Introduction

You will create your first machine in VirtualBox (or UTM if you can’t use VirtualBox) under specific instructions. Then, at the end of this project, you will be able to set up your own operating system while implementing strict rules.

What is a Virtual Machine?

A virtual machine is a software capable of installing an Operating System within itself, making the OS think that it is hosted on a real computer. With virtual machines we can create virtual devices that will behave in the same way as physical devices, using their own CPU, memory, network interface and storage. This is possible because the virtual machine is hosted on a physical device, which is the one that provides the hardware resources to the VM. The software program that creates virtual machines is the hypervisor. The hypervisor is responsible for isolating the VM resources from the system hardware and making the necessary implementations so that the VM can use these resources.

The devices that provide the hardware resources are called host machines or hosts. The different virtual machines that can be assigned to a host are called guests or guest machines. The hypervisor uses a part of the host machine's CPU, storage, etc., and distributes them among the different VMs.

There can be multiple virtual machines on the same host and each of these will be isolated from the rest of the system. Thanks to this, we can run different operating systems on our machine. For each virtual machine, we can run a different operating system distribution. Each of these operating systems will behave as if they were hosted on a physical device, so we will have the same experience when using an OS on a physical machine and on a virtual machine.

How do Virtual Machines work?

Virtualization allow us share a system with multiple virtual environments. The hypervisor manages the hardware system and separate the physical resources from the virtual environments. The resources are managed following the needs, from the host to the guests. When a user from a VM do a task that requires additional resources from the physical environment, the hypervisor manages the request so that the guest OS could access the resources of the physical environment.

Once we know how they work, it is a good idea to see all the advantages we get from using virtual machines:

Different guest machines hosted on our computer can run different operating systems, so we will have different OS working on the same machine.

They provide an environment in which to safely test unstable programs to see if they will affect the system or not.

We get better use of shared resources.

We reduce costs by reducing physical architecture.

They are easy to implement because they provide mechanisms to clone a virtual machine to another physical device.

What is LVM?

LVM (Logical Volume Manager) is an abstraction layer between a storage device and a file system. We get many advantages from using LVM, but the main advantage is that we have much more flexibility when it comes to managing partitions. Suppose we create four partitions on our storage disk. If for any reason we need to expand the storage of the first three partitions, we will not be able to because there is no space available next to them. In case we want to extend the last partition, we will always have the limit imposed by the disk. In other words, we will not be able to manipulate partitions in a friendly way. Thanks to LVM, all these problems are solved.

By using LVM, we can expand the storage of any partition (now known as a logical volume) whenever we want without worrying about the contiguous space available on each logical volume. We can do this with available storage located on different physical disks (which we cannot do with traditional partitions). We can also move different logical volumes between physical devices. Of course, services and processes will work the same way they always have. But to understand all this, we have to know:

Physical Volume (PV): physical storage device. It can be a hard disk, an SD card, a floppy disk, etc. This device provides us with storage available to use.

Volume Group (VG): to use the space provided by a PV, it must be allocated in a volume group. It is like a virtual storage disk that will be used by logical volumes. VGs can grow over time by adding new VPs.

Logical volume (LV): these devices will be the ones we will use to create file systems, swaps, virtual machines, etc. If the VG is the storage disk, the LV are the partitions that are made on this disk.

What is AppArmor?

AppArmor provides Mandatory Access Control (MAC) security. In fact, AppAmor allows the system administrator to restrict the actions that processes can perform. For example, if an installed application can take photos by accessing the camera application, but the administrator denies this privilege, the application will not be able to access the camera application. If a vulnerability occurs (some of the restricted tasks are performed), AppArmor blocks the application so that the damage does not spread to the rest of the system.

In AppArmor, processes are restricted by profiles. Profiles can work in complain-mode and in enforce-mode. In enforce mode, AppArmor prohibits applications from performing restricted tasks. In complain-mode, AppArmor allows applications to do these tasks, but creates a registry entry to display the complaint.

What is the difference between Apt and Aptitute?

In Debian-based OS distributions, the default package manager we can use is dpkg. This tool allows us to install, remove and manage programs on our operating system. However, in most cases, these programs come with a list of dependencies that must be installed for the main program to function properly. One option is to manually install these dependencies. However, APT (Advanced Package Tool), which is a tool that uses dpkg, can be used to install all the necessary dependencies when installing a program. So now we can install a useful program with a single command.

APT can work with different back-ends and front-ends to make use of its services. One of them is apt-get, which allows us to install and remove packages. Along with apt-get, there are also many tools like apt-cache to manage programs. In this case, apt-get and apt-cache are used by apt. Thanks to apt we can install .deb programs easily and without worrying about dependencies. But in case we want to use a graphical interface, we will have to use aptitude. Aptitude also does better control of dependencies, allowing the user to choose between different dependencies when installing a program.

How to use SSH?

SSH or Secure Shell is a remote administration protocol that allows users to control and modify their servers over the Internet thanks to an authentication mechanism. Provides a mechanism to authenticate a user remotely, transfer data from the client to the host, and return a response to the request made by the client.

SSH was created as an alternative to Telnet, which does not encrypt the information that is sent. SSH uses encryption techniques to ensure that all client-to-host and host-to-client communications are done in encrypted form. One of the advantages of SSH is that a user using Linux or MacOS can use SSH on their server to communicate with it remotely through their computer's terminal. Once authenticated, that user will be able to use the terminal to work on the server.

The command used to connect to a server with ssh is:

```
$ ssh {username}@{IP_host} -p {port}
```

There are three different techniques that SSH uses to encrypt:

Symmetric encryption: a method that uses the same secret key for both encryption and decryption of a message, for both the client and the host. Anyone who knows the password can access the message that has been transmitted.

Asymmetric encryption: uses two separate keys for encryption and decryption. These are known as the public key and the private key. Together, they form the public-private key pair.

Hashing: another form of cryptography used by SSH. Hash functions are made in a way that they don't need to be decrypted. If a client has the correct input, they can create a cryptographic hash and SSH will check if both hashes are the same.

How to implement UFW with SSH

UFW (Uncomplicated Firewall) is a software application responsible for ensuring that the system administrator can manage iptables in a simple way. Since it is very difficult to work with iptables, UFW provides us with an interface to modify the firewall of our device (netfilter) without compromising security. Once we have UFW installed, we can choose which ports we want to allow connections, and which ports we want to close. This will also be very useful with SSH, greatly improving all security related to communications between devices.

What is cron and what is wall?

Once we know a little more about how to build a server inside a Virtual Machine (remember that you also have to look in other pages apart from this README), we will see two commands that will be very helpful in case of being system administrators. These commands are:

Cron: Linux task manager that allows us to execute commands at a certain time. We can automate some tasks just by telling cron what command we want to run at a specific time. For example, if we want to restart our server every day at 4:00 am, instead of having to wake up at that time, cron will do it for us.

Wall: command used by the root user to send a message to all users currently connected to the server. If the system administrator wants to alert about a major server change that could cause users to log out, the root user could alert them with wall.