



NOVA: Virtualization With A Small Trusted Computing Base

Udo Steinberg, TU Dresden

Motivation

- x86 virtualization is becoming popular
 - Consolidating multiple OSs on a single machine can save resources: maintenance, power, cooling, floor space
- With virtualization, security of hosted OSs additionally depends on the security of the virtualization layer
 - Successful attack on the virtualization software compromises all guest operating systems at once

Problem Statement

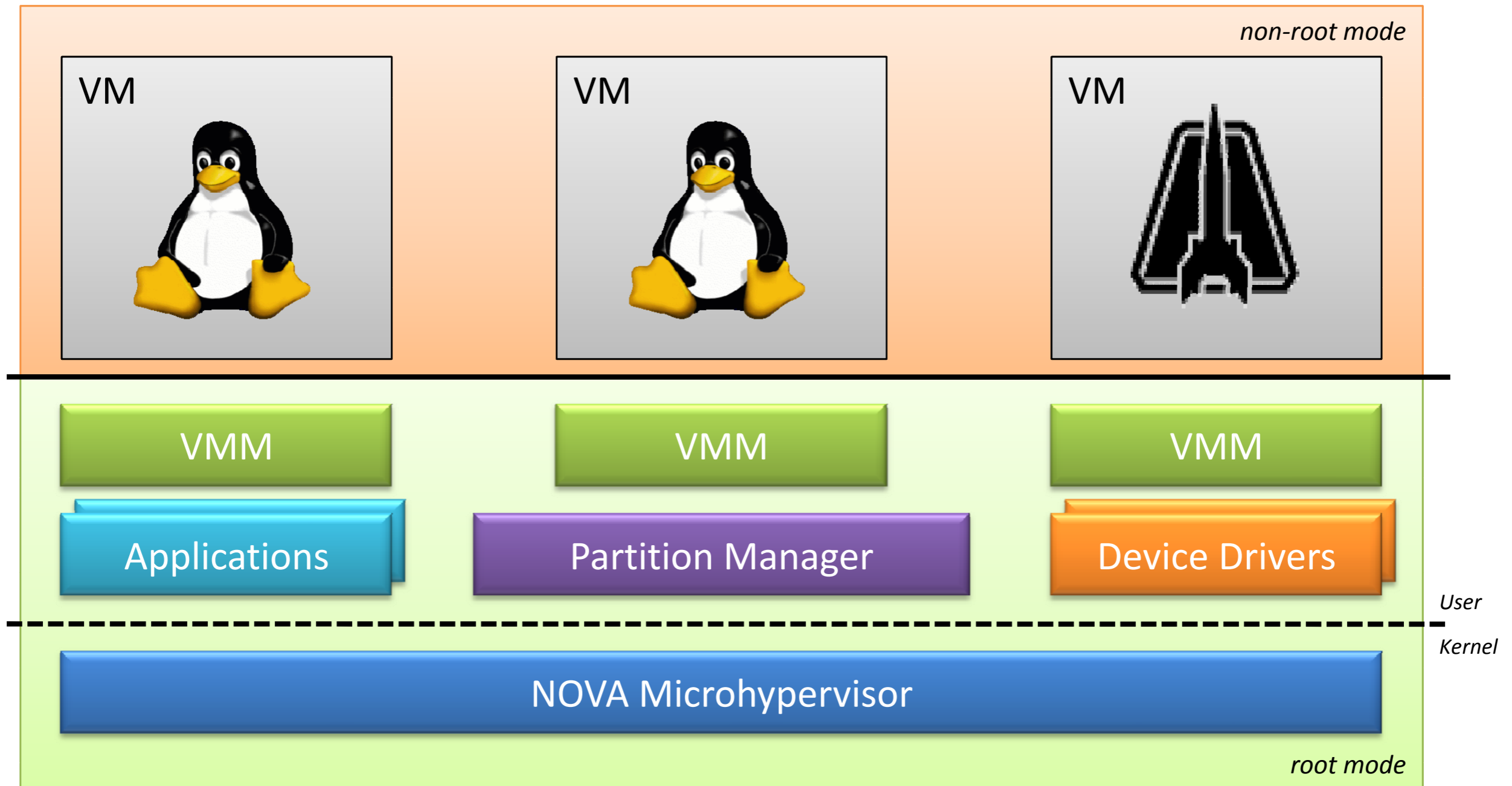
- Current virtualization software is too big
 - Number of defects increases with code size and complexity
- Small hypervisors did not really happen
 - Reuse of legacy OS for device drivers
 - Linux kernel in Dom0 (Xen)
 - Linux kernel as hypervisor (KVM)
 - Virtualization just added on top of old interfaces
 - Driven by time-to-market constraints

NOVA Design Principles

1. Fine-grain functional decomposition of the virtualization layer
2. Enforcement of the principle of least privilege among all these components

Systematic application of both principles can shrink the size of the trusted computing base by at least an order of magnitude

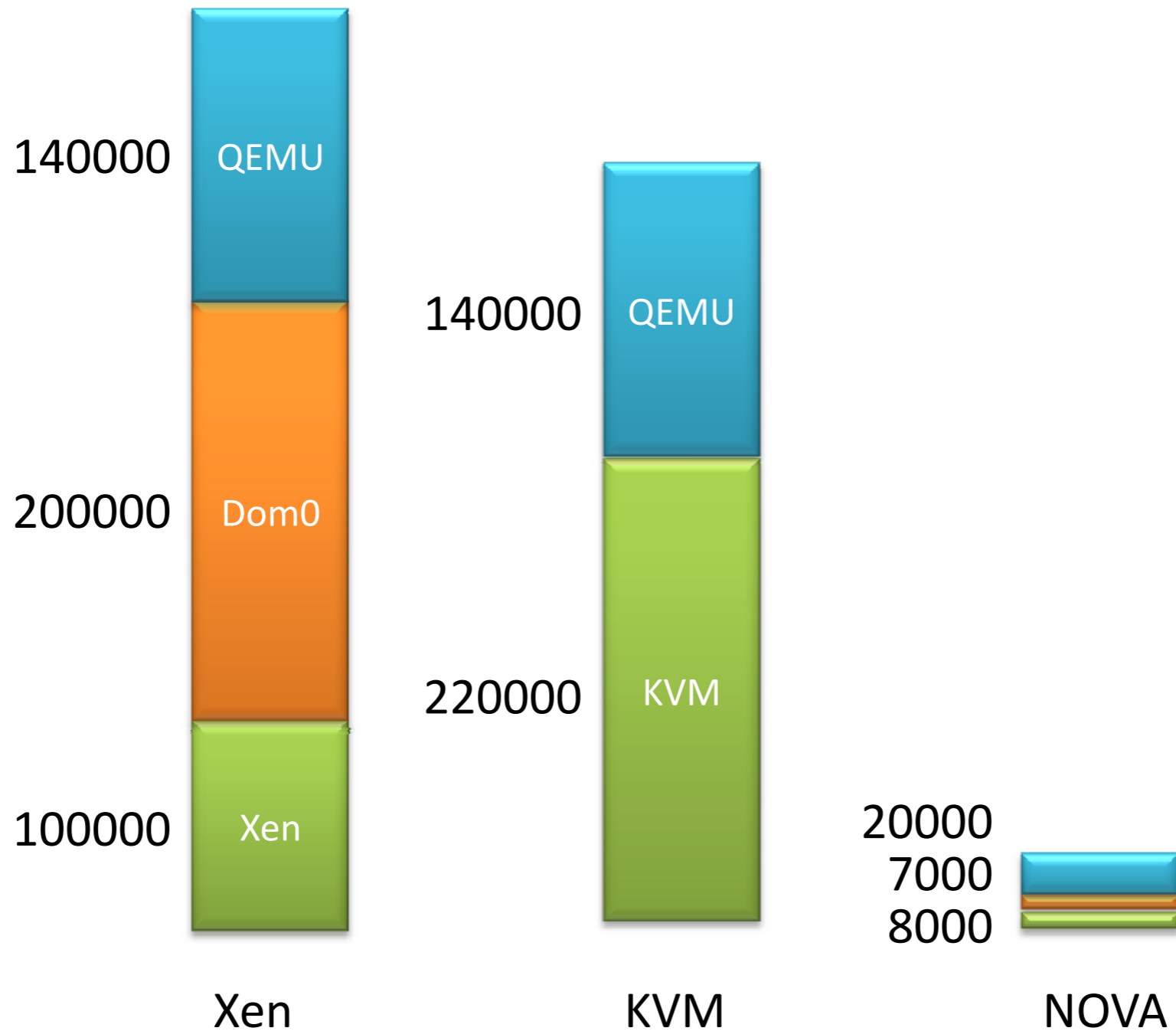
Architecture Overview



Functional Decomposition

- Small microhypervisor
 - ca. 8000 lines of code
- One user-level VMM per virtual machine
 - ca. 20000 lines of code
 - VM escape does not affect the hypervisor or other virtual machines
- User-level device drivers
- Security-sensitive apps can run next to VMs

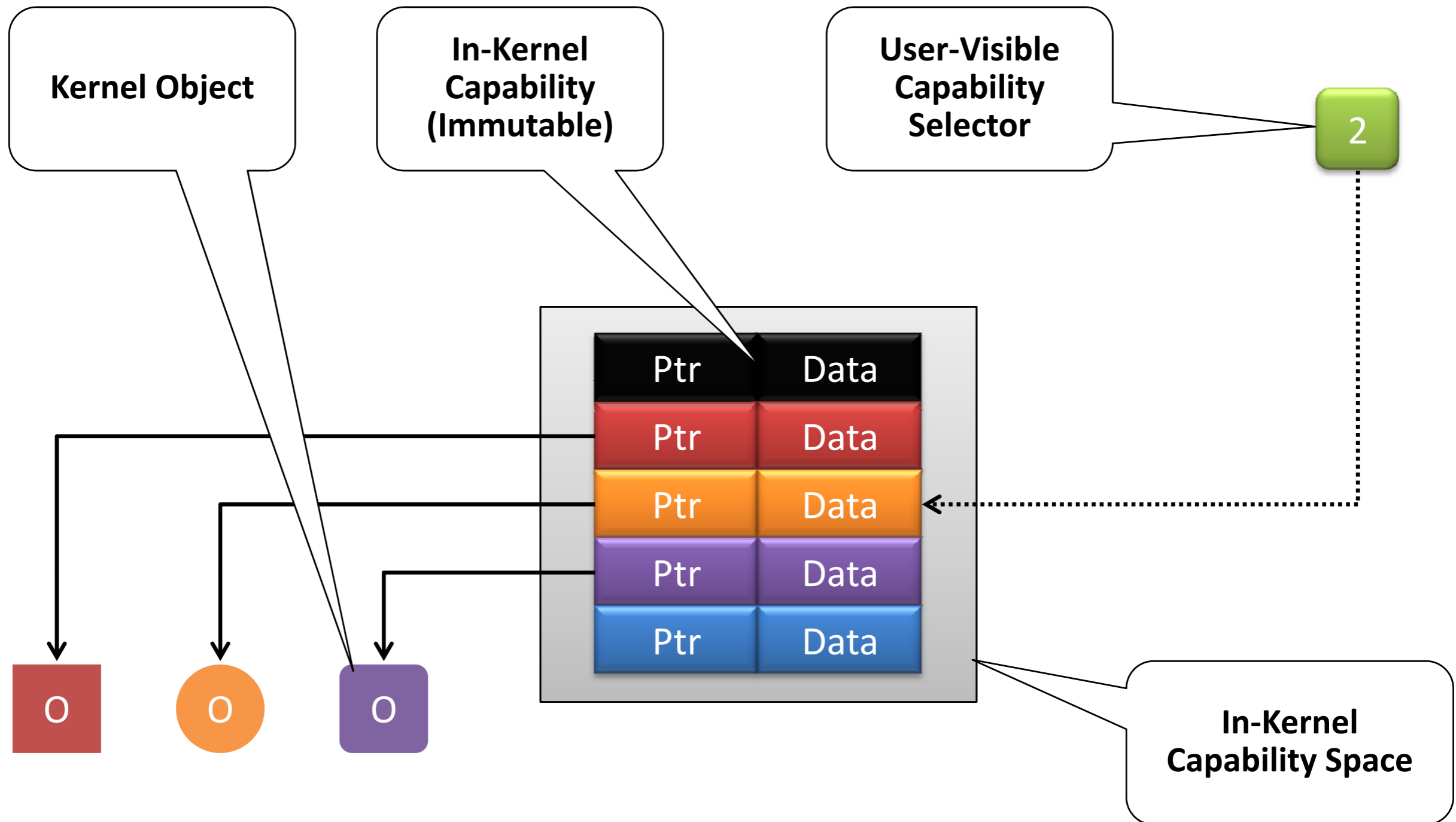
Trusted Computing Base



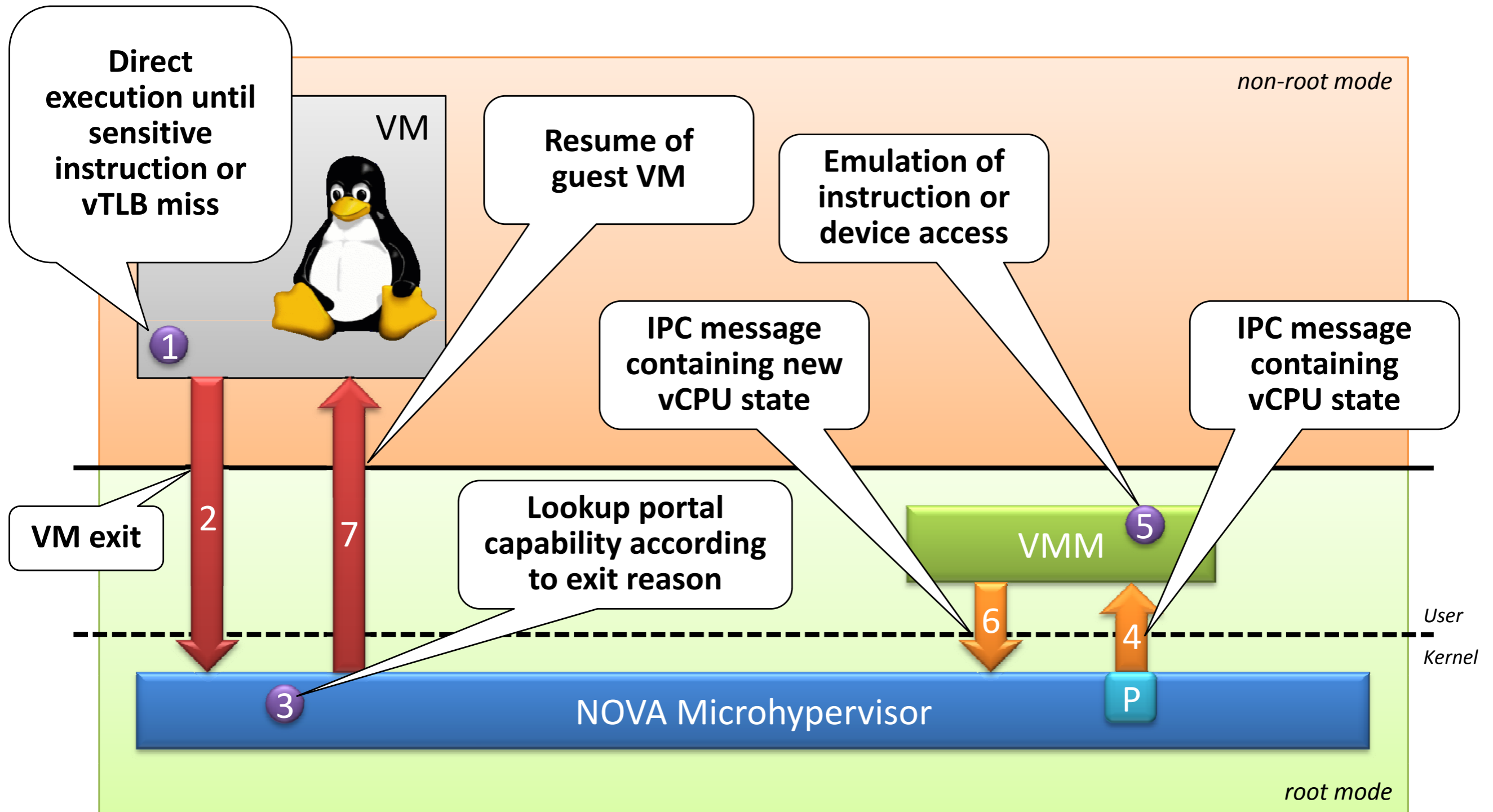
Principle of Least Privilege

- Capability-based hypercall interface
- Creation of a new kernel object yields root cap
 - Capabilities stored in private capability space of the respective protection domain
 - Creator can delegate cap according to its policy
- VMM can only access the memory of its VM
- Device drivers can only DMA into memory regions that have been delegated to them

Capabilities

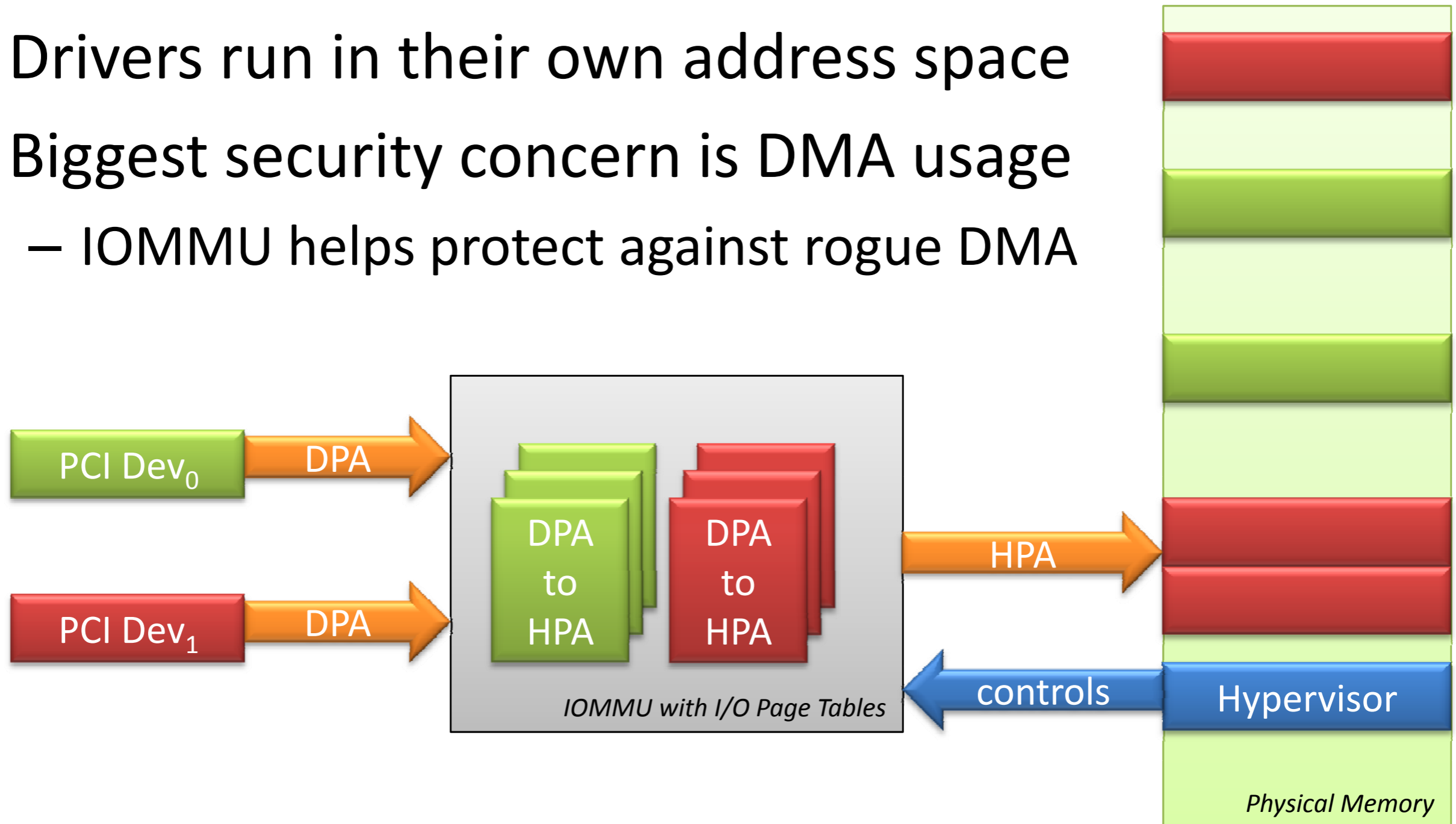


Handling of VM Exits



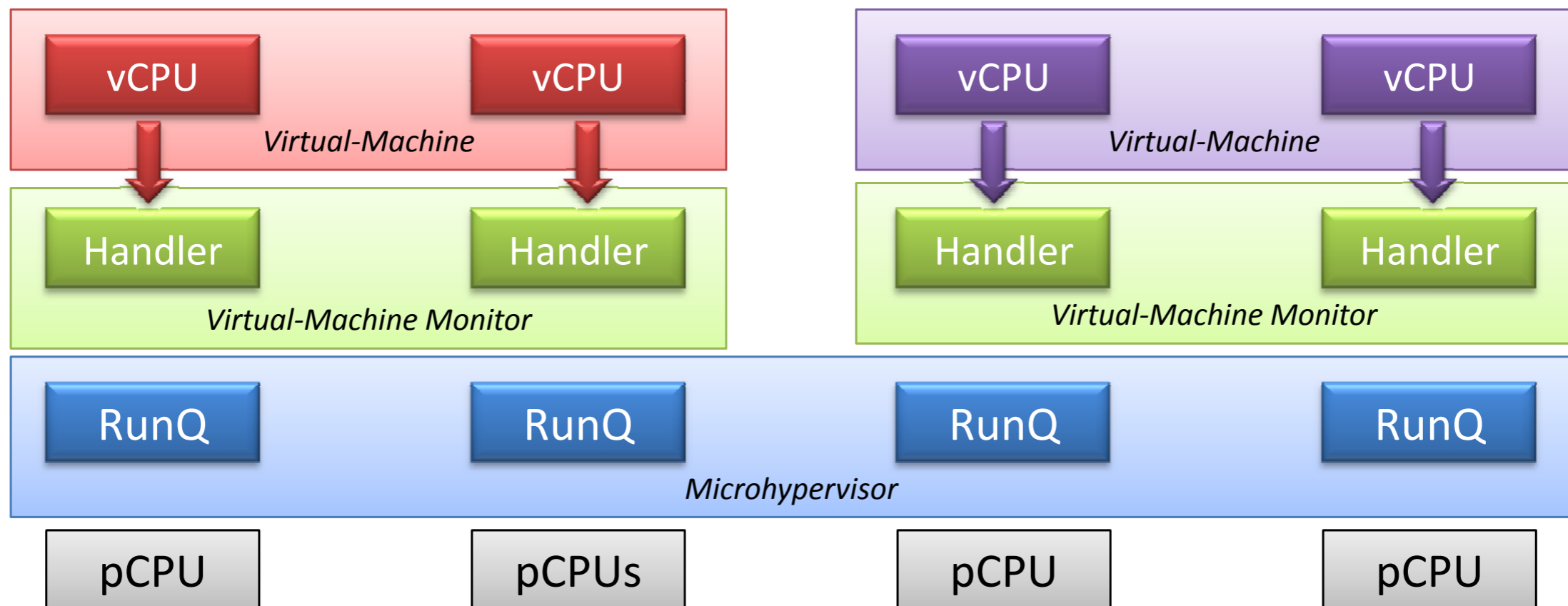
Device Drivers

- Drivers run in their own address space
- Biggest security concern is DMA usage
 - IOMMU helps protect against rogue DMA



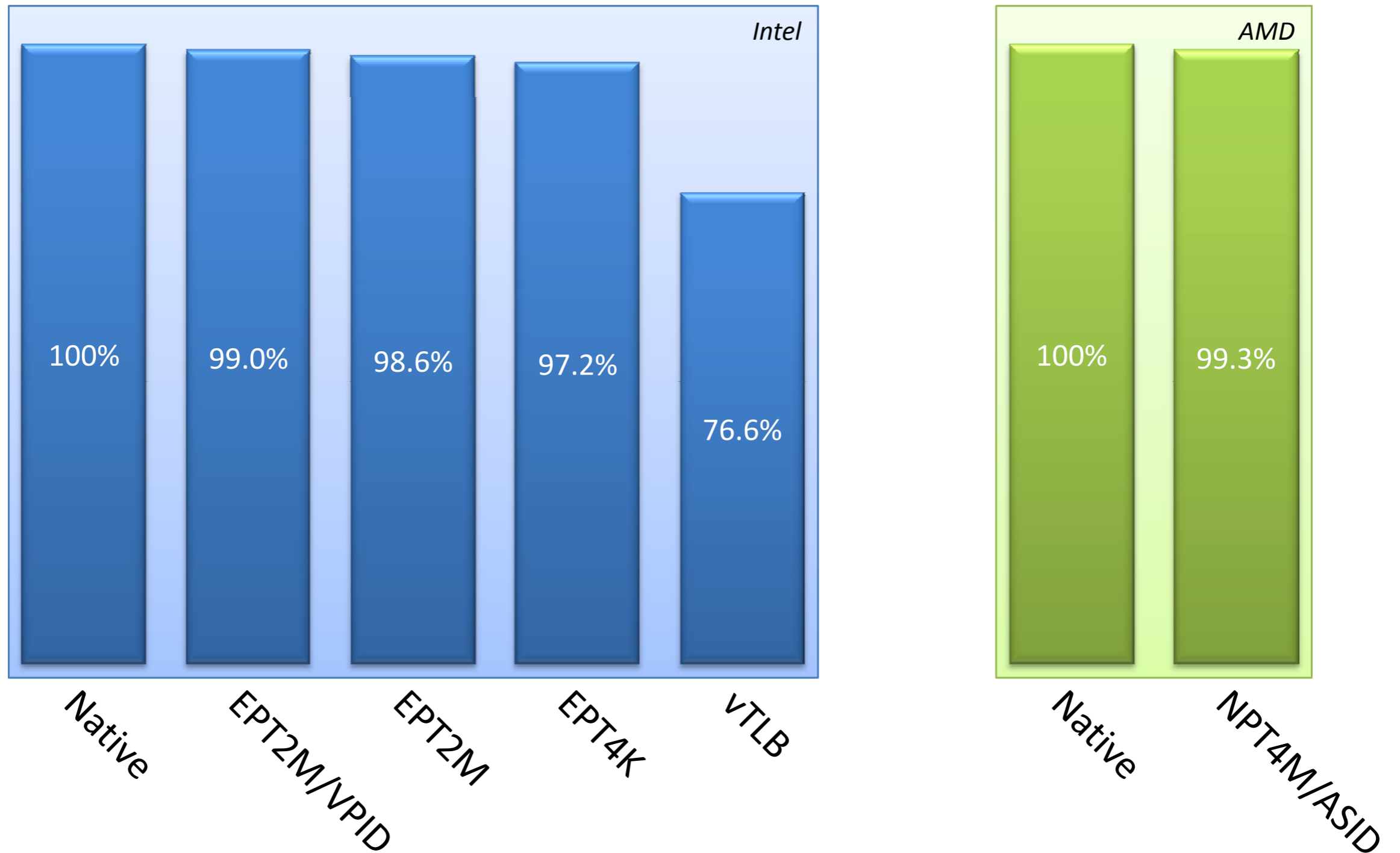
Multi-Core

- Expose as much parallelism as possible
 - CPU-local runqueues
 - One VMM handler thread per virtual CPU



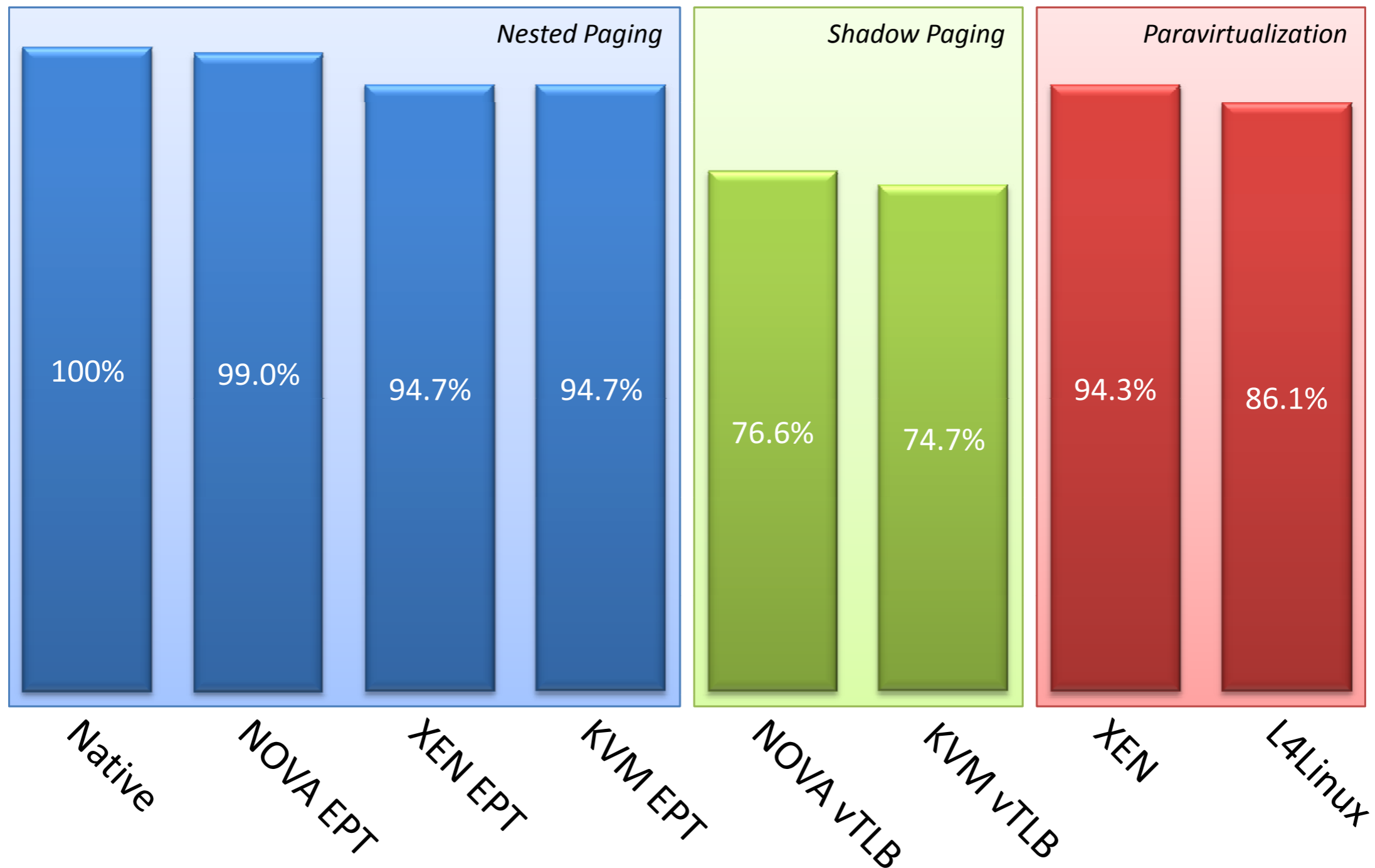
NOVA Performance

Linux Kernel Compile



Performance Comparison

Linux Kernel Compile



Summary

- NOVA provides full virtualization functionality with near-native performance
 - Runs on Intel VT and AMD-V
 - Can fully virtualize x86 with approx. 35000 LOC
- Fine-grain functional decomposition results in minimal application-specific TCB



Questions?