x86 : Virtualization

Gabriel Laskar <gabriel@lse.epita.fr>

LSE Security System

Basics : What is it ?

- Virtual Machine
- Hypervisor
- Virtual machine monitor

LSE Security System

Basics : Virtualization vs Emulation

- CPU Emulation : Interpret code in order to execute the same behavior
- CPU Virtualization : Execute on real hardware, but in a controlled way

LSE Security System

VM Requirements

« For any conventional third generation computer, a virtual machine monitor may be constructed if the set of sensitive instructions for that computer is a subset of privileged instructions »

-- Popek & Goldberg

LSE Security System

Virtualization Solutions

- Xen
- Qemu/KVM
- VMWare ESX
- VMWare Workstation
- VirtualBox

Other Kind of Virtualization

- Paravirtualization
- Containers

LSE Security System

CPU Virtualization

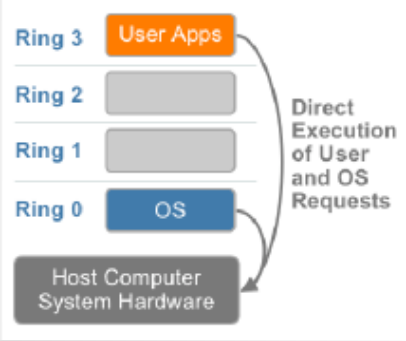
- Run the VMM at a higher level of privilege
- Sensitive instructions will trap and the VMM will emulate them.

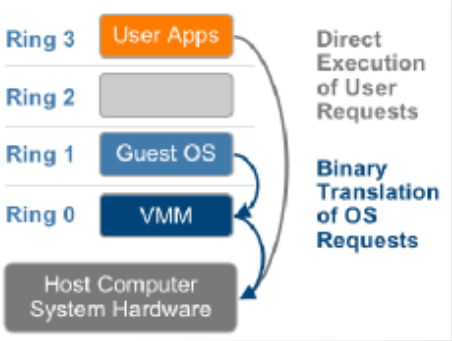
Virtualize the “unvirtualizable”

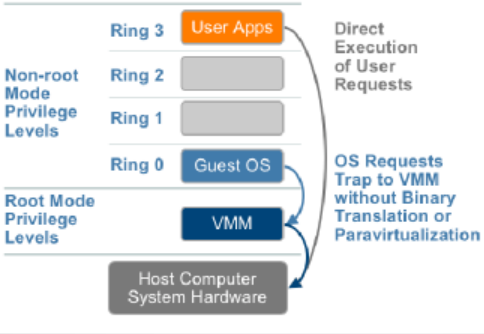
- Binary Rewriting
- Para-virtualization
- HVM

LSE Security System

Rings & Virtualization







LSE Security System

vt-x

- root vs non-root mode
- VMCS
- Instructions
- What can trap ?

LSE Security System

vt-x : instructions

- vmptrld, vmptrst
- vmclear
- vmread, vmwrite
- vmlaunch, vmresume
- vmxoff, vmxon
- invept, invvpid
- vmcall, vmfunc

LSE Security System

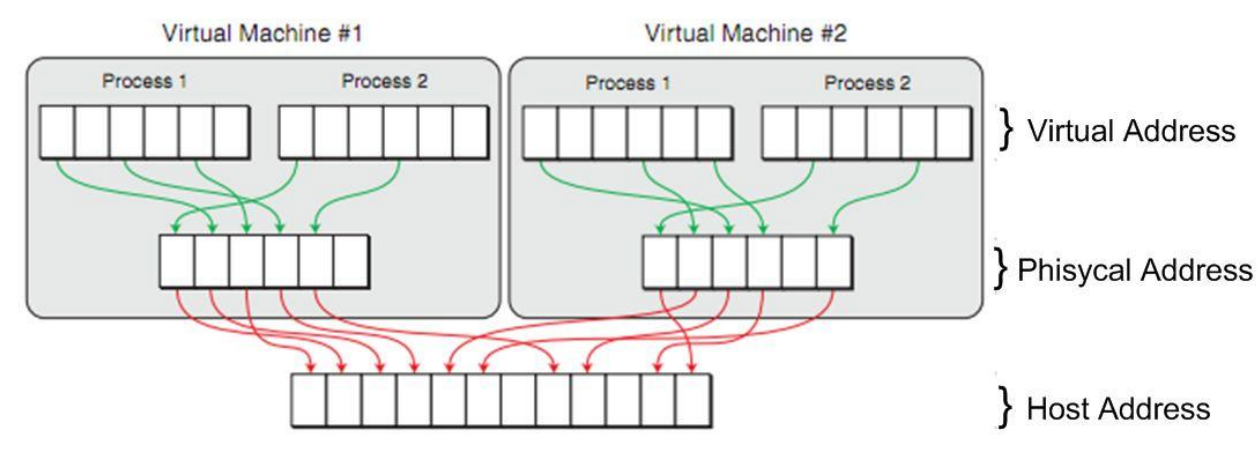
EPT

- No Shadow Page Tables
- A second translation Layer
- translation : physical → guest-physical

LSE Security System

Memory Virtualization

- shadow page tables
- EPT



LSE Security System

vt-d

I/O Virtualization

- HW emulation
- Paravirtualized Devices
  - Xen
  - Virtio
  - Other (vmxnet, synthetic devices, …)
- Hardware Pass through
  - Full Device (pci, vga)
  - Protocol (usb, serial)
  - Other way

LSE Security System

Use case : Qemu/KVM

LSE Security System

Virtio Devices

Questions ?

LSE Security System