Linux Survival Skills

Bob Wakefield
Principal
bob@MassStreet.net
Twitter: @MassStreet

Who is Mass Street?

- Boutique data consultancy
- Hortonworks Partner
- Big data training

Mass Street Analytics

Bob's Background

- IT professional 16 years
- Background is mostly Windows tech
- Education
  - BS Business Admin (MIS) from KState
  - MBA (finance concentration) from KU
  - Coursework in Mathematics at Washburn
  - Graduate certificate data science from Rockhurst
- Addicted to everything data

Administration

- The history of the group
- Day of the week will rotate between Tues, Wed and Thurs
- Weekends for larger topics
- Location?

Mass Street Analytics

This evening's objectives

- NOT to become masters of the Linux OS
- Just be able to move around in Linux
- Linux is the gateway drug to big data

Where Can I Get More Info?

- Udemy!
  - Linux for Beginners by Jason Cannon
  - Whole series on Linux
- Stack Overflow (surprisingly)

Linux: A quick history

- Created by Linus Tovalds et al and released to the wild in 1991. v1 released in 1994.
- Not mini Unix!
- Linux 2002 = yuck
- Linux 2015 = pretty awesome

The Many Flavors of Linux

- There are many Linux distros and each one is a little (or a lot) different.
- DistroWatch.com
- Popular Distros
  - Ubuntu = Linux for the “home”
  - Red Hat (RHEL) = Linux “pro”
  - CentOS = Free RHEL
- Hadoop distros run on RHEL and CentOS (and Windows but meh)
- Just say no to GUIs!

Getting Your VM Set Up

Package Managers

- Ubuntu = apt-git
- CentOS = yum

Let's update and install some software

- Yum update
- Yum install nano

Linux Shell Basics

- [user@server current directory] $|=
- $ = normal user
- = = super user
- Up and down arrows will show command history
- No recycle bin! Changes are permanent.

Linux Shell Basics

- ~ = this accounts home directory
- / = root

Mass Street Analytics

Directory Structure

- / = “Root,” the top of the file system hierarchy.
- /bin = Binaries and other executable programs.
- /etc = System configuration files.
- /home = User Home directories (like Windows documents)
- /opt = Optional or third party software.
- /tmp = Temporary space, typically cleared on reboot.
- /usr = User related programs.
- /var = Variable data, most notably log files.

Useful Linux Commands

- ls - Lists directory contents.
- cd - Changes the current directory.
- pwd - Displays the present working directory.
- man - Displays the online manual.
- exit/logout/ ctrl-d - Exits the shell or your current session.
- clear - Clears the screen.

Useful Linux Commands

- ls - Lists directory contents.
  - Ls – l – long listing
- cd - Changes the current directory.
- pwd - Displays the present working directory.
- man - Displays the online manual.
- exit/logout/ ctrl-d - Exits the shell or your current session.
- clear - Clears the screen.

Navigation

- Cd (nothing) = go home
- CD . = this directory
- CD .. = parent directory
- CD - = previous directory
- CD /bob = change to the bob directory

File Operations

- Shhhh! I don't use these.
- So why are we learning them?
- A lot of HDFS commands are the same as linux commands
- You cannot interact with HDFS using WinSCP or another FTP program

File Operations

- mkdir directory = Create a directory
- rm file = Remove (delete) file
- rm -r directory = Remove the directory and its contents
- recursively
- rm -f file = Force removal of file without prompting for
- confirmation
- rm -rf directory = Forcefully remove directory recursively
- mv file1 file2 = Rename or move file1 to file2 . If file2 is an existing directory, move file1 into directory file2

File Permissions

- Page 5 of cheat sheet
- Permissions
  - R = read
  - W = write
  - X execute
- Permission Categories
  - U = user that owns the file
  - G = the users in that group
  - O = other not the user and current user is not in group

File Permissions Decrypted

- rw-r--rwx
- User/Group/Other
- User has read write
- Group has read only
- Other has read write execute

Changing Permissions

- We're going to learn the easy way.
- Covert symbolic to base 10
  - Three categories UGO
  - Three positions rwx
  - rwx in binary = 111 (rw- = 110)
  - 111 in base 10 = 7 (4 + 2 + 1) (110 = 6)
- rwxrwxrwx = 777 (don't try this at home)
- Chmod 777 filename

Things We Didn't Tackle

- Finding files
- Gzip (compression)
- Tar (archive)
- SCP (secure copy over network)
- Wget (download file from web)
- Sudo (act as a user)
- Ssh (connect securely to another machine)

On Your Own