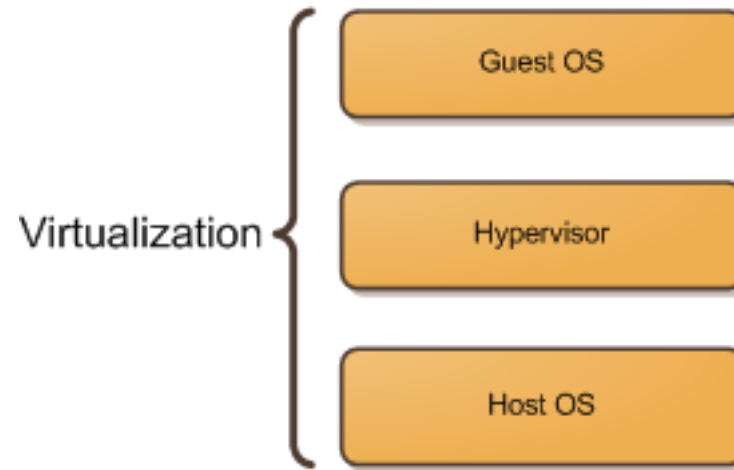# Virtualization (II)

SPD Course
17/03/2010
Massimo Coppola
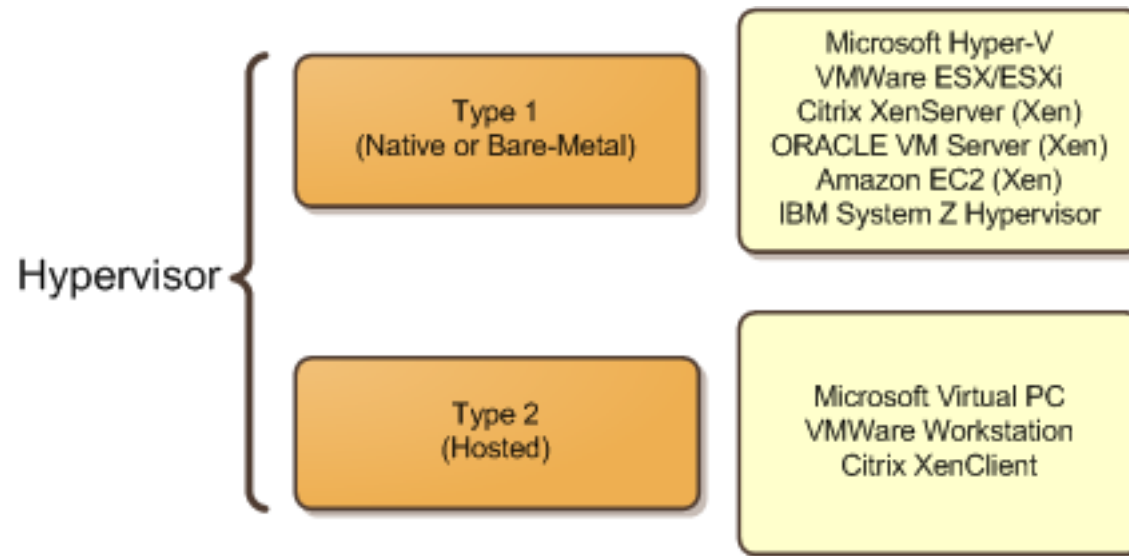
The players

- The Hypervisor (HV) implements the virtual machine emulation to run a Guest OS

- Provides resources and functionalities to the **Guest** OS

- Typical settings: the VM emulation stacks with the **Host** and the **Guest** OS

- System-level VM emulation = emulate each HW transition, including ISA, CPU and device user & privileged state
  - Different ISA → full emulation
 - Guest ISA = Host ISA → a subset of the ISA can be **executed**

Different Kinds of Hypervisor

- In system virtualization, the Hypervisor is not necessarily run on top of a host O.S.
- Even with no Host O.S. , the HV will still need a **Control** O.S.
  - Create, monitor and manage other VMs

Hypervisor {

Type 1
(Native or Bare-Metal)

Microsoft Hyper-V
VMWare ESX/ESXi
Citrix XenServer (Xen)
ORACLE VM Server (Xen)
Amazon EC2 (Xen)
IBM System Z Hypervisor

Type 2
(Hosted)

Microsoft Virtual PC
VMWare Workstation
Citrix XenClient

ISTITUTO DI SCIENZA E TECNOLOGIE
DELL'INFORMAZIONE "A. FAEDO"

Type 1 Hypervisor : bare-metal

- HV runs directly on the hardware
- Essential management functions: memory, CPU, system bus
- "Control" instance of the OS is special VM
  - Privileged link to the HV to steer other VMs
  - *May provide device drivers to Guests*
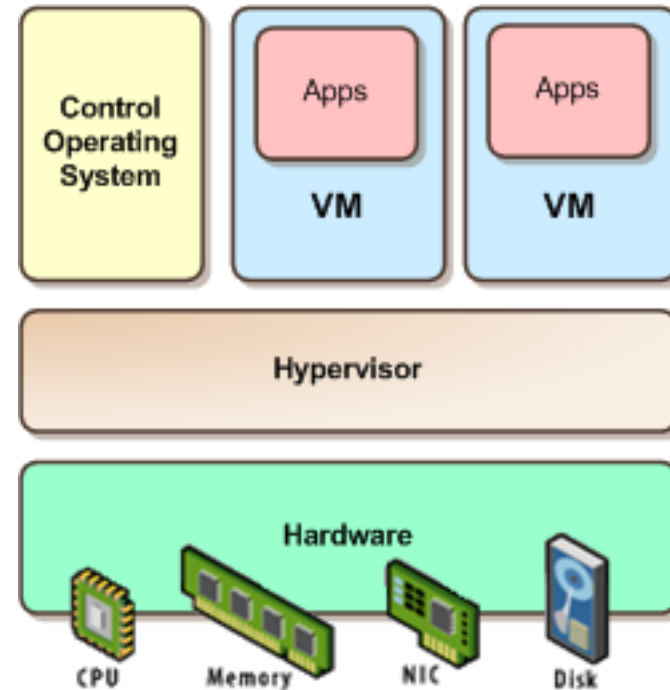
- Pros: complete knowledge of HW, can make HW-aware decisions
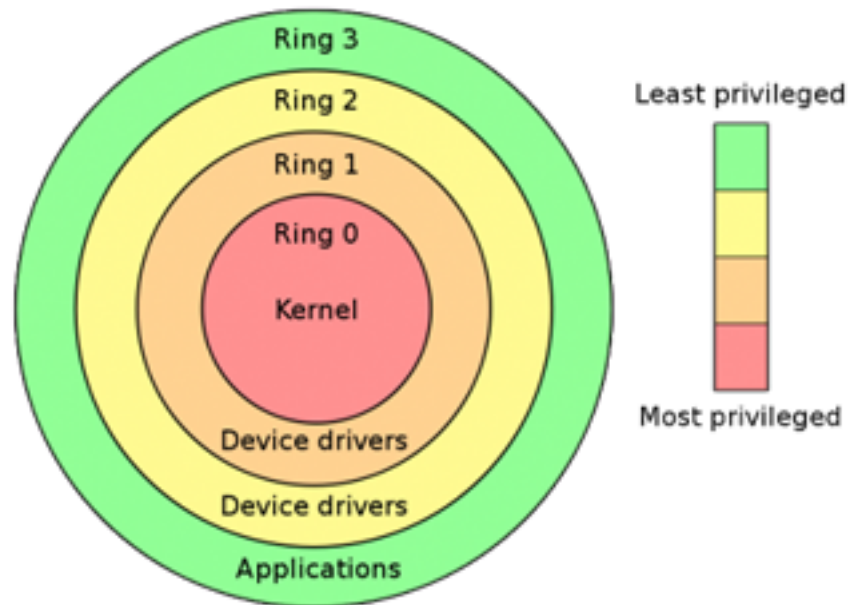- Cons: manage full VM emulation, intercept all HW mechanisms

Type 2 hypervisor : Hosted HV

- HV runs within the Host OS

- Resources controlled and managed by the Host

- Need mechanisms to separate host and guests if same ISA
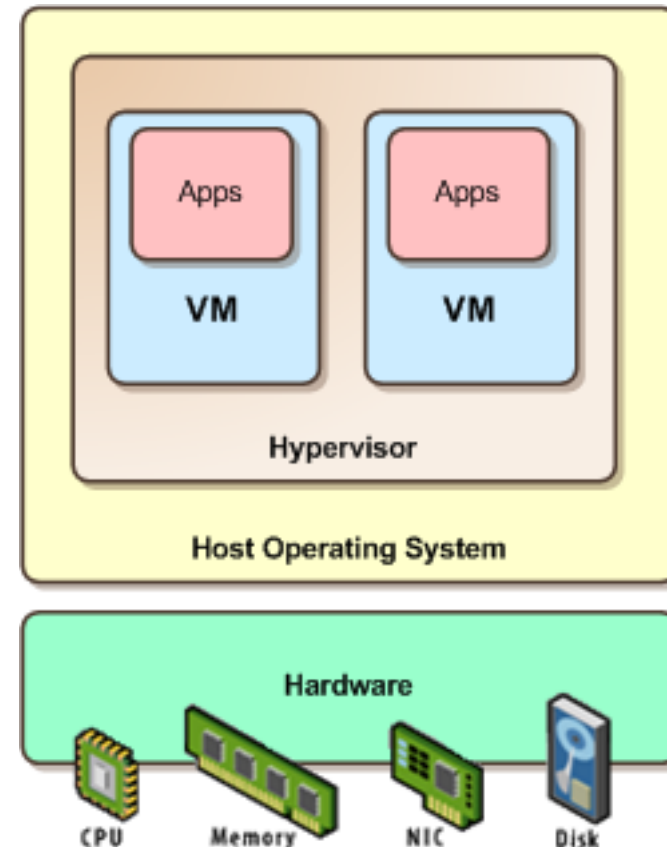
- Pros: less intrusive on Host OS

- Cons: low-level resource management done by host, HV has limited access to HW

Protection rings

- Current X86 architecture
- A generalization of the basic mechanism of the "supervisor" state
- Different classes of machine instructions (subsets of the ISA) are allowed in different levels
- Can be used as an OS containment mechanism
- In practice, most O.S. kernels use just one ring for OS (then it must be ring 0) and one for apps (ring 3)
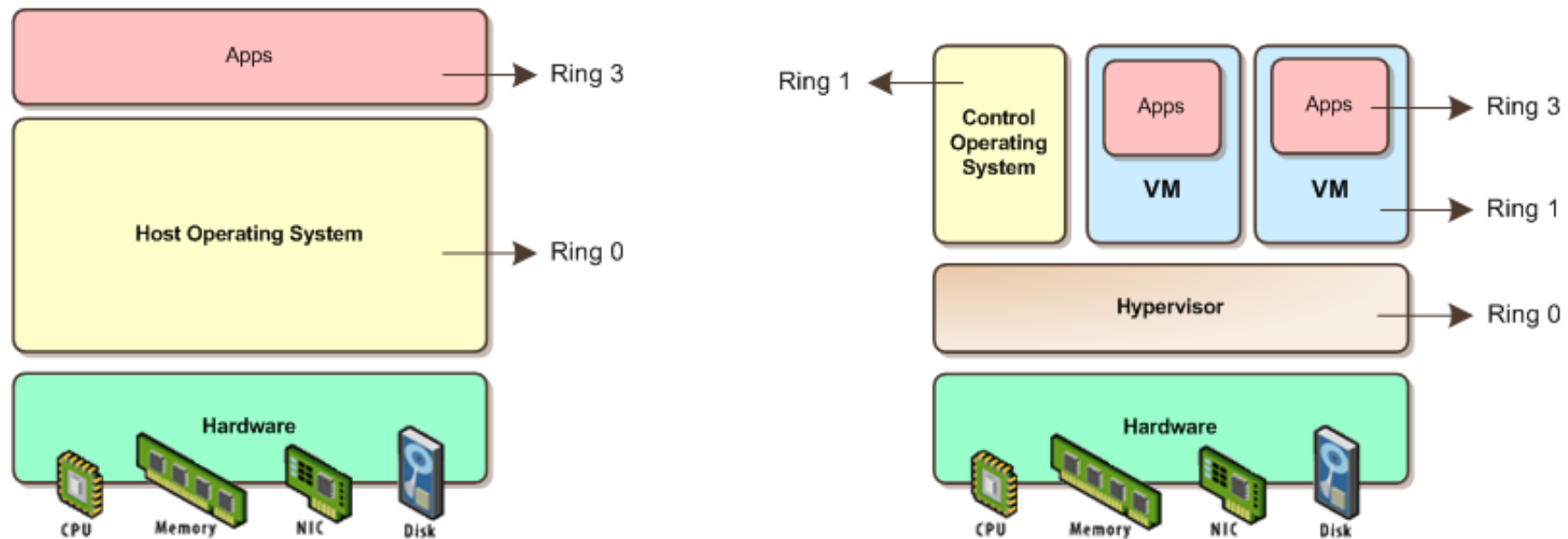
Paravirtualization

- Hypervisor needs to to run in ring 0
- Host O.S. kernel is modified
    - it runs in ring 0 on top of the hypervisor.
- Guest VM kernel is also modified
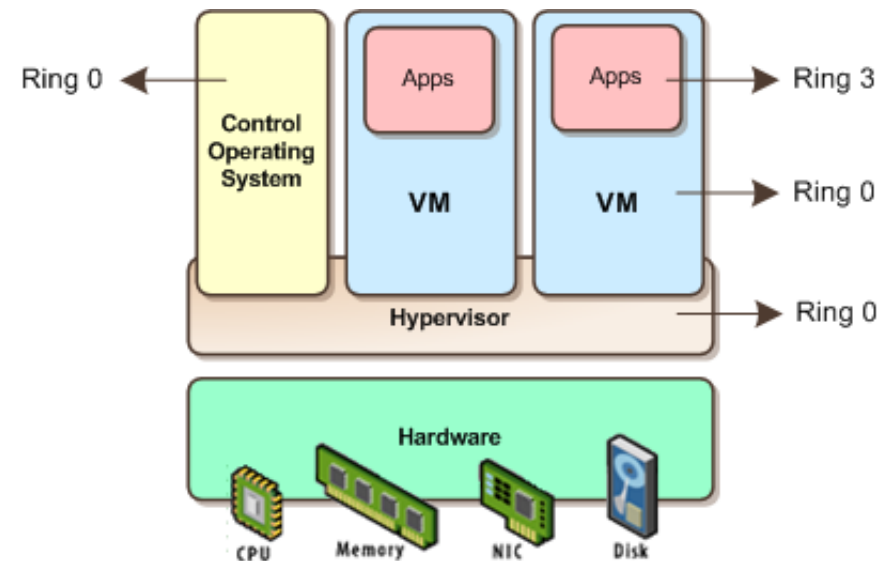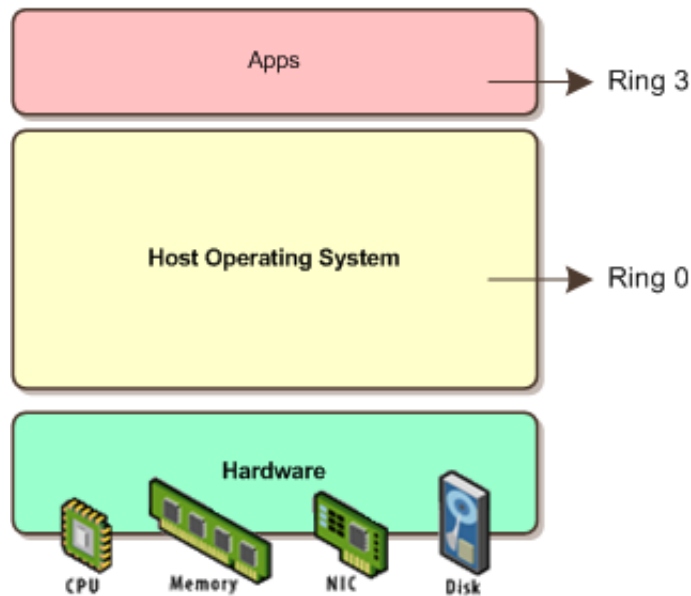
Full virtualization, no HW assist

- The HOST O.S. kernel is modified to run in ring 1
  - Some host kernels are easily ported to ring 1

- Guests may be modified, or run on SW emulation
 - A subset of the ISA trapped and emulated by the HV

Full virtualization with HW assist

- Hypervisor is run in a special ring **-1**
- **Intel and AMD** provide HW support in recent CPUs
- Control O.S. kernel is run in ring 0
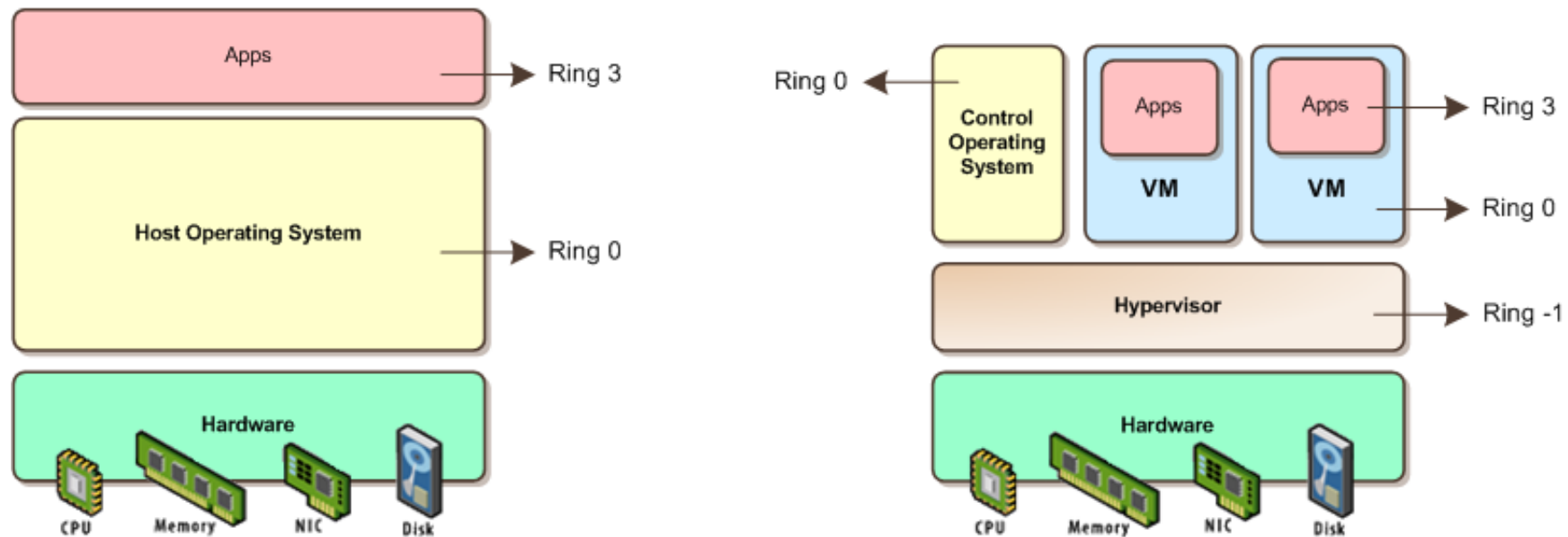
Containers

- Process-level VMs for one (group of) processes
- OS – like abstraction, inside container, no info about the outside
- Isolation and protection are the aim

- Special support into in the Host O.S.
  - No HW virtualization need (same ISA!)

- May exploit system VM in order to build containers
  - OpenVZ technology as used in XtreemOS
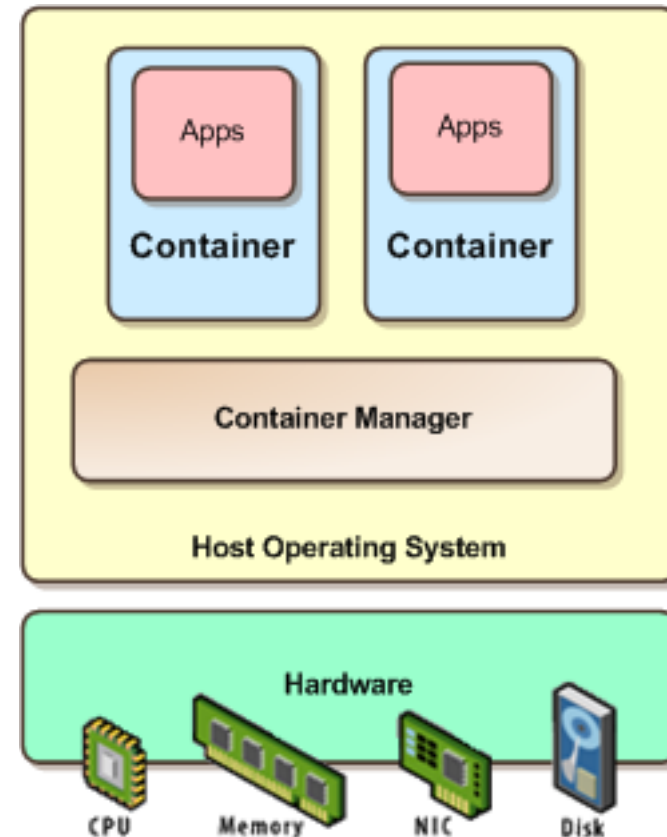
Hosted virtualization
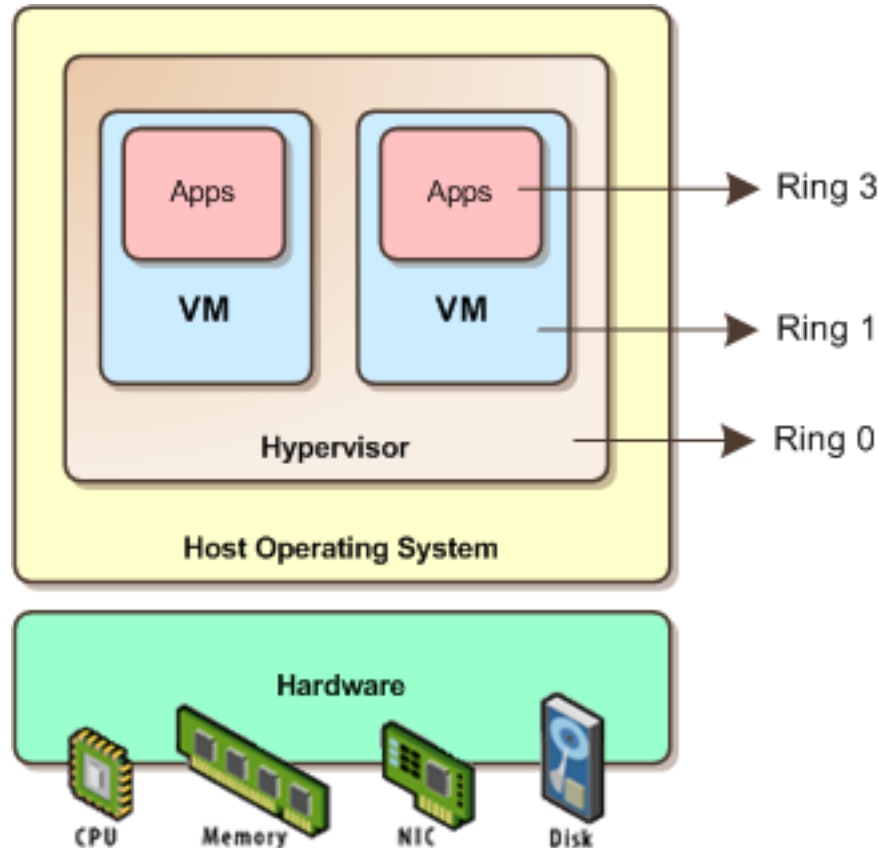
- Common case
- You virtualize also device drivers to enhance VM interoperation
  - Virtual disk drivers
  - Virtual video drivers mapped in the host desktop
  - Inter-VM cut & paste

The MPI Message-passing Standard (V) Practical use

SPD Course

17/03/2010

Massimo Coppola

ISTITUTO DI SCIENZA E TECNOLOGIE
DELL'INFORMAZIONE "A. FAEDO"

Exercises

- Define a program with 2 processes
  - they send back and fort a data buffer, the second process executes an operation on the data (e.g. sum 1).
  - Verify after a given number N of iterations, that the expected result is achieved.

- Define a datatype for a square matrix, with parametric size. Define a datatype for its lower triagular matrix. Define one for its upper triangular.

- In the two-process program, initialize randomly a square matrix, send the lower triangular and receive it back as upper triangular in the same buffer.
  - Is the result a symmetric matrix?
  - Do you need to modify one of the two triangular datatypes?