Hello!
Please make sure you are logged in to Clark via a terminal (Mac) or MobaXterm (Windows).

If you are having trouble doing this, please let us know during this time so we can assist before beginning the tutorial.

-_-_-_-_-
RCSS Linux & Lewis/Clark Introduction

RCSS - mudoitrcss@missouri.edu
(contact us here for any cluster related help)

Asif Ahamed
Predrag Lazic
Brian Marxkors
Ashkan Mirzaee
Christina Roberts
What is a cluster/supercomputer?

- A supercomputer is a computer with a high level of performance as compared to a general-purpose computer
- The world's fastest 500 supercomputers run Linux-based operating systems
- Purpose: massive parallelization because life is too short!
Why learn Linux?

Not only is it a valuable job skill, *It is the only way* to use a super computer!

The learning curve in Linux is very steep, but you only have to climb it once!
Graphical User Interface (GUI) vs the Command Line

“intuitive”

Reproducible, quick, efficient!
## Basic Linux Commands

<table>
<thead>
<tr>
<th>Command</th>
<th>Command</th>
<th>Command</th>
<th>Command</th>
<th>Command</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>ssh</td>
<td>wget</td>
<td>more</td>
<td>cp</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>scp</td>
<td>less</td>
<td></td>
<td>rm</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>rsync</td>
<td>head</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pwd</td>
<td>unzip</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ls</td>
<td>tail</td>
<td></td>
<td></td>
<td></td>
<td>..</td>
</tr>
<tr>
<td>cd</td>
<td>gzip</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mkdir</td>
<td>tar -xzf</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>man</td>
<td>vim</td>
<td>grep</td>
<td>history</td>
<td></td>
<td>&gt;&gt;</td>
</tr>
<tr>
<td>tree</td>
<td>nano</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>clear</td>
<td>emacs</td>
<td>df</td>
<td>exit</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Command prompt

- `ssh USER\NAME@clark.rnet.missouri.edu`

~ (tilde) = home directory
- `pwd`
  Will show what directory you are working in.
  If you are not in home type -> `cd ~`

- `ls`
  Will list everything in the directory.

- `mkdir TUTORIAL`
  To make a subdirectory named “TUTORIAL”.

- `cd TUTORIAL`
  To move into the TUTORIAL subdirectory.
-> cp /group/training/Linux_Cluster_202106.zip.

Other ways to get files to & from server -> scp rsync wget

-> unzip Linux_Cluster_202106.zip

Will unzip this file.
Other ways to unzip files -> gunzip unrar tar -xzf

**We acknowledge usage of this file is provided by http://software-carpentry.org/**

Research Computing Support Services (RCSS)
Will list everything in the directory with extra information.

Shows other options for the commands, such as ls.

Will list everything in the directory with extra information.
-rwxrw-r--. 1 predrag RCSS 5204 Feb 22 16:00 Toyota.sh
drwxrwxr-x. 2 predrag RCSS 5 Aug 7 23:51 practice/

- d – directory
- - - file

U – user/owner permissions     r - read
G – group permissions         w - write
O – other users permissions    x - execute
Research Computing Support Services (RCSS)

```
-> tree Linux_Cluster_202106
  Will list everything in the directory AND subdirectories.

-> clear
  Clears all text from terminal

-> cd Linux_Cluster_202106/data_tab
-> ls
  Change to this directory instead of typing path in front of each command.
```
Research Computing Support Services (RCSS)

Will preview contents of file.
Also try -> less

Will display the last 10/2 lines of the file.
Also try -> head

To view the number of lines, words and bytes of a file.
To search occurrences of patterns in a file.

- `grep Hy Honda_tab.txt`
- `grep Accord\ Hy Honda_tab.txt`
- `grep "Accord Hy" Honda_tab.txt`
- `grep Hy *`

will produce error  

`-> grep Accord Hy Honda_tab.txt`
To REDIRECT output into a new file:

```bash
- grep Hy Honda_tab.txt > hybrid.txt
- grep Hy Toyota_tab.txt > hybrid.txt
```

| (pipeline) allows for usage of multiple commands at once:

```bash
- grep Hy Honda_tab.txt | wc -l
- grep Hy Toyota_tab.txt | wc -l
```

```bash
- grep Hy Honda_tab.txt >> hybrid.txt
```
To remove a file.

To copy contents of a file to a new name.
To move content of a file to a new name and/or directory.

.. accesses the previous directory the sub directory is located in.

.     current directory
..    previous directory
~     home directory
ls ./scripts/Linux/
more ./scripts/Linux/script_2020.sh
sh ./scripts/Linux/script_2020.sh

more ./scripts/Linux/script_insertyear.sh
sh ./scripts/Linux/script_insertyear.sh 2018
sh ./scripts/Linux/script_insertyear.sh 2015
Writing Scripts

**KNOW HOW TO TYPE, SAVE, & EXIT BEFORE USING EDITORS (some tips on following slide)

- `cp ../scripts/Linux/script_2020.sh ./script_2018.sh`
- `vim script_2018.sh`
  
  **change 2020 to 2018 in editor**
- `sh script_2018.sh`

File editor of choice

Also try

- emacs
- nano
- visual studio
<table>
<thead>
<tr>
<th>emacs -nw</th>
<th>nano -z</th>
<th>vim</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctrl-k</td>
<td>Ctrl-k</td>
<td>i</td>
</tr>
<tr>
<td>Ctrl-y</td>
<td>Ctrl-u</td>
<td>Esc dd</td>
</tr>
<tr>
<td>Alt-w</td>
<td>Alt-6</td>
<td>Esc yy</td>
</tr>
<tr>
<td>Shift-(arrow)</td>
<td>Alt-m + Alt-a</td>
<td>Esc p</td>
</tr>
<tr>
<td>Ctrl-x + Ctrl-s</td>
<td>Ctrl-o</td>
<td>Esc u</td>
</tr>
<tr>
<td>Ctrl-x + Ctrl-c</td>
<td>Ctrl-x</td>
<td>Esc + :w</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Esc + :q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Esc + :q!</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
GUI log-on for MAC USERS: ssh -YC USERNAME@clark.rnet.missouri.edu
(MAY require XQuartz download: https://www.xquartz.org/releases/index.html)

RECAP/REVIEW

Linux Commands
Moving Data
Data Search (more/grep)
Text Editors
Creating Scripts
What is a Supercomputer
Partitions

LEWIS PARTITIONS
- Interactive
- General
- hpc5
- Gpu
- Serial
- Dtn

BioCompute
- Lewis
- hpc6
- gpu3
- gpu4

hpc3
hpc4
hpc4rc

CLARK PARTITIONS
- General
- hpc3
- r630-hpc3

Research Computing Support Services (RCSS)
Slurm is for cluster management and job scheduling. All RCSS clusters use Slurm (https://slurm.schedmd.com).

**SLURM SCHEDULER**

Slurm is for cluster management and job scheduling. All RCSS clusters use Slurm (https://slurm.schedmd.com).

User 1: 4 PEs
User 2: 6 PEs
User 3: 2 PEs
User 4: 6 PEs (with higher fairshare)

**EXPLAINED LATER**
More Commands (Linux and Slurm)

top
hostname
module avail
module load
module list
module clear
	srun
sbatch
sacct
scancel

info
sshare
sjstat

sinfo
sshare
sjstat

groups
free -h
df --human-readable
lfs quota -hg

#SBATCH
--mem
--time
--partition
--job-name
All jobs must be submitted to SLURM scheduler. Jobs that are found running on the **login node** will be immediately terminated.
Interactive Node on **Lewis** (minimum necessary)

-> `srun -p Interactive --qos Interactive --pty /bin/bash`

hpc3 Node on **Clark**

-> `srun -p hpc3 --pty bash`

**Examples of other options**

`srun -p Interactive --qos interactive -t 0-2:00 -n 4 --mem 8G --pty bash`

`srun -p Interactive --qos interactive --mem 4G --licenses=matlab:1 --pty /bin/bash`

NOTE: ONLY ONE command on this page should be entered, unless there is an error and any allocated resources are first exited.
module avail
module avail r

module load r/4.0.3
module list

module load rstudio
module list

GUI log-on for MAC USERS:
ssh -YC USERNAME@clark.rnet.missouri.edu
```r
honda_car <- read.csv("Honda_csv.csv")

head(honda_car)
levels(honda_car$Type)
mean(honda_car$City.mpg.)

small_cars <- subset(honda_car, Type=="Hatch" | Type=="Sedan" | Type=="Coupe")
large_cars <- subset(honda_car, Type=="Van" | Type=="Truck" | Type=="SUV")

small_mean <- mean(small_cars$City.mpg.)
large_mean <- mean(large_cars$City.mpg.)

plot(small_cars$City.mpg., col="blue", xlim=c(0,50), ylim=c(0,150))
points(large_cars$City.mpg., col="red")
points(small_mean, pch=7)
points(large_mean, pch=7)

dev.off()
q()
```
```r
loadhistory(file=".Rhistory")

plot(small_cars$City..mpg, col="blue", xlim=c(0,50), ylim=c(0,150))
points(large_cars$City..mpg, col="red")

jpeg('honda_rstudio_plot.jpg')

plot(small_cars$City..mpg, col="blue", xlim=c(0,50), ylim=c(0,150))
points(large_cars$City..mpg, col="red")
points(large_mean, pch=7)
points(small_mean, pch=7)
dev.off()

q()
```
Rscript r_example_script.R
ls
module list
module purge
module list
exit
pwd
Submitting & Monitoring Jobs

- `cd ~/TUTORIAL/ Linux_Cluster_202106/data_tab`
- `more ../scripts/Cluster/script_2017_forSLURM.sh`
- `sbatch ../scripts/Cluster/script_2017_forSLURM.sh`

- `sacct -j JOBID`

- `ls`

- `more cars_slurm_output.txt`
Monitoring & Canceling Jobs

-> sbatch ../scripts/Cluster/script_to_CANCEL.sh

-> sacct -u USERNAME
-> sacct -u USERNAME -S2021-06-19-21:00 --format=user,JobID,Start,End,State,Partition

-> scancel JOBID
Cluster Information

-> sinfo
-> sinfo -s
-> sjstat -c
-> scontrol show partition hpc3

**LEWIS** EXAMPLE:

-> scontrol show node lewis4-z10ph-hpc3-node395

**CLARK** EXAMPLE:

-> scontrol show node clark-r630-hpc3-node909
**Fairshare & User Information**

FOR LEWIS ONLY, NOT FOR CLARK

-> sshare -U

-> sacctmgr show assoc user=USERNAME

-> sacctmgr show assoc user=USERNAME format=acc,user,share,qos,maxj

-> groups

-> df -h /home/USERNAME

-> lfs quota -hg USERNAME /storage/hpc

-> lfs quota -hg GROUPNAME /storage/hpc

There are no backups of any storage. You are responsible for your own data and backups.
REVIEW

What is a Supercomputer
Linux Commands
Moving Data
Data Search (more/grep)
Text Editors
Creating scripts

SLURM Commands
Login & Interactive Nodes
GUI on the Cluster
Scheduling & sbatch Submissions
Monitoring Jobs & Cluster
Fairshare
cd ../R/
more
cp toy_rscript.R hon_rscript.R

vim hon_rscript.R

Use text editor to change all toyota to honda

with vim
  -> %s/toyota/honda/g
  -> %s/Toyota/Honda/g
  -> :wq

more r_multi_script.sh
sbatch r_multi_script.sh
sacct -u USERNAME
ls

rsync USER@CLUSTER.rnet.missouri.edu:/home/USER/TUTORIAL/Linux_Cluster_202106/scripts/R/*.jpg ./Desktop/

^ Transfer all created jpg in directory to own computer
**Enter command with location & destination all as ONE line

Research Computing Support Services (RCSS)
Thank you!

Please send us a message with your username and email to.

mudoitrcss@missouri.edu

(with Subject: Linux & Cluster Training)

Feel free to include any feedback or ask for a copy of these slides.