Virtualisation

Havard Bjerke

CERN openlab

Overview

- Virtual machines
- Benefits of virtualisation
- Computer architecture
  - Memory management
  - Privilege separation
  - Interrupts
- Virtualisation
- Para-virtualisation

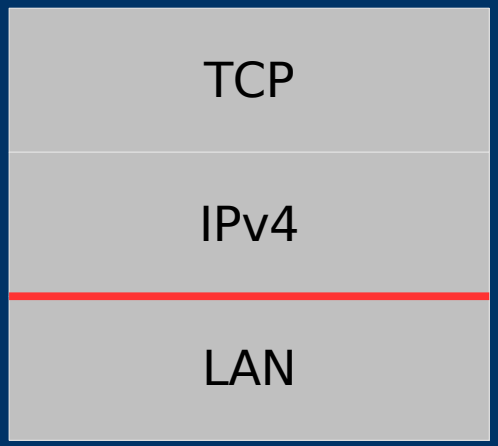
Virtual Machines

- Software level
  - Java
  - Software compatibility
- Hardware level
  - Ex: VMWare
  - Multiple OS instances

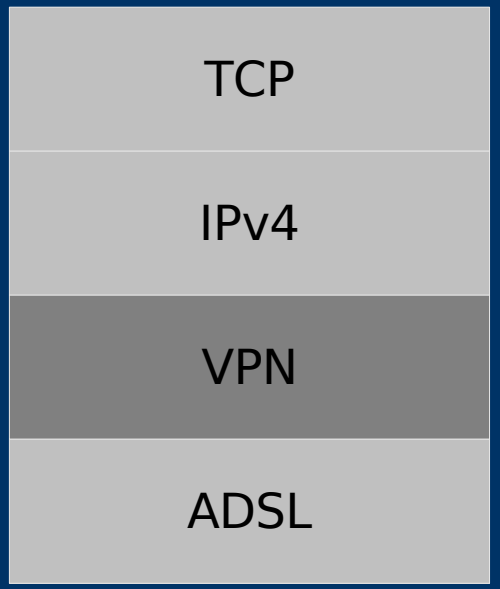
- Encapsulation
- Isolation

Abstraction vs Virtualisation

- Abstraction
  - TCP/IP stack
  - Replaceable layers
  - Friction between layers

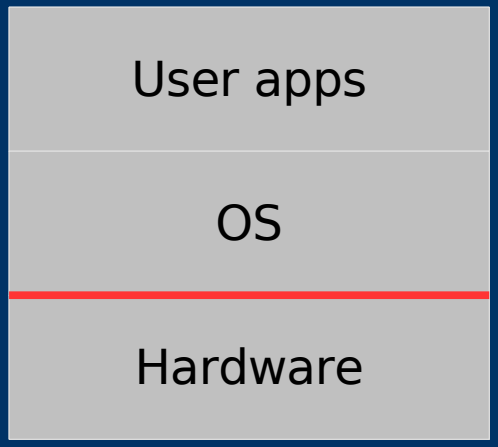


- Virtualisation
  - Virtual Private Networking (VPN)



Abstraction vs Virtualisation

- Computer abstraction layers



- Computer virtualisation

Benefits of HW virtualisation

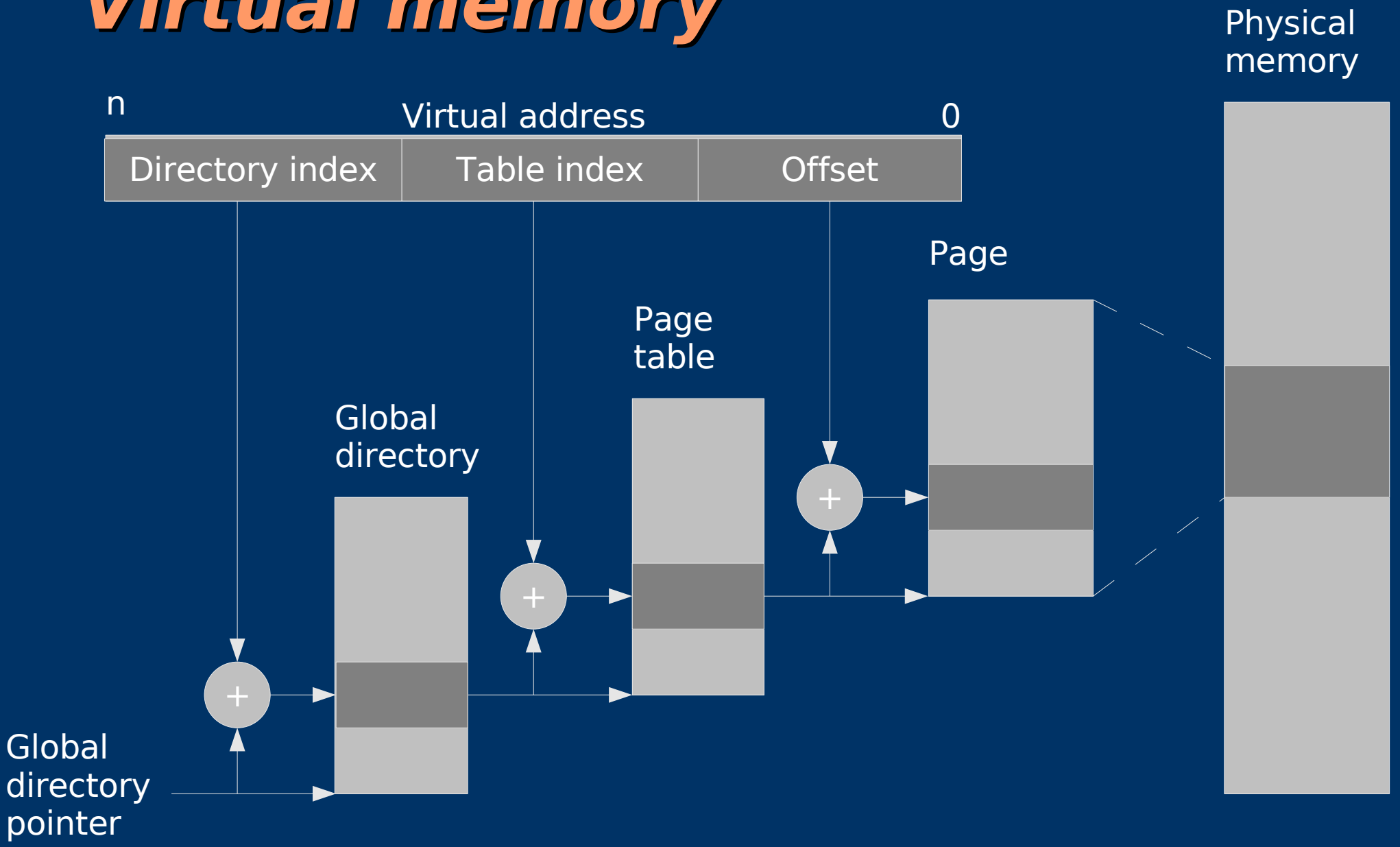
- General application:
  - Server consolidation
- HPC specific:
  - Software flexibility
    - Let each user manage their own OS
    - And satisfy their own software dependencies
  - Utilisation of SMP and multi-core resources.
  - Secure isolation between users
  - Migration between nodes
  - Checkpointing
  - Utilisation of public computing resources

Computer architecture

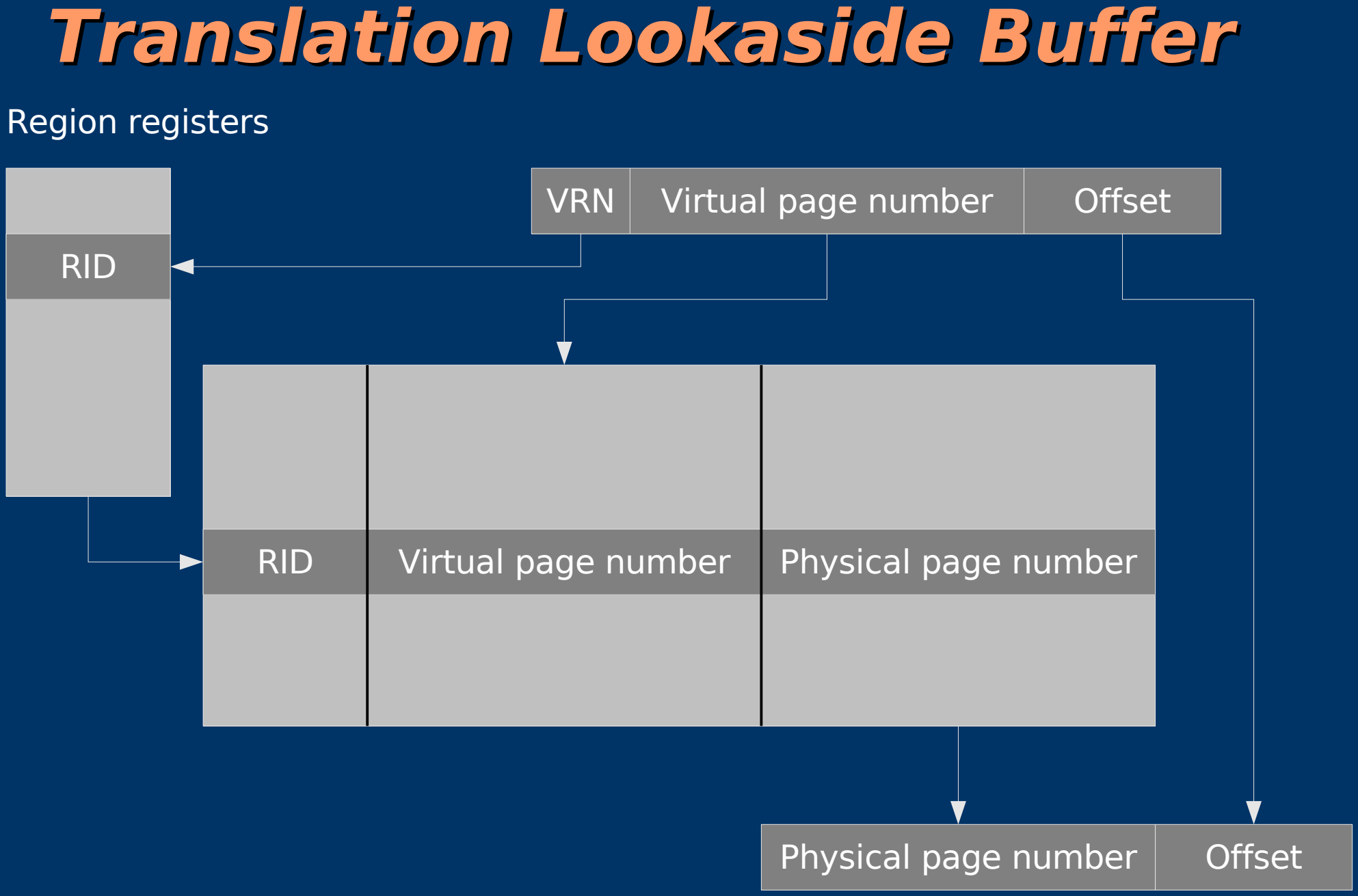
Computer architecture

- X86 – 80386, Pentium, Xeon
- X86_64 – AMD64, EM64T
- IA-64 – Itanium (IPF)

Virtual memory

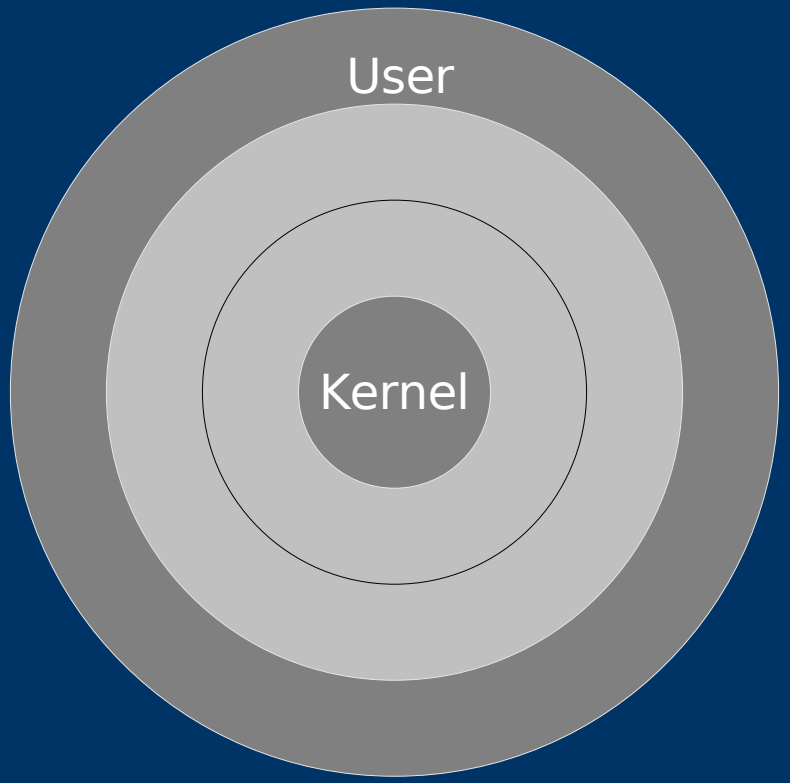


Translation Lookaside Buffer



Protection rings

- Protect kernel from faulty or malicious code
- Protection of
  - Privileged state
  - Privileged instructions
  - Privileged pages or segments



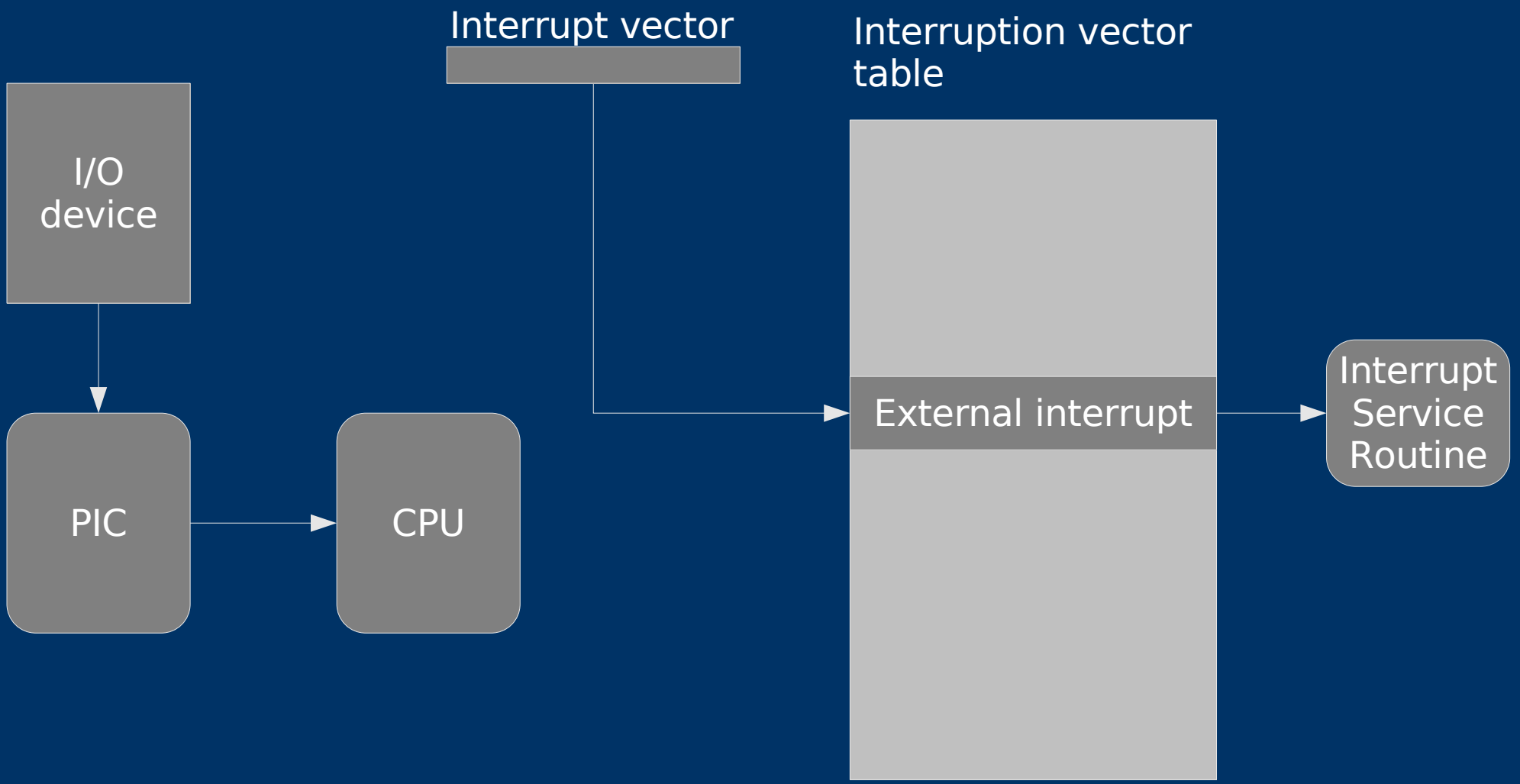
Kernel entry

- From ring 3 to ring 0 – From User space to Kernel space
- System calls
- Interrupt Service Routines
- Device access

Interrupts and exceptions

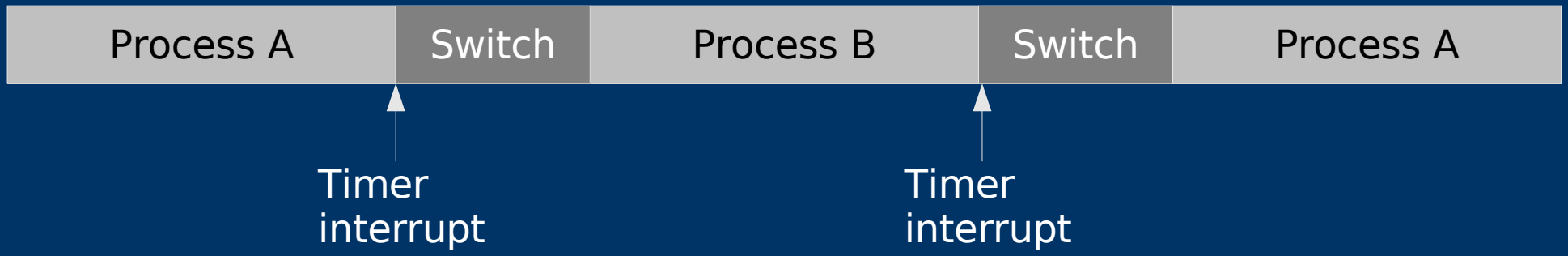
- Kernel entry
  - Exceptions
    - General protection fault
    - Segmentation fault
    - Page fault
    - Divide-by-zero
  - External interrupts
    - Keyboard
    - DMA finished
    - Packet on network
    - Timer

Interrupts and exceptions



Processes

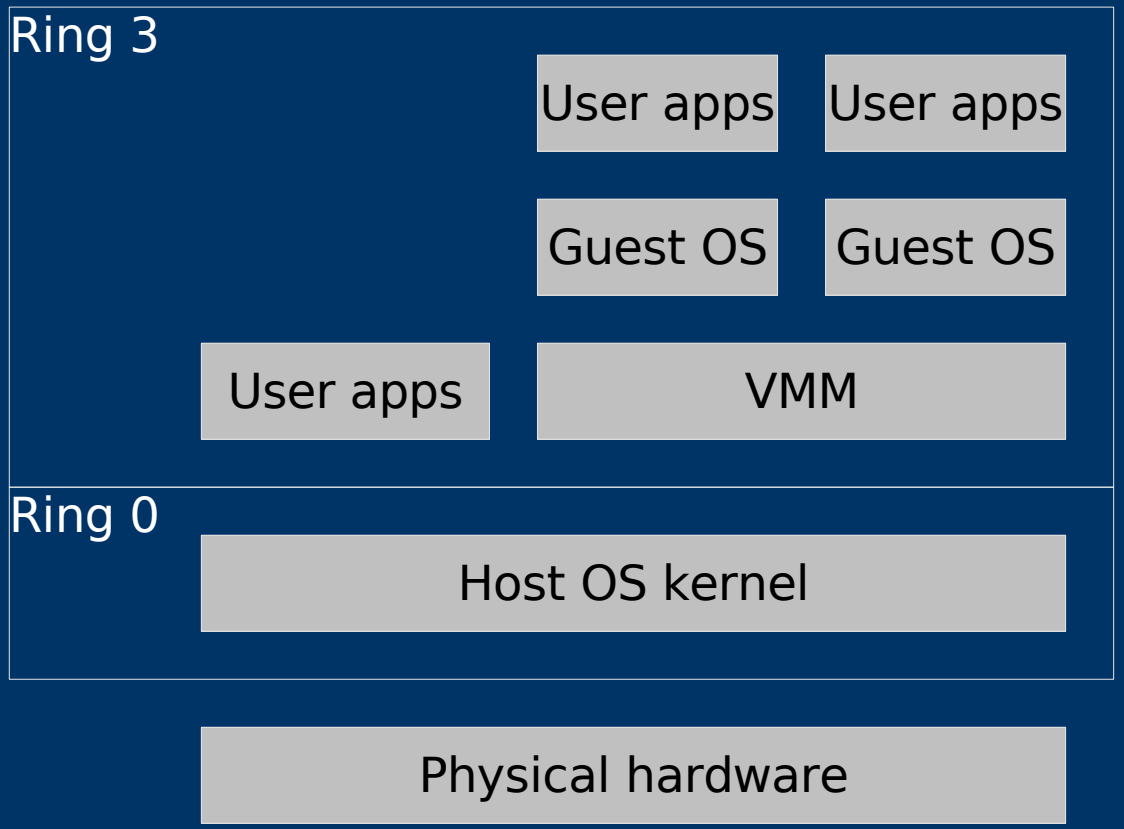
- Multitasking



Hardware Virtualization

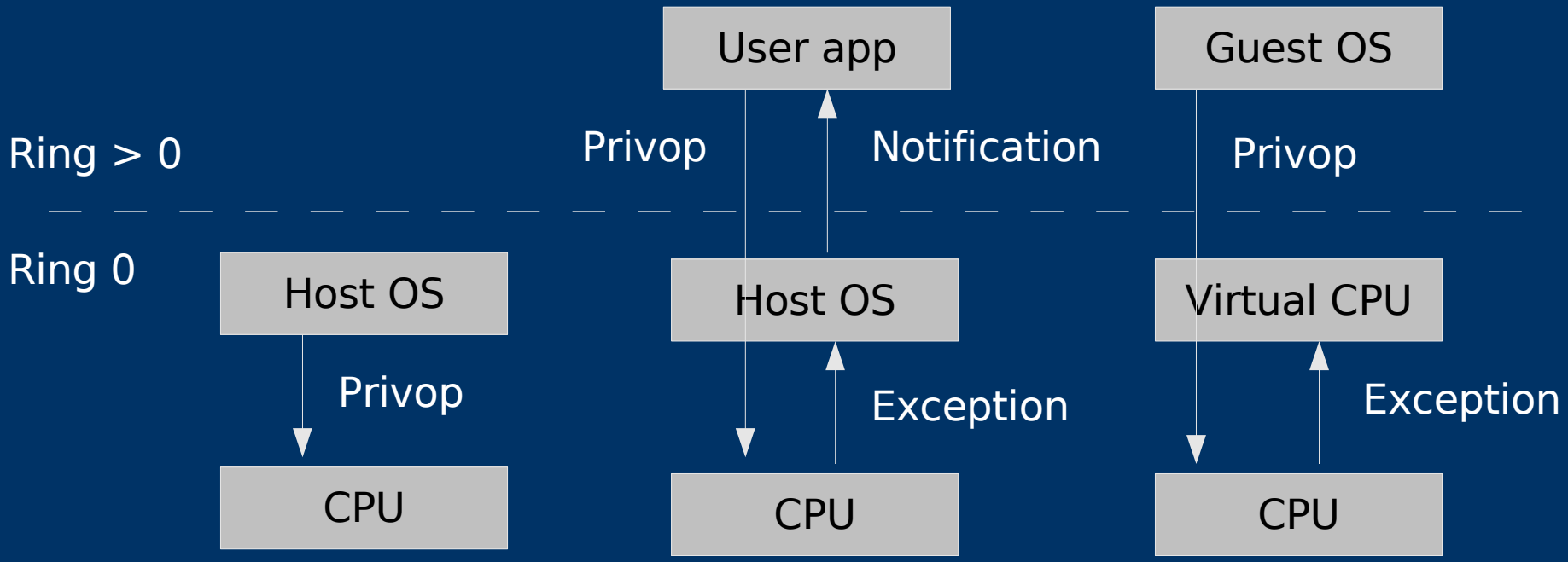
Virtualisation

- Interpretation
- Binary patching or translation
  - Privileged operations
  - Privilege-sensitive operations



Privileged operations

- The guest OS must think that it is privileged

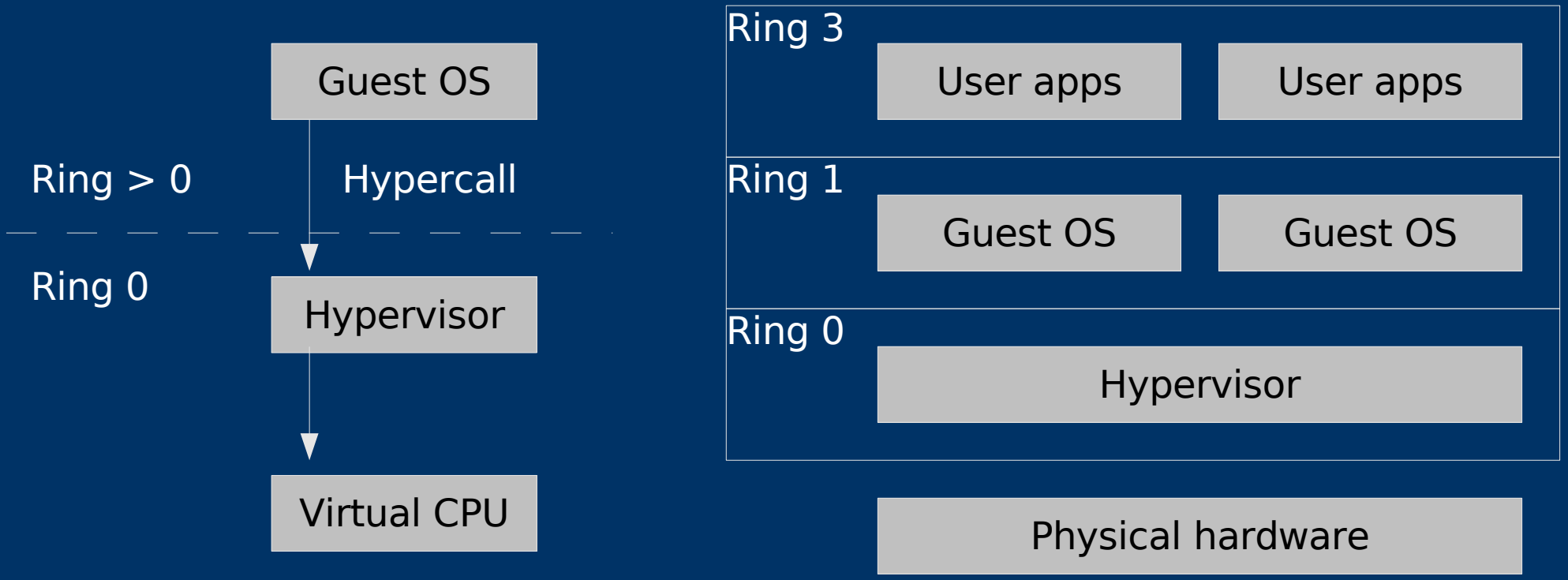


Privilege-sensitive operations

- Operations that are not protected, but
  - Access privileged state or
  - Whose results depend on CPL

Para-virtualisation

- Replace sensitive operations with calls to the Hypervisor - *hypercalls*



Xen memory management

- X86
  - Page table updates through hypercalls
  - Direct mapping between physical and virtual memory space
- IA-64
  - Logically separated address spaces using RIDs
  - Physical memory space has its own RID

Vanderpool (VT)

- VTx, VTi