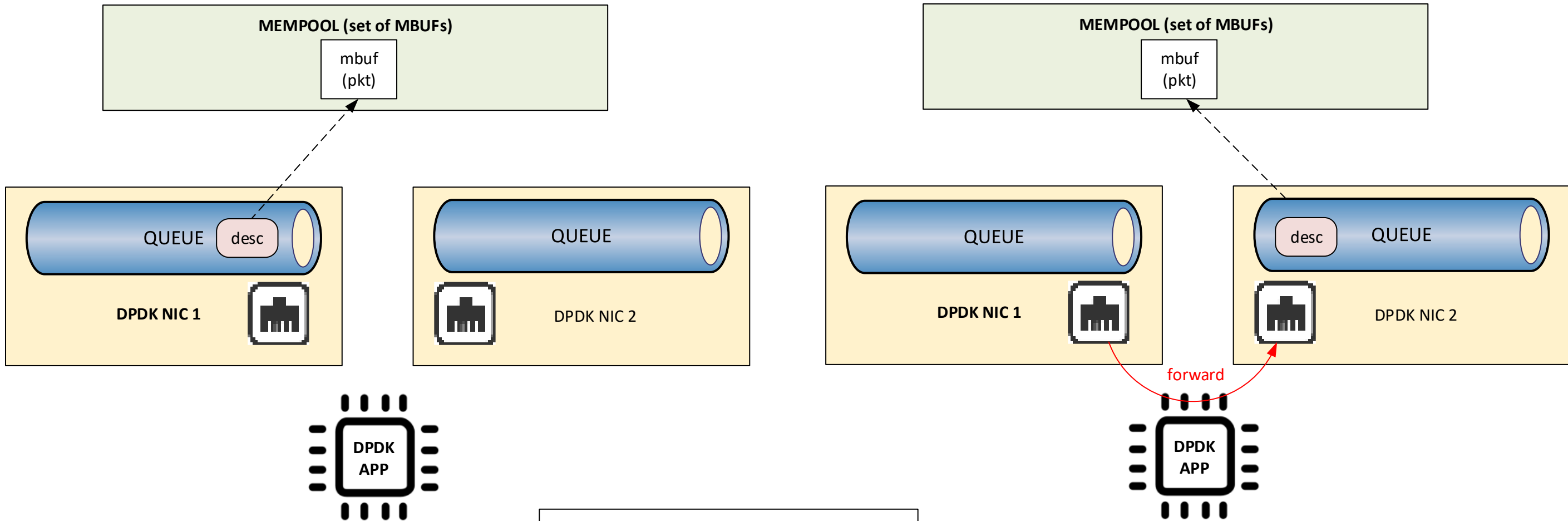
date: 15/06/2020



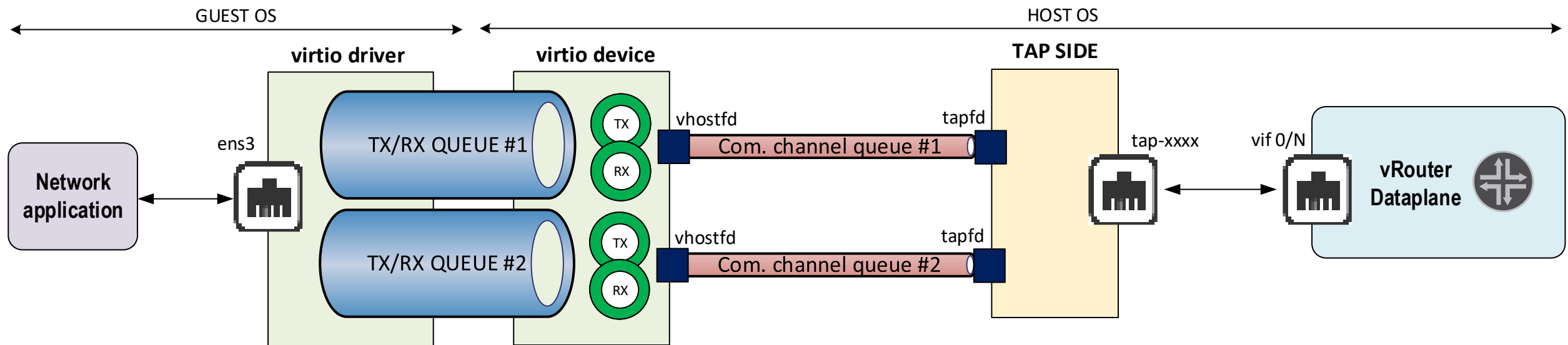
Title: DPDK polling

version: 1.0

date: 02/06/2020

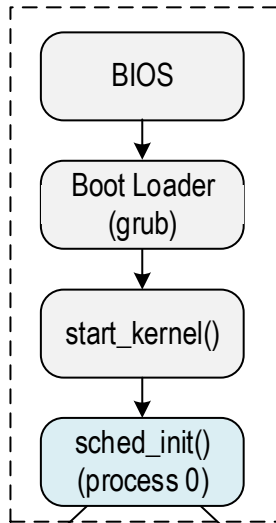


Title: DPDK Packet Forwarding	
version: 1.0	date: 03/06/2020

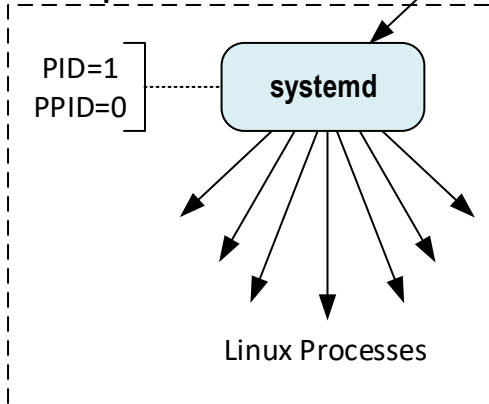


Title: vHost FDs	
version: 1.0	date: 30/09/2020

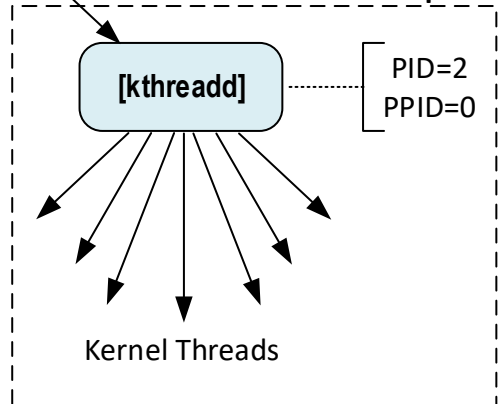
OS startup process



User Space



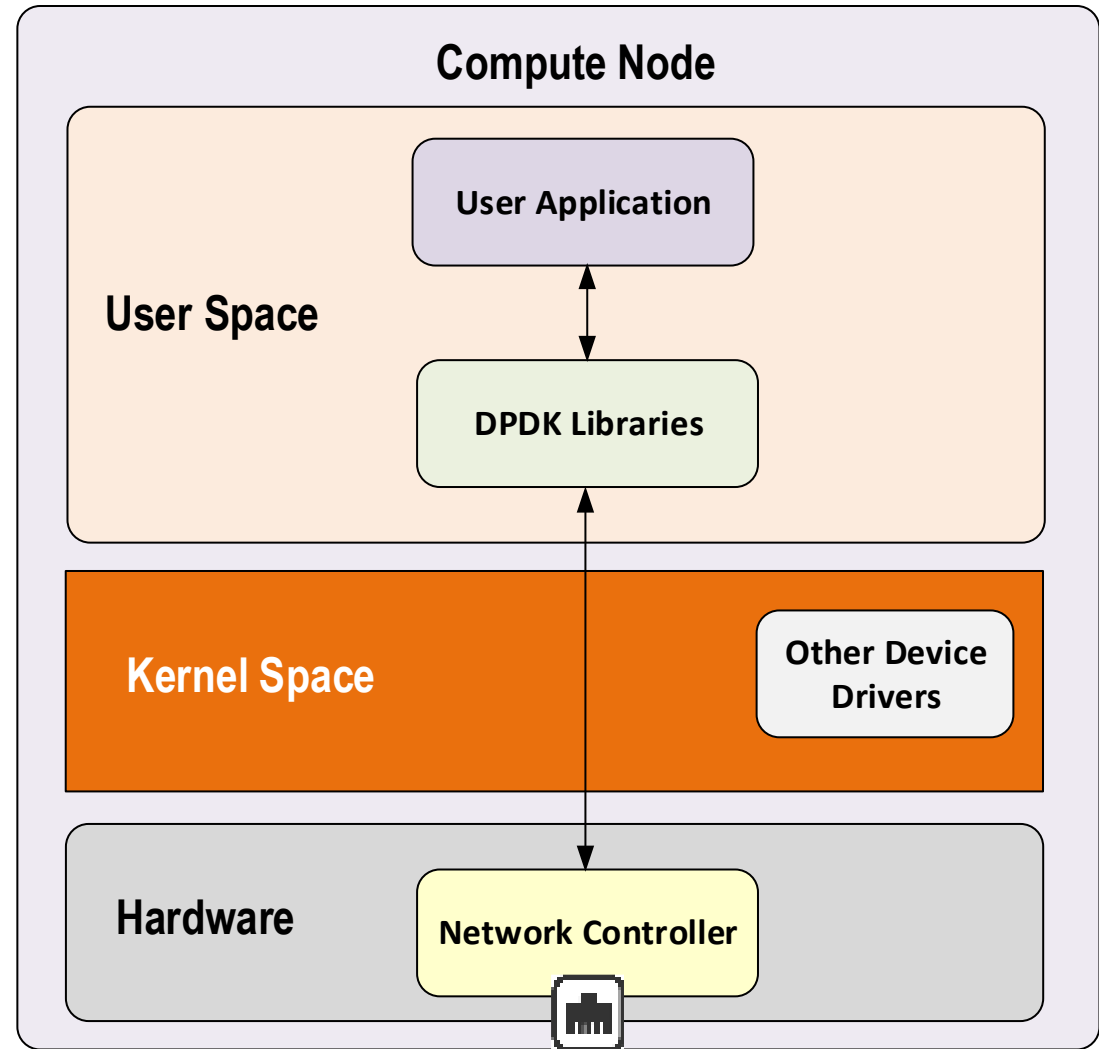
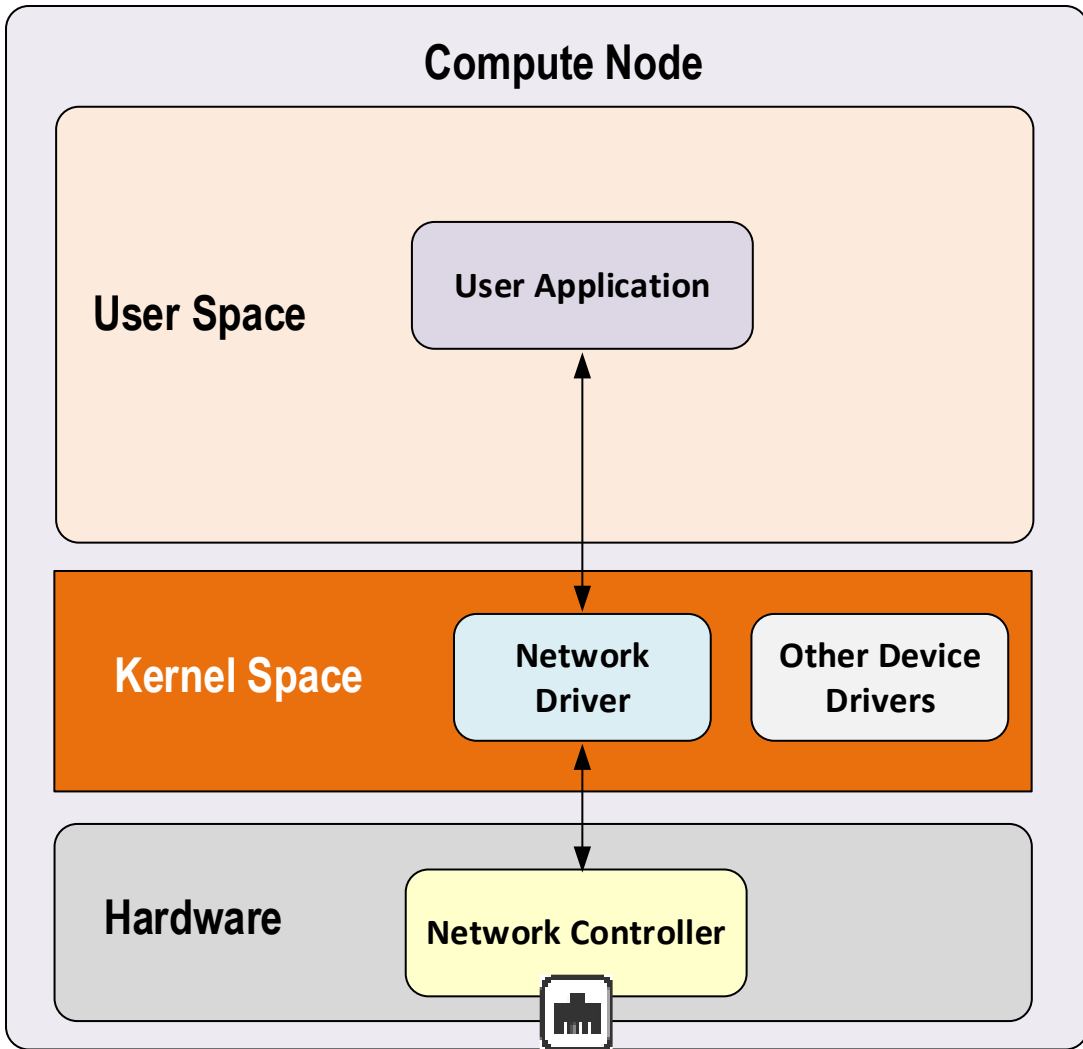
Kernel Space



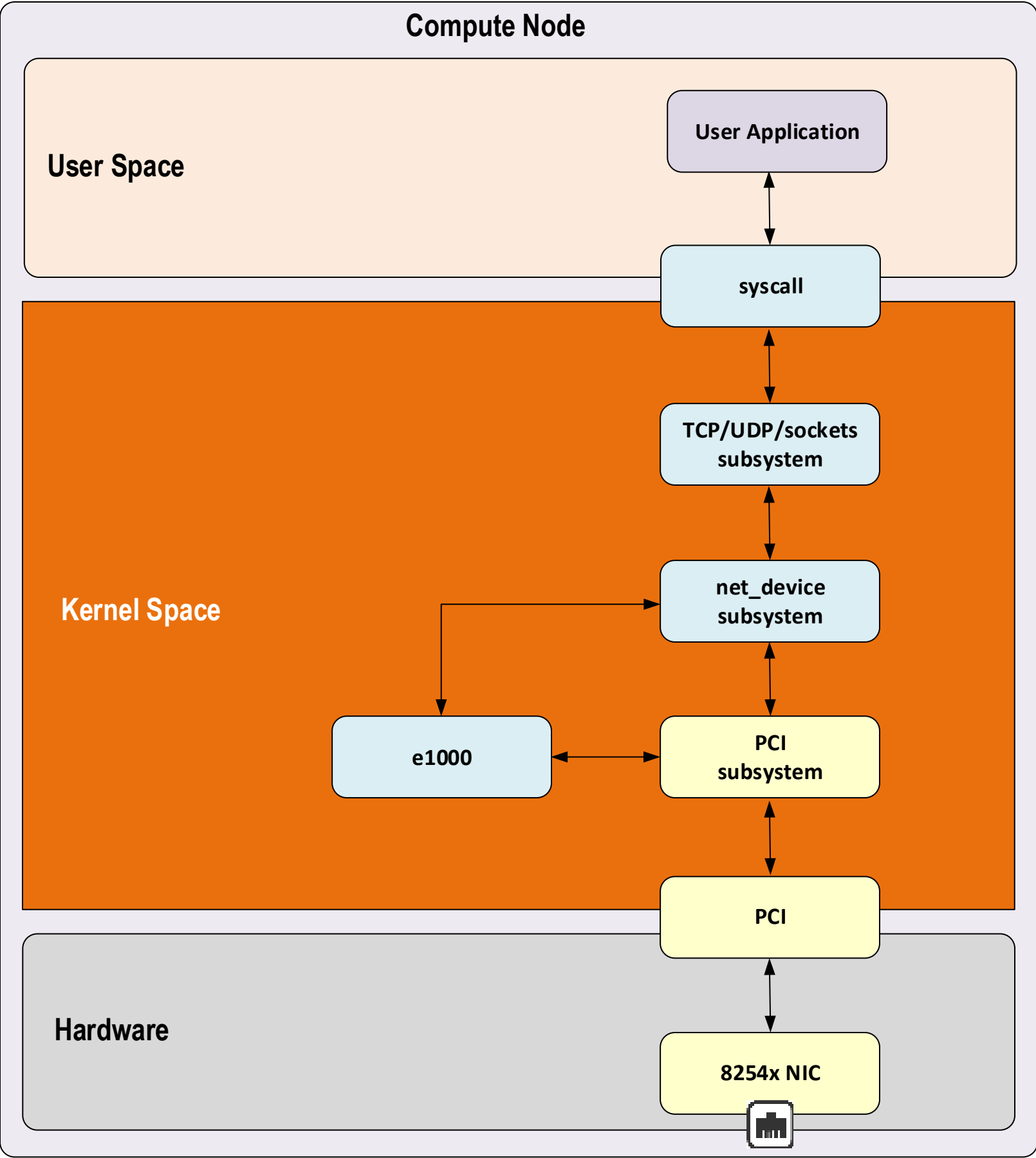
Title: Linux Startup

version: 1.0

date: 30/09/2020



Title: DPDK versus NIC Kernel Overview	
version: 1.0	date: 30/05/2020



Title: NIC Kernel Driver	
version: 1.0	date: 30/05/2020

Compute Node

User Space

User Application

Socket
API

Kernel Space

Driver

Ring
buffers

DMA transfer

Hardware

RX/TX
physical queues



Title: NIC Kernel Driver – Packet Processing

version: 1.0

date: 30/05/2020

Compute Node

User Space

User Application

ioctl()
(Socket API)

Kernel Space

syscalls

Virtual File
System (VFS)

Kernel API
(copy_for_user,
io_write)

Hardware

Configuration
Registers



Title: NIC Kernel Driver – Configuration

version: 1.0

date: 30/05/2020

Compute Node

User Space

User Application

DPDK
API

Ring
buffers

DMA transfer

Kernel Space

user space
device enabler

DMA transfer

Hardware

RX/TX
physical queues



Title: NIC PMD Driver – Packet Processing

version: 1.0

date: 02/06/2020

Compute Node

User Space

User Application

DPDK API

PMD

Memory
Mapped
Registers

Kernel Space

user space
device enabler

Hardware

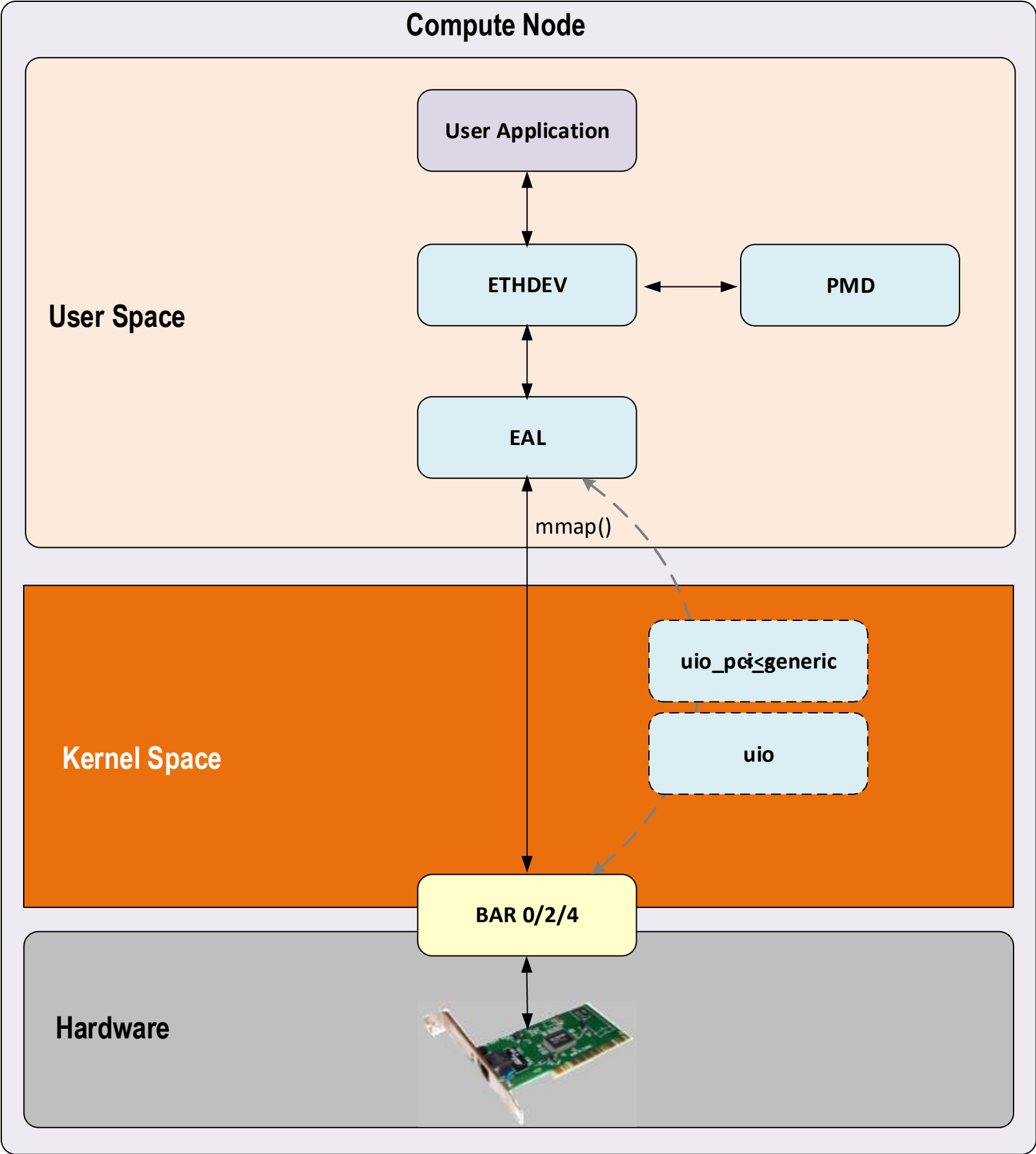
Configuration
Registers



Title: NIC PMD Driver – Configuration

version: 1.0

date: 02/06/2020



Title: DPDK NIC Configuration	
version: 1.0	date: 02/06/2020

Compute Node

User Space

Guest VM

User Space

User Application

DPDK

Vendor specific PMD

Read/Write

Assign

Kernel Space

VFIO

viOMMU

Virtual
PCI NIC

Emulated
Hardware



Assign

Huge Pages Memory

Guest
huge pages

map

NIC
registers and queues

map

Kernel Space

VFIO

IOMMU

Physical
PCI NIC

Hardware



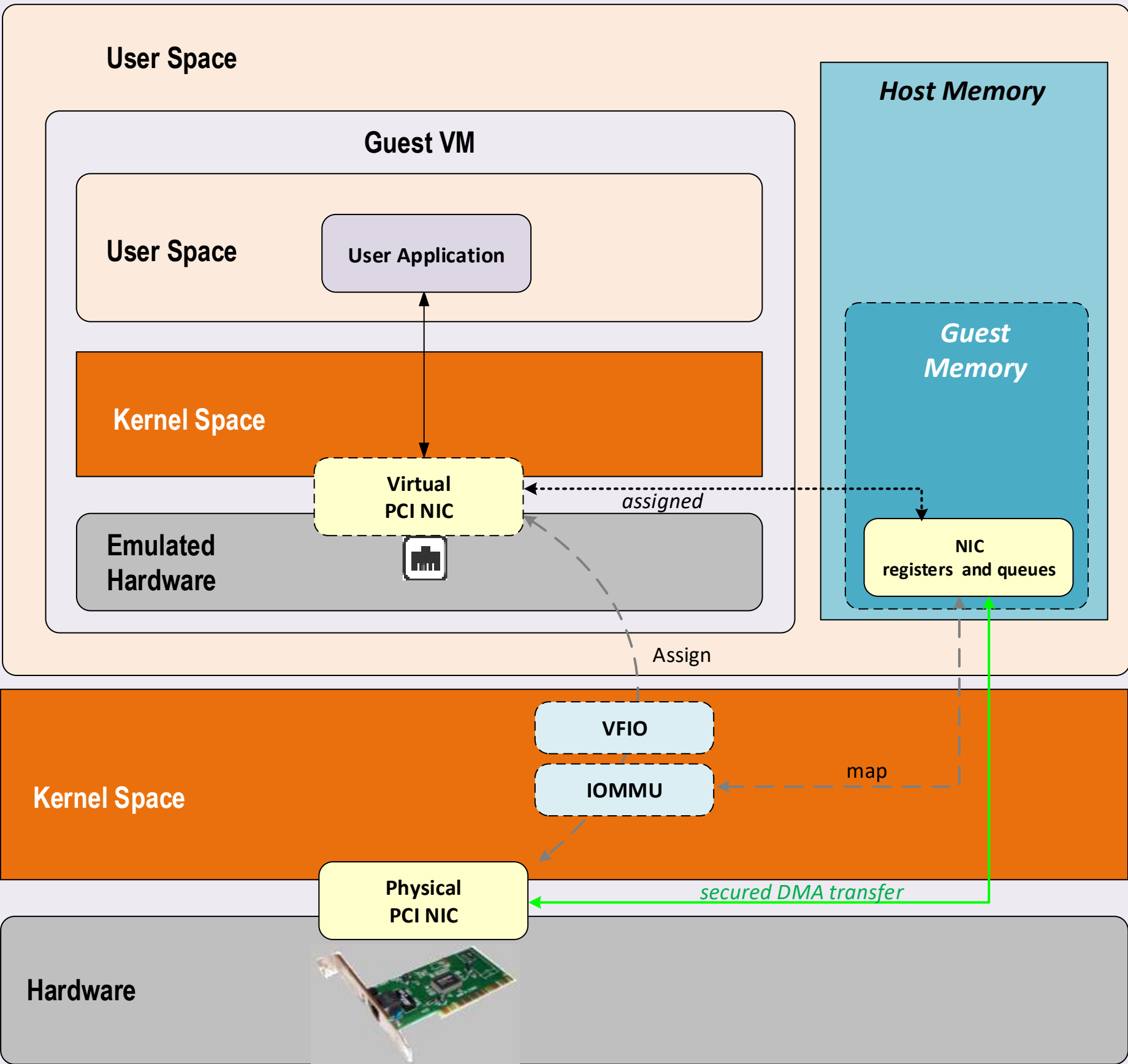
secured DMA transfer

Title: DPDK VFIO and IOMMU (PCI passthrough)

version: 1.0

date: 02/06/2020

Compute Node



Title: VFIO device assignment

version: 1.0

date: 02/06/2020

Compute Node

User Space

Guest VM

User Space

User Application

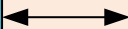
Kernel Space

Virtual
PCI NIC

Emulated
Hardware

QEMU process

LibVirt



KVM

Kernel Space

Physical
PCI NIC

Hardware



Title: QEMU-KVM-LibVirt

version: 1.0

date: 02/06/2020

Compute Node

User Space

Guest VM

User Space

User Application

DPDK
PCI NIC

Read/Write

Assign

Kernel Space

VFIO

viOMMU

Virtual
PCI NIC

emulated DMA transfer

Emulated
Hardware

Huge Pages Memory

*Guest
huge pages*

map

NIC
registers and queues

Kernel Space

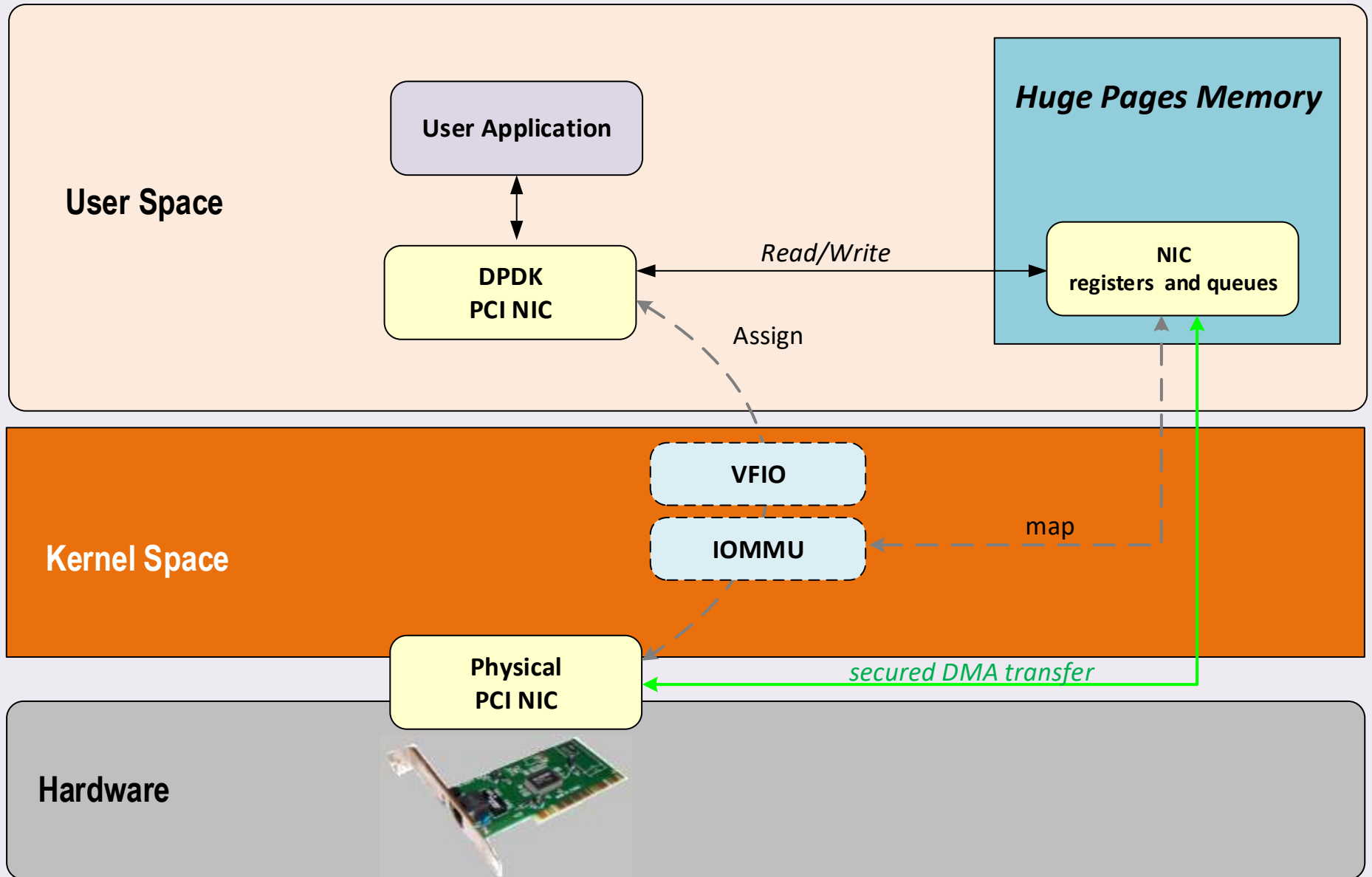
Hardware

Title: DPDK VFIO and emulated NIC

version: 1.0

date: 02/06/2020

Compute Node



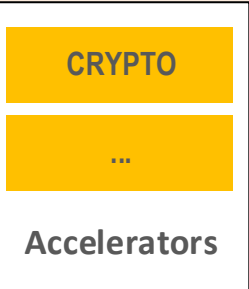
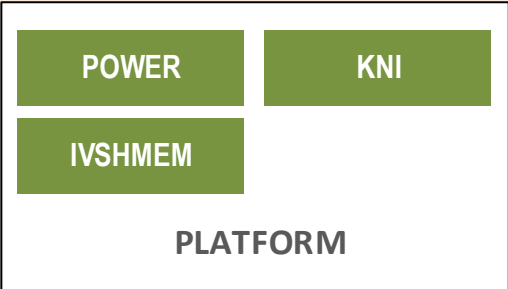
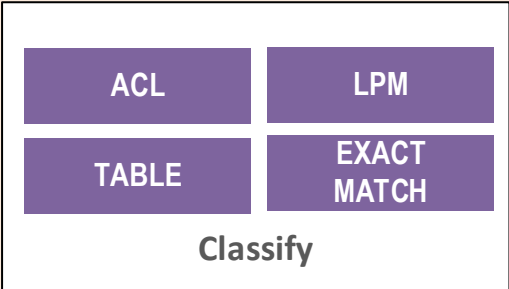
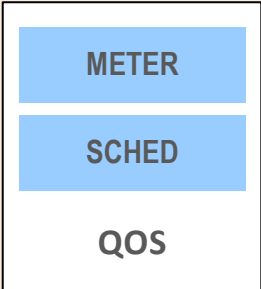
Title: DPDK VFIO and physical NIC

version: 1.0

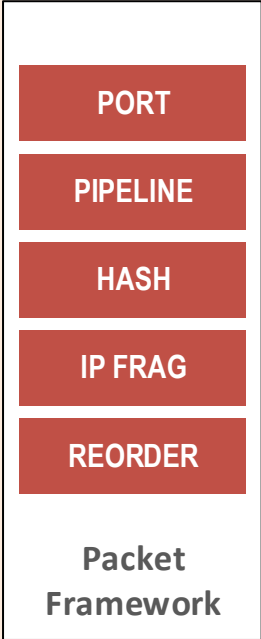
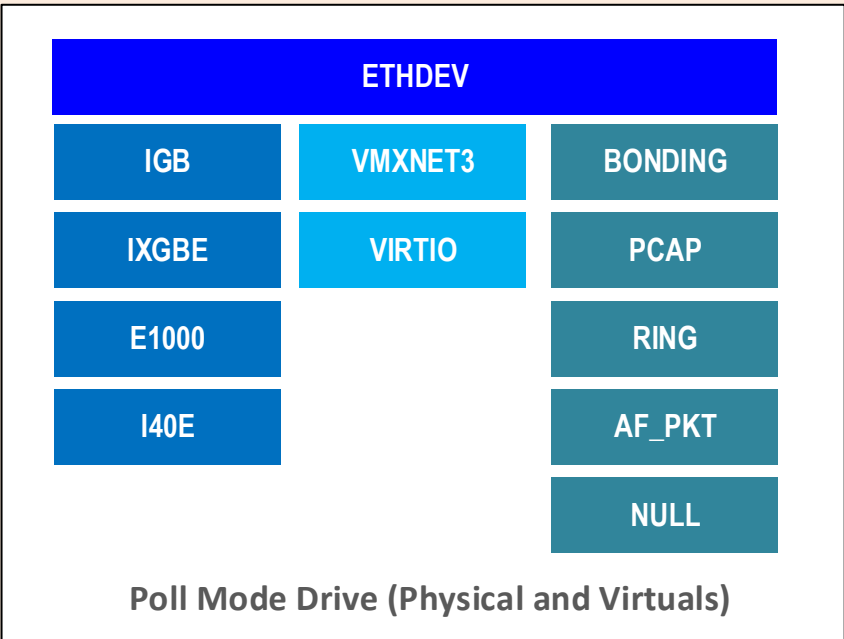
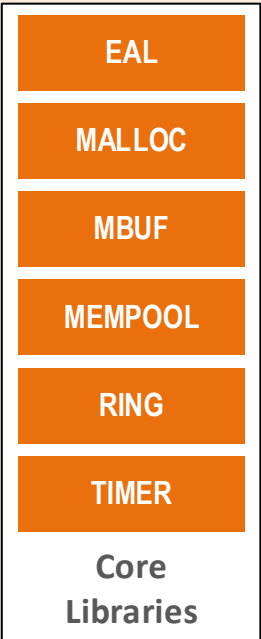
date: 02/06/2020

Compute Node

DPDK Application

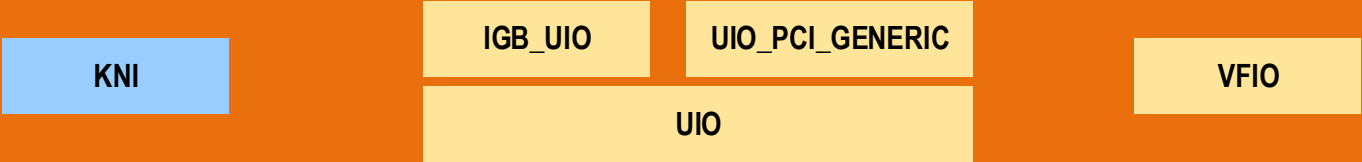


User Space



DPDK Library Framework

Kernel Space



Hardware

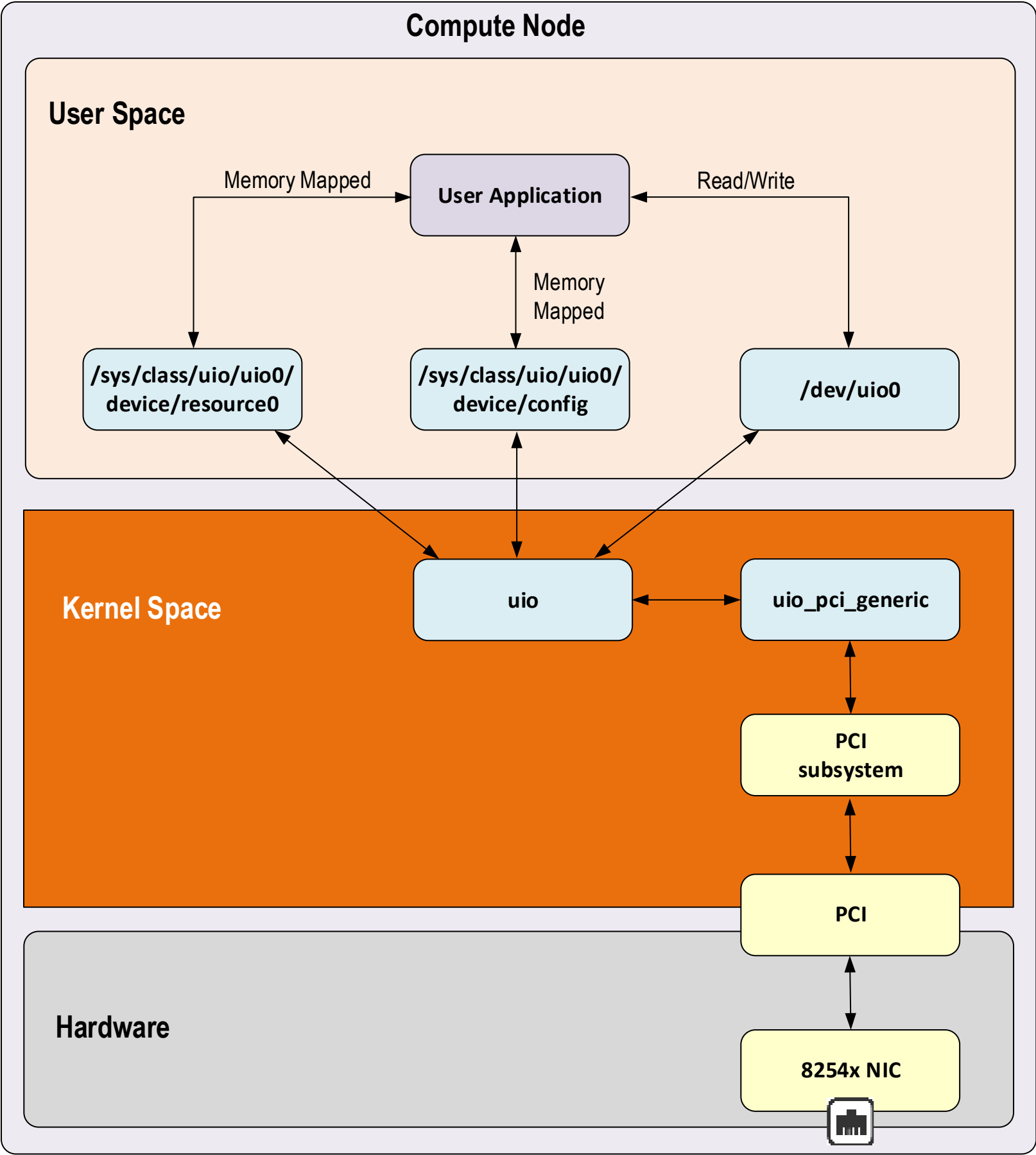
Network Controller

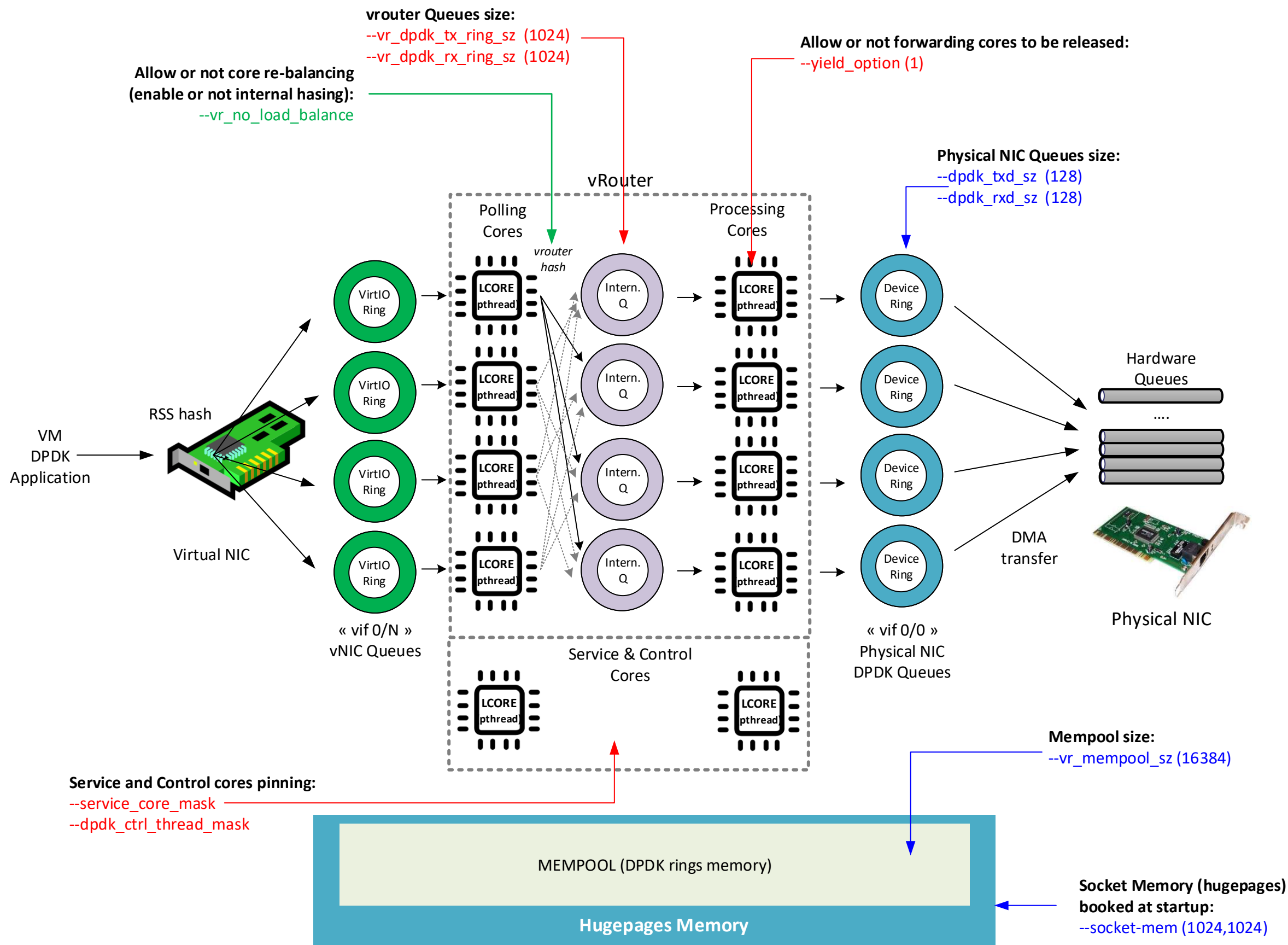


Title: DPDK Library Framework

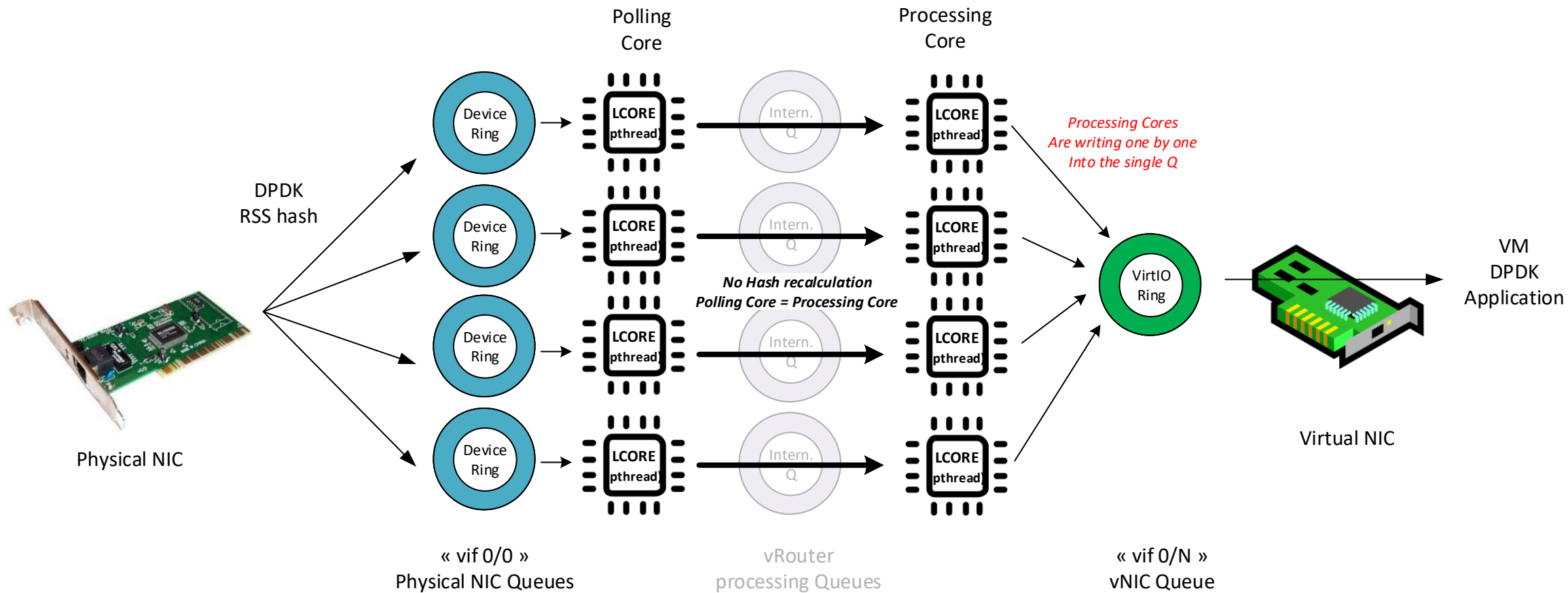
version: 1.0

date: 02/06/2020





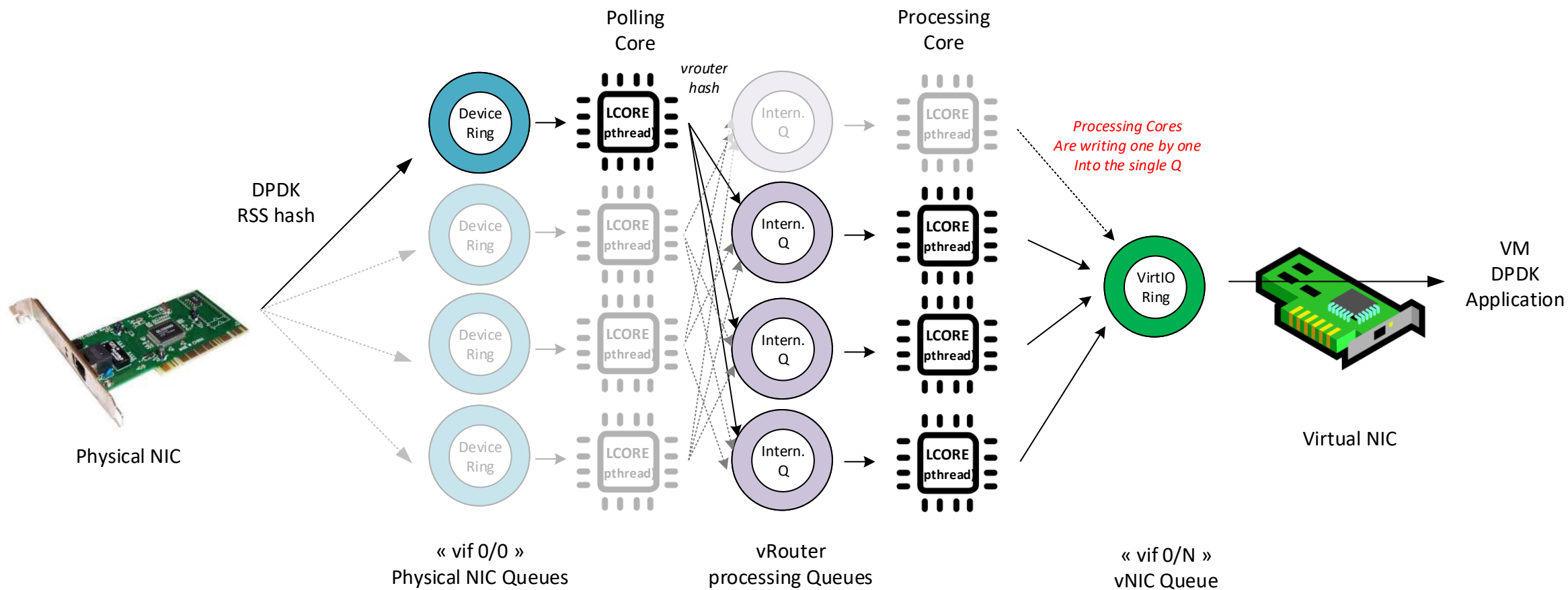
Title: vRouter Architecture	
version: 1.1	date: 22/05/2020



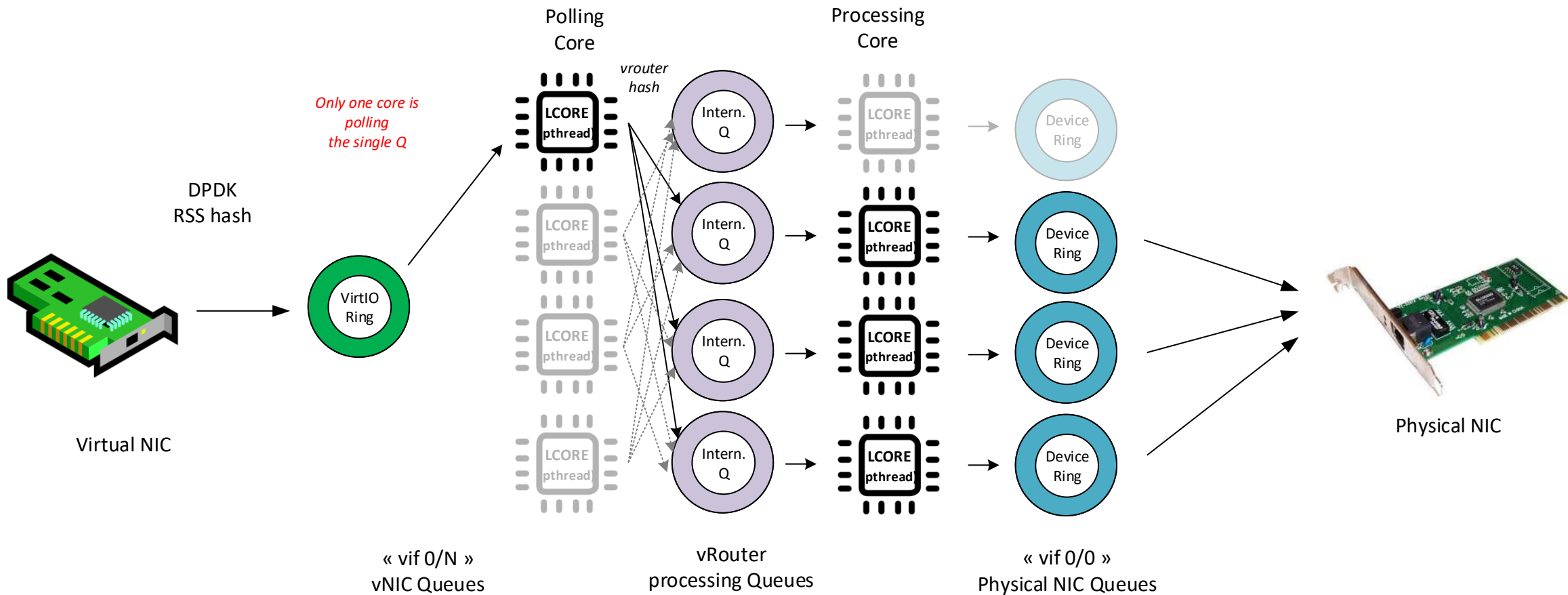
Title: UDP to SingleQ

version: 1.0

date: 25/03/2020



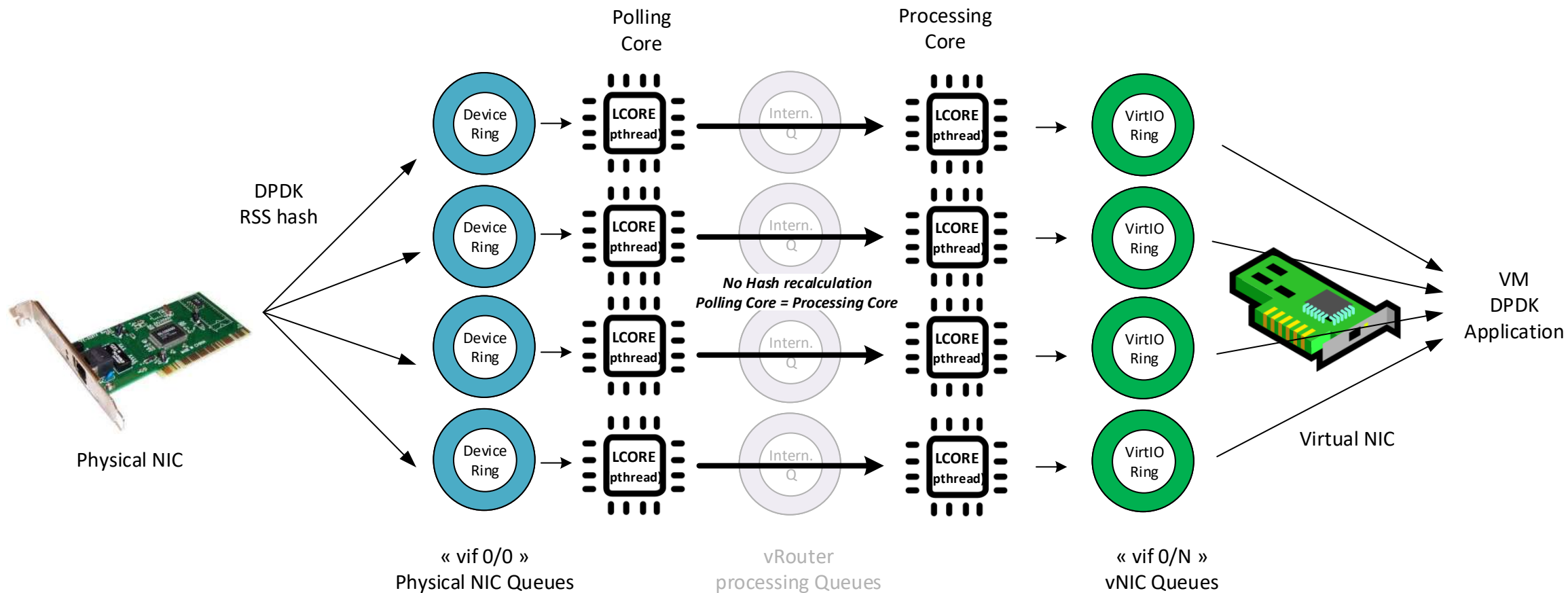
Title: GRE to SingleQ	
version: 1.0	date: 25/03/2020



Title: SingleQ to Underlay

version: 1.0

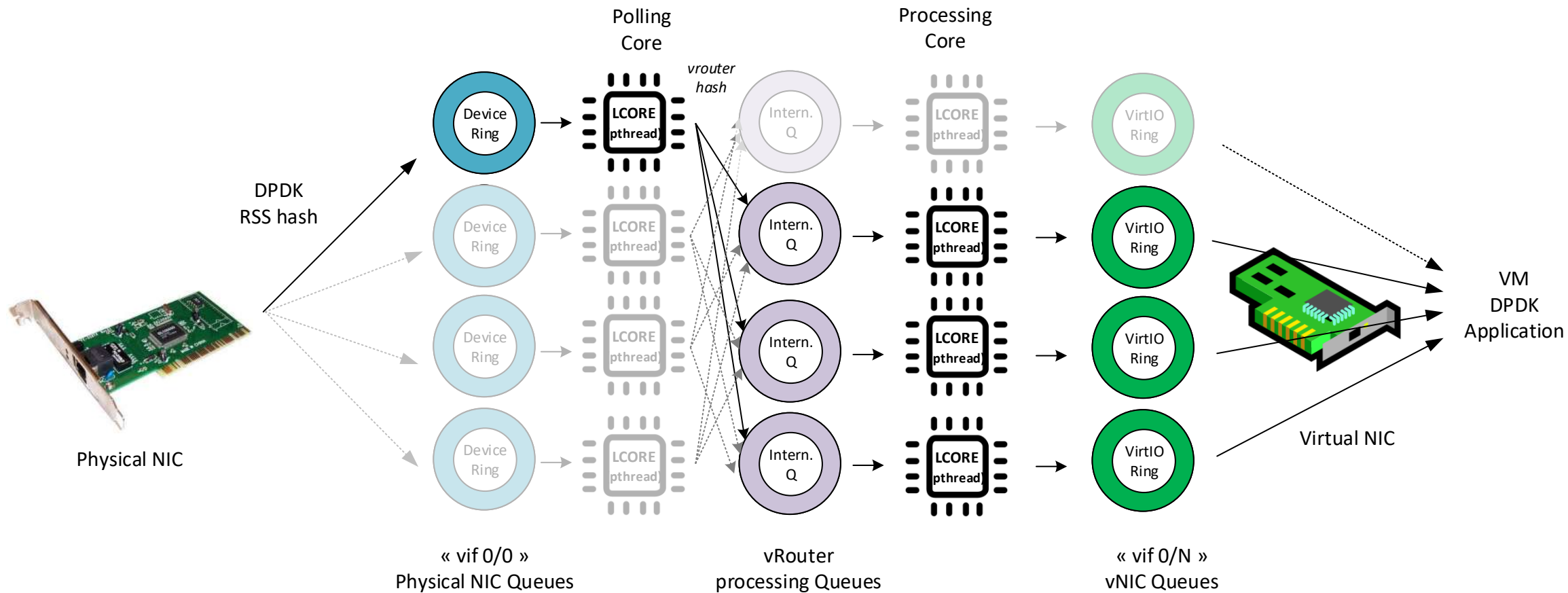
date: 25/03/2020



Title: UDP to MultiQ

version: 1.0

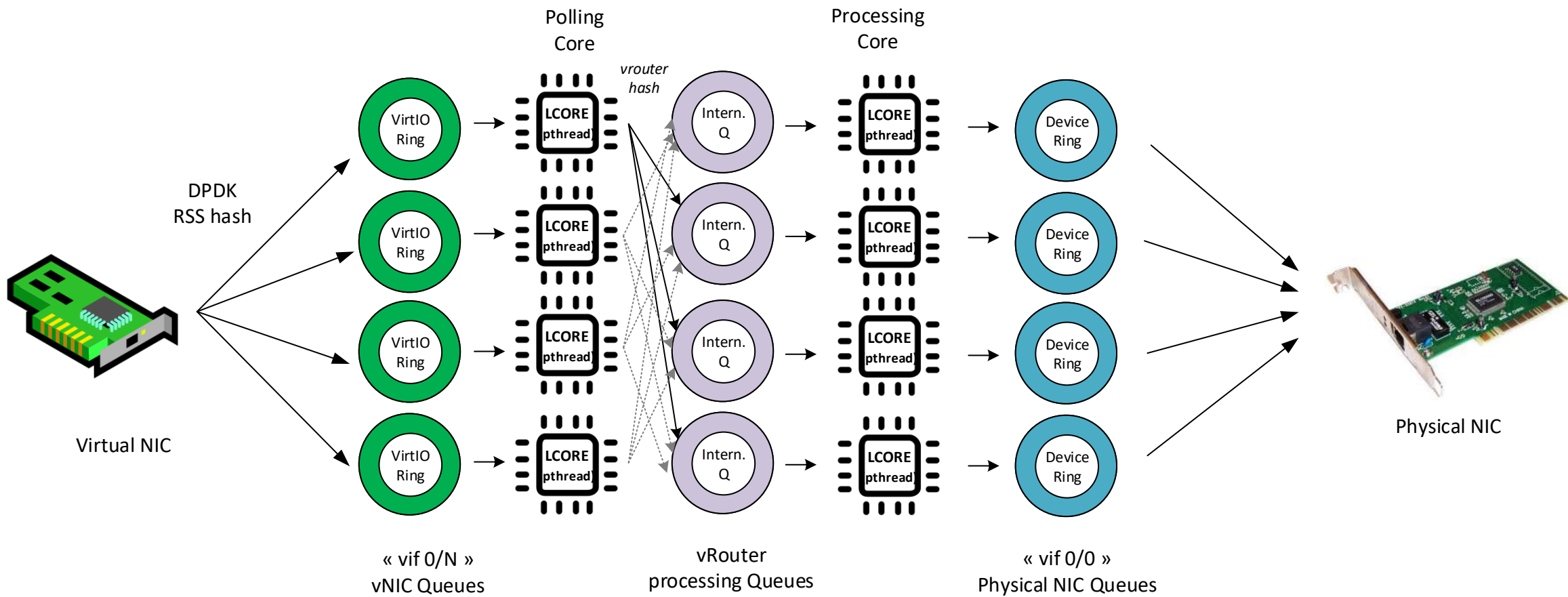
date: 25/03/2020



Title: GRE to MultiQ

version: 1.0

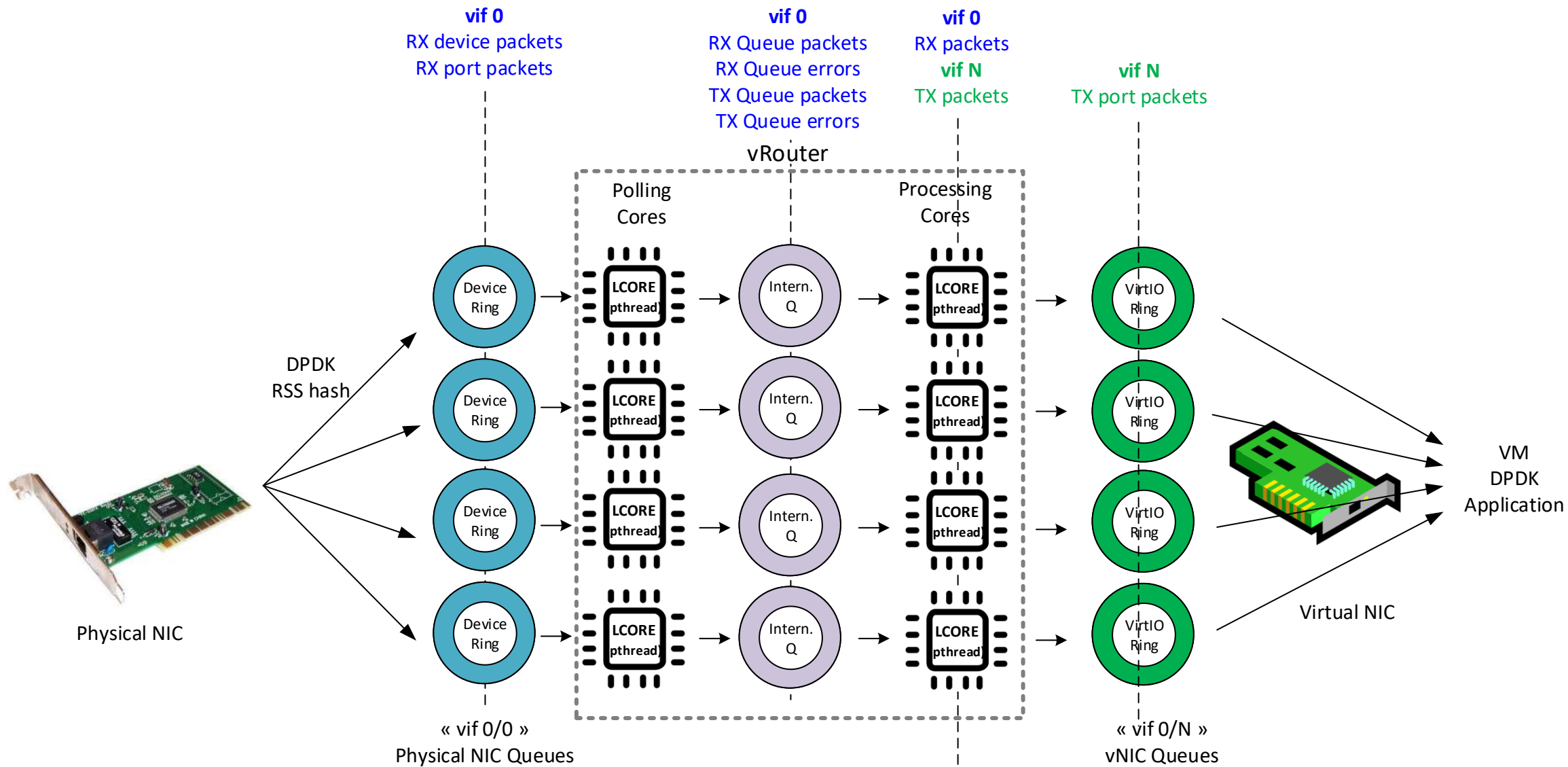
date: 25/03/2020



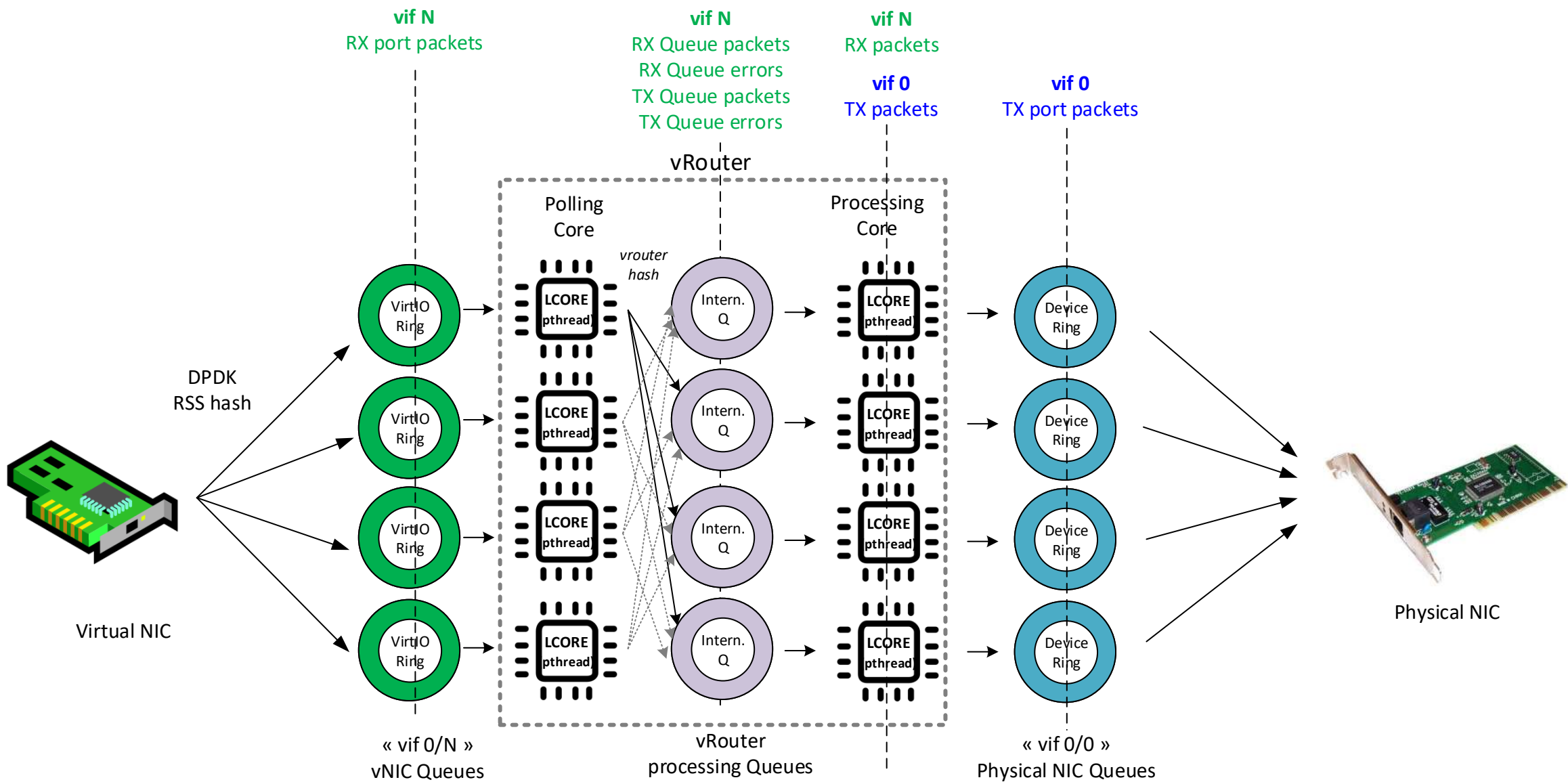
Title: MultiQ to Underlay

version: 1.0

date: 25/03/2020



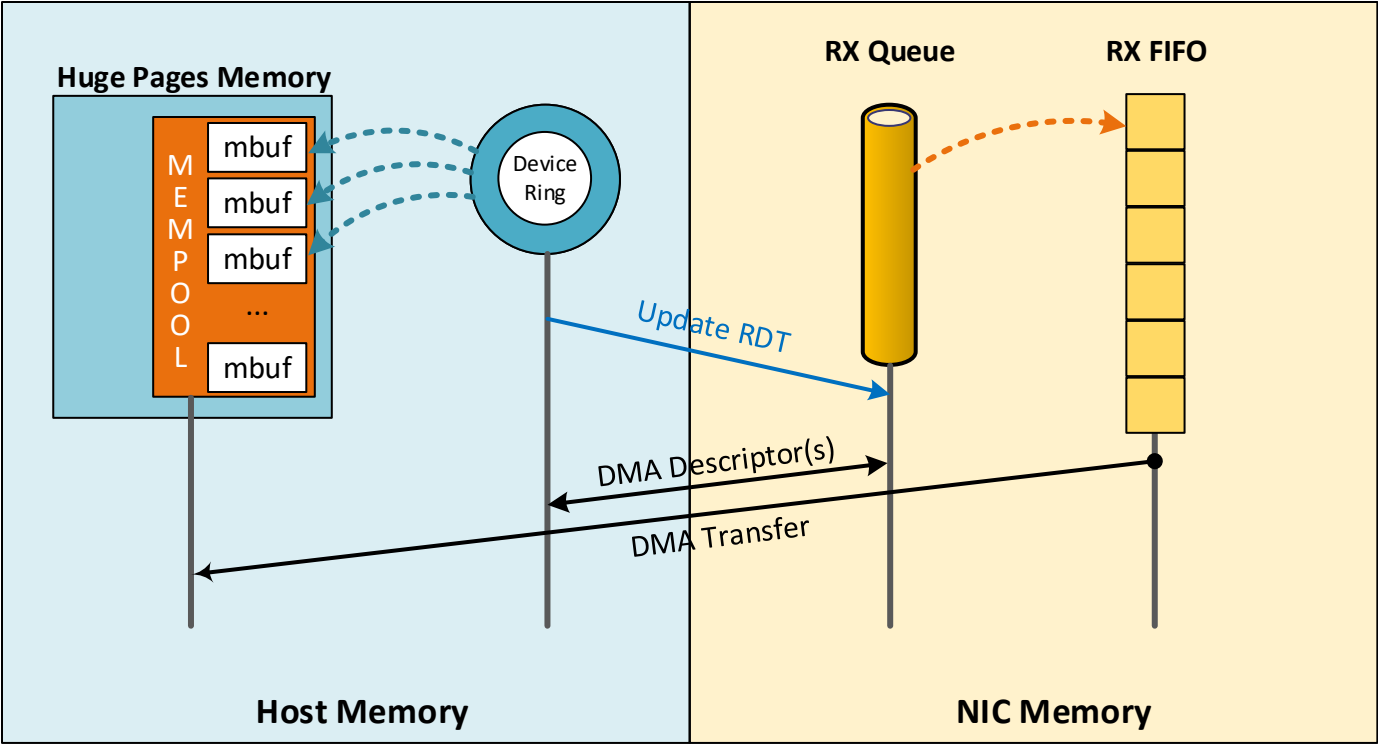
Title: vif counters (phys to virt)	
version: 1.0	date: 25/03/2020



Title: vif counters (virt to phys)

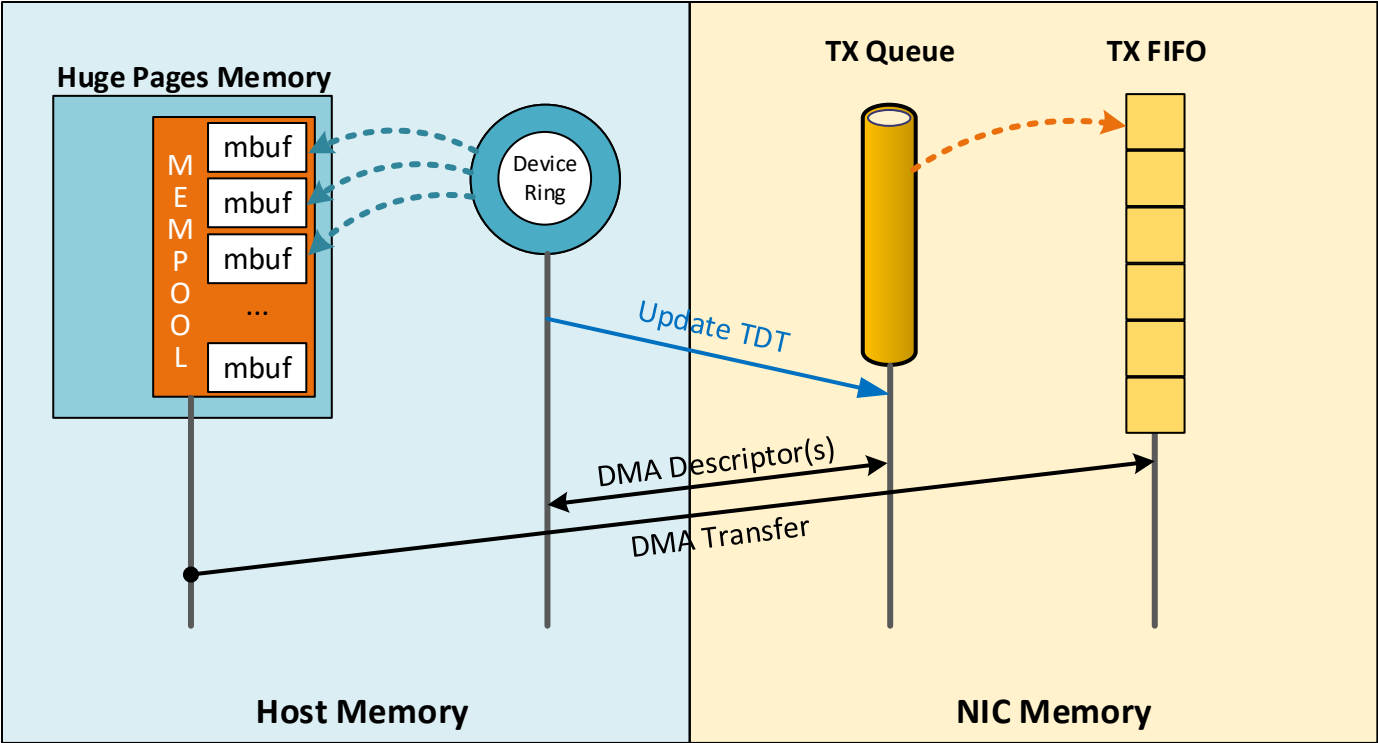
version: 1.0

date: 25/03/2020



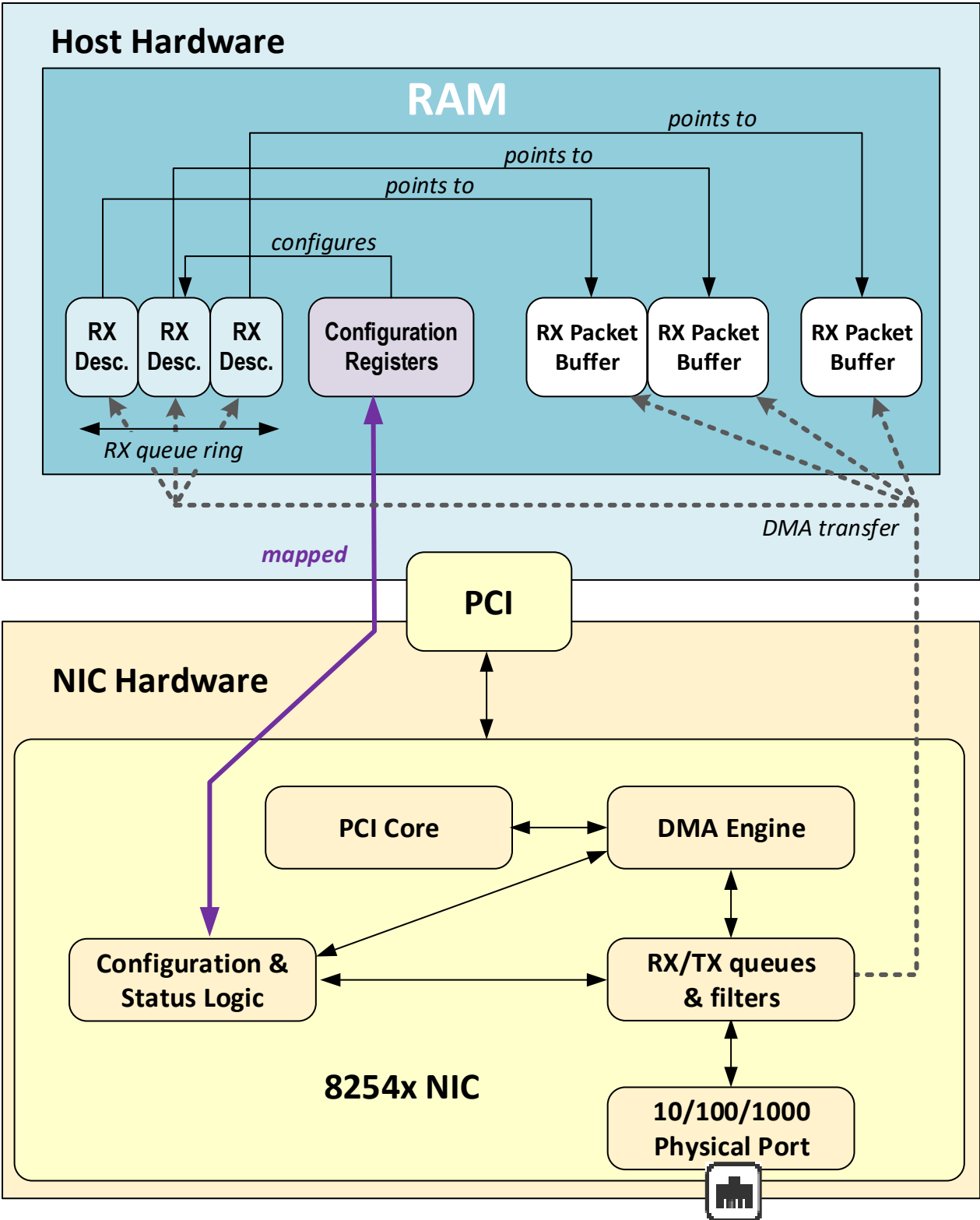
RDT: Receive Descriptor Tail

Title: pNIC incoming packet	
version: 1.0	date: 25/03/2020



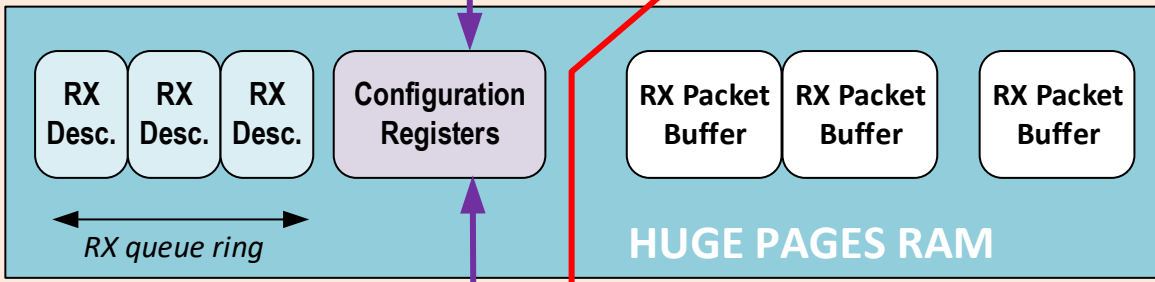
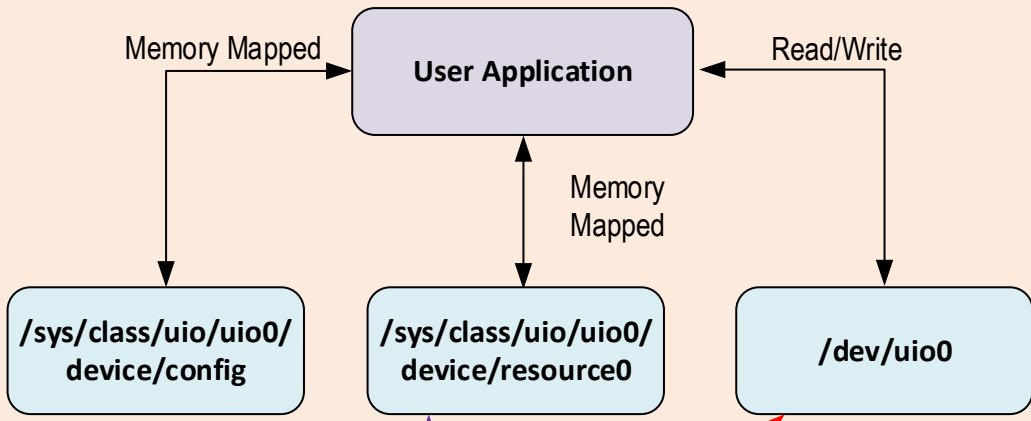
TDT: Transmit Descriptor Tail

Title: pNIC outgoing packet	
version: 1.0	date: 25/03/2020

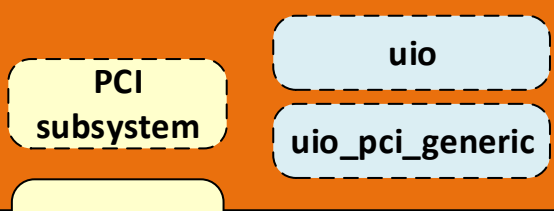


Host Hardware

User Space

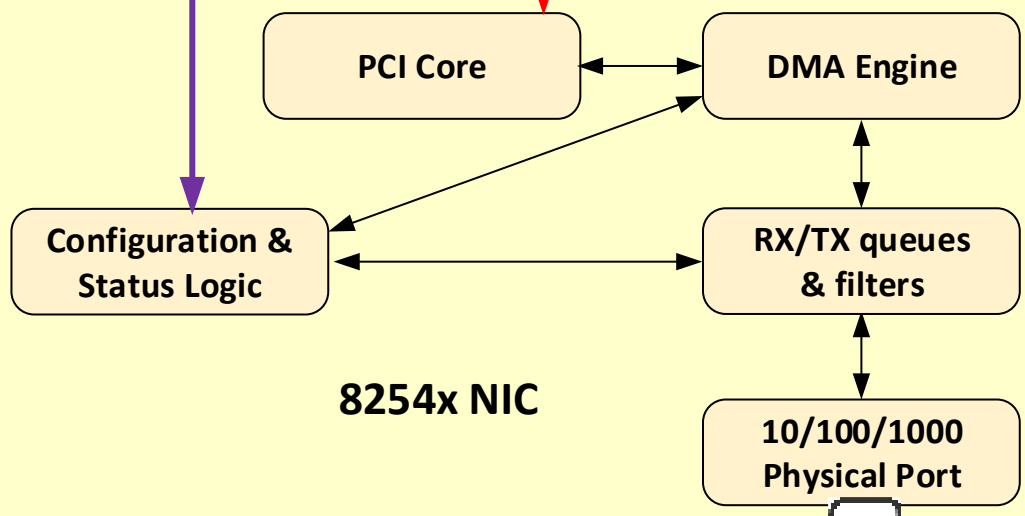


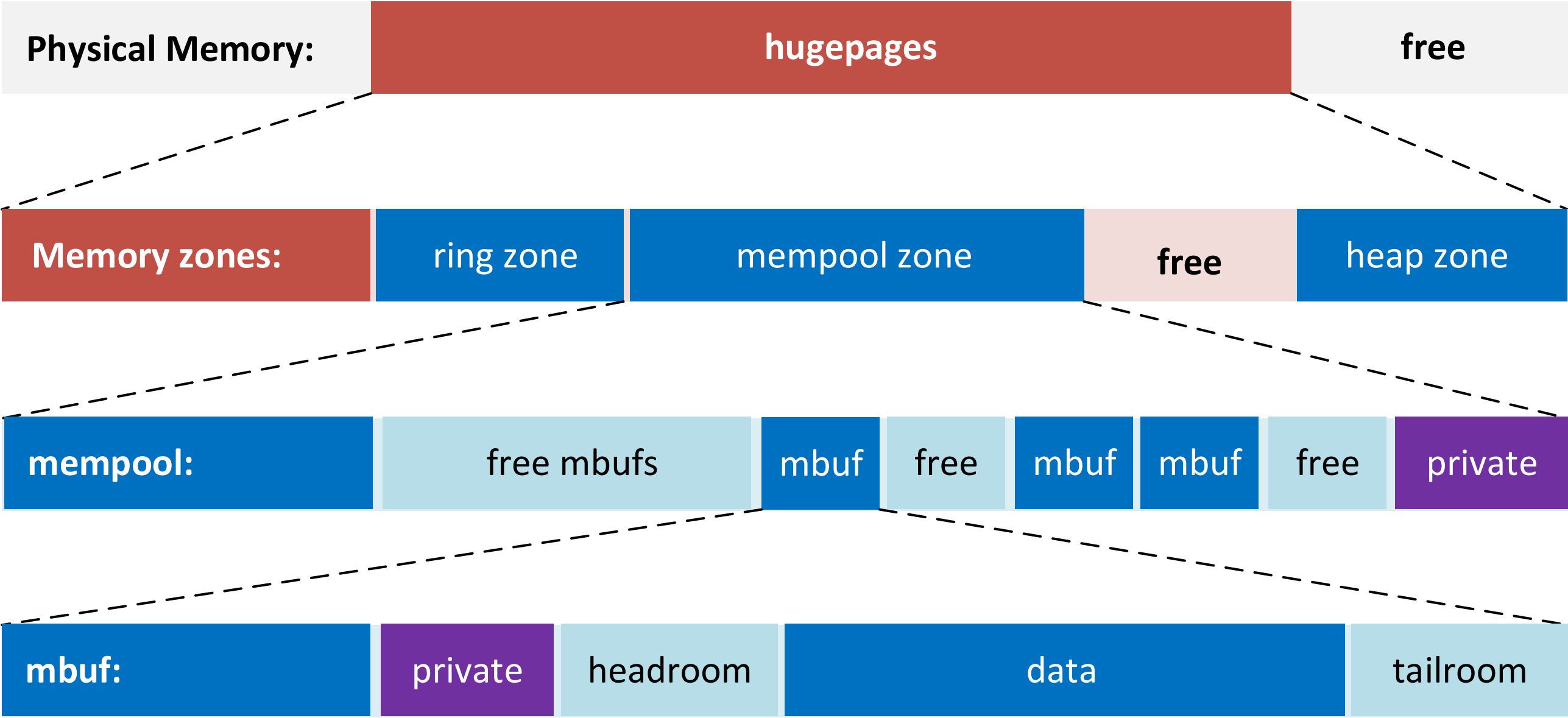
Kernel Space

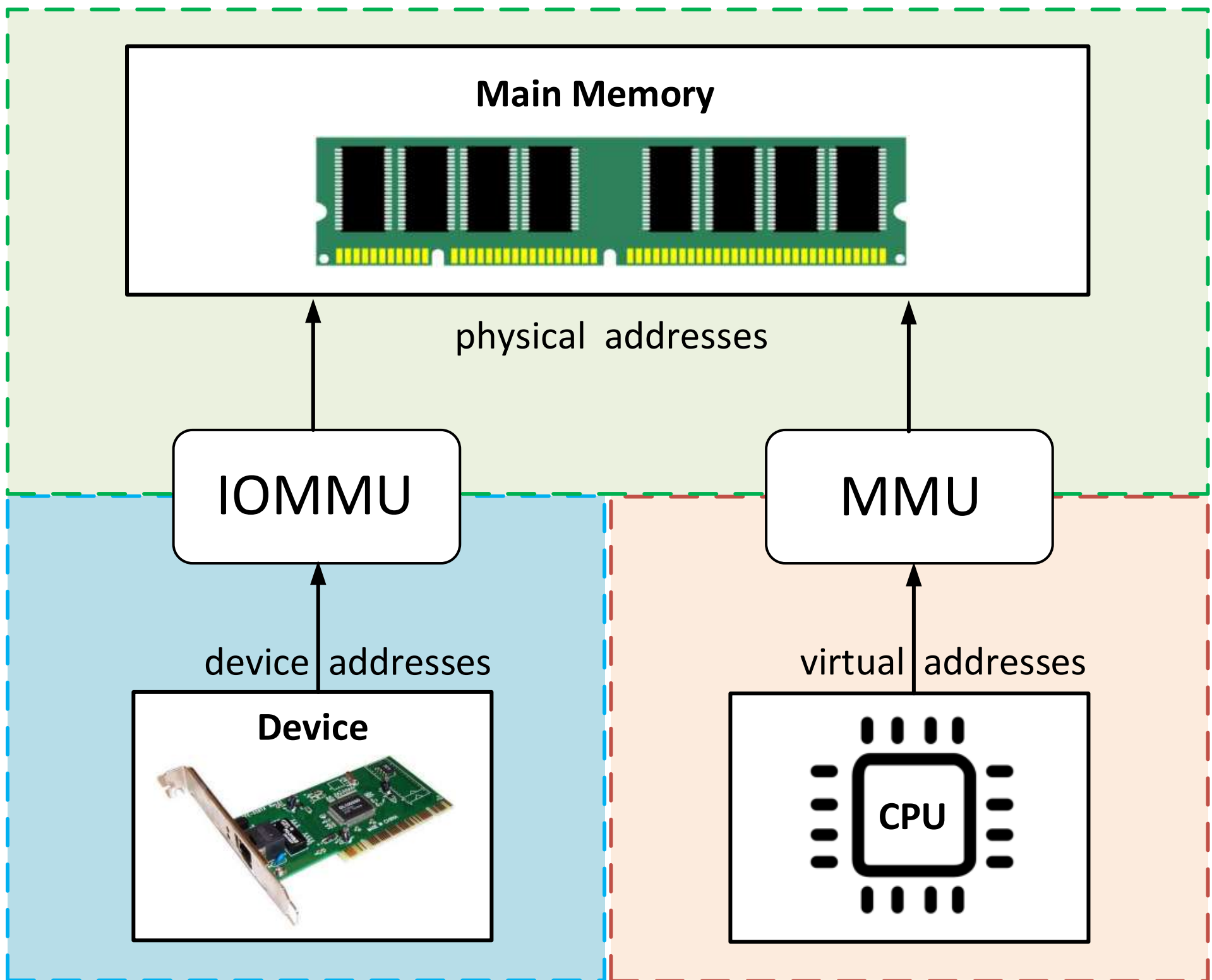


PCI

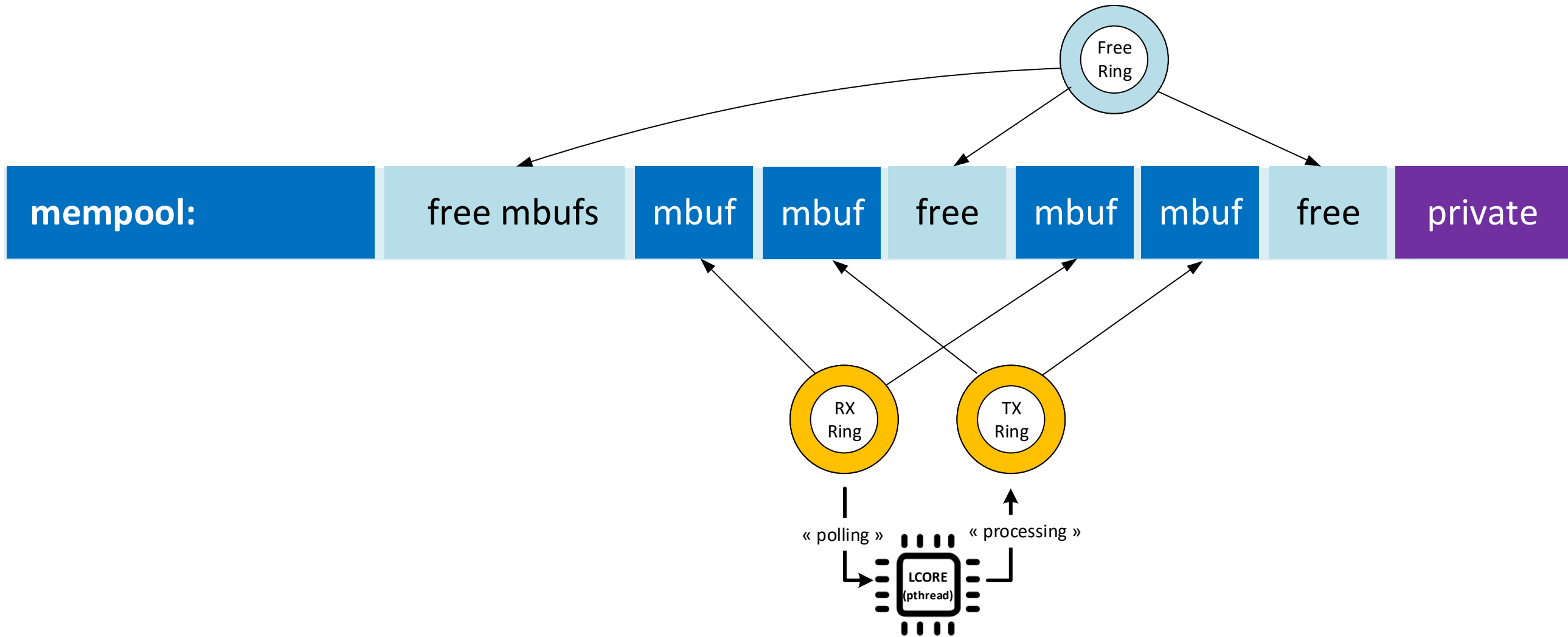
NIC Hardware



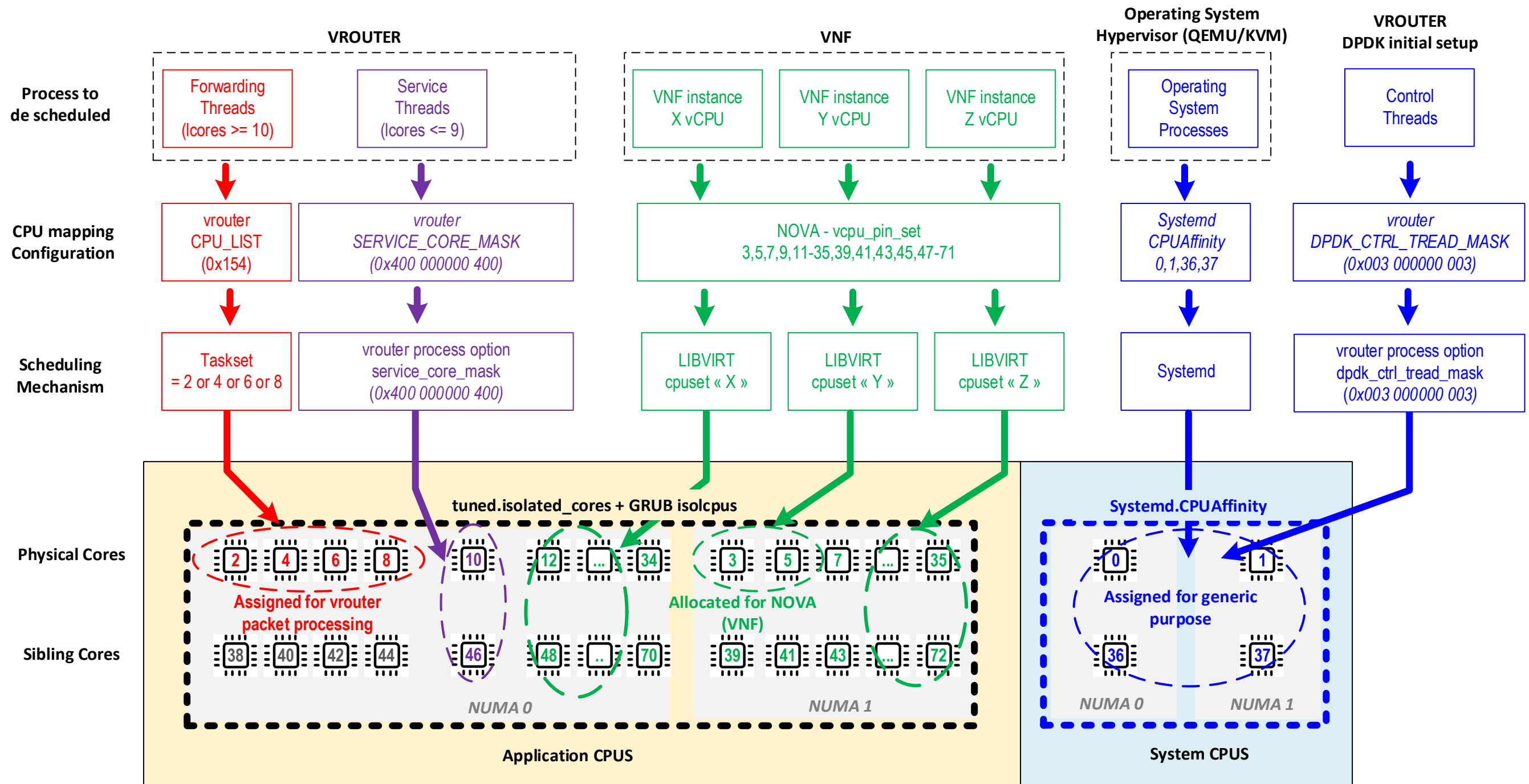




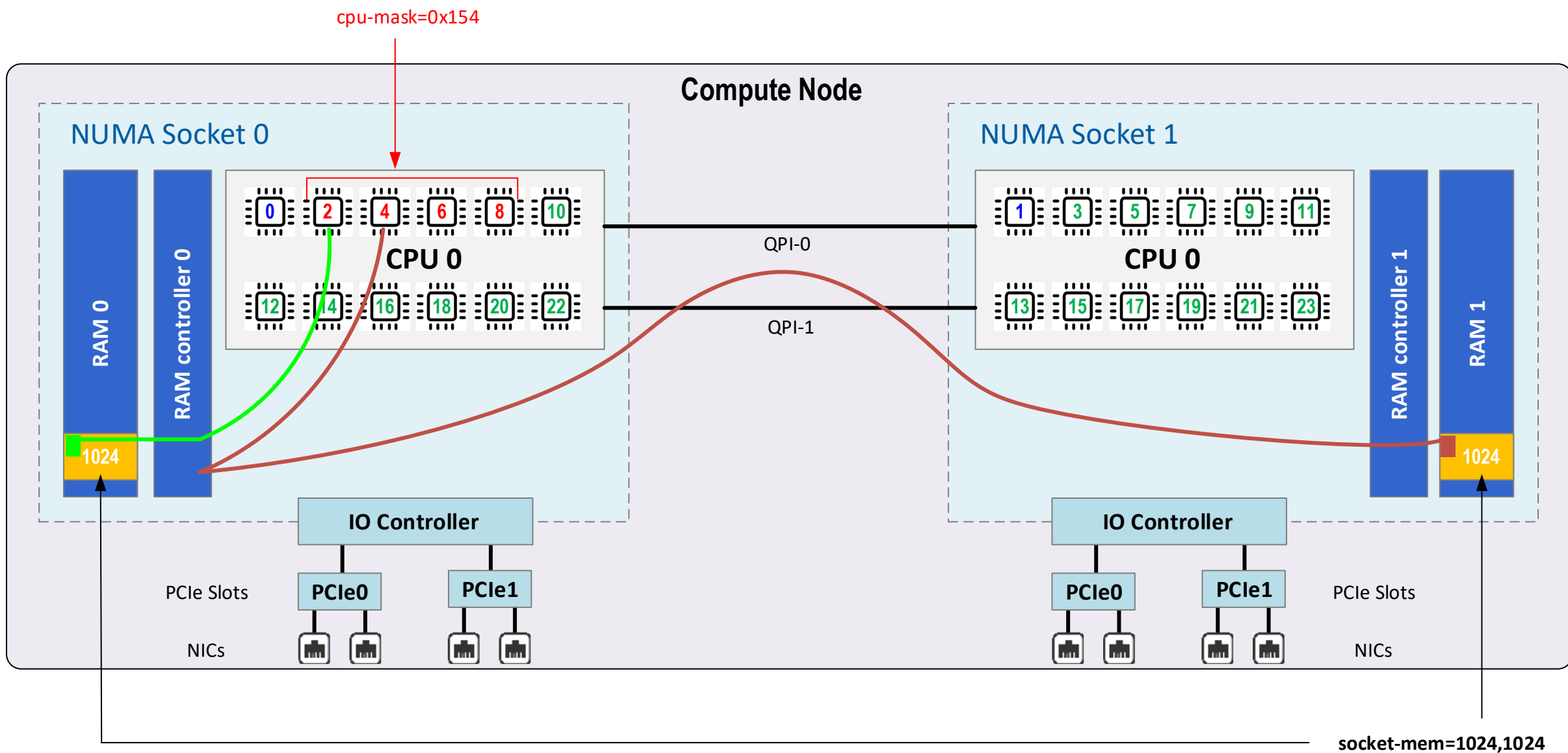
Memory Addressing	
version: 1.0	date: 27/03/2020



Packets Processing and Memory	
version: 1.0	date: 25/03/2020



Title: CPU Pinning	
version: 1.0	date: 25/03/2020



vRouter DPDK PMD allocated CPUs

VM allocated CPUs

Host OS allocated CPUs



vRouter memory access shortest path

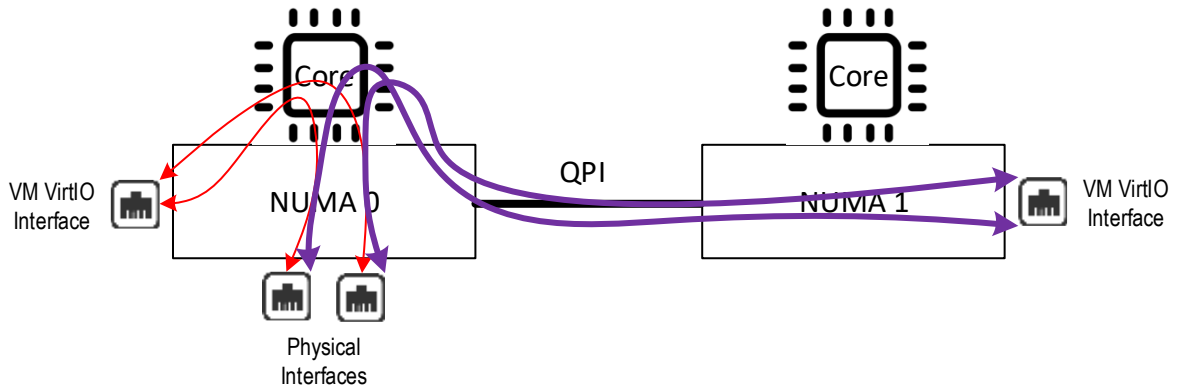
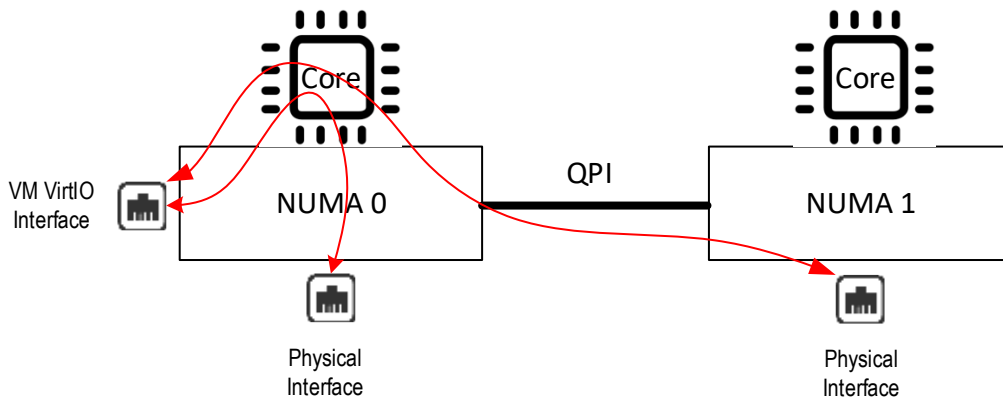


vRouter memory access longer path

Title: vRouter Memory Access and NUMA

version: 1.0

date: 27/03/2020

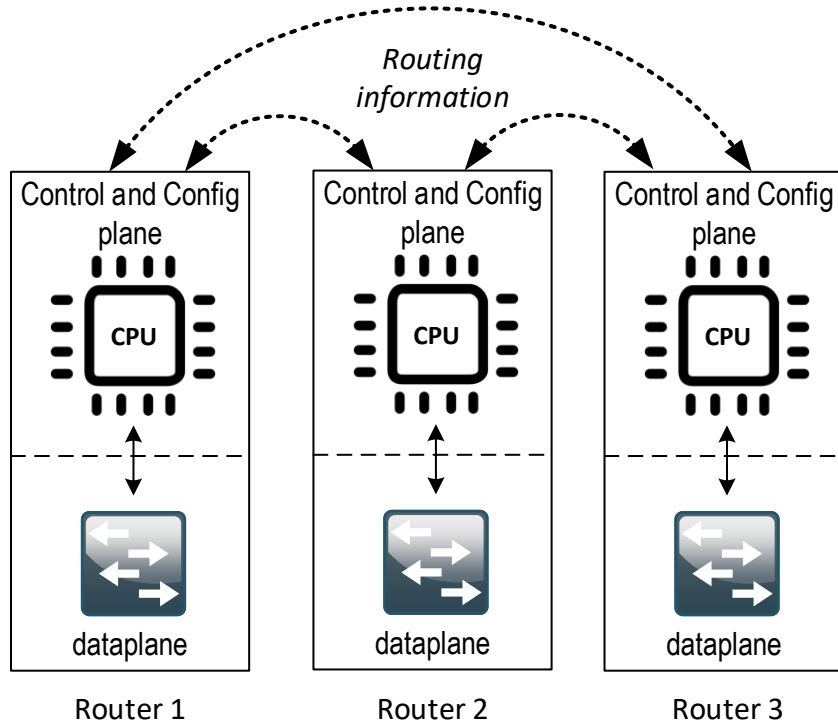


Title: NUMA concern

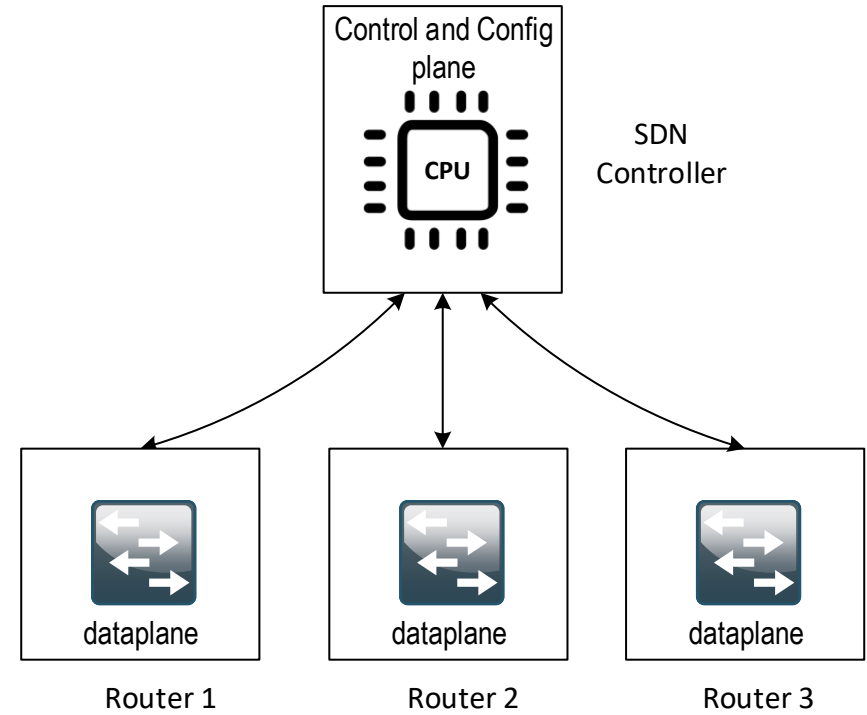
version: 1.0

date: 25/03/2020

Traditional Network Devices



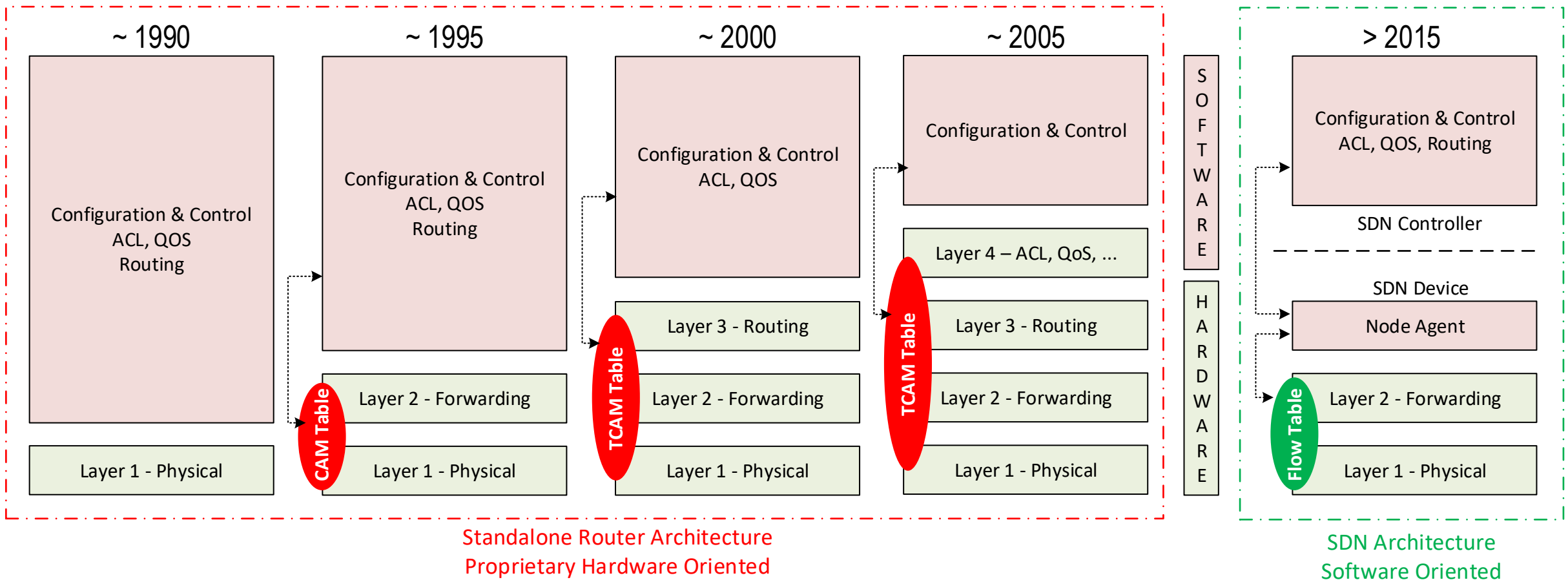
SDN Devices



Title: SDN Overview

version: 1.0

date: 13/04/2020



Title: SDN History	
version: 1.0	date: 13/04/2020

Configuration

Modular Router



Console Port

Router Processor Card

Network Protocols

OSPF

ISIS

BGP

LDP

RSVP-TE

RIB

LIB

FIB

LFIB

Switch Fabric

Line Cards ...

FIB

LFIB

FIB

LFIB

Packets Walk

Control Packets

Traditional Network Node

version: 1.0

date: 13/04/2020

Modular Router

Router Processor Card

Configuration Plane

CLI

SNMP

NetConf

Control Plane

OSPF

ISIS

BGP

LDP

RSVP-TE

Switch
Fabric

Forwarding Plane

TCAM

ASIC

Line Cards
...

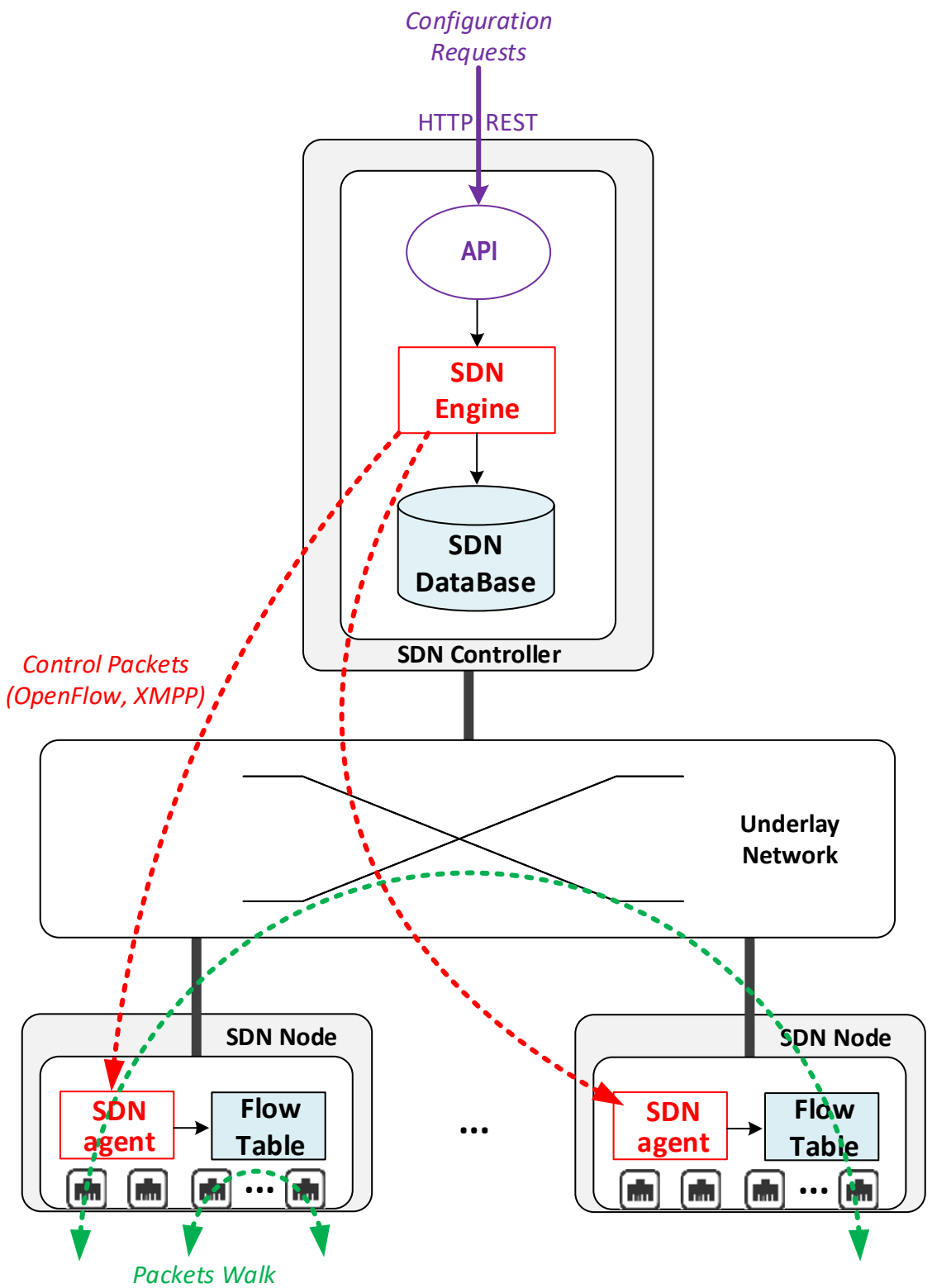
TCAM

ASIC

Traditional Network Planes

version: 1.0

date: 13/04/2020



SDN Detailed Architecture	
version: 1.0	date: 13/04/2020

Admin and Users

OpenStack

Kubernetes

APPLICATION
LAYER

SDN Dashboard
Web GUI
Application

Cloud Network
Interface

Container Network
Interface

CONTROL
LAYER

SDN Controller

Northbound Interfaces

REST
API

Python
API

JAVA
API

SDN Engine

SDN Control
Logic

Routing Info
(flows)

Topology

Statistics

NetFlow

XMPP

NetConf

BGP/LS

PECP

Southbound Interfaces

Underlay
Network

INFRASTRUCTURE
LAYER

SDN Node

SDN
agent

Flow
Table

Packet Engine

Forwarding Plane

SDN Node

SDN
agent

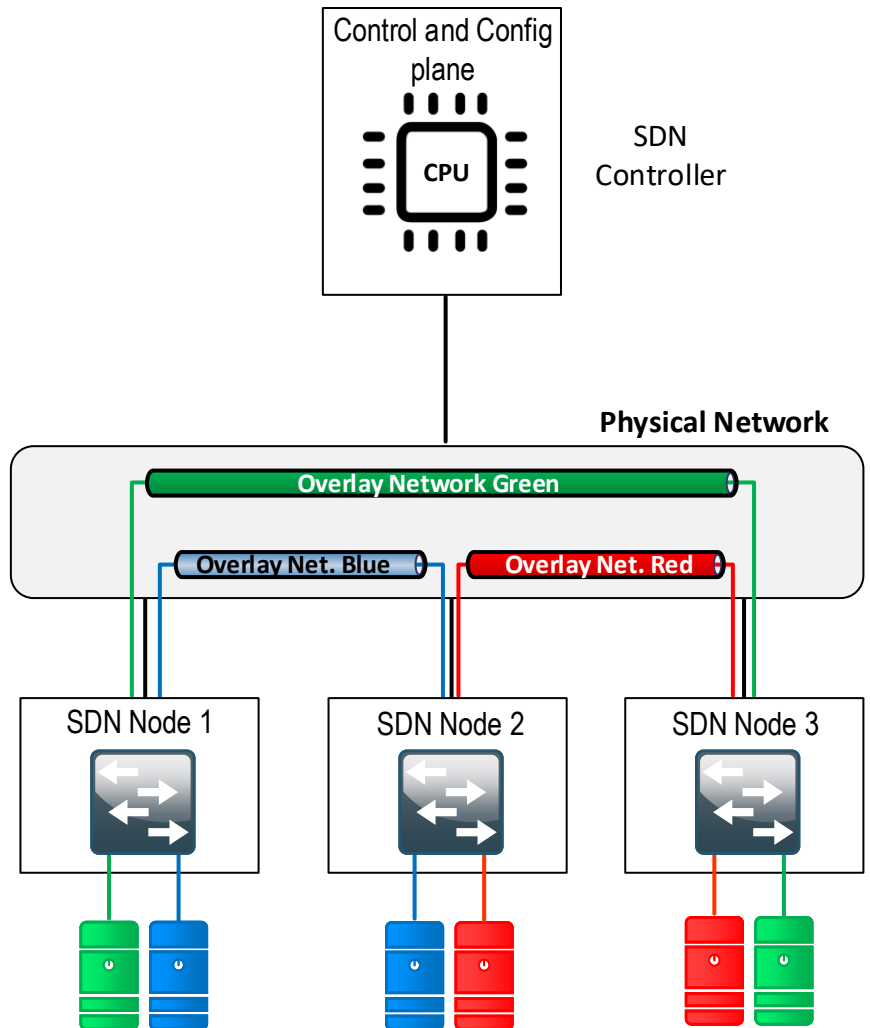
Flow
Table

Packet Engine

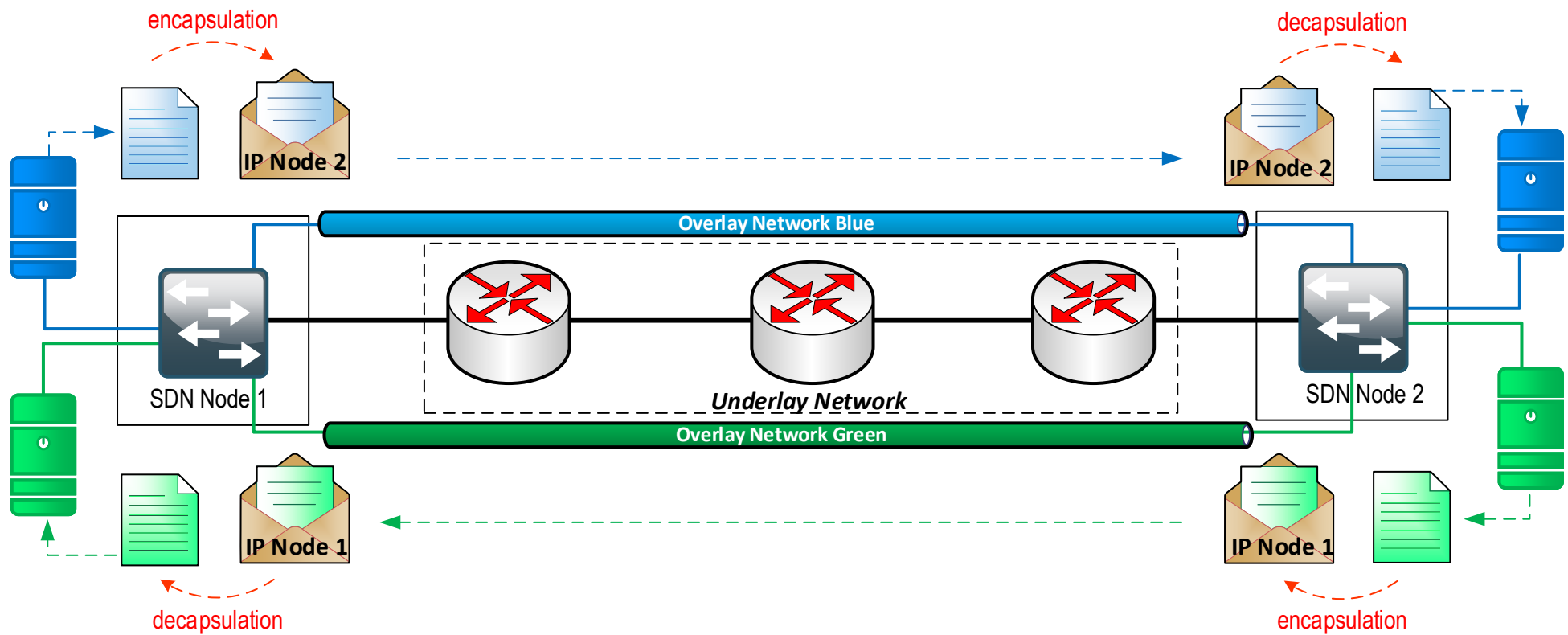
SDN Layers

version: 1.0

date: 13/04/2020



SDN and Overlay	
version: 1.0	date: 13/04/2020



SDN and Encapsulation Principle	
version: 1.0	date: 13/04/2020

Compute Node

User Space

Network Application

NIC
Configuration

Data
Packets
Processing

control path

data path

Host
RAM

Kernel Space

Kernel
Module

Kernel
Module

Hardware

Network
Controller

NIC
Buffer



Title: Control and Data paths

version: 1.0

date: 08/06/2020

Compute Node

User Space

Network Application
Data Packets Processing

Socket API

read
system call

write
system call

Kernel Space

Socket Buffer
(host RAM)

Ethernet Driver API

Hardware

interrupt &
DMA
write

DMA
read &
interrupt

NIC
Buffer

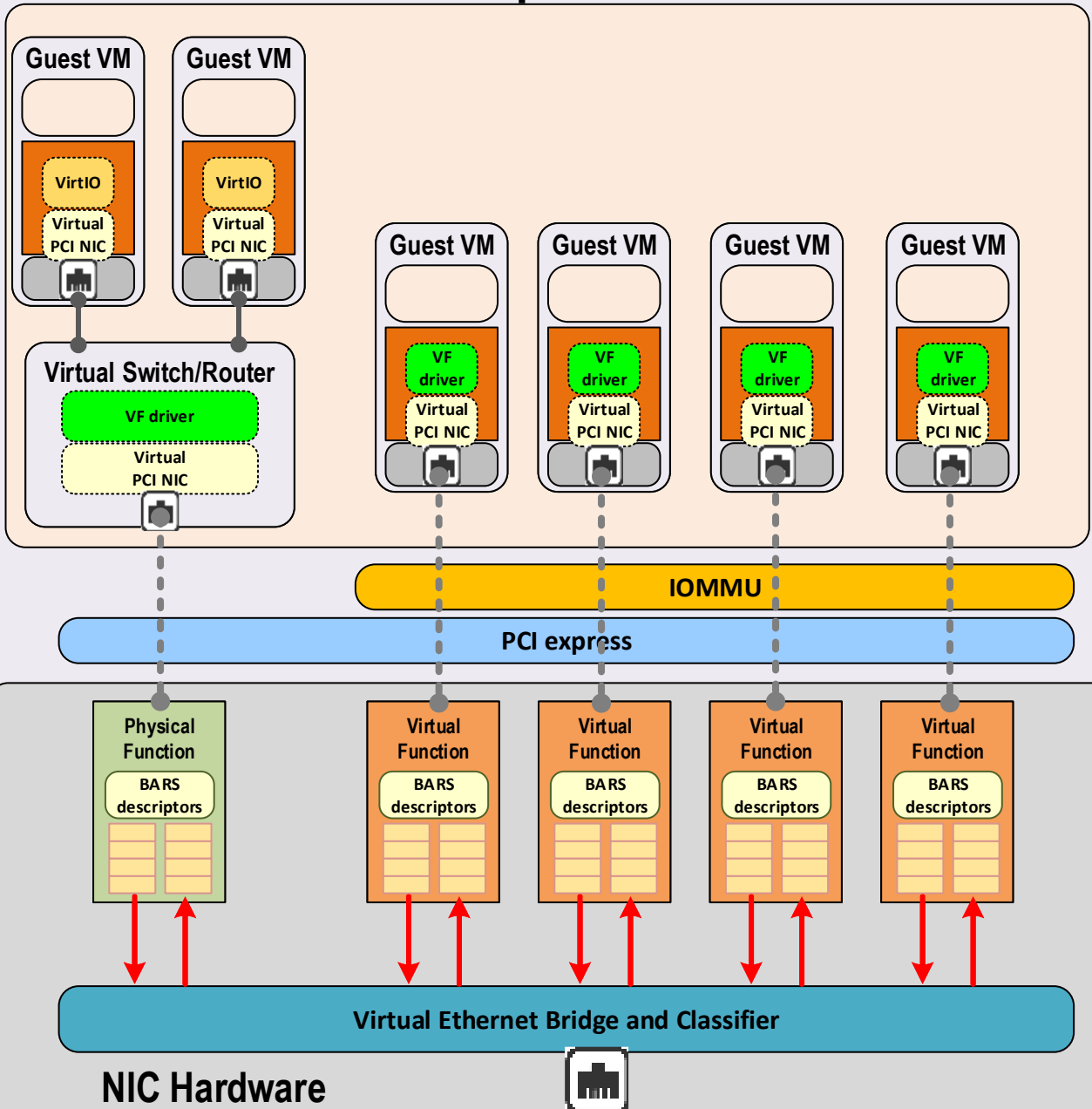


Title: Event Base Packet processing

version: 1.0

date: 08/06/2020

Compute Node



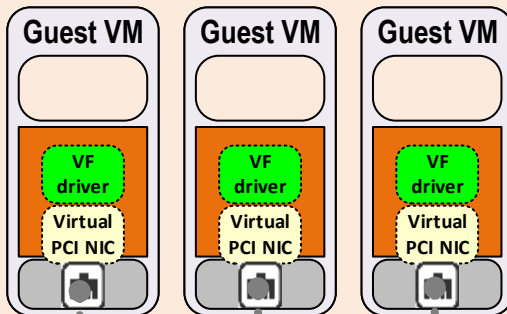
Title: SRIOV - Overview

version: 1.0

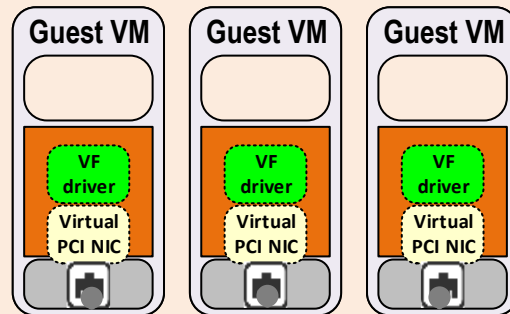
date: 08/06/2020

Compute Node

Direct I/O Assignment PCI Passthrough



Direct I/O Assignment SRIOV



VMM

Hardware

VF1 VF2 VF3

Virtual Ethernet
Bridge and Classifier

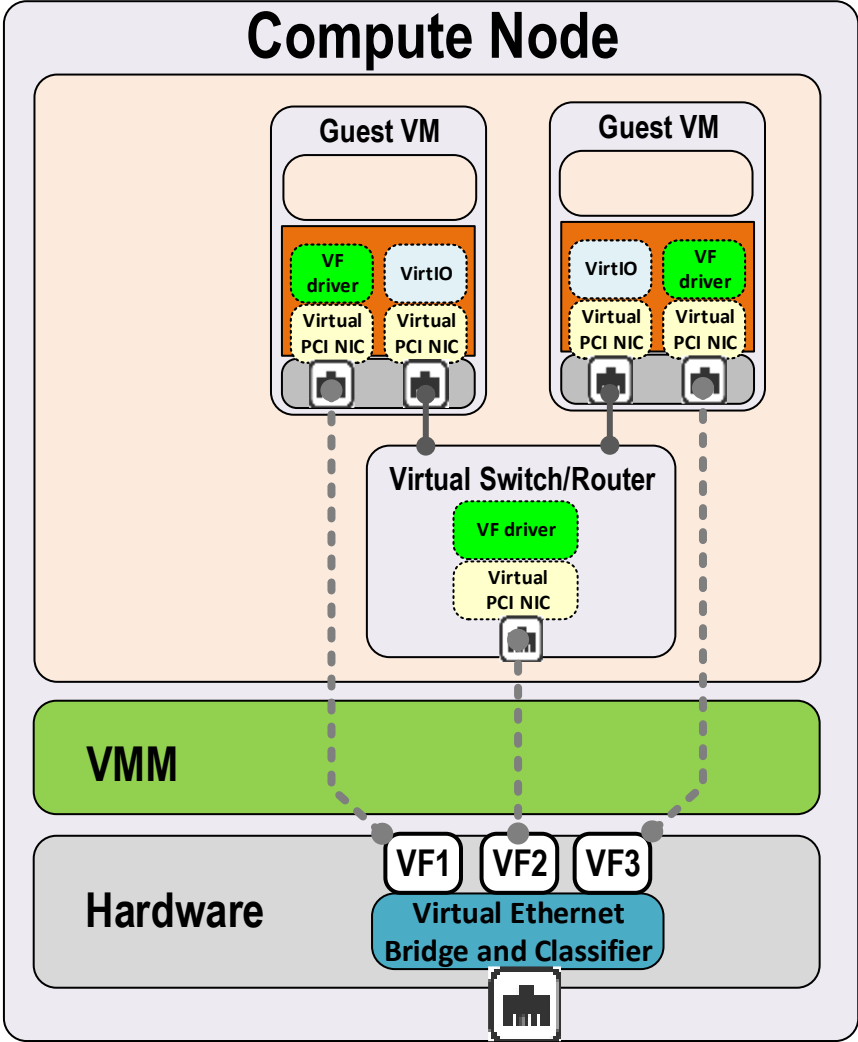
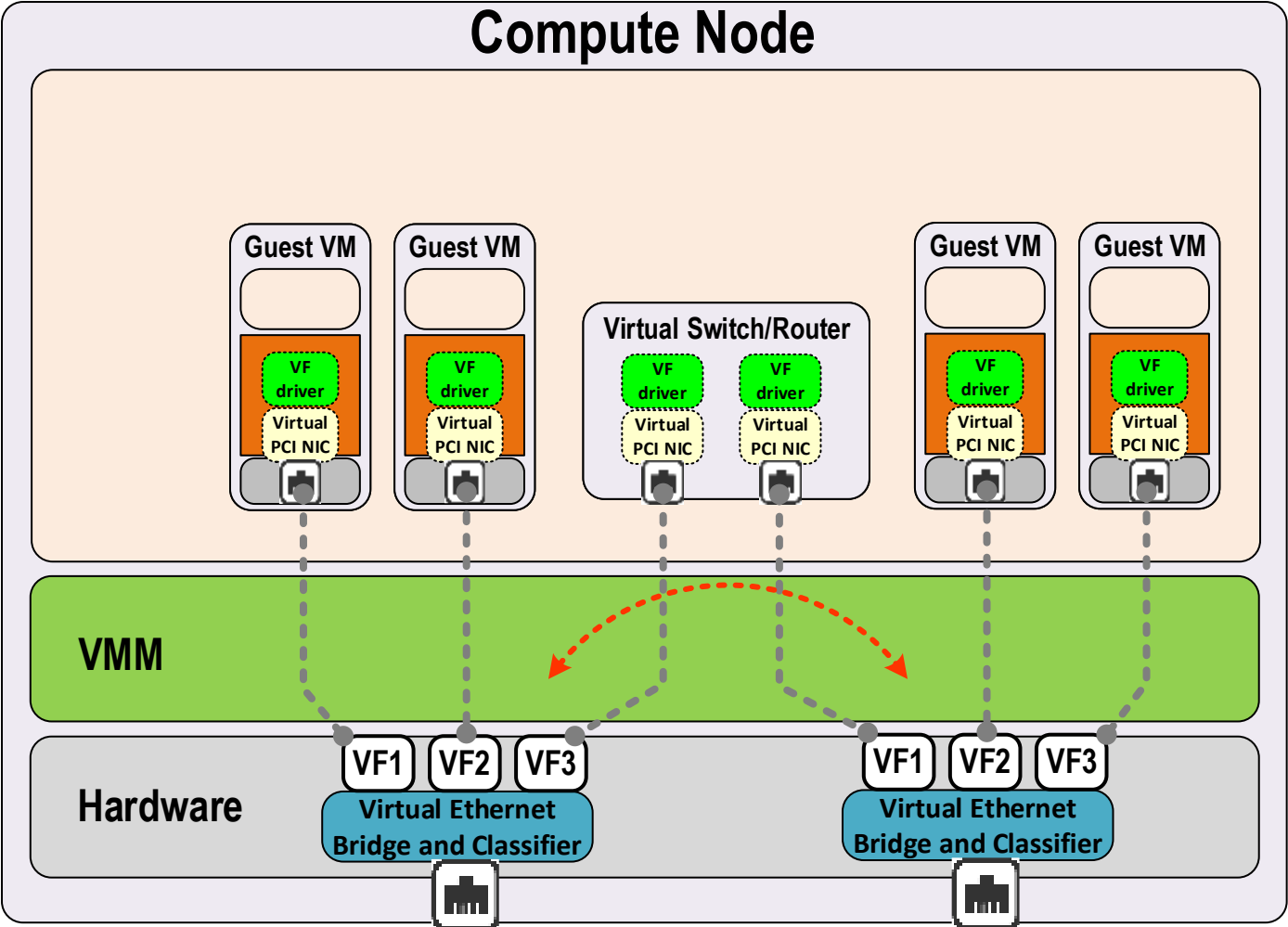
one **dedicated** physical
port per VM

one single physical port
shared between VM

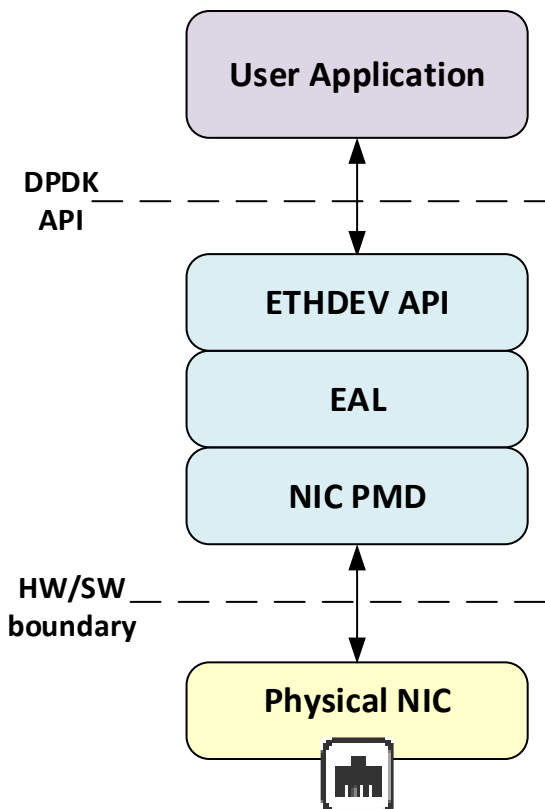
Title: SRIOV versus PCI Passthrough

version: 1.0

date: 08/06/2020



Title: SRIOV and SDN	
version: 1.0	date: 08/06/2020



Title: DPDK model	
version: 1.0	date: 08/06/2020

Compute Node

User Space

Network Application
Data Packets Processing

User
read

User
write

DMA
write

packets buffer
host RAM

DMA
read

Kernel Space

Ethernet Driver API

Hardware

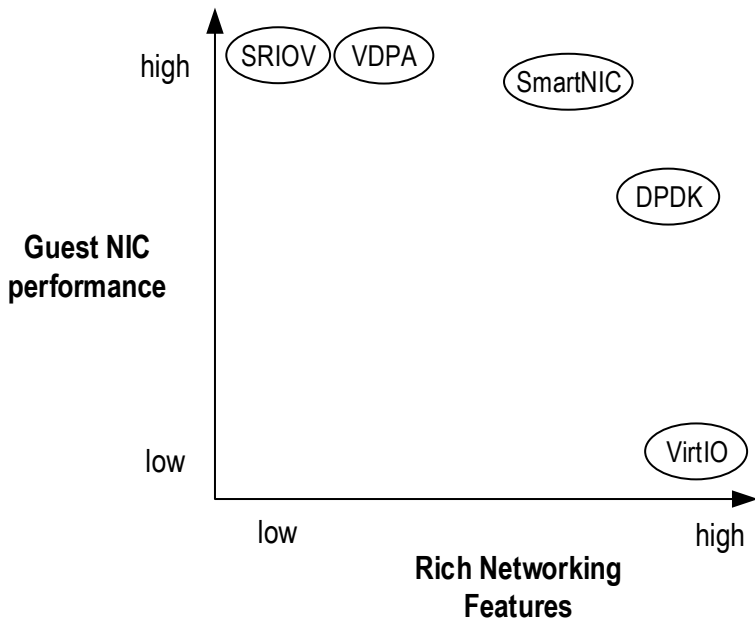
NIC
Buffer



Title: Polling Based Packet processing

version: 1.0

date: 08/06/2020



Title: Comparison Diagram

version: 1.0

date: 08/06/2020

Compute Node

User Space

Guest VM

User Space

User Application

DPDK
Vendor specific PMD

Read/Write

Assign

Kernel Space

VFIO

viOMMU

Virtual
PCI NIC

Emulated
Hardware



Assign

Huge Pages Memory

Guest
huge pages

map

NIC
registers and queues

map

secured DMA transfer

Kernel Space

Vendor
PF driver

Virtual
Function
Creation

VFIO

IOMMU

Physical
PCI NIC

Hardware



Title: Guest VM DPDK and SRIOV

version: 1.0

date: 13/06/2020

Compute Node

User Space

Guest VM

User Space

User Application

DPDK
VirtIO net PMD

Read/Write

Assign

Kernel Space

VFIO

viOMMU

Virtual
PCI NIC

Emulated
Hardware



Assign

Huge Pages Memory

Guest
huge pages

map

NIC
registers and queues

map

secured DMA transfer

Kernel Space

Vendor
PF driver

Virtual
Function
Creation

VFIO

IOMMU

VirtIO Physical
PCI NIC

Hardware



Title: Guest VM DPDK and VirtIO passthrough

version: 1.0

date: 13/06/2020

Compute Node

User Space

Guest VM

User Space

User Application

DPDK
VirtIO net PMD

Read/Write

Assign

Kernel Space

VFIO

viOMMU

Virtual
PCI NIC

Emulated
Hardware



Assign

Huge Pages Memory

Guest
huge pages

map

NIC
registers and queues

map

secured DMA transfer

Kernel Space

Vendor
PF driver

Virtual
Function
Creation

VFIO

VFIO-MDEV

IOMMU

vDPA Physical
PCI NIC

Hardware

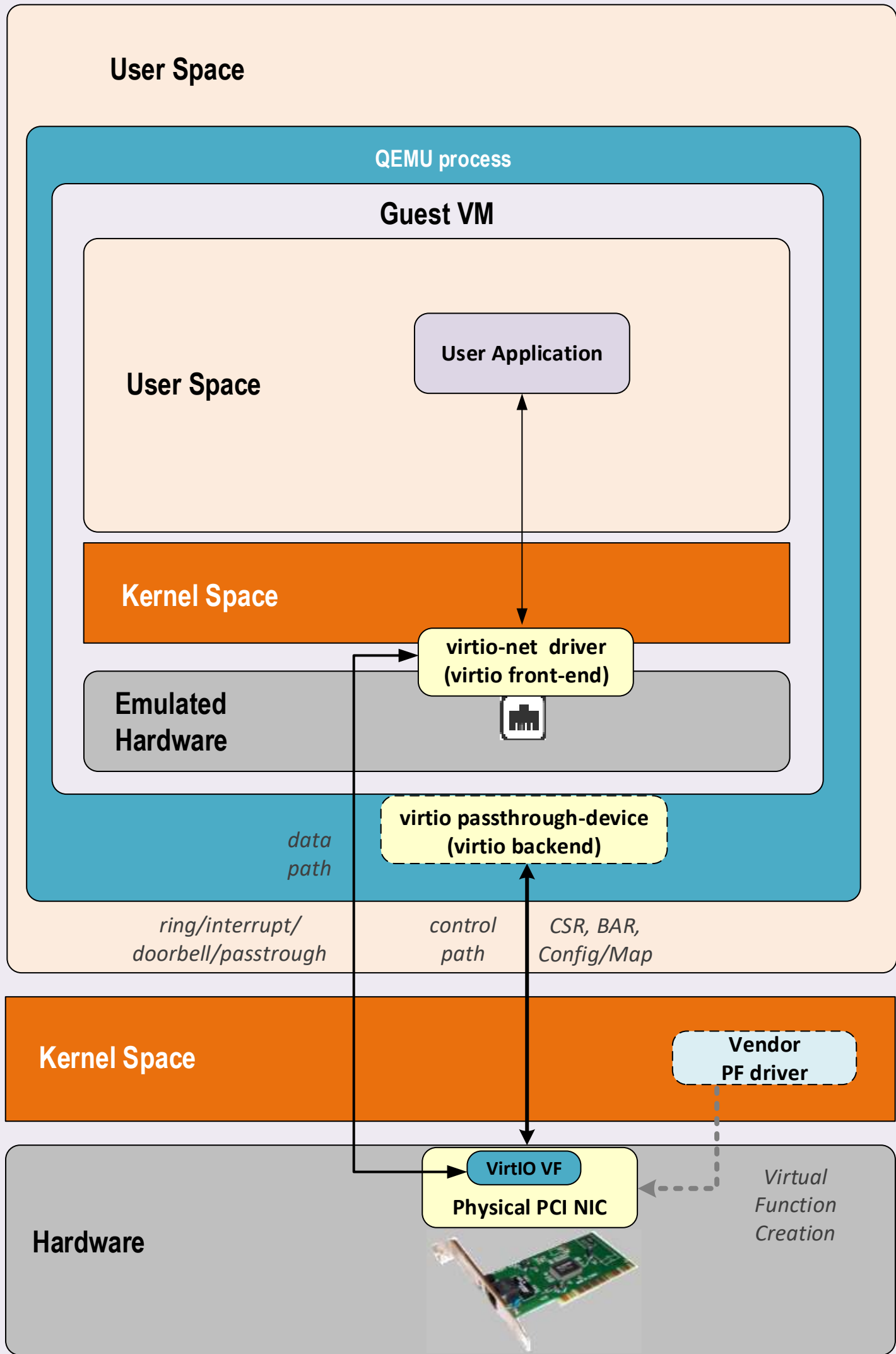


Title: Guest VM DPDK and vDPA

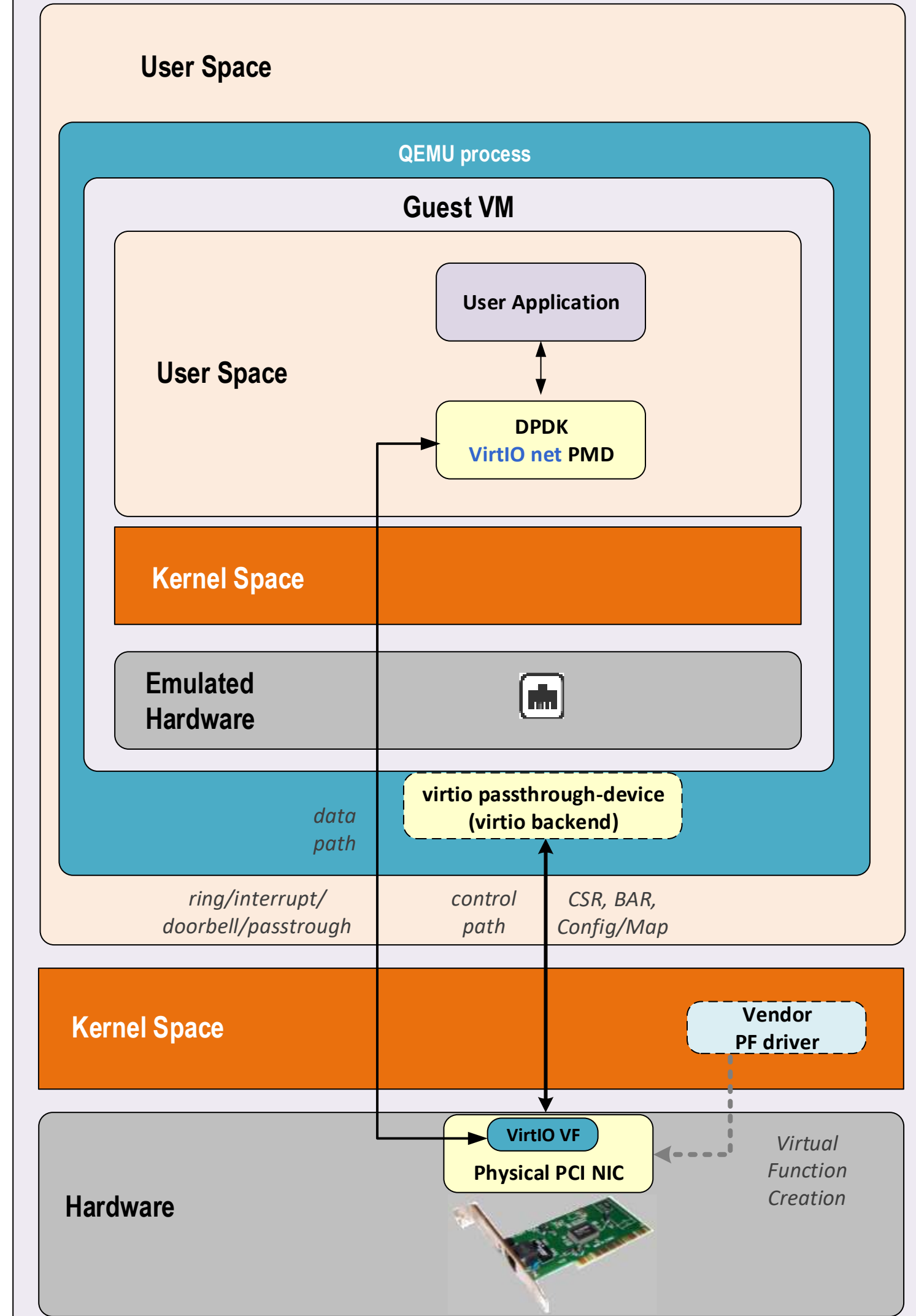
version: 1.0

date: 13/06/2020

Compute Node



Compute Node

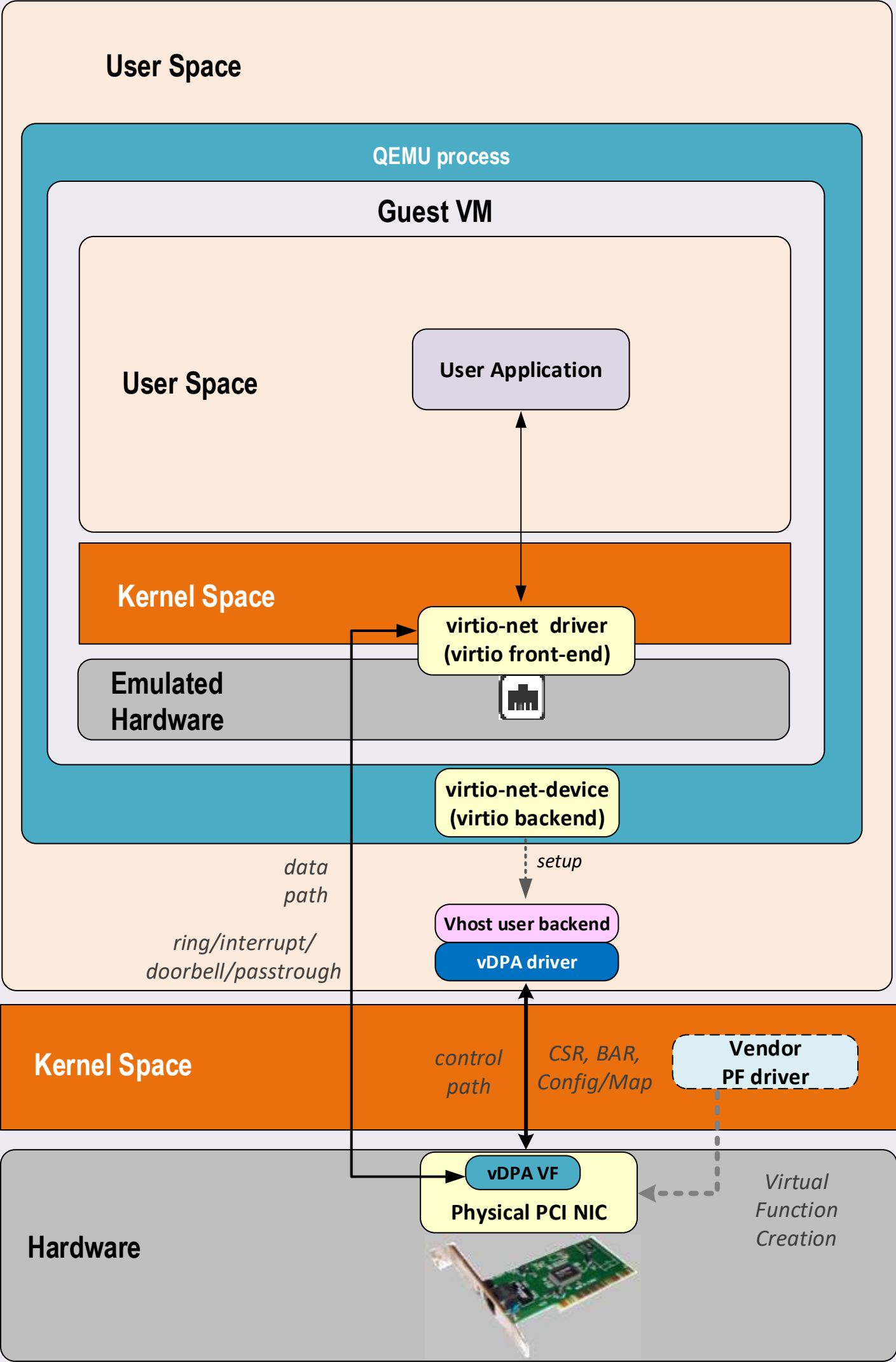


Title: VirtIO full offloading

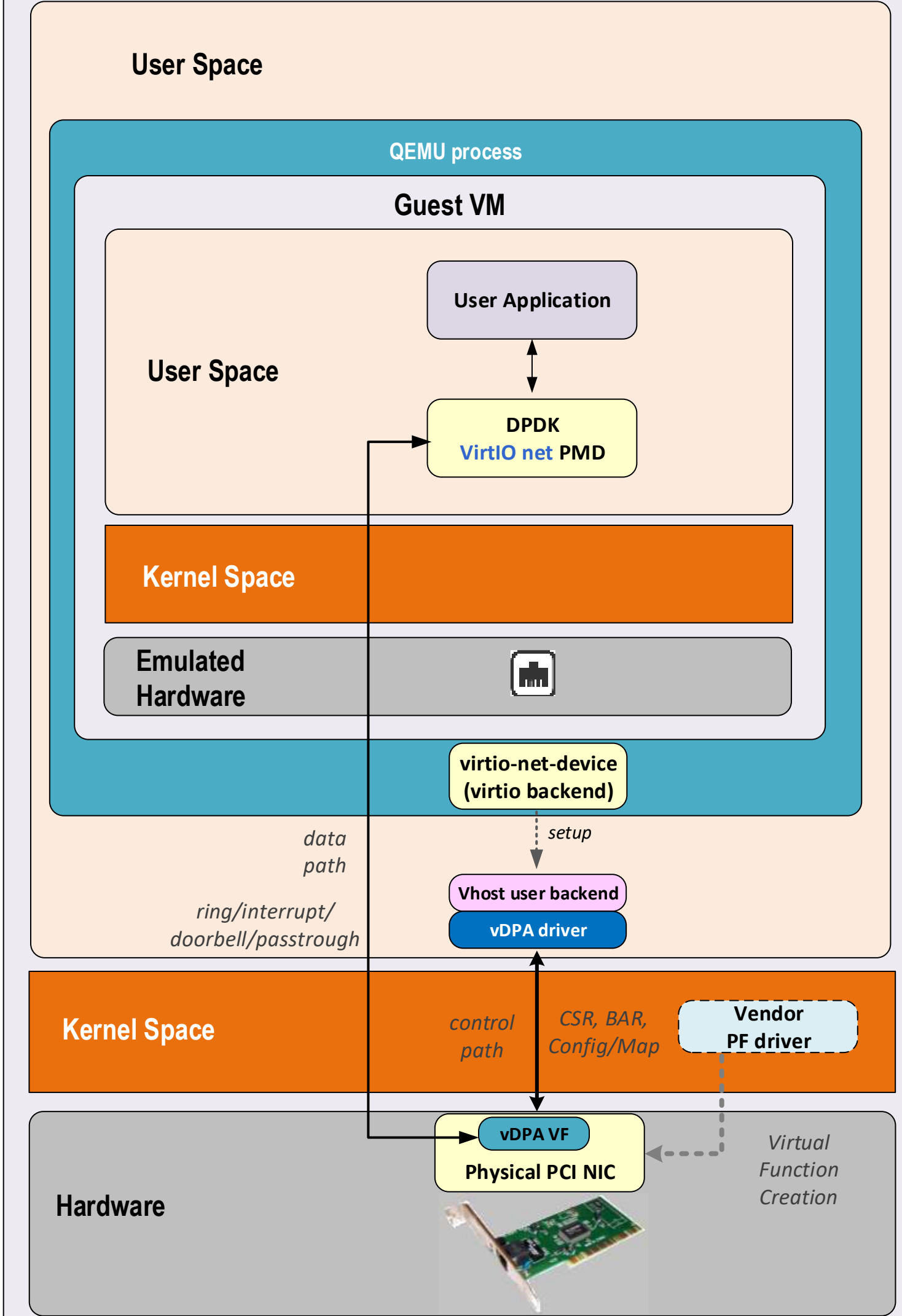
version: 1.0

date: 13/06/2020

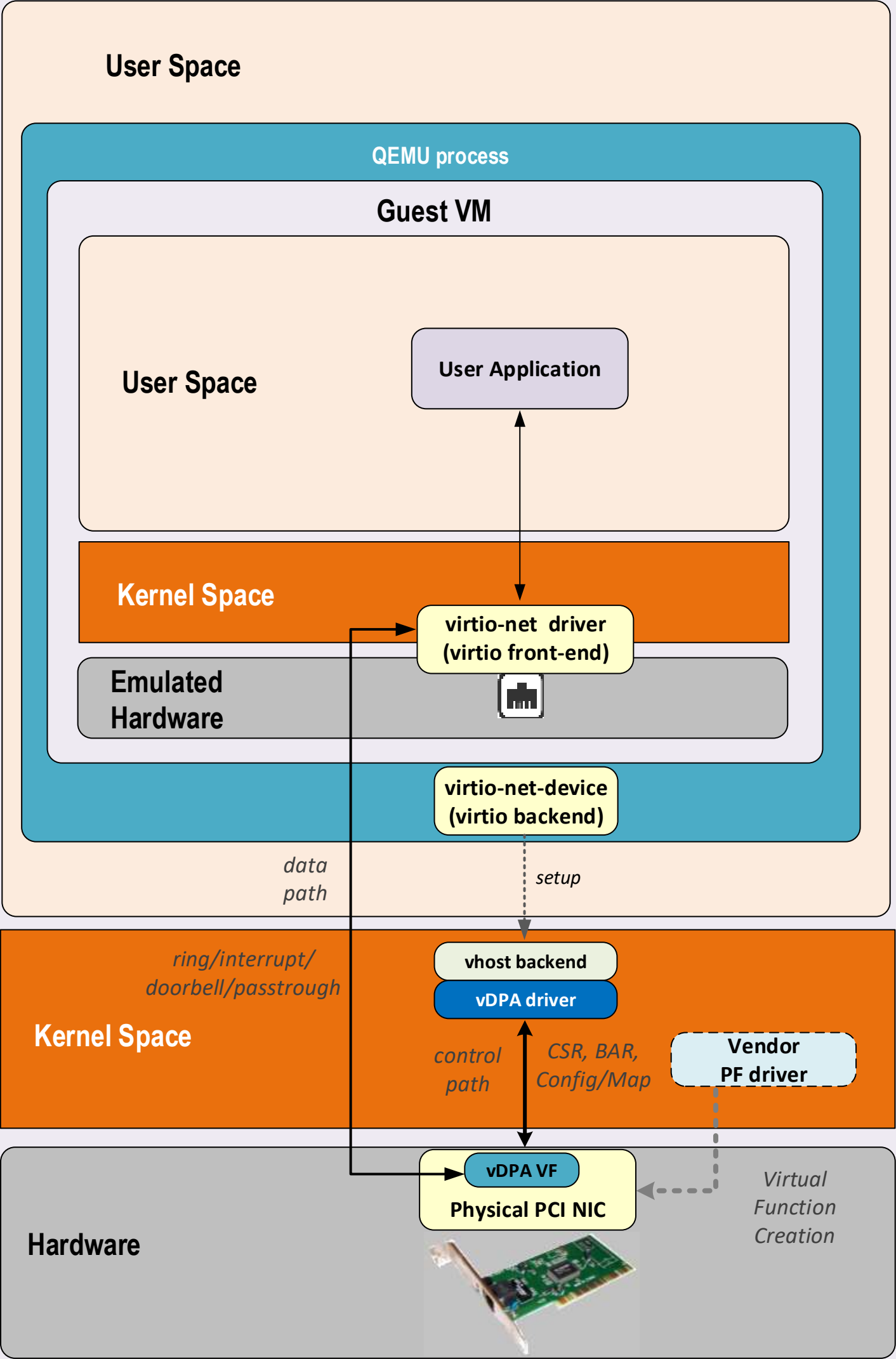
Compute Node



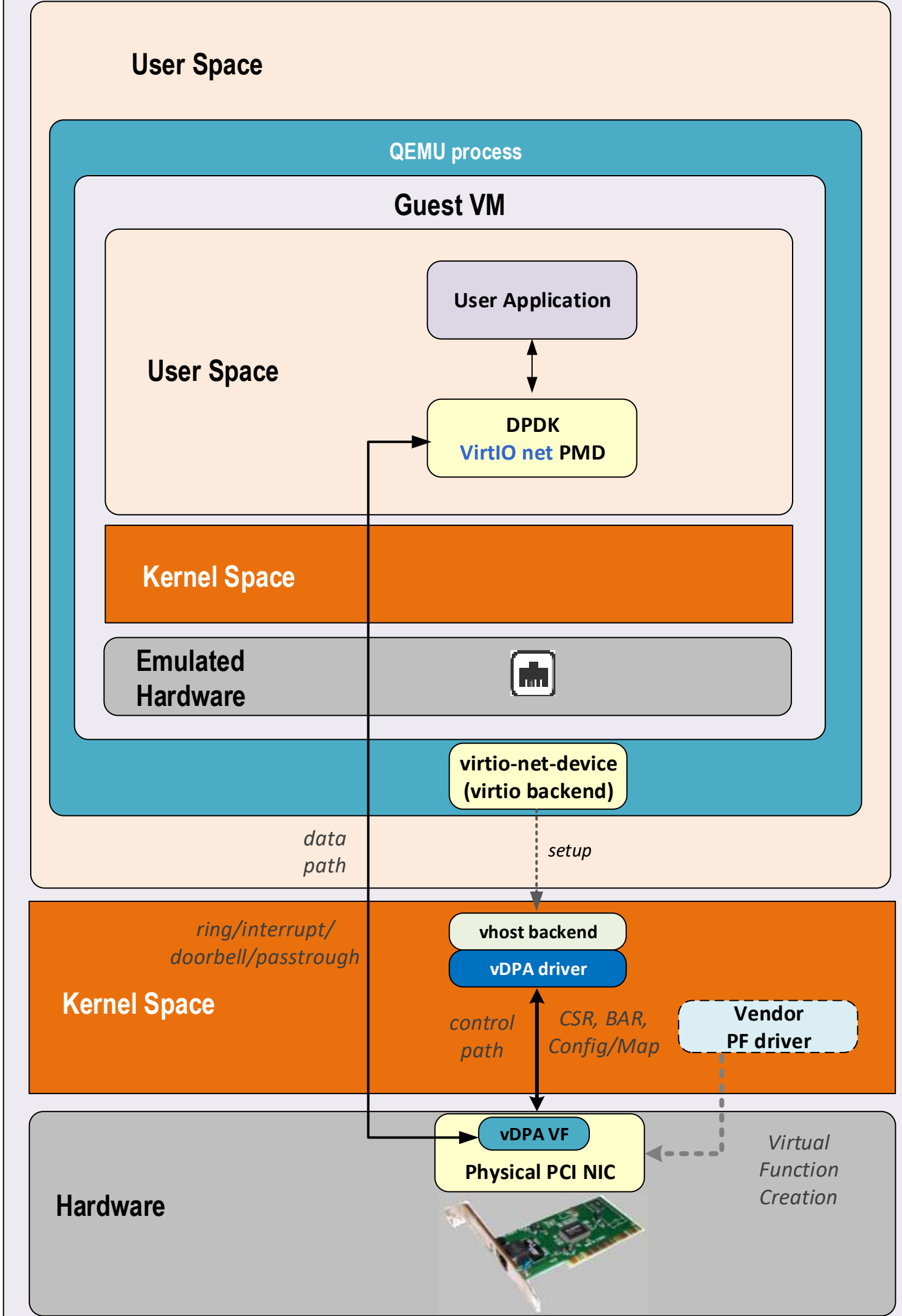
Compute Node

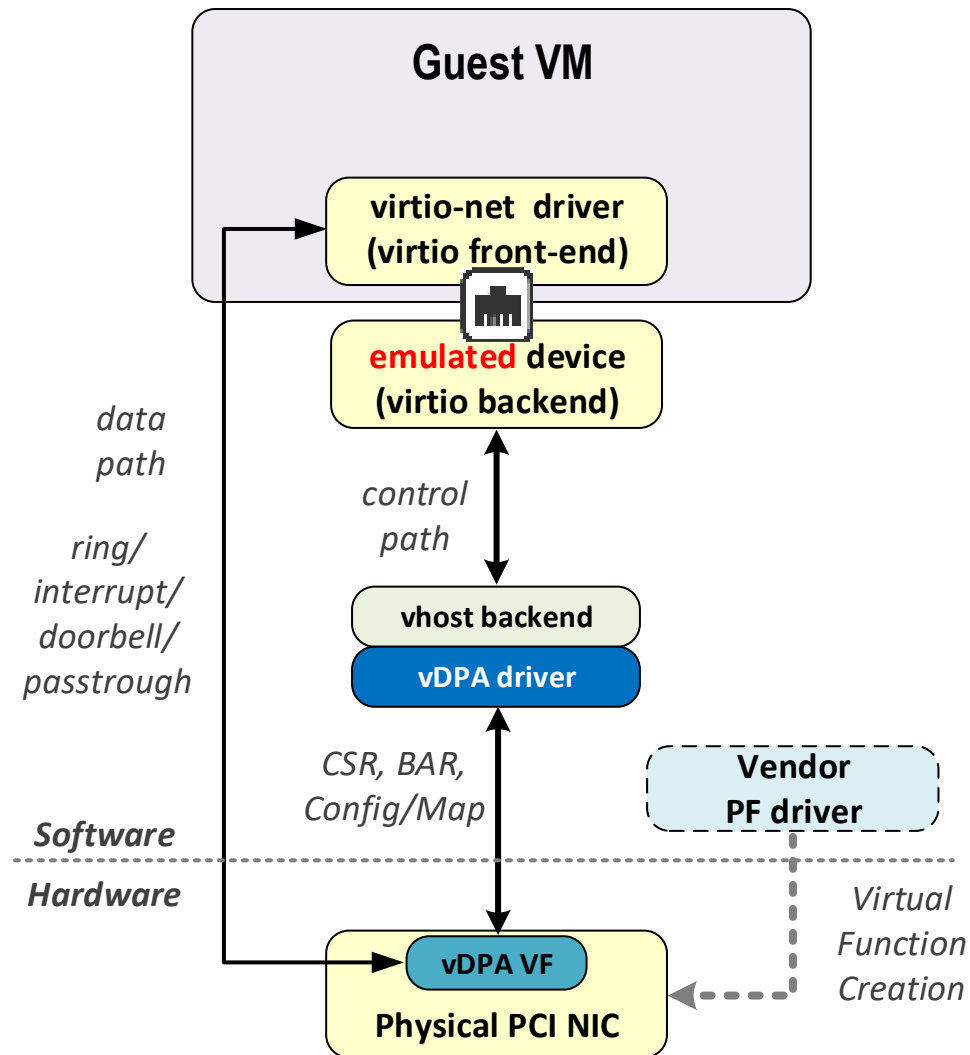
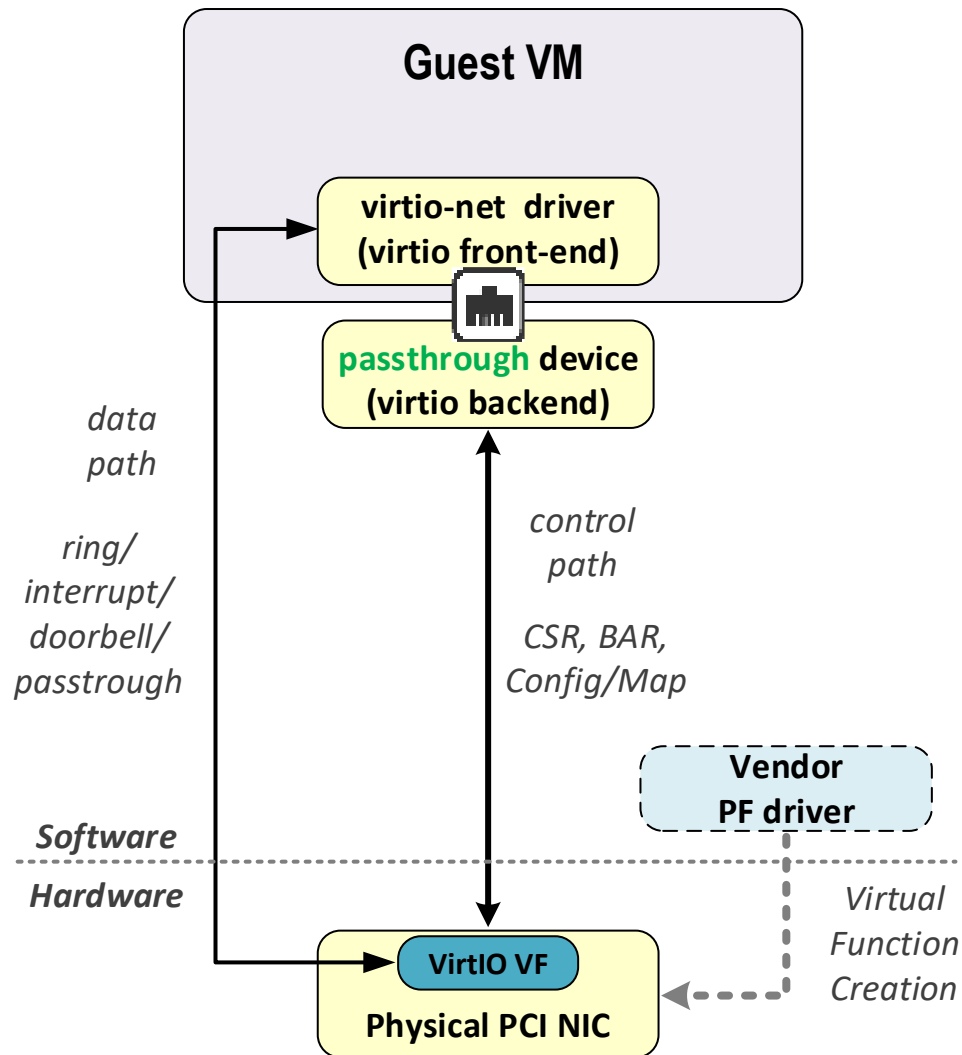


Compute Node



Compute Node





Title: VirtIO Hardware Accelleration	
version: 1.0	date: 13/06/2020

Compute Node

Direct I/O Assignment
(SR-IOV or vDPA)

vRouter/vSwitch
Control Plane Agent

Guest VM

VirtIO
driver
Virtual
PCI NIC

Guest VM

VirtIO
driver
Virtual
PCI NIC

Guest VM

VirtIO
driver
Virtual
PCI NIC

VMM

control
information

Hardware

VF1

VF2

VF3

vRouter/vSwitch
dataplane



one single physical port
shared between VM

Title: Smart NIC

version: 1.0

date: 13/06/2020

Compute Node

User Space

User Application

eBPF program

Socket
API

Kernel Space

Driver

XDP

eBPF rules
DROP,
FORWARD

Ring
buffers

DMA transfer

Hardware

RX/TX
physical queues



Title: XDP and eBPF

version: 1.0

date: 13/06/2020

Compute Node

User Space

User Application

eBPF
program

Kernel Space

Socket
API

Driver

XDP

triggers

Ring
buffers

Hardware

DMA transfer

RX/TX
physical queues

eBPF rules
DROP,
FORWARD

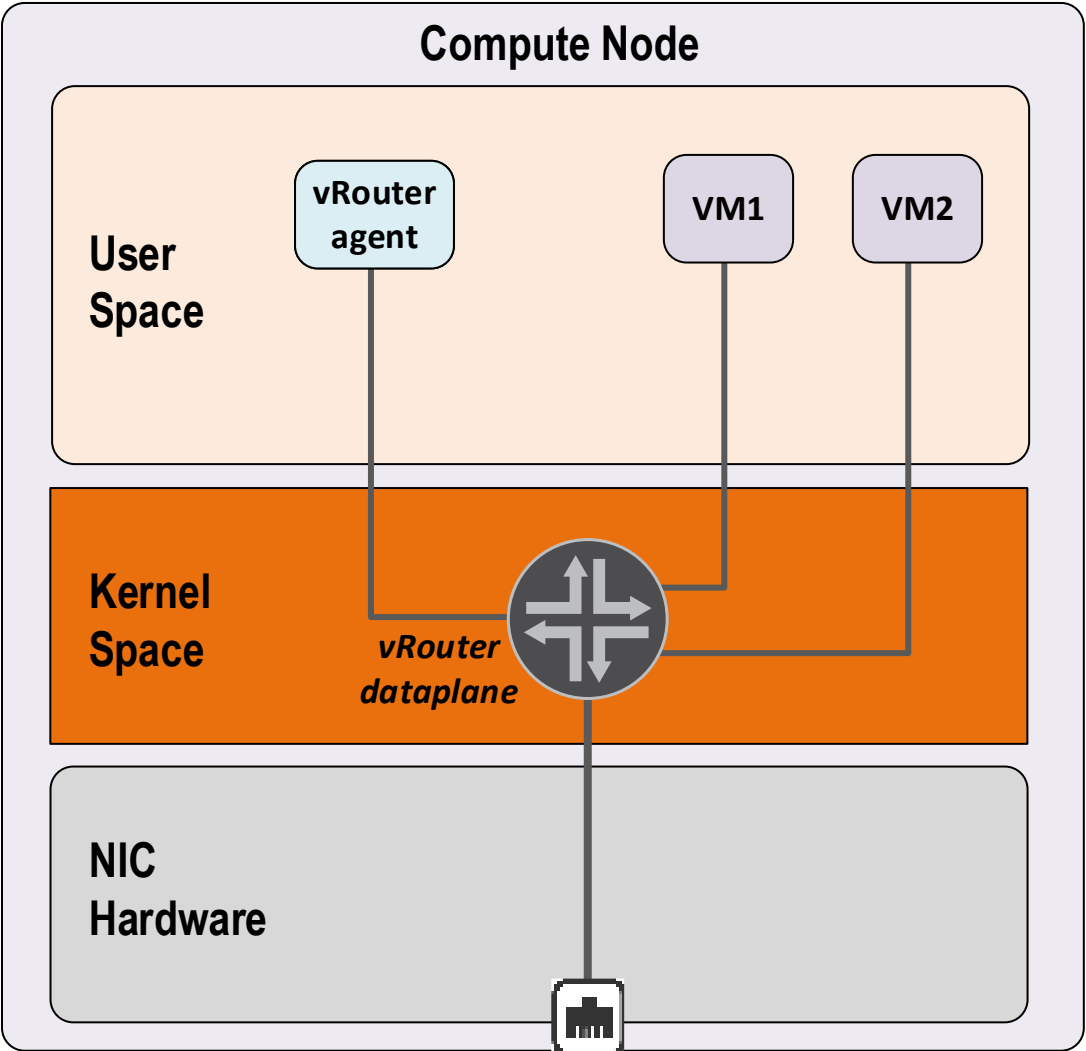
Title: XDP and eBPF offload

version: 1.0

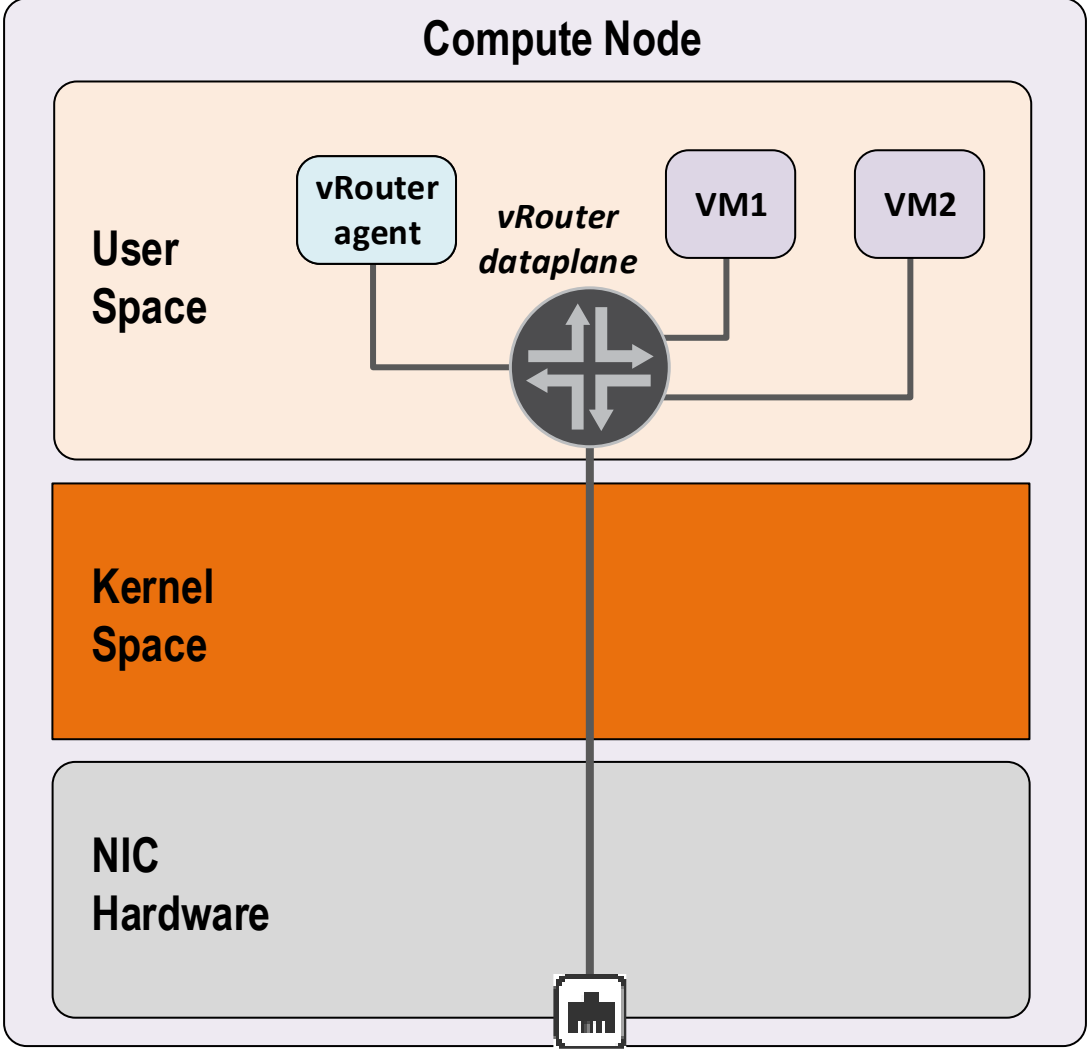
date: 13/06/2020



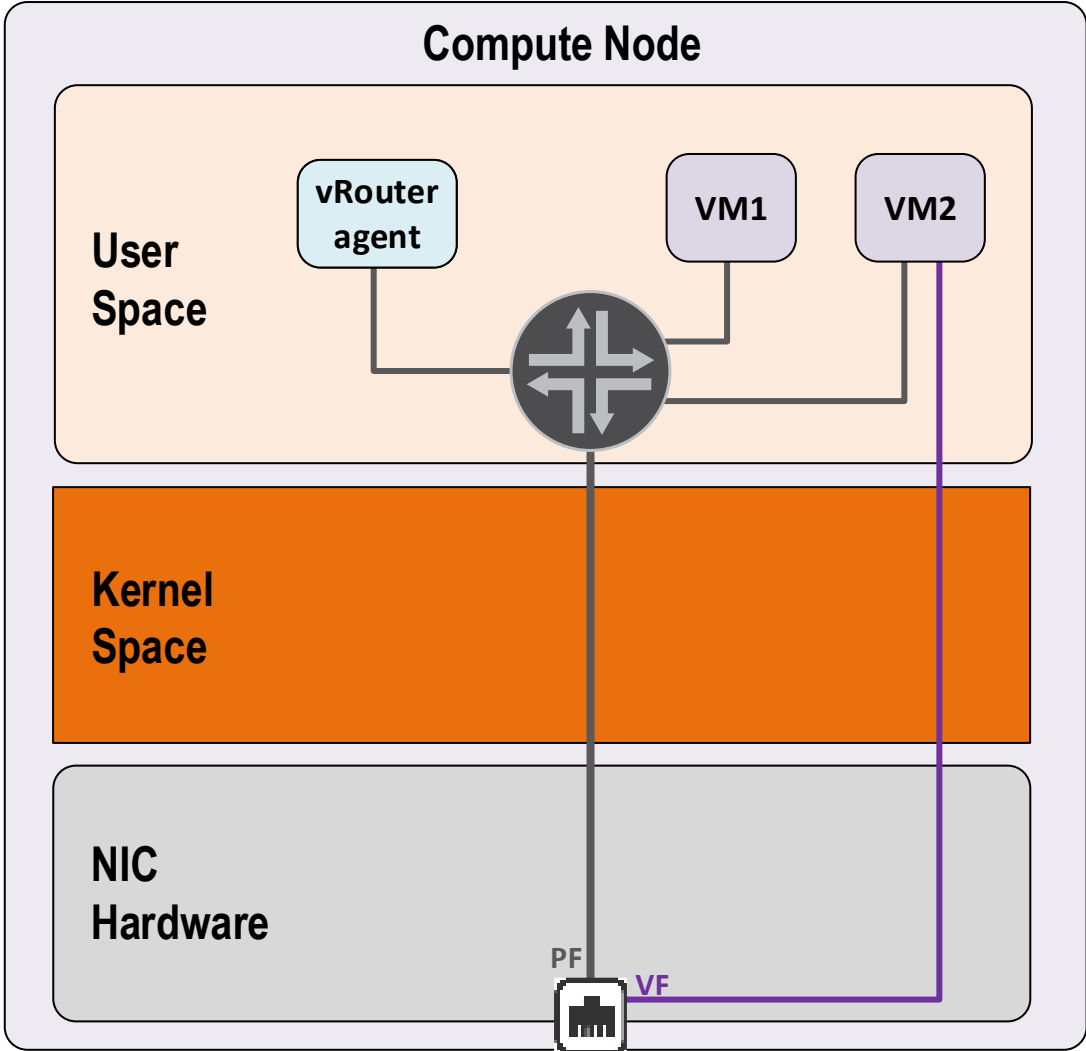
Kernel mode



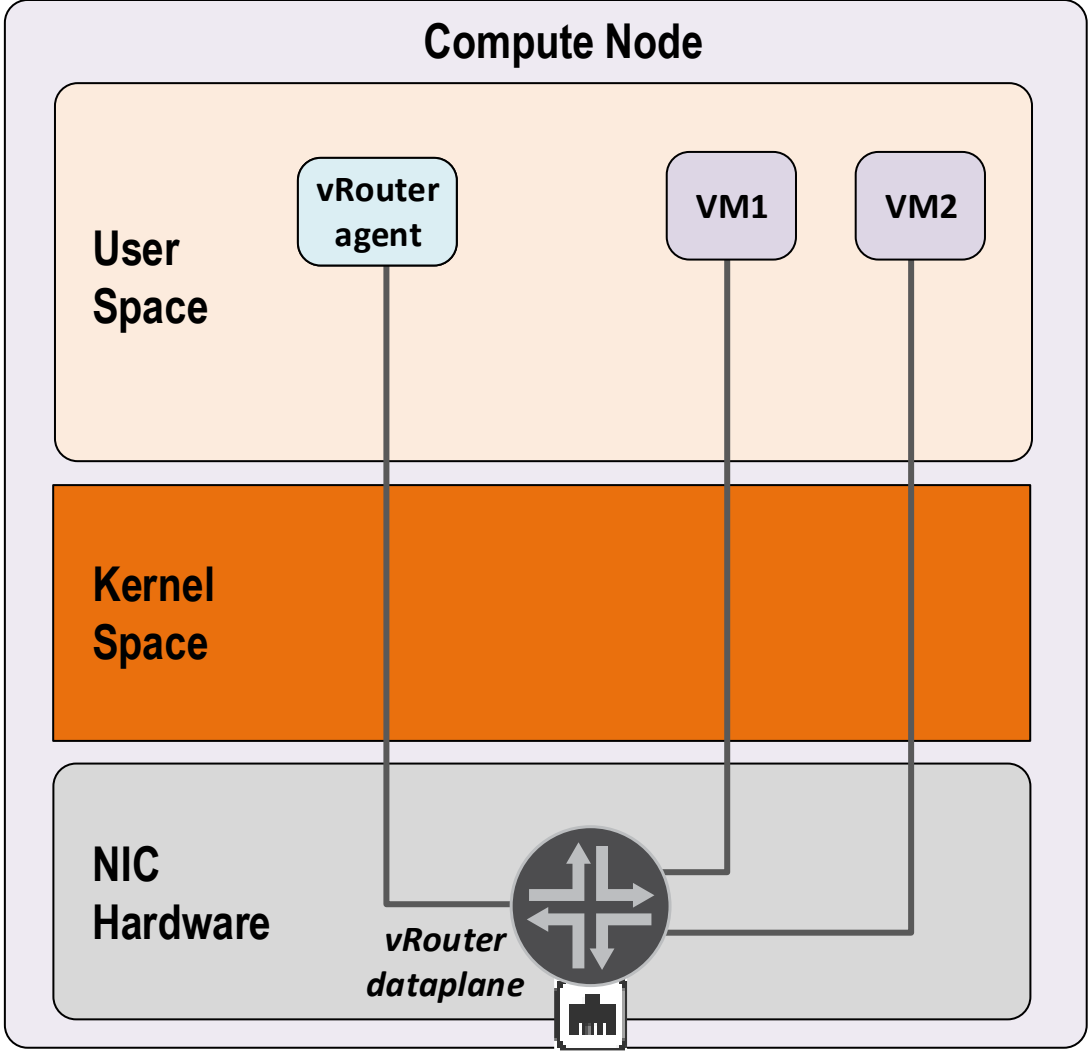
DPDK mode

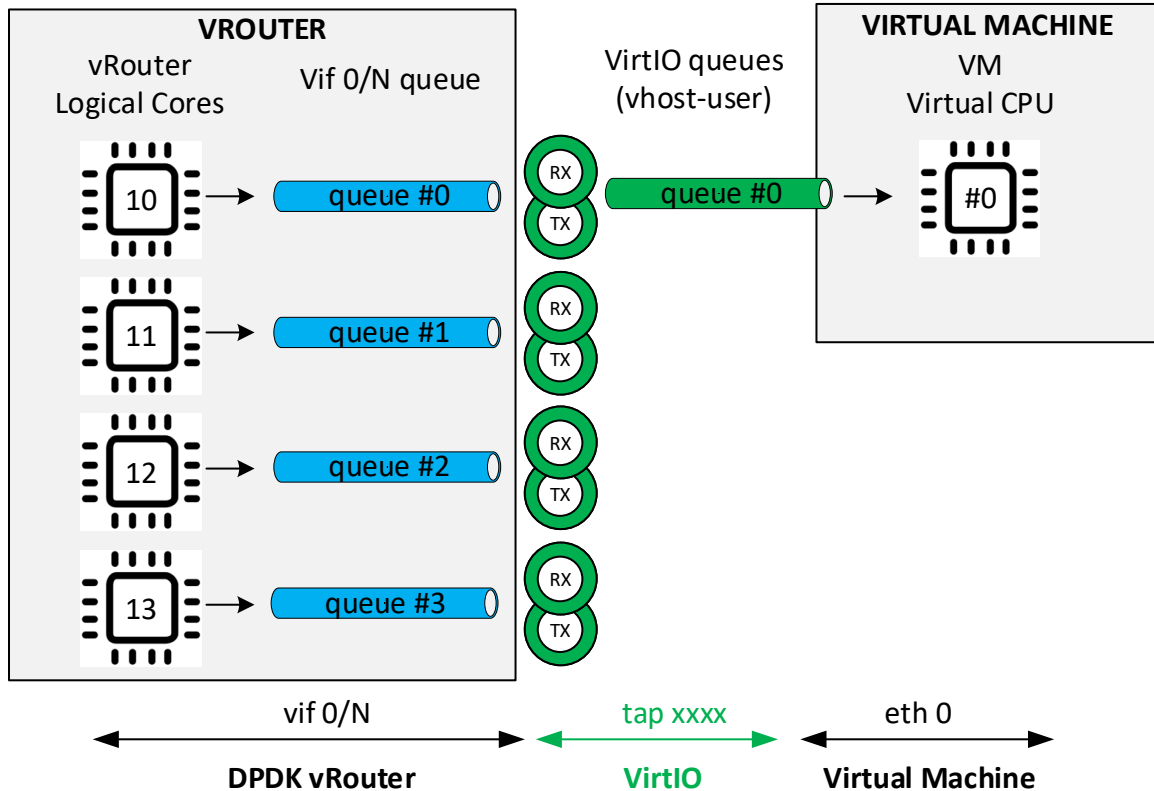


Hybrid SRIOV-DPDK

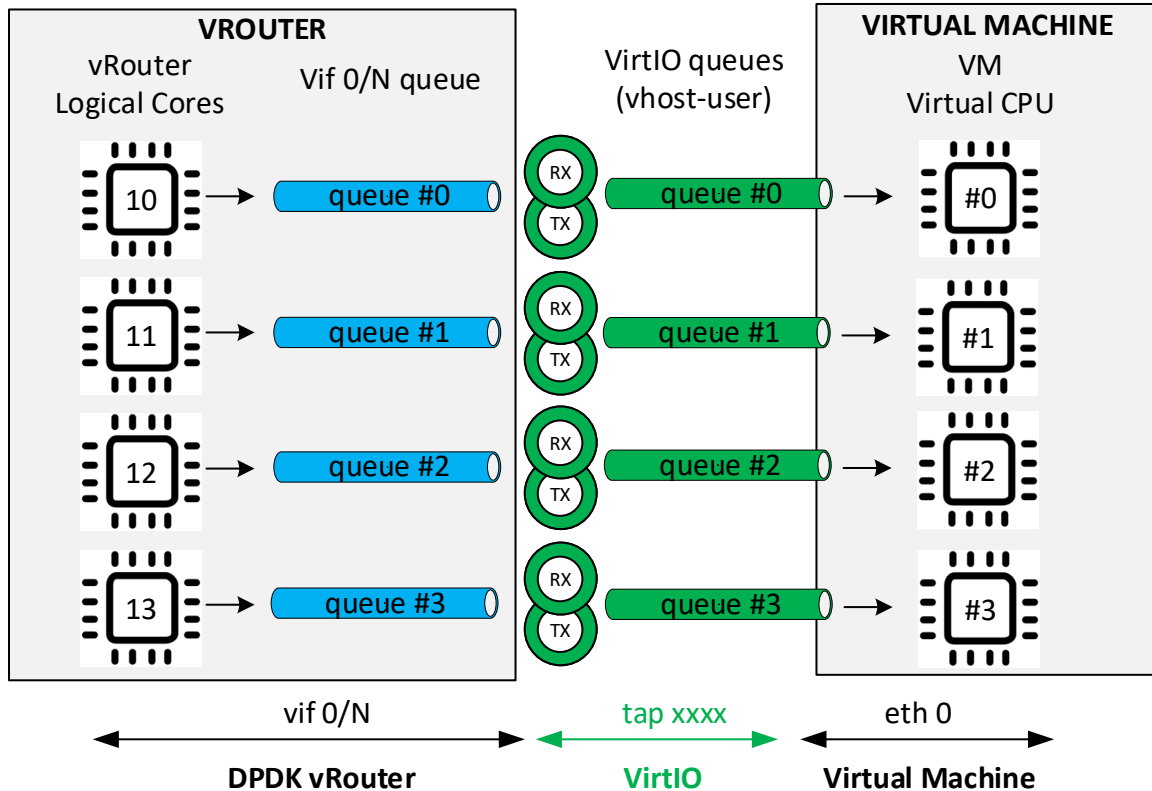


Smartnic

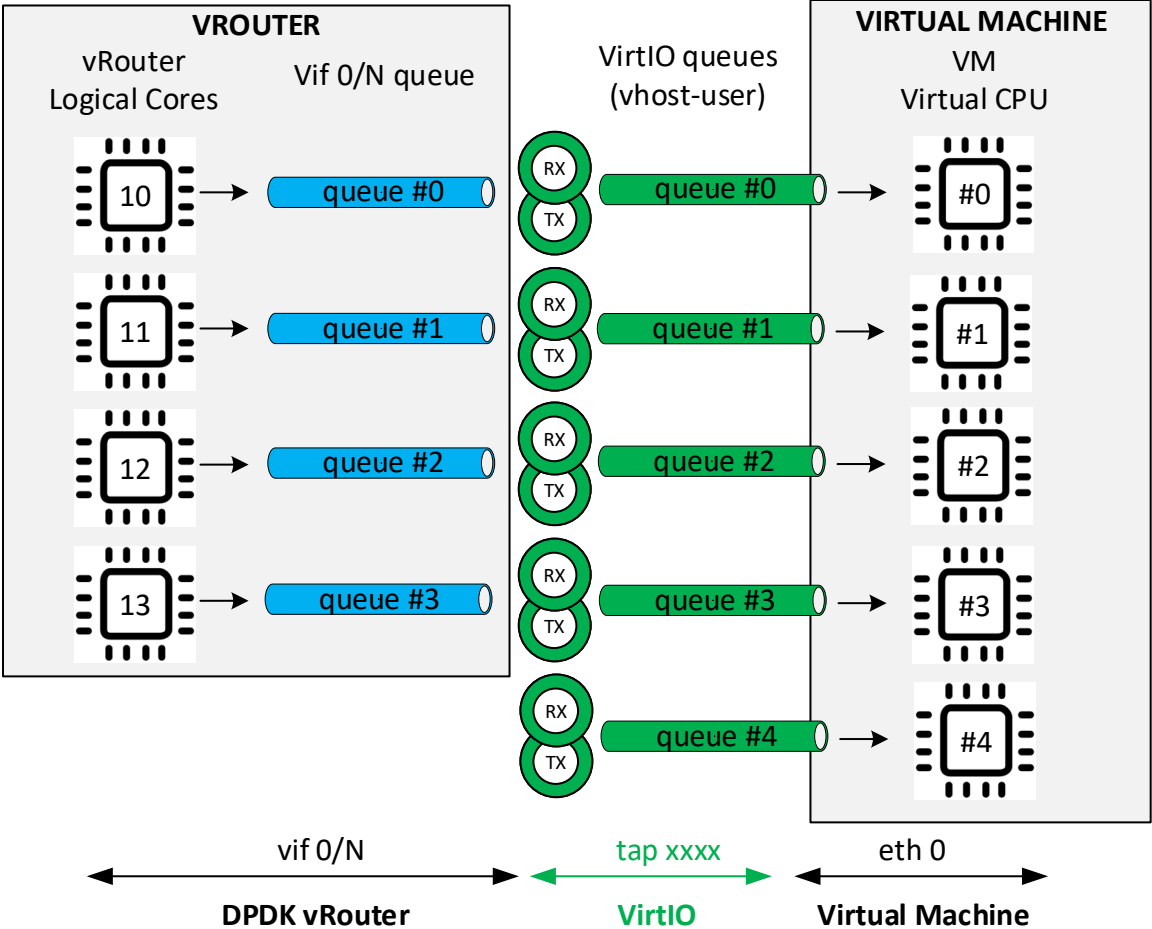




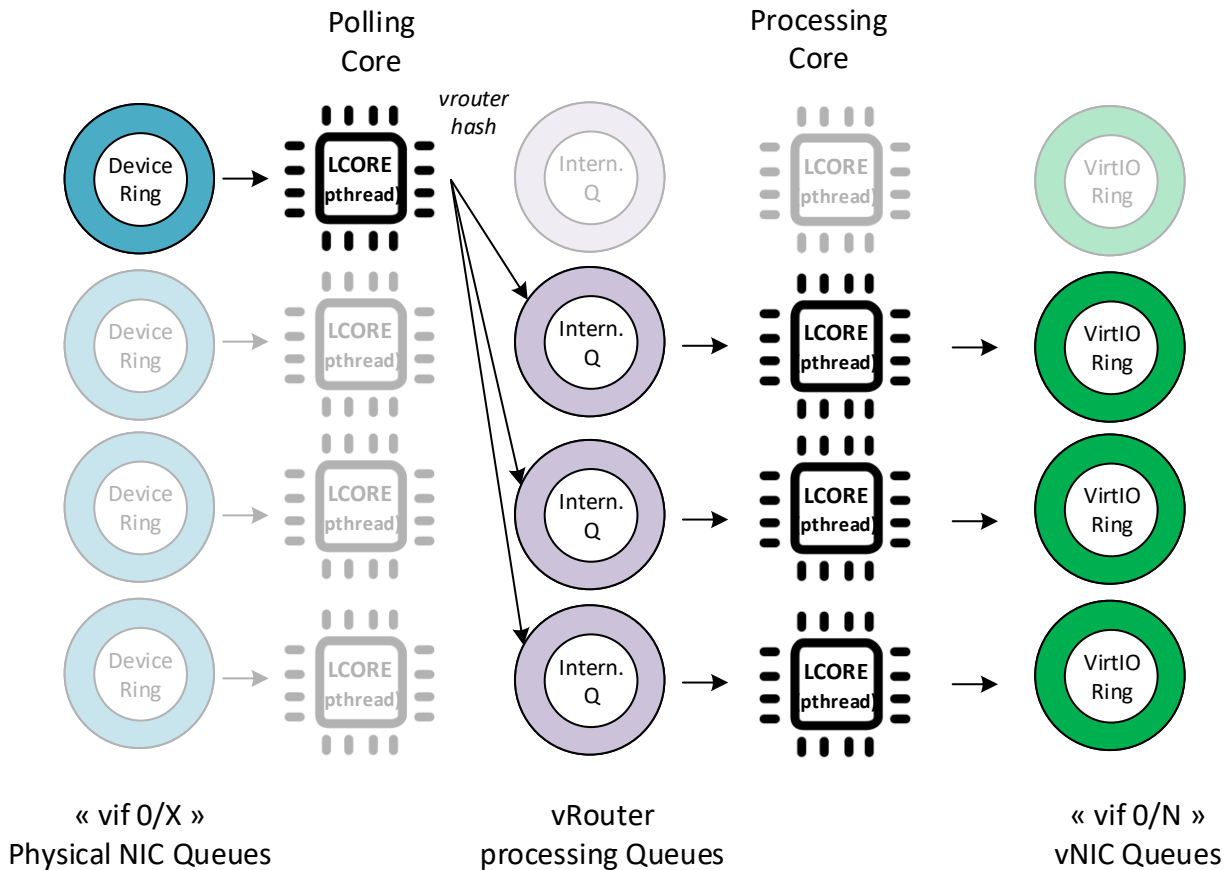
Title: DPDK Single Queue mapping	
version: 1.0	date: 16/09/2020



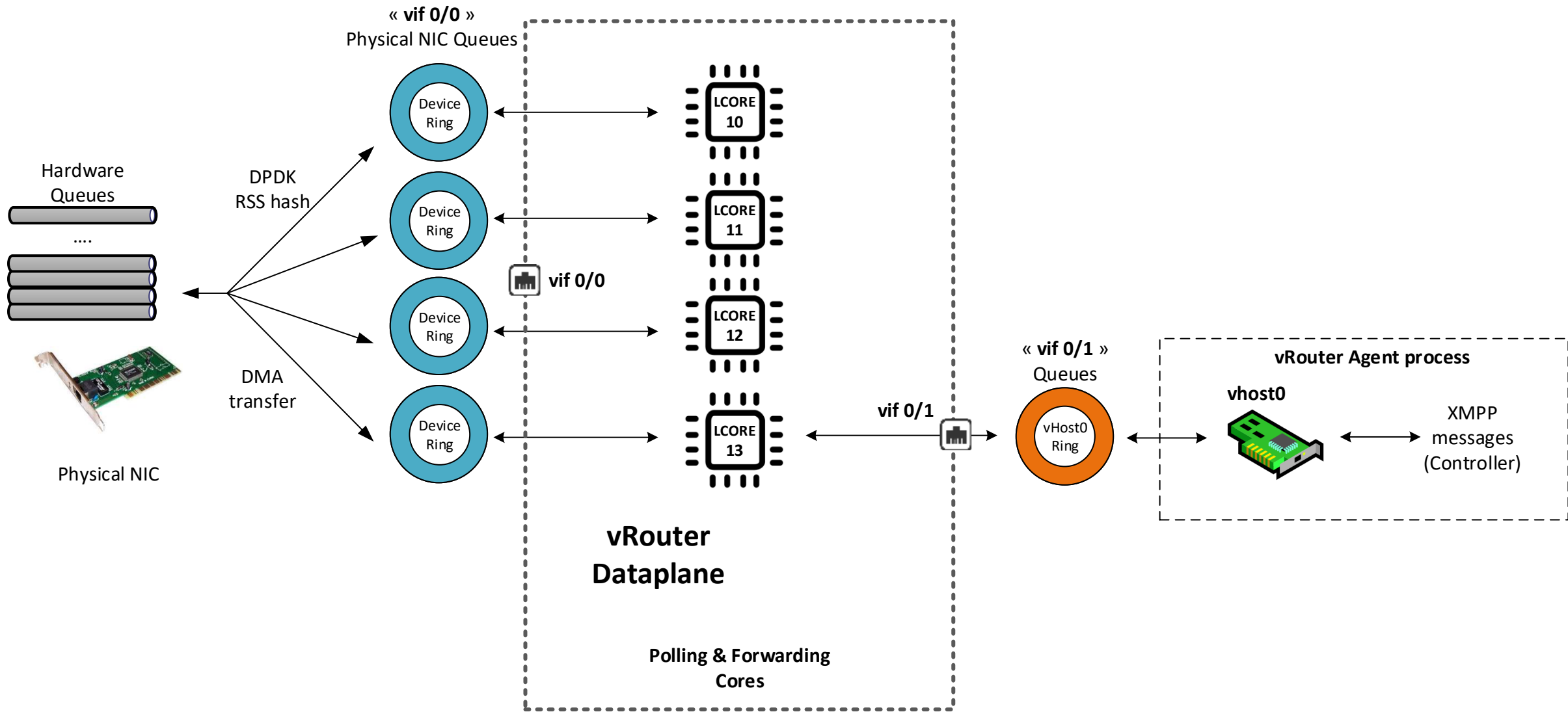
Title: DPDK MultiQueue mapping	
version: 1.0	date: 16/09/2020



Title: DPDK Faulty MultiQueue mapping	
version: 1.0	date: 16/09/2020



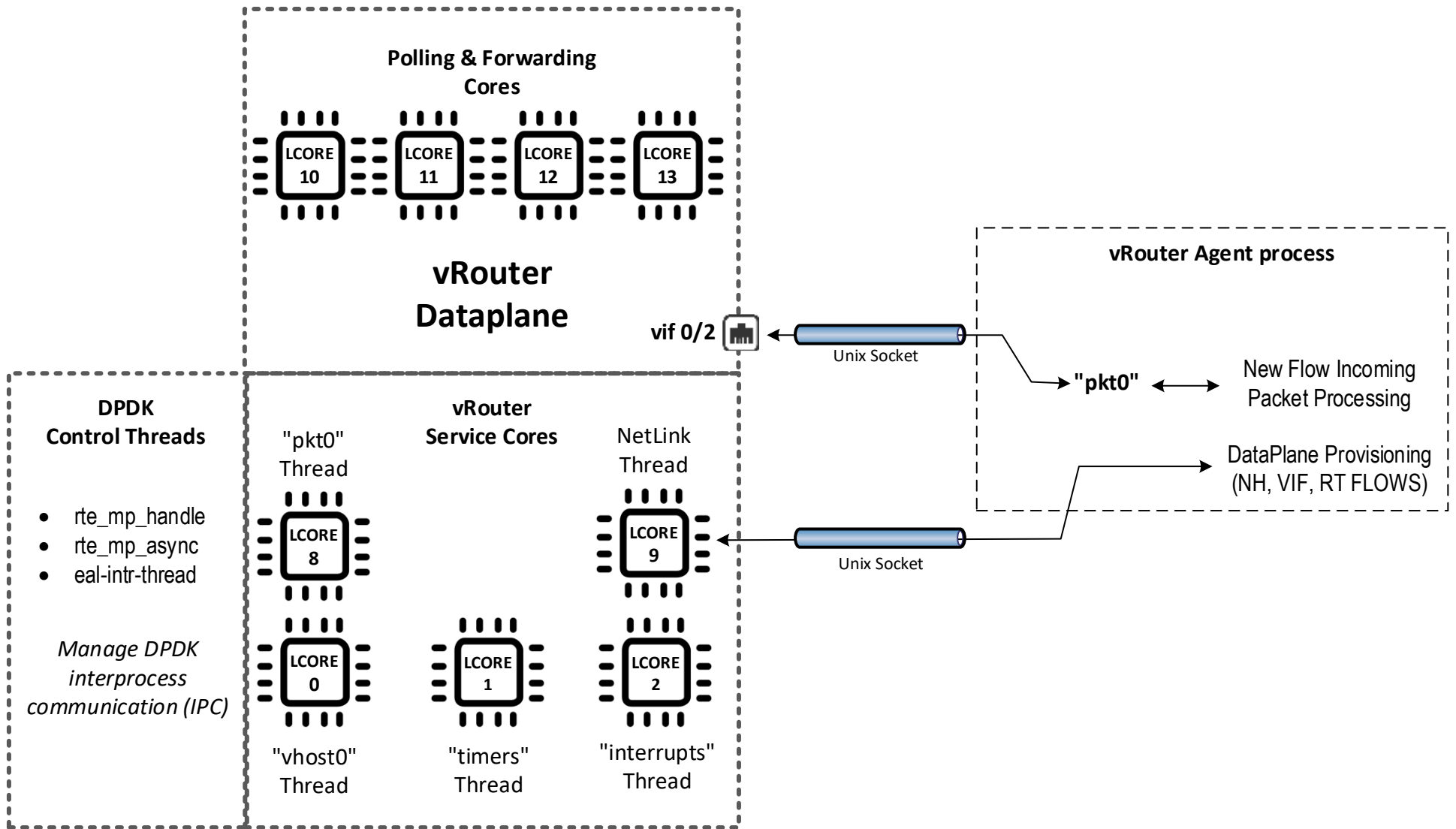
Title: vRouter internal loadbalancing	
version: 1.0	date: 16/09/2020



Title: vRouter Agent XMPP Messages Path

version: 1.0

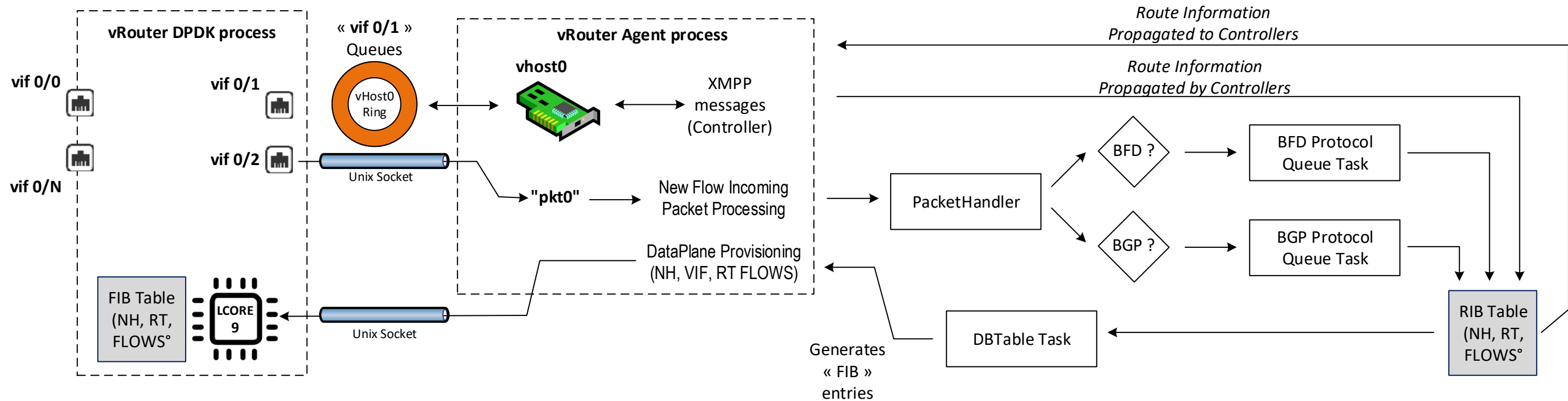
date: 16/09/2020



Title: vRouter Agent New Flow Processing Path

version: 1.0

date: 16/09/2020

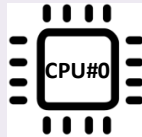


Title: vRouter Agent Protocol Processing

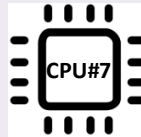
version: 1.0

date: 16/09/2020

Compute Node



...



instance-000000a7 ID:43

g2-vm1

2 vCPUs

Server

iperf -s -u

ens3

120.0.0.3/24

2 queues

169.254.0.3

instance-000000a8 ID:44

g2-vm0

2 vCPUs

Client

**while true; do iperf -c 120.0.0.3 \
-P 2 -u -t 1 -b 20M ; done**

ens3

120.0.0.4/24

1 queue

169.254.0.8

tap295c750f-23

0/3

vRouter

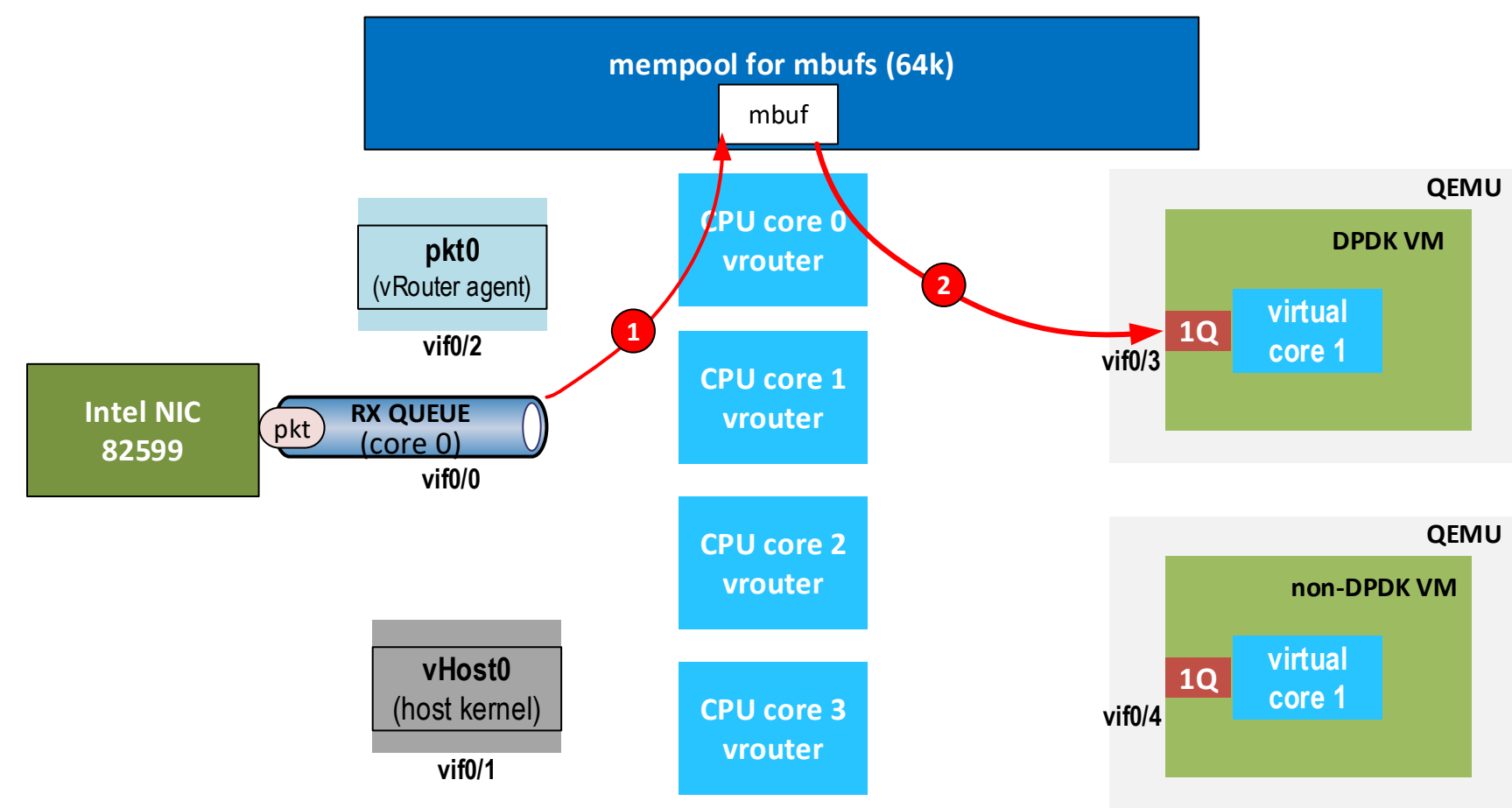
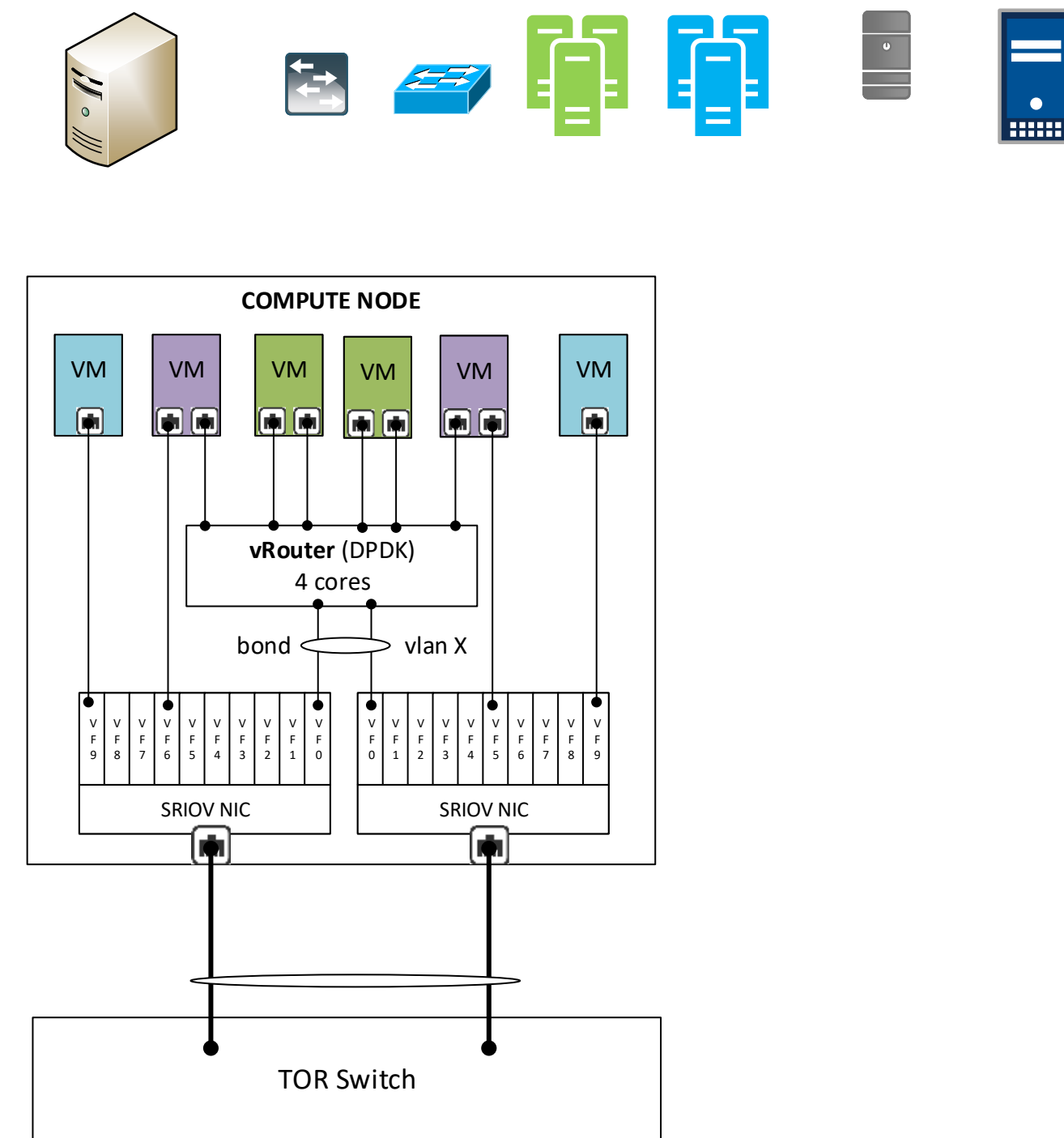
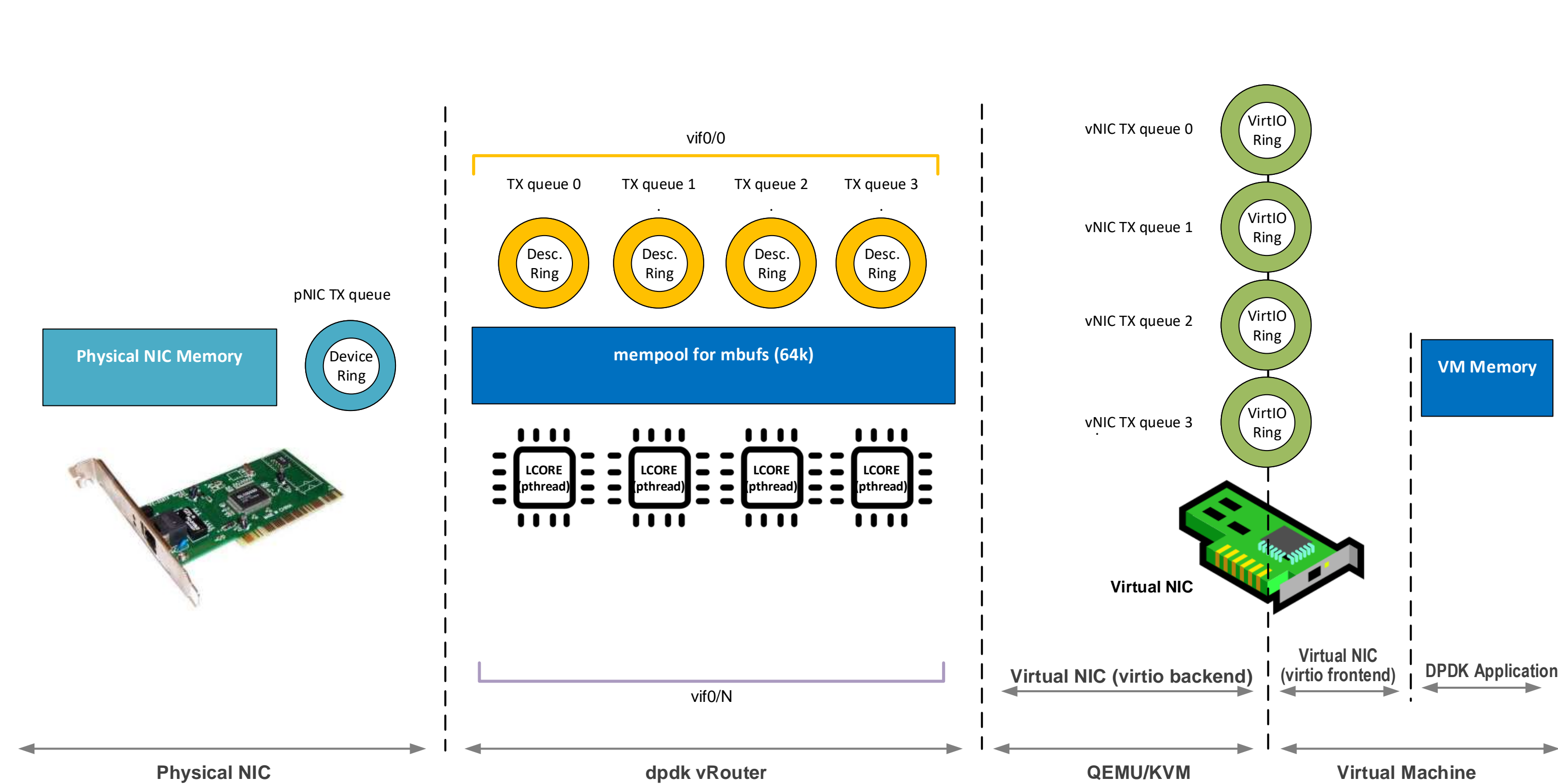
0/8

tapb7188999_d0

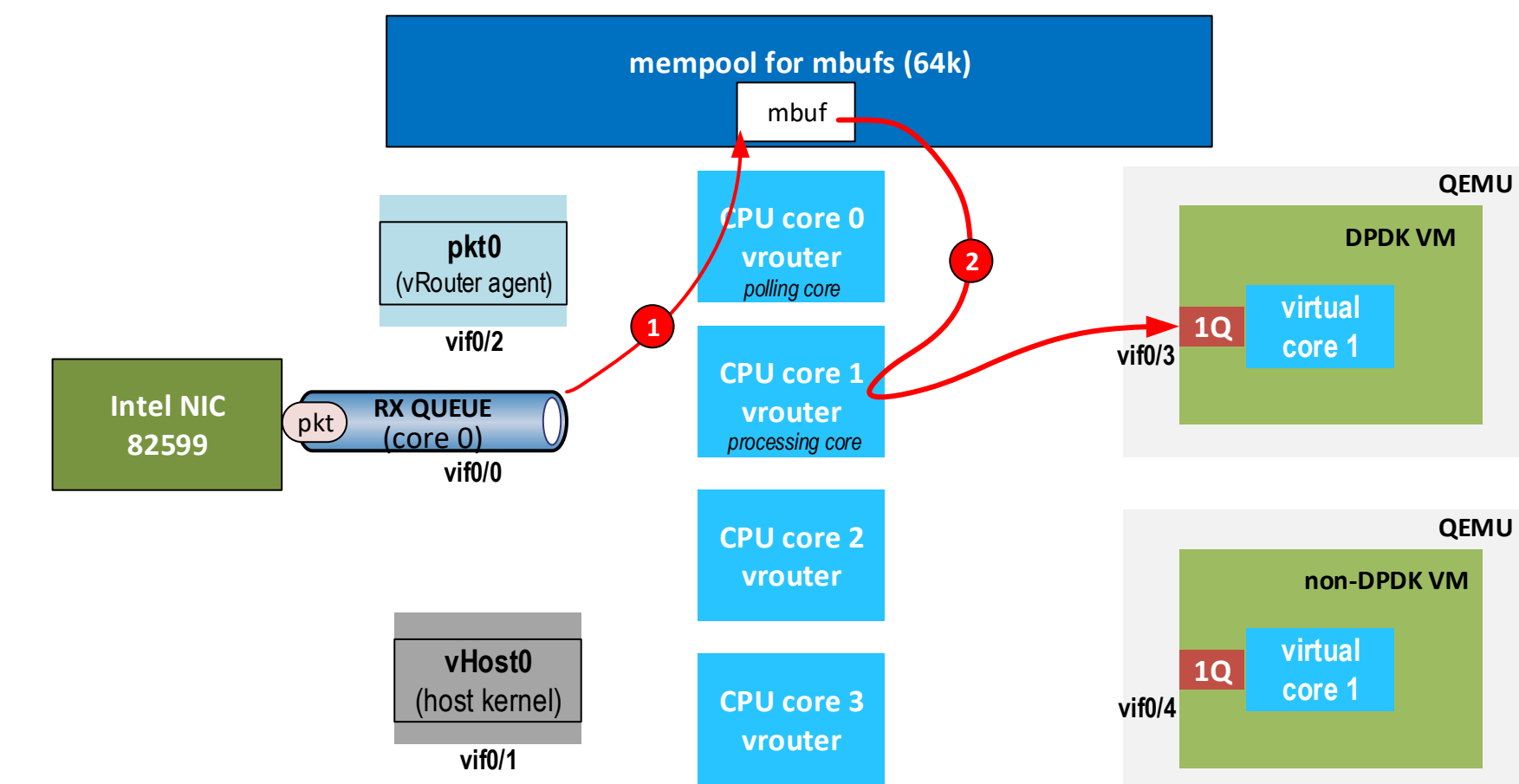
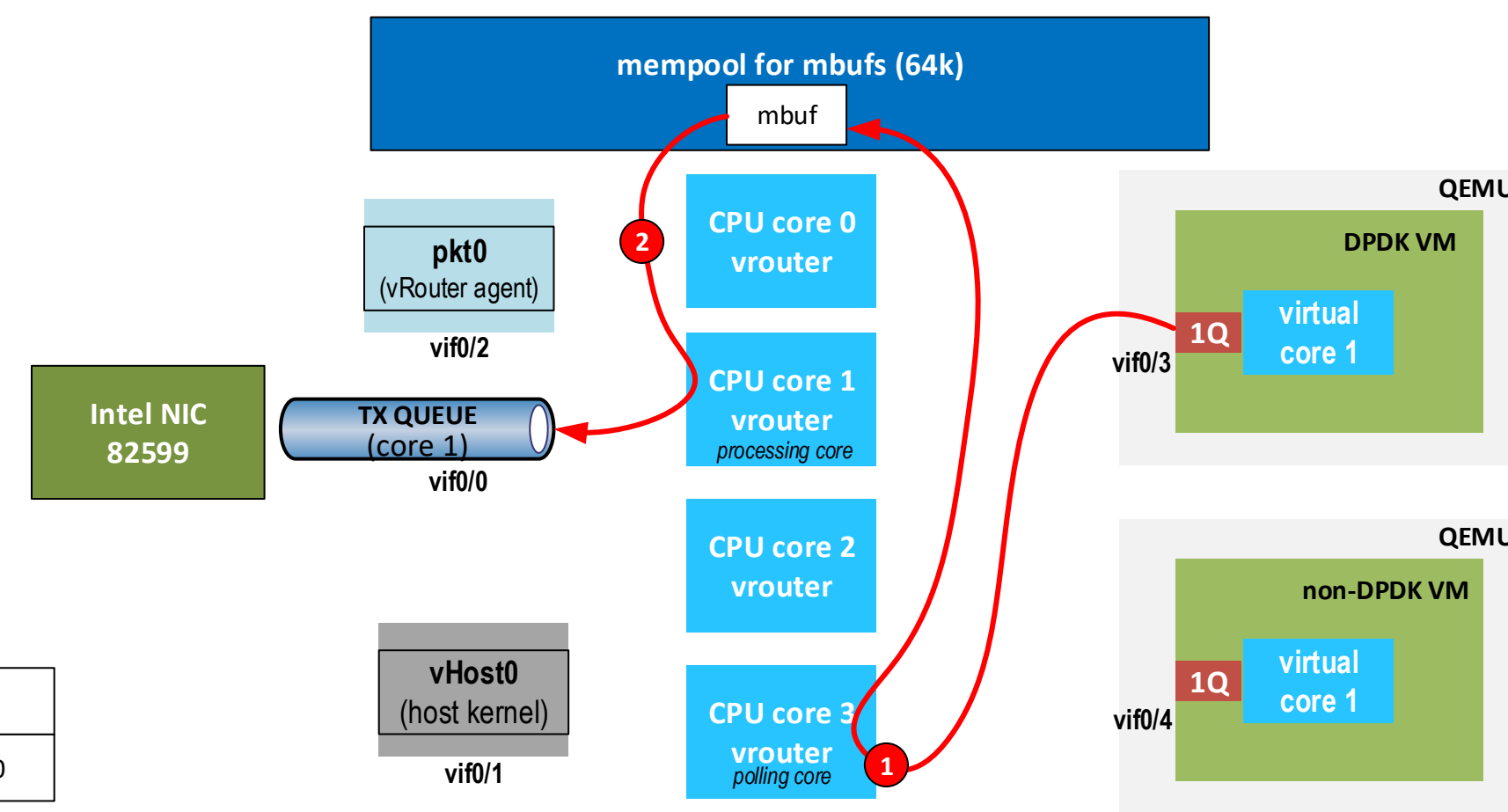
Title: Test Lab

version: 1.0

date: 23/09/2020



Title: Personal Stuff	
version: 1.0	date: 25/03/2020



KERNEL MODE

DPDK MODE

SRIOV (vrouter bypassed)

