Linux Overview

Written by:

Josh Lowery

The Senator Patrick Leahy Center for Digital Investigation

Champlain College

October 29, 2012

Disclaimer:

This document contains information based on research that has been gathered by employee(s) of The Senator Patrick Leahy Center for Digital Investigation (LCDI). The data contained in this project is submitted voluntarily and is unaudited. Every effort has been made by LCDI to assure the accuracy and reliability of the data contained in this report. However, LCDI nor any of our employees make no representation, warranty or guarantee in connection with this report and hereby expressly disclaims any liability or responsibility for loss or damage resulting from use of this data. Information in this report can be downloaded and redistributed by any person or persons. Any redistribution must maintain the LCDI logo and any references from this report must be properly annotated.

Contents

1. Introduction

Linux is a collection of open source operating systems based off of the UNIX systems. It is gaining popularity with average computer users, rather than those who are more computer literate, and enjoy working with operating systems that allow for full customization. There are hundreds of distributions of Linux that stem from the GNU/Linux core, with the two most popular distributions based off of Debian and Red Hat. Each distribution is slightly different from other distributions, but there are aspects that tie them together. All Linux distros have similar directories and commands, allowing for a user that has experience in any distro to be able to quickly learn how to use another. The downside for investigators is that, due to variation in the distributions, the way in which a case is investigated can be changed drastically depending on which distribution is used compared to the one an investigator is used to using. The graphic interface can be changed easily, and many popular distributions have interfaces that look very similar to Windows or Mac environments. The command line is where Linux really shines, though. Nearly anything you could want to do on a Linux computer can be handled through the command line, which is one of the reasons that power users are drawn to Linux.

Linux itself is just an operating system, but there are commands and tools that are available to the variety of Linux systems which will aid in the computer forensics field, by providing simple and free tools that can produce results comparable to purchased software such as EnCase. Alongside the commands and tools specifically for Linux, there are open-sources tools that are available for popular operating systems such as Windows, Apple OS, and Linux. The Linux operating system and many of these open-source tools are included in the GNU license that guarantees quality products are available to everyone. Tools under the GNU license are numerous in functionality and purpose, including programs such as: The Sleuth Kit, File System Investigator, Net Sleuth, tcpdump, SIFT, and Volatility.

1.1 Pros vs. Cons

Here are a few advantages to Linux:

Open source – source code is released for distributions, allowing for full customization of the operating system and easy fixes for bugs.

Price – due to the open source aspect of Linux, most distributions are free as are many of the applications/programs. Linux is also able to run on older hardware while still getting results that can rival Windows and Mac, and it requires significantly less resources than other operating systems, with a number being able to run off of floppy disks or USB devices.

Customization – With some knowledge of programming, you can customize your environment to suit your needs exactly. Thanks to the availability of the source code, every aspect of the operating can be changed; it just depends on the user's intelligence and abilities.

Security –Since there are fewer Linux users, as well as a wide variety of different Linux platforms, Malware/Virus infection rates are minimal. There are also many programmers and security officials that use Linux because of its ability to almost instantly update their OS when a new way of attacking is discovered, rather than wait for Microsoft or Apple to develop and market a patch.

There are a few difficulties that may arise using Linux, such as:

Learning Curve – Using Linux competently requires knowledge of the Linux environment including: commands (different than Windows/Mac as well as different through the distros), file structure, how hardware is handled, processes, etc.. There are many books and tutorials on how to use Linux, but a general understanding of computers is necessary.

Limits on Software – Many software companies do not offer support for Linux environments, and because of this, it is up to the Linux community to write patches that will allow for some programs to work in Linux. Unfortunately, this means that less popular programs may not have the community support needed to patch and be used in Linux.

Hardware – Hardware can be viewed as both a positive and negative aspect of Linux. While it is able to use older hardware, some of the newer hardware may not have the correct drivers to use in Linux, and even if they have support in some distros, they may not be supported in every distro on the market.

Encountering Errors
When beginning to use Linux, encountering errors or being afraid of breaking something while in command line can cause users to feel like they are ruining their system. If they are unsure of what to do, they will be less willing to use the computer to its fullest potential.

1.2. Benefits of Command Line

Command line is invaluable when it comes to working with computers, and it is especially useful when using Linux distributions. Using the command line saves valuable system resources by removing the need for a graphic interface for tasks. Nearly everything that can be done on a computer can be achieved through command line, and it can be done faster and with more options than with a graphic user interface. It is very useful for a forensics investigator to know a variety of commands when performing an investigation. It is also important to have resources on hand to explain exactly what each command does, and to have a reference to other commands that the investigator may not be as familiar with.

2. Setting up an Investigation

In a Linux environment, forensically wiping a drive is possible with a single command: “dd if=/dev/zero of=/dev/sdXX bs=8M,” where the “sdXX” is the name of the device that you will be wiping. You can discover the name of the device by looking for the device in the dev directory both before and after inserting the device into the system, and "sd" can also be substituted with "hd," depending on the type of device. This command works at a low level, meaning that it works with hardware more than it does with software, and because of this, it can overwrite certain functions that can be found in software such as the master boot record of a device. By wiping the device at this low of a level, investigators can be sure of erasing all the information on the device. If more information on the dd command is needed, investigators simply need to input the command "man dd" or "info dd" to learn more uses for the command.

The next step after erasing the target drive is to image the device so that you will not tamper with the evidence on the main device. The command used in this tutorial to image a device is the same base command used to wipe the drive, though the syntax is changed slightly. The command used now will be "dd if=/dev/sdxx of=/dev/sdyy," where "sdxx" is the name of the device that will be imaged and "sdyy" is the device that will have the image placed on it (the device that was wiped previously). To determine what command to use on which device, the investigator simply needs to check the dev directory before and after attaching the device.

Using the dd command instead of EnCase is a tradeoff in some aspects when working on a case. Using the dd command allows the investigators to perform many functions using a single command with different arguments, while taking very few system resources. It is also very simple to perform if the investigator is comfortable with the command. EnCase, on the other hand, uses more resources and involves multiple steps to perform the same function, but allows for reliability and the comfort of knowing that the action will be performed correctly and on the correct device. With dd, if the wrong drive is selected, all the information could be destroyed on that device. In EnCase, the investigator can correctly choose the target and destination drive without having to worry about the program making a mistake.

3. File Structure

Linux has a vastly different file structure than Windows. Linux has a similar design, but it has a different application of the file "tree". Linux considers everything included in the system a "file," and every file can be edited with a text editor. Windows views files, programs, and equipment differently. The way that users can interact with the file system is different in Windows because users are able to access nearly every drive and program with very little trouble. In essence, every user has the ability to perform administrator tasks in a Windows environment. In Linux, the opposite is true; a normal user can only do simple tasks and run programs that the admin ("root" user) allows them too. If they try to run a program above their privileges, they must enter the command "sudo" or "su,"enter in an administrative password, and be in a special group that allows them to use the sudo command. Since most computer users have Windows as their primary operating system, this tutorial is going to assume that a basic knowledge of the Windows file structure is known. The point of this section is to show where information is stored and how devices are viewed in a Linux environment to give investigators an idea where to look for evidence.

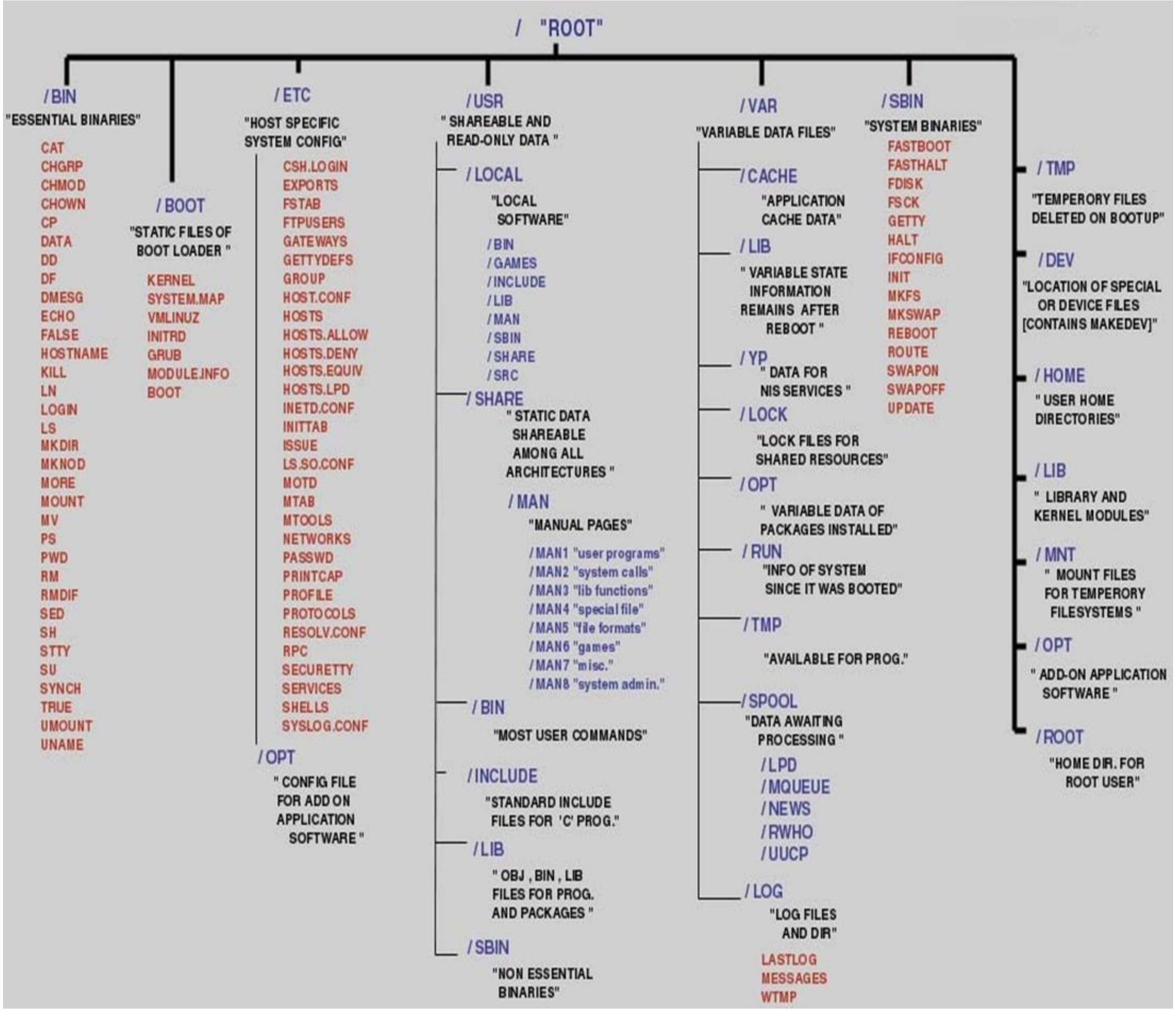

Figure 1 http://www.secguru.com/files/linux_file_structure.jpg

The image above shows an example of a Linux file structure and its layout. Notice how everything stems from the root folder. In Linux command line interface, the root directory will be shown as a forward slash "/".

The following directories are considered top-level directories and should not be confused with sub-directories that contain the same names. For example: /SBIN should not be confused with /USR/SBIN, as they contain two separate types of files. Some distributions of Linux have different directories in the top-level, but these are the core directories that make up a majority of distributions.

/BIN

/BIN is the directory that holds all of the binary files that can be used by every user, rather than a single user. It holds the main binary executable files that make up the major commands used in the Linux command line including: cat, grep, ls, mkdir, and others that will be covered later in the tutorial.

/BOOT

/BOOT is the directory that contains information necessary for booting up the operating system itself, including the kernel, boot loader, configuration files, and others. This tutorial will not cover the /BOOT directory in detail due to the advanced nature of the contents of the directory .

/ETC

The /ETC directory contains the local configuration files about the system such as passwords, services, network information, groups, and users, as well as many other functions. This tutorial will not cover most of the /ETC directory, but will touch on some the files that may prove useful in an investigation.

/USR

USR stands for “Unix System Resources”. When reading the name of this directory, many investigators infer that the user data is inside this directory, but it actually contains read-only files necessary for running programs, manual files (which will be covered later), and software that is installed in the Linux environment. /USR is very helpful when looking for what programs the user has installed.

/VAR

/VAR is a directory that contains files that are written to and read by the system during the course of normal operations. All of the files in this directory will never leave the computer unless specifically targeted by the user, and any information in this directory most likely came from the users of the computer and not from an outside influence. The files also contain information that can link files to programs and the opposite. The tutorial will not cover the information contained in this folder in detail.

/SBIN

/SBIN is a directory that contains administrative programs such as ifconfig, fdisk, route, and init. These tools and programs allow for the maintenance and administration of Linux systems. This directory is very important, and many of the programs in /SBIN will be covered later in the tutorial.

/TMP

/TMP stands for temporary and is filled with files that are erased when the system is shut down. It is safe to think of this folder as volatile in the same way that RAM is. If you are collecting evidence from a suspect and the system is still running, it is preferential to copy all the files out from this directory before shutting down the system as it may contain very important information. Though in most cases the /TMP files are erased at shutdown, there are cases in which some files are retained and the information is kept through multiple sessions. When collecting information from the /TMP file, the investigator should look how the specific operating system stores information in /TMP.

/DEV

The /DEV directory is where all the devices that make up the computer hold their information. In Windows you can only interact with devices in the way that the manufactures and Windows allow you to, but in Linux everything is a file, meaning that every folder contains information that is the heart of a device and its drivers. Every partition is visible, as are speakers, network cards, and removable media.

/HOME

/HOME can be thought of as the Users folder in Windows. It contains the information that the users have saved on the computer, and all the documents and saved files that users have will be located by default in this drive. Each user has their own home folder, and if you are in the root account you will be able to see everyone that has an account on the machine. Normal users will not be able to see other users' folders without root access. This is also where users are able to install programs that they only have access to. This directory will be where most of the evidence will be found in an investigation.

/LIB

The /LIB directory contains kernel modules and other files that are needed to boot the system and run commands. The files contained in this directory can be seen as the Linux equivalent to the Windows .dll files. They are essential for running the system and all devices. This directory should not be modified by investigators, but can be searched for evidence if it is assumed that the suspect has hidden files inside this directory. Use extreme caution when doing anything in this directory, as changing certain files in anyway can result in catastrophic damage to the system.

/MNT

The /MNT directory is where all the mounted devices are listed. If the investigator is going through an imaged drive, he or she will not gain much information here, unless the hard drive has multiple partitions. If the investigator is running an investigation on a live machine, this directory could contain a wealth of information since this is where all connected devices are listed such as USB devices and external media. This could be thought of as the equivalent to the My Computer window in Windows.

/OPT

/OPT is where third-party software is stored in Linux. Programs that are not included in the distribution by default are known as third party software collectively. This includes anything that was used to get the apt-get command, as well as anything downloaded and installed. Information can be found here if the programs installed log information to be stored.

/ROOT

/ROOT is the home directory for the root user. This directory has the same attributes as /HOME, except that it is where the administrator keeps his or her files. It can also be seen as “/”.

4. Basic Linux Environment

4.1 Commands

While in a Linux environment, either in graphic or text mode, there are simple ways to get around. This guide will discuss helpful commands, BASH keys, built-in documents, and basic file structure. Although there are a large variety of Linux distributions, this guide will focus on the common features that most distributions have. The commands have a basic syntax which correspond to the name of the “program” that they are running. In addition, the user can add “arguments,” which specify what the command does. Arguments can be as simple as adding a “–H” after a command or “-mtime +x,” depending on the command used and the available arguments.

The first thing users will notice when they open up a terminal window is the “*user@host:/~$”* in the corner of the window. This lists the user, the hostname of the local machine, and the location of where the user is currently. In the example, user is the name of the person logged into the terminal at the moment, host is the name of the machine currently being used, and ~ is the current location of the user (~ represents the home folder of the current user).

Before any commands are used, the user should first know the full abilities of commands that they are going to use. To achieve this, the user should use the “man” and “info” commands followed by the command that they want to learn about. This will bring up the manual and info pages, respectively, which both contain useful information about the commands, including the available arguments.

“ls” is one of the very first commands that new Linux users should learn. “ls” stands for list, which will list all folders and files located in the current directory. It is the equivalent to the “dir” command in windows. There are arguments that are able to be used with the “ls” command such as “-a,” which will list everything including hidden files. To learn the full capabilities of “ls,” users should consult the man page for “ls”.

Moving around the Linux environment in the terminal is the same as in Windows: use the command “cd” (change directory). The syntax is the same as in windows and can easily be picked up by new users. A command that compliments “cd” and can be used to help to locate the exact position in the file structure is the “pwd” command (Print Working Directory).

To search for a specific file, investigators will need to use the “find” command. The find command is one of the most helpful commands an investigator can use because it allows the investigator to search through every directory for a specific filename, files within a specific period of time, as well as a variety of other options that make it an invaluable tool for any case. The find command can be seen as the same as a file search in EnCase or another software program. Some of the basic syntaxes that will help in an investigation are:

find / -name test (alternatively you can use –iname to ignore letter case)

Searches for a file named test

find /home –user suspect

Searches files in a suspect’s home folder

find /home/suspect –mtime +30

Searches for files in a suspect’s home folder modified in the last 30 days

Additionally, the above syntaxes can be modified to the needs of the investigator by changing the words used. Two examples are changing home to usr to look through the usr directory and changing the +30 to +365 to search for the last year. There are many more arguments that are available by looking at the find man page.

A command that is similar to the “find” command is “grep.” Grep is a tool that searches for strings of text. This is useful when looking for a specific keyword or string of characters within files. There are multiple ways to use the grep command including specifying a specific file, multiple files, insensitive searches, and additional ways that are listed in the man page. Examples of grep commands are:

grep “string” file1

This command will search file1 for the word string and list every line that contains that word

*grep “String” file1**

This command will search for the text “String” in all files that start with file1. This is helpful when looking for a string within files that have similar naming structures, such as templates, budgets, and documents that are constantly updated. The asterisk represents a wildcard placeholder. Any file that starts with the name file1 will be included in the search including names such as file14 or file1-briefing.

grep –I "string" file1

This will search file1 for the text string, and it will also search for String, StRing, or any other combination. The –I tells the system that the text being searched for is case insensitive.

If a file is found that the investigator wants to read, simply use the command "cat" (Concatenation) followed by the file name. An example would be "cat file1" to read what is written inside the file named file1.

"mkdir" creates a directory inside the directory that the user is currently in, unless specified by the user. "mkdir penguin" creates a directory named penguin in the current directory, while "mkdir /home/investigator penguin" creates a directory named penguin in the investigator's home directory.

4.2 BASH Keys

BASH keys are keys and key combinations that will make moving around Linux faster and easier for users. They can be considered shortcut keys much like using the windows key in Windows. A few keys are:

- Ctrl and A will move the cursor to the beginning of the command line
- Ctrl and C will end a running program
- Ctrl and E brings you to the end of the command line
- Ctrl and H generates a backspace
- Ctrl and L will clear the terminal
- Pressing Tab once will finish the possible filename you are typing, or cycle through the options that fit what you have already typed
- Pressing Tab twice will show all possible commands or files that a user is trying to find

These are just a few BASH keys that are available in Linux. These few commands will help you understand how Linux functions and will make movement quicker for easier use.

5. Forensics Tools

There are many forensics tools that are available in the Linux environment that are not available to other operating systems. This could be because of the way that they are programed or that were created under the GNU license. http://www2.opensourceforensics.org/home has many open source tools that are available for review and download. Some programs such as The Sleuth Kit and Autopsy have reputations in the forensics field as reliable forms of open source forensics investigation tools. SANS is a forensics company that releases many papers and studies on computer forensics, as well as their own operating system named SIFT, which runs on a Linux distribution. There are many investigators and general computer users that do not see a benefit to using open source tools in a professional environment, but these programs are helpful to forensics labs that have a small budget or to investigators that are interested in and able to customize their software to meet their needs. Choosing a program like Autopsy over something like EnCase is preferential in situations where purchasing a program that costs as much as EnCase is not possible or if the investigator does not need all the functions that EnCase provides and does not want to spend all the money on a program that offers more than they need.

6. Conclusion

Linux is growing in popularity and functions as time goes on and in a few years will be recognized by the average user as a reliable operating system. It will no longer be “that other OS” besides Windows and Apple if it continues to gain popularity at its current rate. The amount of forensics information that is available on the Linux system is amazing if the investigator knows where to look for it. Linux can be used as a powerful tool in forensics investigations and has the ability to change the face of computer forensics as we know it. Log files are created for everything users do, as well as reports that are sent whenever there are unsuccessful login attempts and errors. Linux does an exceptional job at logging nearly everything that happens on the machine and keeps a record of commands that are entered into the terminal. This allows investigators to go back and find if the suspect has entered in commands that would either hinder or help the investigation.