- All OS-level VMs on the same physical machine share a single operating system kernel; and (2) the virtualization layer can be designed in a way that allows processes in VMs to access as many resources of the host machine as possible, but never to modify them.

Disadvantages of OS Extensions

- The main disadvantage of OS extensions is that all the VMs at operating system level on a single container must have the same kind of guest operating system.
- For example, a Windows distribution such as Windows XP cannot run on a Linux-based container.
- Some prefer Windows and others prefer Linux or other operating systems.

3.6 VIRTUALIZATION STRUCTURES/TOOLS AND MECHANISMS

Part –B	1.What is Virtualization? Describe para and Full virtualization architectures. Compare and Contrast them (AU/Nov/Dec 2017)

- In general, there are three typical classes of VM architecture.
- Before virtualization, the operating system manages the hardware.
- After virtualization, a virtualization layer is inserted between the hardware and the operating system.
- In such a case, the virtualization layer is responsible for converting portions of the real hardware into virtual hardware. Therefore, different operating systems such as Linux and Windows can run on the same physical machine, simultaneously.
- Depending on the position of the virtualization layer, there are several classes of VM architectures, namely the hypervisor architecture, para-virtualization, and host-based virtualization.
- The hypervisor is also known as the VMM (Virtual Machine Monitor). They both perform the same virtualization operations.

Hypervisor and Xen Architecture

- The hypervisor supports hardware-level virtualization on bare metal devices like CPU, memory, disk and network interfaces.
- The hypervisor software its directly between the physical hardware and its OS. This virtualization layer is referred to as either the VMM or the hypervisor.
- The hypervisor provides hyper calls for the guest OS and applications. Depending on the functionality, a hypervisor can assume micro kernel architecture like the Microsoft Hyper-V Or it can assume a monolithic hypervisor architecture like the VMware ESX for server virtualization.
- A micro-kernel hypervisor includes only the basic and unchanging functions (such asphysical memory management and processor scheduling).
- The device drivers and other changeable components are outside the hypervisor.
- A monolithic hypervisor implements all the aforementioned functions, including those of the device drivers.
- Therefore, the size of the hypervisor code of a micro-kernel hypervisor is smaller than that of a monolithic hypervisor. Essentially, a hypervisor must be able to convert physical devices into virtual resources dedicated for the deployed VM to use.

The Xen Architecture

- Xen is an open source hypervisor program developed by Cambridge University.
- Xen is a micro-kernel hypervisor, which separates the policy from the mechanism.
- The Xen hypervisor implements all the mechanisms, leaving the policy to be handled by Domain 0, Xen does not include any device drivers natively.

- It just provides a mechanism by which a guest OS can have direct access to the physical devices.
- As a result, the size of the Xen hypervisor is kept rather small.
- The core components of a Xen system are the hypervisor, kernel, and applications.
- The organization of the three components is important.
- Like other virtualization systems,many guest OS can run on top of the hypervisor. However, not all guest OS are created equal, and one in particular controls the others. The guest OS, which has controlability, is called Domain 0, and the others are called Domain U.
- Domain 0 is a privileged guest OS of Xen. It is first loaded when Xen boots without any file system drivers being available.
- Domain 0 is designed to access hardware directly and manage devices.
- For example, Xen is based on Linux and its security level is C2.
- Its management VM is named Domain 0, which has the privilege to manage other VMs implemented on the same host.
- If Domain 0 is compromised, the hacker can control the entire system.
- So, in the VM system, security policies are needed to improve the security of Domain 0.
- Domain 0, behaving as a VMM, allows users to create, copy, save, read, modify, share,migrate, and roll back VMs as easily as manipulating a file, which flexibly provides tremendous benefits for users.

3.6.1 FULL VIRTUALIZATION

- With full virtualization, noncritical instructions run on the hardware directly while
- critical instructions are discovered and replaced with traps into the VMM to be emulatedby software.
- Both the hypervisor and VMM approaches are considered full virtualization.
- Noncritical instructions do notcontrol hardware or threaten the security of the system, but critical instructions do.

Binary Translation of Guest OS Requests Using a VMM

- This approach was implemented by VMware and many other software companies.
- VMware puts the VMM at Ring 0 and the guest OS at Ring 1.
- TheVMM scans the instruction stream and identifies the privileged, control- and behaviorsensitiveinstructions.
- When these instructions are identified, they are trapped into theVMM, which emulates the behavior of these instructions.
- The method used in thisemulation is called binary translation.
- Therefore, full virtualization combines binarytranslation and direct execution.
- The guest OS is completely decoupled from theunderlying hardware.
- Consequently, the guest OS is unaware that it is being virtualized.

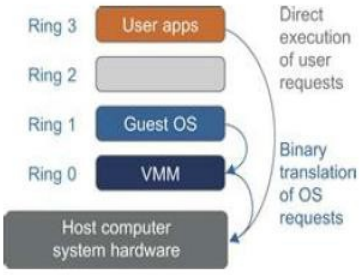


FIGURE 3.6 Indirect execution of complex instructions via binary translation of guest OS requests using the VMM plus direct execution of simple instructions on the same host. Courtesy of VMWare [71]

- The performance of full virtualization may not be ideal, because it involves binary translation which is rather time-consuming.

3.6.2 HOST-BASED VIRTUALIZATION

- An alternative VM architecture is to install a virtualization layer on top of the host OS.
- This host OS is still responsible for managing the hardware.
- The guest OS are installed and run on top of the virtualization layer.
- Dedicated applications may run on the VMs.
- The host-based architecture has some distinct advantages, as enumerated next.
- First, the user can install this VM architecture without modifying the host OS.
- The virtualizing software can rely on the host OS to provide device drivers and other low-level services.
- Second, the host-based approach appeals to many host machine configurations.
- Compared to the hypervisor/VMM architecture, the performance of the host-based architecture may also be low.
- When an application requests hardware access, it involves four layers of mapping which downgrades performance significantly.
- When the ISA of a guest OS is different from the ISA of the underlying hardware, binary translation must be adopted.

3.6.3 PARA-VIRTUALIZATION WITH COMPILER SUPPORT

- Para-virtualization needs to modify the guest operating systems.
- A para-virtualized VM provides special APIs requiring substantial OS modifications in user applications.
- Performance degradation is a critical issue of a virtualized system.
- No one wants to use a VM if it is much slower than using a physical machine.
- The virtualization layer can be inserted at different positions in a machine software stack. However, para-virtualization attempts to reduce the virtualization overhead, and thus improve performance by modifying only the guest OS kernel.
- The guest operating systems are para-virtualized.
- They are assisted by an intelligent compiler to replace the non virtualizable OS instructions by hyper calls.
- The traditional x86 processor offers four instruction execution rings: Rings 0, 1, 2, and 3. The lower the ring number, the higher the privilege of instruction being executed.
- The OS is responsible for managing the hardware and the privileged instructions to execute at Ring0, while user-level applications run at Ring 3. The best example of para-virtualization is the KVM to be described below.

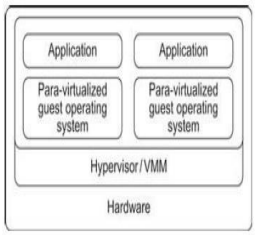
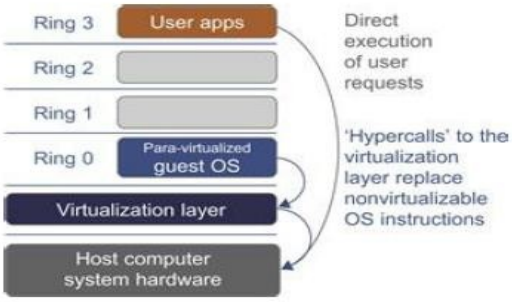


FIGURE 3.7 Para-virtualized VM architecture, which involves modifying the guest OS kernel to replace nonvirtualizable instructions with hypercalls for the hypervisor or the VMM to carry out the virtualization process (See Figure 3.8 for more details.)



3.7 VIRTUALIZATION OF CPU, MEMORY, AND I/O DEVICES

Part –B	1. Discuss how virtualization is implemented in CPU, Memory and I/O Devices.

- To support virtualization, processors such as the x86 employ a special running mode and instructions, known as hardware-assisted virtualization.
- In this way, the VMM and guest OS run in different modes and all sensitive instructions of the guest OS and its applications are trapped in the VMM.
- To save processor states, mode switching is completed by hardware. For the x86 architecture, Intel and AMD have proprietary technologies for hardware-assisted virtualization.

HARDWARE SUPPORT FOR VIRTUALIZATION

- Modern operating systems and processors permit multiple processes to run simultaneously.
- Therefore, all processors have at least two modes, user mode and supervisor mode, to ensure controlled access of critical hardware.
- Instructions running in supervisor mode are called privileged instructions.
- Other instructions are unprivileged instructions.
- In a virtualized environment, it is more difficult to make OS and applications run correctly because there are more layers in the machine stack.

SOFTWARE:

- The VMware Workstation is a VM suite for x86 and x86-64 computers.
- This software suite allows users to set up multiple x86 and x86-64 virtual computers and to use one or more of these VMs simultaneously with the host operating system.
- The VMware Workstation assumes the host-based virtualization.
- Xen is a hypervisor for use in IA-32,x86-64, Itanium, and PowerPC 970 hosts.