Introduction to Linux

Basic GNU/Linux Concepts

First: a small digression

What is parallel programming, and how does in enter into today's discussions?

"The art and science of formulating a solution to complex problems so that the solution takes every advantage of all resources that can be exercised independently, simultaneously, but also coordinated through control and data sharing"

Pararallelism

- Aren't computers fast enough already?

	ENIAC	Intel Core Duo chip
debut	1946	2006 (discontinued 2011)
performance	5,000 additions/s	21.6 billion ops/s
power	170,000 W	31 W
weight	28 tons	few grams
size	80' x 8'	90 sq mm
innards	17,840 vacuum tubes	151.6 M transistors
price	$487,000	$637

$60 online

- Are there limits to speeding up computers?

Consider a simple problem: add numbers

1 + 2 + 3 + 4 + 5 + 6 + 7 + 8

7 operations done in sequence,

Rearrange operations slightly

1 + 2 3 + 4 5 + 6 7 + 8

3 + 7 11 + 15

10 + 26

7 operations, but several can be done in parallel so the time it takes to complete the task is …

Performance measure

time to completion is more important than number of operations

with eight numbers we went from 7 timesteps to 3.

saved 57% is that so impressive?

Performance measure

time to completion is more important than number of operations

number of numbers	sequential steps (time)	parallel steps (time)	savings (exercise)
8	7	3	
32	31	5	
64	63	6	
1024	1023	10	
Million	~Million	20	

This class will….

- Provide some basic concept information for users familiar with MacOS or Windows.
- Get you familiar with Linux commands.
- Get you started in understanding command line interfaces.

Background

- Biowulf
  - is a cluster of many computers (called nodes)
    - connected with high speed communications hardware
    - using a very large shared file system accessible to every node
- Enables you to
  - use multiple computers independently at the same time
  - use *parallel* applications
- Your interaction with Biowulf
  - scripts
    - specify the programs, data files, parameters, etc.
    - submit jobs
  - you will do this on a computer running GNU/Linux

Unix and Unix-like operating systems

- Many Unix-like systems used in research since the last century:
  - Western Digital, Bell Labs, AT&T : Unix (original)
  - Digital Equipment (PDP, Dec Alpha): DEC ULTRIX
  - Silicon Graphics Inc. (now SGI): Irix
  - Sun Microsystems: SunOS/Solaris
  - IBM: Aix
  - UC Berkeley: BSD Unix (Berkeley Software Distribution)
    - FreeBSD, OpenBSD,
    - Darwin: macOS, iOS, watchOS, tvOS
- Open source community: GNU/Linux
  - NIH Biowulf cluster

Linux in Science (why?)

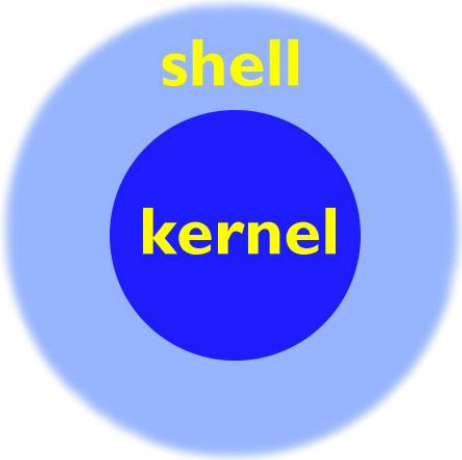
- Source code availability
  - researchers modify anything they want to improve the system
- GNU/Linux became widespread
  - liberal licensing and community expertise make it easy for vendors of HPC equipment to write drivers for their hardware
  - Wide range of tools for free
    - many languages (C, Fortran, Python, and dozens more)
    - scientific libraries (numerical analysis, linear algebra, GUI toolkits)
- **Performance, functionality and portability**
  - e.g.: many things you can write on a Mac will also work under Linux
  - anything that works on one Linux box is going to work on another one … *probably* …

Concepts: Kernel

- Operating system "kernel" is the core software used to "talk" to computer hardware
- Is a core and modular system of drivers used to create a standardized environment for interfacing with hardware
- Responsible for allocating memory and time to system and user processes as well as interacting with files.

The Shell

- The shell is a program that reads what you type, and interprets that as commands that the computer will execute
  - command line interpreter *(similar to cmd on Windows)*
- As soon as you login, a dialog with a shell begins
- The shell is a program named "bash"



Linux accounts

- An account includes
  - username and password
  - home directory
  - default shell

There are too many shells

- sh
  - the original UNIX shell
- bash
  - written as a replacement/extension of sh
  - used extensively with Biowulf scripts
- csh
  - some like it, you should choose bash for scripting
- and the rest… (for information only)
  - csh
  - tcsh
  - ksh
  - zsh
  - dash

Use `ssh` to log in

- ssh stands for secure shell
  - encypted all dialog between your terminal and the remote host
  - prevents a third party from eavesdropping

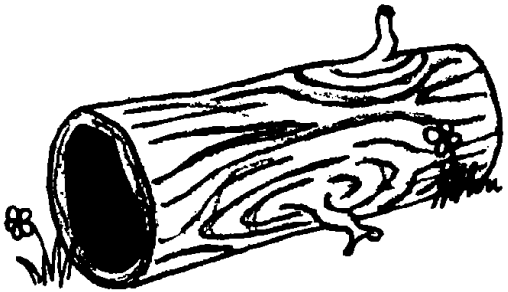
Logging in

- macOS:
  - Finder -> Applications -> Utilities -> Terminal
  - Type this: `ssh username@helix.nih.gov`
  - At the prompt, enter your password

- Windows:
  - Launch PuTTY. Under "Host Name (or IP address), type:

    `username@helix.nih.gov`

  - …and click "Open"
  - At the prompt, enter your password



More on shells

- What shell am I in?

  "`echo $SHELL`" displays your *default* shell's full pathname: `/bin/bash`

  "`echo $0`" displays your current shell: `bash`
- `$SHELL` and `$0` are shell expressions
  - *more about this and shell variables in a later course*
- List of available shells on the system can be displayed by typing
  - `chsh --list-shells`
- Use `chsh` command to change your default shell. **WARNING!** on Helix and Biowulf never change it to a shell that ends in LOCKED – you will lock yourself out of your account!

Shell preferences

- When you login, startup scripts are run to setup your environment
- For bash, you can customize your environment by adding or modifying environment variables and aliases in the `.bashrc` file in your home directory.
- Examples:

```
alias list='ls'
export PATH=$PATH:/scratch/myusername
export EDITOR=/usr/bin/vim
export PS1="[\u@\h \w \# ]"
set -o noclobber
```

Summary of Linux commands

awk	allows manipulation of text
bg	place suspended job into background
cat	view contents of a file
cd	change directory
chmod	change permissions on a file/directory
cp	copy a file
cut	extract a field of data from text output
echo	output text to the terminal or to a file
emacs	text editor
fg	bring suspended job to foreground
file	display file type
find	search for files
grep	search a file or command output for a pattern
head	view beginning of file
history	display list of most recent commands
less	scroll forward or back through a file
ln	create a link to a file
ls	list files in a directory
man	view information about a command
mkdir	make directory
more	scroll through file a page at a time
mv	change the name of a file (move)
nano/pico	text editors
printenv	display shell variables
ps	show current process information
pwd	print current working directory
rm	delete or remove a file
rmdir	delete or remove a directory
sed	stream editor
sleep	pause
sort	perform a sort of text
stat	display file status info
tail	view end of the file
touch	create an empty file or update timestamps
tr	character substitution tool
uniq	remove identical, adjacent lines
vi/vim	text editor
wc	print number of lines, words or characters
which	shows full path of a command
whoami	displays username

Our cast! (of characters)

Character	Name/Location
\	Backslash (above the enter key)
/	Slash (left of right shift key)
`	Back quote (left of the number 1, above the tab key)
|	Pipe (shift-\)
[and]	Brackets (left of the backslash)
{ and }	Braces or “curly” brackets (shift-[and shift-])
< and >	Angle brackets (left of the right shift key)
~	Tilde (shift-`)
! @ # $ % ^ & * ( )	Bang/exclamation mark, at sign, hash, ($) dollar percent, caret, ampersand, asterisk/star left and right parenthesis

Linux Command Basics

- Linux commands are case-sensitive

`ls` is not the same as `LS`

- Linux commands may have options that come after the command that start with a "–" and followed by a letter or "- -" and a word:

```
$ ls –r
$ ls --reverse
```

- Linux commands may allow for arguments:

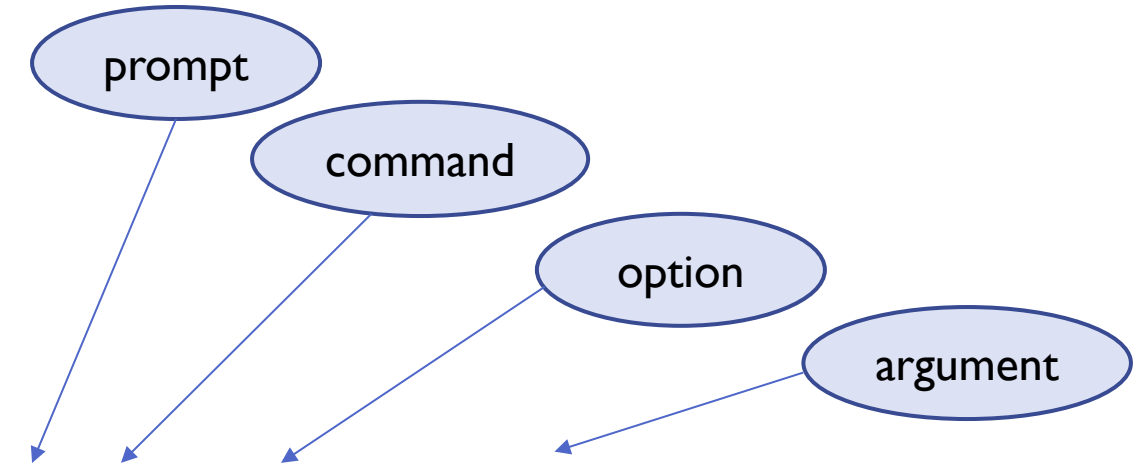
```
$ ls /scratch
```

- You can run more than one command on the same line by separating the commands with a semicolon (;)

```
$ date; ls
```

- Most Linux commands have a manual page or help to describe how they can be used….more about this later!

Linux Command Example




```
$ ls  -r /home/$USER
```

Exercise #1: First commands

```
[username@helix ~]$
```

- Type "`whoami`", press Enter
  - your user id
- Type "`pwd`" press Enter
  - *your working directory*

What do the following show?

- `echo $HOME`
- `echo $USER`
- `echo $PWD`

- HOME, USER and PWD are three more examples of shell variables (like SHELL). These are *evaluated* with the '$' prefix

Concepts: Files and Processes

In Linux, (almost) everything is either a file or a process. In other words, *processes* interact file *files*

- Regular files: text, data, documents, …
- Directories: special files that contain a bunch stuff to organize other files
- Devices: disks, video and audio hardware, processors, memory, I/O (USB) ports
  - look in `/dev` : special files that facilitate interactions with devices (hardware)
- Processes: things that get executed by a processor (CPU, core) and are "running" or "sleeping"
  - identified by a number in `/proc`

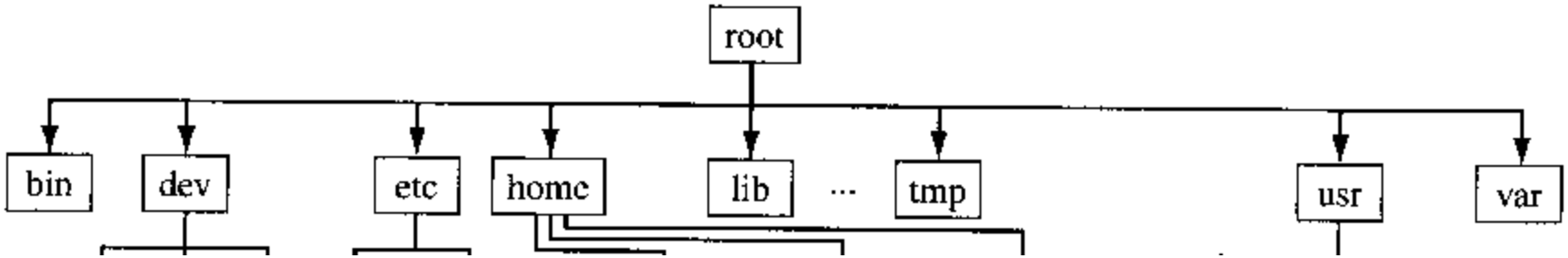
More on Files

- Each file (and directory) has a name
- The filename can contain letters, numbers and special characters
- Every file has a unique path to its location
  - e.g.: `/home/student2/read-write.txt`
- A filename MUST be unique within a directory…though files with the same filename can exist in different directories
- Filenames are also case sensitive
  - "`myfile`" and "`Myfile`" are distinct filenames
  - In practice, avoid using similar names
- Filenames can be lengthy

More on Directories

- A directory is a special type of file that can hold other files
- The "working directory" is the directory with which your shell is currently associated…where you currently are! When you first login, you will normally be in your home directory, /home/username
- Use the **pwd** command to print working directory
- Special directory notations:
  - **.** refers to the current working directory
  - **..** refers to the parent directory
    - the parent directory of **/home/username** would be **/home**

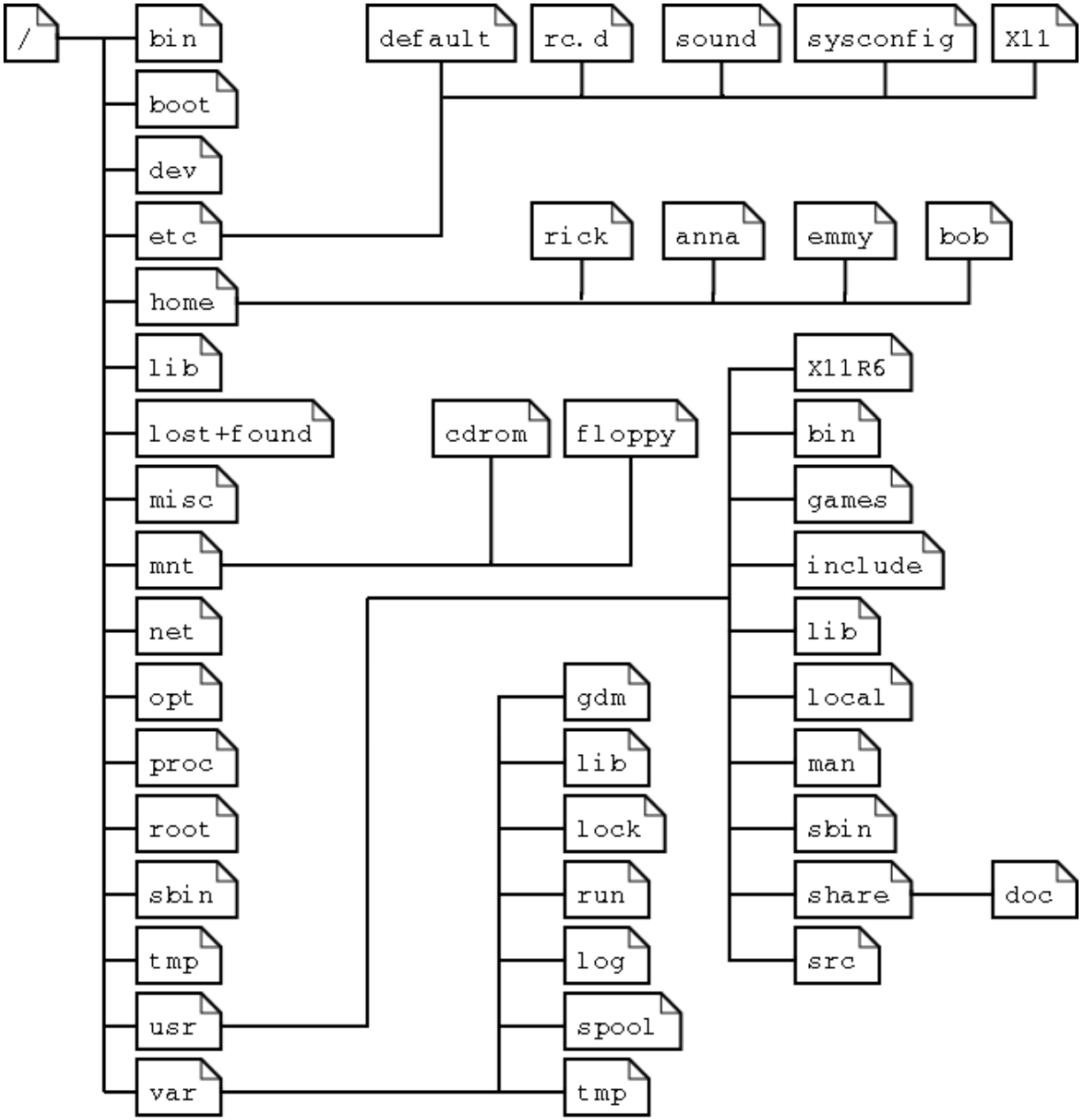
Concepts: The File System



- Linux file systems are arranged in a tree structure. There is a single root node called `'/'`
  - Note to Windows users: there are no drive letters here.
- Mapping disks to file systems is transparent to the user. Any disk can have multiple file systems and any file system can be "mounted" in any directory. Normal user activity does not care about the hardware layout of disk storage devices

Directory Structure

`/`	root
`/bin`	bare essential commands
`/boot`	OS Kernels
`/dev`	hardware device drivers
`/etc`	system files, configuration
`/home`	user home directories
`/lib`	Libraries needed by the system
`/opt`	3rd party applications
`/proc`	Running process IDs and other hardware stuff
`/sbin`	administrative commands
`/tmp`	temporary space
`/usr`	operating system applications
`/var`	system written potentially volatile files



File system layout

Common directory layout on a Linux system

Note that /var is frequently a mount-point to a separate file system. This is often true of /home, /tmp and /usr as well.

On Helix/Biowulf, /home is on a network file system as are data directories

`cd` and `ls` commands

- The **`cd`** command is used to change directory location. Without an argument, cd takes you to your home directory
- The **`ls`** command is used to list the files in a directory. Like many Linux commands, it can take a number of *flags* as options to change the behavior of the command

```
$ cd /home/$USER
$ ls
$ cd /etc
$ pwd
$ ls -l
$ cd ..
$ pwd
$ ls -l
$ cd
$ pwd
```

Exercise #1: ls, cd & man commands

- Type `cd` to get to your home directory
- `ls -l`
- `ls -a`
- `ls -la`
- How are the above outputs different?
- `ls -l /data/classes/linux`

ls will display the directory named in the argument, or the current working directory without it

- `cd /scratch` change to the /scratch directory
- `ls -lt`
  - How is the output ordered?
- `ls -help`

- `"man"` displays a ***user manual*** for a command
- `man ls`
  - Scroll with arrow keys
  - "`q`" to quit

pico – a simple editor

$ pico filename

The essentials:

- Just start typing – can use arrow keys to position, backspace or delete key to delete characters to the left
- Keystrokes for basic commands at bottom of the screen
- ^G – help screen (^C to exit help)
- ^O – save the file
- ^W – search for a string
- ^X – exit pico

```
$ pico  --help
```

Exercise #2: Edit a file using pico

```
$ cd  /home/$USER/LinuxClass
$ pico bashrc
```

1) Using the RIGHT arrow key, position the cursor at the end of the first line,
2) Use the Backspace key to remove the bracketed text from the first line and then simply start typing $HOME after the colon
3) Press ^W (to search), type PICOPATH and hit Enter – this should place you on the last line
4) Hit ^E to get to the end of the line
5) Use the Backspace key to remove everything after the '=' sign and type '/usr/bin/pico'
6) Use the up & the right arrow keys to get to the @ on the 2nd line
7) Backspace to remove <USERNAME> and type your username
8) Use the down arrow key to get to the 3rd line
9) Hit ^K to cut the 3rd line
10) Hit the Up arrow to get to the 1st line & ^A to get to the start of the line
11) Hit ^U to paste the text – the 3rd line should now be the first
12) Hit ^X to exit – type Yes to save the file when prompted and hit Enter when prompted for the name

```
$ cat bashrc
```

Users and Groups

- Users are associated with a unique user identification (UID) number that the system uses internally
- Users can be real people
- Users can be system entities
- Users can be herded via groups
- Groups also are associated with a unique group identification (GID) number by the system
- Groups allow multiple users to access/share the same files

Ownership & Permissions

- Linux systems are multi-user environments that allow users to create files, run programs and share data.
- Files and directories have two types of ownership – the user and group. A Linux group consists of one or more users.

- Files and directories have three types of access permissions:

read permission (r)

write permission (w)

execute permission (x)

- Every files and directory has permissions for three levels or entities of permissions:

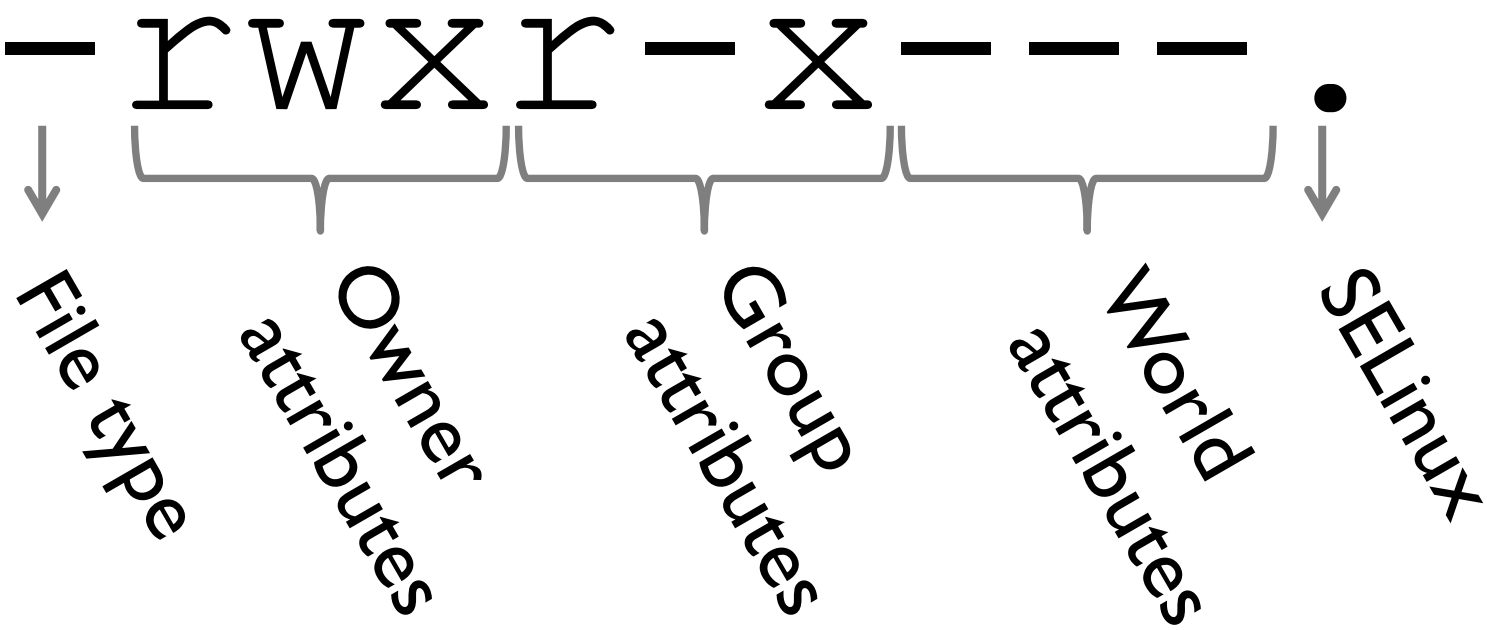
a) user or owner

b) group (one or more users)

c) others or world

Permissions triplets

Each triplet indicates the access permissions for that level – in the example below, the user/owner has read, write & execute permission, other group members only have read and execute permissions and all others have no access permissions.



Long List Output Explained (a little)

- From left to right:

Unix permissions

Hard links

Owner

Group ownership

File size in bytes

Modification date

Name of file

```
$ ls -la
drwxrwx---  104 patkus    staff   110592 Jan  6 13:02 .
drwxr-xr-x 2510 root      root    196608 Jan  6 12:58 ..
-rw-r--r--    1 patkus    patkus    1051 May  8  2012 ad-week
-rwxr--r--    1 patkus    staff      239 May 11  2007 alias.pl
-rw-r-----    1 patkus    staff     1185 Jun 22  1998 bp.txt
-rwxr-xr-x    1 root      root     27320 Mar 29  2012 getpass.awk
-rw-rw-r--    1 susanc    staff    20529 Aug  7  2009 httpd.conf
-rwxr--r--    1 root      staff   136236 Sep 10  2013 memcon
drwxr-x---    2 patkus    staff     4096 Jun 24  2010 misc
drwx------    3 patkus    staff     4096 Jun 24  2008 test
-rwx------    1 patkus    staff      493 Feb 10  2009 unlock
-rw-r-----    1 patkus    staff       38 Oct 20  2010 world.c
-rwxr-x---    1 patkus    staff     6703 Jan  8  2013 world.exe
-rwxrwx---    1 patkus    staff     2350 May 22  2009 year.pl
lrwxrwxrwx    1 patkus    staff        7 Jan  6 15:30 year2 -> year.pl
```

Special Directories:

. is the current working directory

. . is the "parent" directory, one level "up"

Permissions described:

File Type:

-	regular file
d	directory
l	symlink
b	block device
c	character device
p	named pipe
s	socket

Permissions

r	read
w	write
x	execute

Special values:

x	executable
s or t	executable and setuid/setgid/sticky
S or T	setuid/setgid or sticky, but not executable.

Changing Permissions and Ownership

- Use "**chmod**" to change the file permissions:

chmod [ugoa][+/-][rwx] filename

where u=owner, g=group, o=others or world and a=all three

For example, to provide group read access to a file:

$ chmod g+r myfile

Or to remove file access to another than the owner or group members (in other words, others):

$ chmod o-rwx myfile

- The "**chown**" command is used to change file ownership and the "**chgrp**" command can change group ownership of a file. As a regular user, you can not change the ownership of a file, but you can change the group ownership if you are a member of the group to which you are changing the group ownership

- Use the –R option on any of the above to recursively make changes on a directory of files

Exercise #3 pre-exercise

- We'll talk about some of these commands shortly, but you each need to make a copy of files needed for the rest of class.

- First go to your home directory

```
$ cd  /home/$USER
```

- Make a directory called "LinuxClass" in your HOME and go into that directory:

```
$ mkdir  LinuxClass
$ cd  LinuxClass
```

- Copy the files from the exercise file to your directory:

```
$ cp   -r   /data/classes/linux/*  .
$ ls  -l
```

- Create a shell variable to your class directory:

```
$ export  class=$PWD
$ cd  $class
```

Exercise #3: File Permissions

Read Permissions

The cat command displays contents of a file

```
$ cd /home/$USER/LinuxClass
$ cat  read-write.txt
```

Change the read permission

```
$ chmod  u-r  read-write.txt
$ cat  read-write.txt
```

What happened? Now restore the read permission

```
$ chmod  u+r  read-write.txt
$ cat  read-write.txt
```

Execute Permissions

```
$ cd /home/$USER/LinuxClass
```

Run the myhostname file to see the system name

```
$ ./myhostname
```

Change the file permissions:

```
$ chmod  u-x  myhostname
$ ./myhostname
$ chmod  u+x  myhostname
$ ./myhostname
```

Change permissions on the directory dir-perms:

```
$ chmod  u-x  dir-perms
$ ls dir-perms
$ ls -l dir-perms
```

What happened and why?

Shell Variables

Show all currently assigned variables

```
$ printenv
HOSTNAME=helix.nih.gov
TERM=xterm
SHELL=/bin/bash
HISTSIZE=500
SSH_CLIENT=165.112.93.227 49886 22
QTDIR=/usr/lib64/qt-3.3
QTINC=/usr/lib64/qt-3.3/include
SSH_TTY=/dev/pts/286
HISTFILESIZE=500
USER=patkus
…
```

Useful predefined and important variables

HOSTNAME	System hostname
USER	Your Username
SHELL	Your shell
HOME	Home directory
PWD	Current directory
PATH	Command paths

Shell Variable Scope

Variables are available only to your immediate shell environment by default.

```
$ MYVAR="This is my var"
$ echo $MYVAR
$ printenv | grep MYVAR
```

Is MYVAR listed?

Export makes variables available to subprocesses (like /usr/bin/printenv)

```
$ MYVAR="exported"
$ printenv | grep MYVAR
$ export MYVAR
$ printenv | grep MYVAR
```

Or

```
$ export MYVAR2=“exported2”
$ printenv | grep MYVAR
```

Your PATH

Execution path

In BASH, execution of a program happens when you enter the program name. Your PATH variable keeps you from having to enter the full path to the program

```
$ echo $PATH

$ which date
$ which whoami
$ which perl
```

Setting your PATH

```
$ PATH=/home/$USER:$PATH
$ export PATH
```

Or

```
$ export
PATH=/home/$USER:$PATH

$ echo $PATH
```

Wildcards

* (asterisk) Match zero or more characters

```
$ ls  bear*
bears  bears7  bears_chicago
$ ls  *bear*
bears bears7 bears_chicago polarbears
`
```

? (question mark) Match exactly one character

```
$ ls  bears?
bears7
```

Special Keys

Tab allows for command auto-completion

Arrow keys allow you to:

a) horizontal arrows: edit command without deleting and retyping everything

b) vertical arrows: browse your history of previously typed commands

Ctrl-C and ESC

- If you get into a situation where you can't seem to get back to a command line prompt, try pressing the Ctrl-C combination or the Esc key – often one of those will abort whatever you are currently doing.

- The Ctrl character is often denoted by the caret symbol:

^C **press Ctrl and c**
(no shift)

Let's go home!

- ~ is a special character (in bash) that is short-hand for your home directory (like $HOME)
- Several ways to get to your home directory:

```
$ cd ~
$ cd $HOME
$ cd /home/username
$ cd /home/$USER
$ cd
```

- Ok to use ~ and $HOME as arguments with other commands:

```
$ ls  ~/tmp
$ ls  $HOME/LinuxClass
```

Absolute and Relative paths

- The starting '/' in the directory argument explicitly spells out a pathname – specifying an absolute or full path
- No leading '/' means you are specifying a path that is relative to the current working directory.

```
$ cd  /home/username

# Absolute path:
$ cd  /home/username/tmp

# Relative path:
$ cd  tmp

# Using ~
These do the same:
$ cd  ~/tmp
$ cd  /home/username/tmp
```

What is that file?

file

Tells us what type of file it is:

e.g: text, executable, PDF, jpg, gzip, tar, directory, etc.

```
$ file read-write.txt

$ file world.exe

$ file examples
```

apropos

Apropos will search documentation for keywords. If you don't know what man page to look at, use apropos to search for potentially related material

```
$ apropos  editor

$ apropos  "text editor"
```

Cat and Echo

Use cat to display file contents to the terminal:

```
$ cat bears
$ cat bears7
$ cat bears bears7
```

"cat" is short for concatenate. The "cat" command takes one or more files and concatenates their contents to standard output.

Use echo to output arbitrary text to the terminal:

```
$ echo 'Hello World!'
$ echo without single quotes
$ echo 'output' > Myoutput
```

Output Redirection to Files

Redirect output:

```
$ cat bears > Newbears
$ cat Newbears
$ cat bears bears7 > newbears
$ cat newbears

$ echo 'Hi there!' > greeting
$ cat greeting
```

Append files:

```
$ echo 'Hi yourself!' >> greeting
$ cat newbears >> greeting
$ cat greeting
```

Exercise #4: cat and echo

cat a file to view contents

```
$ pwd
$ cd  /home/$USER/LinuxClass
$ cat  lions
$ cat  tigers
$ cat  bears
$ cat  lions tigers > animals
$ cat  animals
$ cat  bears >> animals
$ cat  animals
```

Using echo

```
$ echo my name is Chris
$ echo “my name is $USER” > myname
$ cat myname
$ echo Hello $USER >> myname
$ cat myname
```

Creating and deleting files and directories

Using touch and mkdir

To create an empty file, use the touch command:

```
$ touch  emptyfile
```

You can also create a file using an editor such as pico, nano, vi or emacs:

```
$ pico  penguinfile
```

To create a directory:

```
$ mkdir  Mydirectory
$ mkdir  2012
$ mkdir  -p  2014/Jan/stats
```

Using rm to remove files and directories

To remove a file:

```
$ rm emptyfile
$ touch myFile
$ chmod  u-rwx  myFile
$ rm myFile
$ rm -f myFile
```

To remove a directory:

```
$ rmdir Mydirectory
$ rm -r 2012
```

rmdir only works if the directory is empty!

Dangerous:

```
$ rm  -rf  *
```

Exercise #5: Creating and deleting files

Creating a file and directory

```
$ cd  /home/$USER/LinuxClass
$ echo  'I love genomic research!'  >  science
$ touch  science_project
$ mkdir  scienceclass
$ ls  -ld  science*
```

Deleting a file and directory

```
$ rm  science*
$ ls  -ld  science*
What happened?
$ rmdir  scienceclass
```

Displaying/Editing Files

more, less

head, tail

```
$ more  colleges.txt

$ less  colleges.txt

$ head  colleges.txt
```

prints out the first 10 lines by default. Can use the -n argument to specify the number of lines

```
$ tail  -20  colleges.txt
```

prints out the last 20 lines

Text editors:

Good simple editors:

- pico (pine composer)
- nano (pico clone)

More powerful and complex editors:

- vim (vi-improved)
- emacs (Editor MACroS)

Moving /Copying files

Move (mv)

Move a file or rename a file

```
$ touch football
$ ls -l football
$ mv football volleyball
$ ls -l *ball
$ mkdir sports
$ mv sports sportsNEW
$ ls -ld sports*
$ mv volleyball
sportsNEW
$ ls -la sportsNEW
```

Copy (cp)

Copy a file or directory with the cp command

```
$ echo 'Goal!!' > soccer
$ cp soccer soccerball
$ ls -la soccer*
$ cp -p soccer soccerball
$ ls -la soccer*
$ cp -p soccer sportsNEW
```

Archival copy:

```
$ cp -a 2012tax 2012save
$ cp -pr 2012tax 2012save
```

Exercise #6: Moving/Copying Files

Move (mv)

```
$ cd /home/$USER/LinuxClass
$ touch  baseball
$ mv  baseball  hockey
$ ls  -la  baseball
$ ls  -la  hockey
```

Create a directory path using mkdir with the -p option:

```
$ mkdir  -p 2012/tax/forms
$ mv  2012  2013
$ ls  -Rl  2012
$ ls  -Rl  2013
```

Copy (cp)

```
$ cp  hockey  icehockey
$ mkdir  hockeypuck
$ mv  icehockey  hockeypuck
$ cp -pr hockeypuck hockeystick
$ ls  -la  hockey*
$ ls  -la  *hockey
$ cp  2013  2012
```

What did you see?

```
$ cp  -r  2013  2012
```

Archival copy:

```
$ cp  -a  2013  2014
```

grep – pattern matching search of a file

grep – global/ regular expression/ print

```
$ grep  cat  nonsense.txt
$ grep  dog  nonsense.txt
$ grep  -i  dog  nonsense.txt
$ grep  -v  dog  nonsense.*
$ grep  oc  nonsense.txt
$ grep  ^oc  nonsense.txt
$ grep  oc$  nonsense.txt
```

grep is a powerful tool. Use it !

```
$ grep  --help
$ man  grep
```

find – where are my files?

find - search for files (subject to various criteria)

```
$ find  [path_to_search]  [expression/options]
$ find  .   –name  "*.txt"
$ find  /home/$USER/LinuxClass  –iname  "capital*"
$ find  /home/$USER/LinuxClass  –type  f  -mmin  20
$ find  /home/$USER/LinuxClass  –type  f  -mmin -20
$ find  /home/$USER/LinuxClass  –type  f  -mmin +20
$ find  /home/$USER/LinuxClass  –type  f  –mtime 1
$ find  /home/$USER/LinuxClass  –name  "*.bak"  –delete
$ find  .  –name  "*.txt"  –exec  ls  \-la  {}  \;
$ man  find
```

Exercise #7: Using find

- Search for files from the class directory as the root

```
$ cd ~/LinuxClass
$ pwd
```

- Use `find`
  - locate the file named 'colors'
  - locate the files that **contain** the word 'bear' where the match is **case insensitive** (hint: -iname option)
  - locate the files that were modified LESS than 90 minutes ago
    (hint: -mmin option)

- How many files did you find for each?

wc – count lines, words and bytes in a file

wc (word count)

```
$ wc  nonsense.txt
 19 95 505 nonsense.txt
Output shows the number of lines, words and characters
in the file


Can use argument to only get one of the three values:


$ wc  -l  nonsense.txt
$ wc  -w  nonsense.txt
$ wc  -m  nonsense.txt


$ wc  -help
```

uniq – show or remove duplicate lines

- uniq

  show either unique or duplicate **consecutive** lines in a file or output. Default behavior is to merge adjacent matching lines into one, but can be used to print just the matching lines or provide a count of matching lines…most effective with the sort command

```
$ uniq  bears     # supress consecutively repeated lines
$ uniq  -d  bears  # show only duplicate lines
$ uniq  -c  bears  # show a count of each unique line
```

Sorting

Sort command

sort and output to the terminal

```
$ cat  baseball.txt
$ sort  baseball.txt
$ sort  -r  baseball.txt
$ sort  -b  -k2  baseball.txt
$ sort  -bn  -k2  baseball.txt
$ sort  -bnr -k2  baseball.txt
$ sort  -help
```

Pipes (redirect to other processes)

Connect the output of one process to the input of another process.

'|'

remember, '>' redirects to a file

```
$ cat college1 | sort | uniq
$ cat college2 | sort | uniq
$ cat college1 college2 | sort | uniq  -c
```

Write to a file at the end:

```
$ cat college1 college2 | sort | uniq |grep  ^B > Colleges
```

Exercise #8: sort, grep and redirection

- `cd ~/LinuxClass`
- Look at the contents of two files, `grocery1` and `grocery2` (use `cat` command)
- Concatenate the two files with `cat` command and then use the `sort` and `uniq` commands to get a list of sorted, unique items for the grocery list
- Now redirect the output to a file named `grocery3`
- Use the `wc` command to determine how many unique items are on the list (in the `grocery3` file).
- Use `grep` and `wc` to determine how many items in the `grocery3` list start with the letter 'c'

Exercise #8 continued

```
$ cat  grocery1
$ cat  grocery2
$ cat  grocery1 grocery2 | sort | uniq
$ cat  grocery1 grocery2 | sort | uniq | wc  -l
```

You should have 32 items

```
$cat  grocery1  grocery2 | sort | uniq > grocery3
$ grep ^c grocery3
```

7 items start with the letter c

Other useful commands

- history – displays a history of commands
- alias – list aliases or create a new one

Example:

```
$ alias hi="history 20"
```

- cut – print out selected fields

Example:

```
$ cat famousdogs | cut –f1,4 –d:
```

Input, Output and Error

- Commands (processes) read from input and write to output.
  - Standard input is your keyboard by default
  - Standard output is your terminal screen by default
  - Standard error is also your terminal screen by default
    - this is where error messages appear

```
$ echo foobar > mystuff
$ echo foobar > /mystuff
bash: /mystuff: Permission denied
```

Input, Output and Error (cont.)

- STDIN, STDOUT and STDERR have numbers associated with each:

```
STDIN   0       Standard input
STDOUT  1       Standard output
STDERR  2       Standard error
```

```
$ chmod  ugo-r  /scratch/Class/exercises/dogfile2
```

- Redirect STDERR to a file:

```
$ grep  dog  $HOME/LinuxClass/exercises/dogfile*   2>  errors
$ grep  dog  $HOME/LinuxClass/exercises/dogfile*   2>  /dev/null
```

- Redirect STDOUT and STDERR (both of them) to a file:

```
$ grep  dog  $HOME/LinuxClass/exercises/dogfile*  >  out_errs  2>&1
```

or

```
$ grep  dog  $HOME/LinuxClass/exercises/dogfile*  &>  out_errs
```

Combining a chain of commands

Read from a file and do some stuff

Do some more stuff

Write to a different file

```
$ sort -r < foodfile | grep -i ^s > s-foods
```

- sort reads from a file; its output is piped to grep which then write to a file
- Most programs will read from standard input if no file is specified in arguments

awk – text manipulation (if time allows)

- awk treats a line of input as numbered fields (words) separated by white space. Each field is referenced by its number like this: $1, $2, etc..
- It is a feature rich program, this is only the simplest, but most often used application

```
$ cat  hare_tortoise
```

The hare beat the tortoise handily.

- We tell awk to print the words in a different order:

```
$ cat hare_tortoise | awk '{print $1,$5,$3,$4,$2,$6}'
```

The tortoise beat the hare handily.

- Use only the fields you need

sed – stream editor

- sed allows you to modify text based on a pattern match

```
$ cat  hare_tortoise
```

The hare beat my tortoise handily.

- replace 'beat' with 'defeated':

```
$ cat hare_tortoise | sed  's/beat/defeated/g'
```

The hare defeated my tortoise handily.

- The '/g' at the end indicates that the change is global. without it, only the first instance (per line) would be changed

- We can use both awk & sed on the same command line:

```
$ cat hare_tortoise | awk ‘{print $1,$5,$3,$4,$2,$6}’ | \
 sed 's/beat/defeated/g'
```

- The tortoise defeated my hare handily.

tr – translation: character substitution

- tr lets you replace all instances of one character with a different one– often used to change case of letters

```
$ echo 'I love Linux!' | tr "a-z" "A-Z"
I LOVE LINUX!
```

More Linux Command Basics…Quotes

- bash treats single, double and back quotes in commands differently
- Contents within a pair of single quotes are used verbatim

```
$ echo '$USER'
$USER
```

- Contents within double quotes allows shell variable evaluation

```
$ echo "The home directory of $USER is $HOME"
The home directory of user is /home/user
```

- Contents within back quotes are treated as a command; the result is a string formed from the output of the command.
- can be assigned to a variable:

```
$ NOW=`date`; echo $NOW
Thu Nov 21 14:38:13 EDT 2014
```

Processes

Show processes

Show your processes:

```
$ ps
$ ps –f      # full format

$ sleep 5    # do nothing for five seconds

$ sleep 25 &   # do nothing for 25 seconds in the background
$ ps -f
$ ps –f --forest
```

Show all processes:

```
$ ps –e
$ ps –ef --forest
$ man ps
```

More on Processes

Background and Foreground processes

- A command/job can be run in the background by adding '&' to end of the command:

```
$ sleep 50 &
[1]+ Done        sleep 50
```

- **^Z** suspends a running job/process
- **bg** resume a suspended job in the background. you get a prompt and can type new command while the bg'd job runs
- **fg** resume a suspended job in the foreground. continues to run normally
- **^C** interrupts or kills the currently running process

- **Warning:** Backgrounded processes will die when you log out of your session unless you use something like nohup or screen.

More on Processes

Suspend and Resume processes

^Z suspends an active job

```
$ sleep 300
^Z (process is suspended)

$ bg
$ ps -f
$ fg
```

Killing a process

```
$ sleep 300
^Z
$ bg
$ ps
  PID TTY          TIME CMD
 6686 pts/0    00:00:03 bash
 8298 pts/0    00:00:00 sleep
 8299 pts/0    00:00:00 ps
```

find the PID of the process you want to kill

```
$ kill 8298
```

Processes: terminate (kill) them

"kill" only requests that the program exit. Use a signal 9 to force it to exit

```
$ sleep 300
^Z
$ ps
  PID TTY          TIME CMD
 6686 pts/0    00:00:03 bash
 8298 pts/0    00:00:00 sleep
 8299 pts/0    00:00:00 ps
```

find the PID of the process to be terminated

```
$ kill -9 8298
```

- The kill command is slightly misnamed, what it actually does is send a *signal* to a process
- Most signals are interpreted by the application receiving the signal, so behavior is consistent only by convention
- Signal 9 is the exception: it will terminate the process with extreme prejudice

Processes: Make them Nicer

"nice" a process. Reduce the priority of your process to reduce its overall impact on the system. Valid values for *n* are between 0 (highest priority) and 19 (lowest priority).

```
$ nice -n 10 cat /dev/urandom > /dev/null &
$ top -u `whoami`
…
$ renice -n 15 -p [pid of "cat" command]

Now kill the process:
$ ps -ef|grep urandom
$ kill -9 [pid of urandom process]
```

Exercise #10: terminate process with "kill"

- Start a 'sleep' process that will run in the background for 300 seconds:

```
$  sleep 300
```

- Type **^Z** to suspend the **sleep** process
- Type **bg** resume the **sleep** in the background
- Check that the process is running by using the **ps** command and note the pid, process identification number
- Using the **kill** command with the pid of the **sleep** process from the previous step, terminate the sleep process
- How can you check that the sleep process is gone?

uptime

- **uptime** -- summary of the system status

```
$ uptime
 14:39:46 up 14 days,  7:00, 305 users
 load average: 39.18, 40.68, 38.68
```

- In order (left to right):
  - Current time
  - Length of time since last boot
  - Number of users currently logged on
  - The average system load for the past 1, 5 and 15 minutes
    - load = number of running proceses
    - if load < number of cores, then the system is not yet stressed
    - Helix has 128 cores

Who is doing what…using top

top makes continuous output about running processes

```
$ top
top - 16:19:54 up 28 days,  9:07, 255 users,  load average: 32.18, 32.79, 33.22
Tasks: 4749 total,   8 running, 4733 sleeping,   7 stopped,   1 zombie
Cpu(s):  9.6%us,  5.8%sy,  6.0%ni, 78.2%id,  0.2%wa,  0.0%hi,  0.2%si,  0.0%st
Mem:  1058656848k total, 955041356k used, 103615492k free,    79064k buffers
Swap: 67108856k total,   547376k used, 66561480k free, 89619996k cached

   PID USER      PR  NI  VIRT  RES  SHR S %CPU %MEM    TIME+  COMMAND
120202 johanesb  39  19  235m 180m 1432 R 96.4  0.0 170:21.86 merlin
252158 liqingli  39  19 58496  26m  756 S 95.0  0.0  17141:15 moe
170176 bozser    33  13  407m 117m 2588 S 60.1  0.0  62:30.33 ascp
218983 jrussler  20   0 18532 4704  872 R 22.3  0.0   0:00.38 top
127988 elliottm  39  19  223m 3544 1064 S 16.8  0.0 782:02.42 sshd
198816 wenxiao   20   0  4280  792  416 D 14.0  0.0  24:50.19 gzip
```

press 'q' to quit

Looking at file system (disk) Space 1.

To see local file system space:

```
$ df -l
Filesystem           1K-blocks      Used Available Use% Mounted on
/dev/mapper/vg_helix-root
                      51403396   8502228  40289968  18% /
tmpfs                529355640      2216 529353424   1% /dev/shm
/dev/sda2               495844    180879    289365  39% /boot
/dev/sda1               204580     33228    171352  17% /boot/efi
/dev/mapper/vg_helix-tmp
                      51606140    973788  48010912   2% /tmp
/dev/mapper/vg_helix-var
                      32253856  19349996  11265460  64% /var
```

Looking at file system (disk) Space 2.

To see numbers in human readable format:

$ df -lh

```
Filesystem            Size  Used Avail Use% Mounted on
/dev/mapper/vg_helix-lv_root
                       50G  8.9G   38G  19% /
tmpfs                 505G  5.6M  505G   1% /dev/shm
/dev/sda2             485M  142M  318M  31% /boot
/dev/sda1             200M  256K  200M   1% /boot/efi
/dev/mapper/vg_helix-lv_tmp
                       50G  613M   47G   2% /tmp
```

All filesystems, including network file systems:

$ df -h

Directory size

Estimate disk usage (du)

```
$ cd /home/$USER
```

Estimate a file size:

```
$ du LinuxClass/pbs_user_guide.pdf
```

Summary:

```
$ du -s LinuxClass
```

Summary in human-readable format:

```
$ du -sh LinuxClass
```

Default output

```
$ du
```

Checking Quotas on Helix/Biowulf

checkquota

- The **checkquota** command will query all network storage devices to find the applicable quota(s) for you
- This command is specific to Helix and Biowulf and is not available to Linux in general since it relies on information that is site-specific to this infrastructure.

```
$ checkquota
Mount                    Used       Quota  Percent     Files
/data:                 94.2 GB   200.0 GB   47.12%     70424
/home:                  5.2 GB     8.0 GB   64.71%    133607
mailbox:              347.9 MB     1.0 GB   33.98%
```

File Transfer

SCP, SFTP and clients

- SCP and SFTP are file transfer protocols that run over SSH, the same protocol that you used to log in
- They are very secure and encrypt both the log-in and content of any transfer

Clients

Linux/MacOS:

- `scp` secure copy
- `sftp` secure FTP
- `fuze-ssh` (Linux only)

Windows:

- WinSCP
- Filezilla
- Swish

Using OpenSSH (Linux/macOS)

scp: securely copy files

```
$ scp user@helix.nih.gov:/tmp/file ~
```

Recursive copy (whole directory)

```
$ scp -r user@helix.nih.gov:/tmp/dir ~
```

Preserve time stamps

```
$ scp -p user@helix.nih.gov:/tmp/file ~
```

From local host to remote.

```
$ scp ~/file user@helix.nih.gov:/tmp/
```

As usual

```
$ scp --help
$ man scp
```

Using OpenSSH (Linux/macOS)

sftp: secure file transfer protocol

```
$ sftp user@helix.nih.gov
sftp> cd /tmp
sftp> get file
Fetching /tmp/file to file
/tmp/file  100% 2048KB   2.0MB/s   00:00
sftp> put file newfile
Uploading file to /tmp/file
file  100% 2048KB   2.0MB/s   00:00
sftp> exit

$ man sftp
```

Exercise #11: Using scp

- Open a new terminal window on your local machine (not Helix)

- Use scp to copy the file read-write.txt from Helix to your local system:

```
$ scp user@helix.nih.gov:/data/classes/linux/read-write.txt .
```

- Advanced : copy a whole directory:

```
$ scp -pr user@helix.nih.gov:/data/classes/linux/examples  .
```

Logging out

```
$ exit
```

Review

- History – Linux Torvalds, 1991
- Why Linux? Performance, functionality and portability
- Bash shell & shell variables
- Files and directories – permissions & ownership
- Linux file system
- Paths to files/directories
- Basic Linux commands to create & access files & directories
- pico editor
- sort, grep & find
- pipe & file redirection
- processes
- transferring data to and from a Linux system
- cron

Resources

- Linux Documentation Project: http://tldp.org/
  - **Introduction to Linux - A Hands on Guide**
  - **Bash Guide for Beginners**

- Helix Web Site: http://helix.nih.gov