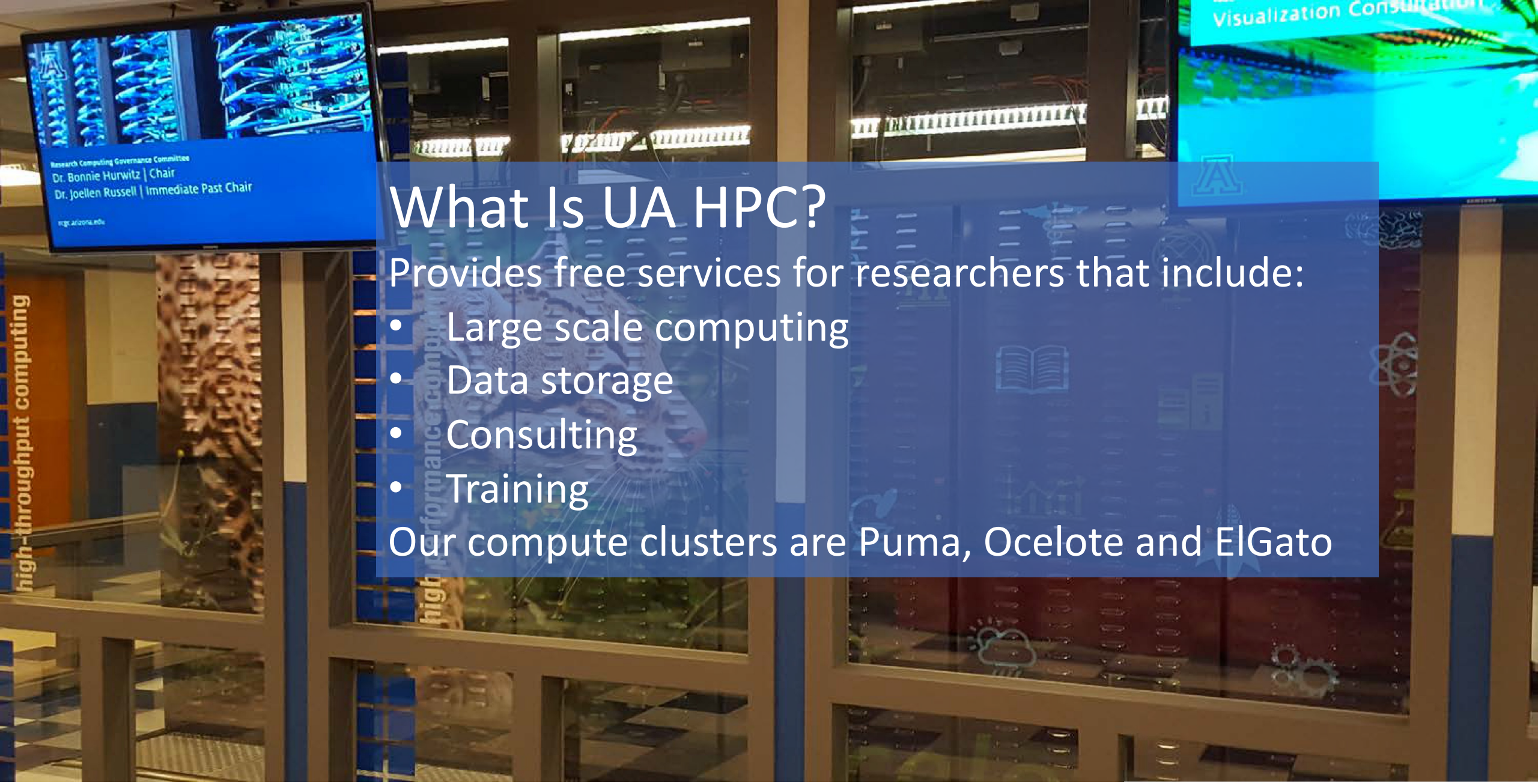
Introduction to Linux on HPC

UNIVERSITY INFORMATION TECHNOLOGY SERVICES
Research Computing

Research Computing
UNIVERSITY OF COLORADO BOULDER



What Is UA HPC?

Provides free services for researchers that include:

- Large scale computing
- Data storage
- Consulting
- Training

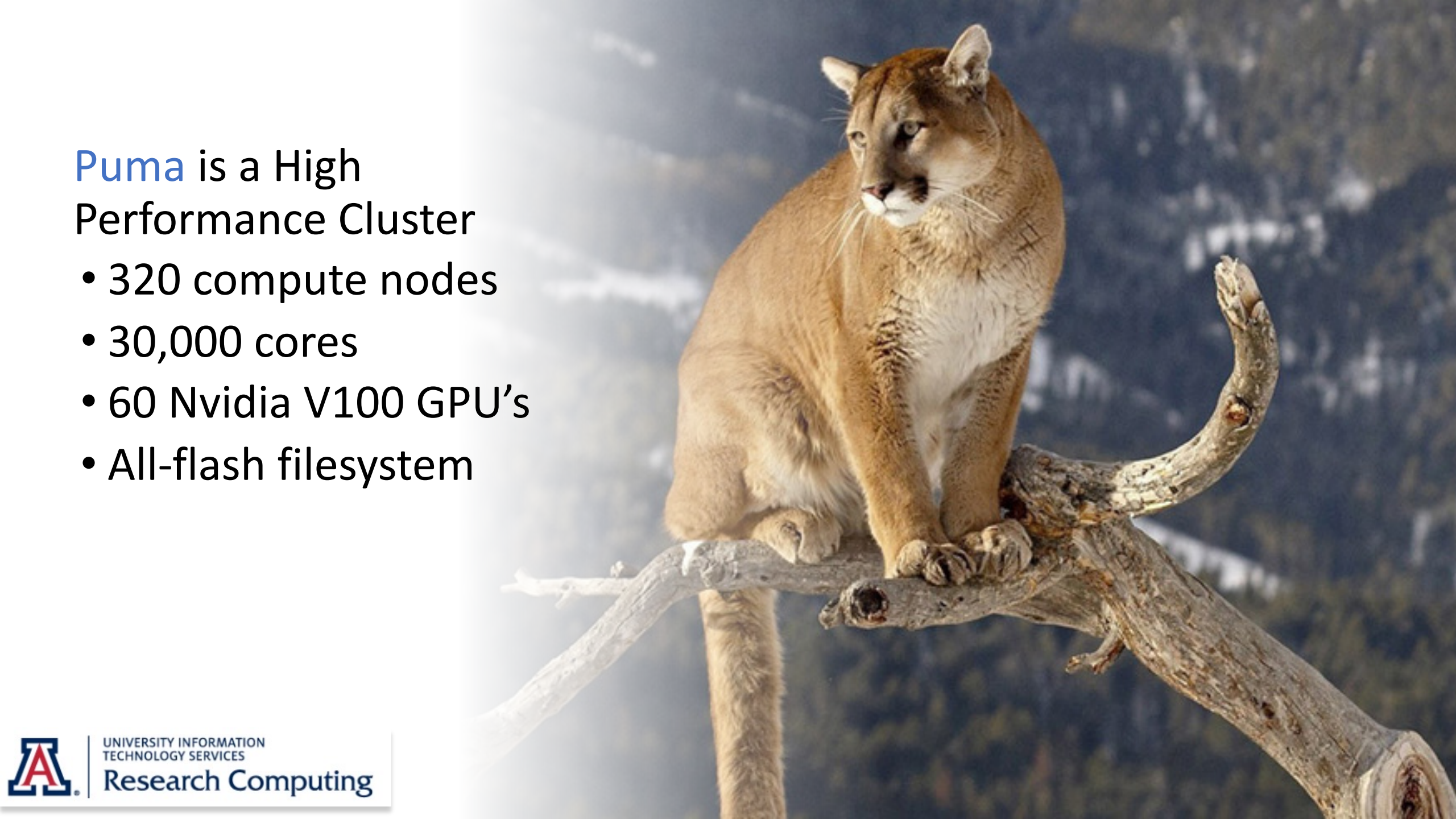
Our compute clusters are Puma, Ocelote and ElGato

https://ua-researchcomputing-hpc.github.io/Intro-to-HPC/

UNIVERSITY INFORMATION TECHNOLOGY SERVICES
Research Computing

Puma is a High Performance Cluster

- 320 compute nodes
- 30,000 cores
- 60 Nvidia V100 GPU’s
- All-flash filesystem



UNIVERSITY INFORMATION TECHNOLOGY SERVICES
Research Computing

Opening a Terminal

- Mac: Go to Applications → Utilities → Terminal
- Windows: Download a terminal emulator
  - PuTTY: https://www.putty.org
  - Git BASH: https:gitforwindows.org
- Open OnDemand: https://ood.hpc.arizona.edu/

ARIZONA Research Technologies – HPC Systems

Research Computing
UNIVERSITY OF COLORADO BOULDER

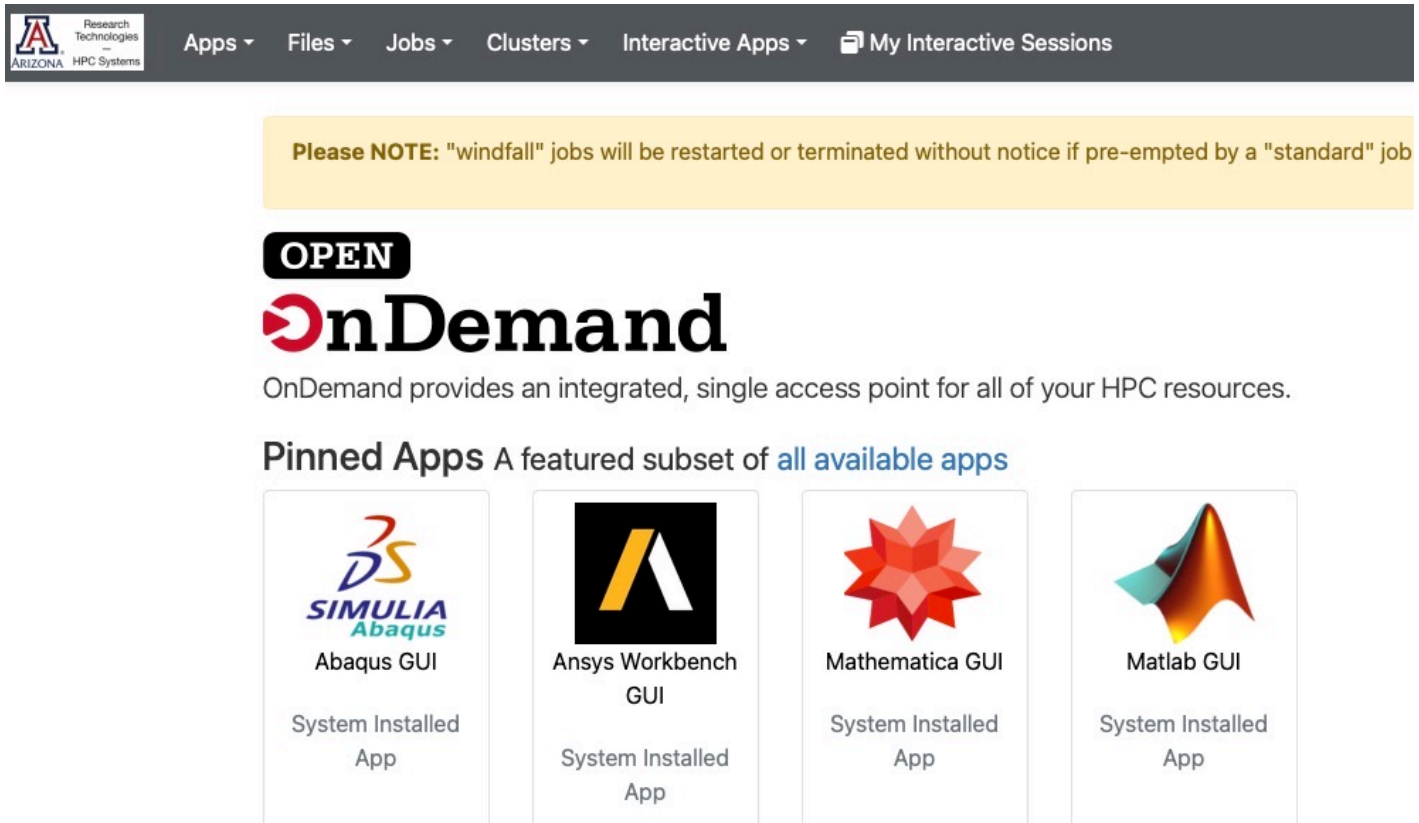
Logging into HPC

- Request an Account Sponsored by a PI
  https://public.confluence.arizona.edu/display/UAHPC/Account+Creation

- Access Your Account
  https://public.confluence.arizona.edu/display/UAHPC/System+Access

OnDemand GUI Interface

- Open **ood.hpc.arizona.edu** in your web browser and login with your NetID and password.
- From the “Clusters” drop-down menu choose which HPC cluster you would like to access:



Research Technologies — HPC Systems

Command Line Interface

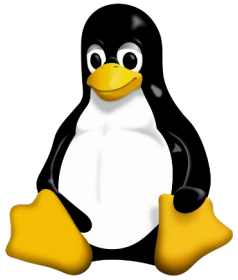
```
Laptop $ ssh netid@hpc.arizona.edu
This is a bastion host used to access the
rest of the RT/HPC environment.
Type "shell" to access the job submission hosts
for all environments
----------------------------------------
[netid@gatekeeper ~]$ shell
Last login: Mon Nov  8 20:16:14 2021 from
gatekeeper.hpc.arizona.edu
***
----------------------------------------
(puma) [netid@junonia 08:35:32 ~]$
```

bastion

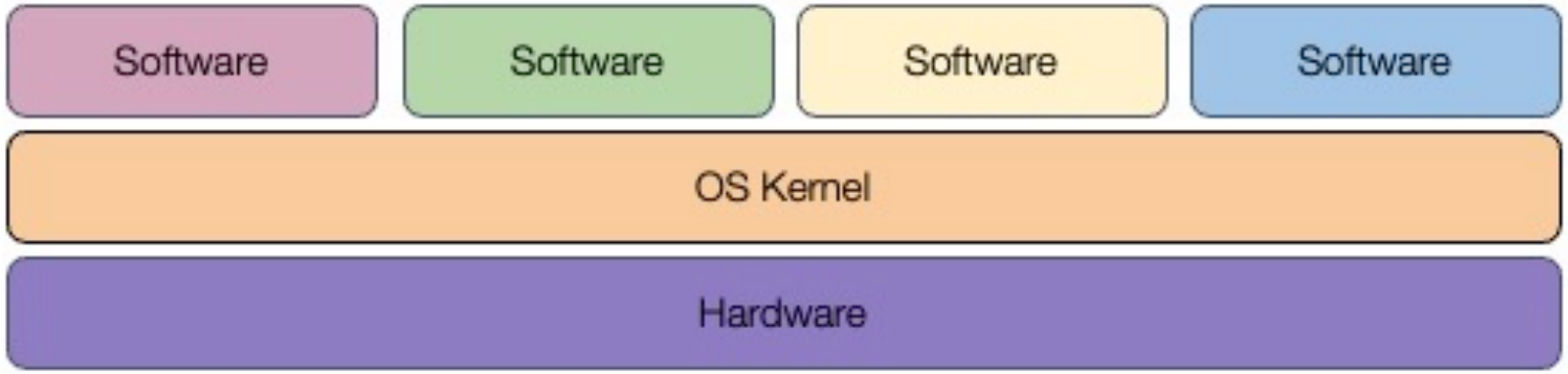
shell

prompt

What is Linux?



- Part of the Unix-like family of operating systems.
- Started in early ‘90s by Linus Torvalds.
- Typically refers only to the kernel with software from the GNU project and elsewhere layered on top to form a complete OS. Most is open source.



What is Linux?

- Several distributions are available from enterprise-grade, like RHEL or SUSE, to more workstation-focused like Ubuntu.
- Runs on everything from embedded systems to supercomputers.

Why use Linux?

- Default operating system on virtually all HPC systems and the foundation for many business services globally
- Extremely flexible
- Fast and powerful
- Frequently open source and free
- Many tools for software development

Anatomy of a Linux Command

- command [flags] [target(s)]
  ls –l myworkdir
  - A long list of directory called myworkdir
  - A directory is a collection of files

- Case is important!
- Help on commands is available through the “man” command (short for manual)
  man ls

The Linux Filesystem

- System of arranging files on disk
- Consists of directories (folders) that can contain files or directories
- Levels in file paths separated by forward slashes:

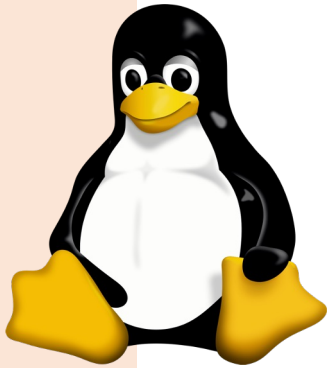
  e.g. /home/user/scripts/analyze_data.sh
- Case-sensitive; spaces in names discouraged
- Some shorthand:

  `.` (the current directory)

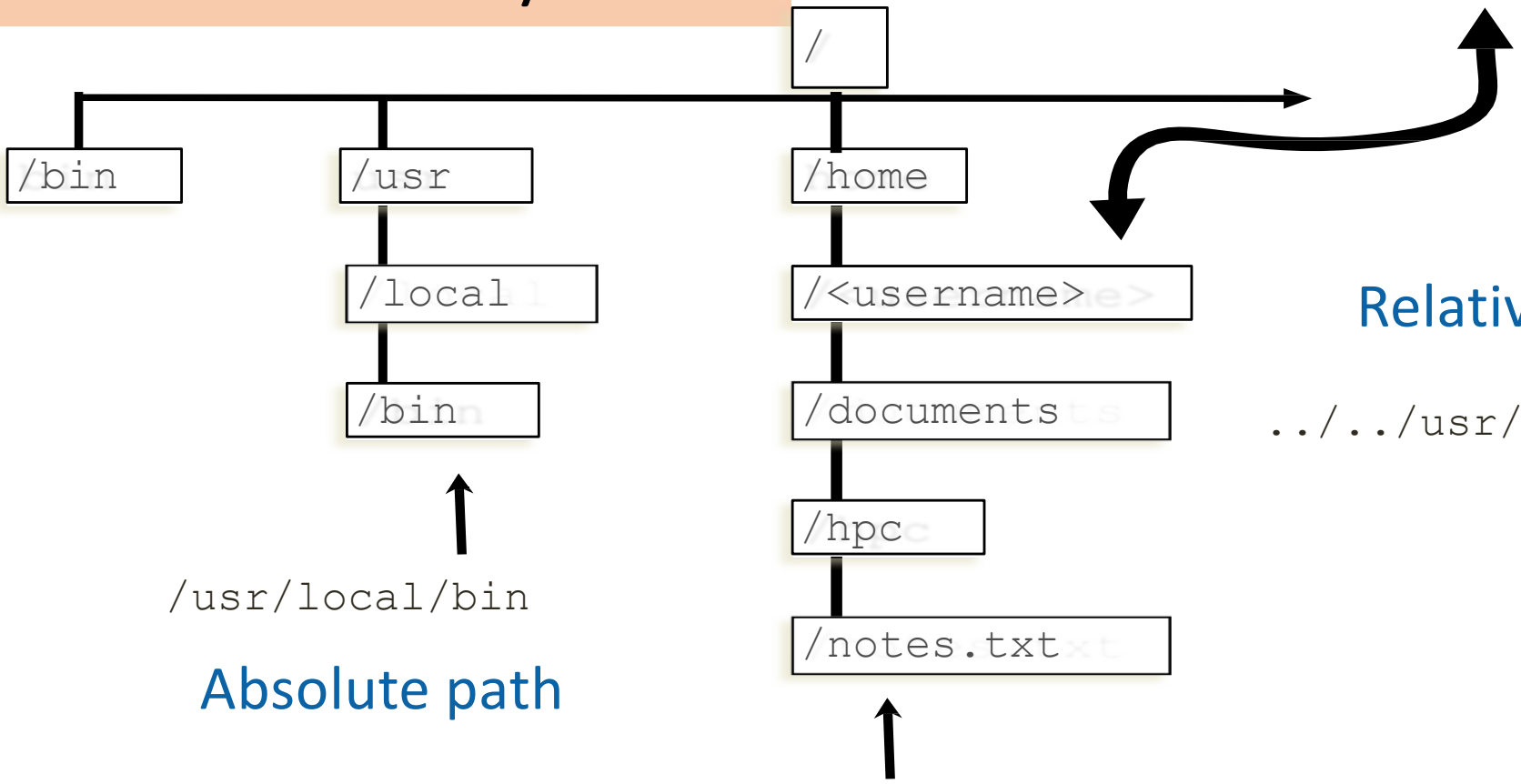
  `..` (the directory one level above)

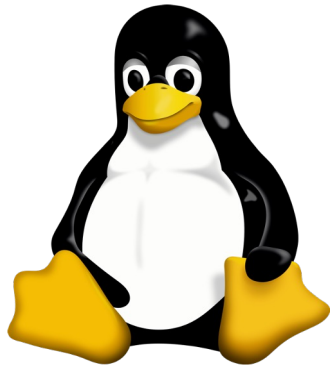
  `~` (home directory)

  `-` (previous directory, when used with cd)



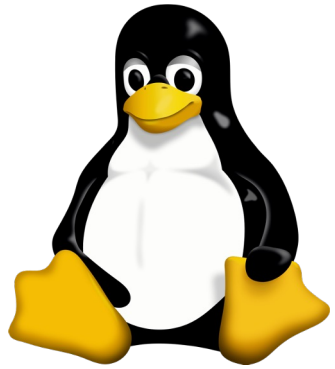
The Linux Filesystem





The Filesystems on HPC

User with netid kim



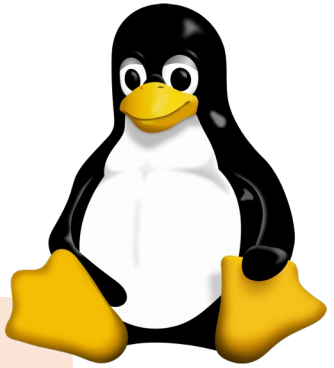
[kim@r3u13n1 kim]$ cd /groups/PI

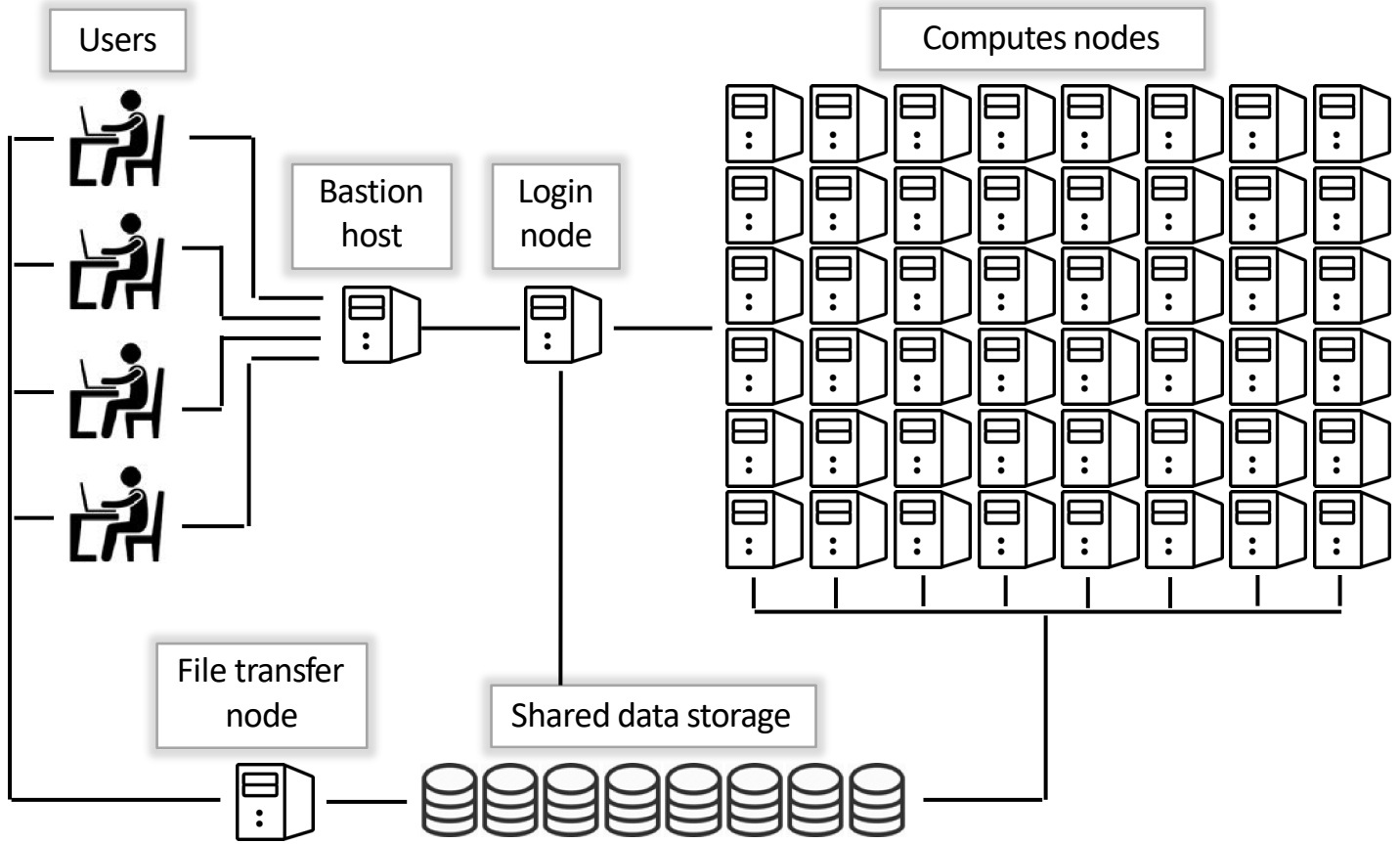
[kim@i5n5 kim]$ cd /xdisk/PI

(puma)[kim@junonia time ~]$ pwd
home/u13/kim

[kim@cpu38 kim]$ cd /tmp

Where Filesystems are Mounted





Bastion host
- Nothing here

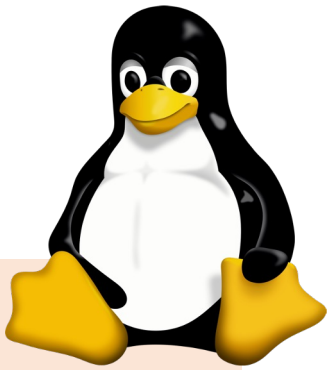
Login node
- /home
- /groups/PI
- /xdisk/PI

Compute nodes
- /home
- /groups/PI
- /xdisk/PI
- /tmp

File transfer node
- /rental/PI

Linux jargon: mounted means accessible

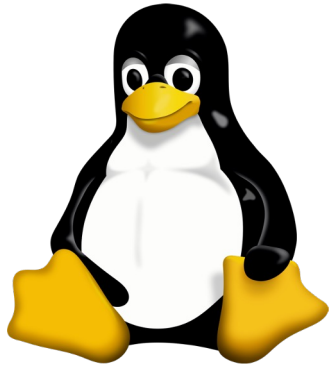
Environment Variables



- Environment variables are important for Linux users and programs
- Type env to see your currently set up environment variables
- Useful environment variables:

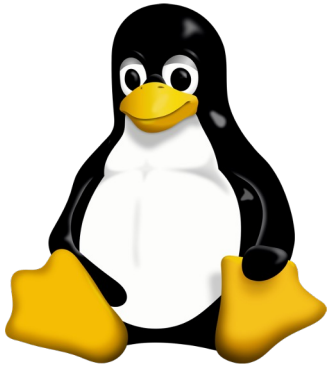
PATH	directories to search for commands
HOME	home directory
PWD	current working directory
USER	username
LD_LIBRARY_PATH	directories to search for dynamically-loaded libraries

File and Directory Commands



pwd– prints full path to current directory	**cd** – changes directory; can use full or relative path as target	**mkdir** – creates a subdirectory in the current directory	**touch** – creates an empty file
rm – removes a file (rm -r removes a directory and all of its contents)	**cp** – copies a file	**mv** – moves (or renames) a file or directory	**ls** – lists the contents of a directory (ls -l gives detailed listing)

File Viewing Commands



less– displays a file one screen at a time

cat – prints entire file to the screen

head – prints the first few lines of a file

grep – prints lines containing a string or other regular expression

ps –ef | grep xx

diff– shows differences between two files

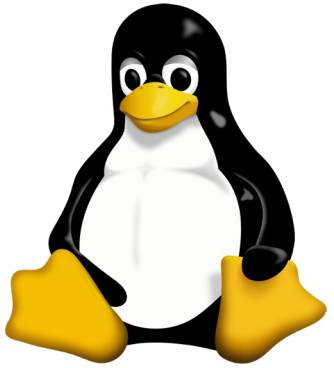
tail – prints the last few lines (with -f in real-time the end of a file that may be changing)

sort – sorts lines in a file

find – searches for files that meet specified criteria

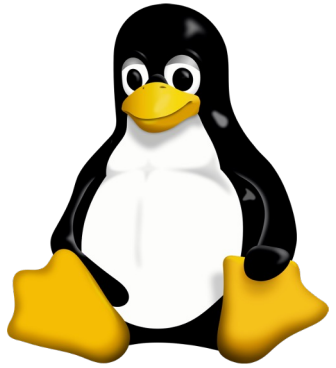
wc – count words, lines, or characters in a file

Exercise 1: Navigation



1. Sign-in to your account
2. Print the path to your current directory
3. List the contents of this directory
4. From your home directory create a new directory. How can you be sure the new directory is there?
5. Change to your new directory and create a file.
6. Remove the file you just created.

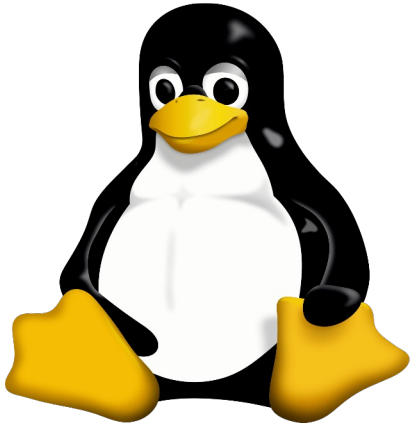
Access the example scripts



- How to get there: github.com/ResearchComputing/Supercomputing_Spinup
- From home, create a “linux” directory
- Change to this directory for the exercises
- Git clone the repository:

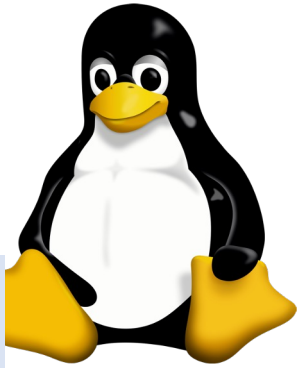
git clone https://github.com/ResearchComputing/Supercomputing_Spinup.git

Exercise 2: File Viewing



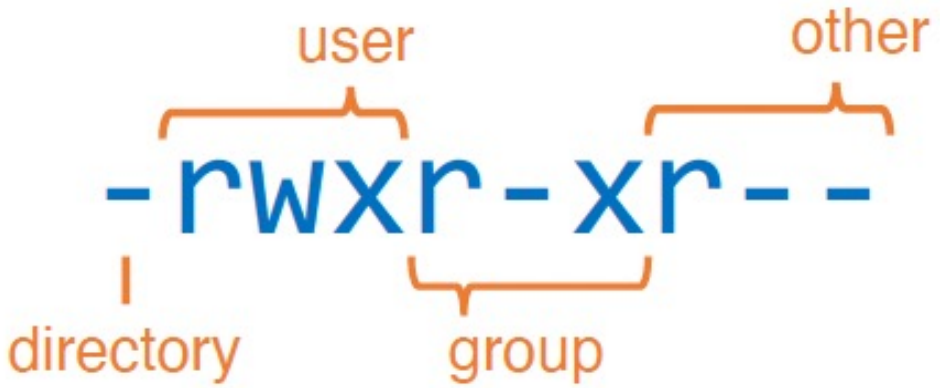
1. First change to the Supercomputing_Spinup directory
   Hint: try tab complete
2. Change to the "linux_bash_spinup/scripts" directory
3. Print out the entire "test.sh" file
4. Print out the last 3 lines of "local_vs_global.sh" file
5. Find how many words are in "case_example.sh"

Modes aka Permissions



```
drwxrwsr-x  2 sarawillis chrisreidy   3072 Nov  8 14:50 system-scripts
-rw-r--r--  1 baylyd     chrisreidy      0 Oct 18  2021 test
```

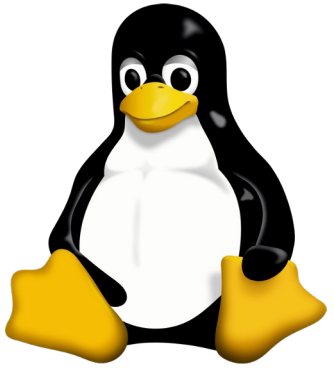
- 3 classes of users:
  - User (u) aka “owner”
  - Group (g)
  - Other (o)
- 3 types of permissions:
  - Read (r)
  - Write (w)
  - Execute (x)



Note:

- One is a file, the other is a directory
- The first name is the owner
- The second name is the group

Modes aka Permissions



```
-rw-r--r--  1 baylyd     chrisreidy      0 Oct 18  2021 test
```

chmod changes mode

To add write and execute permissions for the group

chmod g+wx test or

chmod 674 test

```
-rw-rwxr--  1 baylyd     chrisreidy      0 Oct 18  2021 test
```

To remove read permission for all others

chmod o-x test

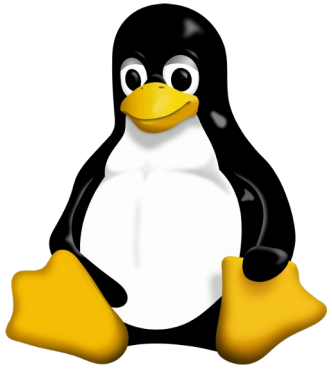
chmod 670 test

```
-rw-rwx---  1 baylyd     chrisreidy      0 Oct 18  2021 test
```

File Types and Dots

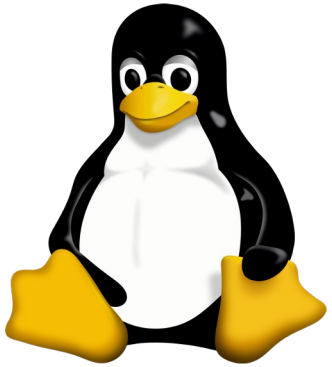
```
drwxr-xr-x.   3 chrisreidy chrisreidy      2048 Feb 10 09:17 bayes
-rw-r--r--.   1 chrisreidy tmerritt        3579 Oct 23  2019 conda-bash.sh
drwxr-xr-x.   3 chrisreidy staff            512 Oct 22  2019 .anaconda
-rwxr-xr-x.   1 chrisreidy staff         121902 Dec 16  2020 nettest
```

- ls short list
- ls –l long list
- ls –la list hidden files also
- cd .. change to parent dir
- cd ~ change to home dir



Profiles

```
-rw-------. 1 chrisreidy staff 2574 Oct 22  2021 .bashrc
```

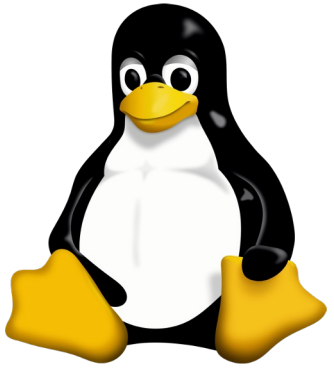


```
Sample Lines

PATH="~/perl5/bin${PATH+:}${PATH}:$HOME/.local/bin"; export PATH;
# Forces ~/.bashrc resource after cluster switch
alias puma=". /usr/local/bin/slurm-selector.sh puma && source ~/.bashrc"
#-# >>> conda initialize >>>
# Python virtualenv
# source ~/tflow/tensorflow_virtual_env/bin/activate
```

```
-rw-------. 1 chrisreidy staff 177 Aug 27  2020 .bash_profile
-rw-------. 1 chrisreidy staff 21335 Feb 20 13:26 .bash_history
```

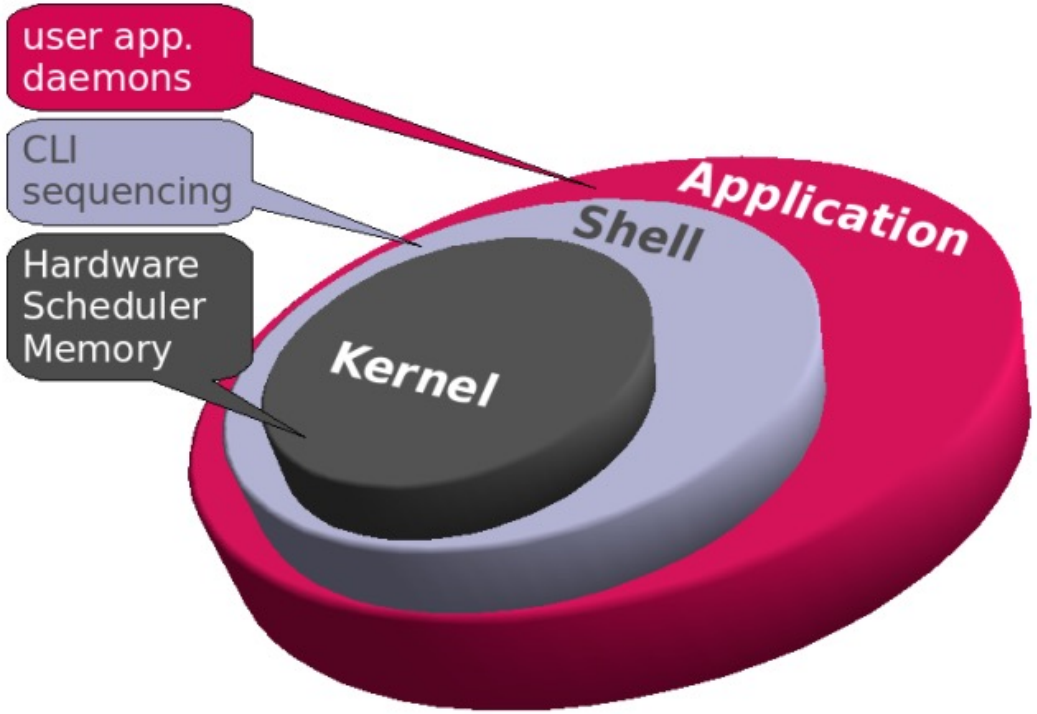
Shells and Shell Scripts



A **shell** is the environment in which commands are interpreted in Linux.

On HPC we prefer bash (Bourne Again Shell)
Other shells include: sh, csh, tcsh, ksh, zsh

Shell scripts are files containing collections of commands for Linux systems that can be executed as programs.

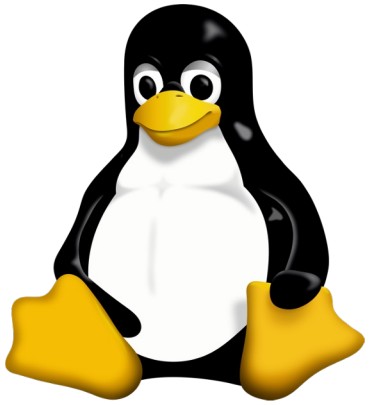


Shells cont'd

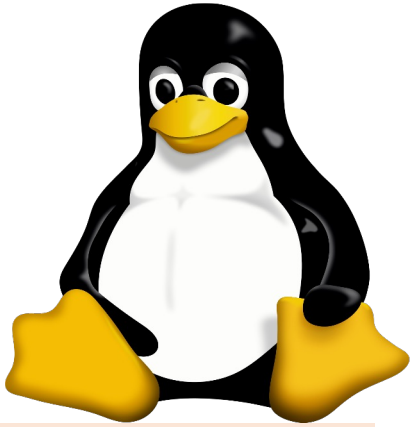
- Can be programmed interactively, directly on the terminal.
- It can also be programmed by script files. The first line of the file must contain #!/bin/bash
- The program loader recognizes the #! and will interpret the rest of the line (/bin/bash) as the interpreter program.
- If a line starts with #, it is a comment and is not run.

```
#!/bin/bash
# the files in /tmp.
cd /tmp
ls
```

Shell to run
Comments
Change directories
List everything in /tmp

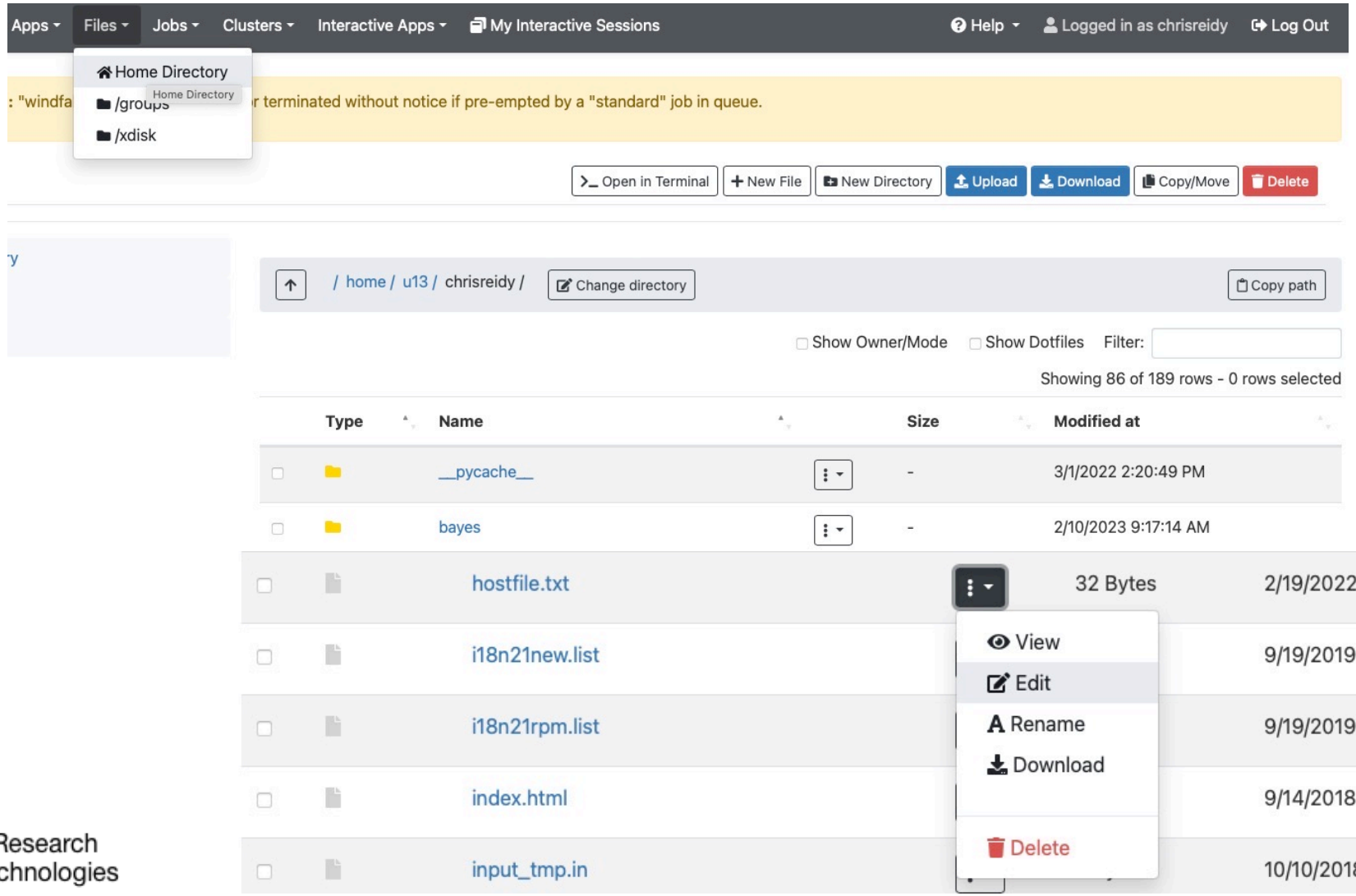


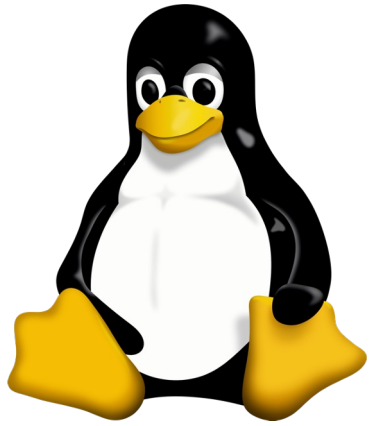
Exercise 3: Permissions and Running Bash Scripts



1. Ensure you are in the “scripts” directory
2. Use `less` to view the contents of `hello_world.txt`
3. Use `cat` to show the contents of `hello.sh` in `bash_spinup/scripts`
4. Try to run `hello.sh` by typing `./hello.sh` at the command line
5. Add execute permission to `hello.sh` using `chmod`
6. Try to run `hello.sh`

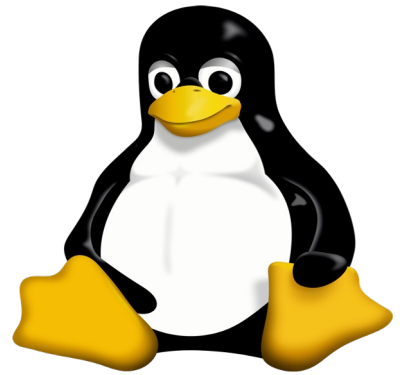
File Editing – the easy way





Research Technologies — HPC Systems

File Editing – command line



- **nano** – simple and intuitive to get started;
  - not very feature-rich;
  - keyboard driven
- **vi/vim** – universal; keyboard-driven;
  - powerful but has a learning curve
- **emacs** – keyboard or GUI versions;
  - helpful extensions for programmers;
  - well-documented
- **LibreOffice** – for WYSIWYG (what you see is what you get)

Editing with Nano

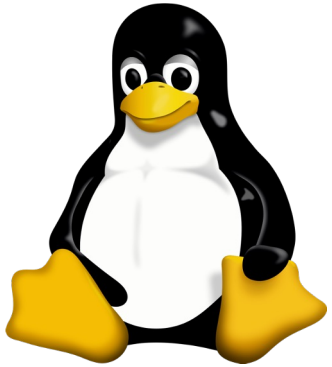
$ nano trademarks.txt

```
  GNU nano 2.3.1                File: trademarks.txt

MATLAB and Simulink are registered trademarks of The MathWorks, Inc.
Please see mathworks.com/trademarks for a list of additional trademarks.
Other product or brand names may be trademarks or registered trademarks of their respective h$




                [ Read 3 lines (Warning: No write permission) ]
^G Get Help    ^O WriteOut    ^R Read File   ^Y Prev Page   ^K Cut Text    ^C Cur Pos
^X Exit        ^J Justify     ^W Where Is    ^V Next Page   ^U UnCut Text  ^T To Spell
```



Exercise 4: File Editing with Nano

1. Edit the contents of `hello_world.txt` contents with `nano` (you can edit it to say anything!)
2. Run the program "`hello.sh`" by typing `bash hello.sh` or `./hello.sh` at the command line

More Resources

Additional Bash learning resources:

http://tldp.org/HOWTO/Bash-Prog-Intro-HOWTO.html (general)

https://www.shell-tips.com/2010/06/14/performing-math-calculation-in-bash/ *(math)*

Bash kernel for jupyter notebooks *(install anaconda first)*:

https://github.com/takluyver/bash_kernel