## Hardware and Software Support for Virtualization; Edouard Bugnion, Jason Nieh, Dan Tsafrir; Synthesis Lectures on Computer Architecture, 2017

1. Definitions:
   1. Virtualization

The application of the layering principle through enforced modularity, whereby the exposed virtual resource is identical to the underlying physical resource being virtualized

* 1. Virtual Machine

An abstraction of a compute environment with all resources

A complete compute environment with its own isolated processing capabilities, memory, and communication channels

* 1. Hypervisor / Virtual Machine Monitor (VMM)

A specialized piece of system software that manages and runs virtual machines

* 1. Type-1 Hypervisor

Hypervisor runs on a bare machine; has direct control of all resources of the physical computer



* 1. Type-2 Hypervisor

Hypervisor operates as part of an existing host operating system; runs on an extended host



1. Three Specific Criteria for System Virtualization?

Equivalence:

The exposed resource (i.e., the VM) is equivalent with the underlying computer. A strong requirement, relaxed when the architecture demands it.

Safety:

Isolation requires that the VMs are isolated from each other and the hypervisor. Safety is enforced by the hypervisor without it making any assumptions about the software in the VM (including the guest OS)

Performance:

The virtual system must show at worst a minor decrease in speed.

Differentiates hypervisors from machine simulators

1. Resources
   1. CPU: Direct Execution

Hypervisor lets the virtual CPU directly execute instructions on the real processor for good performance

* 1. Trap-and-emulate

Safety: When guest OS attempts to execute privileged instructions 🡪 trap into hypervisor which emulates instruction

* 1. I/O: Front-end device and back-end emulation

Front-end device: Canonically device abstraction, visible as hardware to VM

Back-end emulation: Uses underlying physical resources to implement functionality needed by front-end device

* 1. Memory: virtual memory, guest-physical memory, host-physical memory

Disco definitions?

Guest-physical memory 🡪 physical memory

Host-physical memory 🡪 machine memory

1. Approaches to Virtualization
   1. Full software virtualization

Run unmodified OSes on architectures lacking the full support for hardware virtualization

Must sometimes translate guest instructions

* 1. Hardware Virtualization (HVM or HV)

Run unmodified OSes; Hypervisors built for architectures with full support for virtualization (all recent processors); rely exclusively on direct execution

* 1. Paravirtualization

Values simplicity and overall efficiency over full compatibility with underlying hardware

Introduced by Denali, popularized by early Xen hypervisor

Requires changes to guest OS (now often in device drivers)

1. Benefits of Virtual Machines?

OS Diversity: Run different OSes

Server consolidation: Enterprise IT best practices say run each application on its own machine

Rapid provisioning: VMs can be created quickly and software stack can be deployed as a virtual appliance

Security: VMs enable introspection, intrusion analysis, logging of OS activity; can also insert firewalls

High-availability: After hardware crash, can move VM quickly to new server

Distributed resource scheduling: Use live migration to turn cluster into single resource pool

Cloud computing: Different customers can operate own VMs in isolation from one another

1. Popek/Goldberg Theorem: A VMM may be constructed if the set of sensitive instructions for that computer is a subset of the set of privileged instructions. What are sensitive instructions?

Control sensitive: Can update the system state

Behavior sensitive: Semantics depend on the actual values set in the system state

Privileged: Can only be executed in supervisor mode and causes a trap when attempted from user mode

Violations:

Direct access to physical memory (KSEG0 in MIPS)

Able to read sensitive data structures via unprivileged instructions (interrupt flag, segment descriptors in x86)

Different semantics for same instruction, depending on privilege level