Cloud Networking
Host Virtualization

Philip Godfrey and Ankit Singla
Department of Computer Science

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

The cloud depends on it, virtually

The cloud depends on it, virtually

- Sharing of physical infrastructure

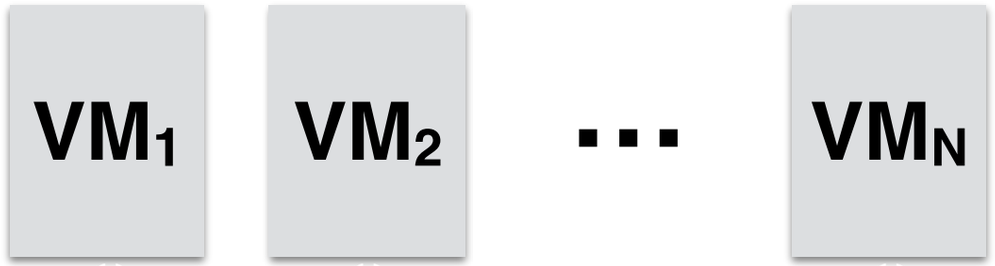
The cloud depends on it, virtually

- Sharing of physical infrastructure
- Spin-up a virtual machine in seconds

The cloud depends on it, virtually

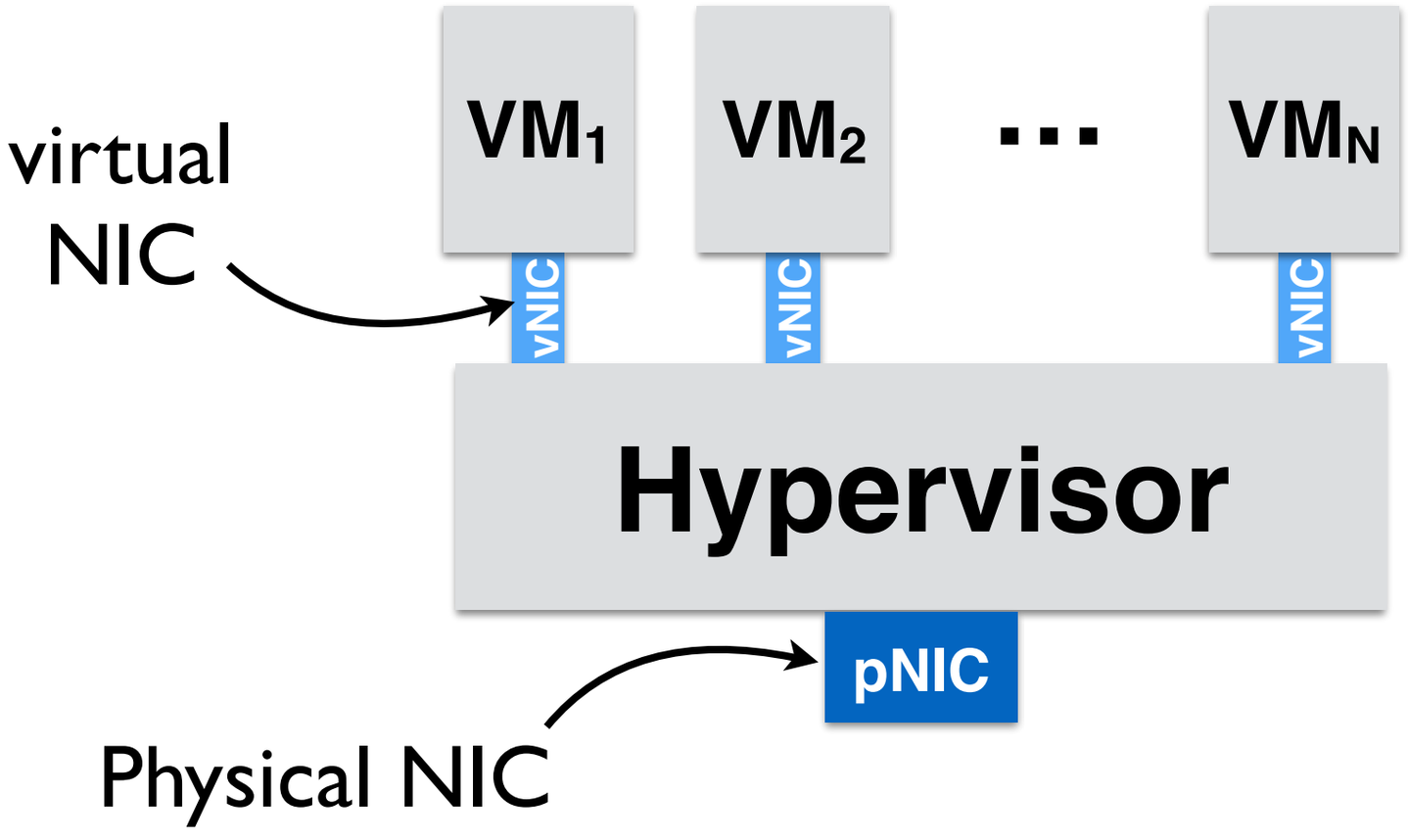
- Sharing of physical infrastructure
- Spin-up a virtual machine in seconds
- Live VM migration

Server virtualization

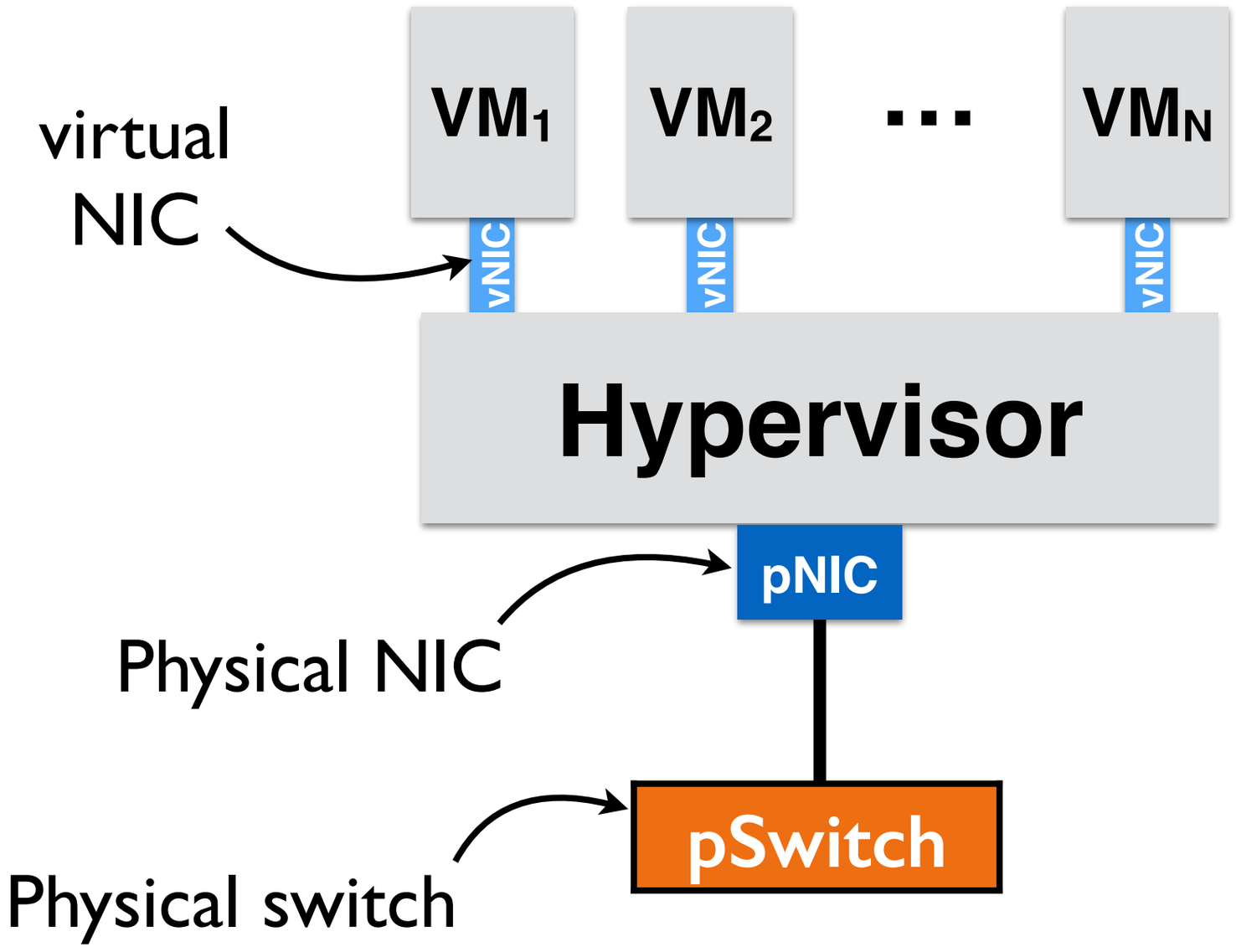


Hypervisor

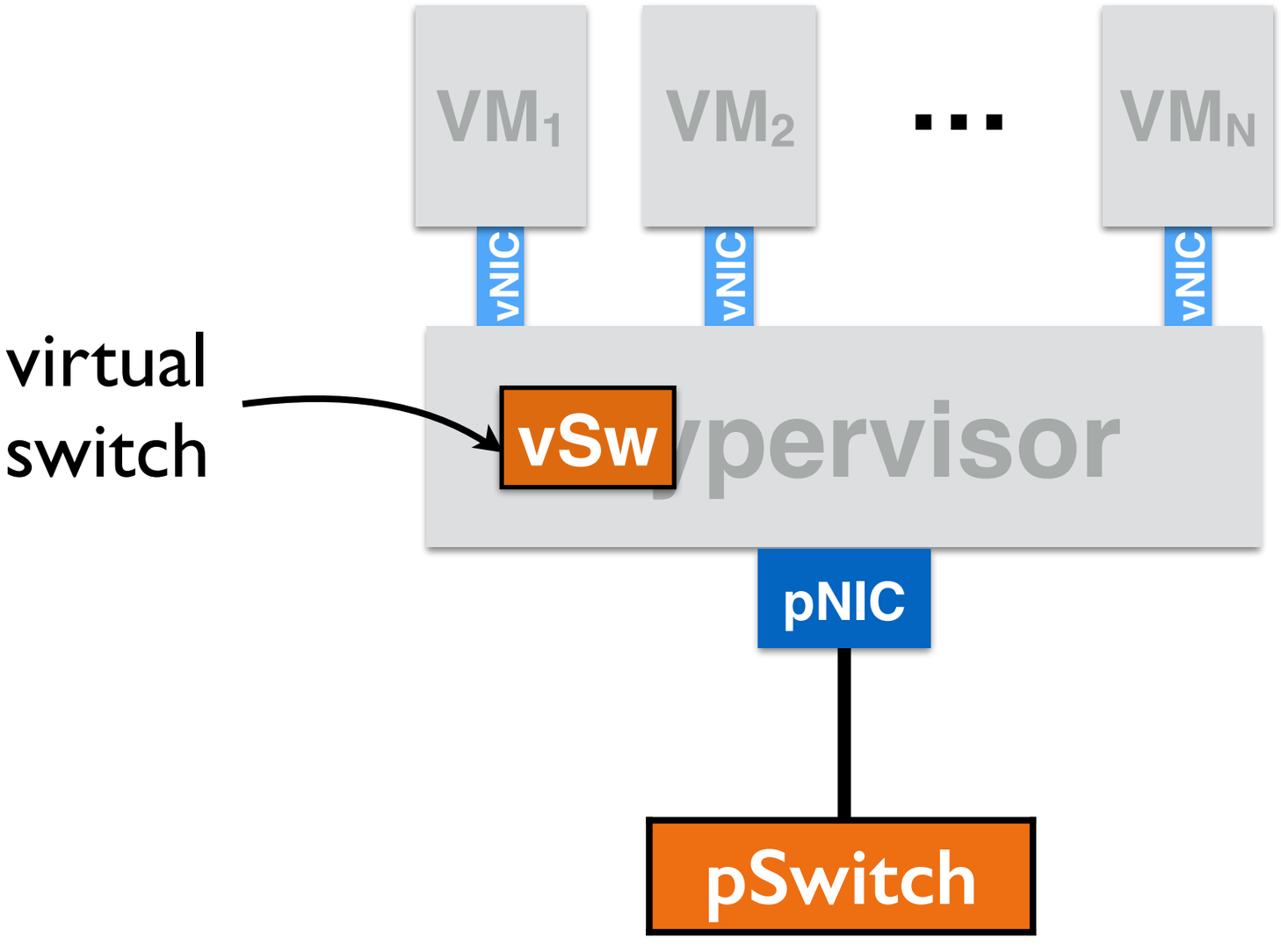
Server virtualization



Server virtualization

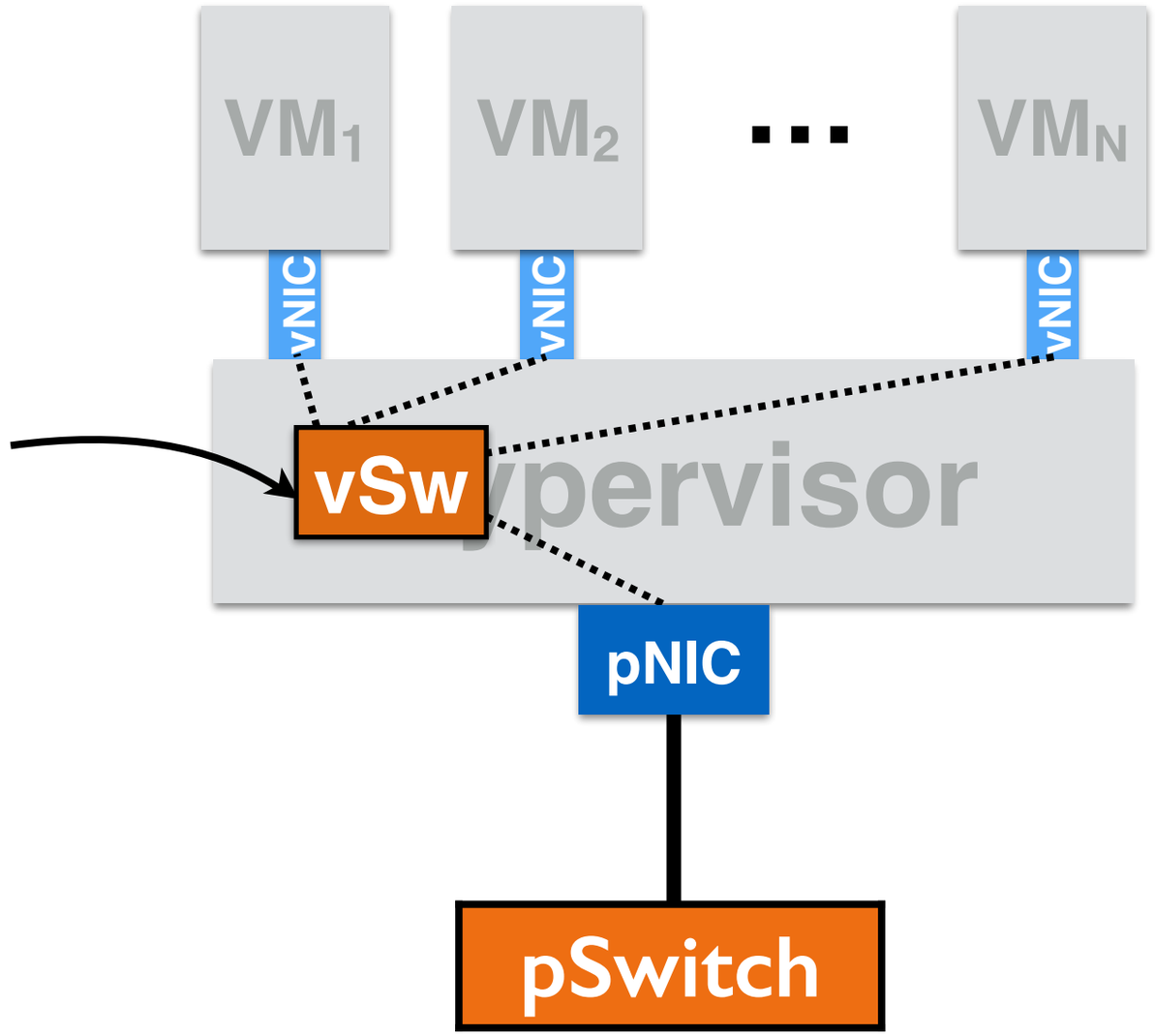


Networking VMs

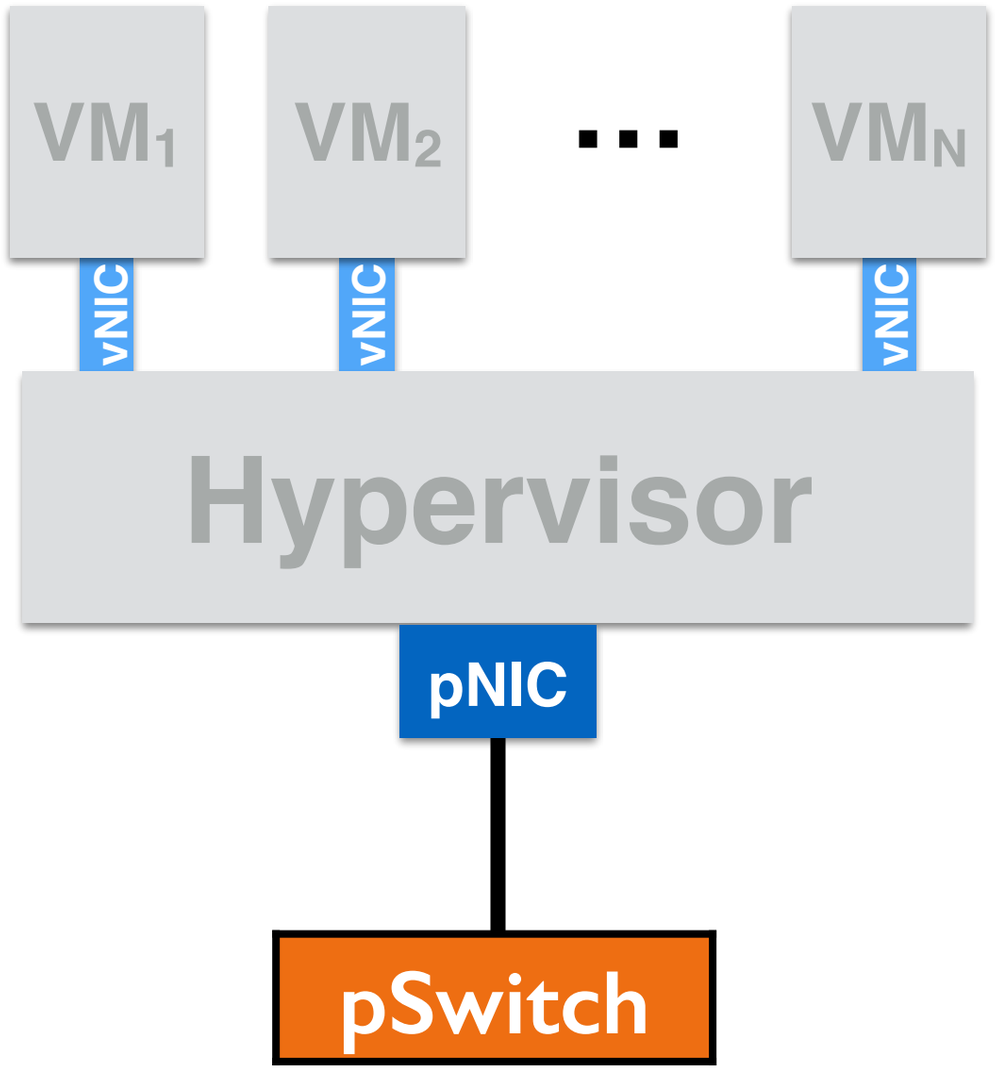


Networking VMs

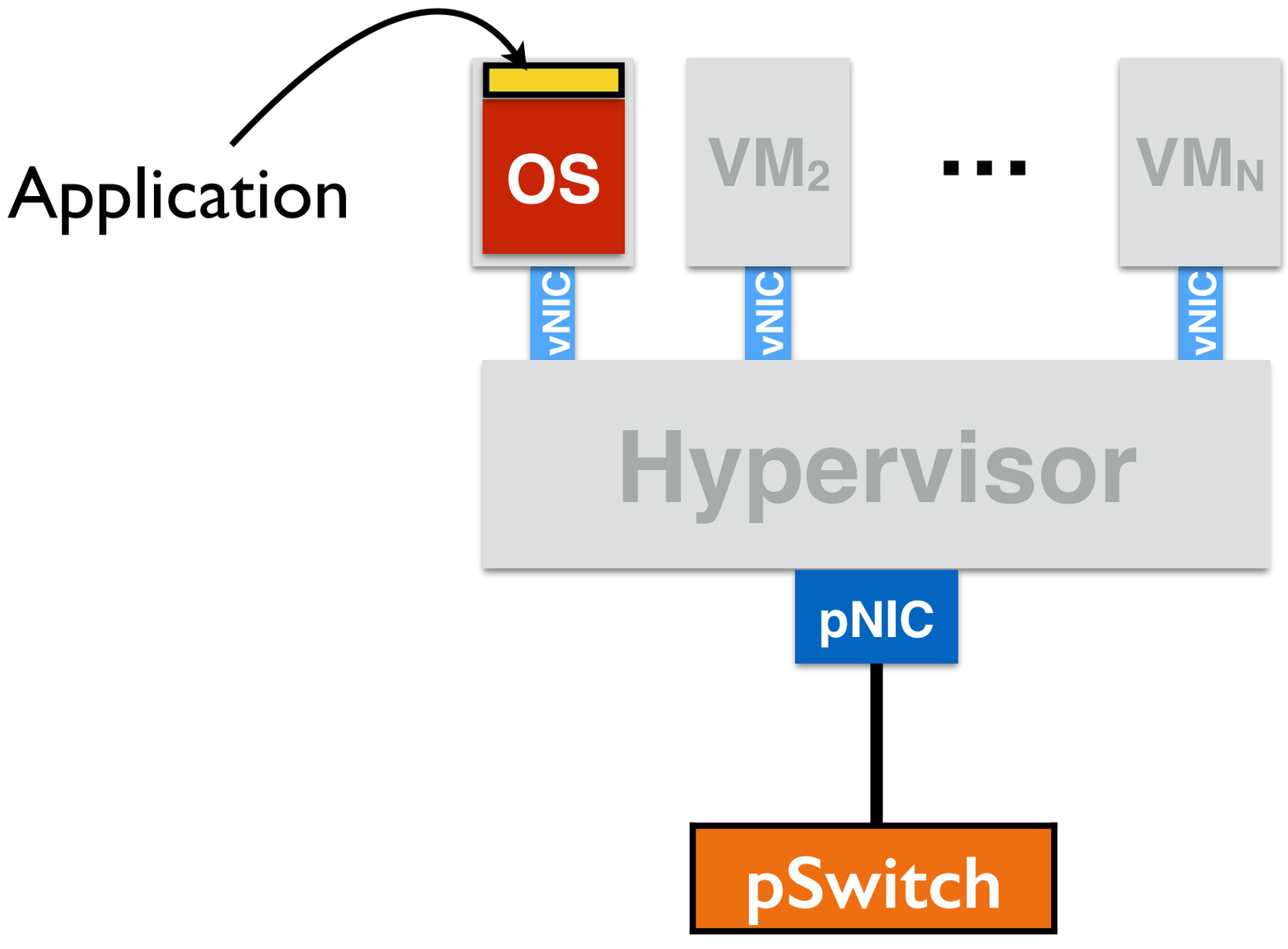
virtual switch



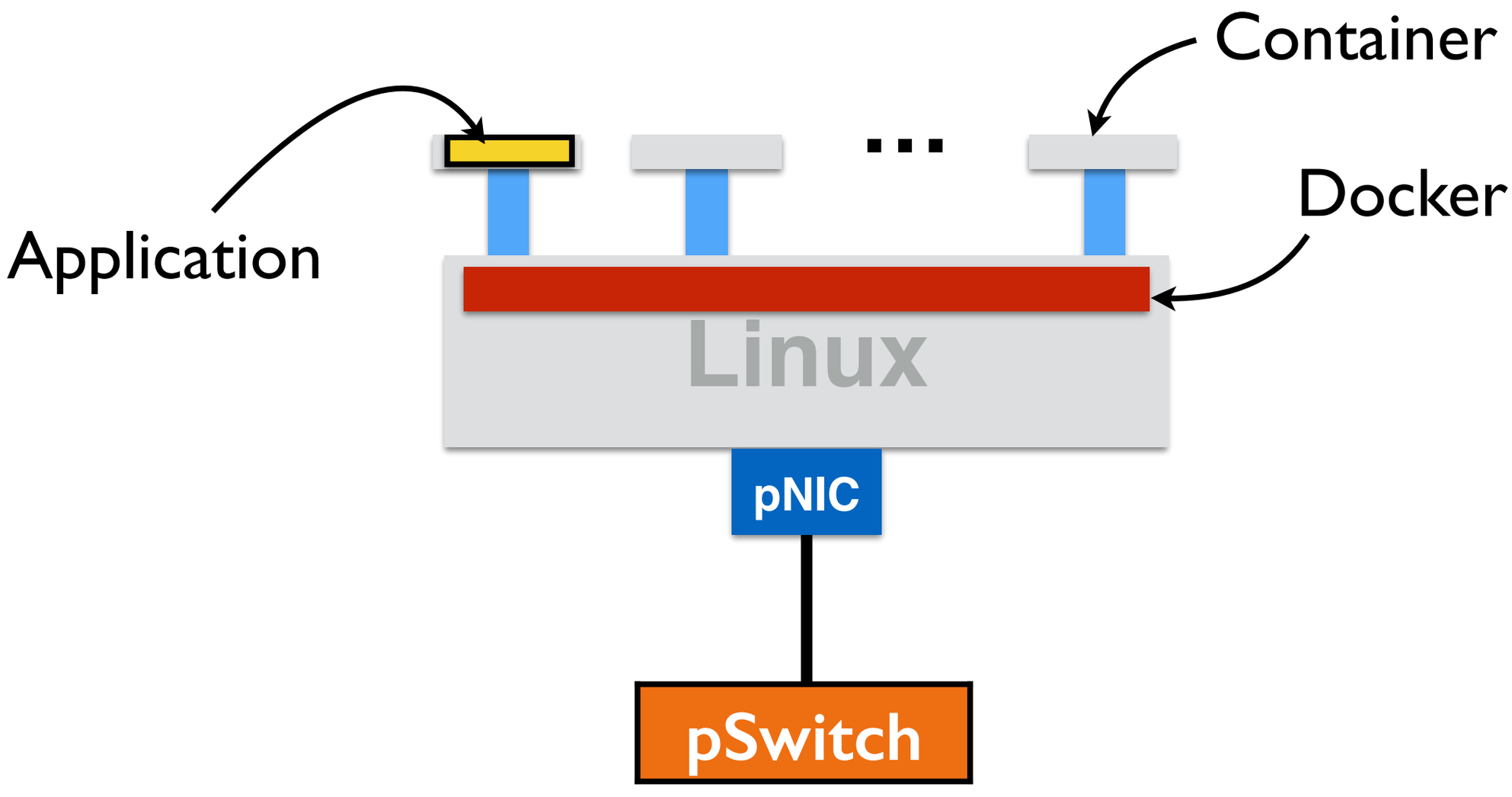
Sidenote: Docker



Sidenote: Docker



Sidenote: Docker

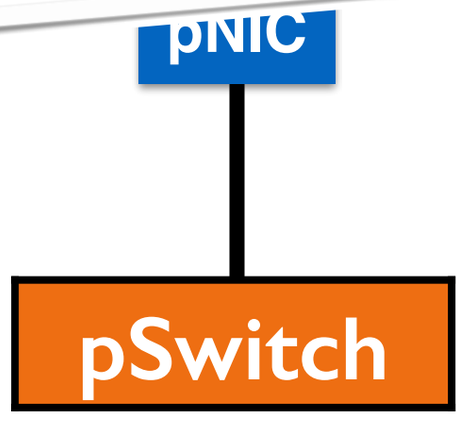


Sidenote: Docker

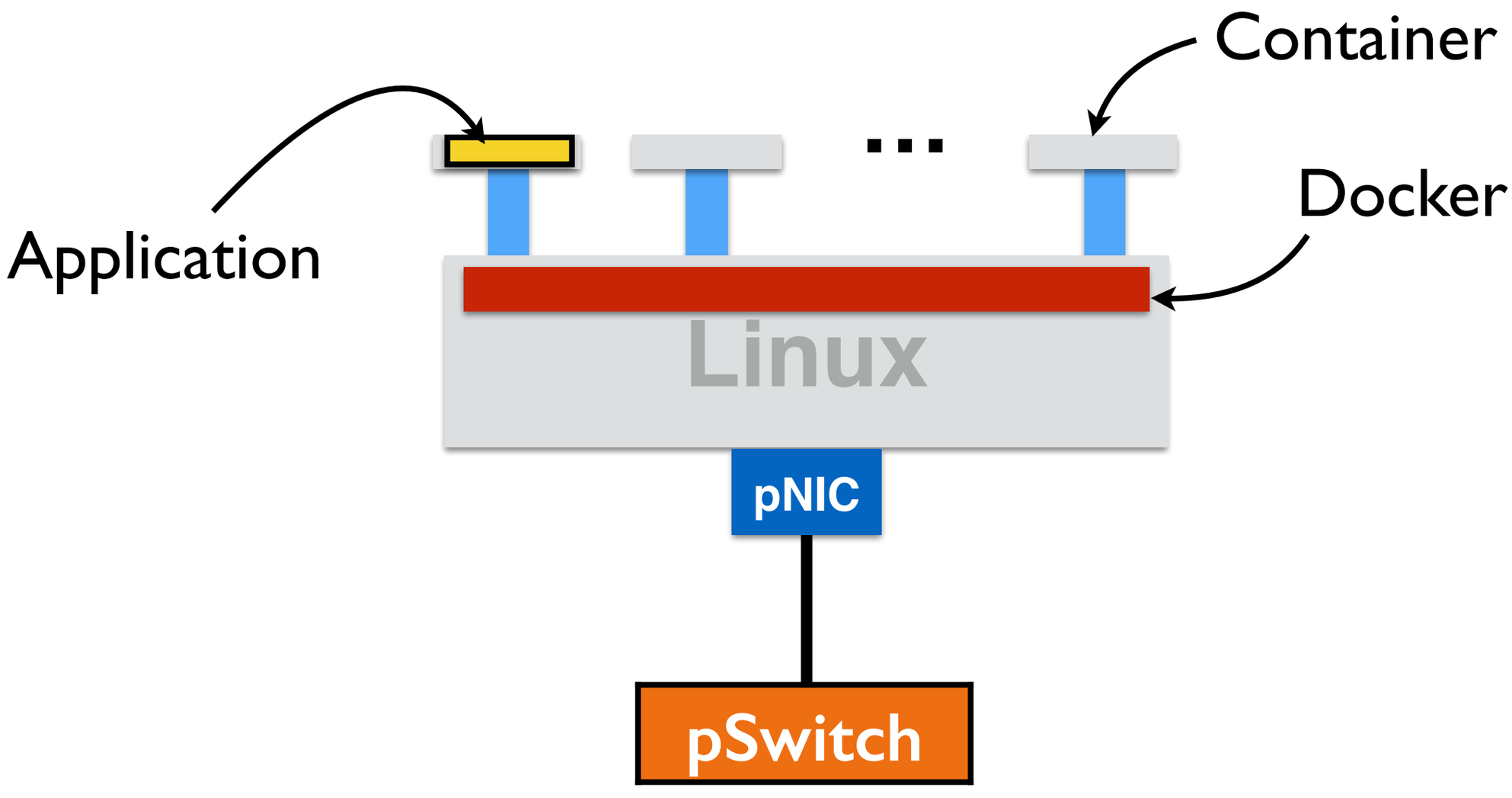
IBM Research Report, 2014

An Updated Performance Comparison of Virtual Machines and Linux Containers

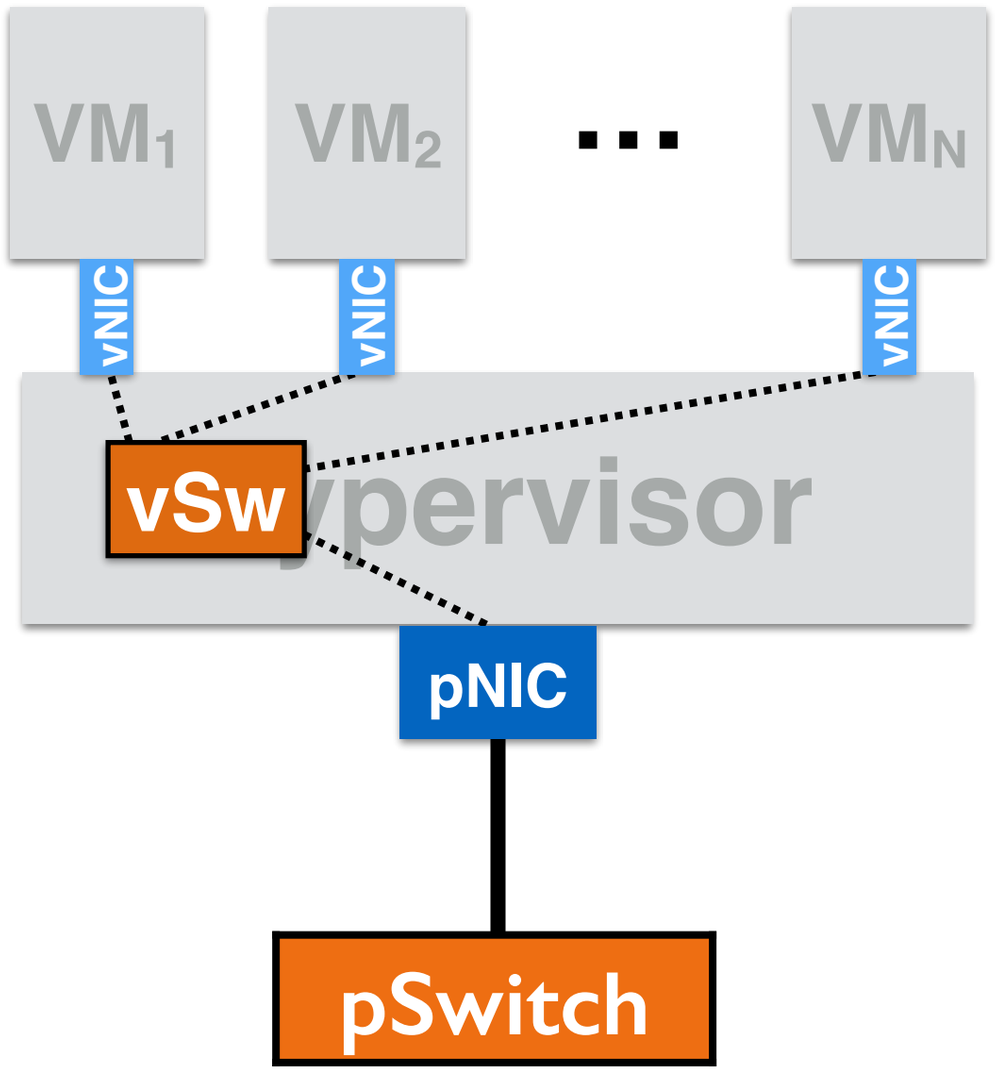
Wes Felter, Alexandre Ferreira, Ram Rajamony, Juan Rubio
IBM Research, Austin, TX



Sidenote: Docker

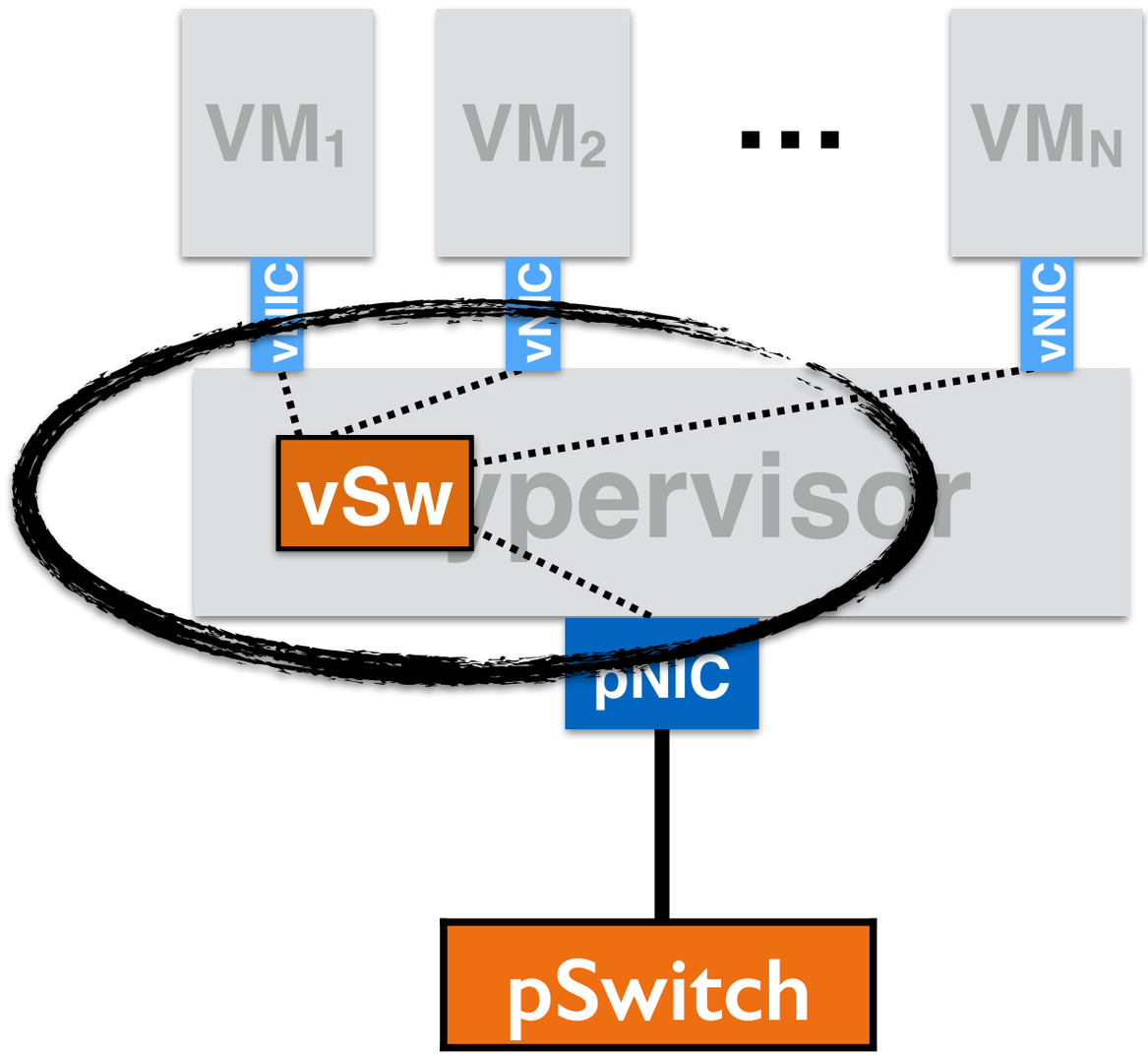


Improving networking performance



Improving networking performance

CPU does the work!



Packet processing on CPUs

Flexible

Packet processing on CPUs

Flexible **slow, CPU-expensive**

Packet processing on CPUs

Flexible **slow, CPU-expensive**

10Gbps, 84 Byte packets ⇒ 67ns time budget

Packet processing on CPUs

Flexible **slow, CPU-expensive**

10Gbps, 84 Byte packets ⇒ 67ns time budget

Context: CPU-memory takes tens of ns

Packet processing on CPUs

Flexible **slow, CPU-expensive**

Packet processing on CPUs

Flexible **slow, CPU-expensive**

- Packet I/O

Packet processing on CPUs

Flexible **slow, CPU-expensive**

- Packet I/O
- Userspace overheads

Packet processing on CPUs

Flexible **slow, CPU-expensive**

- Packet I/O
- Userspace overheads
- Packet classification

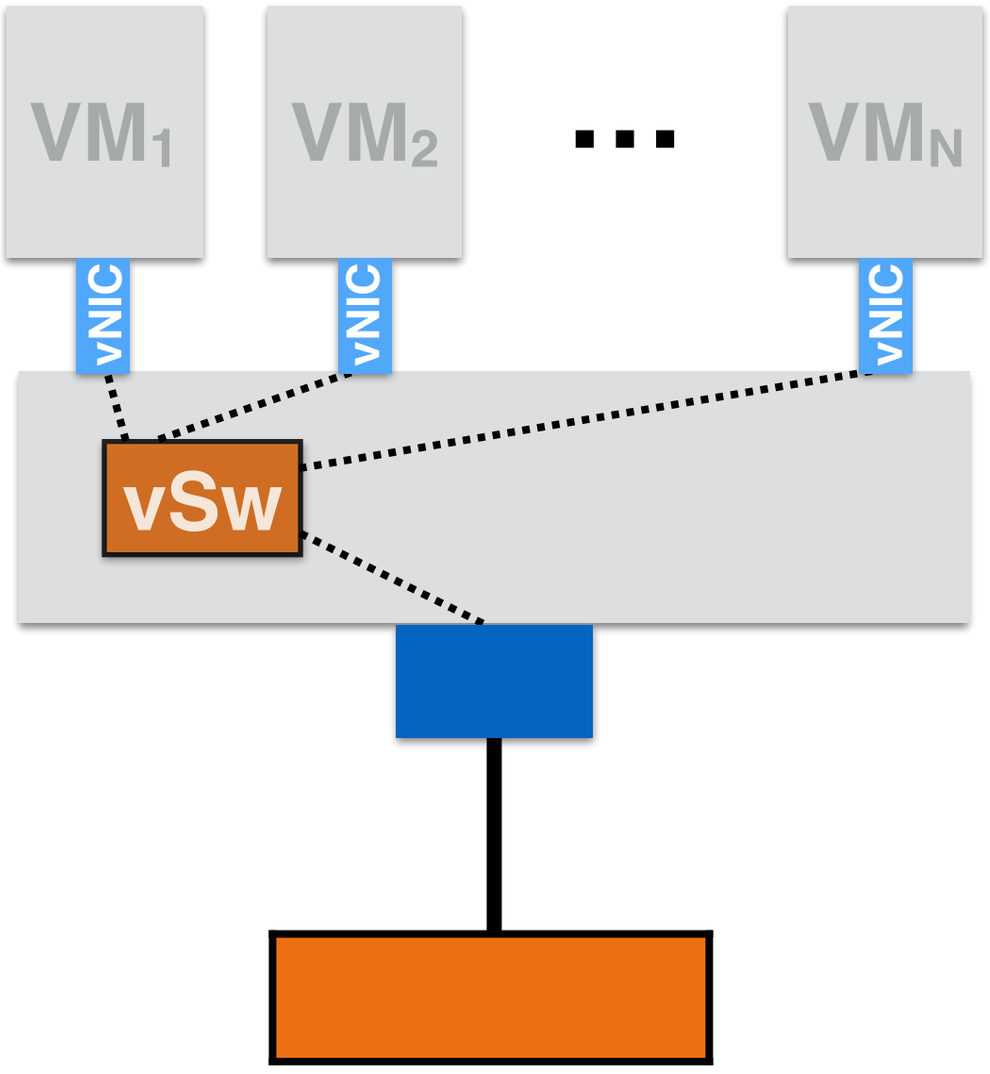
Packet processing on CPUs

Flexible **slow, CPU-expensive**

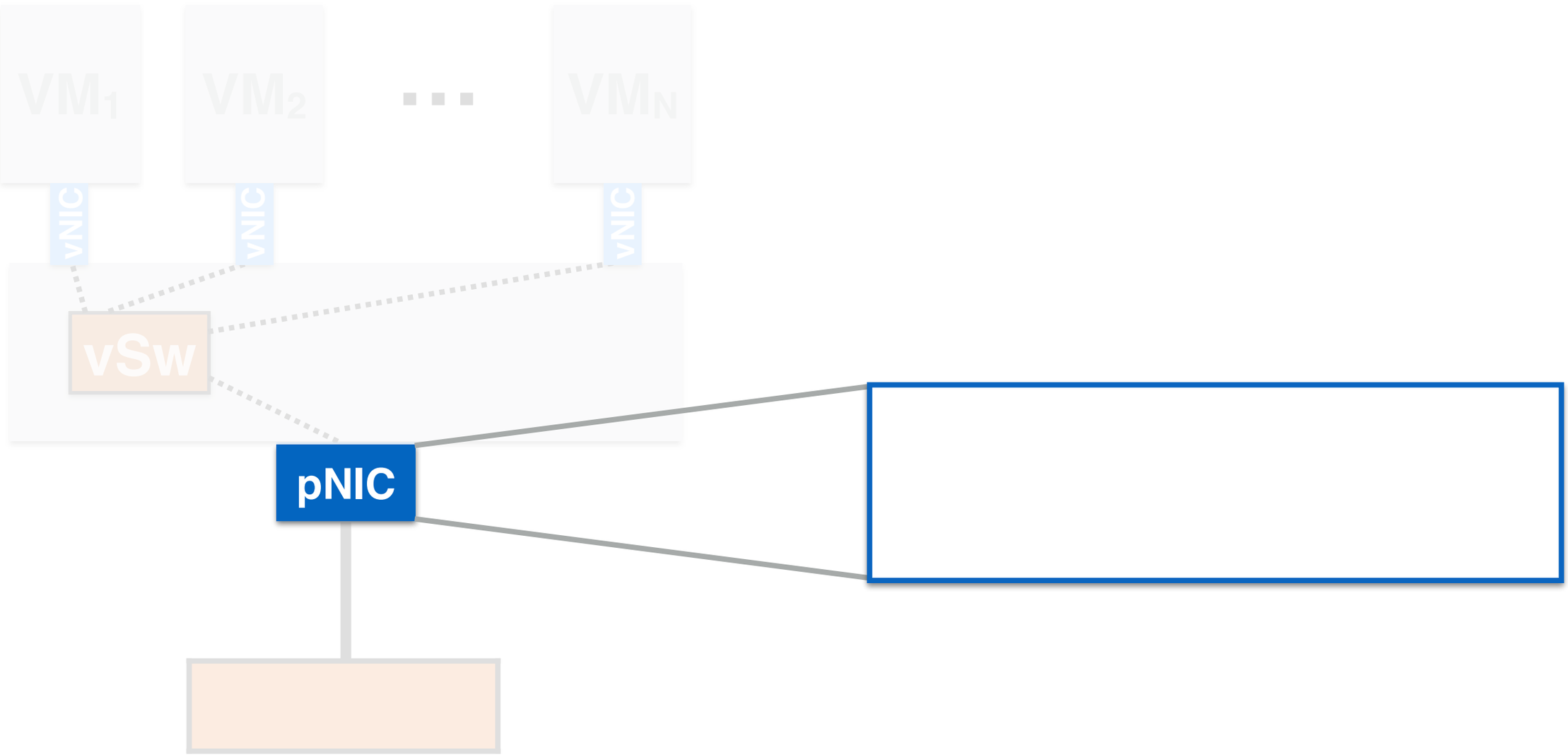
- Packet I/O
- Userspace overheads
- Packet classification

Progress is being made!

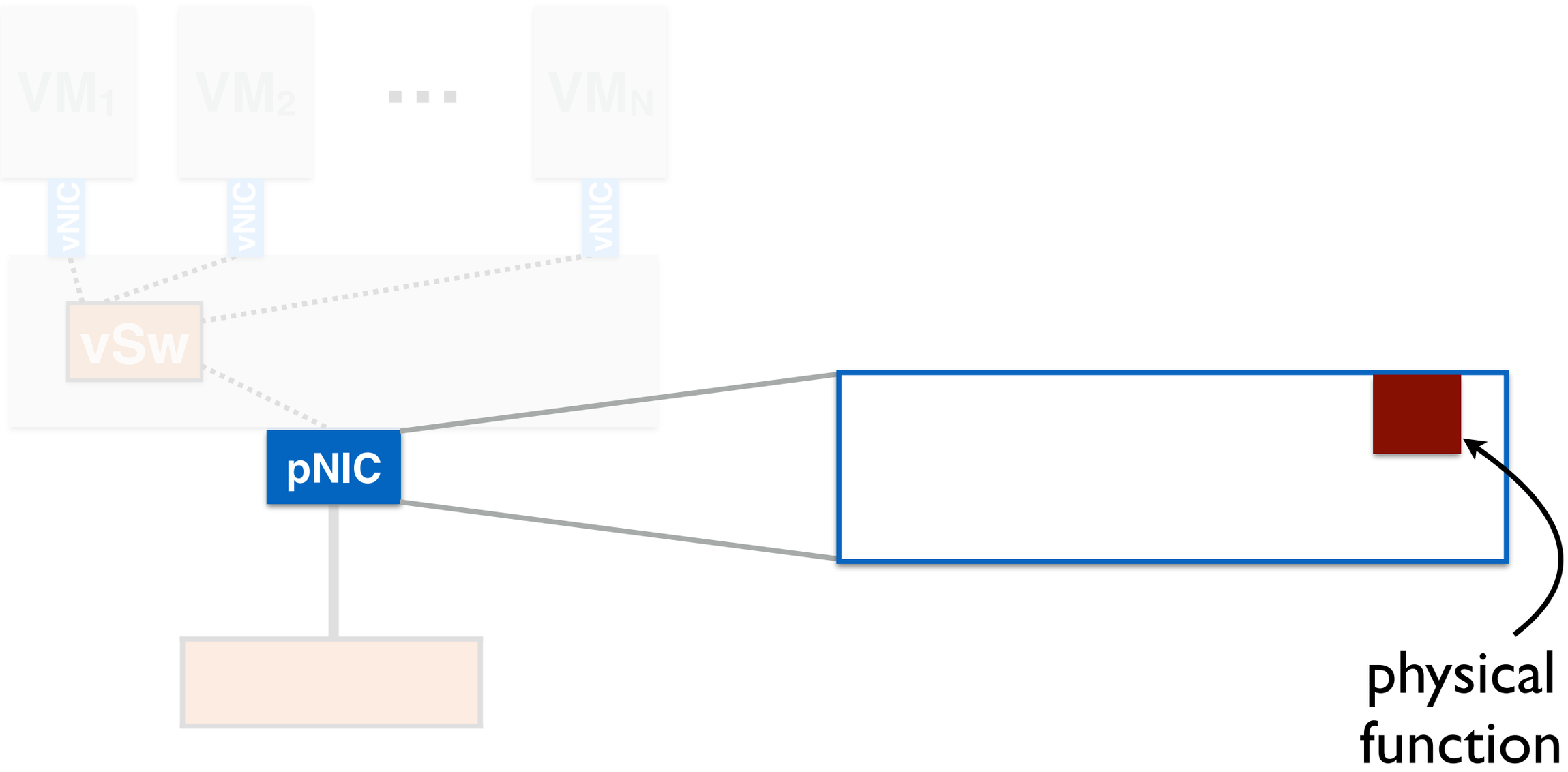
SR-IOV: Single-root I/O Virtualization



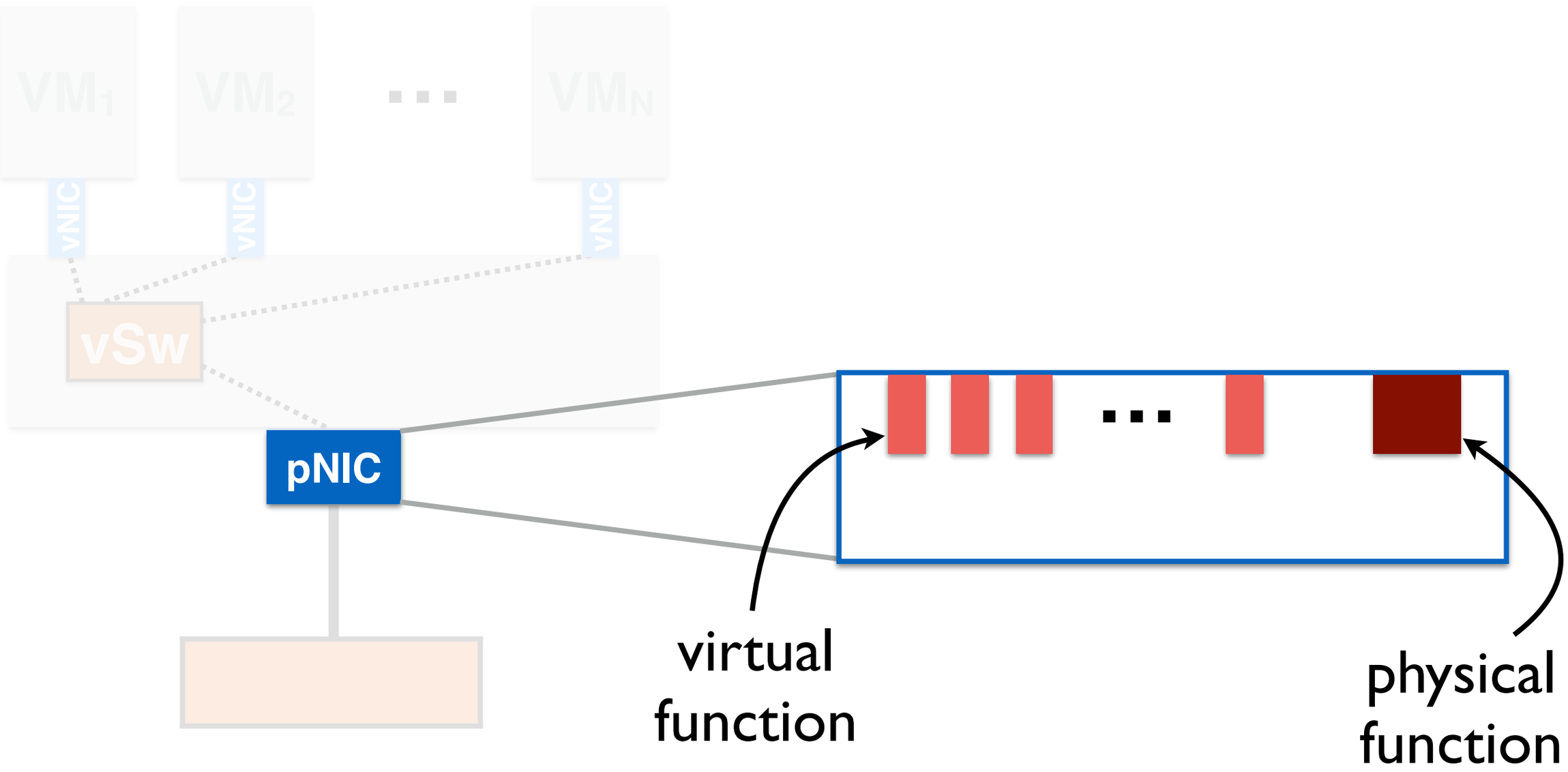
SR-IOV: Single-root I/O Virtualization



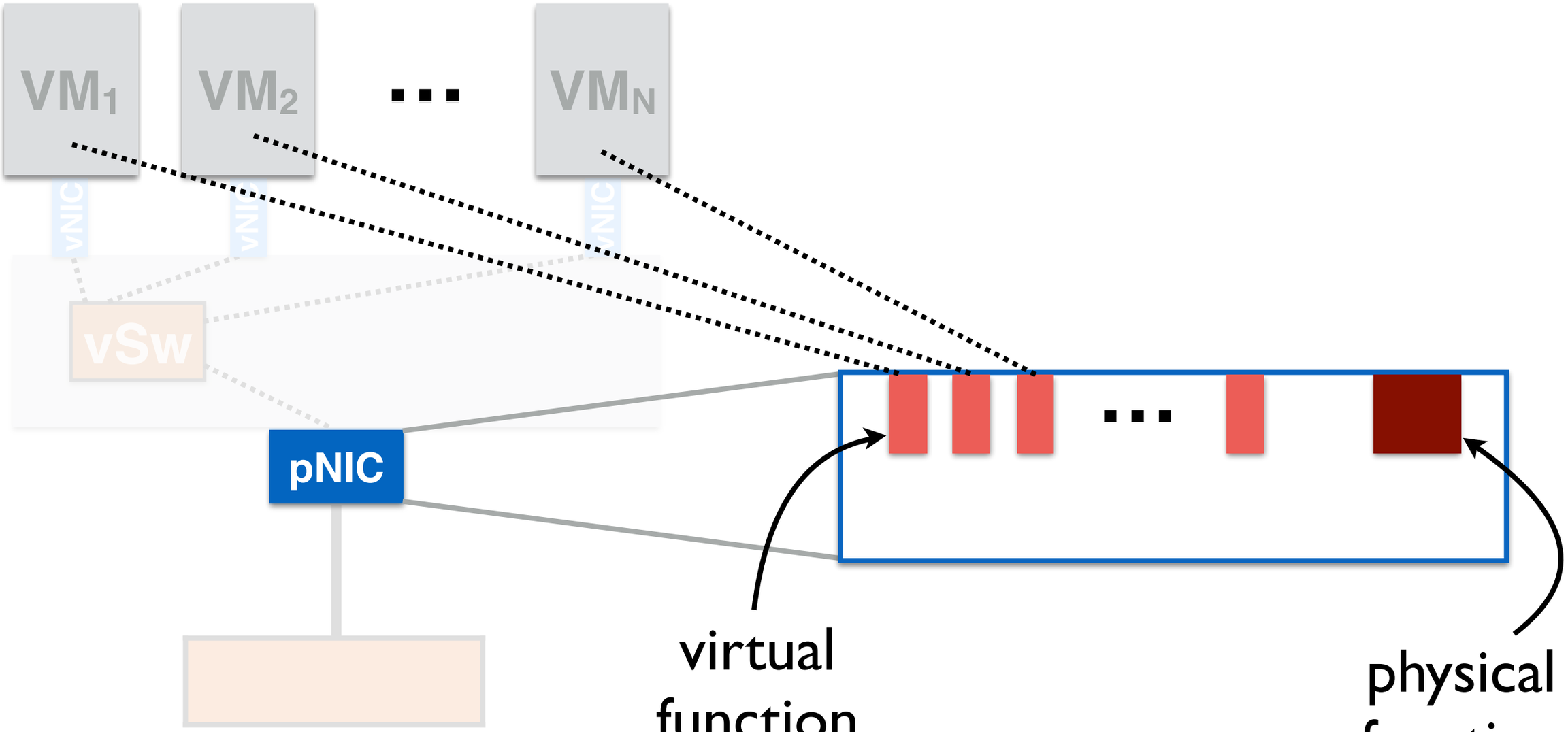
SR-IOV: Single-root I/O Virtualization



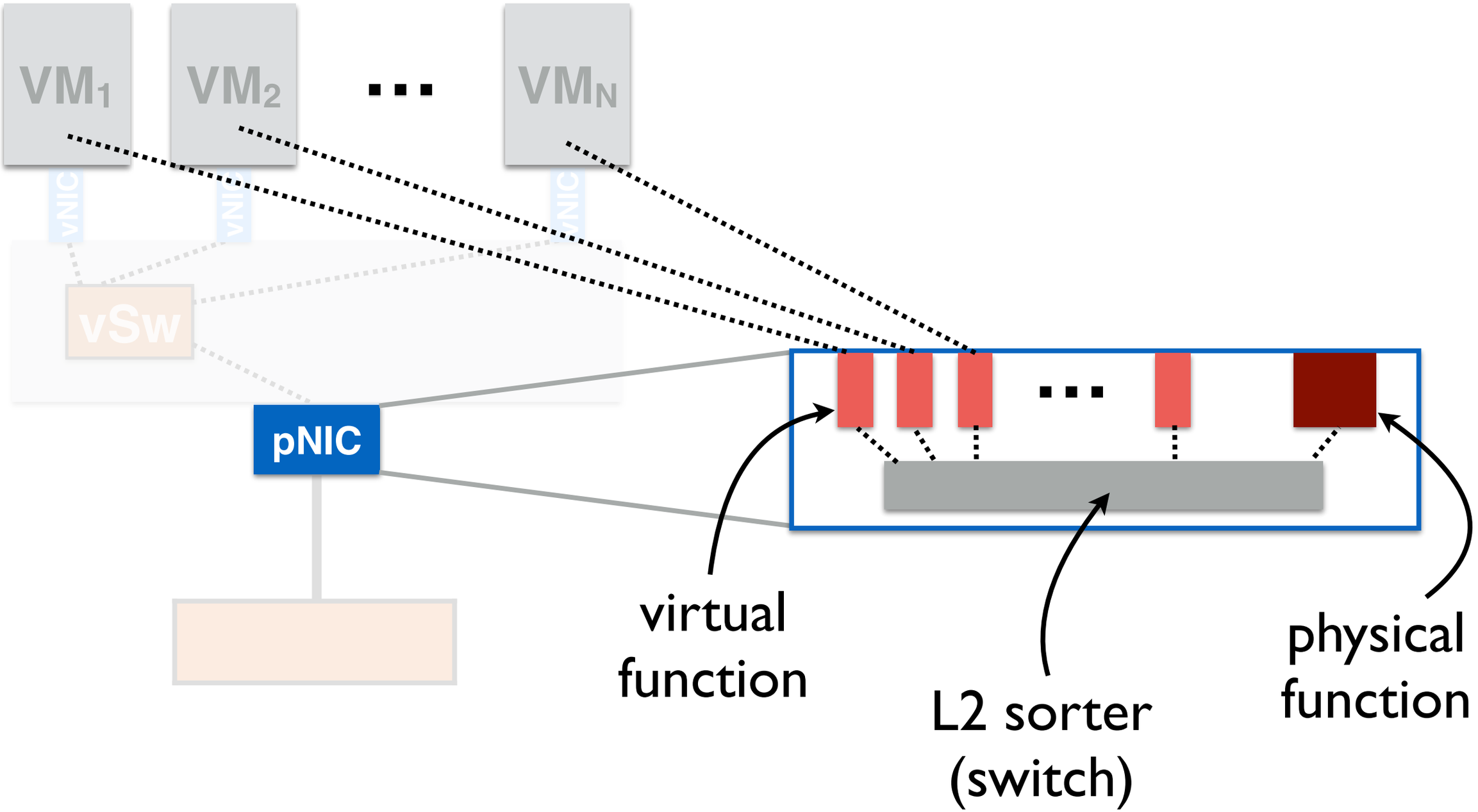
SR-IOV: Single-root I/O Virtualization



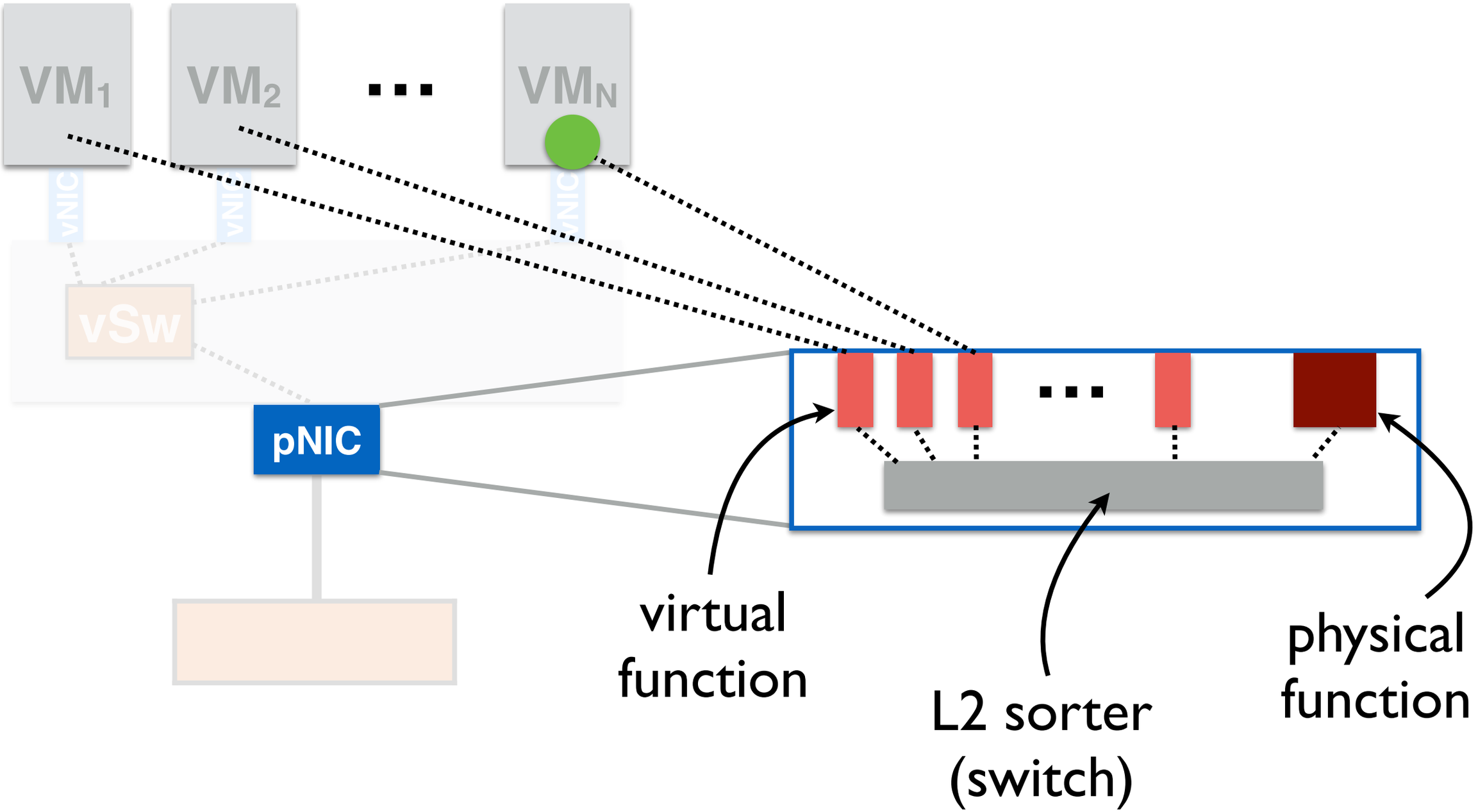
SR-IOV: Single-root I/O Virtualization



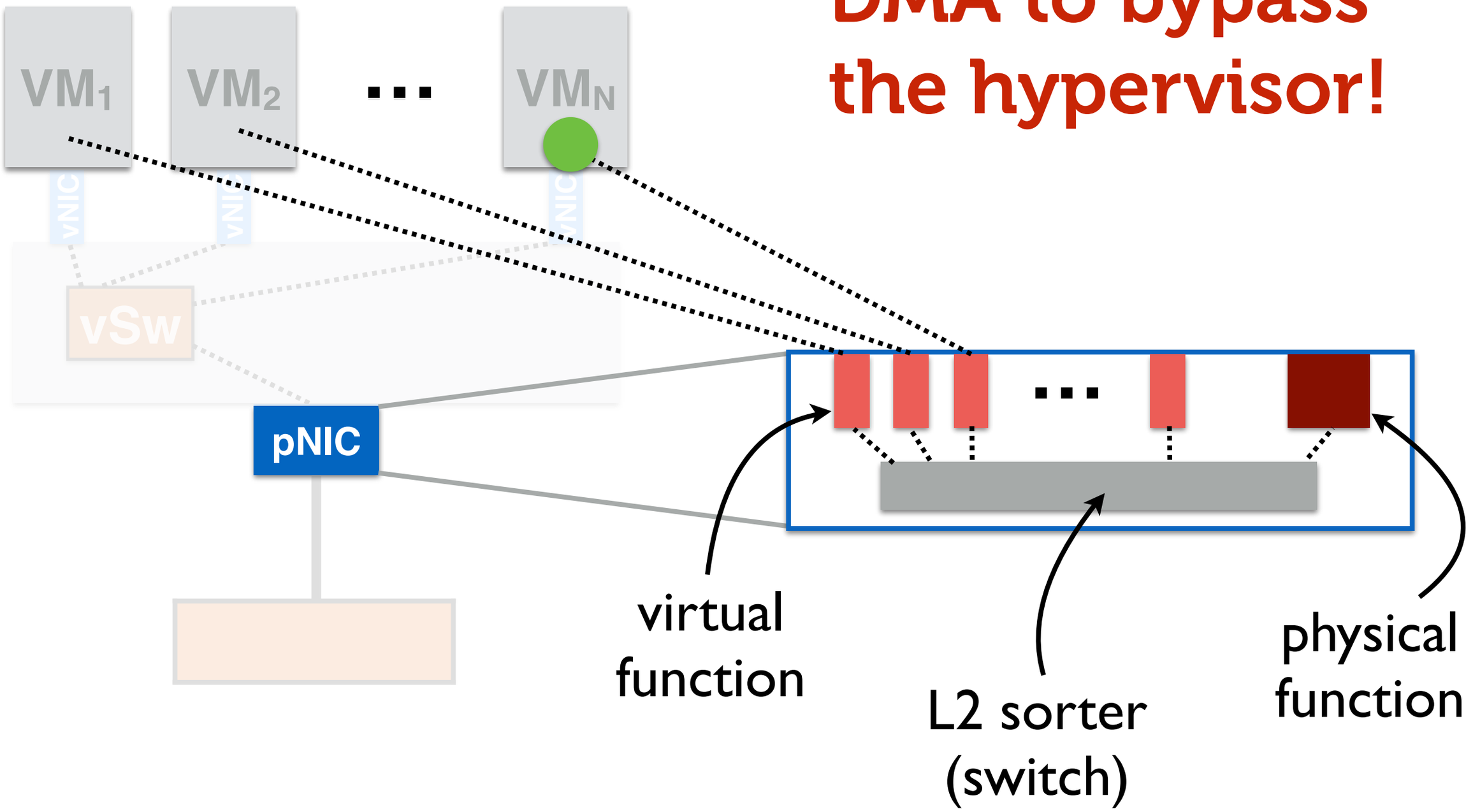
SR-IOV: Single-root I/O Virtualization



SR-IOV: Single-root I/O Virtualization



SR-IOV: Single-root I/O Virtualization



DMA to bypass the hypervisor!

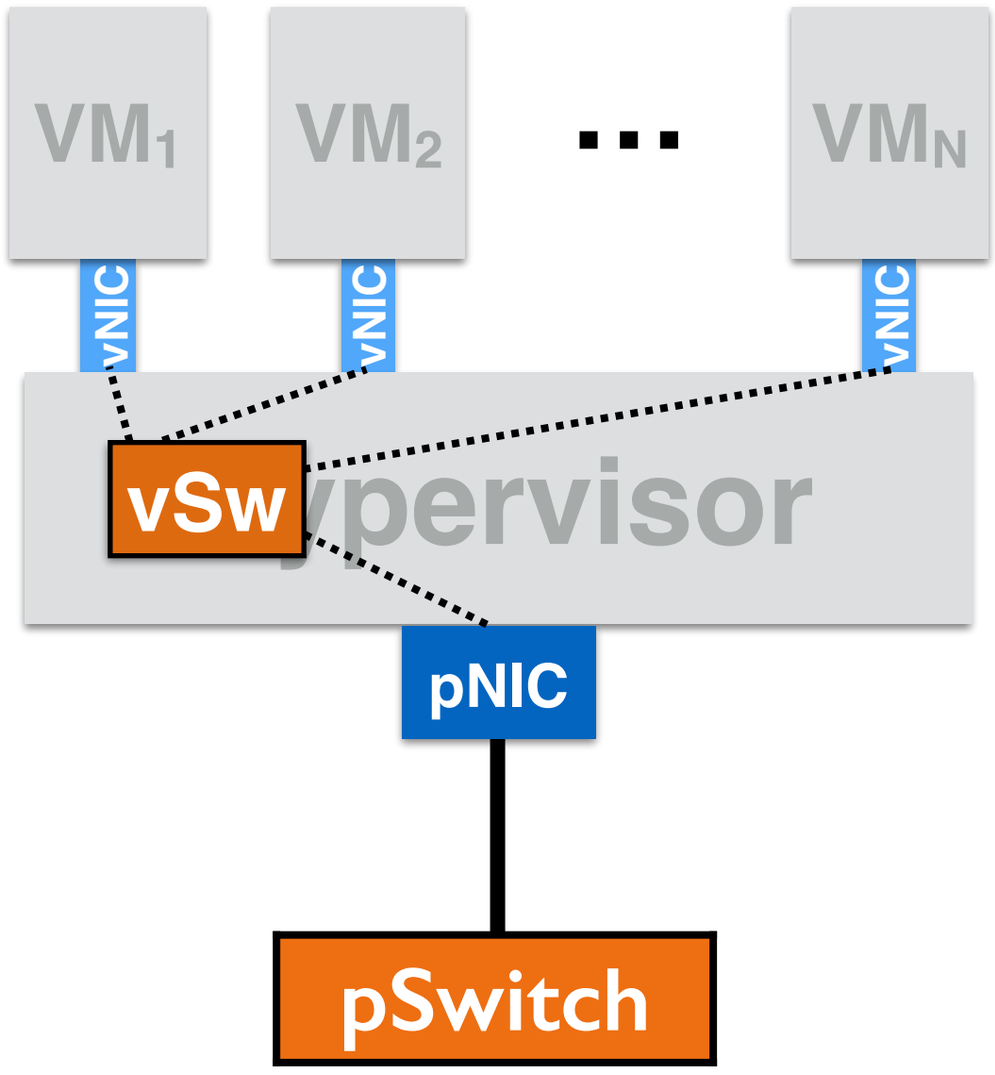
Open vSwitch

USENIX NSDI, 2015

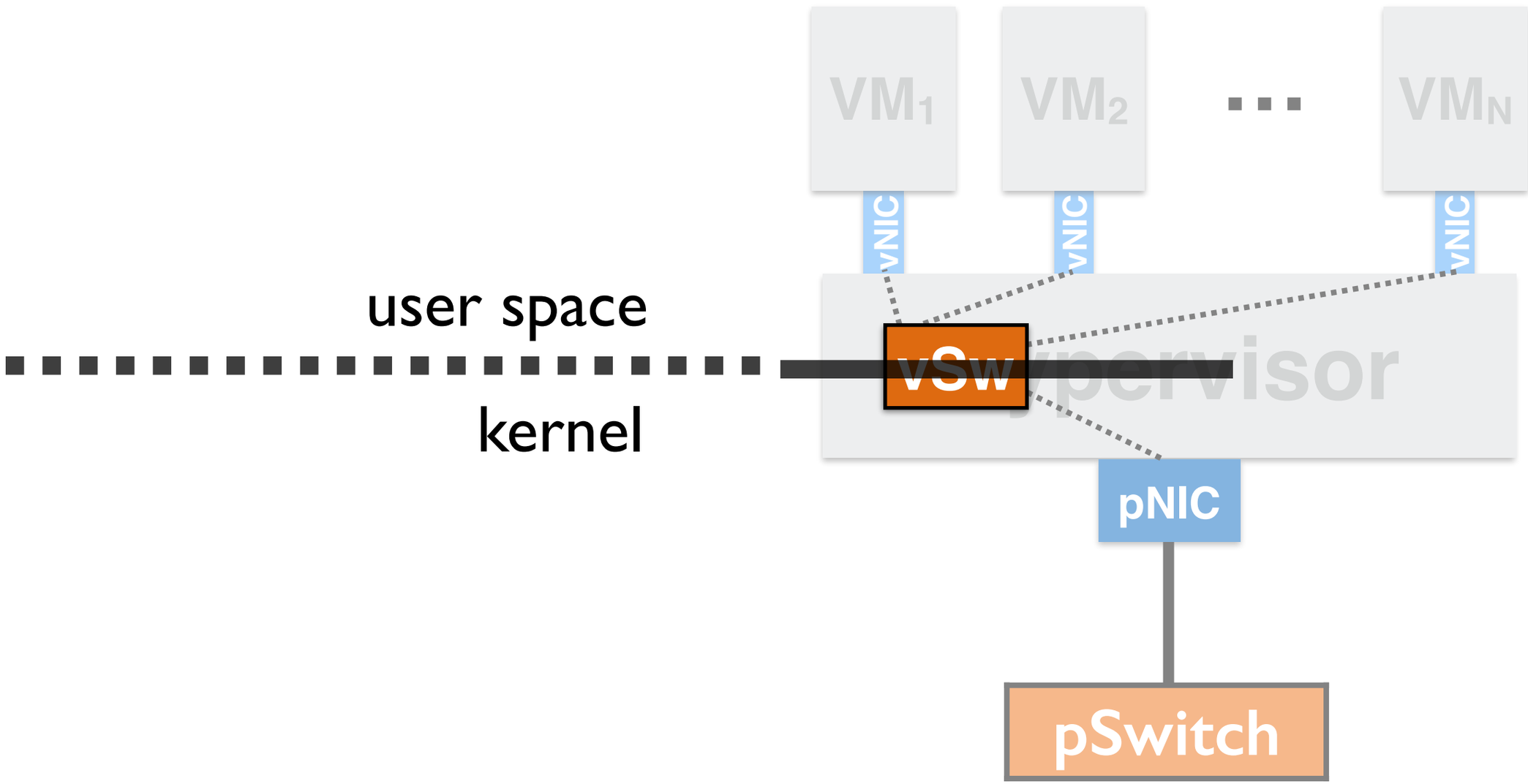
The Design and Implementation of Open vSwitch

*Ben Pfaff**, *Justin Pettit**, *Teemu Koponen**, *Ethan J. Jackson**,
*Andy Zhou**, *Jarno Rajahalme**, *Jesse Gross**, *Alex Wang**,
*Jonathan Stringer**, *Pravin Shelar**, *Keith Amidon*†, *Martín Casado**
*VMware †Awake Networks

Open vSwitch



Open vSwitch

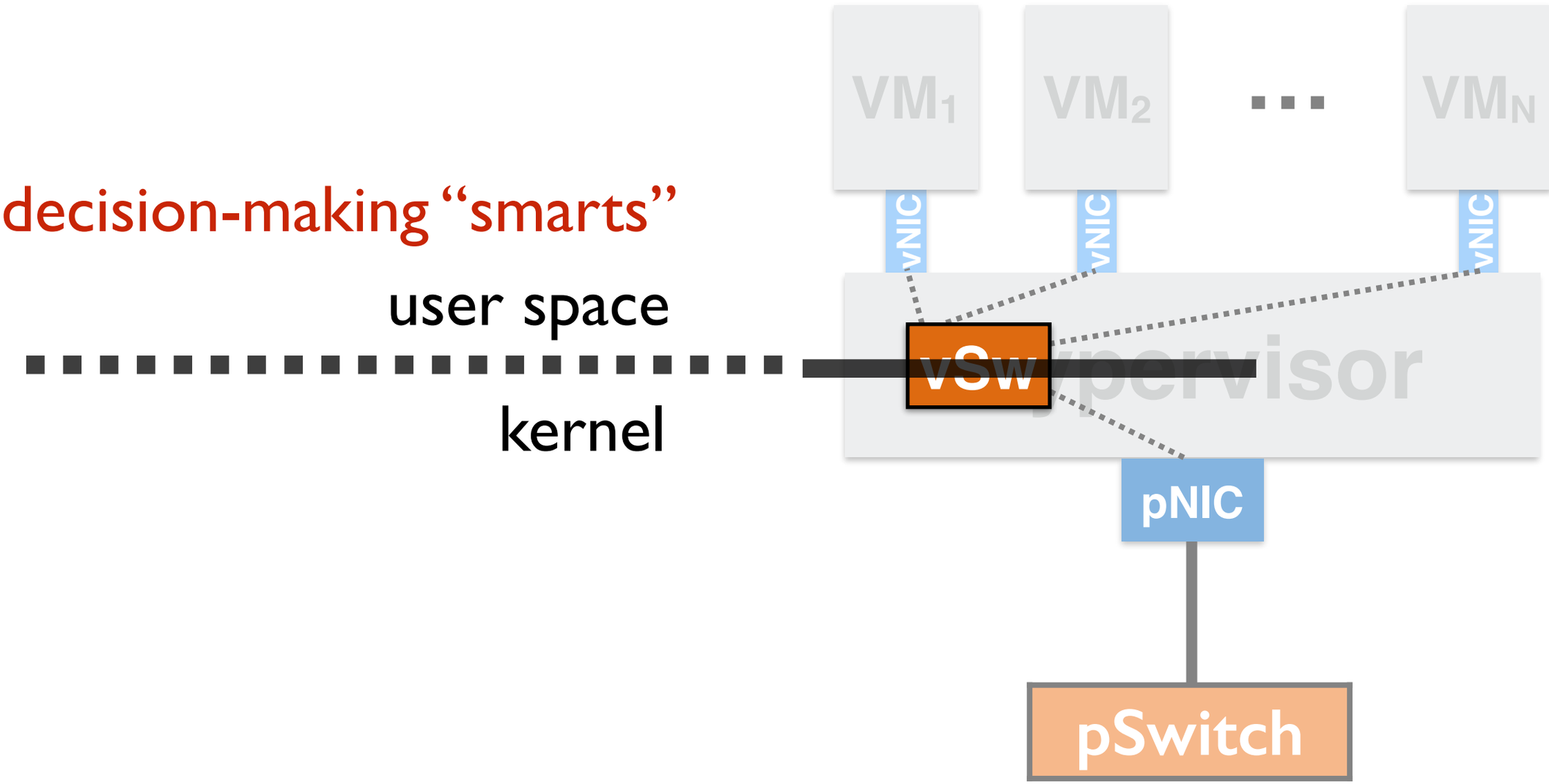


Open vSwitch

decision-making "smarts"

user space

kernel



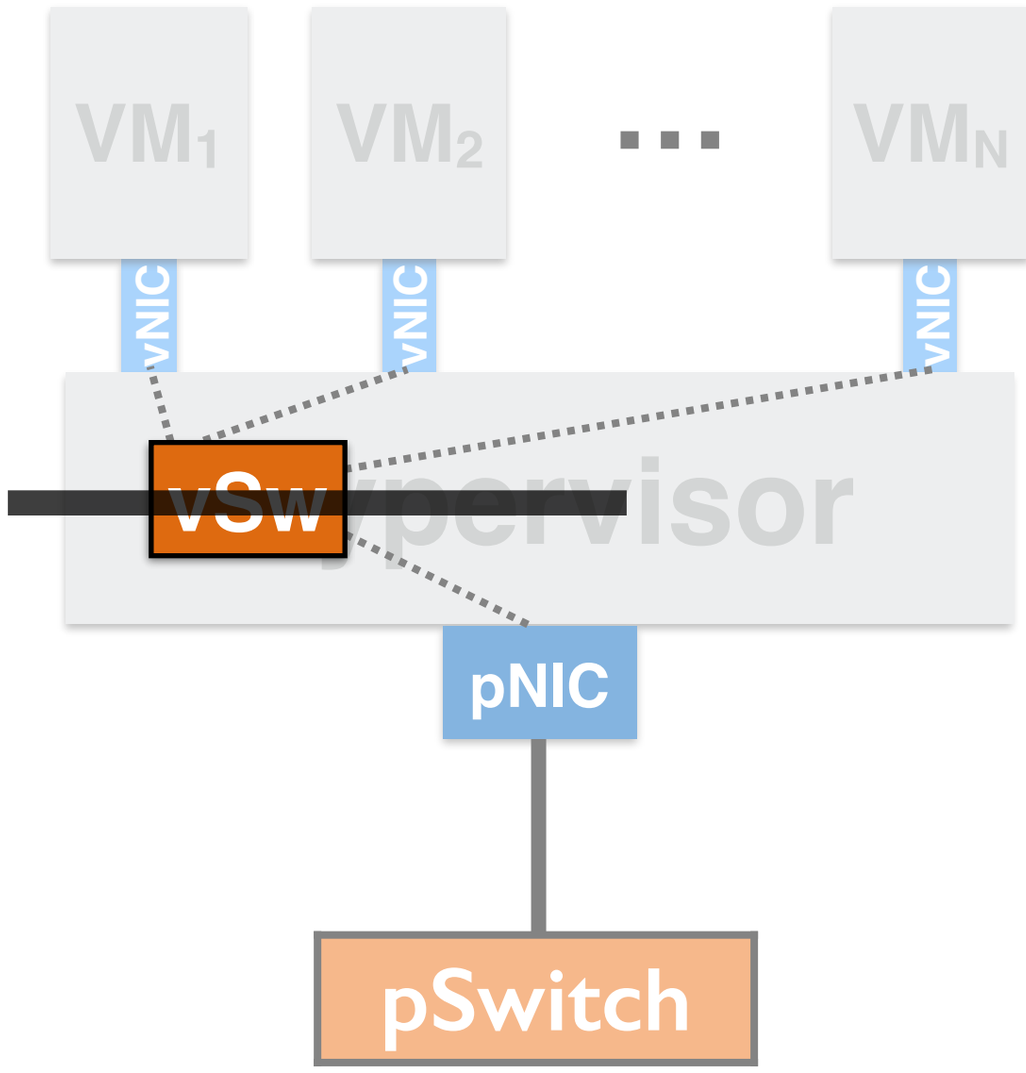
Open vSwitch

decision-making "smarts"

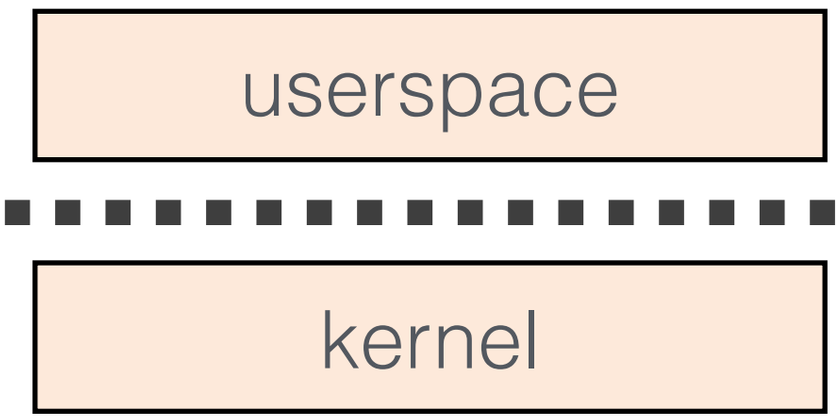
user space

kernel

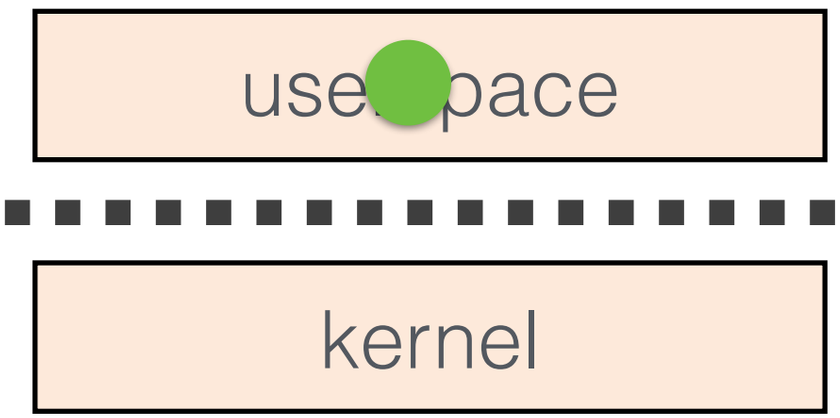
simple, fast forwarding



Inside Open vSwitch



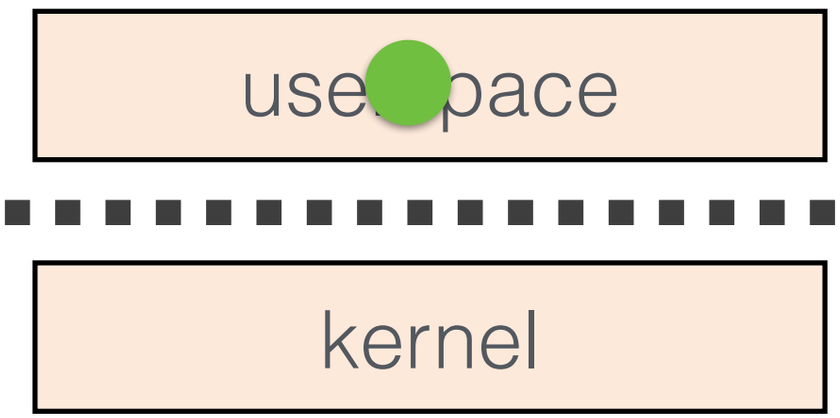
Inside Open vSwitch



Inside Open vSwitch

srcMAC	dstMAC	action

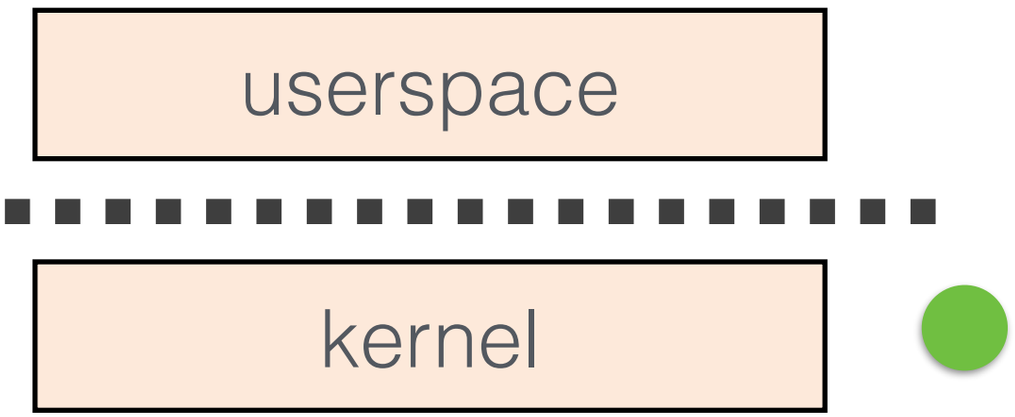
dstIP	dstTCPPort	action



Inside Open vSwitch

srcMAC	dstMAC	action

dstIP	dstTCPPort	action



Inside Open vSwitch

srcMAC	dstMAC	action

dstIP	dstTCPPort	action

userspace

kernel

srcMAC	dstMAC	dstIP	dstTCPPort	action

Inside Open vSwitch

srcMAC	dstMAC	action

dstIP	dstTCPPort	action

userspace

kernel

srcMAC	dstMAC	dstIP	dstTCPPort	action

Inside Open vSwitch

srcMAC	dstMAC	action

dstIP	dstTCPPort	action

userspace

kernel

srcMAC	dstMAC	dstIP	dstTCPPort	action

Inside Open vSwitch

srcMAC	dstMAC	action

dstIP	dstTCPPort	action

userspace

kernel

cache

hash-key	flow-table-entry

srcMAC	dstMAC	dstIP	dstTCPPort	action

Inside Open vSwitch

srcMAC	dstMAC	action

dstIP	dstTCPPort	action

userspace

kernel

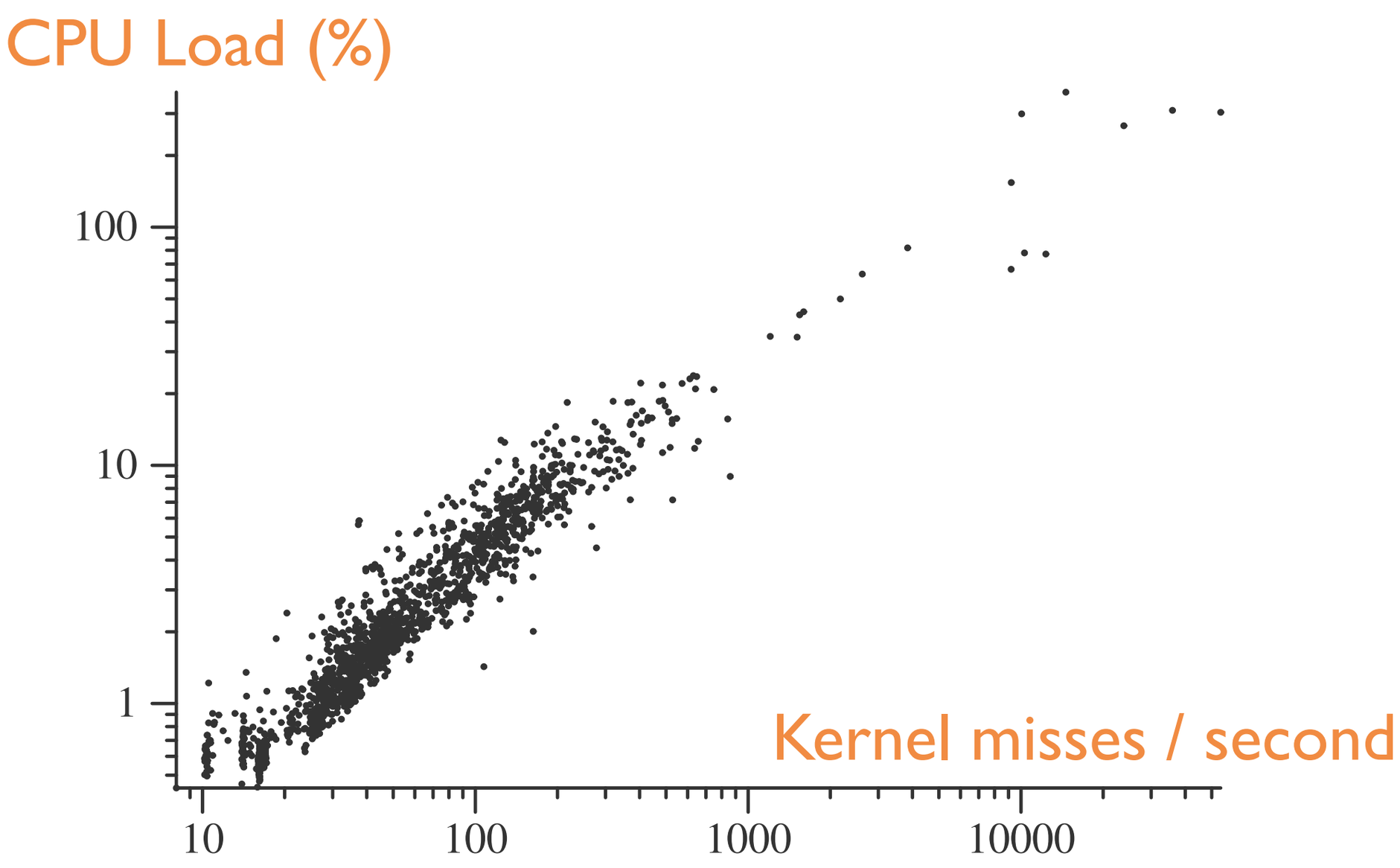
cache

hash-key	flow-table-entry

srcMAC	dstMAC	dstIP	dstTCPPort	action

hash-key = **hash** (srcMAC, dstMAC, dstIP, dstTCPPort)

Inside Open vSwitch



[The Design and Implementation of Open vSwitch. Pfaff et. al, USENIX NSDI 2015]