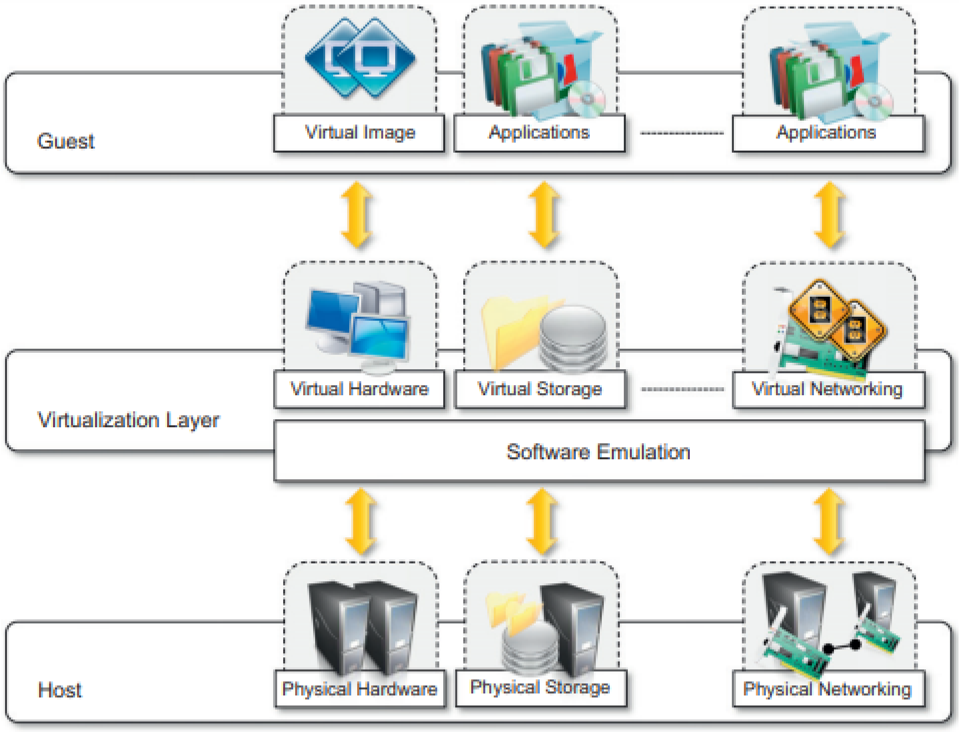
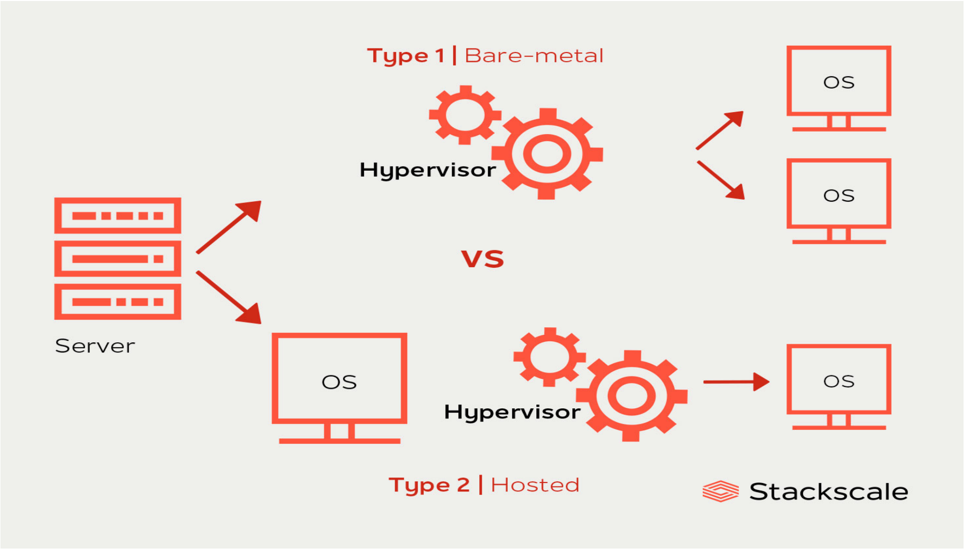
Introduction to Virtualization

1. Virtualization is a broad concept that refers to the creation of a virtual version of something, whether hardware, a software environment, storage, or a network.
2. In a virtualized environment there are three major components: guest, host, and virtualization layer.
3. The guest represents the system component that interacts with the virtualization layer rather than with the host, as would normally happen.
4. The host represents the original environment where the guest is supposed to be managed.
5. The virtualization layer is responsible for recreating the same or a different environment where the guest will operate.
6. Virtualization technologies provide a virtual environment for not only executing applications but also for storage,memory,and networking.



7. The most intuitive and popular is represented by hardware virtualization, which also constitutes the original realisation of the virtualization concept.
8. In the case of hardware virtualization, the guest is represented by a system image comprising an operating system and installed applications.
9. These are installed on top of virtual hardware that is controlled and managed by the virtualization layer, also called the virtual machine manager.

10. The host is instead represented by the physical hardware, and in some cases the operating system, that defines the environment where the virtual machine manager is running.
11. In the case of virtual storage, the guests might be client applications or users that interact with the virtual storage management software deployed on top of the real storage system.
12. The case of virtual networking is also similar: The guest— applications and users—interacts with a virtual network, such as a virtual private network (VPN), which is managed by specific software (VPN client) using the physical network available on the node.
13. The main common characteristic of all these different implementations is the fact that the virtual environment is created by means of a software program also known as hypervisor.
14. A hypervisor, also known as a virtual machine monitor or VMM, is software that creates and runs virtual machines (VMs).
15. There are two main hypervisor types, referred to as “Type 1” (or “bare metal”) and “Type 2” (or “hosted”). A type 1 hypervisor acts like a lightweight operating system and runs directly on the host’s hardware, while a type 2 hypervisor runs as a software layer on an operating system, like other computer programs.
16. The most commonly deployed type of hypervisor is the type 1 or bare-metal hypervisor, where virtualization software is installed directly on the hardware where the operating system is normally installed.
17. Because bare-metal hypervisors are isolated from the attack-prone operating system, they are extremely secure. In addition, they generally perform better and more efficiently than hosted hypervisors.



Need of Virtualization

Virtualization technologies have gained renewed interested recently due to the confluence of several phenomena:

- **Increased Performance And Computing Capacity:**

> Nowadays,the average end-user desktop PC is powerful enough to meet almost all the needs of everyday computing, with extra capacity that is rarely used.

>Almost all these PCs have resources enough to host a virtual machine manager and execute a virtual machine with by far acceptable performance.

>The same consideration is to the high-end side of the PC market, where super computers can provide immense compute power that can accommodate the execution of hundreds or thousands of virtual machines.

- **Underutilized Hardware And Software Resources:**

> Hardware and software underutilization is occurring due to
(1) increased performance and computing capacity,
(2) the effect of limited or sporadic use of resources.

> Computers today are so powerful that in most cases only a fraction of their capacity is used by an application or the system.

> Moreover,if we consider the IT infrastructure of an enterprise,many computers are only partially utilized where as they could be used without interruption on a 24/7/365 basis.

> For example, desktop PCs mostly devoted to office automation tasks and used by administrative staff are only used during work hours, remaining completely unused overnight.

> Using these resources for other purpose after hours could improve the efficiency of the IT infrastructure.

>To transparently provide such service, it would be necessary to deploy a completely separate environment, which can be achieved through virtualization.

- **Lack of space:**

> The continuous need for additional capacity, whether storage or compute power, makes data centers grow quickly.

> Companies such as Google and Microsoft expand their infrastructures by building data centers as large as football fields that are able to host thousands of nodes.

> Although this is viable for IT giants, in most cases enterprises cannot afford to build another data center to accommodate additional resources capacity.
> This condition,along with hardware under utilization, has led to the diffusion of a technique called server consolidation,for which virtualization technologies are fundamental.

- **Greening Initiatives:**

> Recently,companies are increasingly looking for ways to reduce the amount of energy they consume and to reduce their carbon foot print.
> Datacenters are one of the major power consumers; they contribute consistently to the impact that a company has on the environment.
> Maintaining a datacenter operation not only involves keeping servers on,but a great deal of energy is also consumed in keeping them cool. Infrastructures for cooling have a significant impact on the carbon foot print of a datacenter.
>Hence,reducing the number of servers through server consolidation will definitely reduce the impact of cooling and power consumption
of a datacenter. Virtualization technologies can provide an efficient way of consolidating servers.

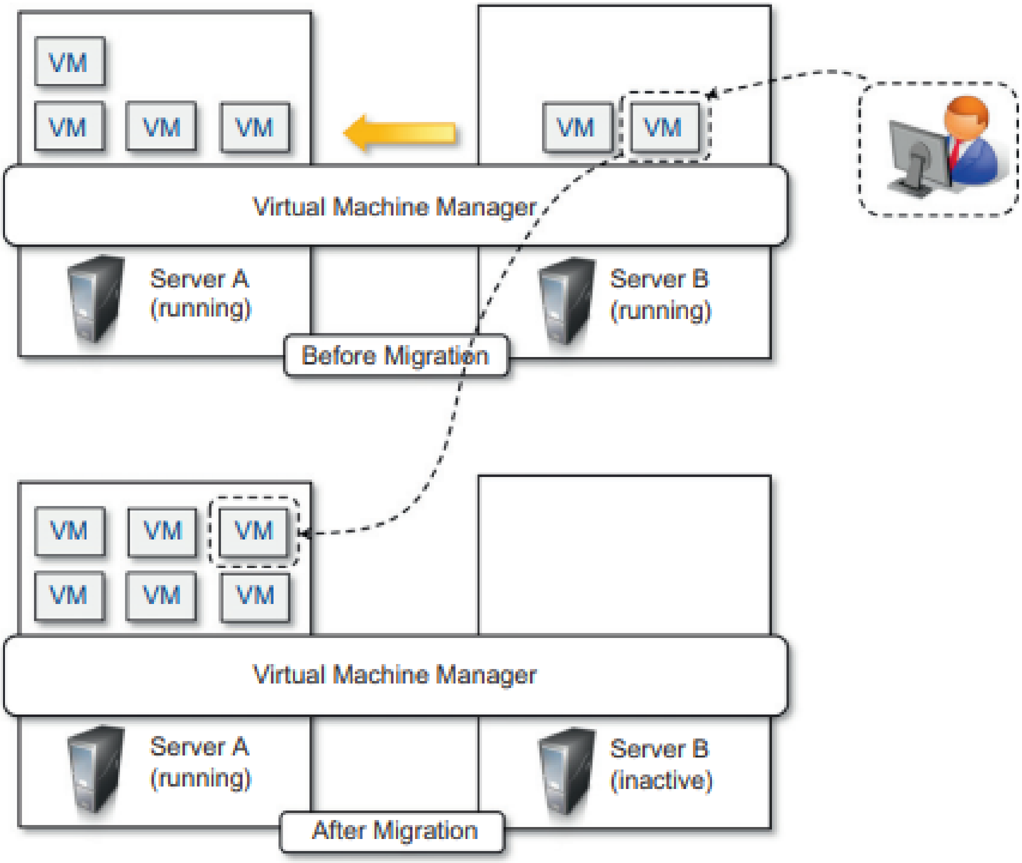
- **Rise of administrative costs:**

> Power consumption and cooling costs have now become higher than the cost of IT equipment.
> Moreover, the increased demand for additional capacity, which translates into more servers in a datacenter, is also responsible for a significant increment in administrative costs.
> Computers—in particular, servers—do not operate all on their own,but they require care and feeding from system administrators.
> Common system administration tasks include hardware monitoring, defective hardware replacement, server set up and updates, server resources monitoring,and backups.
>These are labour-intensive operations, and the higher the number of servers that have to be managed, the higher the administrative costs.
> Virtualization can help reduce the number of required servers for a given workload, thus redrqucing the cost of the administrative personnel.

Virtualization and cloud computing

1. Virtualization plays an important role in cloud computing since it allows for the appropriate degree of customization, security, isolation, and manageability that are fundamental for delivering IT services on demand.
2. Virtualization technologies are primarily used to offer configurable computing environments and storage.
3. Network virtualization is less popular and, in most cases, is a complementary feature, which is naturally needed in build virtual computing systems.
4. Virtual computing environments and execution virtualization techniques play an important role. Among these, hardware and programming language virtualization are the techniques adopted in cloud computing systems.
5. Hardware virtualization is an enabling factor for solutions in the Infrastructure-as-a-Service (IaaS) market segment, while programming language virtualization is a technology leveraged in Platform-as-a-Service (PaaS) offerings.
6. Virtualization also allows isolation and a finer control, thus simplifying the leasing of services and their accountability on the vendor side.
7. Besides being an enabler for computation on demand, virtualization also gives the opportunity to design more efficient computing systems by means of consolidation, which is performed transparently to cloud computing service users.
8. Since virtualization allows us to create isolated and controllable environments, it is possible to serve these environments with the same resource without them interfering with each other.
9. If the underlying resources are capable enough, there will be no evidence of such sharing.
10. This opportunity is particularly attractive when resources are underutilized, because it allows reducing the number of active resources by aggregating virtual machines over a smaller number of resources that become fully utilized.
11. This practice is also known as server consolidation, while the movement of virtual machine instances is called virtual machine migration.
12. Because virtual machine instances are controllable environments, consolidation can be applied with a minimum impact, either by temporarily stopping its execution and moving its data to the new resources or by performing a finer control and moving the instance while it is running.

13. This second technique is known as live migration and in general is more complex to implement but more efficient since there is no disruption of the activity of the virtual machine instance.



Pros and cons of virtualization

Advantages

1. Managed execution and isolation:

These are perhaps the most important advantages of virtualization. In the case of techniques supporting the creation of virtualized execution environments,these two characteristics allow building secure and controllable computing environments. A virtual execution environment can be configured as a sandbox, thus preventing any harmful operation to cross the borders of the virtual host. Moreover, allocation of resources and their partitioning among different guests is simplified, being the virtual host controlled by a program. This enables fine-tuning of resources, which is very important in a server consolidation scenario and is also a requirement for effective quality of service.

2. Portability:

This is another advantage of virtualization, especially for execution virtualization techniques. Virtual machine instances are normally represented by one or more files that can be easily transported with respect to physical systems. Moreover, they also tend to be self-contained since they do not have other dependencies besides the virtual machine manager for their use. Portability and self-containment simplify their administration. Portability and self-containment also contribute to reducing the cost of maintenance,since the number of hosts is expected to be lower than the number of virtual machine instances.

3. Efficient use of resources:

By means of virtualization it is possible to achieve a more efficient use of resources. Multiple systems can securely coexist and share the resources of the underlying host,without interfering with each other. This is a prerequisite for server consolidation, which allows adjusting the number of active physical resources dynamically according to the current load of the system, thus creating the opportunity to save in terms of energy consumption and to be less impacting on the environment.

Disadvantages:

1. Performance degradation

Performance is definitely one of the major concerns in using virtualization technology. Since virtualization interposes an abstraction layer between the guest and the host, the guest can experience increased latencies. For instance, in the case of hardware virtualization, where the intermediate emulates a bare machine on top of which an entire system can be installed, the causes of performance degradation can be traced back to the overhead introduced by the following activities: • Maintaining the status of virtual processors • Support of privileged instructions (trap and simulate privileged instructions) • Support of paging within VM • Console functions

2. Single point of failure

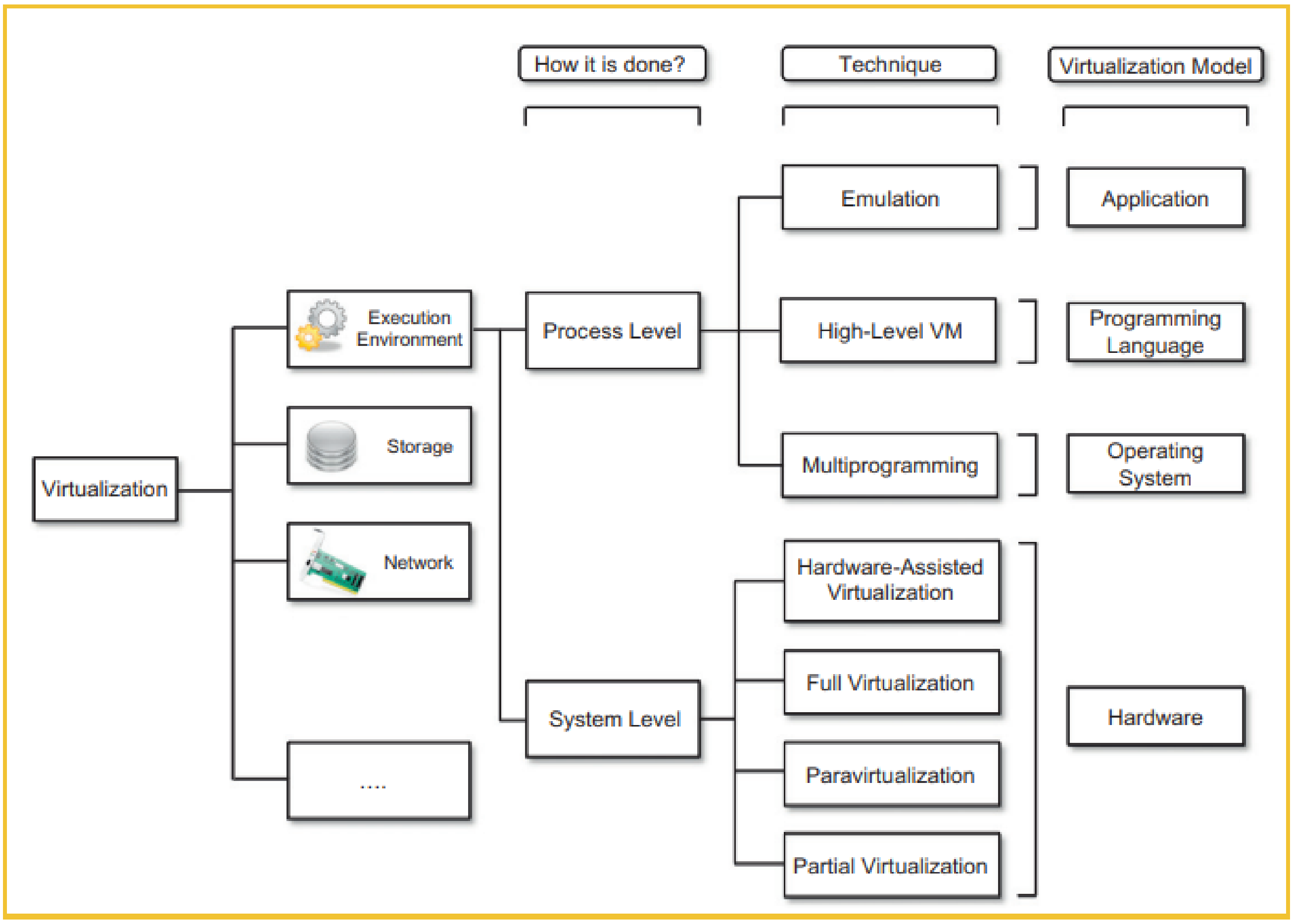
The ability to run many servers on one piece of hardware is one of virtualization's most tangible benefits, but it also creates a single point of failure. If the physical server hosting the virtual servers fails, it results in the loss of a large chunk of data center operations. Single point of failure also applies to the storage system supporting the virtual servers; if several VMs are using the same RAID array and it fails, data might be lost in addition to the interruption of service. Clustering virtual and physical servers might provide enough support to overcome a hardware failure.

3. Security holes and new threats

Virtualization opens the door to a new and unexpected form of phishing.The capability of emulating a host in a completely transparent manner led the way to malicious programs that are designed to extract sensitive information from the guest. In the case of hardware virtualization, malicious programs can preload themselves before the operating system and act as a thin virtual machine manager toward it. The operating system is then controlled and can be manipulated to extract sensitive information of interest to third parties. Examples of these kinds of malware are BluePill and SubVirt. BluePill, malware targeting the AMD processor family, moves the execution of the installed OS within a virtual machine. The original version of SubVirt was developed as a prototype by Microsoft through collaboration with Michigan University. SubVirt infects the guest OS, and when the virtual machine is rebooted, it gains control of the host.

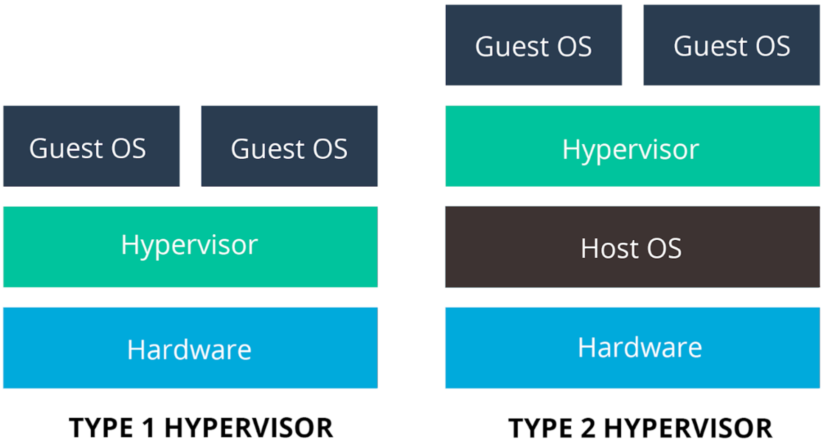
Types of Virtualization

K&U



I. Hardware Level Virtualization

1. Hardware-level virtualization is a virtualization technique that provides an abstract execution environment in terms of computer hardware on top of which a guest operating system can be run.
2. In this model, the guest is represented by the operating system, the host by the physical computer hardware, the virtual machine by its emulation, and the virtual machine manager by the hypervisor.
3. The hypervisor is generally a program or a combination of software and hardware that allows the abstraction of the underlying physical hardware.
4. Hardware-level virtualization is also called system virtualization, since it provides ISA(Instruction Set Architecture) to virtual machines, which is the representation of the hardware interface of a system.
5. A fundamental element of hardware virtualization is the hypervisor, or virtual machine manager (VMM). It recreates a hardware environment in which guest operating systems are installed.
6. There are two major types of hypervisor: Type I and Type II.



- **Type I** hypervisors run directly on top of the hardware. Therefore, they take the place of the operating systems and interact directly with the ISA interface exposed by the underlying hardware, and they emulate this interface in order to allow the management of guest operating systems. This type of hypervisor is also called a native virtual machine since it runs natively on hardware
- **Type II** hypervisors require the support of an operating system to provide virtualization services. This means that they are programs managed by the operating system, which interact with it through the ABI and emulate the ISA of virtual hardware for guest operating systems. This type of hypervisor

is also called a hosted virtual machine since it is hosted within an operating system.

7. Three main modules, dispatcher, allocator, and interpreter, coordinate their activity in order to emulate the underlying hardware.

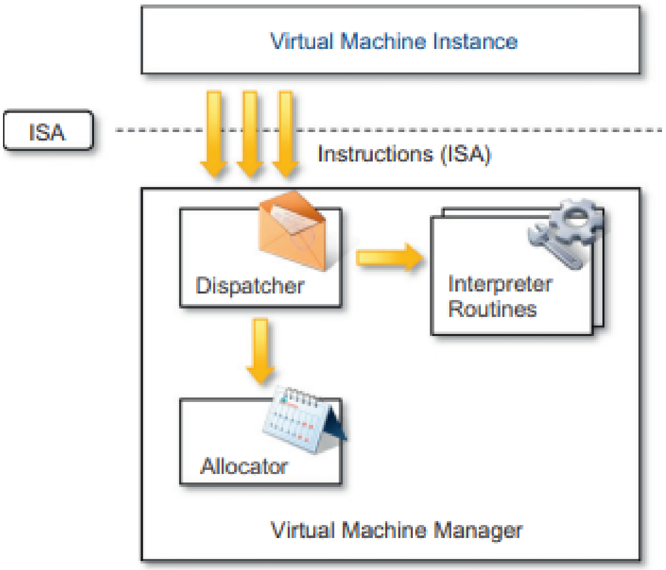


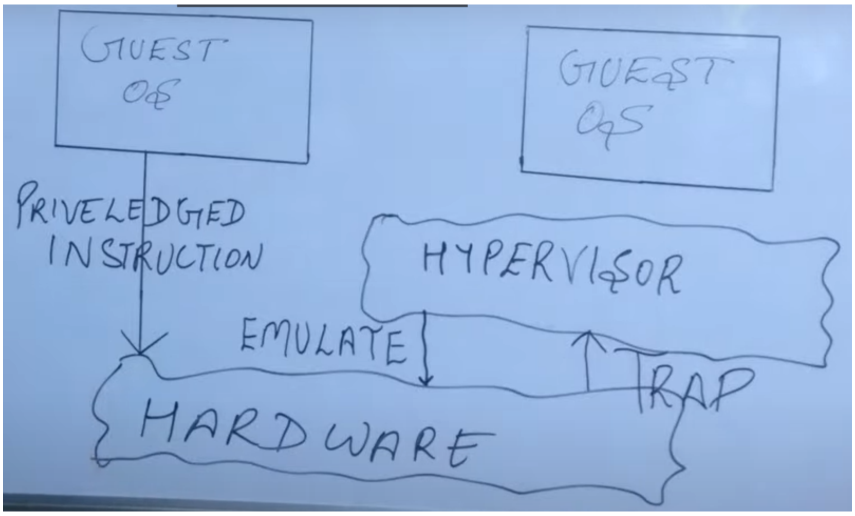
Fig. Hypervisor reference architecture

8. The dispatcher constitutes the entry point of the monitor and reroutes the instructions issued by the virtual machine instance to one of the two other modules.
9. The allocator is responsible for deciding the system resources to be provided to the VM: whenever a virtual machine tries to execute an instruction that results in changing the machine resources associated with that VM, the allocator is invoked by the dispatcher.
10. The interpreter module consists of interpreter routines. These are executed whenever a virtual machine executes a privileged instruction: a trap is triggered and the corresponding routine is executed.
11. Three properties have to be satisfied by a VMM to efficiently support virtualization:
    - Equivalence. A guest running under the control of a virtual machine manager should exhibit the same behavior as when it is executed directly on the physical host.
    - Resource control. The virtual machine manager should be in complete control of virtualized resources.
    - Efficiency. A statistically dominant fraction of the machine instructions should be executed without intervention from the virtual machine manager.

Types of Hardware Virtualization

1. Full virtualization
2. Partial virtualization
3. Para Virtualization

Full virtualization

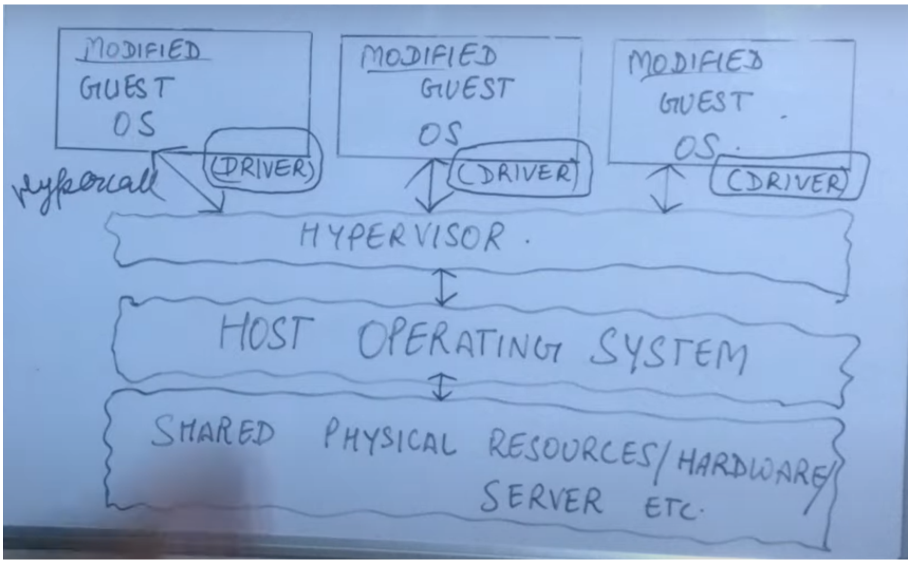


1. Full virtualization refers to the ability to run a program, most likely an operating system, directly on top of a virtual machine and without any modification, as though it were run on the raw hardware.
2. To make this possible, virtual machine managers are required to provide a complete emulation of the entire underlying hardware.
3. The principal advantage of full virtualization is complete isolation, which leads to enhanced security, ease of emulation of different architectures, and coexistence of different systems on the same platform.
4. Whereas it is a desired goal for many virtualization solutions, full virtualization poses important concerns related to performance and technical implementation.
5. A key challenge is the interception of privileged instructions such as I/O instructions: Since they change the state of the resources exposed by the host, they have to be contained within the virtual machine manager.
6. A simple solution to achieve full virtualization is to provide a virtual environment for all the instructions, thus posing some limits on performance.
7. A successful and efficient implementation of full virtualization is obtained with a combination of hardware and software, not allowing potentially harmful instructions to be executed directly on the host. This is what is accomplished through hardware-assisted virtualization.

Partial Virtualization

1. Partial virtualization provides a partial emulation of the underlying hardware, thus not allowing the complete execution of the guest operating system in complete isolation.
2. Partial virtualization allows many applications to run transparently, but not all the features of the operating system can be supported, as happens with full virtualization.
3. An example of partial virtualization is address space virtualization used in time-sharing systems; this allows multiple applications and users to run concurrently in a separate memory space, but they still share the same hardware resources (disk, processor, and network).
4. Address space virtualization is a common feature of contemporary operating systems.

Para Virtualization



1. This is a not-transparent virtualization solution that allows implementing thin virtual machine managers.
2. Paravirtualization techniques expose a software interface to the virtual machine that is slightly modified from the host and, as a consequence, guests need to be modified.
3. The aim of paravirtualization is to provide the capability to demand the execution of performance-critical operations directly on the host, thus preventing performance losses that would otherwise be experienced in managed execution.

4. This allows a simpler implementation of virtual machine managers that have to simply transfer the execution of these operations, which were hard to virtualize, directly to the host.
5. To take advantage of such an opportunity, guest operating systems need to be modified and explicitly ported by remapping the performance-critical operations through the virtual machine software interface. This is possible when the source code of the operating system is available.
6. This technique has been successfully used by Xen for providing virtualization solutions for Linux-based operating systems specifically ported to run on Xen hypervisors.

Desktop virtualization

1. Desktop virtualization abstracts the desktop environment available on a personal computer in order to provide access to it using a client/server approach.
2. Desktop virtualization provides the same outcome of hardware virtualization but serves a different purpose. Similarly to hardware virtualization, desktop virtualization makes accessible a different system as though it were natively installed on the host, but this system is remotely stored on a different host and accessed through a network connection.
3. Moreover, desktop virtualization addresses the problem of making the same desktop environment accessible from everywhere.
4. Although the term desktop virtualization strictly refers to the ability to remotely access a desktop environment, generally the desktop environment is stored in a remote server or a data center that provides a high-availability infrastructure and ensures the accessibility and persistence of the data.
5. In this scenario, an infrastructure supporting hardware virtualization is fundamental to provide access to multiple desktop environments hosted on the same server; a specific desktop environment is stored in a virtual machine image that is loaded and started on demand when a client connects to the desktop environment.
6. This is a typical cloud computing scenario in which the user leverages the virtual infrastructure for performing the daily tasks on his computer. The advantages of desktop virtualization are high availability, persistence, accessibility, and ease of management.

7. The basic services for remotely accessing a desktop environment are implemented in software components such as Windows Remote Services, VNC, and X Server.
8. Infrastructures for desktop virtualization based on cloud computing solutions include Sun Virtual Desktop Infrastructure (VDI), Parallels Virtual Desktop Infrastructure (VDI), Citrix XenDesktop, and others.

Storage virtualization

1. Storage virtualization is the practice of grouping physical storage servers from multiple network storage devices into a single virtual storage device managed from a central console.
2. It is the process of pooling various physical disks and making them appear as a single virtual storage device.
3. It logically integrates storage hardware from data centers, networks and vendors into a single glass panel.
4. Using this technique, users do not have to be worried about the specific location of their data, which can be identified using a logical path.
5. There are different techniques for storage virtualization, one of the most popular being network-based virtualization by means of storage area networks (SANs).
6. SANs use a network-accessible device through a large bandwidth connection to provide storage facilities.
7. Below are some of the main advantages of storage virtualization.

- It is highly scalable.
- It allows easy addition and deletion of storage without affecting any application.
- Easy data migration.
- Easy storage management.

Network virtualization

1. Network Virtualization is a process of logically grouping physical networks and making them operate as single or multiple independent networks called Virtual Networks.
2. There are two kinds of network virtualization: external virtualization and internal virtualization.
3. External network virtualization can combine systems physically attached to the same local area network (LAN) into separate virtual local area networks

(VLANs), or conversely divide separate LANs into the same VLAN. This allows service providers to improve a large network’s efficiency.

4. Internal virtualization uses network-like functionality in software containers on a single network server, enabling VMs to exchange data on a host without using an external network.
5. The result of external network virtualization is generally a virtual LAN (VLAN). A VLAN is an aggregation of hosts that communicate with each other as though they were located under the same broadcast domain.
6. One of the options for implementing internal network virtualization: the virtual machine manager can emulate, and install on the host, an additional network device, together with the driver; or the guest can have a private network only with the guest.