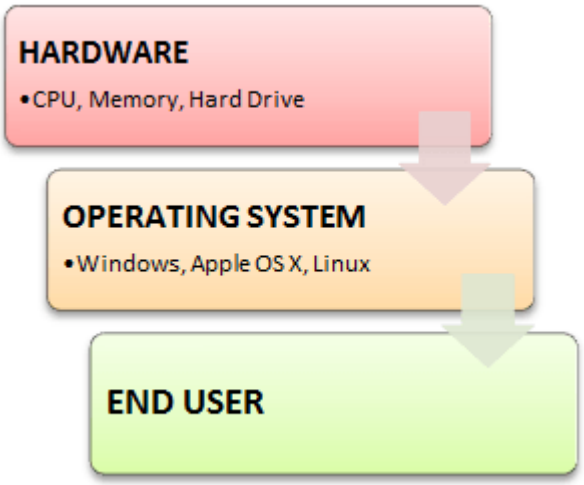
Introduction to the Linux Operating System

What is an Operating System?

Every time you switch on your computer, you see a screen where you can perform different activities like write, browse the internet or watch a video. What is it that makes the computer hardware work like that? How does the processor on your computer know that you are asking it to run a mp3 file?

Well, it is the operating system or the kernel which does this work. A kernel is a program at the heart of any operating system that takes care of fundamental stuff, like letting hardware communicate with software.

So, to work on your computer, you need an Operating System(OS). In fact, you are using one as you read this on your computer. Now, you may have used popular OS's like Windows, Apple OS X but here we will learn what Linux is and what benefits it offers over other OS choices.



What is Linux?

Linux is an operating system or a kernel. It is distributed under an open source license. Its functionality list is quite like UNIX.

Who created Linux?

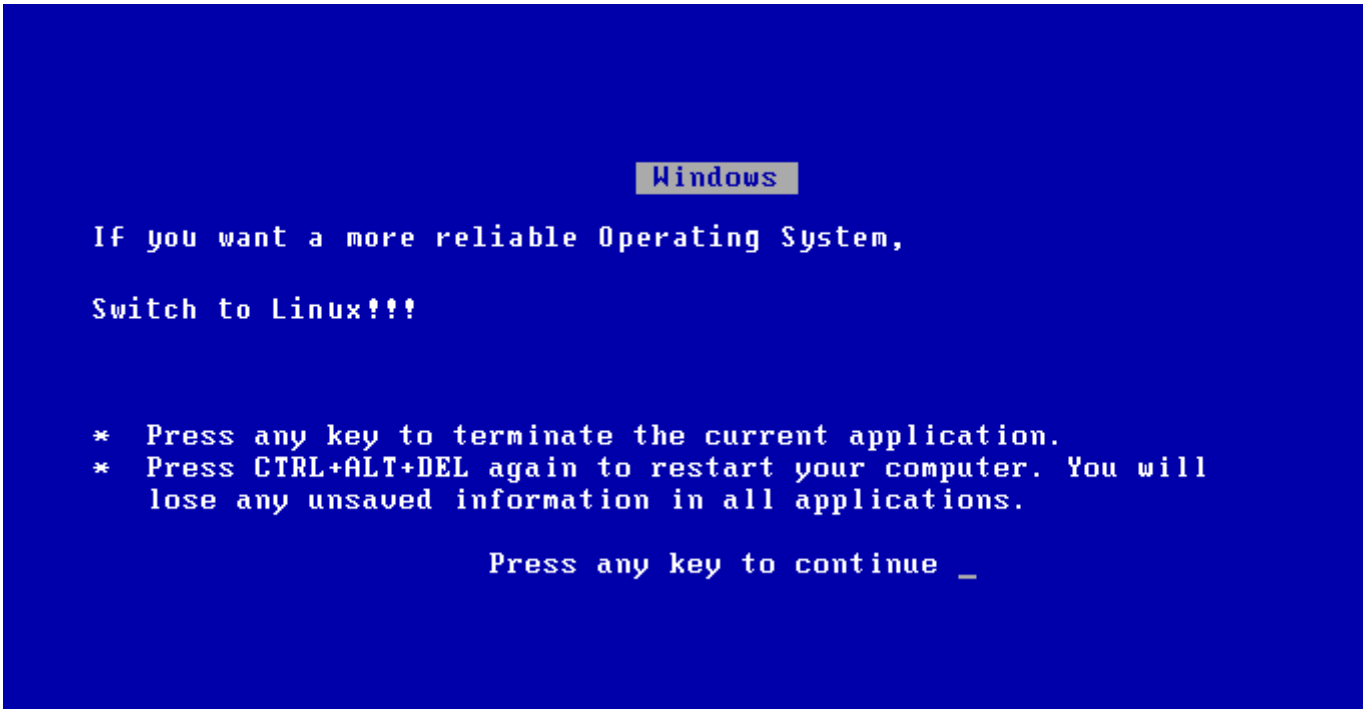
Linux is an operating system or a kernel which germinated as an idea in the mind of young and bright Linus Torvalds when he was a computer science student. He used to work on the UNIX OS (proprietary software) and thought that it needed improvements.

However, when his suggestions were rejected by the designers of UNIX, he thought of launching an OS which will be receptive to changes, modifications suggested by its users.

The benefits of using Linux

Linux now enjoys popularity at its prime, and it's famous among programmers as well as regular computer users around the world. Its main benefits are -

It offers a free operating system. You do not have to shell hundreds of dollars to get the OS like Windows!



- Being open-source, anyone with programming knowledge can modify it.
- The Linux operating systems now offer millions of programs/applications to choose from, most of them free!
- Once you have Linux installed you no longer need an antivirus! Linux is a highly secure system. More so, there is a global development community constantly looking at ways to enhance its security. With each upgrade, the OS becomes more secure and robust
- Linux is the OS of choice for Server environments due to its stability and reliability (Mega-companies like Amazon, Facebook, and Google use Linux for their Servers). A Linux based server could run non-stop without a reboot for years on end.

I am asked to Learn Unix? Then why Linux?

UNIX is called the mother of operating systems which laid out the foundation to Linux.Unix is designed mainly for mainframes and is in enterprises and universities. But, the commands used on both the operating systems are usually the same. There is not much difference between UNIX and Linux. Though they might seem different, at the core, they are essentially the same. Since Linux is a clone of UNIX.

How many distributions are out there?

There are hundreds of Linux operating systems or Distributions available these days. Many of them are designed with a specific purpose in mind. For example, to run a web server or to run on network switches like routers, modems, etc.

The latest example of one of the most popular smartphone-based Linux Distribution is Android!

Many of these Distributions are built to offer excellent personal computing.

Here, are a few popular Linux Distributions (also called Linux Distro) -

Linux Distribution	Name	Description
archlinux	Arch	This Linux Distro is popular amongst Developers. It is an independently developed system. It is designed for users who go for a do-it-yourself approach.
CentOS	CentOS	It is one of the most used Linux Distribution for enterprise and web servers. It is a free enterprise class Operating system and is based heavily on Red Hat enterprise Distro.
debian	Debian	Debian is a stable and popular non-commercial Linux distribution. It is widely used as a desktop Linux Distro and is user-oriented. It strictly acts within the Linux protocols.
	Fedora	Another Linux kernel based Distro, Fedora is supported by the Fedora project, an endeavor by Red Hat. It is popular among desktop users. Its versions are known for their short life cycle.

Linux Distribution	Name	Description
	Gentoo	It is a source based Distribution which means that you need to configure the code on your system before you can install it. It is not for Linux beginners, but it is sure fun for experienced users.
	LinuxMint	It is one of the most popular Desktop Distributions available out there. It launched in 2006 and is now considered to be the fourth most used Operating system in the computing world.
	OpenSUSE	It is an easy to use and a good alternative to MS Windows. It can be easily set up and can also run on small computers with obsolete configurations.
	RedHat enterprise	Another popular enterprise based Linux Distribution is Red Hat Enterprise.It has evolved from Red Hat Linux which was discontinued in 2004. It is a commercial Distro and very popular among its clientele.
	Slackware	Slackware is one of the oldest Linux kernel based OS's. It is another easy desktop Distribution. It aims at being a 'Unix like' OS with minimal changes to its kernel.
	Ubuntu	This is the third most popular desktop operating system after Microsoft Windows and Apple Mac OS. It is based on the Debian Linux Distribution, and it is known as its desktop environment.

Linux vs Windows: What's the Difference?

It's time to make the big switch from your Windows or Mac OS operating system.

Mac OS uses a UNIX core. Your switch from Mac OS to Linux will be relatively smooth.

It's the Windows users who will need some adjusting. In this tutorial will introduce the Linux OS and compare it with Windows.

Windows Vs. Linux:

Windows	Linux
Windows uses different data drives like C: D: E to stored files and folders.	Unix/Linux uses a tree like a hierarchical file system.
Windows has different drives like C: D: E	There are no drives in Linux
Hard drives, CD-ROMs, printers are considered as devices	Peripherals like hard drives, CD-ROMs, printers are also considered files in Linux/Unix
There are 4 types of user account types 1) Administrator, 2) Standard, 3) Child, 4) Guest	There are 3 types of user account types 1) Regular, 2) Root and 3) Service Account
Administrator user has all administrative privileges of computers.	Root user is the super user and has all administrative privileges.
In Windows, you cannot have 2 files with the same name in the same folder	Linux file naming convention is case sensitive. Thus, sample and SAMPLE are 2 different files in Linux/Unix operating system.
In windows, My Documents is default home directory.	For every user /home/username directory is created which is called his home directory.

KEY DIFFERENCE

- Linux is an open source operating system so user can change source code as per requirement whereas Windows OS is a commercial operating system so user doesn't have access to source code.
- Linux is very well secure as it is easy to detect bugs and fix whereas Windows has a huge user base, so it becomes a target of hackers to attack windows system.
- Linux runs faster even with older hardware whereas windows are slower compared to Linux.
- Linux peripherals like hard drives, CD-ROMs, printers are considered files whereas Windows, hard drives, CD-ROMs, printers are considered as devices
- Linux files are ordered in a tree structure starting with the root directory whereas in Windows, files are stored in folders on different data drives like C: D: E:
- In Linux you can have 2 files with the same name in the same directory while in Windows, you cannot have 2 files with the same name in the same folder.
- In Linux you would find the system and program files in different directories whereas in Windows, system and program files are usually saved in C: drive.

Why learn Command Line Interface?

Even though the world is moving to GUI based systems, CLI has its specific uses and is widely used in scripting and server administration. Let's look at it some compelling uses -

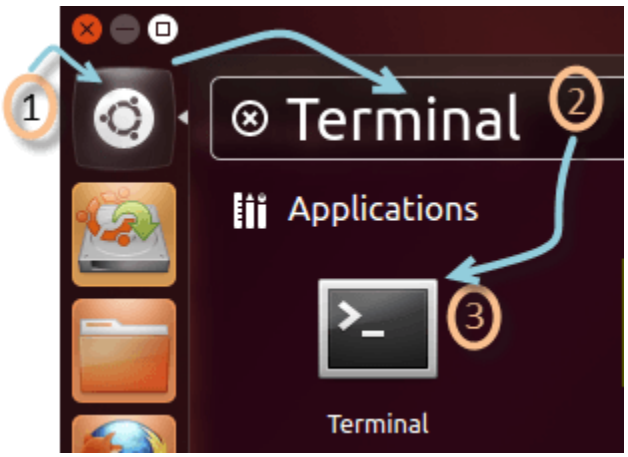
- Comparatively, Commands offer more options & are flexible.

- Moving, renaming 1000's of the file in GUI will be time-consuming (Using Control /Shift to select multiple files), while in CLI, using regular expressions so can do the same task with a single command.
- CLI load fast and do not consume RAM compared to GUI. In crunch scenarios this matters.

Launching the CLI on Ubuntu

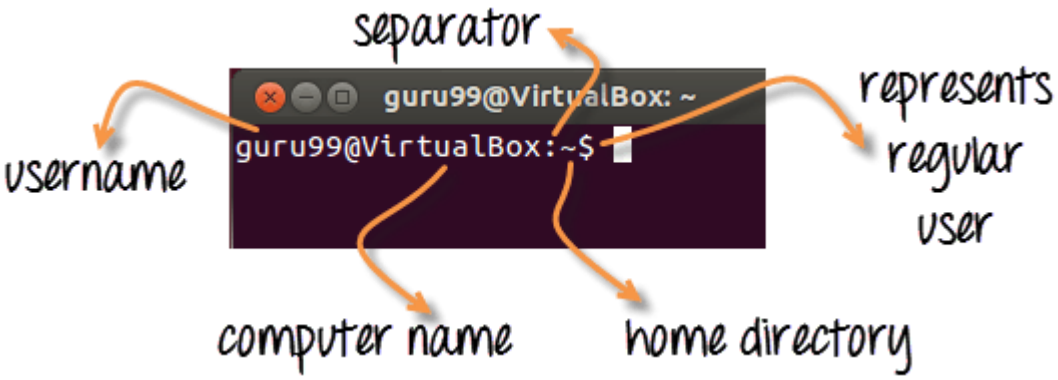
There are 2 ways to launch the terminal.

1) Go to the Dash and type terminal



2) Or you can press CTRL + Alt + T to launch the Terminal

Once you launch the CLI (Terminal), you would find something as guru99@VirtualBox(see image) written on it.

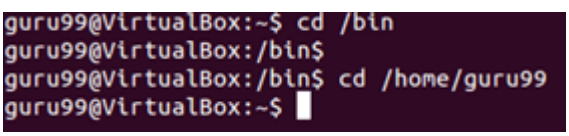


1) The first part of this line is the name of the user (bob, tom, ubuntu, home...)

2) The second part is the computer name or the host name. The hostname helps identify a computer over the network. In a server environment, host-name becomes important.

3) The ':' is a simple separator

4) The tilde '~' sign shows that the user in working in the home directory. If you change the directory, this sign will vanish.



In the above illustration, we have moved from the /home directory to /bin using the 'cd' command. The ~ sign does not display while working in /bin directory. It appears while moving back to the home directory.

5) The '$' sign suggests that you are working as a regular user in Linux. While working as a root user, '#' is displayed.

Why Linux for HADOOP??

There are many answers which point to just these:
Being opensource, meant for each other etc

Although Apache Hadoop (2.2 onwards) supports windows, its still not deployed in that configuration.

Reasons:

1. Maturity of Apache Hadoop on Windows
2. Cost effectiveness (imagine few k nodes X licensing)
3. That too for running an unsupported software!
4. This causes enterprises to shy away from Hadoop itself(CDH, HDP, MapR, Pivotal HD are filling that gap).
5. Those organizations who don't shy away from Apache Hadoop, don't do that for Linux either.

I think reasons 2 to 4 contribute to become reason 1

LINUX COMMANDS