People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
Larbi Ben M'hidi University - Oum El Bouaghi
Faculty of Exact Sciences and Natural and Life Sciences
Department of Mathematics and Computer Science

Cloud and Virtualization Course Handbook

Level: 2nd year Master
Specialty: Distributed Architecture
Proposed by
Dr. ZERTAL Soumia

2025/2026

Preamble

This document is a course support for the subject entitled "**Cloud and Virtualization** ", taught in the Department of Mathematics and Computer Science at the University of Oum El Bouaghi and intended for second year Master's students in computer science, specializing in distributed architecture.

The objective of this course is to provide students with:

- Understanding of the principles of virtualization;
- Manipulating the different concepts of virtualization through practical tools;
- Discovering the concept of Cloud Computing and its applications;
- Knowledge of the most well-known Cloud platforms and the ability to manipulate the services offered by at least one of these platforms.

To achieve the stated objective, we made every effort to approach this work from multiple perspectives and synthesized the most relevant information using a variety of sources (books, lecture notes, articles, websites, etc.), while adhering to the official framework defined by the Ministry of Higher Education and Scientific Research. To make the course more engaging, we added practical sessions designed to facilitate the application of the theoretical concepts covered in the lectures.

Nevertheless, we are aware that this document will remain partial, incomplete, and not exhaustive. For this reason, we strive to ensure continuous updates, with the aim of enriching its content. Therefore, we would be grateful if readers would point out any errors or offer suggestions in this regard.

FEBRUARY 2026

DR. ZERTAL SOUMIA.

MATERIAL DESCRIPTION

ACCORDING TO THE MASTER'S ACADEMIC TRAINING OFFER

Course Title : Cloud Computing and Virtualization

Weekly teaching hours: 3 hours

Credits: 5

Coefficients: 3

Content

1. Virtualization
    - 1.1. Hypervisors
    - 1.2. The different types of virtualization
    - 1.3. The main solutions
    - 1.4. Areas of application
    - 1.5. Storage virtualization
2. Cloud Computing
    - 2.1. The different types of Cloud
    - 2.2. Cloud architectures
    - 2.3. The various players in the Cloud

Bibliography

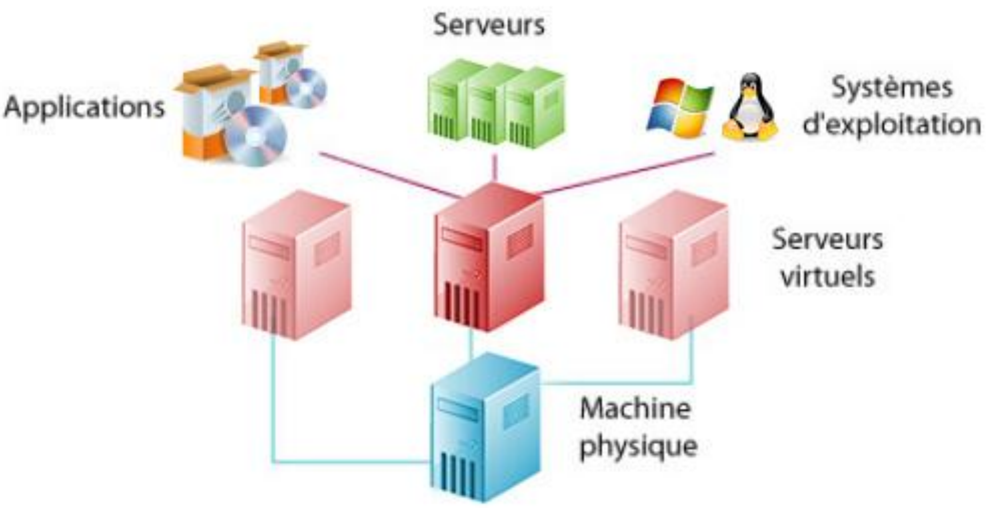
D. Bruley: VMware: VSphere and View. ENI

Linux magazine: Test the Cloud. GNU/Linux magazine 142

Noureddine GRASSA: Virtualization and Cloud Course. ISET Kairouan

Chapter I

VIRTUALIZATION



1. Introduction

Information technology has developed considerably over the last twenty years, to the point of becoming virtually indispensable for each of us and even more so for businesses. Among the new technologies that have a significant impact today are personal computers, mobile phones, tablets, and connected devices. It is once they are connected to the internet that the use of these devices becomes truly worthwhile. The graph below (see Figure 1) shows us that the number of internet users increased very rapidly from the beginning of the 2000s. That year, the number of internet users worldwide reached approximately 3.9 billion.

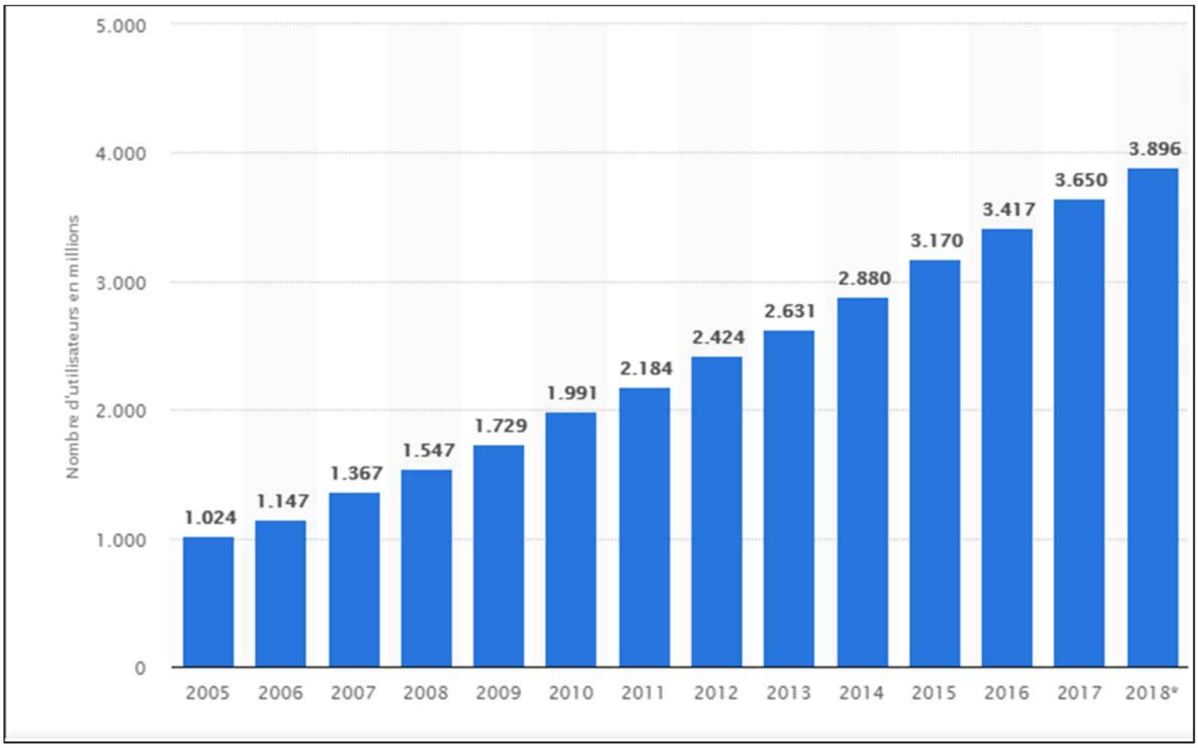


Figure 1: Evolution of the number of Internet users in the world.

(https://fr.statista.com/statistiques/571074/nombre-d-utilisateurs-d-internet-dans-le-monde-2005--/)

Easier access to computers, the modernization of countries worldwide, and the wider use of smartphones have enabled people to use the internet more frequently and conveniently. However, internet penetration often relates to the current state of development of communication networks. As of January 2016, there were approximately 680 million total internet users in China and 282 million total internet users in the United States. However, access to broadband internet varies by region due to infrastructure, the development of internet markets, and mobile connections.

Cisco predicts that global mobile traffic growth (from 2014 to 2019) will outpace global fixed-line traffic growth by a factor of three (see Figure 2). Among the factors expected to drive this trend, Cisco cites the ever-increasing number of mobile users. The equipment manufacturer estimates the number of mobile users at 5.2 billion in 2019, representing more than 69% of the global population (compared to 4.3 billion in 2014, or nearly 59% of the global population).

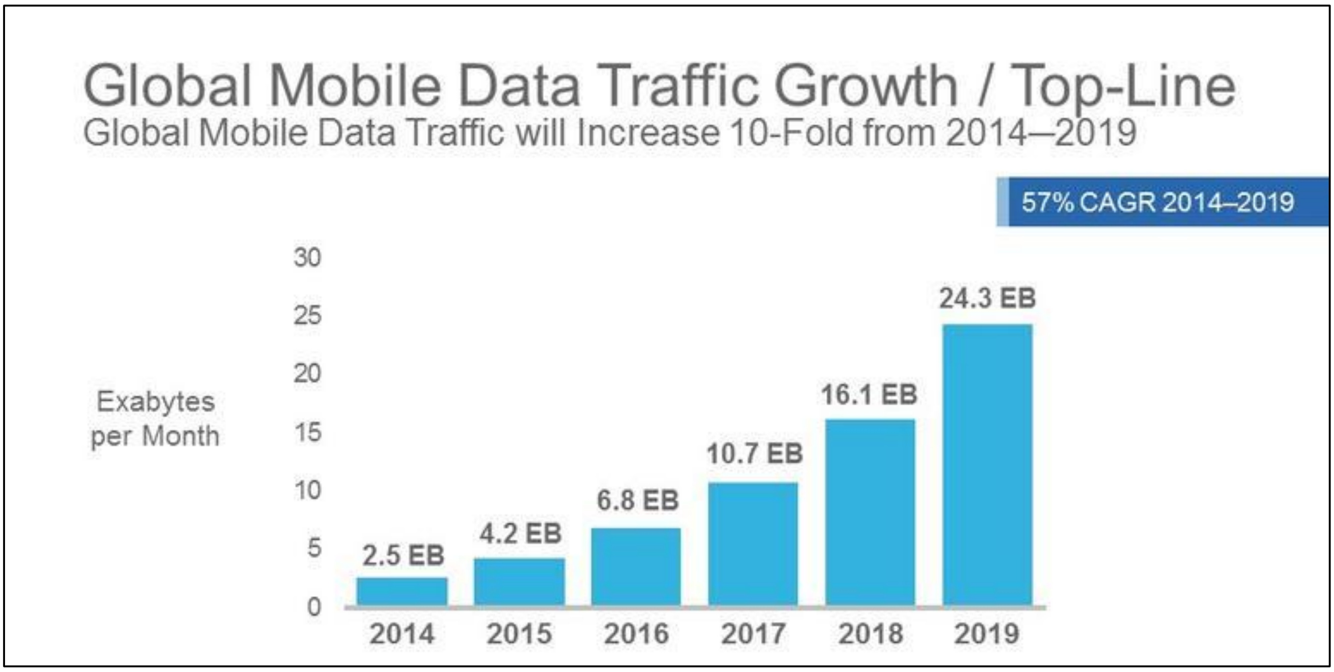


Figure 2: Evolution of data traffic from fixed and mobile devices.

Therefore, solutions had to be found to store the enormous amount of data on the internet and to design new ways to access this data easily. For several years now, virtualization and cloud computing have been among the solutions for storing large quantities of data and making them available to a client regardless of their location or the device they are using.

The rapid evolution of Internet services, social media platforms, big data analytics, and artificial intelligence applications has dramatically increased the demand for scalable computing infrastructures. Traditional physical infrastructures are no longer sufficient to satisfy modern requirements related to elasticity, availability, and resource optimization. Consequently, virtualization technologies emerged as a strategic solution allowing organizations to maximize hardware utilization while reducing operational costs. Today, virtualization represents one of the fundamental pillars of cloud computing infrastructures and modern datacenters.

2. Terminology

2.1.Host System

The host system is the physical machine that hosts the virtual machines.

2.2.Hypervisor

The hypervisor is a virtualization platform that abstracts the hardware layer of the host machine. It also allows the execution of different guest systems on the host system. It manages their operation and provides isolation between guest systems.

2.3.Guest System

The guest system is the instance of the operating system that runs in the virtual environment of the host system. It is also called a virtual machine (VM).

In modern enterprise environments, additional terminology is frequently associated with virtualization technologies. For instance, the term “snapshot” refers to the process of saving the current state of a virtual machine in order to restore it later if necessary. Similarly, “cloning” allows administrators to duplicate a virtual machine rapidly without reinstalling the operating system. These mechanisms simplify backup operations, disaster recovery, and large-scale deployment.

3. Virtualization

Virtualization is used to generate a simulated physical system on top of a real physical system. It allows the use of virtual computing resources from a real physical machine. In most cases, there are multiple simulated physical systems. Virtualization is used to create system density. We can have several virtual systems, called virtual machines, running on a single physical system. These virtual systems share the use of physical resources such as a processor, network interface, or hard drive; these resources are allocated to a virtual machine so that it functions like a physical machine. When a virtual system is not using the resources of a physical system, those resources can be used by another virtual system. In a non-virtualized environment, system resources can be idle for periods of time.

Virtualization also improves hardware abstraction by decoupling software environments from physical infrastructures. This separation enables organizations to migrate workloads between servers without interrupting services. Moreover, virtualization enhances flexibility because virtual machines can easily be created, paused, duplicated, or deleted according to organizational needs. Such flexibility is essential for companies operating dynamic applications requiring high scalability.

Today, for a computer to optimally host multiple virtual machines, it needs sufficiently powerful hardware resources. In short, virtualization involves using information and communication technologies, both hardware and software, to host several different operating systems on a single physical machine.

Virtualization allows adding an abstraction layer that separates the operating system from the hardware in order to deliver better utilization and flexibility of processing resources. Virtualization can operate at different levels of abstraction within the computing architecture. Hardware-level virtualization consists of isolating physical resources (CPU, RAM, storage, network) in order to share them among several independent execution environments. Operating system-level virtualization, on the other hand, relies on the concept of containers, which share the same kernel while isolating user processes. Finally, application-level virtualization allows applications to run in encapsulated environments, independently of the host system. Each of these levels addresses specific needs in terms of performance, isolation, and flexibility.

4. Historical

Virtualization is not a new concept:

4.1.Appeared in the 1960s with CP/CMS

Virtualization dates back to the 1960s, a time when computers were still relatively uncommon and hardware was still very rare in businesses. At that time, virtualization focused on mainframes, powerful computers capable of performing a large number of tasks. Given the immense computing power of these large machines, it was crucial to optimize their use to prevent unnecessary workloads. This is a problem that remains relevant today with modern servers.

To this end, the first company to develop virtualization was IBM®. The company developed the IBM System/360 Model 67, which enabled time-sharing, but this feature was not fully implemented. However, IBM® launched a new system that used virtualization to implement time-sharing and observe the interaction between operating systems and computer hardware. This system, which appeared in 1967, was the CP/CMS (Control Program/Cambridge Monitor System, later renamed Console Monitor System). Its defining characteristic was that its source code was open source, meaning all IBM® customers could access it free of charge. The term CP (Control Program) referred to a control program used to create the environment of a virtual machine. CP allowed a user to manage a virtual machine from a terminal. The CMS (Console Monitor System) was defined as a simple, interactive, single-user operating system. The quality of this CP/CMS system was evident because, at that time, it was already possible to run multiple copies of CMS simultaneously on CP virtual machines without impacting performance. The CP/CMS system evolved considerably during this period and was therefore a significant innovation at the time.

The historical development of virtualization illustrates the continuous evolution of computing paradigms. Initially limited to large mainframe systems because of their high cost, virtualization progressively became accessible to personal computers and enterprise servers due to hardware advancements and processor virtualization extensions such as Intel VT-x and AMD-V. This democratization contributed significantly to the expansion of cloud computing platforms worldwide.

Indeed, this system greatly influenced virtualization as we know it today. The environment was highly virtualized with CP/CMS; each virtual machine had its own virtual devices based on the system's physical hardware type. There were several versions of CP/CMS:

- The first was CP-40/CMS, which enabled the implementation of the CP/CMS virtual machine architecture.
- Subsequently, the CP-67/CMS was developed for the IBM System/360 model 67.
- Finally, the CP-370/CMS was developed for the IBM System/370 but was never released to the market.

4.2.The commercial launch of the VM/370 operating system in 1972

The VM/370 operating system was commercialized in 1972 based on IBM's CP-370/CMS. In the "modern" version of virtualization, the hypervisor has taken over the role of IBM's CP. VMware introduced and popularized the concept of virtualization as we know it today.

4.3.The development of VMware

- During the 1980s and 90s, the x86 architecture emerged, and PCs became widespread among a large number of users. The need for virtualization to optimize machines diminished.
- It was in 1999 that this new company released VMware Workstation™ 1.0, and it succeeded in virtualizing an x86 workstation. This opened the door to more possibilities and revived the desire for IT companies to develop new features to optimize and offer more flexibility.

4.4. The 2000s

- Starting in 2000, IBM®, Dell® and Compaq® joined VMware® as partners to release VMware GSX Server™ 1.0 and ESX Server™ 1.0 in 2001. VMWare® grew rapidly and reached one million users in 2002.
- The VMotion™ product, launched in 2003, is considered a major advance in the field because it allows virtual machines to be transferred from one physical machine to another.
- Intel® and Cisco® became shareholders of the company in 2007, following its initial public offering. Beyond the shared financial interests, this will allow for the establishment of technology-based partnerships in the future.

4.5.Currently

- VMware and Microsoft are the two major players in server virtualization. Today, the most advanced companies have virtualization rates exceeding 75%. This market currently represents approximately €5.5 billion and experienced 6% growth in 2015-2016.
- Virtualization is very well known today. We hear about server virtualization, Virtualbox, bare metal, but also desktop virtualization, VDI, and virtualization in video games with emulators.

5. Why virtualize?

What do companies expect from virtualization?

- Reducing the number of servers.
- Reducing the space occupied in data centers.
- Reducing the energy consumption of data centers.
- Reduction of administrative costs.
- Improved flexibility and speed of services.
- Improvement in the quality of services.

Another important motivation behind virtualization is business continuity. Organizations must ensure that their critical services remain operational even during hardware failures or maintenance periods. Virtualization technologies facilitate disaster recovery strategies by enabling fast backup restoration and live migration of virtual machines between physical hosts. As a result, service downtime is considerably minimized.

6. Advantages of virtualization

For several years now, the concept of virtualization has been developing considerably, becoming a very important element in the field of information technology. Virtualization has gained such prominence today because it offers several significant advantages, particularly for companies with substantial IT resource needs. Indeed, virtualization is primarily beneficial for achieving various types of cost savings (hardware, energy, and financial). Several factors demonstrate why virtualization has become so widespread in recent years.

6.1.Improve deployment

Integrating hardware or software components and migrating to new operating systems can often become sources of problems for system administrators. Indeed, installing and configuring new physical servers and operating systems can be time-consuming, not to mention migrating applications hosted on older servers. Furthermore, in some cases, a single application can prevent a server's operating system from being updated, and consequently, all other applications are also stuck on the current operating system. In the example explained above, virtualization would allow the system that cannot be updated to become a virtual machine hosting the application that only runs on that operating system. Conversely, the other applications could be hosted on a separate virtual machine running the updated operating system.

6.2.Optimize portability

The speed of virtual machine management is one of the advantages of using virtualization. With virtualization, it is possible to move virtual machines from one physical server to another quickly and easily. This is crucial during maintenance or in the event of unexpected downtime of a physical machine. Furthermore, virtual machines adapt easily to different types of physical infrastructure. In the case of virtualization on a computer, virtual machines run and function independently of the type of physical hardware, such as motherboards or other peripherals.

6.3.Use software designed for various environments

By virtualizing, most operating systems can run on other operating systems. This principle allows a user to access applications originally designed for a large number of other operating systems. This is achieved by deploying multiple virtual machines on a single physical computer. Based on this principle, the possibilities for working with a multitude of software programs designed for different operating systems are vast and become accessible to everyone without having to purchase physical hardware dedicated to each system used. To illustrate this, let's take the case of a user who has a computer running Windows®. With virtualization, their computer can run a Linux™ virtual machine and, by extension, the software designed for it.

6.4.Reduce costs

By choosing virtualization for their technical infrastructure, companies are not overlooking the significant cost savings they can potentially achieve. First, the cost reduction is clearly noticeable at the hardware level, because with the use of virtual machines, the need to acquire physical machines logically tends to decrease. Consequently, since virtual machines use the components of the same physical machine and communicate through them, the purchase of network equipment such as routers is also considerably reduced. Furthermore, having fewer physical machines also means saving space in the company's premises as well as significant energy savings, particularly in the areas of cooling and hosting. Thanks to the simplification of deployment, maintenance, backup, and other operations, virtualization allows administrators to reduce the time spent on these tasks in order to participate in other projects.

6.5. Environmental impact reduction

Virtualization contributes significantly to environmental sustainability. By consolidating several virtual machines on fewer physical servers, organizations reduce electricity consumption and cooling requirements inside datacenters. This reduction lowers carbon emissions and supports green computing initiatives. Consequently, virtualization is considered an environmentally responsible technology aligned with sustainable IT strategies.

7. What is the difference between consolidation, rationalization, and concentration?

Since these terms can be used inappropriately, let us redefine each of the following terms: consolidation, rationalization and concentration.

7.1.Consolidation

This involves optimizing server utilization. Running only a single application on servers results in wasted performance because the servers are only operating at 10% of their capacity (or even much less in some cases). Consolidation allows for significantly higher utilization rates. (see Figure 3) .

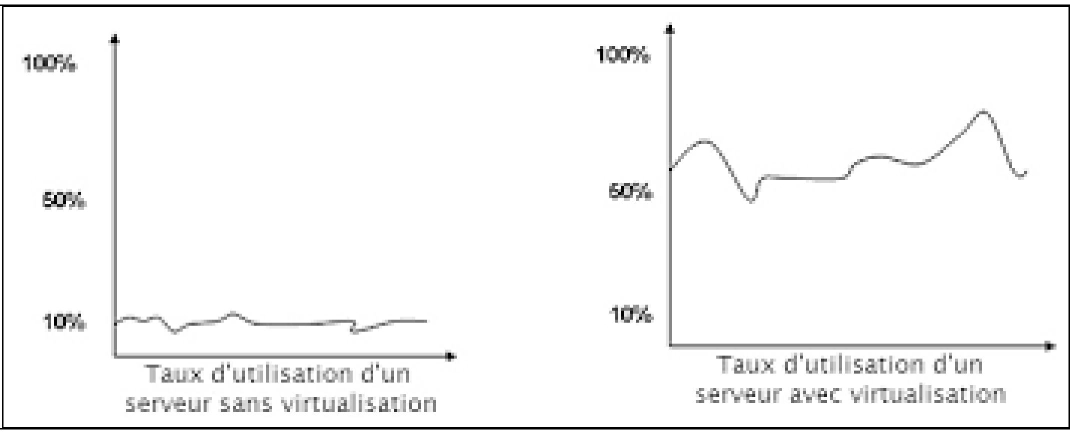


Figure 3: Server Consolidation
(http://yannickprie.net/archives/VEILLE-2009-2012/2012/cloudsec/partie1/01.html)

7.2.Rationalization

This involves removing unnecessary and redundant equipment. The most striking example concerns the various components of a server, such as RAID cards, HBA cards, and hard drives. Streamlining the infrastructure drastically reduces the number of all these hardware components.

Besides the financial aspect and the real savings brought about by rationalization, the advantages also include a significant reduction in the daily management of this equipment, which is often a waste of time.

7.3.Concentration

Server stacking allows you to reduce the footprint and fit more servers into a smaller space. There are several stacking options: Rack, Blade, and Tower. The height of a server is determined by the number of "U"s (1U equals 44.45 millimeters).

Blade servers are the most optimized systems in terms of concentration since today it is possible to host up to 16 Blade servers in a 10U chassis. They also allow for equipment rationalization because there is much less cabling and redundant equipment.

On the other hand, they do not offer consolidation because the utilization rate on a Blade server is the same as on a Tower or Rack server.

LAME servers combined with virtualization offer consolidation, streamlining, and centralization. For some companies hosting a large number of servers (several hundred), combining LAME with virtualization is a fundamental technical choice.

8. Hypervisor

Virtualization is implemented using a hypervisor, which provides the actual virtualization capabilities. It acts as an intermediary between the physical system (host system) and the virtualized system (guest system). Hypervisors require various components to be installed on the host system to ensure virtualization. Furthermore, hypervisors offer several options for guest operating systems.

Hypervisors are responsible not only for resource allocation but also for ensuring isolation and security between virtual machines. Without proper isolation mechanisms, one compromised virtual machine could potentially affect others hosted on the same physical server. Therefore, modern hypervisors integrate advanced security features including memory isolation, secure boot mechanisms, encryption support, and access control policies. The hypervisor has two major roles:

8.1.Security and Isolation in Hypervisors

Security is a fundamental concern in virtualized environments. The hypervisor is responsible for ensuring strict isolation between the virtual machines hosted on the same physical server. A vulnerability in the hypervisor, known as a "VM escape," could allow an attacker to gain access to the host system or other virtual machines, thereby compromising the entire infrastructure. To mitigate these risks, modern hypervisors integrate advanced security features such as integrity verification at boot time (Secure Boot), encryption of virtual machine memory, and granular access control policies. Technologies such as Intel TXT (Trusted Execution Technology) and AMD SEV (Secure Encrypted Virtualization) also contribute to strengthening security at the hardware level.

8.2.Create virtual resources specific to each VM

Initially, it creates virtual resources for each VM (see Figure 4) . The VMs will not directly access the host machine's resources, simply because they are unaware of being VMs hosted on a host. And like any physical machine, a VM will therefore have its own hard drive, memory, processor, and peripherals, the difference being that all of this will be virtual.

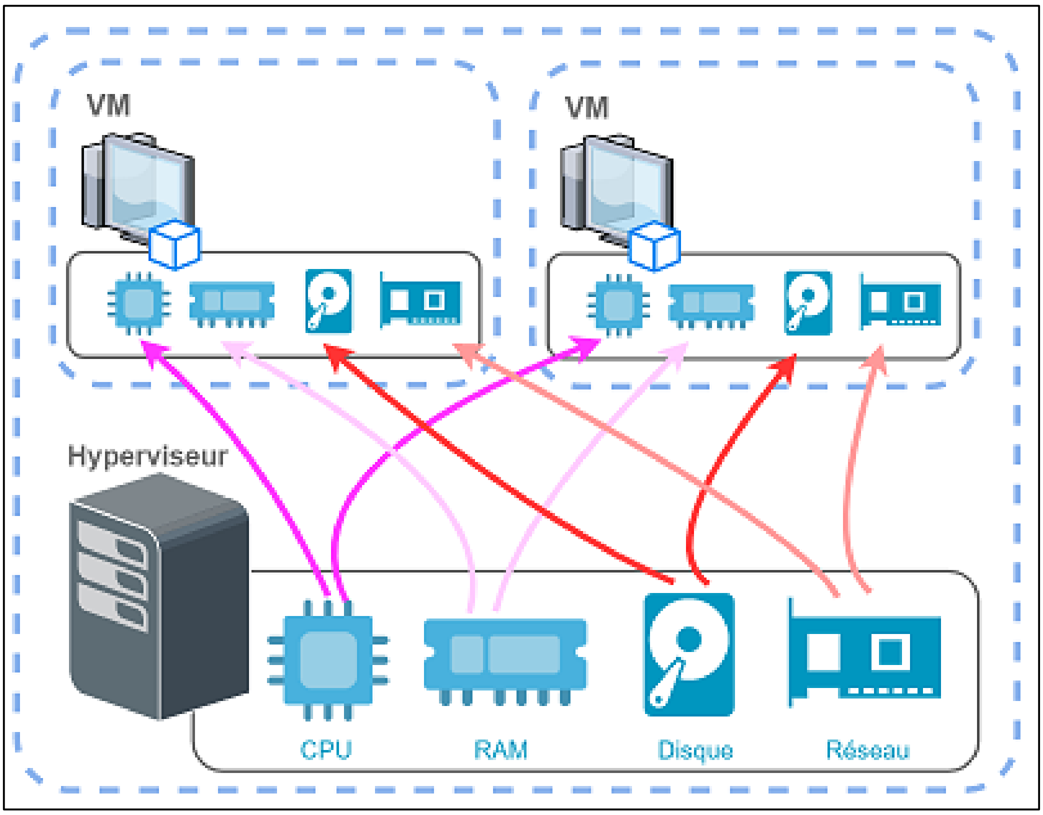


Figure 4: Creation of virtual resources specific to each VM

8.3.Distribute one's own resources

- In a second step, the hypervisor acts as a conductor to allocate to each VM the resources it needs, at the right time and in the right quantities (but within a certain limit that will have been set).

8.4.Types of hypervisors

There are two types of hypervisor:

- *Type 1 hypervisors* : they are particularly useful in **large enterprise network architectures** , which require cost and maintenance optimizations, while improving robustness against failures.
- *Type 2 hypervisors* : These are indeed suitable for small infrastructures. Generally speaking, they are suitable when you have **a single machine** and want to perform cross-platform testing (application, OS, communication, etc.).

These two types of hypervisors correspond to two distinct uses and are therefore not interchangeable.

8.4.1. Type-1 Hypervisor

Type 1 hypervisors run directly on the hardware platform, without an intermediate OS. They manage access for the kernels of guest operating systems to the underlying hardware architecture. As a result, multiple operating systems can be run almost directly on top of the hardware, without depending on a host OS (see Figure 5) .

For example, when installing the **ESXi hypervisor** , it completely erases the operating system and data before replacing them.

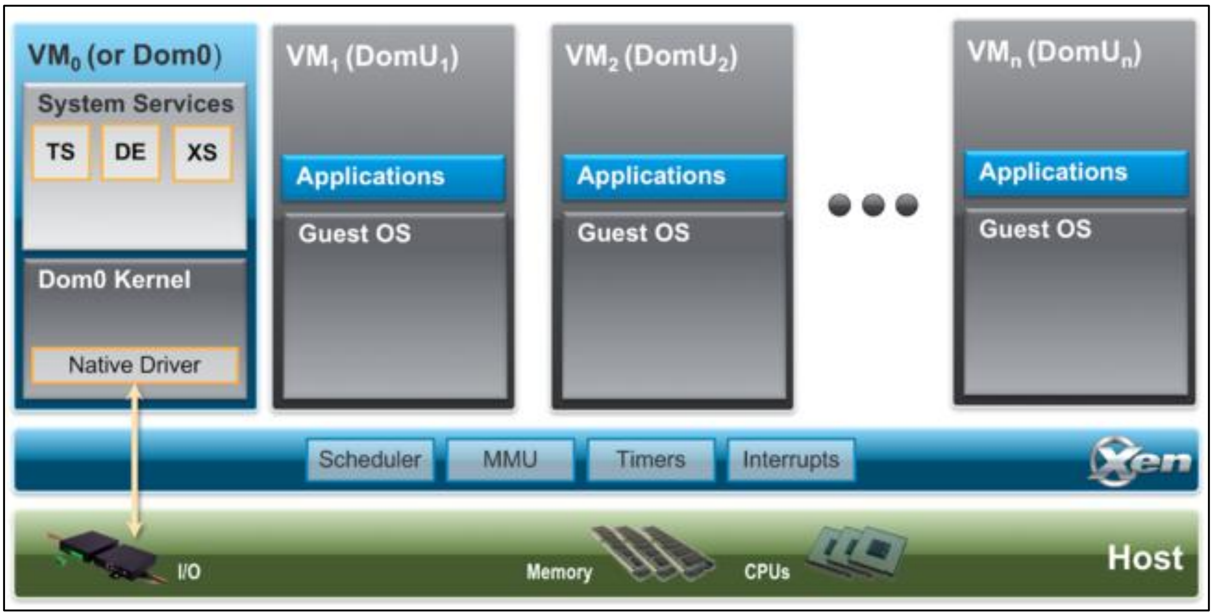


Figure 5: Type 1 hypervisor (the Xen hypervisor)
(https://wiki.xen.org/wiki/Xen_Project_Software_Overview)

This type of hypervisor is used in a completely different context than a type 2 hypervisor, mainly due to its significantly superior performance, made possible because:

- On the one hand, the hypervisor has direct access to resources (without going through an OS);
- On the other hand, all resources are dedicated to VMs.

This type of hypervisor is designed to host between ten and twenty VMs. It therefore requires very robust and high-performance servers, and the cost to use them is often very high.

Type 1 hypervisors are used in enterprise environments for several reasons, such as:

- reduce material and maintenance costs;
- optimize physical resources;
- distribute the load dynamically;
- to enable high server availability;
- create pre-production VMs to test them in a real environment before putting them into production.

This virtualization technique is used by, for example: Xen, Hyper-V, vSphere .

8.4.2. Type-2 Hypervisor

Type 2 hypervisors, or *hosted* hypervisors, are the easiest to set up. They are installed like any other application, running on top of the host operating system. The operating system controls access to the physical hardware. The hypervisor acts as a control system between the host operating system and guest operating systems (see Figure 6) . Once installed, it allows the creation of virtual machines (VMs) that are independent of the host OS.

It is therefore treated like any other application and **has no priority** over host resources.

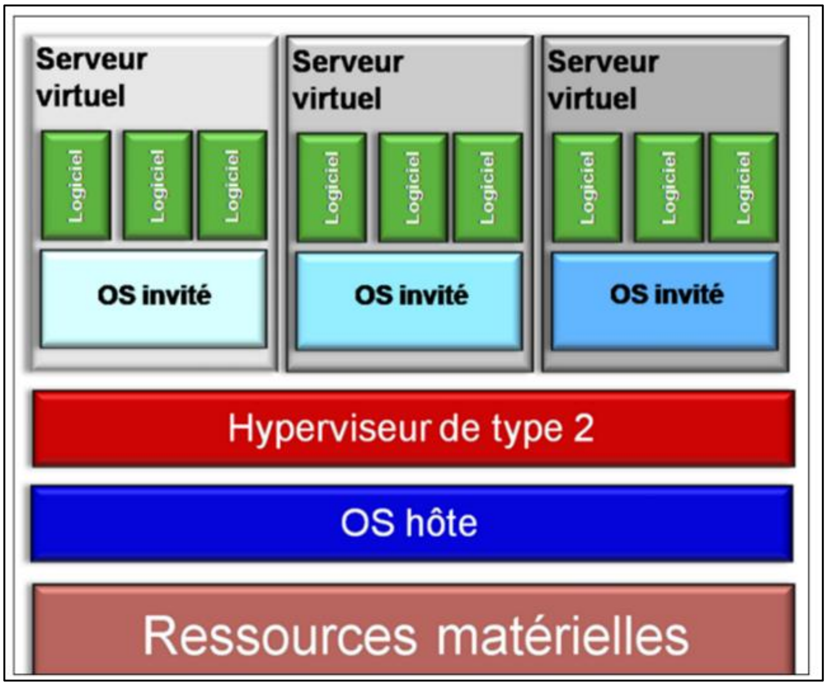


Figure 6: Type-2 Hypervisor

(http://www-igm.univ mlv.fr/~dr/XPOSE2008/virtualisation/definitions.html)

A type 2 hypervisor is an application that is quite memory and CPU intensive. It is necessary to close as many other applications as possible while using it.

The uses of a type 2 hypervisor are numerous. They are fairly easy to implement and very effective in meeting the needs of the following types of systems:

1. test an OS without formatting your physical machine;
2. to test or regularly use an application on a particular OS;
3. simulate a second machine and perform simple communication tests;
4. create a small network of several VMs to test network protocols, firewall rules, configure a monitoring server or other.

Because these hypervisors are installed on a host OS, they have **limited performance** and host VMs which can therefore be **unstable.**

These hypervisors are not designed to create production VMs, but only for **development,** testing **,** and **personal use.**

The main type 2 hypervisors are as follows:

1. **Oracle VirtualBox** ;
2. **VMWare Workstation** (Player and Pro) and VMware Fusion (for Mac).

9. The different types of virtualization

To be able to run different operating systems simultaneously on the same hardware, hypervisors use different virtualization technologies.

The most commonly used technologies are:

- Full Virtualization
- Paravirtualization
- Hardware-level assisted virtualization

9.1.x86 processor architectures

Before understanding how these techniques work, it is necessary to know **the architectures of x86 processors** .

x86 processor architectures offer four levels of privileges (Ring 0 to 3).

Ring 0 = Kernel mode, Ring 3 = user mode (see Figure 7) .

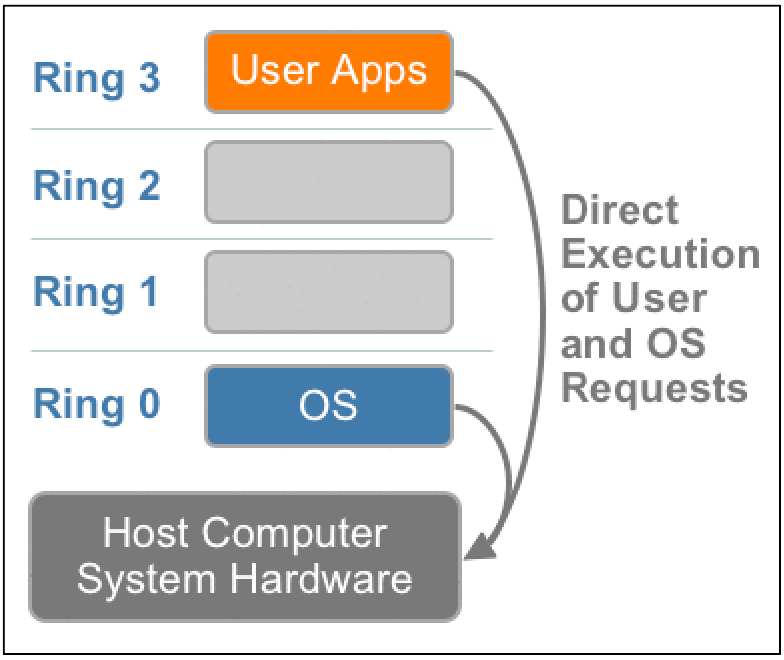


Figure 7: The architecture of x86 processors

Execution levels, or Rings, define the execution privileges of programs. The lower the level a program is installed on, the more control it exerts over the system. The OS has the highest level of control and direct access to resources when running on Ring 0.

Applications run on Ring 3, the highest ring. They cannot modify what runs on rings lower than their own. An application cannot stop the operating system, but the operating system can stop an application.

Rings 1 and 2 define privileges of lesser importance than those of Ring 0.

9.2.Full Virtualization using binary translation

Full virtualization involves emulating a complete hardware environment on each virtual machine (VM). Each VM (guest system) therefore has its own allocation of virtual hardware resources provided by the hypervisor and can run applications on this basis.

VMware developed a technique called **binary translation in 1988 (Binary Translation)** allowing the hypervisor to be placed in Ring 0, moving the OS to a higher level (Ring 1) while guaranteeing them a higher privilege level than the applications (Ring 3) (see Figure 8) .

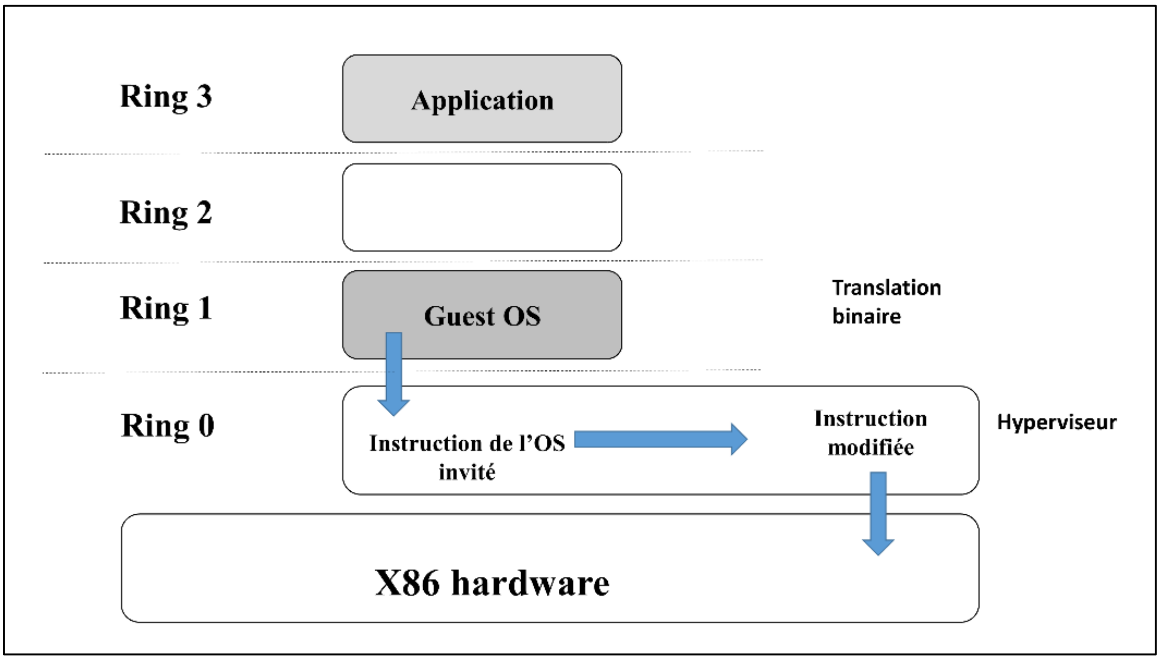


Figure 8: Complete virtualization

Since operating systems are designed to run on Ring 0, they regularly check their positioning, as some instructions only execute if they come from Ring 0. VMware uses binary translation by intercepting certain requests, which allows it to deceive the guest OS about the place it actually occupies on the system.

Binary translation modifies certain instructions coming from the Guest OS before sending them for processing to the physical processors.

The system running within the virtual machine is a fully functional operating system, just like one you would install on a physical machine: Microsoft Windows, GNU/Linux, Mac OS X, etc. This is the key characteristic of full virtualization: guest operating systems do not need to be modified to be used in a virtual machine using virtualization technology. In practice, this is the case for the most widespread operating systems and virtual machines.

The advantage of this technique is that it does not require any modification to the Guest OS kernel because the binary translation is performed at the binary code level by the processor.

Examples:

- Microsoft VirtualPC and Microsoft VirtualServer
- VirtualBox
- VMware Server, VMware Player, VMware Workstation, VMware Fusion

Advantage: The host operating system is unaware of being virtualized, and no modifications to the operating system are necessary. This allows for compatibility with many operating systems.

Disadvantage: binary translation requires additional work from the CPU to perform the binary translation (called **overhead**).

9.3.Para-virtualization

Paravirtualization is another technique developed notably by Citrix's XenServer. It avoids using a complete host system for virtualization. Instead, a very lightweight host operating system kernel is used. The hypervisor simply provides an application programming interface (API) that allows guest operating systems to directly access the physical hardware of the host system (see Figure 9).

It involves modifying the guest OS (the kernel layer) to allow execution outside of Ring 0.

The Guest OS is aware that it is virtualized and modifies certain low-level instructions before sending them to the hardware. Therefore, there is no instruction interception or binary translation.

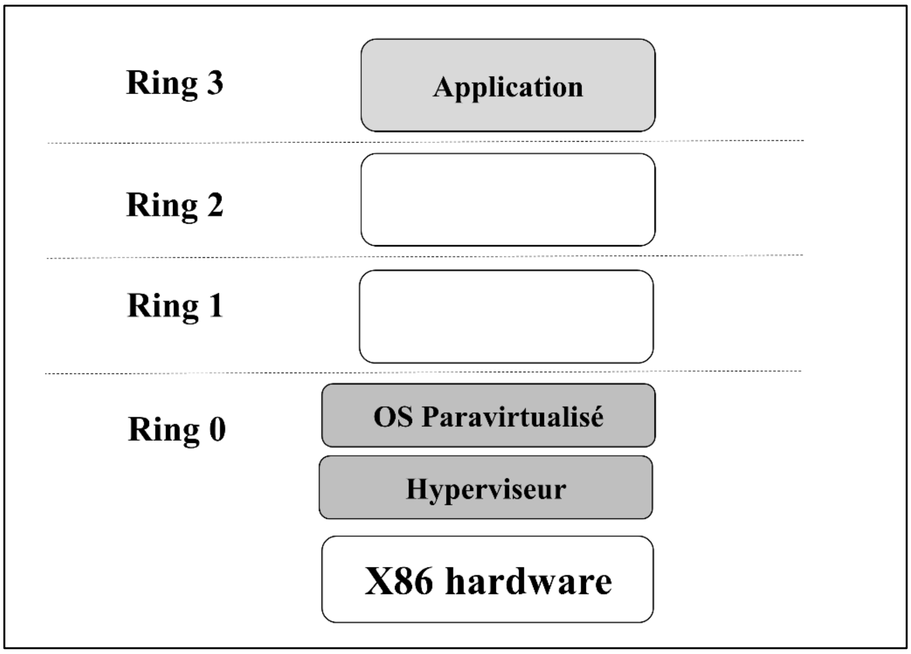


Figure 9: Para-virtualization

Paravirtualization therefore offers a performance advantage over full virtualization. However, this requires that the guest operating system kernel be ported to the API beforehand. Thus, only modified guest systems can be paravirtualized . Proprietary system vendors—such as Microsoft Windows—generally do not allow such modifications .

Among the hypervisors that allow paravirtualizations are Xen and Oracle VM Server for SPARC.

Para-virtualized drivers

VMware introduced certain aspects of paravirtualization techniques through the use of paravirtualization drivers. These specific drivers, developed by VMware, are aware of the virtualization layer and communicate more easily with the hypervisor. They are added via VMware Tools for certain guest operating systems, thus improving performance.

9.4. *Hardware-Assisted Virtualization*

To simplify the task of the hypervisor by avoiding putting the OS on a Ring that is not designed for it or modifying the OS kernel, Intel processors with Intel VT (Virtualization Technology) and AMD-V (Virtualization) introduce a new execution mode, this is called Hardware Assisted Virtualization.

It comprises a root level **,** corresponding to Rings less than 0, and a normal level, corresponding to the older Rings from 0 to 3. This privileged level has direct access to the hardware. This allows it to receive certain instructions directly from the Guest OS to limit the binary translation work (see Figure 10) .

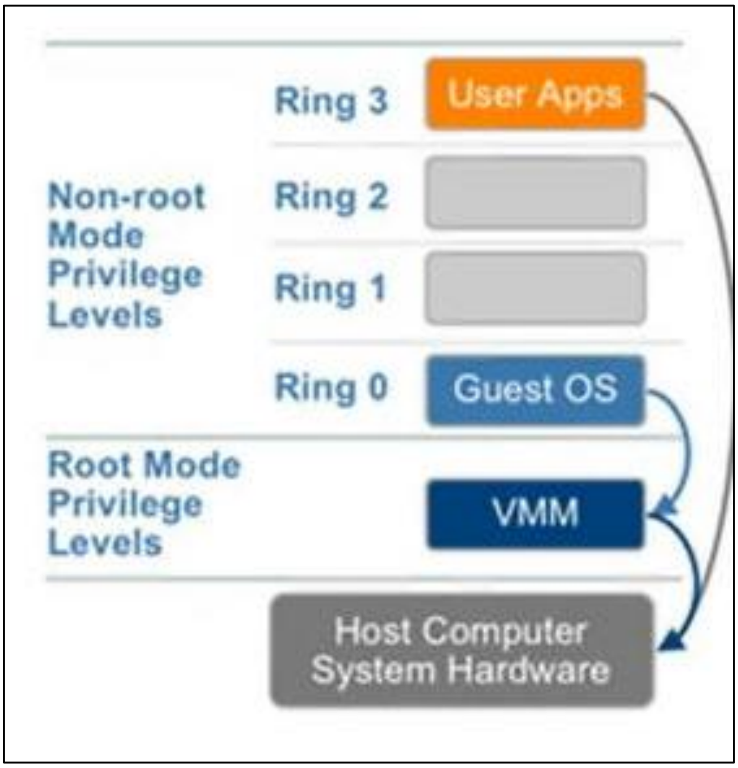


Figure 10: Hardware-assisted virtualization

The hypervisor operates in **Root Mode** with the highest level of control. Guest operating systems run on Ring 0. They occupy the space for which they were designed.

There is no longer any need to modify the Guest OS, nor to use binary translation (although binary translation is still necessary for some instruction sets).

This new root level significantly reduces overhead. This evolution of the instruction set also streamlines the sharing of physical resources between virtual machines.

Thanks to this hardware assistance in the processor, x86 architectures overcome a number of technical barriers.

9.5. Network Virtualization (NFV)

Network virtualization, or NFV (Network Function Virtualization), consists of decoupling network functions, such as firewalls, load balancers, routers, or switches, from the dedicated hardware on which they traditionally run. These functions are then implemented as software running on virtualized standard servers. This approach enables a significant reduction in infrastructure costs, greater agility in the deployment of network services, and improved scalability. NFV is today closely linked to the concept of Software-Defined Networking (SDN), which aims to centralize the network control plane in order to simplify its management and automation. Together, NFV and SDN form the pillars of the modern network infrastructure of datacenters and telecommunications operators.

9.6. Application Virtualization

Application virtualization is a technique that allows applications to run on a client workstation without being directly installed on its operating system. The application runs in an isolated virtual environment, which avoids dependency conflicts between different software versions. Solutions such as Microsoft App-V or VMware ThinApp illustrate this approach. It is particularly advantageous in enterprise environments where many applications must coexist without interference. Furthermore, the centralized management of applications simplifies updates and reduces support costs.

9.7. Desktop Virtualization (VDI)

Desktop Infrastructure Virtualization, or VDI (Virtual Desktop Infrastructure), consists of hosting user workstations on centralized servers in a datacenter. Each user accesses their virtual desktop through a thin client, a web browser, or a dedicated application, from any terminal. This model offers numerous advantages: it simplifies workstation administration, strengthens data security (since data remains within the datacenter), facilitates business continuity, and allows organizations to better control their IT resources. The most widely used solutions include VMware Horizon, Citrix Virtual Apps and Desktops, and Microsoft Remote Desktop Services.

10.The main solutions

The virtualization market has evolved considerably over the last decade. Today, virtualization platforms are integrated into larger cloud ecosystems offering orchestration, monitoring, automation, and security management services. Open-source solutions such as KVM and Xen are increasingly adopted in academic and enterprise environments because they reduce licensing costs while maintaining high performance. Virtualization has gained more importance over the last 15 years since the emergence of VMware® on the market.

10.1. Xen

Xen is free and open-source virtualization software. It is developed by the University of Cambridge in the United Kingdom. Xen allows multiple operating systems (and their applications) to run in isolation on the same physical machine on x86, x86-64, IA-64, and PowerPC platforms (soon to be supported on SPARC). The guest operating systems thus share the host machine's resources. Xen is a "paravirtualizer" or a "Type 1 hypervisor." The guest operating systems are "aware" of the underlying Xen system; they need to be "ported" (adapted) to run on Xen.
Linux, NetBSD, FreeBSD (porting in progress), Plan 9 and GNU Hurd can already run on Xen.

10.2. KVM

KVM (Kernel-based Virtual Machine) is a free virtual machine for Linux. It runs on x86 architectures with Intel VT or AMD SVM (AMD-V) technologies. The module has been integrated into the Linux kernel since version 2.6.20.
Unlike programs like VirtualBox, KVM uses the host's operating system kernel to emulate the physical computer or server. Simply put, KVM is a loadable kernel module for Linux that leverages hardware virtualization technologies like Intel VT and AMD-V. Each virtual machine has its own dedicated CPU, RAM, and network interface.

10.3. vSphere ESXi 6

vSphere ESXi 6 is a hypervisor developed by VMware. Version 6 of **vSphere** was released on February 3, 2015, and represents the latest stable version of **vSphere** . Furthermore, it is one of the most comprehensive hypervisor solutions on the market. **vSphere ESXi 6** is a **Type 1 hypervisor** and allows for the management and virtualization of computers or servers.

vSphere ESXi 6 server can support up to **1024** virtual machines and up to **480** CPUs, **6 TB** of RAM and up to **2048** virtual hard disks per host.

A Virtual Machine can support up to **128** Virtual **CPUs** , **4 TB** of RAM and virtual disks up to **62 TB** .

In addition to creating virtual machines, **vSphere ESXi** offers numerous features that allow for optimal management of different VMs. These features include:

- **vMotion** : vMotion is a feature that enables live migration (without having to shut down the VM) between two **ESXi hosts** . This eliminates service interruptions during maintenance operations. The same principle applies to the storage portion of VMs (Storage vMotion).
- **vSphere HA** : High Availability is a feature that allows for automatic restart of VMs after a failure on the host.
- **vShield Endpoint** : This is an antivirus/antimalware system that secures VMs on the host.
- **Etc** …

These features represent only a small fraction of those offered by **vSphere** . Furthermore, additional features, in the form of Appliances, allow for even more tasks, such as **vSphere Replication** .

vSphere ESXi 6 has three types of editions (**Standard** , **Enterprise** , **Enterprise Plus).**

10.4. Hyper-V

Also known as Windows Server Virtualization, it is a system of virtualization based on a hypervisor 64-bit version of Windows Server 2008. It allows a physical server to become a Hypervisor and thus manage and host virtual machines commonly called VMs (*virtual machines*).

Thanks to this technology, it is possible to run multiple operating systems virtually on the same physical machine and thus isolate these operating systems from each other.

The hypervisor's resources are then shared between different VMs, which is economically advantageous because previously a physical machine had to be considered for each server.

It is possible to use the Hyper-V console on Windows 7. Conversely, many operating systems can run inside Hyper-V:

- Obviously for the operating systems Microsoft Windows 10 (except Home Edition), Windows 8.1, Windows 8, Windows 7 (except Home Edition), Windows Vista SP1/SP2 (except Home edition), Windows Server 2012, Windows Server 2008 x64 SP1/SP2 & R2, Windows Server 2003 x64 SP2 & R2 SP2, Windows 2000 SP4, Windows XP Professional SP2/SP3 & x64 SP2
- For operating systems Linux:
  - SUSE Linux Enterprise Server 10 SP1/SP2 & 11.
  - Red Hat Enterprise Linux 5.2 x64 and later versions.
  - Ubuntu 12.04 LTS and later versions.

10.5. OpenVZ

OpenVZ is an operating system-level virtualization technology based on the Linux kernel. OpenVZ allows a physical server to run multiple isolated operating system instances, known as virtual private servers (VPS) or virtual environments (VE).

Compared to virtual machines like VMware and paravirtualization technologies like Xen, OpenVZ offers less flexibility in operating system choice: both the guest and host operating systems must be Linux-based (although different Linux distributions can be used in different virtual machines). However, OpenVZ's OS-level virtualization offers better performance, scalability, density, dynamic resource management, and ease of administration than its alternatives. According to the OpenVZ website, this virtualization method introduces a very low performance penalty: only 1 to 3% loss compared to a physical computer.

OpenVZ is the basis of Virtuozzo, a proprietary product provided by SWsoft, Inc. OpenVZ is distributed under the GNU General Public License version 2. OpenVZ includes the Linux kernel and a set of user commands.

10.6. LXC

The Linux 2.6.24 kernel integrates fundamental support for containerization to ensure virtualization at the operating system level and allow a single host to run multiple isolated Linux instances, called "Linux containers", or LXC (Linux Containers).

LXC is based on the concept of Linux control groups, or cgroups. Here, each control group offers applications complete isolation of resources (including CPU, memory, and I/O access), without resorting to full-fledged virtual machines.

Linux containers also offer complete isolation of their namespace. Functions such as file systems, network IDs, and user IDs, as well as any other elements typically associated with operating systems, can therefore be considered "unique" from the perspective of each container.

10.7. Docker

Docker is an open-source platform launched in 2013 that automates the building, shipping, and running of applications inside lightweight, portable containers. It was created by Solomon Hykes at the company dotCloud and rapidly became the most widely adopted containerization technology in the software industry. Docker introduced a simple and standardized way to package an application along with all its dependencies — libraries, configuration files, and runtime environment — into a single portable unit called a container image.

Unlike traditional virtual machines, Docker containers share the host operating system's kernel, making them significantly faster to start and far more efficient in their use of hardware resources. Docker relies on a layered image system in which each image is built from a base layer to which additional layers are progressively added, each representing a step in the build process such as installing a dependency or copying application files. This layered approach enables efficient reuse of common layers across multiple images, reducing both build time and storage consumption.

Docker images are stored and distributed through registries, the most prominent of which is Docker Hub, a public repository hosting millions of community-maintained and officially verified images. The Docker Engine, the core component of the Docker platform, runs as a background service on the host machine and exposes a command-line interface (CLI) and a REST API through which containers can be created, started, stopped, and deleted. Docker also provides Docker Compose, a tool that allows developers to define and manage multi-container applications using a declarative YAML configuration file. Docker integrates natively with orchestration platforms such as Kubernetes and Docker Swarm, enabling the large-scale deployment and management of containerized workloads in production environments.

11.Virtualization vs. Containers

Virtualization uses a hypervisor to emulate hardware, allowing multiple operating systems to run in parallel. When resources and features are limited, applications need to be lightweight and deployable in a dense manner. Linux containers run natively on their operating system, which they share. This keeps our applications and services lightweight and allows them to run quickly in parallel.

Containers have become extremely popular in DevOps and cloud-native environments because of their lightweight architecture and rapid deployment capabilities. Unlike traditional virtual machines, containers start within seconds and consume fewer hardware resources. However, containers provide less isolation than virtual machines since they share the same operating system kernel. For this reason, virtual machines remain preferable for highly sensitive or isolated workloads.

Containers have become extremely popular in DevOps and cloud-native environments because of their lightweight architecture and rapid deployment capabilities. Unlike traditional virtual machines, containers start within seconds and consume fewer hardware resources. However, containers provide less isolation than virtual machines since they share the same operating system kernel. For this reason, virtual machines remain preferable for highly sensitive or isolated workloads.

Linux containers represent a new evolution in how we develop, deploy, and manage applications. Linux container images ensure application portability and version control. This guarantees developers that what works on their laptop will also work in the production environment. A Linux container uses fewer resources than a virtual machine. It offers a standard interface (startup, shutdown, environment variables, etc.), provides application isolation, and can be more easily managed as a module of a larger application (multiple containers) (see Figure 10).

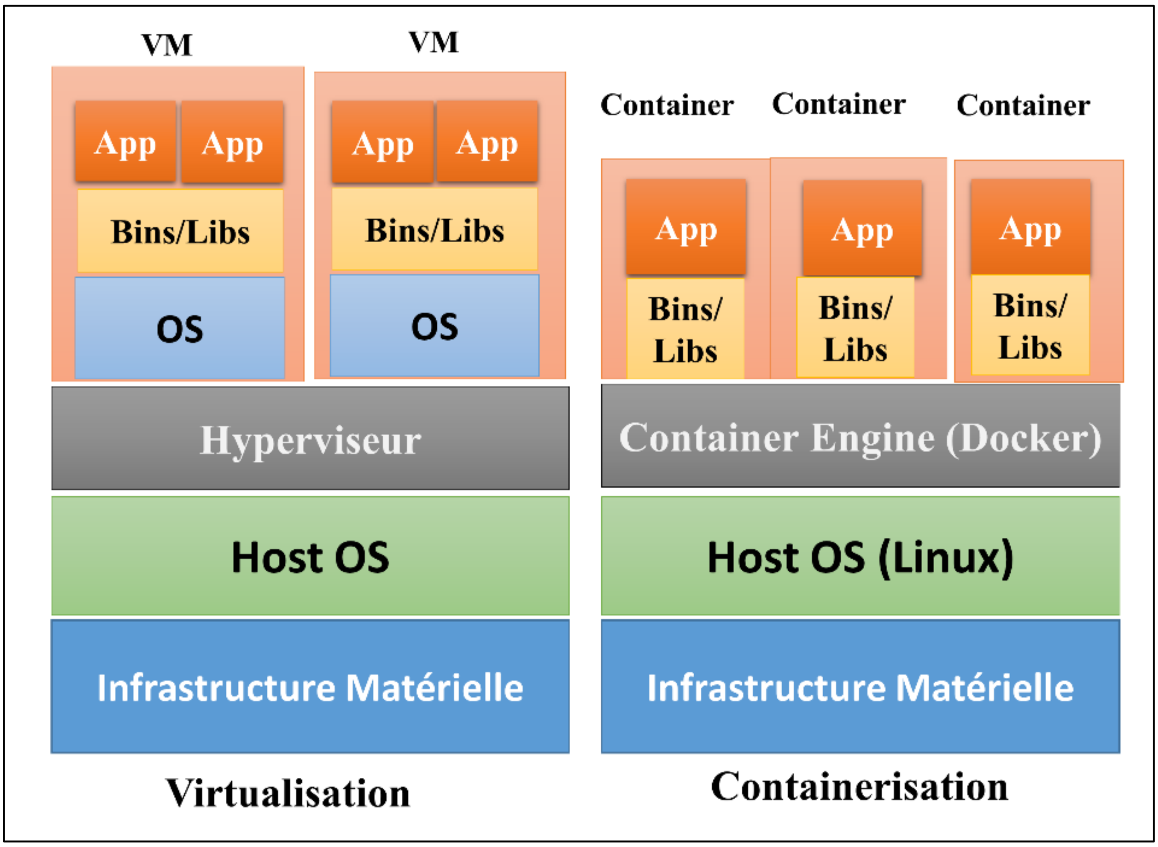


Figure 10: Virtualization vs. Container

11.1. Technical Comparison Between Virtual Machines and Containers

From a technical standpoint, virtual machines and containers differ in their level of abstraction and their footprint on system resources. A virtual machine includes a complete operating system, including the kernel and all system libraries, making it heavier but also more isolated. A container, on the other hand, shares the host system's kernel and only encapsulates the libraries and dependencies specific to the application. This results in much shorter startup times (a few seconds for a container versus several minutes for a VM), reduced memory consumption, and a higher hosting density. However, this proximity to the host kernel implies a lower level of isolation than virtual machines, which can raise security concerns in certain contexts.

11.2. Container Orchestration with Kubernetes

Faced with the proliferation of containers in production environments, managing these entities manually quickly became unmanageable. It is in this context that Kubernetes emerged, an open-source container orchestration system originally developed by Google and now managed by the Cloud Native Computing Foundation (CNCF). Kubernetes automates the deployment, scaling, and management of containerized applications. It provides essential features such as load balancing, automatic service discovery, fault management, and continuous application updates without service interruption. Kubernetes has today become the de facto standard for container orchestration in enterprise environments, and integrates natively with all major public cloud providers.

12.Storage virtualization

With current computing practices, data storage can be problematic. Data volumes are constantly growing, applications demand higher performance, and new hardware is difficult to integrate. Typical storage environments use multiple devices from different manufacturers, but they cannot communicate with each other. This makes centralized management difficult, resulting in siloed operations that quickly become obsolete within a few years. One of the most effective solutions to this problem is storage virtualization.

Storage virtualization also simplifies backup and disaster recovery operations. Administrators can replicate virtual storage volumes between geographically distant datacenters in real time. In case of hardware failure or natural disaster, data can rapidly be restored from replicated storage systems, ensuring high availability and business continuity.

12.1. Definition of storage virtualization

Storage virtualization is a concept that aims to virtually map the various storage resources of a company such as hard drives, flash memory or tape drives and make them available as **a linked storage pool** (data store).

A virtualization solution thus establishes a virtualization layer between existing application servers and storage hardware so that applications no longer need to know on which disks, partitions or storage subsystems their data reside.

12.2. Advantages of storage virtualization

All storage equipment can now coexist and work together:

- Availability also increases with storage virtualization because applications are not limited to specific storage resources and are therefore isolated from most interruptions. If a host fails, a third-party system can detect the failure and transfer all VMs running on the failed host to a working one (High Availability: HA).
- In addition, storage virtualization allows data to be automatically migrated to another virtual machine that continuously ensures the performance required for applications (Vmotion).

- Furthermore, storage virtualization typically automates storage capacity expansion, reducing the need for manual provisioning. Storage resources can be updated on the fly without impacting application performance, thus minimizing downtime.

The difference between Vmotion and High Availability (HA) is that Vmotion is launched by the administrator while HA is launched by the system automatically as soon as a host fails (it must be enabled).

Thus, administrators can identify, provision and manage distributed storage as if it were a single consolidated resource, thereby eliminating wasted capacity.

In an enterprise context, storage virtualization is typically implemented using blocks. Data is divided into equal-sized blocks, and each block is assigned a unique address. This address is stored by the virtualization software (hypervisor) in the central mapping table. The mapping table contains all the metadata necessary to locate the physical position of a data block. This mapping allows virtual data to be managed independently of the physical storage controller, for example, to move, copy, mirror, or replicate it.

12.3. Storage architectures

12.3.1. Direct Attached Storage (DAS)

Direct Attached Storage (DAS) refers to a computer storage system that is directly connected to the server or computer instead of being transmitted over a network. Virtual machines are stored locally and are only accessible by the server hosting the VMs. For example, a hard drive on a computer or server is the usual form of DAS storage, as are groups of external disks connected directly to the server via a SCSI (Small Computer System Interface), SATA (Serial Advanced Technology Attachment), or Serial Attached SCSI (SAS) interface (see Figure 11).

In this type of storage:

- Server storage space is non-dynamic.
- Utilization rate is not optimal.
- If a server fails, no other server will be able to access the VMs. Therefore, you must wait until the server is operational again before you can restart the VMs.

12.3.2. ***Software-Defined Storage (SDS)***

Software-Defined Storage (SDS) is an evolution of storage virtualization that fully decouples storage management software from the underlying physical hardware. Unlike traditional storage systems where software and hardware are tightly coupled, SDS allows organizations to manage heterogeneous storage resources through a unified software layer. This approach provides greater flexibility, simplified management, and the ability to use commodity hardware instead of expensive proprietary storage arrays. Solutions such as Ceph, VMware vSAN, and GlusterFS are representative examples of SDS implementations widely deployed in modern datacenters and cloud environments.

Advantage :

- This is the most economical solution, and it's also very simple and quick to implement; no advanced storage skills or expertise are required. This architecture is feasible in a production environment, but a suitable backup solution must be implemented. It can be the ideal solution for small and medium-sized businesses.

Inconvenience :

Service levels are lower. Advanced features such as VMotion, HA, etc., are not available. There is no centralized VM management.

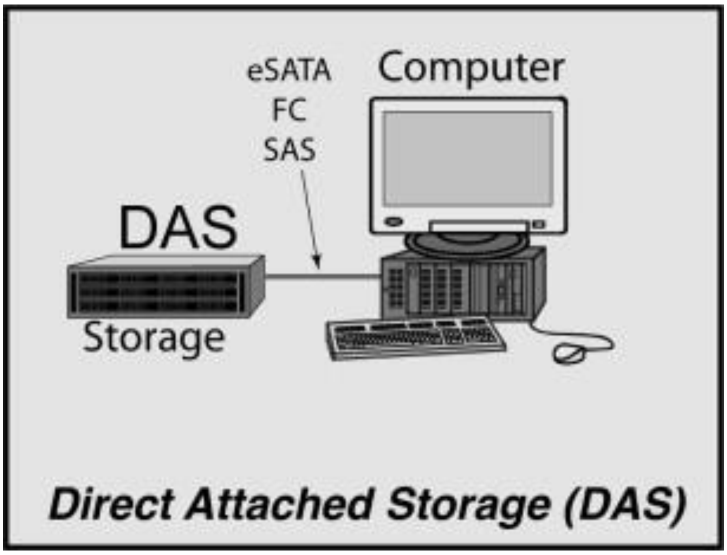


Figure 11: DAS storage architecture

12.3.3. SAN (Storage Area Network)

A storage area network (SAN) is a dedicated network that allows storage resources to be shared. In this architecture, virtual machines are stored on the storage array and are accessible by multiple servers. If a server fails, another server connected to the storage can access the VMs and quickly and automatically bring them back online without having to wait for the failed server to become operational again. This minimizes service interruptions. This architecture also allows for the use of advanced features such as VMotion.

Furthermore, architectures of this type offer excellent performance, a high level of security with the implementation of redundancy for VM access, and thanks to the integrated features of storage arrays such as mirroring and data replication.

It differs from other storage systems by its block-mode access to disks. In the case of a SAN, storage arrays do not appear as shared volumes on the network; rather, they are directly accessible in block mode by the server's file system, introducing a layer of abstraction between the server and the storage system, thus providing administrators with greater flexibility. In short, each server sees the disk space of a SAN array to which it has access as its own hard drive (see Figure 12) .

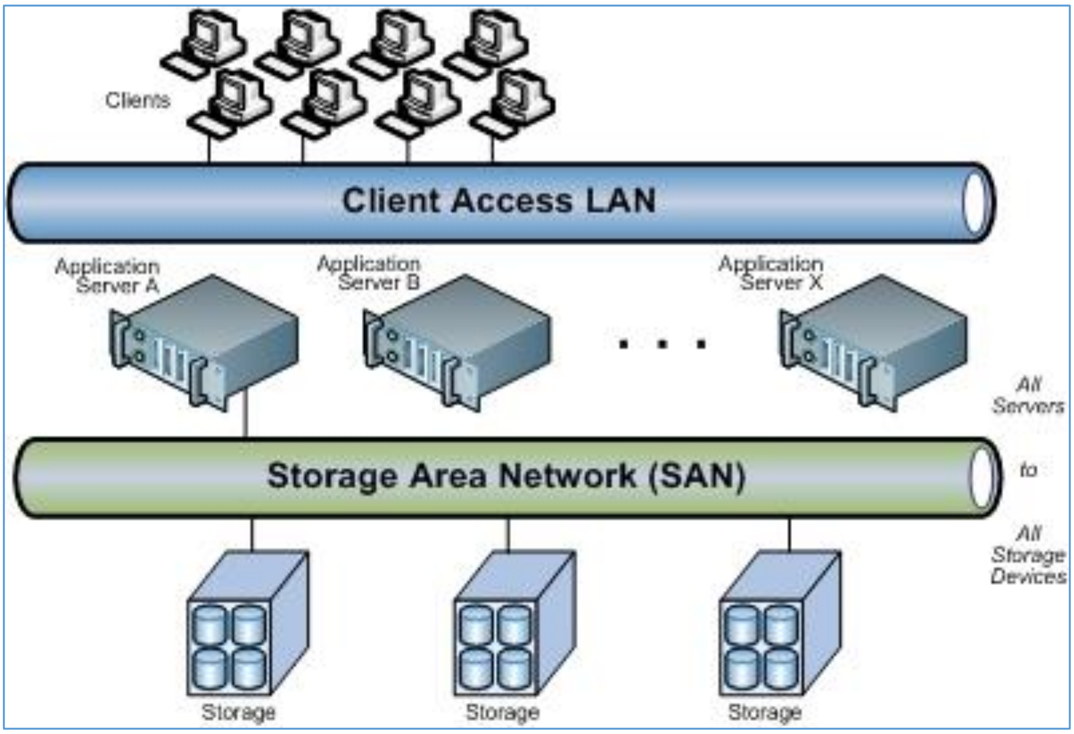


Figure 12: SAN storage architecture

Protocols are used for data transit on the SAN network. The most common are Fibre Channel and iSCSI.

Advantage :

- One of the advantages of a SAN is the scalability of storage space. Disk space is no longer limited by the specifications of the servers and can be expanded as needed by adding disks or storage arrays to the SAN. Shared physical storage space for the servers thus optimizes disk management.

- Data rates on a SAN using Fibre Channel technology can reach 32Gbit/s (previously 1Gbit/s, 2, 4, 8 and 16) per fibre optic link, and ensure that the request sent by a server has been received and taken into account by the storage systems.
- Another advantage of a SAN is its ability to provide storage redundancy, meaning that the storage system remains accessible even if one component fails, by duplicating at least each component of the system (high availability). Furthermore, it can operate in a completely heterogeneous environment, making the server operating system irrelevant.

Inconvenience :

The cost is higher than local storage and requires storage expertise. This necessitates more upfront work.

12.4. Storage management

1. *The Datastore*

The storage space is viewed as a **datastore** . The datastore is a virtual representation of storage resources where virtual machines are stored. It hides the complexity of the various storage technologies and solutions on the market by offering the server a uniform model regardless of the storage infrastructure in place.

2. *Access to data via the virtual machine*

- When a VM communicates with its virtual disk (vmdk), it sends SCSI commands.
- The Guest OS driver will communicate with its SCSI controller (vSCSI).
- The SCSI virtual controller transmits commands to the VMkernel
- VMkernel locates the file on the Datastore that corresponds to the virtual disk vmdk as well as the location of the blocks to be modified.

12.5. SAN storage protocols

12.5.1. iSCSI (Internet Small Computer System Interface)

iSCSI is an IP-based network storage protocol designed to connect data storage facilities. iSCSI encapsulates SCSI commands within TCP packets and then transmits them over the Ethernet network to storage devices. This allows clients (called initiators) to send SCSI commands (CDBs) to SCSI storage devices (targets) on remote servers. It is a protocol used in SANs (Single Area Networks) that aggregates storage resources in a data center while giving the illusion of local storage. Unlike Fibre Channel, which requires dedicated hardware infrastructure, iSCSI can be used while maintaining existing infrastructure.

Advantage :

iSCSI has been widely adopted in many industries because it uses the company's existing TCP/IP network. Therefore, there's no need to set up a dedicated storage architecture with FC switches, specific cards, and cables. For this reason, in certain environments, this solution is ideal because it's much simpler to implement and doesn't require advanced storage expertise. Furthermore, this protocol offers very good performance.

Inconvenience :

Tests have proven that this protocol uses the most CPU resources (up to 60% more CPU time than the FC process in the case of software iSCI).

12.5.2. Fibre Channel

Fibre Channel is another storage protocol that corresponds to the underlying transport layer used by SANs to transmit data. It's the language used by the Host Bus Adapters (HBAs), switches, and controllers within a storage array in a SAN to communicate with each other. Fibre Channel is a low-level language, meaning it's used simply as a communication language between the hardware itself, not the applications running on it. It's the protocol most commonly implemented in production environments.

Advantage :

It seems to be accepted today (in 2009) that the Fibre Channel protocol is the most efficient protocol in I/O and it is the one that consumes the least CPU resources of the host server compared to the iSCSI protocol.

Inconvenience:

FC is the most expensive solution because it requires building the storage architecture and investing in HBA cards, FC switches, and cables, most often fiber optic. Furthermore, this solution is more complex to implement and requires advanced storage expertise.

13. Virtualization and Artificial Intelligence

The convergence of virtualization technologies and artificial intelligence represents one of the most significant developments in modern computing. AI workloads, particularly those involving deep learning model training, require massive computational resources — typically GPU clusters — that are both expensive and energy-intensive. Virtualization addresses this challenge by enabling GPU virtualization, a technique that allows multiple virtual machines or containers to share a single physical GPU. Technologies such as NVIDIA vGPU and AMD MxGPU partition a physical graphics processor into multiple virtual GPU instances, each allocated to a separate virtual machine. This approach enables organizations to maximize GPU utilization while reducing hardware costs. Furthermore, AI-driven platforms are increasingly deployed in cloud environments, where virtualized infrastructure provides the elasticity required to scale training jobs on demand and release resources once the computation is complete. Conversely, artificial intelligence is also being applied to improve virtualization management itself: modern hypervisors and cloud orchestrators increasingly rely on machine learning algorithms to optimize resource scheduling, predict workload peaks, detect anomalies, and automate infrastructure scaling decisions.

14. Virtualization in Practice: Performance and Monitoring

14.1. Performance Metrics in Virtualized Environments

Monitoring and optimizing performance in virtualized environments is a critical responsibility for system administrators. Unlike physical environments where resources are dedicated to a single workload, virtualized environments involve multiple virtual machines competing for shared physical resources. This competition can lead to performance degradation if not carefully managed. The key performance metrics to monitor in a virtualized infrastructure include CPU utilization, memory consumption, disk I/O throughput, and network bandwidth. CPU ready time, the amount of time a virtual machine is waiting for physical CPU access because all available CPUs are busy, is a particularly important indicator of resource contention. A high CPU ready time typically indicates that the physical host is over-committed and that additional physical processors or a reduction in the number of hosted virtual machines is required.

13.2 Memory Management Techniques

Hypervisors employ several sophisticated memory management techniques to maximize the efficiency of physical RAM usage across multiple virtual machines. Memory ballooning is a mechanism by which the hypervisor reclaims memory from virtual machines that are not actively using it and reallocates it to those with higher demand. This is achieved through a special driver installed inside the guest operating system that inflates or deflates a "balloon" of memory within the VM. Transparent Page Sharing (TPS), also known as memory deduplication, allows the hypervisor to identify identical memory pages across different virtual machines and maintain only a single physical copy, reducing overall memory consumption. Memory compression is another technique used when physical memory is under pressure:

instead of immediately writing memory pages to disk (a slow operation), the hypervisor compresses them and stores them in a compressed cache, which is faster to access than disk-based swap space.

13.3 Live Migration and its Implications

Live migration refers to the process of moving a running virtual machine from one physical host to another without any interruption to the services it provides. This capability is essential for load balancing, planned hardware maintenance, and energy optimization. During a live migration, the hypervisor copies the virtual machine's memory pages from the source host to the destination host while the VM continues to run. Once the majority of memory has been transferred, the hypervisor briefly pauses the VM to synchronize the remaining dirty memory pages, resumes execution on the destination host, and updates the network to redirect traffic to the new location. In VMware environments, this feature is implemented through vMotion, while in KVM environments it is supported through QEMU's live migration framework. The total downtime experienced by end users during a live migration is typically measured in milliseconds, making it imperceptible in most applications.

14. Virtualization in Cloud Datacenters

14.1 Datacenter Architecture

Modern cloud datacenters are designed around the principle of massive horizontal scalability, in which compute, storage, and networking resources are organized into standardized building blocks that can be added or removed independently. A typical cloud datacenter is structured around pods or racks of servers, each containing dozens of physical machines connected through high-speed networking equipment. Virtualization is applied at every layer of this architecture: compute virtualization allows physical servers to host dozens of virtual machines; network virtualization decouples the logical network topology from the physical cabling; and storage virtualization presents heterogeneous storage devices as unified, policy-driven storage pools. The orchestration layer, typically implemented using platforms such as OpenStack, VMware vCenter, or a cloud provider's proprietary management system, provides a unified control plane through which administrators can provision, monitor, and manage the entire virtualized infrastructure.

14.2 Resource Scheduling and Automation

One of the most powerful capabilities of virtualized datacenter environments is the ability to automate resource scheduling decisions. Modern hypervisor management platforms continuously monitor the resource utilization of all virtual machines across the cluster and automatically migrate workloads between physical hosts to maintain optimal balance. In VMware environments, this feature is implemented by Distributed Resource Scheduler (DRS), which evaluates cluster-wide CPU and memory utilization at regular intervals and issues migration recommendations or performs automatic migrations when thresholds are exceeded. Similarly, Distributed Power Management (DPM) can automatically consolidate workloads onto fewer physical hosts during periods of low demand and power off idle servers to reduce energy consumption, powering them back on when demand increases. These automation capabilities dramatically reduce the manual administrative burden on IT teams and enable large-scale infrastructure to be managed with relatively small operations staff.

14.3 Disaster Recovery in Virtualized Environments

Disaster recovery planning is significantly simplified in virtualized environments compared to traditional physical infrastructures. Virtual machines, being software-defined entities represented as collections of files, can be replicated to remote locations using storage replication or dedicated VM replication tools. In the event of a disaster affecting the primary datacenter, replicated virtual machines can be powered on at the recovery site within minutes, restoring service availability far more rapidly than would be possible with physical server replacement. VMware Site Recovery Manager (SRM) automates the orchestration of disaster recovery workflows, executing predefined recovery plans that power on virtual machines in the correct order, update network configurations, and validate the health of recovered services. Recovery Time Objectives (RTO) and Recovery Point Objectives (RPO) are both substantially improved through virtualization-based disaster recovery strategies.

Conclusion

Virtualization technologies have revolutionized modern computing infrastructures by improving flexibility, scalability, and resource optimization. They constitute the technological foundation of cloud computing platforms and modern datacenters. Understanding the principles and mechanisms of virtualization is therefore essential for designing efficient distributed systems and cloud environments.

Practical Activity 1

Case study: Virtualization

General objective	■ To study and practically manipulate virtualization tools .
Specific objectives	■ To study and compare virtualization tools. ■ Create, configure and make virtual machines communicate.
Keywords	■ virtualization, type 1 hypervisor, hypervisor of type 2…

Practical Activity 1
Case study: Virtualization

Goals :

- ⇨ To study and compare virtualization tools;
- ⇨ Create, configure and make virtual machines communicate;

General context:

After theoretically studying the main concepts related to virtualization, we will try, through this study, to study and manipulate some virtualization tools.

This activity has two parts:

- The first part is purely theoretical in which you will study a virtualization tool (each student chooses a tool/hypervisor) to carry out at the end a comparative study of the main virtualization tools according to certain criteria.
- The second part is devoted to a practical use of these tools chosen by the students.

Conditions for implementation:

- Completion time: 2 weeks➔ Part 1 : 1 week; Part 2 : 3 weeks
- Theoretical and practical presentations: 2 sessions = 3 hours
- Work to be done in pairs or groups of three.
- To be submitted:
  - → 1. Theoretical report of the first part
  - → 1. Manual for creating, configuring, and using virtual machines

Part 1 (the theoretical part):

The task is to answer the following questions:

- After choosing a virtualization tool, conduct a presentation and analysis of this tool, based on criteria such as: types of virtualization, supported operating systems, security, names of commercial solutions, major clients, selling points, client examples, market share…

Part 2 (Practical Work):

You will need to complete your practical work by choosing one of the case studies listed below:

1. **Case study #01 (manipulating a type 2 hypervisor)**

You are a developer at a small company working on a **cross-platform application** for internal use only. The company has Android, Windows, and Linux machines, and you want to test if your application functions correctly in this complex environment, knowing that to work it must be able to:

- **send** requests to other machines;
- **retrieve** database updates from a remote server accessible via a URL.

Work requested

To complete the required task, you will need to follow these steps:

1. Hypervisor installation.
2. Creating the VM: we will create 3 VMs here:
- 1 **Ubuntu VM** ;
- 1 **Windows 10 VM** ;
- 1 VM **Android** .
3. Network configuration
   You will try different network modes and for each mode you will retrieve the IP address of the VMs
   - Explain the link between the IP address and the chosen network mode.
   - Which network mode is best suited to your case study?

4. Duplicate, transport, import and export a VM.

- Duplicate: Right-click on the VM -> manage -> clone -> linked clone (requires the original to function) or full clone (takes up more space)
- Transporter: Right-click on the VM -> manage -> share -> move the virtual machine
- Import: This is done when creating a new machine -> choose an existing file
- Export: File -> Export to OVF

5. Sharing documents between the VM and the host **.**

VMware Tools installation required on both the host and client.

Right click -> Settings -> Options -> Shared folders Folder Sharing -> always enabled + Folder (select a folder + name, read-only option) Launch the VM -> The shared folder is found.

6. Share devices between the VM and the host.

VM -> Removable Devices -> TOSHIBA -> Connect

7. Using snapshots:

Create and restore. A snapshot is used to save the entire state of the machine at a given point in time. You can create multiple snapshots one after the other. Once the VM is running, go to the VM tab -> Snapshot -> Take Snapshot (enter the name and description). To view the snapshots, go to the VM tab -> Snapshot -> Snapshots Manager. You can also perform these operations using buttons located on the VMware Workstation taskbar.

8. Place the VMs in a private network.

Case study #02 (manipulating a type 1 hypervisor)

A small online sales company with 5 employees has the following in its network infrastructure (see Figure 13):

- **mail** server ;
- **web** server to host your website;
- 1 **NAS** to store employee work files and documents.

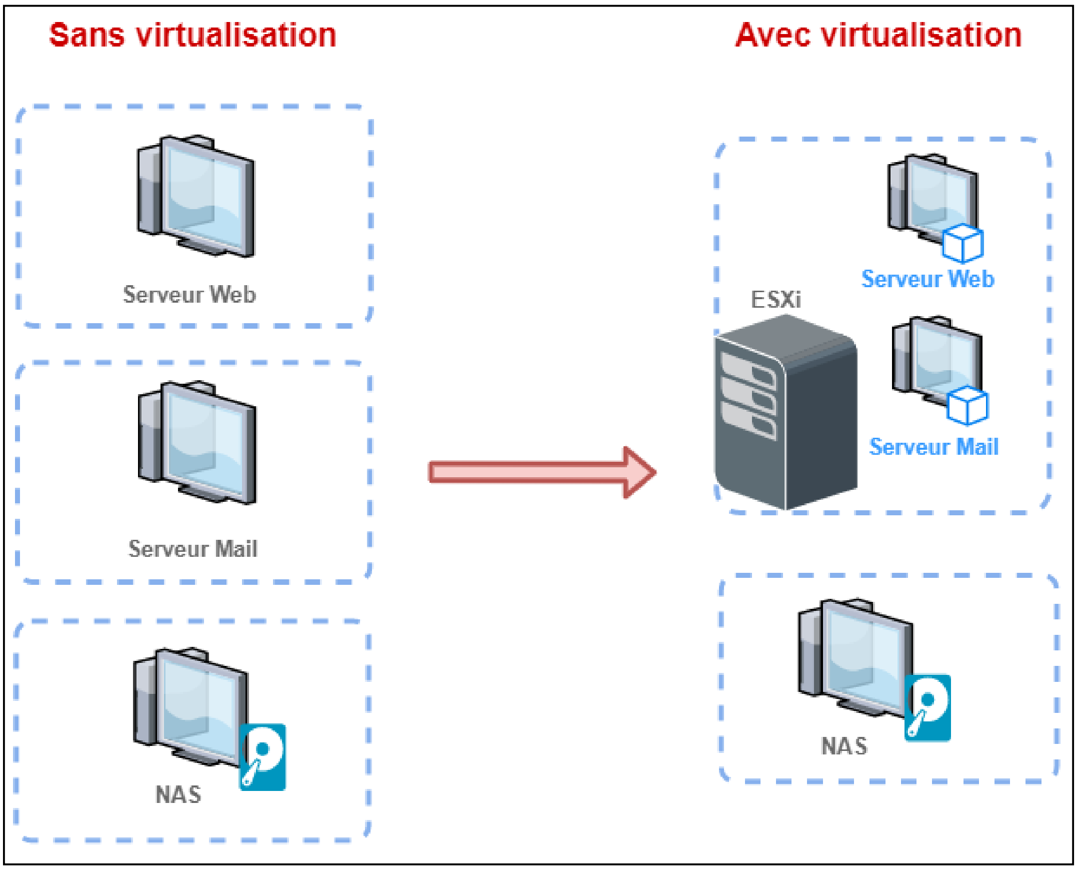


Figure 13: The required architecture of TP

Work requested

1- Install ESXi + network configuration (see Figure 14)

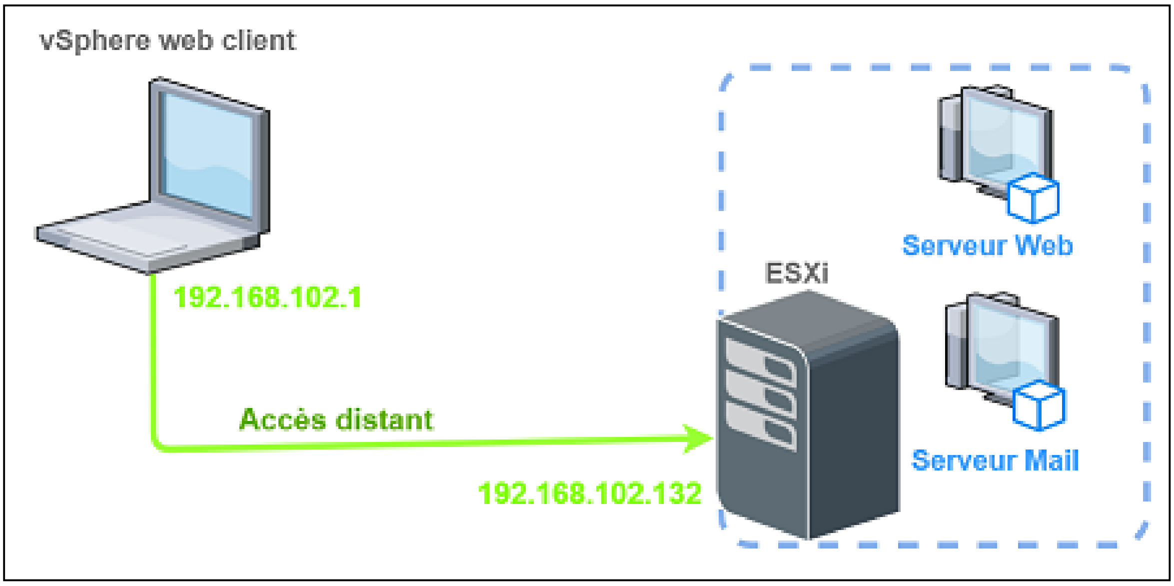


Figure 14: The requested installation

2- Accessing ESXi via a web client to create VMs
3- Enable the VMs to communicate with each other.

Chapter II

CLOUD COMPUTING

1. Introduction

Advances in information technology demand a new computing paradigm that supports the delivery of IT services without requiring on-premises infrastructure, at minimal cost. Cloud computing offers the same model described above, providing services elastically and on-demand over the internet, with fees paid only when resources are released. Generally, a cloud is a technological paradigm that extends several existing technologies, including parallel and distributed computing, Service-Oriented Architecture (SOA), virtualization, networking, and more.

Distributed computing, virtualization, and the internet function as essential building blocks of cloud computing. It's a highly shared computing paradigm where processing, storage, networking, applications, and more are shared. The goal of cloud computing is to provide on-demand, secure, high-quality, scalable, fast, more responsive, cost-effective, and automatically provisioned services. These services are delivered seamlessly (location-independently). Cloud computing can help improve business performance while also contributing to controlling the cost of IT resources for any organization.

Thus, the main objective of this chapter is to clarify the concept of Cloud Computing by presenting its history and different definitions, its main characteristics, its different layers and deployment methods , its advantages and disadvantages and a description of its main players.

2. Historical

The concept of cloud computing is not entirely new, having been introduced by John McCarthy as early as 1961 with the idea of a "computer system available as a public tool." According to Microsoft, it can be considered the **fifth** evolution of computer architectures. This evolution can be presented in the following phases (see Figure 15):

1- It began in 1980 with companies providing access to regional research networks in the United States. It wasn't until 1989 that the first Internet Service Provider (ISP 1.0) via the telephone network became operational. These providers only offered limited access to the Internet through the telephone network to connect individuals and businesses.

2- The second phase is the emergence of the second generation of ISPs 2.0, which allows web browsers to use email and access mail servers;

3- The third generation, ISP 3.0, saw the emergence of colocation centers, which allow web browsers to communicate and connect with other internet networks, access storage and processing servers, and secure physical resources and data transfer. It offers providers the ability to host business applications for companies;

4- The fourth generation, ISP 4.0. An Application Service Provider (ASP) (also called an online application provider) is a company that provides software or IT services to its customers via the internet. It sets up application generation platforms on demand, of which "Software as a Service" is a derivative (see Section 5.1). The advantage of this model is to provide access to specific applications (such as a medical billing program) using a standard protocol like the HTTP protocol;

5- The fifth generation, ISP 5.0, allows a group of users to share the same physical (under Infrastructure-as-a-Service) and logical (Software-as-a-Service) resources simultaneously and at a lower cost. ISP 5.0 also provides a business model that allows users to pay only for the resources they use, much like Google and Salesforce launched the first public website for cloud services in 1999.

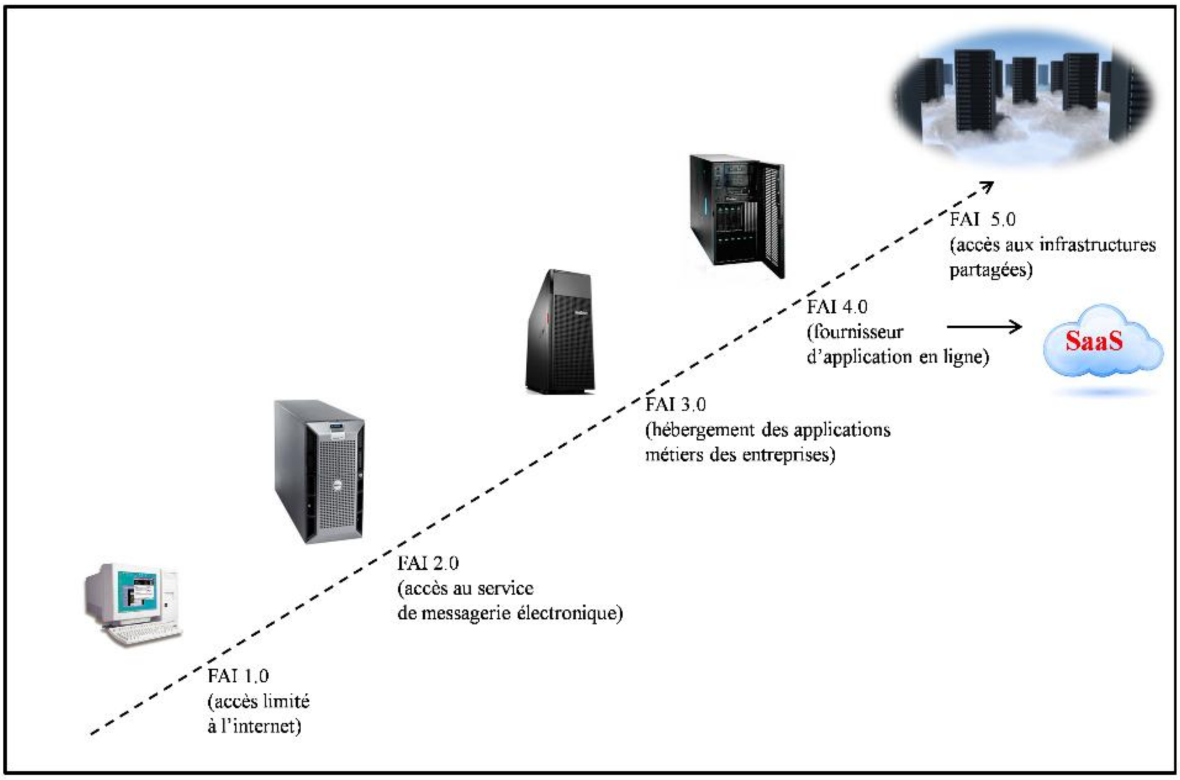


Figure 15: Evolution towards the Cloud .

3. Definition of Cloud Computing

There is currently no universally accepted definition of the concept of Cloud Computing among researchers. Research has been conducted simply to clearly define the concept of the Cloud; for example, Vaquero et al. dedicated an entire article to comparing more than 20 different definitions of the Cloud.

We found several definitions. Below, we present some quotes relating to its definition that seem to us the most important:

- According to Cisco Systems, the world leader in networking technologies, cloud computing can be defined as follows: " *IT resources and services that are abstracted from the underlying infrastructure and provided 'on-demand' and 'at scale' in a multitenant environment."* In other words, cloud computing is a shared IT platform providing businesses with on-demand and scalable services, creating the illusion of unlimited resources.
- According to NIST (National Institute of Science and Technology), cloud computing is defined as follows: *"Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."* In other words, cloud computing is a model that allows on-demand access, via the network, to a set of shared and configurable computing resources (networks, servers, storage capacity, applications, and services) that can be quickly provisioned and released with minimal administrative effort or service provider intervention.
- According to Gartner, the Cloud is defined as follows: " *Cloud is a style of computing where scalable and elastic IT-related capabilities are provided as a service to external customers using Internet technologies* . " In other words, Cloud Computing is a type of computing in which highly scalable capabilities are provided as a service to multiple customers via Internet technologies.
- Amazon refers to cloud computing as the on-demand delivery of computing power, database storage, applications, and other computing resources through a cloud service platform via the internet, with payment based on actual usage. Whether we're using photo-sharing apps across millions of mobile users or performing critical business

operations, the cloud offers rapid access to flexible and cost-effective computing resources .

From these four definitions, we can see that cloud computing is based on virtualization technology, which involves sharing configurable computing resources from a data center. Examples include networks, servers, database storage, computing power, applications, and so on. The cloud also provides its customers with on-demand, instant, and elastic services, with billing based on actual usage.

4. The characteristics of Cloud Computing

According to NIST, Cloud Computing includes the following five characteristics:

4.1. Self-service and on-demand access

The service is provided to the customer at the time of their need and request without any interaction with the service provider. In other words, on-demand service allows users to utilize cloud resources, such as server time and network storage capacity, according to their needs, and provides open access through resource provisioning without human interaction between the user and the service provider.

4.2.Ubiquitous access to the network

The resources are accessible on the network from heterogeneous platforms (tablets, workstations, smartphones, etc.).

4.3.***Pooling of resources***

The same resources are pooled and dynamically allocated to different clients. In the cloud, a single application can be used by multiple clients simultaneously, while preserving the security and privacy of each client's data. This is made possible by using virtualization tools that allow a server to be shared among several users. Various physical and virtual resources (e.g., storage, processing power, memory, and network bandwidth) are dynamically allocated and reassigned to clients based on consumer demand. The client cannot control the exact location of the resources provided. However, they can specify a preferred location for their resources, such as a country or continent, through the integration of a higher level of abstraction .

4.4.Rapid elasticity

Rapid elasticity refers to the cloud's ability to adapt (increase or decrease) resource allocation quickly, flexibly, and efficiently, according to user needs. Compute capacity appears to the user as a vast reservoir of dynamic resources that can be paid for when needed .

4.5.Measured service

Resource usage can be monitored, controlled, and reported, providing transparency for both the service provider and the consumer. Cloud computing allows customers to pay only for what they consume, based on the type of service (storage, processing, bandwidth, active user accounts, etc.) and the time spent using the service.

4.6.Resilience and High Availability in the Cloud

Beyond the five characteristics defined by NIST, Cloud Computing is also distinguished by its ability to ensure resilience and high availability of services. Cloud providers deploy their infrastructures across multiple geographic locations, called regions and availability zones, in order to guarantee service continuity in the event of a component or physical location failure. Through automatic data replication and system redundancy, cloud-hosted applications can maintain availability levels exceeding 99.9%, which corresponds to less than nine hours of downtime per year. This inherent resilience is a major argument for organizations seeking to ensure the continuity of their critical business operations.

5. Service models

The NIST working group classified Cloud services into three models (layers), based on the nature of the service, namely software, platform or infrastructure (see Figure 16).

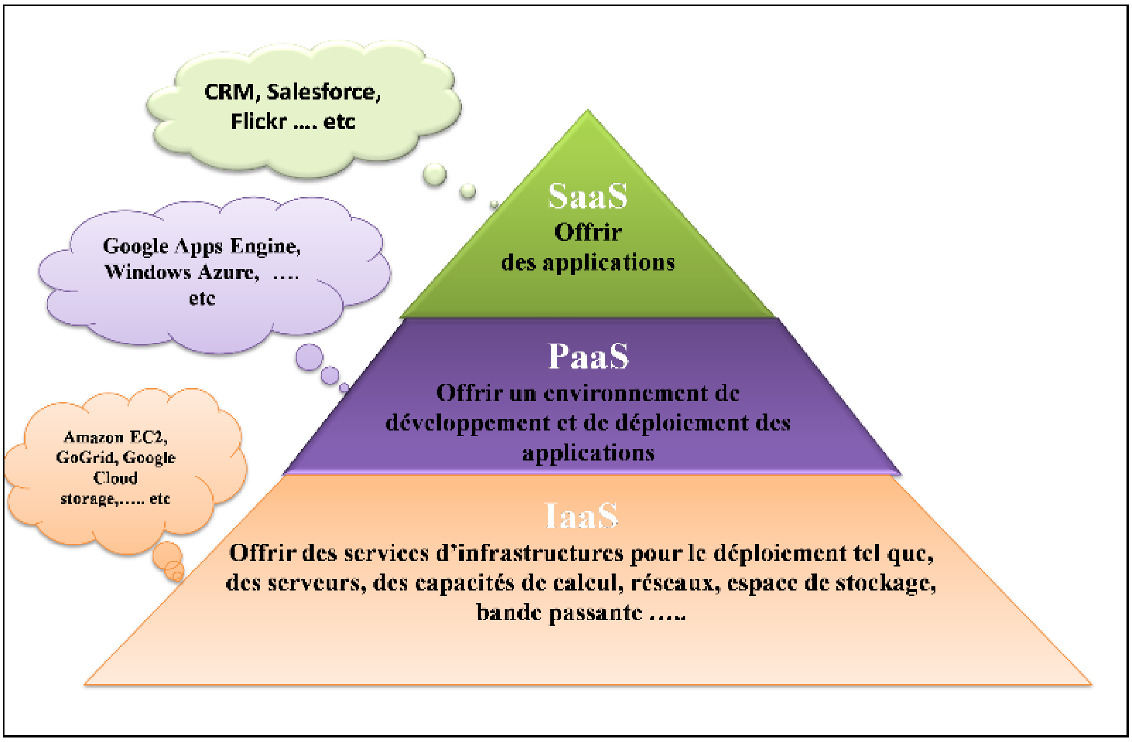


Figure 16: Reference model of Cloud Computing .

5.1.SaaS (Software as a Service)

This is an application deployment model in which a provider (*SaaS operator*) rents a software application to its customers as a service, on demand. Many software applications are offered as a service, such as online messaging, customer relationship management (CRM) systems, storage software like Flickr, human resources management, accounting, Salesforce, etc.

The setup of infrastructure, hosting servers, and operating systems is entirely handled by cloud service providers and is completely abstract for the user. However, the user can configure or customize their application via an API (Application Programming Interface) that the provider sets up for its clients, allowing them to configure the SaaS according to their needs. Furthermore, the user is free to choose not to perform any updates; it is the responsibility of the providers to control access to and updates of the SaaS they provide.

5.2.PaaS (Platform as a Service)

This model provides users with a set of programming languages and tools to develop software applications (for example, web applications, mobile applications, or desktop applications) without needing to install any tools on their local workstation. These applications can then be deployed to the provider's cloud infrastructure. The provider maintains complete control and management of the underlying application infrastructure; however, the user/developer only has control over the deployed applications.

The services offered by a platform as a service are generally accessible via public web interfaces. The best-known development platforms are offered as services by major software publishers, such as Google App Engine.[1] and Windows Azure [2].

5.3.IaaS (Infrastructure as a Service)

This model provides infrastructure as a service for deploying and running applications, such as servers, computing power, networks, storage, bandwidth, etc. These services are offered over the internet on a pay-as-you-go basis. Infrastructure customers only pay for the resources they consume. The user controls the operating systems, storage capacity, and applications deployed . Amazon EC2 (Amazon Elastic Compute Cloud) [3]is an example of IaaS.

To simplify these concepts, we can remember that with:

- SaaS: we can use an application;
- PaaS: we can develop and deploy an application; and
- IaaS: we can host our applications.

5.4 FaaS (Function as a Service) / Serverless Computing

Function as a Service (FaaS), commonly known as serverless computing, represents the most granular level of cloud service abstraction available today. In this model, developers write and deploy individual functions, small, discrete units of code designed to perform a single, well-defined task, without managing or provisioning any underlying infrastructure whatsoever. The term "serverless" is somewhat misleading: servers still exist and execute the code, but they are entirely managed by the cloud provider, completely invisible to the developer.

When a function is triggered by a specific event, such as an HTTP request, a file upload, a database change, or a scheduled timer, the cloud provider automatically instantiates a runtime environment, executes the function, and releases the resources immediately upon completion. The client is billed exclusively for the actual execution time of the function, measured in increments as small as one millisecond, and for the number of invocations. This pay-per-execution model makes FaaS extremely cost-efficient for workloads that are sporadic, event-driven, or highly variable in intensity. AWS Lambda, introduced by Amazon in 2014, was the

---

[1] http://appengine.google.com
[2] https://azure.microsoft.com/fr-fr/
[3] https://aws.amazon.com/fr/ec2/

pioneering FaaS platform and remains the market leader. Google Cloud Functions and Microsoft Azure Functions are the other major offerings in this space.

Serverless computing is particularly well-suited for use cases such as processing API requests in web applications, transforming data in real-time pipelines, executing background tasks triggered by user actions, and implementing lightweight microservices. However, the serverless model also presents notable limitations. Cold start latency, the delay incurred when a function is invoked after a period of inactivity, during which the provider must initialize a new runtime environment, can be problematic for latency-sensitive applications. Additionally, the maximum execution time per invocation is limited (typically between 5 and 15 minutes depending on the provider), making FaaS unsuitable for long-running computational tasks.

5.5 XaaS: Everything as a Service

The proliferation of cloud computing service models has given rise to the broader concept of XaaS, or "Everything as a Service," a term used to describe the growing tendency to deliver any IT capability, tool, or process over the internet as an on-demand, subscription-based service. Beyond the foundational IaaS, PaaS, and SaaS models, numerous specialized service categories have emerged to address specific organizational needs.

Database as a Service (DBaaS) allows organizations to provision, manage, and scale relational or NoSQL databases without deploying or maintaining database server software, with examples including Amazon RDS, Google Cloud SQL, and MongoDB Atlas. Security as a Service (SECaaS) provides cloud-delivered security capabilities such as identity management, encryption, intrusion detection, and vulnerability scanning. Desktop as a Service (DaaS) delivers virtual desktops to end users over the internet, eliminating the need for local workstation management. The XaaS paradigm reflects a fundamental shift in the IT industry away from capital expenditure on owned hardware and software toward operational expenditure on flexible, consumption-based service subscriptions.

6. Cloud Deployment Models

According to the NIST working group , Cloud services can be deployed according to four models: private, public, community, and hybrid cloud, corresponding to different uses:

6.1.The Public Cloud

The public cloud is generally open for use by the general public on the internet. The general public can be defined as users, a commercial, academic or governmental organization, or a combination thereof (see Figure 17).

Amazon, Google, Microsoft, and Salesforce offer a public cloud in which any individual or company can host their applications, services, or data.

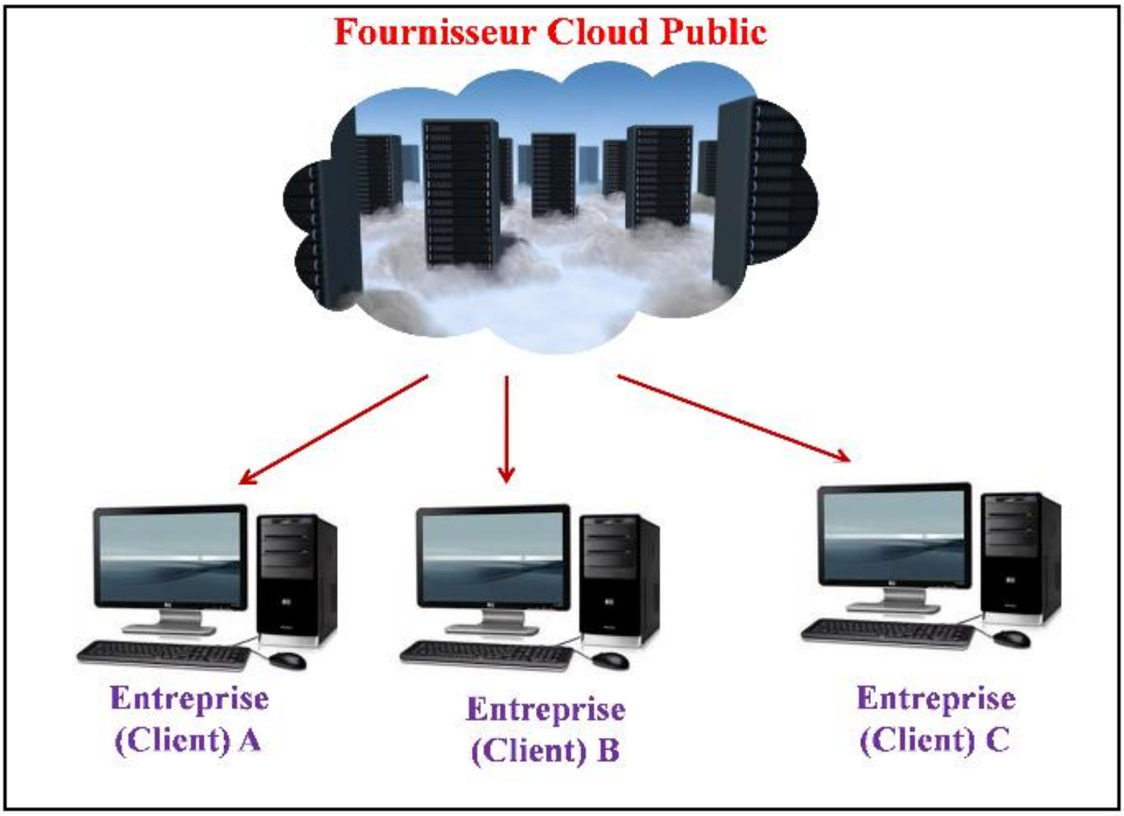


Figure 17: Diagram of the public cloud architecture .

6.2.Private Cloud

Cloud infrastructure is provided for the exclusive use of a single organization. It can be owned, managed, and operated by the organization, a third party, or a combination of both. It can reside on the organization's premises or remotely. Eucalyptus, OpenNebula, and OpenStack are examples of solutions for implementing a private cloud.

So the main difference between a public cloud and a private cloud is that the resources of a private cloud are intended only for clients authorized by the organization that owns the resources, and they cannot be shared with other clients from outside, unlike the public cloud where resources can be shared by the general public.

We can distinguish between two types of private cloud:

1- **Internal private clouds :** in this type of cloud, applications, physical and virtual resources are grouped and managed within the infrastructure owned by the organization.
2- **External private clouds :** in this type of cloud, resources are intended for the needs of the company, but they are hosted and deployed at a provider .

6.3.The Community Cloud

Cloud infrastructure is shared by multiple organizations. This community of organizations may share the same concerns, for example, resource management tasks, data security, application deployment, authentication, etc.

These resources may be owned, managed, and operated by one or more community organizations, a third party, or a combination thereof. They may exist within or outside the premises of the various community organizations.

Examples of community clouds [4]:

- The GSA (General Services Administration) in the United States, which launched a community site for American government organizations.
- Amadeus, the leading provider of IT solutions to the tourism and travel industry, was created 20 years ago by Air France, Lufthansa, Iberia and SAS. Today it is the world's leading player in the travel sector, with more than 150 airline clients, 280 million daily transactions and 2,500 IT professionals.
- CMed, a startup launched in 2010 to create a new community cloud for pharmaceutical laboratories, sought to support laboratories worldwide in their work by creating its Timaeus application following the standardization and harmonization of marketing authorization processes for new drugs around the world.

6.4.Hybrid cloud

A hybrid cloud infrastructure is composed of two or more distinct cloud infrastructures (private, community, or public). These remain unique entities but are linked by standardized or proprietary technologies that enable the portability of data and applications between the different cloud infrastructures. For example, applications can be exported to a public cloud, but these applications can also use data stored on a private cloud, or two applications hosted in two separate clouds can communicate with each other.

7. The advantages and disadvantages of the Cloud

Any serious analysis of cloud computing must consider the advantages and disadvantages offered by this technology. In what follows, we will outline the advantages and disadvantages of cloud computing.

[4] https://blog.outscale.com/fr/le-cloud-communautaire-nouveau-modele-en-vogue

7.1.Benefits

Cloud computing offers many advantages and flexibility to its users. Users can operate securely from anywhere, at any time. Given the increasing number of web-enabled devices in use today (e.g., tablets, smartphones, etc.), access to information and data needs to be faster and simpler. Some of these highly relevant advantages of using a cloud include:

1- Reduce management costs and initial investment: with the Cloud, companies don't have to worry about managing resources or the staff needed to monitor their platforms. The Cloud minimizes business risks;
2- To provide a dynamic infrastructure that offers reduced costs and improved services with lower development and maintenance costs;
3- Providing on-demand, flexible, scalable, improved and adaptable services through the " *Pay-as-you-go "* usage-based payment model ;
4- Provide consistent availability and performance with automatically provisioned maximum loads;
5- Recover quickly and improve restoration capabilities to enhance business resilience;
6- To provide unlimited processing, storage, network capacity, etc., in an elastic manner;
7- Offering automatic software updates, improved document format compatibility, and improved compatibility between different operating systems;
8- To offer easy group collaboration, meaning flexibility for users worldwide to work on the same project; and
9- Offers environmentally friendly computing because it only uses the server space required by the application.

7.2.The disadvantages of Cloud Computing

Some of the disadvantages of using a cloud service are as follows :

1- The Cloud requires a network with high-speed communication and constant connectivity;
2- Data and applications on a public cloud may not be very secure, which raises privacy and security concerns; and
3- Requires constant monitoring and enforcement of service level agreements (SLAs).
Security represents one of the main barriers to Cloud Computing adoption, particularly for sensitive data. When data is hosted on servers belonging to a third-party provider, legitimate questions arise regarding its confidentiality, integrity, and accessibility.

4- Security Challenges in the Cloud

Regulations in force, such as the General Data Protection Regulation (GDPR) in Europe, impose strict constraints on the location and processing of personal data, which can complicate the use of certain non-European cloud providers. Organizations must therefore carefully evaluate the security policies of their service providers, the encryption mechanisms offered, and the incident management procedures before entrusting their data to the cloud.

5- The Shared Responsibility Model

A fundamental concept in cloud security is the Shared Responsibility Model. This model defines the distribution of security obligations between the cloud provider and the client. As a general rule, the provider is responsible for the security of the underlying infrastructure (physical hardware, network, hypervisor), while the client remains responsible for the security of what is deployed on top of that infrastructure (operating systems, applications, data, identity and access management). This boundary varies depending on the service model adopted: in an IaaS context, the client assumes more responsibilities than in a SaaS context, where most management is delegated to the provider.

8. The main players in Cloud Computing

According to recent research by Merrill Lynch, the cloud market is estimated to reach $160 billion, with $95 billion dedicated to business applications and enterprise productivity, and $65 billion to online advertising production. Another study by Morgan Stanley also identified the cloud as a major technological trend. As the IT industry has recently shifted towards providing Platform as a Service (PaaS) and Software as a Service (SaaS) to consumers and businesses, regardless of location or time, the number of available cloud platforms will increase. In this section, we compare some representative cloud platforms:

Amazon Elastic Compute Cloud [5]provides a virtual computing environment that allows a user to run Linux-based applications. The user can either create a new Amazon Machine Image (Amazon AMI) containing the applications, libraries, data, and associated configuration settings, or select a library of available AMIs from the AMIs library. The user must then upload the created or selected AMIs to Amazon Simple Storage Service (S3) before starting, stopping,

[5] https://aws.amazon.com/fr/ec2/

and monitoring the uploaded AMI instances. Amazon EC2 charges the user from the moment the instance is activated, while Amazon S3 charges for all data transfers (upload and download).

Google App Engine [6]allows users to run web applications written using the Python programming language. In addition to supporting the standard Python library, Google App Engine also supports APIs (Application Programming Interfaces) for databases, Google accounts, URL lookup, image manipulation, and email services. Google App Engine also provides a web-based administration console that allows users to easily manage their running web applications. Currently, Google App Engine is free to use with 500 MB of storage and approximately 5 million page views per month.

Microsoft Windows Azure[7] is a cloud-based operating system development platform that allows developers to create applications using a standard framework such as REST, SOAP, or HTTP. It offers hosting (applications and data) and services (workflow, data storage and synchronization, message bus, contacts, etc.). A set of APIs is also available. This allows the use and access of this platform and its associated services. A runtime environment (the "Live Operating Environment") enables tight integration with the main existing operating systems (Windows, Mac OS, and Windows Phone).

Salesforce.com [8]: Salesforce is a software publisher, based in San Francisco to United States. It distributes management software based on the Internet and hosts enterprise applications. The company is best known internationally for its SaaS-based customer relationship management solutions; it has become one of the pioneers of the SaaS model, particularly thanks to its long-standing CRM tool.

OpenStack: The Open-Source Cloud Platform: OpenStack is a free, open-source cloud computing platform that enables organizations to build and manage their own private or public cloud infrastructure. Originally developed as a joint project between NASA and Rackspace in 2010, OpenStack has grown into one of the most widely deployed cloud platforms in the world, supported by a large community of contributors and backed by major technology vendors including Red Hat, SUSE, IBM, and Canonical.

OpenStack is composed of a collection of interoperable services, each responsible for a specific aspect of cloud infrastructure management. The Nova compute service manages the provisioning and lifecycle of virtual machine instances. The Neutron networking service provides software-defined networking capabilities, enabling the creation of virtual networks,

[6] http://appengine.google.com
[7] https://azure.microsoft.com/fr-fr/
[8] https://www.salesforce.com/

routers, subnets, and security groups. The Swift object storage service offers a scalable, highly available storage system for unstructured data such as images, backups, and log files. The Cinder block storage service provides persistent block storage volumes that can be attached to virtual machine instances. The Keystone identity service manages authentication and authorization across all OpenStack components. Horizon provides a web-based graphical dashboard through which administrators and users can interact with the cloud environment. OpenStack is particularly popular among telecommunications operators, research institutions, and large enterprises that require full control over their cloud infrastructure without dependence on commercial hyperscaler vendors.

9. Cloud and virtualization

This renowned technology (virtualization) abstracts away the physical details of the hardware and provides resource virtualization at a very high application level. Virtualization forms the foundation and basis of cloud computing, as it enables the pooling of computing resources in server clusters and their dynamic allocation/releasing within the virtual environment intended for applications.

In summary, the Cloud relies on virtualization technology to achieve the goal of dynamically providing computing resources according to their use.

9.1. Hyperconvergence as an Evolution of Virtualization

Hyperconvergence represents a significant evolution of traditional virtual infrastructure. A Hyper-Converged Infrastructure (HCI) integrates within a single software platform the compute, storage, and networking resources that are typically managed separately in a conventional architecture. This approach considerably simplifies datacenter management, reduces hardware costs, and improves scalability. Solutions such as VMware vSAN, Nutanix, or Microsoft Azure Stack HCI illustrate this trend. Hyperconvergence also facilitates the transition to hybrid cloud by enabling unified management of both local and remote resources from a centralized interface.

9.2. Edge Computing as a Complement to the Cloud

Faced with the limitations of centralized Cloud Computing in terms of latency and bandwidth, a complementary approach has emerged: Edge Computing. This consists of bringing computing and storage capabilities closer to the sources of data, that is, at the edge of the network (at the level of factories, connected vehicles, surveillance cameras, etc.). This distributed architecture enables real-time processing of data generated by connected devices (IoT), without the need to send it to a central datacenter. Edge Computing and Cloud Computing are not opposed; they are complementary: lightweight, real-time processing is performed at the edge, while complex analyses and long-term storage remain in the cloud.

10. Cloud Security

Cloud security encompasses all the policies, technologies, controls, and practices designed to protect cloud computing environments, data, applications, and infrastructure from threats and vulnerabilities. As organizations increasingly migrate critical workloads to the cloud, security considerations have become paramount. Cloud environments introduce unique security challenges that differ fundamentally from traditional on-premises infrastructures, primarily because data is stored and processed on third-party systems and accessed over public networks.

10.1. *Principal Threats to Cloud Environments*

Among the most significant threats faced by cloud environments are data breaches, in which unauthorized parties gain access to sensitive organizational or customer data stored in the cloud. Account hijacking through credential theft or phishing attacks represents another major risk, potentially allowing attackers to manipulate cloud services, redirect transactions, or access confidential data. Insecure APIs constitute a further vulnerability: since cloud services are managed and accessed through APIs, poorly secured interfaces can expose organizations to unauthorized access, data injection, and denial-of-service attacks. Finally, misconfiguration of cloud resources, such as leaving storage buckets publicly accessible or granting excessive permissions, remains one of the leading causes of cloud-related security incidents.

10.2 *Encryption in the Cloud*

Encryption is one of the most fundamental mechanisms for protecting data in cloud environments. Data can be encrypted both at rest (when stored on cloud servers) and in transit (when transmitted between the client and the cloud provider). Most major cloud providers offer built-in encryption services, but organizations may also choose to manage their own encryption keys through dedicated key management systems (KMS) to retain greater control over access to their data. End-to-end encryption ensures that data remains unreadable to unauthorized parties, including, in some configurations, the cloud provider itself. This is particularly critical for healthcare, financial, and government organizations that handle highly sensitive information.

10.3. *Identity and Access Management (IAM)*

Identity and Access Management (IAM) is a framework of policies and technologies that ensures the right individuals have the appropriate level of access to cloud resources. IAM systems allow administrators to define granular access control policies, specifying which users or services can perform which operations on which resources. Multi-factor authentication (MFA) adds an additional layer of verification beyond passwords, significantly reducing the risk of unauthorized account access. Role-based access control (RBAC) assigns permissions based on organizational roles rather than individual users, simplifying access management at scale and reducing the risk of excessive privilege assignments.

11. Emerging Trends in Cloud Computing

11.1. *Multi-Cloud Strategy*

A multi-cloud strategy refers to the use of cloud services from multiple providers simultaneously, rather than relying on a single vendor. Organizations adopt multi-cloud approaches to avoid vendor lock-in, optimize costs by choosing the most cost-effective provider for each workload, improve resilience by distributing services across multiple environments, and leverage the unique capabilities of different cloud platforms. For example, an organization might use AWS for its compute-intensive machine learning workloads, Google Cloud Platform for its data analytics pipelines, and Microsoft Azure for its enterprise productivity applications. While multi-cloud strategies offer significant flexibility benefits, they also introduce complexity in terms of management, security governance, and data integration between different cloud environments.

11.2. *Cloud-Native Development*

Cloud-native development is an approach to building and running applications that fully exploits the advantages of cloud computing. Cloud-native applications are designed from the ground up to be scalable, resilient, and manageable in dynamic cloud environments. They are typically built using microservices architectures, where a large application is decomposed into small, loosely coupled, independently deployable services. Each microservice handles a specific business function and communicates with others through lightweight APIs. This approach enables development teams to update, scale, and deploy individual components without affecting the entire application, significantly accelerating the software delivery cycle. Cloud-native development relies heavily on containerization (Docker), orchestration (Kubernetes), continuous integration and continuous delivery (CI/CD) pipelines, and declarative infrastructure management through Infrastructure as Code (IaC) tools such as Terraform or Ansible.

11.3 *Serverless and Event-Driven Architectures*

As organizations seek to reduce operational overhead and accelerate time-to-market, serverless and event-driven architectures are gaining widespread adoption. In an event-driven model, application components communicate by producing and consuming events — discrete units of information that signal a change in state or the occurrence of an action. Message brokers such as Apache Kafka, Amazon SQS, or Google Pub/Sub serve as intermediaries that decouple event producers from consumers, enabling highly scalable and fault-tolerant systems. Combined with serverless computing functions, this architecture enables organizations to build responsive, cost-efficient systems capable of handling highly variable workloads without provisioning dedicated infrastructure.

12. Cloud Cost Management and Optimization

12.1 Understanding Cloud Pricing Models

One of the most significant financial considerations in cloud computing is understanding and managing the cost of cloud services. Unlike traditional IT infrastructure, where costs are primarily capital expenditures incurred at the time of hardware purchase, cloud computing introduces operational expenditure models in which costs are incurred continuously based on resource consumption. The most common pricing model is pay-as-you-go, in which resources are billed by the hour, minute, or second based on actual usage. This model offers maximum flexibility but can result in unexpected costs if resources are not properly monitored and controlled.

Reserved instances or committed use discounts offer significant price reductions, typically between 30% and 72% , in exchange for a commitment to use a specific resource configuration for a fixed period of one to three years. This model is appropriate for predictable, stable workloads. Spot instances or preemptible virtual machines offer the lowest prices by allowing cloud providers to interrupt workloads when the underlying capacity is needed for higher-priority customers. These are suitable for fault-tolerant, batch-processing workloads that can be interrupted and resumed without data loss.

12.2 FinOps: Financial Operations for the Cloud

FinOps, short for Financial Operations, is an emerging discipline that combines financial management principles with cloud engineering and operations practices to help organizations manage, optimize, and predict their cloud spending. The FinOps Foundation defines it as a cultural practice that enables engineering, finance, and business teams to collaborate and make informed, data-driven decisions about cloud expenditure.

In practice, FinOps involves establishing cloud cost visibility through detailed tagging and allocation of cloud resources to specific teams, projects, or business units; identifying and eliminating waste such as idle virtual machines, unattached storage volumes, and over-provisioned resources; optimizing resource sizing by right-sizing instances to match actual workload requirements; and implementing governance policies that enforce cost controls without impeding development velocity. FinOps tools such as AWS Cost Explorer, Google Cloud Cost Management, and third-party platforms like CloudHealth or Apptio Cloudability provide dashboards, anomaly detection, and recommendations to support these activities.

12.3 Cloud Resource Optimization Strategies

Organizations frequently over-provision cloud resources out of caution, leading to significant waste. Several strategies can be applied to optimize cloud resource usage. Auto-scaling is a mechanism that automatically adjusts the number of compute instances in response to real-time demand, scaling out during traffic peaks and scaling in during quiet periods, ensuring that the organization pays only for what it actually needs at any given moment. Scheduled scaling extends this concept by applying predefined scaling schedules based on known demand patterns, such as increasing capacity before business hours and reducing it overnight.

Container-based workloads, particularly when orchestrated with Kubernetes, enable a higher density of application deployment on a given set of virtual machine instances, reducing per-application infrastructure costs. Storage optimization strategies include automatically migrating infrequently accessed data to lower-cost storage tiers, such as Amazon S3 Glacier or Google Cloud Nearline, while keeping frequently accessed data on higher-performance, higher-cost storage classes.

13. Cloud Computing and Big Data

13.1 The Relationship Between Cloud and Big Data

Big Data refers to datasets so large, fast-moving, or diverse that traditional data processing tools are insufficient to handle them effectively. The three defining characteristics of Big Data, commonly described as the three Vs, are Volume (the sheer quantity of data generated), Velocity (the speed at which data is produced and must be processed), and Variety (the wide range of data types and formats, including structured, semi-structured, and unstructured data).

Cloud computing and Big Data are deeply complementary technologies: the cloud provides the elastic, on-demand infrastructure necessary to store and process massive datasets, while Big Data analytics generate the insights that justify the investment in cloud-scale infrastructure. Cloud providers offer a comprehensive ecosystem of managed Big Data services. Amazon EMR (Elastic MapReduce), Google Dataproc, and Azure HDInsight are managed cluster services that simplify the deployment of distributed data processing frameworks. Google BigQuery and Amazon Redshift are cloud-native data warehousing services capable of executing complex SQL queries across petabyte-scale datasets within seconds.

13.2 Data Lakes and Data Warehouses in the Cloud

A data lake is a centralized repository that allows organizations to store all their structured and unstructured data at any scale, in its raw native format, without requiring a predefined schema. Unlike traditional data warehouses, which require data to be transformed and structured before ingestion, data lakes adopt a schema-on-read approach in which the structure is applied only when the data is queried or analyzed. This flexibility makes data lakes particularly suitable for storing diverse data types such as log files, sensor readings, images, videos, social media streams, and transactional records. Cloud object storage services, such as Amazon S3, Google Cloud Storage, and Azure Blob Storage, are the most common foundations for cloud-based data lakes.

A data warehouse, in contrast, is a structured repository optimized for analytical query processing. Data is ingested from operational systems, cleaned, transformed, and organized into a consistent schema before being loaded into the warehouse. Cloud data warehouses such as Snowflake, Google BigQuery, and Amazon Redshift provide columnar storage architectures and massively parallel processing engines that deliver exceptional query performance on large analytical datasets.

Practical Activity 2

Case study: Cloud Computing

General objective	■ Discover and practically try out Clouds .
Specific objectives	■ Explore the possibility of migrating an application to the Cloud. ■ Discover and practically try out clouds .

Practical Activity 2

Case study: Cloud Computing

Goals :

- ⇨ Explore the possibility of migrating to the Cloud , identifying opportunities and benefits
- ⇨ Discover and practically try out Clouds ;

General context:

After theoretically studying the main concepts related to the Cloud, we will try, through this study, to discover and practically try out Clouds.

This activity has two parts:

- The first part is purely theoretical in which you will conduct a study to test the ability to migrate your company's services (the town hall according to our case study) to the Cloud.
- The second section focuses on the practical use of cloud services.

Conditions for implementation:

- Completion time: 4 weeks➔ Part 1 : 1 week; Part 2 : 3 weeks
- Practical presentations: 2 sessions = 6 hours
- Work to be done in pairs or groups of three.
- Deliverables:
  - → 1. Theoretical report of the first part
  - → 1 Cloud service usage report.

Practical Activity 2

Case study: Cloud Computing

Part 1:

Your city hall is constantly debating whether or not to migrate its application portfolio to the cloud. This city hall has approximately 2,000 employees, serving a city of 50,000 inhabitants that continues to grow. City hall employees use nine applications, each for a specific purpose, including cemetery management, elections, and even fire risk analysis. A significant portion of the IT department's budget is allocated to maintaining these applications. The city hall wants to reduce this budget by focusing less on the infrastructure managing the applications and more on high-value applications; this is why it is considering migrating some of its applications to the cloud.

Before deciding to migrate to the Cloud, the town hall must conduct a study by analyzing criteria.

1- From the criteria you are familiar with, select those that seem relevant to the situation, and then analyze these criteria .

The various applications of the town hall can be divided into two groups:

- Applications dedicated to managing a town hall, such as:

1- openARIA is a free software enabling Fire Risk and Accessibility Analysis for Public Access Buildings (ERP).

2- openCimetière is a free software for managing cemetery plots.

3- openElec is free software for managing electoral lists: It allows for the complete management of political elections, from voter registration…

4- openScrutin is free software for managing the composition of polling stations

5- openRésultat is free software for managing election results. It allows for the entry of results, the organization of election night events...

- Other applications, such as:
  1- Internal messaging
  2- Human Resources Management (HRM)
  3- CRM

By creating a database dependency tree for applications, what is/are the ideal candidate(s) for migration (it can be replaced by a SaaS)?

What solutions do you propose for applications that cannot be migrated as a SaaS?

Part 2:

This activity involves studying and actually working with cloud computing. Below are three examples of cloud services available on the market. You are asked to visit the websites listed and try out their cloud services.

1. Use cases for AWS cloud computing services

Amazon Web Services [9]offers a comprehensive suite of cloud computing services that enable you to develop sophisticated and scalable applications. Currently, hundreds of thousands of enterprise customers of all sizes benefit from these cloud computing services across numerous industries, including healthcare, media, finance, insurance, real estate, retail, education, and the public sector.

AWS products include:

- Web, mobile and social applications
- Big Data
- Backup and storage
- Digital multimedia
- Enterprise applications
- Games

[9] http://aws.amazon.com/fr/

2. Google Cloud Case Study

Google Cloud Platform [10]allows developers to build, test, and deploy applications on Google's highly scalable and reliable infrastructure.

The products offered are:

- Compute Engine
- App Engine
- SQL Coverage
- Cloud Storage
- Database Coverage
- Big Data
- BigQuery
- Google Cloud SDK
- Push-to-Deploy
- Playground cover
- Android Studio
- Google Plugin for Eclipse

3. Rackspace Case Study [11]

This public cloud can power the most critical workloads. Its products work together; we can easily manage them all from the control panel or APIs.

Rackspace products:

[10] https://cloud.google.com

[11] http://www.rackspace.com/cloud/

COMPUTE RACKSPACE CLOUD SERVERS POWERED BY OPENSTACK	**On-Demand Server Offer**
STORAGE RACKSPACE CLOUD BLOCK STORAGE POWERED BY OPENSTACK	**Offers fast and reliable storage**
STORAGE RACKSPACE CLOUD FILES POWERED BY OPENSTACK	**Offers easy online storage for files and media**
DATABASES RACKSPACE CLOUD DATABASES POWERED BY OPENSTACK	**Offers high-performance MySQL databases in the cloud**
APPLICATION & PLATFORM RACKSPACE CLOUD MONITORING	**Cloud Monitoring :** allows you to monitor your application deployed on an IaaS, receiving alerts at any time.

Series of exercises

Chapter I: Virtualization

Exercise 1

State whether each of the following statements is true or false. Justify your answer.

- A Type 1 hypervisor runs on top of an existing operating system.
- A virtual machine is unaware that it is running in a virtualized environment in the case of full virtualization.
- Paravirtualization requires modifying the guest operating system kernel.
- Containers and virtual machines use the same level of isolation.
- KVM is a Type 2 hypervisor integrated into the Linux kernel.
- VMware ESXi requires a host operating system to function.
- The VMotion feature allows a running virtual machine to be migrated from one physical host to another without interruption.
- Storage virtualization hides the complexity of physical storage resources from the applications that use them.

Exercise 2

Choose the correct answer(s) for each question.

Q1. Which of the following is a Type 1 hypervisor?

a) VMware Workstation
b) Oracle VirtualBox
c) VMware ESXi
d) Microsoft Virtual PC

Q2. What does the acronym VDI stand for?

a) Virtual Data Infrastructure
b) Virtual Desktop Infrastructure
c) Virtualized Disk Image
d) Variable Data Interface

Q3. Which technology allows multiple virtual machines to share a single physical GPU?

a) vMotion

b) NVIDIA vGPU

c) OpenVZ

d) Docker Swarm

Q4. Which of the following protocols is used in SAN environments?

a) HTTP

b) FTP

c) iSCSI

d) SMTP

Q5. What is the main advantage of paravirtualization over full virtualization?

a) It does not require hardware support

b) It eliminates the need to modify the guest OS

c) It offers better performance by reducing binary translation overhead

d) It is compatible with all proprietary operating systems

Exercise 3

Explain in your own words the difference between a snapshot and a clone of a virtual machine. In which situations would you use each?

1. What is the role of the VMkernel in a VMware ESXi environment?
2. Describe the Shared Responsibility Model as it applies to a company using IaaS. Who is responsible for what?
3. What are the main differences between a SAN and a DAS storage architecture? Give one use case for each.
4. A company wishes to deploy 50 virtual machines on a single physical server. What hardware characteristics should the server have, and which type of hypervisor would you recommend? Justify your answer.

Exercise 4

Complete the following comparison table between the main virtualization solutions studied in this chapter:

Criterion	VMware ESXi	Hyper-V	KVM	VirtualBox	Xen
Hypervisor Type					
License (free / commercial)					
Host OS required					
Supported guest OS					
Main use case					
Notable feature					

Exercise 5

A medium-sized company with 200 employees operates the following services: a web server, a mail server, a database server, a file server, and a development environment used by 10 developers. Currently, each service runs on a dedicated physical server, leading to very low utilization rates (between 5% and 15% per server).

Questions:

1. Propose a virtualized architecture for this company. Specify the number of physical servers needed, the hypervisor type chosen, and the distribution of virtual machines.
2. What storage architecture would you recommend (DAS, SAN, or NAS)? Justify your choice.
3. What high-availability mechanisms should be implemented to ensure business continuity?
4. Estimate the benefits (in terms of cost, energy, and administration) of transitioning to a virtualized infrastructure.

Chapter II: Cloud Computing

Exercise 1

State whether each of the following statements is true or false. Justify your answer.

1. SaaS gives the user full control over the underlying operating system.
2. A private cloud can be hosted by a third-party provider outside the organization's premises.
3. The "Pay-as-you-go" model means users pay a fixed monthly fee regardless of usage.
4. Edge Computing and Cloud Computing are mutually exclusive technologies.
5. In the Shared Responsibility Model under IaaS, the provider is responsible for the operating system.
6. Multi-factor authentication (MFA) is a best practice for securing cloud accounts.
7. FaaS (Function as a Service) is billed based on the duration of function execution.
8. A hybrid cloud combines at least two distinct cloud deployment models.

Exercise 2

Q1. According to NIST, which of the following is NOT one of the five essential characteristics of Cloud Computing?

a) On-demand self-service

b) Measured service

c) Dedicated hardware

d) Rapid elasticity

Q2. Which cloud service model gives the developer control over the runtime environment but not the underlying hardware?

a) IaaS

b) SaaS

c) PaaS

d) FaaS

Q3. Which of the following is an example of a community cloud?

a) Amazon Web Services

b) Google Drive

c) A shared government cloud for public administrations

d) A private company datacenter

Q4. What does the acronym GDPR stand for?

a) General Data Privacy Regulation

b) Global Data Protection Rule

c) General Data Protection Regulation

d) Generalized Digital Privacy Requirement

Q5. Which of the following best describes the concept of "vendor lock-in" in cloud computing?

a) The inability to share data between different users on the same cloud

b) The difficulty of migrating services from one cloud provider to another

c) The legal restriction preventing cloud providers from storing data abroad

d) The technical limit on the number of virtual machines a provider can host

Exercise 3

1. Explain the difference between rapid elasticity and resource pooling as defined by NIST. Give a concrete example of each.
2. A startup wishes to launch a mobile application that will process real-time user data. Which cloud service model (IaaS, PaaS, SaaS, or FaaS) would you recommend, and why?
3. What are the main advantages of a hybrid cloud over a purely public or purely private cloud?
4. Explain the concept of "cold start" in serverless computing. Why is it a concern for latency-sensitive applications?
5. What measures should an organization take to protect its data when migrating to a public cloud environment?

Exercise 4

For each of the following services or scenarios, identify the corresponding cloud service model (IaaS, PaaS, SaaS, or FaaS):

Service / Scenario	**Cloud Model**
A company uses Gmail for employee email communication	
A developer deploys a web application on Google App Engine	
An organization rents virtual machines on Amazon EC2	
A retail company triggers an image resizing function each time a product photo is uploaded	
A university uses Microsoft 365 for document editing and collaboration	
A startup builds and deploys a machine learning model on a managed ML platform	
A bank provisions its own virtual network and servers on Azure	
An IoT company uses a function triggered by sensor data to send alerts	

Exercise 5

A regional hospital with 1,500 staff members currently manages all its IT infrastructure on-premises. Its systems include a patient records management system (Electronic Health Record — EHR), a medical imaging storage system (PACS), internal email, a payroll application, and a public-facing appointment booking website. The hospital's IT director is evaluating a migration to the cloud.

Questions:

a. For each of the five systems listed, propose the most appropriate cloud deployment model (public, private, hybrid, or community) and justify your choice, taking into account data sensitivity and regulatory constraints.
b. For each system, identify the most suitable cloud service model (IaaS, PaaS, or SaaS) and explain your reasoning.
c. What are the main security risks associated with hosting patient health records in a public cloud? What technical and organizational measures would you recommend to mitigate these risks?
d. Propose a phased migration plan (in three stages) that minimizes disruption to hospital operations.

1- When we talk about Cloud Computing, the resources are generally located:
   - On the user's workstation.
   - On remote servers.
   - On a dedicated server.

2- Match the different concepts with their definitions (Public Cloud, Private Cloud, Hybrid Cloud)?
   - The company has chosen to host part of its infrastructure on a public platform, while the other part is hosted internally and is used primarily for critical company data.
   - Cloud computing infrastructure is hosted on servers within the company.
   - The cloud computing infrastructure is hosted on a platform located outside the company.

3- What are the advantages of Cloud Computing?
   - Outsourcing of IT services.
   - High security.
   - A reduction in costs
   - Easy access
   - An infinite capacity
   - Data outsourcing.

4- What is Cloud Computing?

- A software package
- A spreadsheet
- A way to store data remotely
- A technique for managing computer applications in a business setting

5- Which of the following statements is true for Cloud Computing?

- It will always be cheaper and safer than local computing.

- Your data can be accessed from any computer in the world, as long as you have an internet connection.
- Only a few courageous companies are investing in the cloud; others hesitate because of the risk of data leaks.

6- The cloud does not need very high-speed networks to function.

  - TRUE
  - FAKE

7- What does Cloud Computing billing include?

- Calculation time, generally per hour
- The amount of storage measured in gigabytes
- The number of servers dedicated to the company
- The number of transactions represents the number of accesses to storage services
- Input and output bandwidth
- A fixed price

8- What are the disadvantages of Cloud Computing?

- Security flaws
- Dependence on its supplier
- Subscription prices
- The cost of managing cloud IT infrastructure
- A limited offer

9- What are the modes of use?

- It allows you to outsource servers, network, and storage to remote data centers. Companies start or stop virtual servers hosted on the cloud computing platform.
- It is the platform for executing, deploying, and developing applications on the Cloud Computing platform.
- This is the application platform that provides complete applications delivered on demand. It includes various types of applications ranging from CRM and human

resources management to accounting, collaborative tools, messaging, and other business applications.

10- Ordering the steps for integrating Cloud Computing into a company:

- Planning
- Identify the desired applications and processes
- Launch of the Cloud solution
- Define the need
- Development and configuration

11- Google Docs is a type of cloud computing.

- TRUE
- Fake

References

1. Bakshi, K. (2009). Cisco cloud computing-data center strategy, architecture, and solutions point of view white paper for us public sector 1st edition.
2. Benfenatki, H., Da Silva, CF, Benharkat, AN, & Ghodous, P. (2014). Cloud-Based Business Applications Development Methodology. In WETICE Conference (WETICE), 2014 IEEE 23rd International (pp. 275-280). IEEE.
3. Buyya, R., Yeo, C.S., & Venugopal, S. (2008). Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities. In High Performance Computing and Communications, HPCC'08. 10th IEEE International Conference on, IEEE, 5-13.
4. Clearley, D. W. (2010). Cloud Computing 'Key initiative overview' by Gartner.
5. Eric. M (2011). Virtualization with VMware vSphere 4. ISBN:978-2-7460-6079-1, ISSN: 1767-1590.
6. GRASSA, N. (2014). Virtualization and Cloud Course.
7. Hess, K., & Newman, A. (2010). Virtualization in practice. Pearson.
8. M. Vaquero, L., R. Merino, L., Caceres, J., & Lindner, M. (2009). A Break in the Clouds: Towards a Cloud Definition. ACM SIGCOMM Computer Communication Review, 39(1), 50-55. doi: 10.1145/1496091.1496100 .
9. M. Vaquero, L., R. Merino, L., Caceres, J., & Lindner, M. (2009). A Break in the Clouds: Towards a Cloud Definition. ACM SIGCOMM Computer Communication Review, 39(1), 50-55. doi: 10.1145/1496091.1496100
10. Mathew, S., & Varia, J. (2014). Overview of amazon web services. Amazon Whitepapers.
11. Mell, P. M., & Grance, T. (2011). Sp 800-145. The NIST Definition of Cloud Computing.

12. Miller, M. (2008). Cloud computing: Web-based applications that change the way you work and collaborate online. That publishing.
13. Portnoy, M. (2012). Virtualization essentials (Vol. 19). John Wiley & Sons.
14. Rountree, D., & Castrillo, I. (2013). The basics of cloud computing: Understanding the fundamentals of cloud computing in theory and practice. Newnes.
15. Sajid, M., & Raza, Z. (2013). Cloud computing: Issues & challenges. In International Conference on Cloud, Big Data and Trust. 20(13), 13-15.
16. Saouli, H. (2015). Discovery of web services via Cloud Computing based on mobile agents. Doctoral thesis. University of Mohamed Khider, Biskra, Algeria.