



HOLLAND COMPUTING CENTER

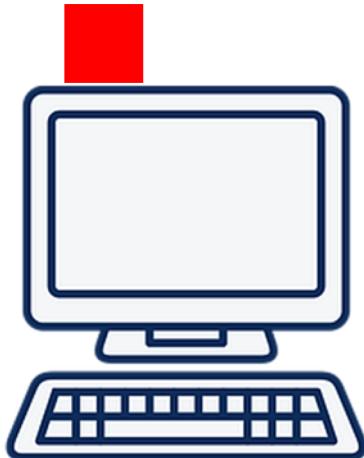
QIIME2 Workshop: Day 1
Introduction to HCC

Setup Instructions

Please complete these steps before the lesson starts at 9:00 AM:

Setup instructions: <http://hcc.unl.edu/qiime2-workshop#setup>

If you need to use a demo account – please speak with one of the helpers



If you need help with the setup, please put a **red** sticky note at the top of your laptop.



When you are done with the setup, please put a **green** sticky note at the top of your laptop.

WiFi Information

Connect to UNL-Conference

Login: QIIME2

Password: Welcome2018!

Schedule

9:00 – 10:30	Introduction to HPC Connecting to the Clusters Navigating in Bash
10:30 – 10:45	Break
10:45 – 12:00	File Manipulation Wildcards and Pipes
12:00 – 1:00	Lunch
1:00 – 2:15	Writing Reusable Scripts Running Applications on the Clusters
2:15 – 2:30	Break
2:30 – 4:00	Submitting Jobs Transferring Data with Globus

Logistics

- Name tags, sign-in sheet
- Sticky notes: **Red = need help, Green = all good**
- Link to slides:
 - [http://hcc.unl.edu/qiime2 workshop](http://hcc.unl.edu/qiime2_workshop)
 - Keep the slides open on your computer for reference.
 - **Terminal commands are in this font**

What is High Performance Computing?

High Performance Computing most generally refers to the practice of

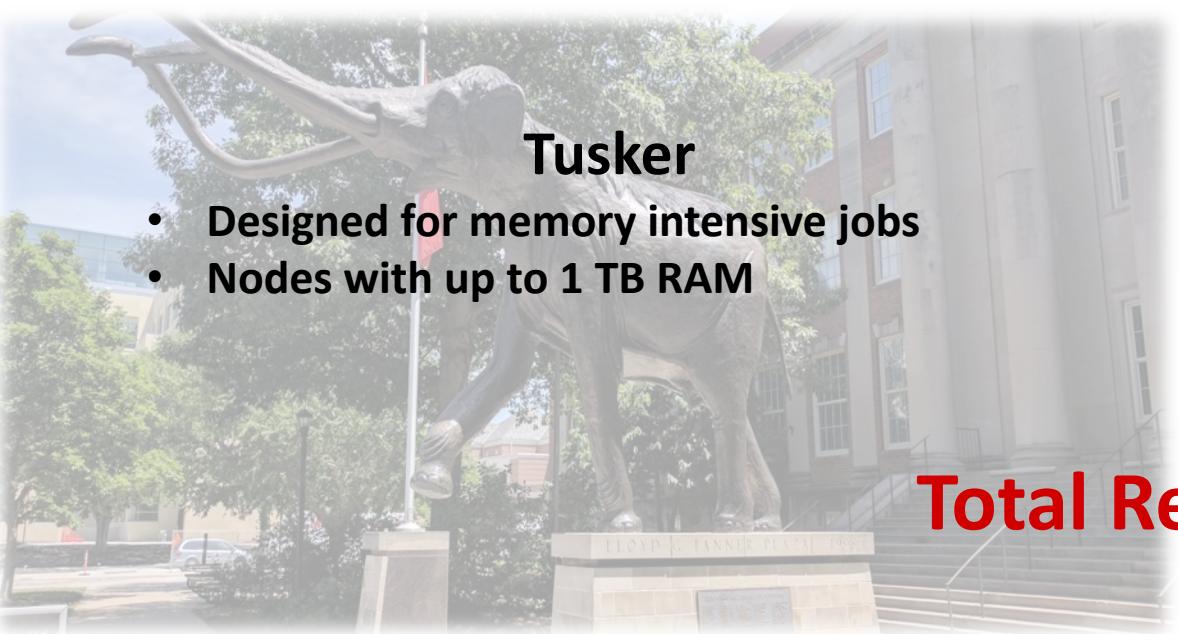
aggregating computing power

in a way that delivers much **higher performance**

than one could get out of a typical desktop computer or workstation

in order **to solve large problems** in science, engineering, or business.

<https://insidehpc.com/hpc-basic-training/what-is-hpc/>



Tusker

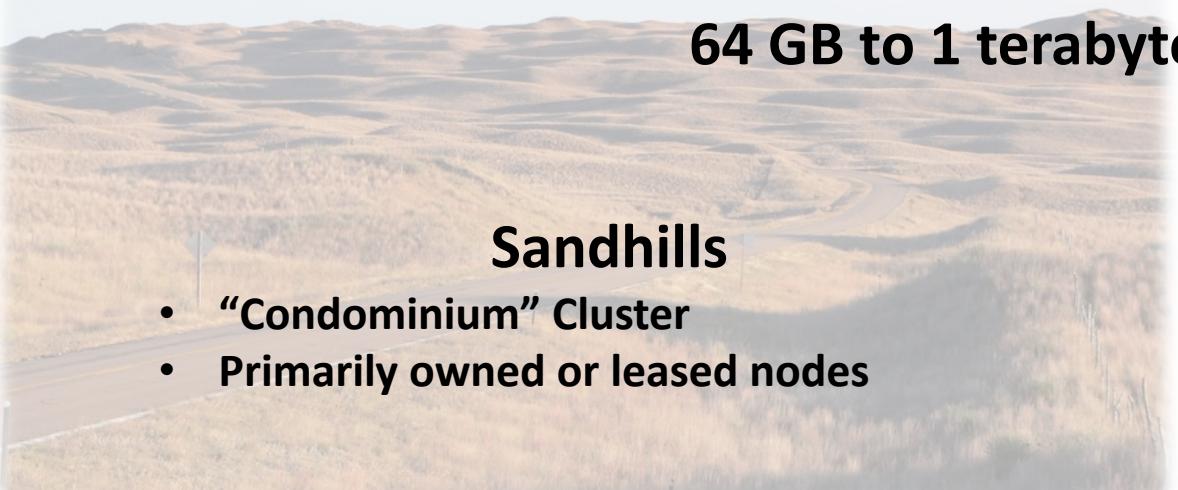
- Designed for memory intensive jobs
- Nodes with up to 1 TB RAM

Total Resources

almost 30k cores

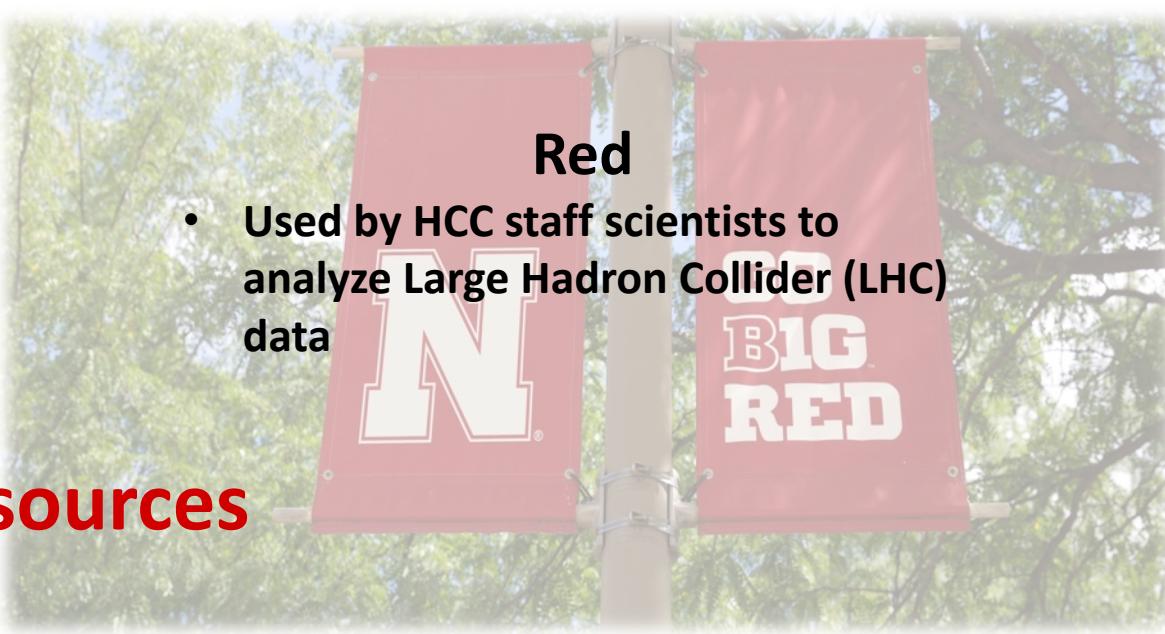
approximately 10 PetaBytes of storage

64 GB to 1 terabyte memory per node



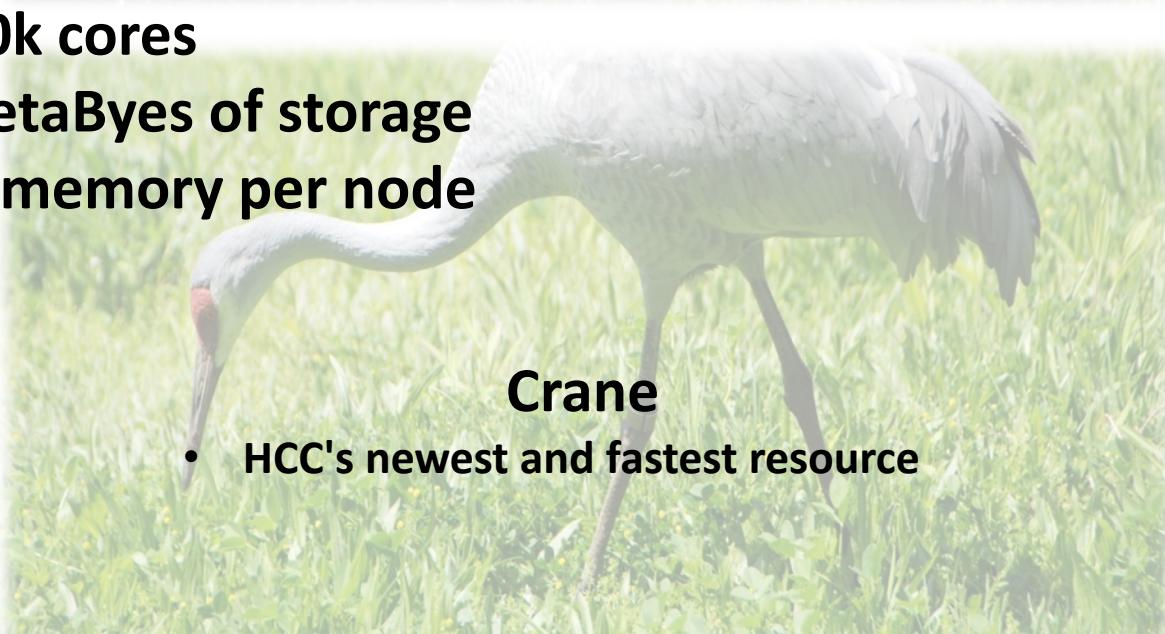
Sandhills

- “Condominium” Cluster
- Primarily owned or leased nodes



Red

- Used by HCC staff scientists to analyze Large Hadron Collider (LHC) data



Crane

- HCC's newest and fastest resource

Anvil



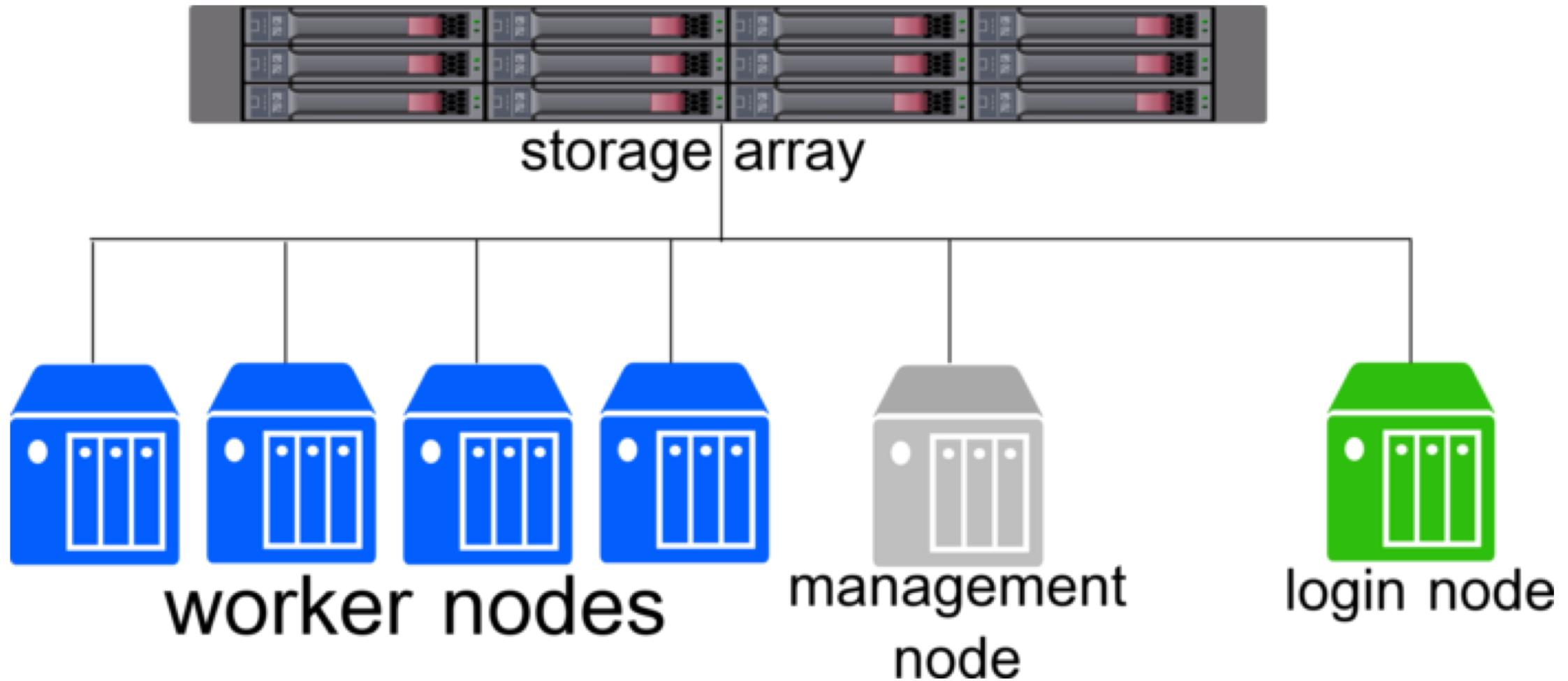
- Customizable virtual machines
- For projects not well served by a traditional Linux environment:
 - interactive environments or alternate operating systems
 - projects that require root access or dedicated resources
 - test cluster environments

Attic

- Near-line data archive
- Backed up in Lincoln and Omaha for disaster tolerance
- 10 GB/s transfer speed to and from the clusters when using Globus Connect
- Cost lower than commercial cloud services



What is a Computing Cluster?



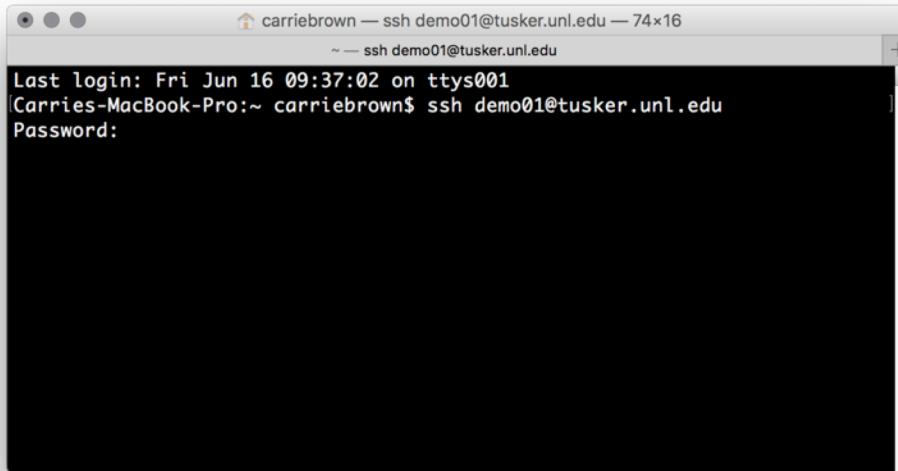
How to connect

Mac OS / Linux

- Open Terminal
- Type in the following command and press Enter:

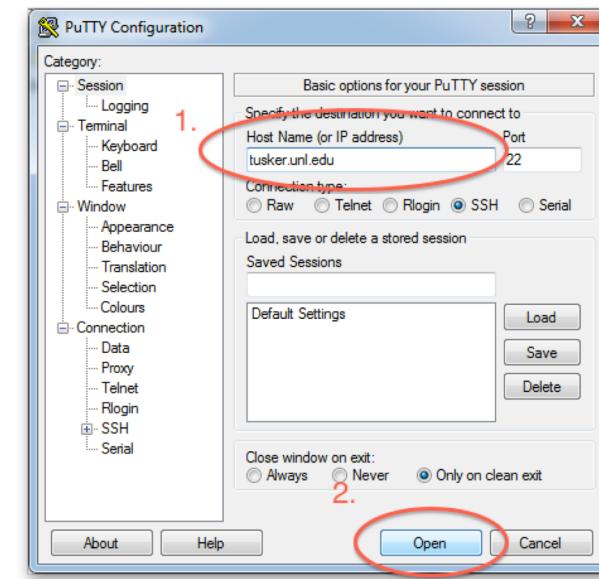
```
ssh <user_name>@tusker.unl.edu
```

(Replace <user_name> with your HCC login)



Windows

- Open PuTTY
- Type **tusker.unl.edu** for Host Name and click Open



- On the second screen, click Yes

**Once you connect, you will be prompted for your password.
Nothing will appear on your screen as you type your password. This is normal.**

Two-Factor Authentication

Enter a passcode or select one of the following options:

1. Duo Push to XXX-XXX-4833
2. Phone call to XXX-XXX-4833
3. SMS passcodes to XXX-XXX-4833

Passcode or option (1-3):

If you are using a cell phone:

- Press 1 to receive a push notification on your device
- Press 2 to receive an automated phone call. Then enter the passcode provided.
- Press 3 to receive a list of passcodes via text message. Enter the passcode which starts with the number indicated

If you are using a hardware token (Yubikey):

- Insert the Yubikey into the computer's USB slot on your computer. A green light should come on inside the gold circle.
- Press your finger against the gold circle and a passcode will be generated.

Exercises: Connecting to the clusters

1. Type `exit` to logoff of the Tusker cluster.

2. Connect to the Crane cluster
 - Try using a different Two-factor authentication method if you have another option.

Once you have finished, put up your green sticky note.

**If you have issues, put up your red sticky note and
one of the helpers will be around to assist.**

Exercises: Navigating in Bash

- Look at the man page for `ls` or using `ls --help`, what does the `-h` (`--human-readable`) option do?
- Starting from `/home/bioresearch/amy/data`, which of the following commands could Amy use to navigate to her home directory which is `/home/bioresearch/amy`?
 1. `cd .`
 2. `cd /`
 3. `cd /home/amanda`
 4. `cd ../../..`
 5. `cd ~`
 6. `cd home`
 7. `cd ~/data/..`
 8. `cd`
 9. `cd ..`

Exercises: Navigating in Bash (cont.)

- What does the command **ls** do when used with the **-l** and **-h** arguments?

Note: some of its output is about properties we do not cover in this lesson (such as file permissions and ownership), but the rest should be useful nevertheless.

- The command **ls -R** lists the contents of directories recursively, i.e. lists their sub-directories, sub-sub-directories, and so on in alphabetical order at each level.

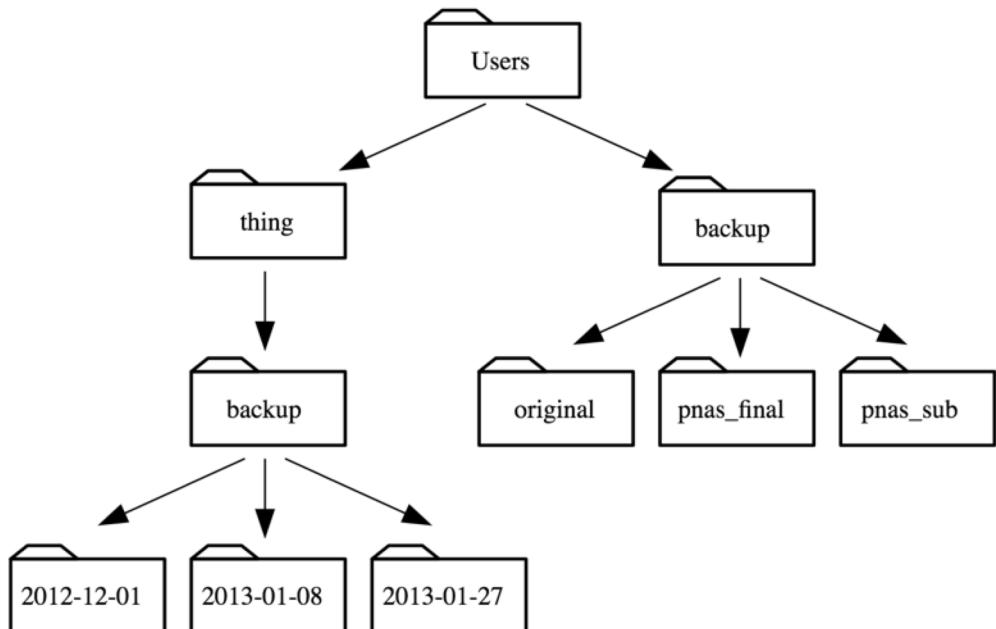
The command **ls -t** lists things by time of last change, with most recently changed files or directories first.

In what order does **ls -R -t** display things?

Hint: **ls -l** uses a long listing format to view timestamps.

Exercises: Navigating in Bash (cont.)

Using this filesystem diagram:



- if `pwd` displays `/Users/thing`, what will `ls -F .../backup` display?
 1. `../backup: No such file or directory`
 2. `2012-12-01` `2013-01-08` `2013-01-27`
 3. `2012-12-01/` `2013-01-08/` `2013-01-27/`
 4. `original/` `pnas_final/` `pnas_sub/`
- If `pwd` displays `/Users/backup`, and `-r` tells `ls` to display things in reverse order, what command will display:
`pnas_sub/` `pnas_final/` `original/`
 1. `ls pwd`
 2. `ls -r -F`
 3. `ls -r -F /Users/backup`
 4. Either #2 or #3 above, but not #1