



Linux workshop series office hours

Amy Stonelake
George Zaki
Andrew Weisman

November 2, 2021

Notes

- **Content overview:** Linux file permissions, ownership, group membership; more common way of running Bash scripts; cut and paste commands; awk and sed commands; Linux GUIs; text editors; Bash variables, environment variables; .bashrc and .bash_profile files; Biowulf
- **There are many excellent explanations and tutorials online for every topic we will discuss... just Google it!**
 - Feel free to ask the instructors about particular webpages, books, etc.
- **Note:** The default background color (blue) of Windows PowerShell appears to get in the way of the output of **grep**; this can be solved by changing the background color to, e.g., black:
 - Right click on PowerShell's title bar, click on Properties, click on the Colors tab, and change the Screen Background color to black

Linux file permissions

- Unix operating systems break up file/directory access in three ways:

- **U**ser (-AAA-----)
- **G**roup (----AAA---)
- **O**thers (-----AAA)

- Files can be:

- **R**eadable
- **W**ritable
- e**X**ecutable

```
shell-lesson-data $ ls -l
total 60
drwxrwxr-x. 2 weismanal weismanal 4096 Jul 30 04:41 creatures
drwxrwxr-x. 6 weismanal weismanal 4096 Oct 12 11:54 data
drwxrwxr-x. 2 weismanal weismanal 4096 Oct 26 11:25 molecules
drwxrwxr-x. 3 weismanal weismanal 4096 Jul 30 04:41 north-pacific-gyre
-rw-rw-r--. 1 weismanal weismanal 86 Jul 30 04:41 notes.txt
-rw-rw-r--. 1 weismanal weismanal 13 Aug 4 02:34 numbers.txt
-rw-rw-r--. 1 weismanal weismanal 32 Jul 30 04:41 pizza.cfg
-rw-rw-r--. 1 weismanal weismanal 21583 Jul 30 04:41 solar.pdf
drwxr-xr-x. 3 weismanal weismanal 4096 Oct 12 11:44 thesis_backup
drwxrwxr-x. 5 weismanal weismanal 4096 Jul 30 04:41 writing
```

- You can change file permissions using the **chmod** command, e.g.:

- **chmod u+x my_script.sh**
- **chmod 664 notes.txt**

- Change ownership of a file using **chown**, e.g., **chown weismanal notes.txt**
- Change the group that a file belongs to using **chgrp**, e.g., **chgrp students hello.pdf**
- Do both at once using **chown** with a dot separating the user and group, e.g.,
chown weismanal.students hello.pdf

Linux file permissions ctd.

- In order to enter a directory (which has a **d** in the leftmost bit), it must be executable
- In order to run a file (such as a Bash script), it must be executable
- If the script is executable and starts with **#!/bin/bash** then you can run it like **./my_script.sh** instead of **bash my_script.sh**

```
shell-lesson-data $ ls -l
total 60
drwxrwxr-x. 2 weismanal weismanal 4096 Jul 30 04:41 creatures
drwxrwxr-x. 6 weismanal weismanal 4096 Oct 12 11:54 data
drwxrwxr-x. 2 weismanal weismanal 4096 Oct 26 11:25 molecules
drwxrwxr-x. 3 weismanal weismanal 4096 Jul 30 04:41 north-pacific-gyre
-rw-rw-r--. 1 weismanal weismanal 86 Jul 30 04:41 notes.txt
-rw-rw-r--. 1 weismanal weismanal 13 Aug 4 02:34 numbers.txt
-rw-rw-r--. 1 weismanal weismanal 32 Jul 30 04:41 pizza.cfg
-rw-rw-r--. 1 weismanal weismanal 21583 Jul 30 04:41 solar.pdf
drwxr-xr-x. 3 weismanal weismanal 4096 Oct 12 11:44 thesis_backup
drwxrwxr-x. 5 weismanal weismanal 4096 Jul 30 04:41 writing
```


Miscellaneous notes

- We learned about **cut** to extract columns from a file; you can use **paste** to “paste” together columns from two different files
- **awk** and **sed** are useful command-line text processors, like **cut** but much beefier, e.g.:

```
animal-counts $ cat animals.txt | awk -v FS=, '{x=$3*2; printf("I saw %s/2 %s(s)\n", x, $2)}'
```

```
I saw 10/2 deer(s)
I saw 44/2 rabbit(s)
I saw 14/2 raccoon(s)
I saw 38/2 rabbit(s)
I saw 4/2 deer(s)
I saw 8/2 fox(s)
I saw 32/2 rabbit(s)
I saw 2/2 bear(s)
```

- Linux installations usually come with graphical user interfaces (GUIs); even if you cannot access the GUI part directly (as in our class virtual machine), you can use GUI software on your own computer and connect to the Linux distribution remotely using SSH
- **vim** and **emacs** are more powerful non-GUI text editors than **nano**; VS Code is an excellent GUI option

Miscellaneous notes ctd.

- Just like a `for` loop defines a variable that accesses the elements in a list (e.g., `for filename in *.txt ...`), variables can generally be defined in Bash using `=` signs, e.g.:
 - Define a variable: `my_name=Kid`
 - Access the variable: `echo "My name is $my_name"`
- Many variables, called environment variables, are set in your shell by the operating system; type `env` to see a list
 - One special environment variable is `$PATH`, which tells the shell where to look for executables
- Special files called `~/ .bashrc` and `~/ .bash_profile` are scripts that are automatically run for you when you log in
 - Here you can define environment variables, aliases, functions, etc.

Biowulf

- **Everything we learned in this workshop can be done on Biowulf**
 - Just log in to `biowulf.nih.gov` instead of `ncias-d2811-v.nci.nih.gov`
- **Unlike the virtual machine we used for this course, Biowulf is a real-life, gigantic set of machines called a supercomputer**
- **Upon logging into Biowulf, you will be in a Bash shell... now you know how to navigate it!**
- **Biowulf allocates “compute nodes” on which you can run computations or processes using Bash or other software such as Python, R, GROMACS, etc.**
 - In order to submit such a “batch job,” you need to write a “batch script”... which is simply a script written in Bash!
- **For smaller, less-demanding jobs, you can use your own computer**
- **Command-line transfers between your computer and Helix:**
 - Use `readlink -f <FILE>` to determine a file’s location on any system
 - Use `scp` to copy between systems, e.g.:
`scp andrew@helix.nih.gov:/gpfs/gsfsl0/users/weismanal/transfer/datafiles.tar.gz ~/Downloads`
- **You can request a student account on Biowulf; email staff@hpc.nih.gov**

Instructors overview and questions?

- **Amy: Bioinformatics Training and Education Program (BTEP)**
 - amy.stonelake@nih.gov
- **George and Andrew: Strategic and Data Science Initiatives (SDSI)**
 - george.zaki@nih.gov
 - andrew.weisman@nih.gov

Questions?