Getting Started with the
Arkansas Research Platform

SAU Campus
February 20th

Day 1

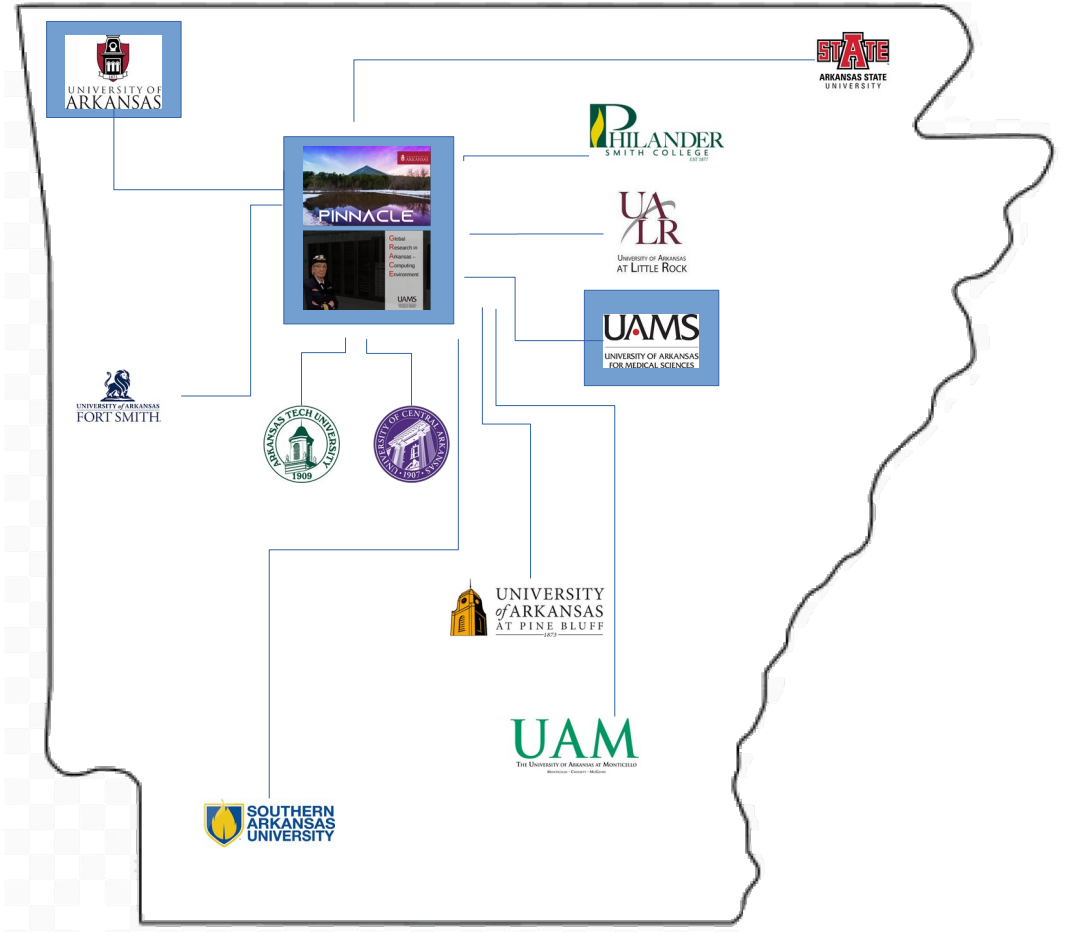
supported by UAMS High Performance Computing AHPCC Arkansas High Performance Computing Center

ARKANSAS RESEARCH PLATFORM ARP

Day 1 Schedule

8:30 am	**Registration & Set-Up** Check with each participant to ensure ability to connect to internet Make sure all required programs have been downloaded and installed
9:00 am	**Welcome & Introductions** Brief introductions: Name, institution, 1-sentence research interests
9:15 am	**HPC Overview** High Performance Computing **-** explained in 30 min
9:45 am	**What is Arkansas Research Platform (ARP)?** Brief description of ARP hardware resources When to use HPC resources Programs that can leverage HPC resources Getting familiar with the vocabulary Cluster Functional diagrams
10:15 am	**The Essentials** Getting an account on Pinnacle / Grace Logging into Open OnDemand portal OOD Overview
10:45 am	**Running Jobs – Open Ondemand** Batch job (run, modify, rerun) Interactive jobs (Jupyter notebook, VMD)
11:45 am	**LUNCH**
12:45 pm	**Moving Data** Open Ondemand File Upload/Download Terminal access in Open OnDemand wget, scp, rsync, FileZilla (GUI)
2:00 pm	Globus Log in/create Globus ID account Globus endpoints Transfer data between Pinnacle and Public endpoint Install Globus Connect Personal Transfer data between Pinnacle and personal device
4:00 pm	**Reflection and Closing Remarks** for Day 1 Where are you stuck?

Arkansas Research Platform



Arkansas Research Platform is a collaboration of higher education institutions within the state of Arkansas which provides computing resources and data storage to all its members (free of charge to students, faculty and staff). The core components are the **Pinnacle cluster** managed by the Arkansas High Performance Computing Center (UAF, Fayetteville) and the **Grace cluster** managed by UAMS High Performance Computing Center (UAMS, Little Rock).

Arkansas Research Platform – Support Staff

David Chaffin
dchaffin@uark.edu

Lawrance Tarbox
lrtarbox@uams.edu

Jeff Pummill
jpummil@uark.edu

Kaleb Abram
kzabram@uams.edu

Pawel Wolinski
pwolinsk@uark.edu

hpc-support@listserv.uark.edu

hpcadmin@uams.edu

supported by UAMS High Performance Computing AHPCC Arkansas High Performance Computing Center

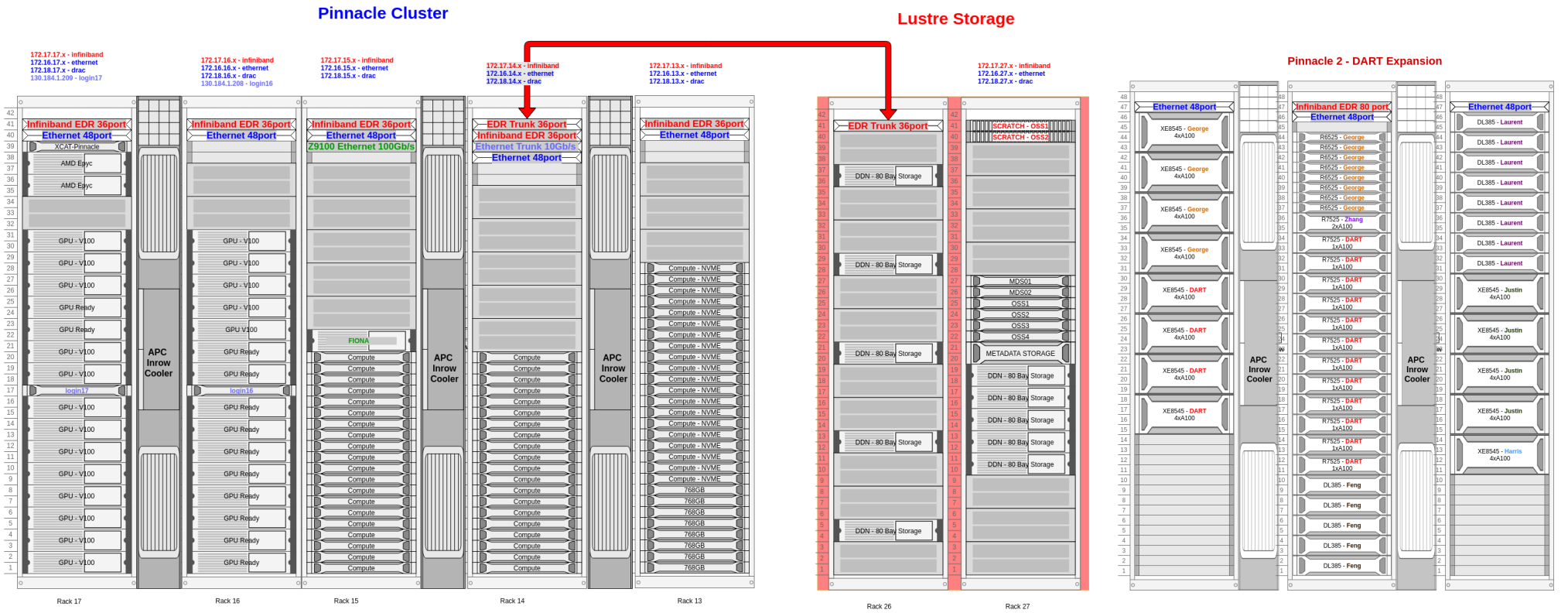
ARKANSAS RESEARCH PLATFORM ARP

Pinnacle Cluster - Fayetteville

supported by UAMS High Performance Computing AHPCC Arkansas High Performance Computing Center

ARKANSAS RESEARCH PLATFORM ARP

Pinnacle Cluster - Fayetteville



supported by UAMS High Performance Computing AHPCC Arkansas High Performance Computing Center

ARKANSAS RESEARCH PLATFORM ARP

Pinnacle Cluster – Fayetteville

Compute Nodes

100	2x Intel Xeon 6130, 2.1GHz, 32cores, 192GB RAM
50	2x AMD Epyc 7543, 2.8GHz, 64cores, 1024GB RAM
248	other/misc configuration

GPU equipped Nodes: 56

13	4x NVidia A100
1	2x NVidia A100
18	1x NVidia A100
1	2x NVidia V100
20	1x NVidia V100
1	4x NVidia TitanV
2	4x NVidia K80

Large Memory Nodes (>1TB): 4

3	2TB RAM
1	3TB RAM

Storage

2.3	PB	main storage (DDN SFA14k exascaler)
0.02	PB	scratch storage (HPE nvme/lustre)
1	PB	archive storage (Dell/lustre)
2	PB	object store (Dell/EMC ECS)

Network

Internal	100Gbps Infiniband interconnect
External	100Gbps Ethernet connection

Pinnacle Cluster – Fayetteville

Node Count	CPUs		Memory	GPUs
124	AMD	a6136 32cores	64GB	None
70	Intel	i6130 32cores	192GB	None
55	Intel	i2650v2 16cores	64GB	None
19	Intel	i6130 32cores	192GB	1xNVidia V100
18	AMD	a7543 64cores	1024GB	1xNVidia A100
14	Intel	i2650v4 24cores	128GB	None
13	AMD	a7543 64cores	1024GB	4xNVidia A100
12	Intel	i2620 8cores	32GB	None
10	Intel	i6230 40cores	384GB	None
8	AMD	a7543 32cores	512GB	None
8	AMD	a7543 64cores	256GB	None
7	Intel	i6128 24cores	768GB	None
5	AMD	a7543 64cores	512GB	None
3	AMD	o6378 64cores	512GB	None
3	Intel	i2667v4 16cores	256GB	None
3	AMD	a7543 32cores	1024GB	None
2	Intel	i2650v2 16cores	128GB	4xNVidia K80
2	AMD	a7351 32cores	256GB	None
2	Intel	a7110 64cores	112GB	None
1	Intel	i4166 24cores	192GB	4xNVidia TitanV
1	Intel	i6130 32cores	768GB	2xNVidia V100
1	AMD	a7543 64cores	512GB	2xNVidia A100
1	Intel	i6230 40cores	192GB	1xNVidia V100
1	Intel	i2620 8cores	92GB	None
1	Intel	i2620 8cores	8GB	None
1	Intel	i2620 8cores	82GB	None
1	Intel	i4627v2 32cores	768GB	None
1	Intel	i2620 8cores	72GB	None
1	Intel	i2620 8cores	62GB	None
1	Intel	i2620 8cores	52GB	None
1	Intel	i2650v4 24cores	512GB	None
1	Intel	i2620 8cores	42GB	None
1	Intel	i4860v2 96cores	3072GB	None
1	AMD	a7402 48cores	256GB	None
1	Intel	i4627v2 128cores	2048GB	None
1	AMD	a7543 64cores	2048GB	None
1	AMD	a7402 48cores	2048GB	None
1	Intel	i6128 24cores	192GB	None
1	AMD	a7452 64cores	1024GB	None

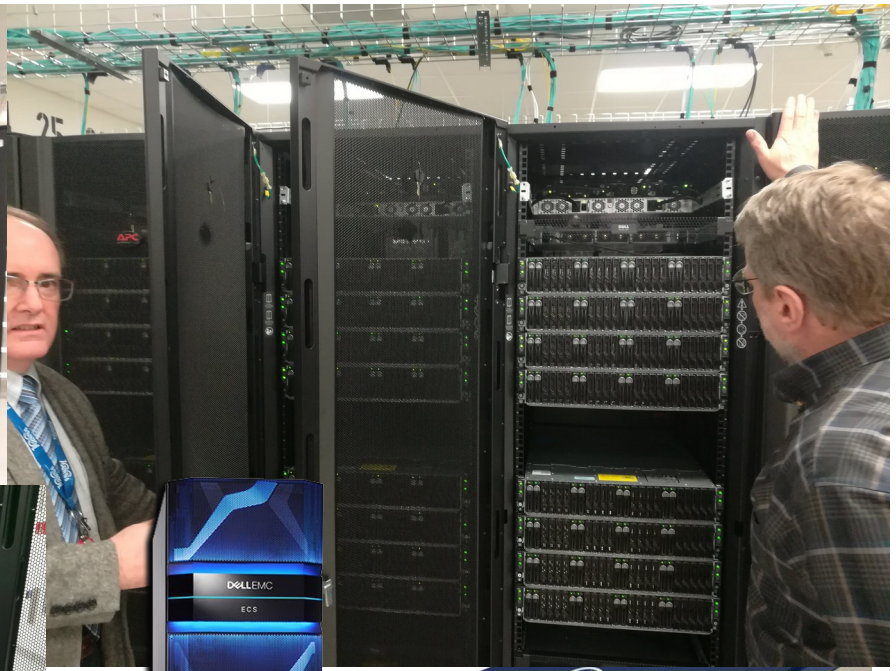
Total #of compute nodes: 398

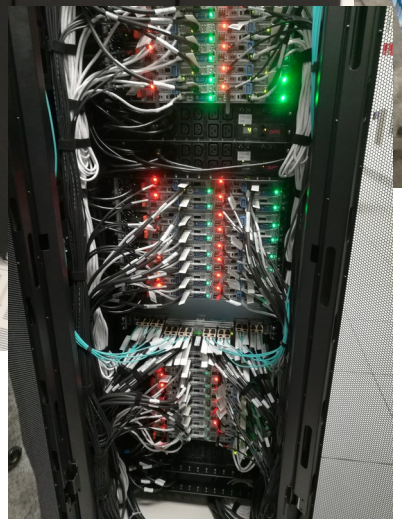
Total #of cores: 12,912

supported by UAMS High Performance Computing AHPCC Arkansas High Performance Computing Center

ARKANSAS RESEARCH PLATFORM ARP

Grace Cluster – Little Rock





supported by UAMS High Performance Computing AHPCC Arkansas High Performance Computing Center

ARKANSAS RESEARCH PLATFORM ARP

Grace Cluster – Little Rock

Compute Nodes

96	2x Intel Xeon E5-2680v4, 2.40GHz, 56cores, 128GB RAM
80	Intel Xeon Phi 7230F, 1.30GHz, 64 cores, 384GB RAM
16	Intel Xeon Phi 7230F, 1.30GHz, 64 cores, 192GB RAM
3	2x Intel Xeon E5-2680v3, 2.50GHz, 48cores, 128GB RAM, 2x NVidia P100 GPUs
1	2x Intel Xeon E5-2680v3, 2.50GHz, 48cores, 128GB RAM, 2x NVidia K40 GPUs

Other Nodes

7	management/login/storage interface

Storage

1.9PB	high speed storage (DDN GS14KX GridScaler. IBM Spectrum Scale 5)
1.0PB	NAS intermediate storage
4.2PB	object store (Dell/EMC ECS)

Network

Internal	100Gbps Omnipath interconnect
External	redundant 10Gbps Ethernet connections

Total #of compute nodes: 196

Total #of cores: 11,712

Coming soon:
10 additional nodes, 640 AMD Epyc cores, 9 A100 GPUs, 100Gbs ethernet

When to use HPC resources

? - word processing
? - editing spreadsheets
? - editing images
? - plotting graphs
? - storing personal images
? - storing personal movies
? - bitcoin mining
? - playing multi-player online games
? - GPU accelerated programs
? - machine learning
? - running parallel multi-node simulations
? - visualizing large data sets
? - moving large data sets
? - processing large data sets

When to use HPC resources

NO - word processing
NO - editing spreadsheets
NO - editing images
NO - plotting graphs
NO - storing personal images
NO - storing personal movies
NO - bitcoin mining
NO - playing multi-player online games
YES - GPU accelerated programs
YES - machine learning/AI
YES - parallel multi-node simulations
YES - visualizing large data sets
YES - moving large data sets
YES - processing large data sets

HPC Applications

Software on Pinnacle

Molecular dynamics: namd, lammps

Quantum chemistry: quantum espresso, VASP, Gaussian, PQS, NWChem

Multiphysics: COMSOL, Ansys

Data processing programming languages: Python, R

Simmulation frameworks: Mathemaica, Matlab

Statistics: SAS

Bioinformatics: abyss, bamtools .. spades, samtools, trinity...
(Bioinformatics Software on Pinnacle)

HPC Vocabulary

cluster – multiple computers connected by a network, with access to the same file system.

node – a single computer

login node – a gateway node to the cluster. Shared by all users. No computationally intensive tasks are permitted

compute node – node which can run computationally intensive tasks

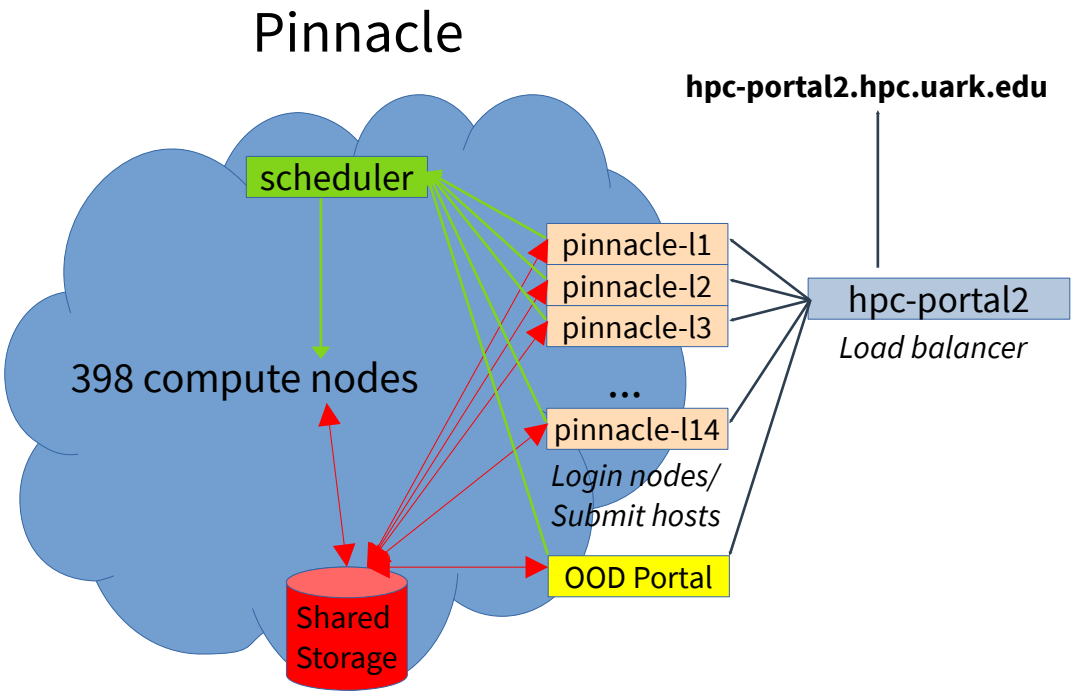
queueing system – software which organizes and schedules jobs submitted by users to be executed on compute nodes

job – a computational task submitted by a user to the queueing system

cpu/core/thread/socket – central prcossing unit (CPU)is a mircoprocessor on a single integrated circuit with 2 or more processing units (cores). Core can run a single instruction at a time. With Hyperthreading the same core can run 2 instructions at once. Socket is a interface on the motherboard which receives a CPU. HPC motherboards usually have 2 – 4 sockets.

HPC Vocabulary - fuller list

Pinnacle Cluster – functional diagram



scheduler – coordinates all of the nodes in the queue (SLURM queueing system)
submit hosts – nodes from which users submit jobs to the queues
login hosts – nodes accessible from the external network
compute nodes – nodes that assigned to queues and run jobs

Getting Accounts on ARP (Pinnacle and Grace)

Pinnacle (UAF, Fayetteville)

Account application:	https://hpc.uark.edu/hpc-support/user-account-requests
Access:	https://hpcwiki.uark.edu/doku.php?id=off-campus_access_dmz
Website:	https://hpc.uark.edu/

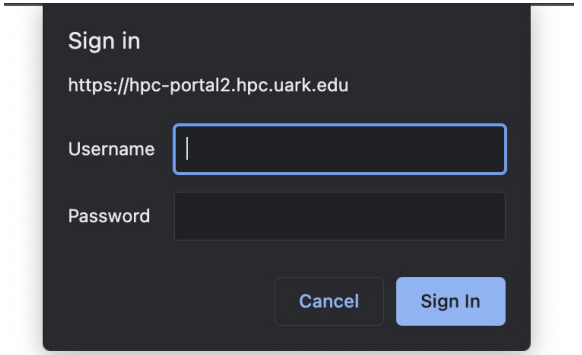
Web Portal - Open Ondemand

Open OnDemand (OOD) is a web portal interface to HPC resources, which hides the complexities of the HPC scheduling system and Linux/bash command line interface. Also provides a convenient method of displaying desktop/GUI applications running on compute nodes on the users client desktop.

Pinnacle OOD portal

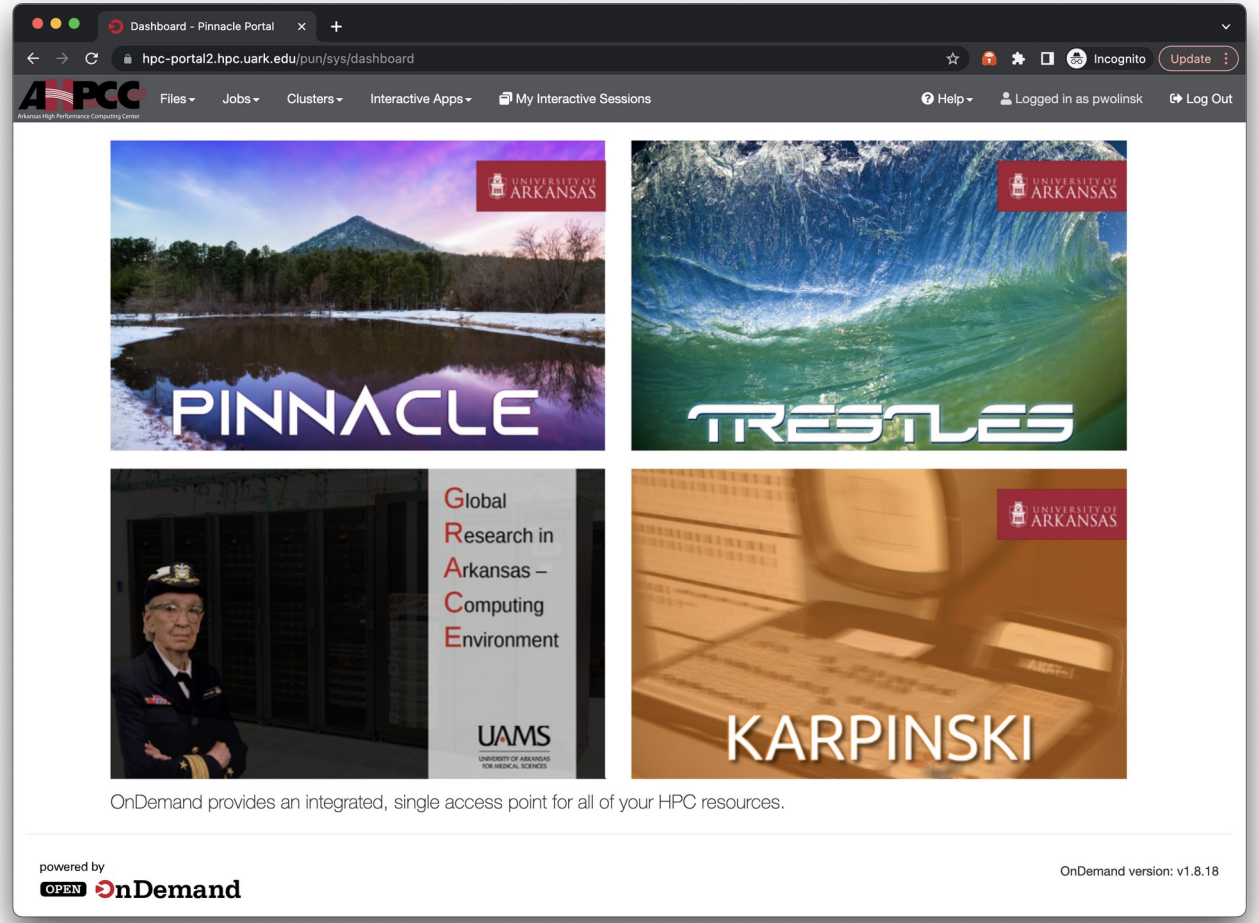
https://hpc-portal2.hpc.uark.edu/

Access from outside of UAMS campus network: I
https://hpcwiki.uark.edu/doku.php?id=off_campus_access_dmz



Usernames and passwords supplied in the login instruction emails when Pinnacle accounts were created.

Open Ondemand Portal – Pinnacle



Open Ondemand Portal – First HPC job

Submit batch job

Top Menu: Jobs→Job Composer

- +New Job button
- Edit Files button
-
- Select ‘myjob.sh’
- View button

```
#!/bin/bash
# JOB HEADERS HERE

echo "Hello World"
```

- Close browser tab
- Submit button

View results

Wait for “completed status”

- Edit Files button
- Select ‘slurm-??????.out’
- View button

```
Hello World
```

Open Ondemand Portal – First HPC job

Modify job script and resubmit

- Edit Files button
- Select 'myjob.sh'
- Edit button

```
#!/bin/bash
# JOB HEADERS HERE

echo "Hello World"
echo "running on host"
hostname
echo "on "
date
```

- Save button
- Close browser tab
- Submit button

View Results

Wait for "completed status"

- Edit Files button
- Select 'slurm-??????.out'
- View button

```
Hello World
running on host
c1405
on
Sun Sep 25 19:21:58 CDT 2022
```

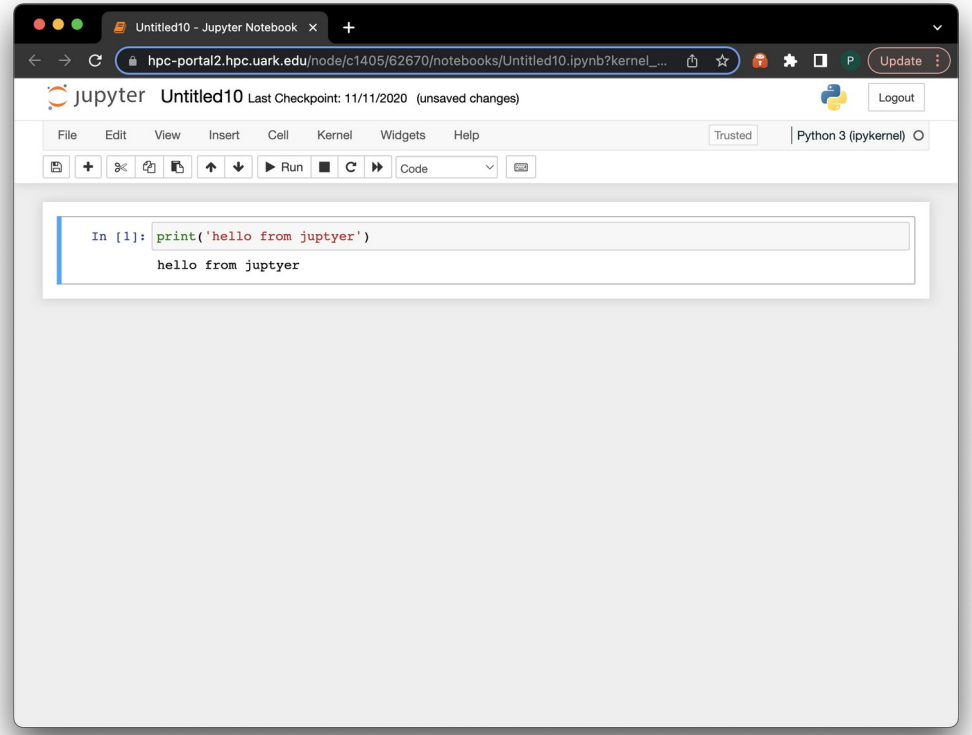
Open Ondemand Portal – interactive job

Interactive Jupyter Notebook job

- Top Menu: Interactive Apps → Jupyter Notebook
- Select “comp01, 1 hour, 1 node”
- Launch button
- Wait for the status “Running”
- Connect to Jupyter button

Interact with the Notebook

- New button → Python3 (ipykernel)
- Type in “print(‘hello’)” into In [] field
- <CTRL>-Enter to run



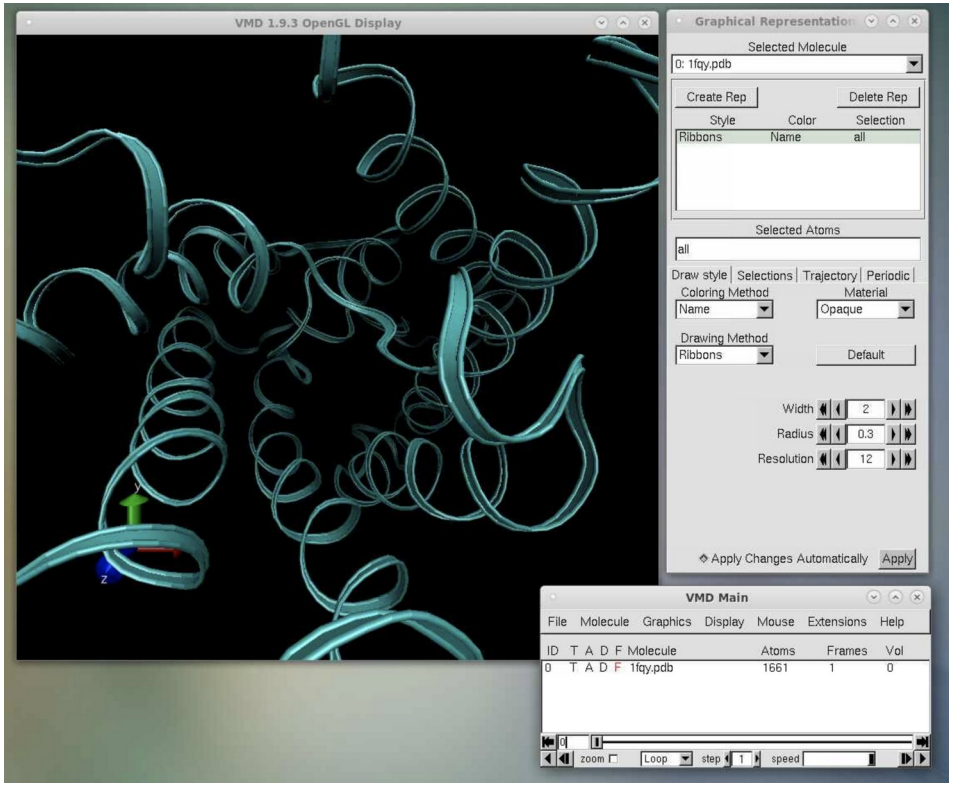
Open Ondemand Portal – interactive job

Interactive VMD job

- Top Menu: Interactive Apps → VMD
- Select "Queue: comp01, Number of hours: 1"
- Launch button
- Wait for the status "Running"
- Launch VMD button

Interact with VMD

- At the VMD prompt Type: "menu main on"
- In the menu "File→New Molecule"
- In the "Molecule File Browser", "File Name": */scrfs*/ahpcc/ARP-workshop/1fqy.pdb
- Load button
- In the menu "Graphics→Representations:" Drawing method: Ribbons



Open Ondemand Portal – Moving Data

File dialog window

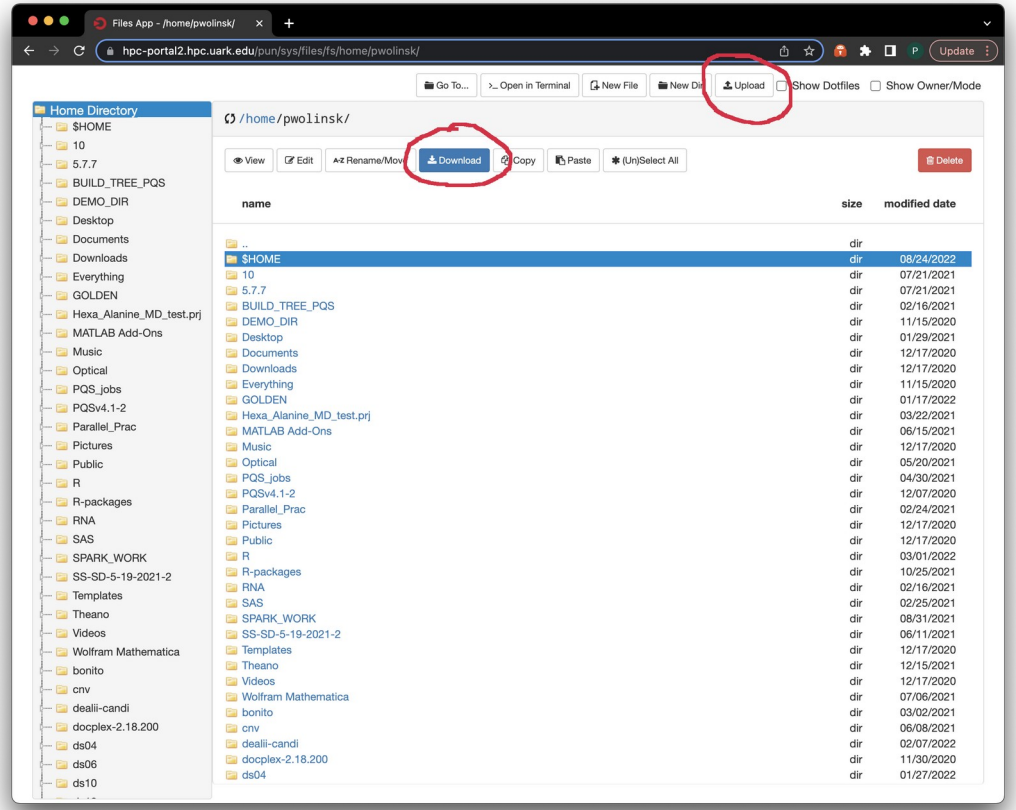
- Top Menu: **Files** → **Home Directory**
  New browser tab opens with the contents of your home directory

- **Upload** button – select file from your local machine and upload it to Pinnacle

- **Download** button – select file from your Pinnacle account and download it to your local machine

  Using those two buttons:

1) *download file etc/redhat-release to your laptop*
2) *Upload any text file to your home directory on Pinnacle*

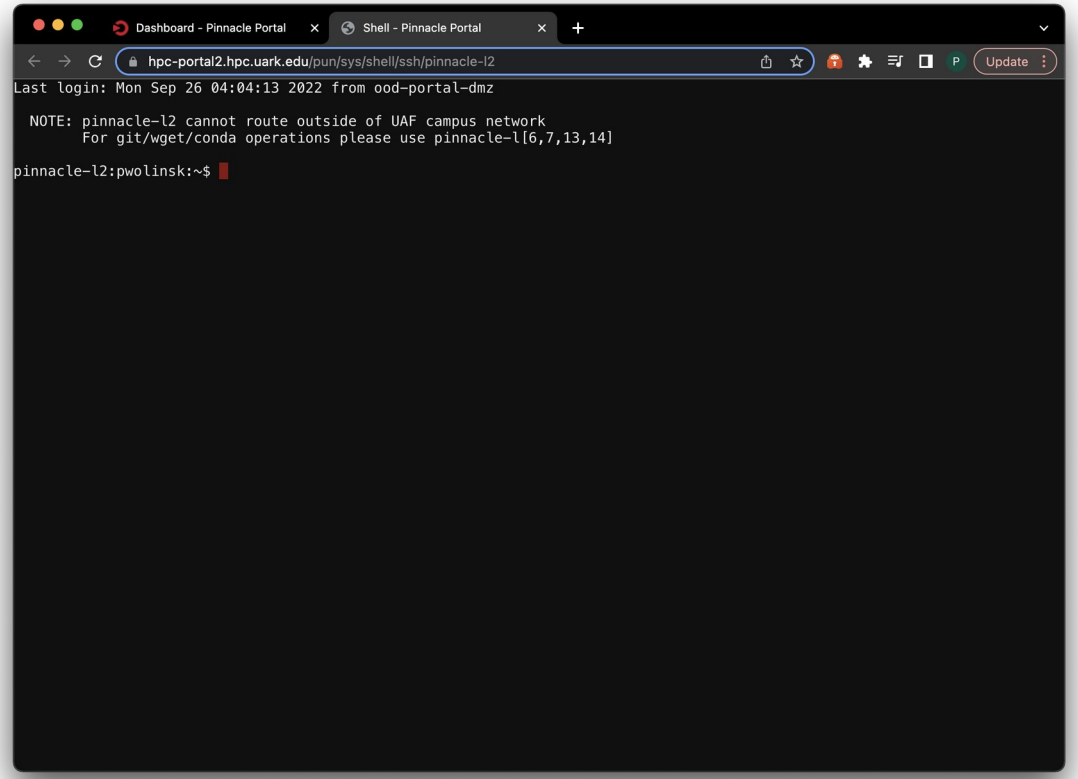


Open Ondemand Pinnacle – Terminal

Terminal

OOD Portal on Pinnacle: https://hpc-portal2.hpc.uark.edu

- Top Menu: **Clusters** → **Pinnacle**
  New browser tab opens with a terminal on a Pinnacle login node



Moving Data – Pinnacle/Grace - rsync

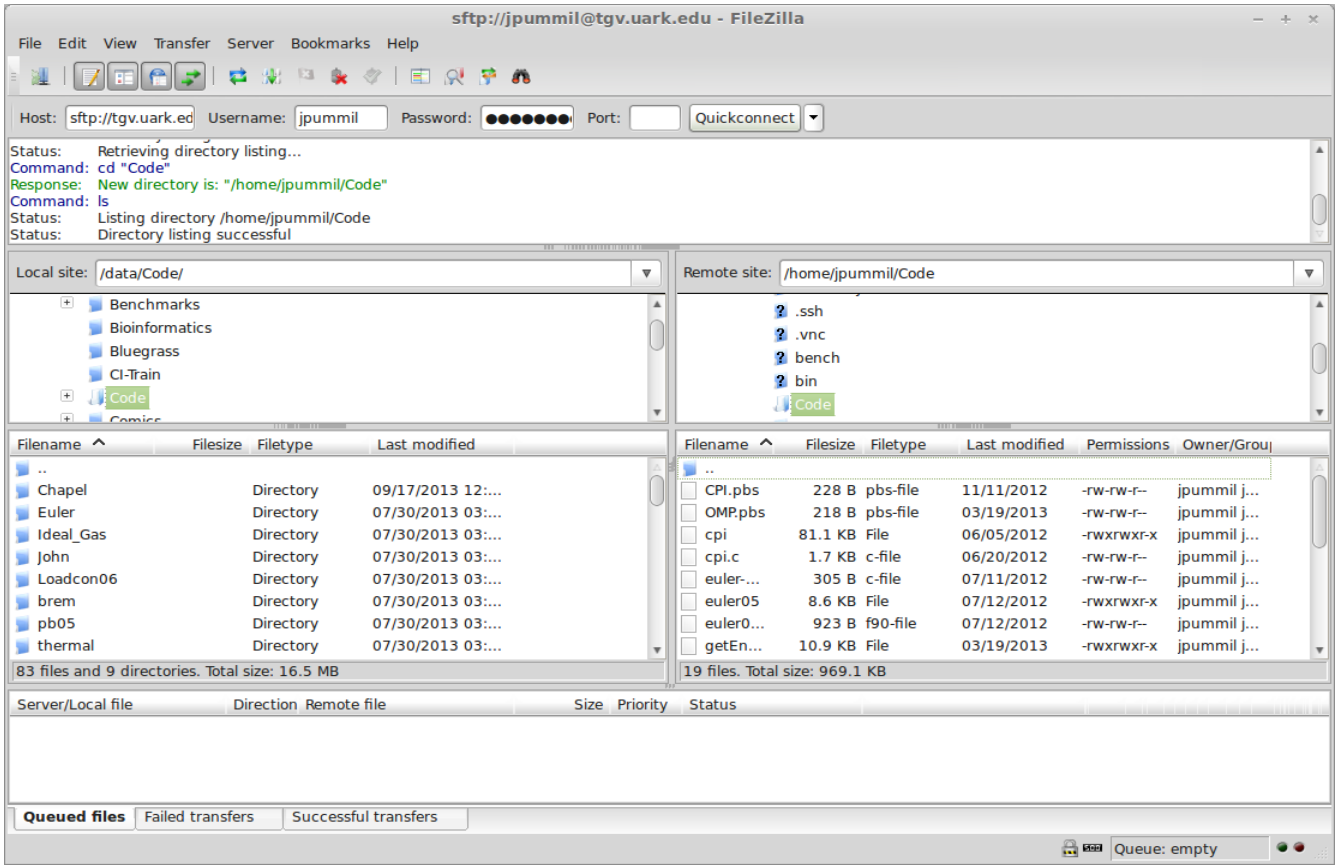
rsync - synchronize files between local and remote hosts. Only moves files which have been updated.

rsync -avz <source> <destination>

-avz options stand for archive, verbose, compress during transfer. Either source or destination can be of from **username@host:file_path**

```
On Grace [login1]
[pwolinsk@login1 ~]$ scp -r slurm-pinnacle/ pwolinsk@hpc-portal2.hpc.uark.edu:
cgroup.conf.example                                          100%  216    20.1KB/s   00:00
first-slurm-startup.sh                                       100%  244     4.2KB/s   00:00
gres.conf                                                    100%   15KB 713.6KB/s   00:00
prolog                                                       100% 3157   169.9KB/s   00:00
…
epilog                                                       100% 1338   133.1KB/s   00:00
slurmdbd.conf.example                                        100%  739    12.5KB/s   00:00
cgroup.conf                                                  100%  216     3.7KB/s   00:00
[pwolinsk@login1 ~]$ echo "some text" >> slurm-pinnacle/prolog
[pwolinsk@login1 ~]$ rsync -avz slurm-pinnacle pwolinsk@hpc-portal2.hpc.uark.edu:
sending incremental file list
slurm-pinnacle/
slurm-pinnacle/cgroup.conf
...
slurm-pinnacle/slurm.conf.example
slurm-pinnacle/slurmdbd.conf.example
slurm-pinnacle/layouts.d/
slurm-pinnacle/layouts.d/power.conf.example
slurm-pinnacle/layouts.d/power_cpufreq.conf.example
slurm-pinnacle/layouts.d/unit.conf.example
sent 1,058 bytes  received 1,310 bytes  4,736.00 bytes/sec
total size is 118,072  speedup is 49.86
[pwolinsk@login1 ~]$
```

Moving Data – Pinnacle/Grace – GUI



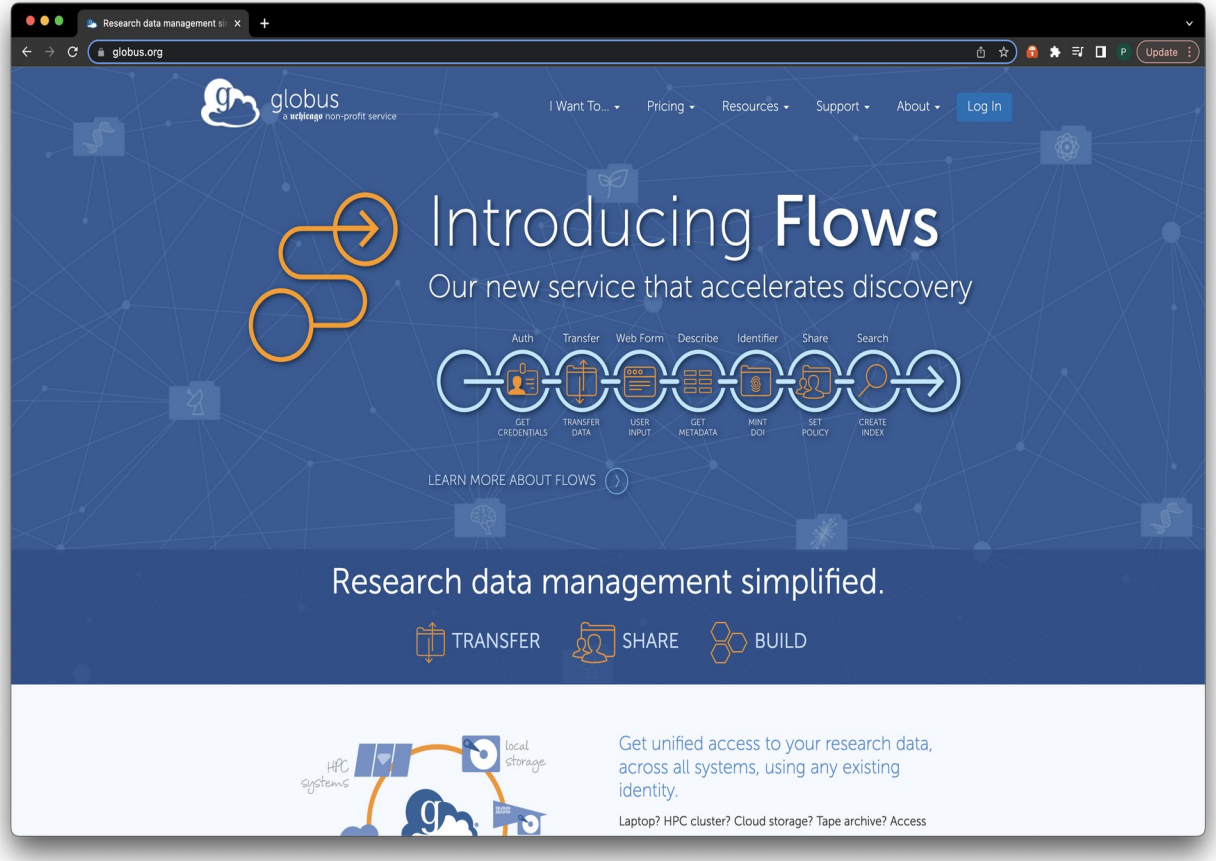
https://filezilla-project.org/download.php

Moving Data – Large Data Transfers - Globus

Globus – efficient, large data transfer and data sharing service.

Process:

1. log into Globus
2. connect to two Globus end points
3. transfer data between end points

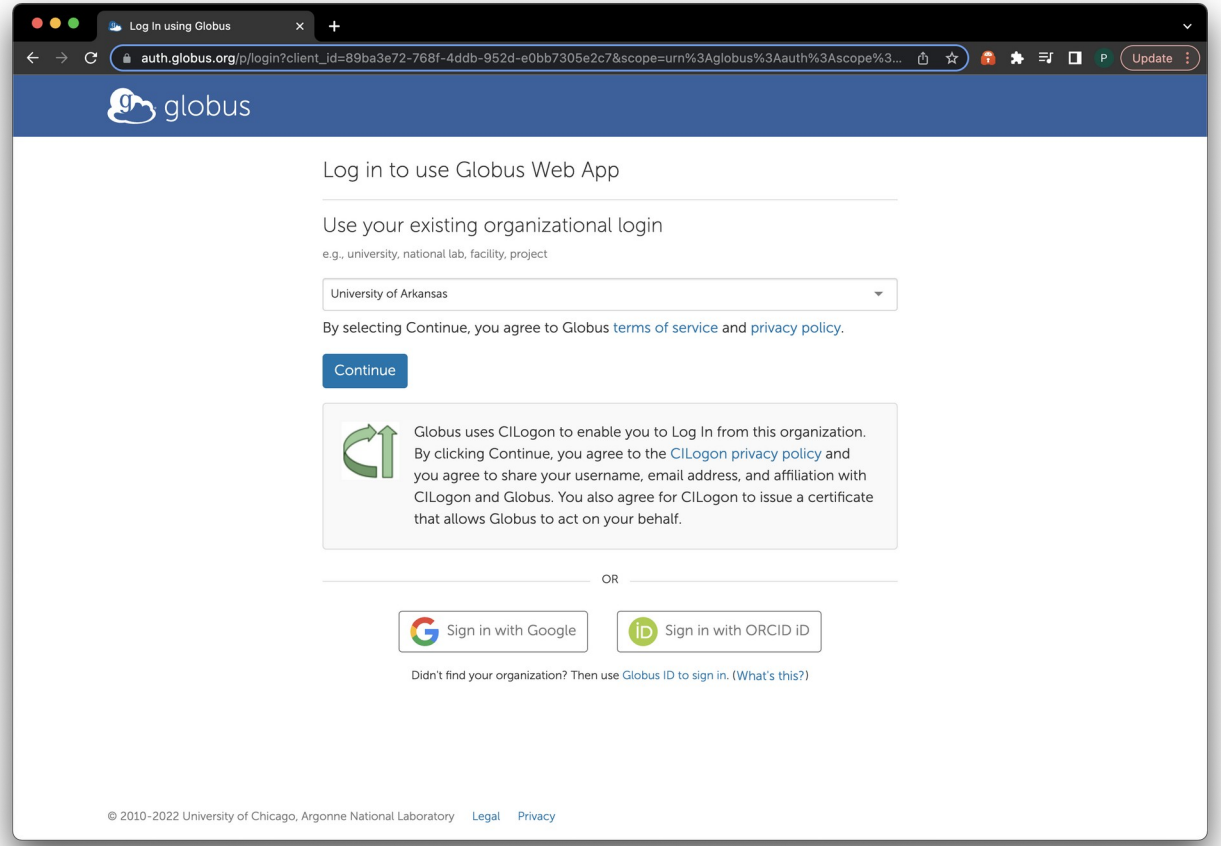


https://globus.org

Moving Data – Large Data Transfers – Globus – Log in

If your institution is listed in the drop down menu, login with those credentials. Google accounts and ORCID ID accounts are an option. Otherwise create a Globus ID account (link at the bottom of page).

In addition to other services, Globus is also an identity provider and maintains its own user accounts.

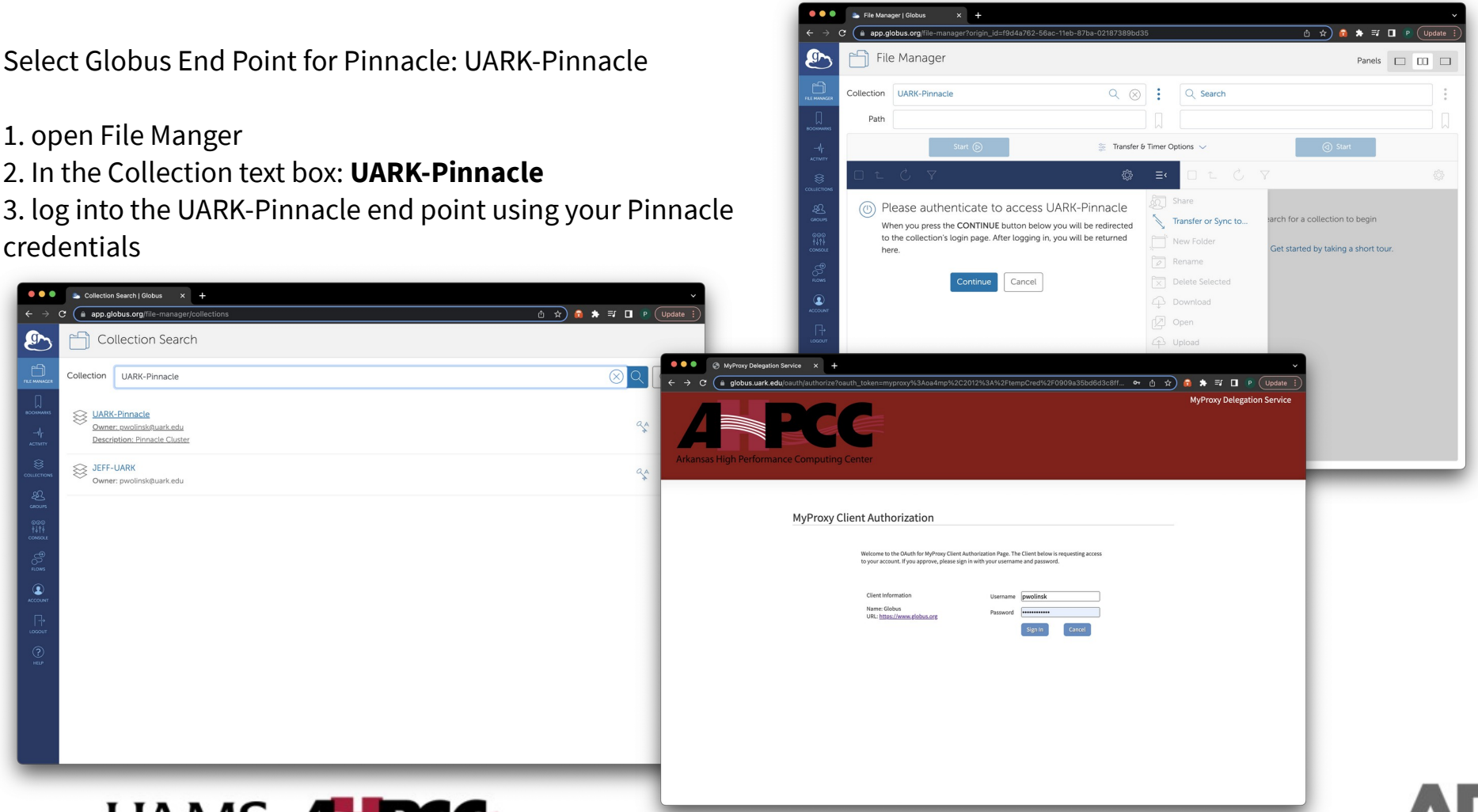


https://globus.org

Moving Data – Large Data Transfers – Globus – Log in

Select Globus End Point for Pinnacle: UARK-Pinnacle

1. open File Manger
2. In the Collection text box: **UARK-Pinnacle**
3. log into the UARK-Pinnacle end point using your Pinnacle credentials

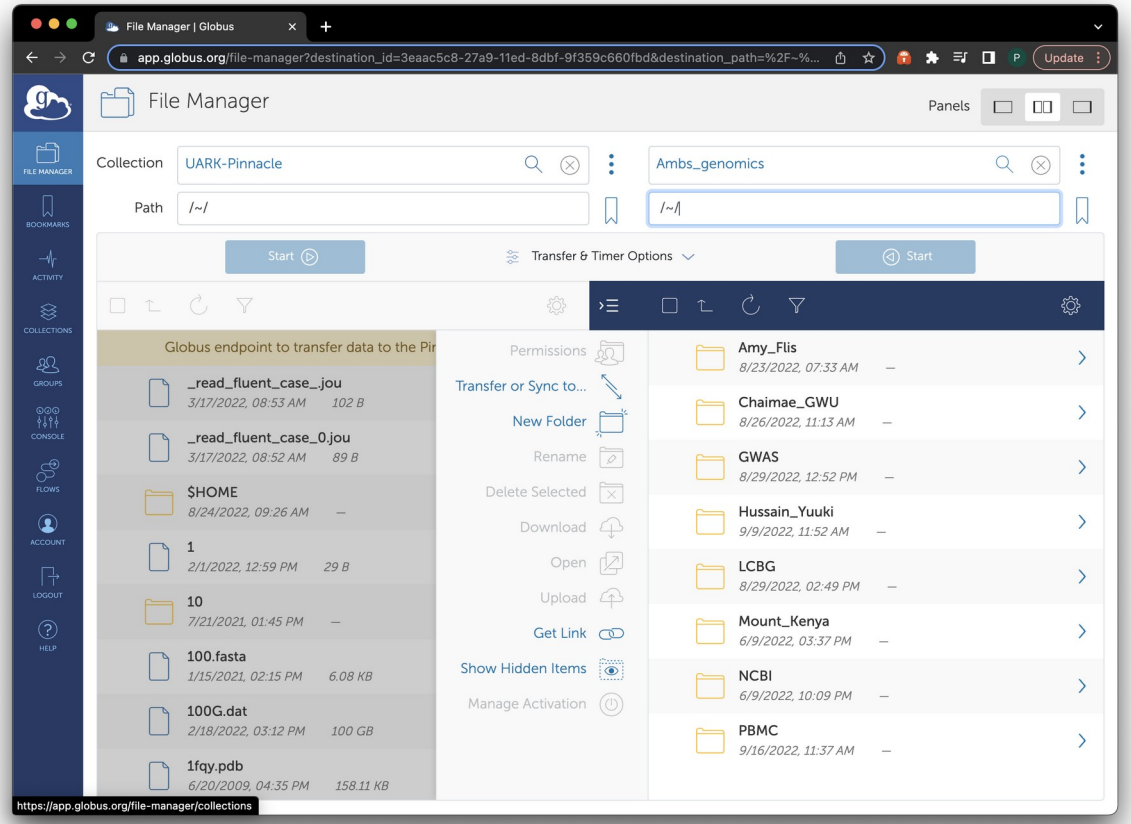


Moving Data – Large Data Transfers – Globus

Select another end point: Ambs_genomics

1. In the other Collection text box:
Ambs_genomics
2. double click on NCBI folder
3. select file “README”
4. Start button

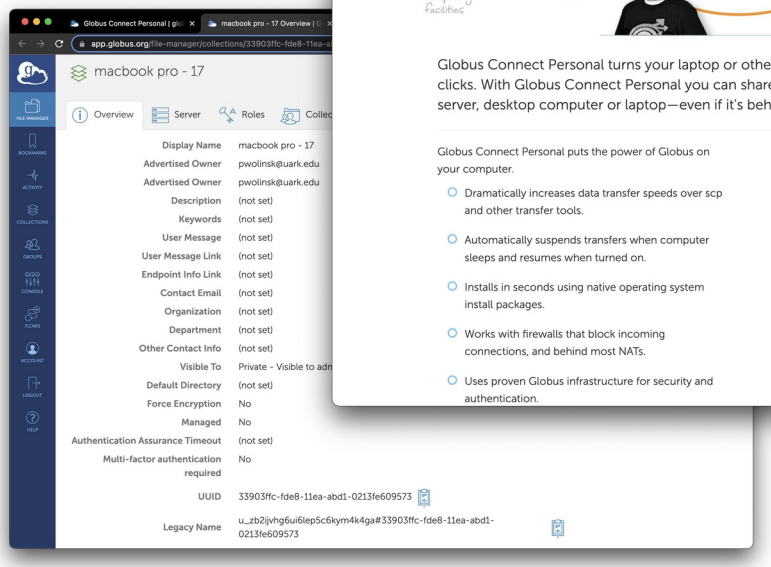
When the data transfer is complete, you’ll see the README file on the UARK-Pinnacle side (and in your HOME directory on Pinnacle)



Globus – Creating a Personal Endpoint

To move data using Globus between your personal device and any Globus end point you have access to, you have to install Globus Connect Personal server on your personal device.

It's a server that runs on your local machine and registers with your Globus account. Once set up and manually started, a new endpoint for your device will be available to you (only you) in the Globus collections.

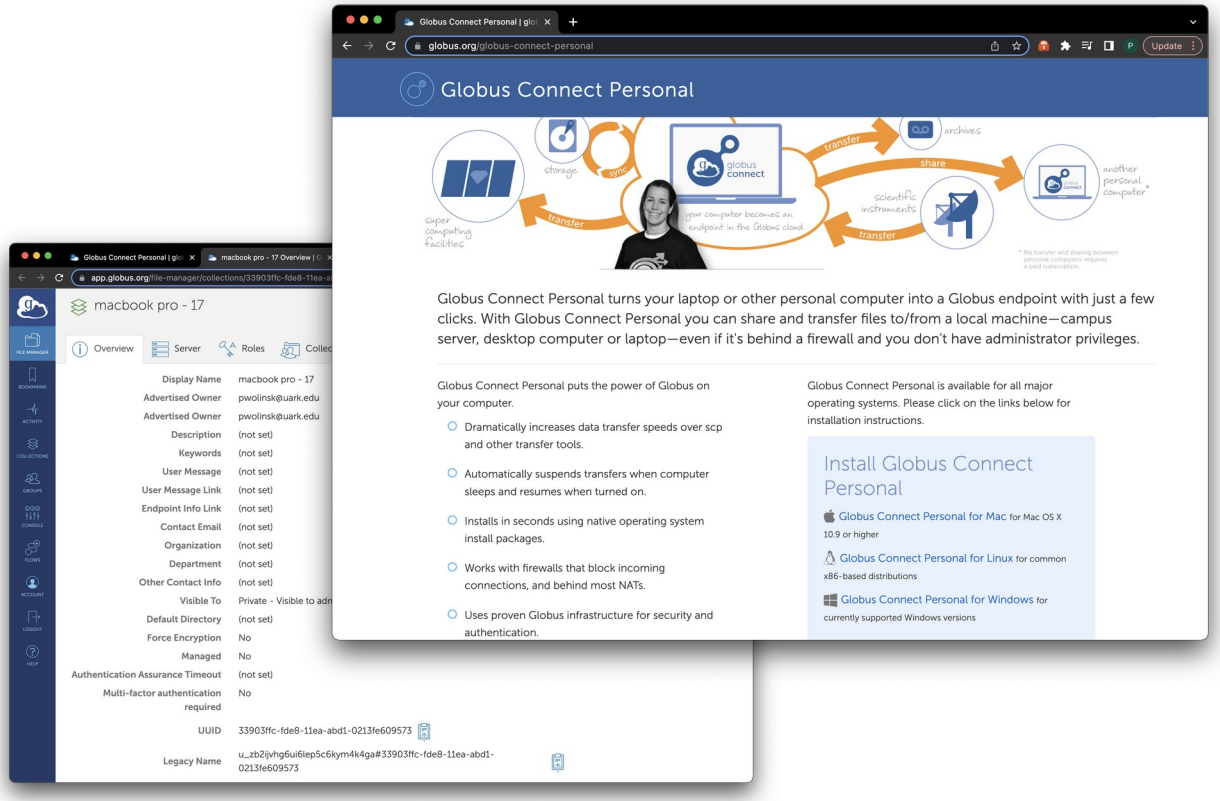


http://globus.org/globus-connect-personal

Globus – Using a Personal Endpoint/Pinnacle

To move data using Globus between your personal device:

1. start your Globus Connect Personal server on your device
2. click on the Glonus Connect Personal server icon and select **Web: Transfer Files**
3. A browser window with Globus File manager will appear with your local endpoint opened on one side
4. select another Globus end point and start moving data



http://globus.org/globus-connect-personal

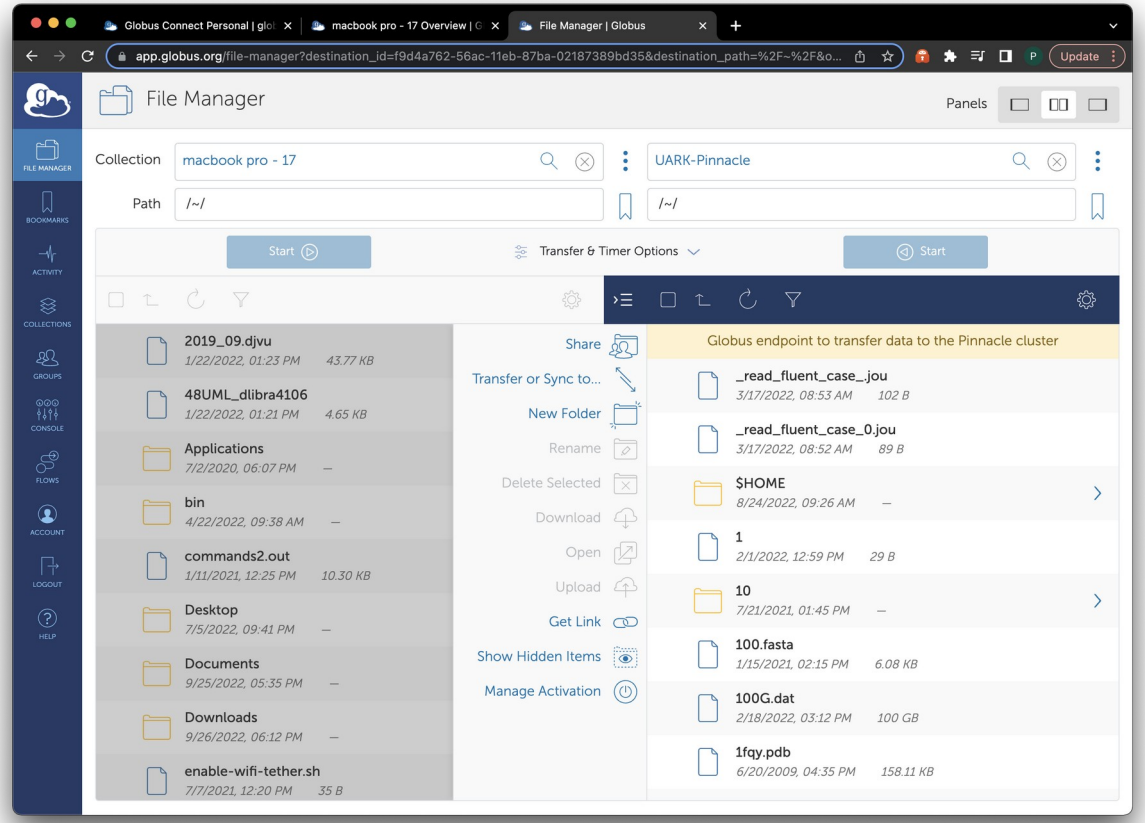
supported by UAMS High Performance Computing AHPCC Arkansas High Performance Computing Center

ARKANSAS RESEARCH PLATFORM ARP

Globus – Using a Personal Endpoint/Pinnacle

To move data using Globus between your personal device:

1. start your Globus Connect Personal server on your device
2. click on the Glonus Connect Personal server icon and select **Web: Transfer Files**
3. A browser window with Globus File manager will appear with your local endpoint opened on one side
4. select another Globus end point and start moving data



http://globus.org/globus-connect-personal

Getting Started with the Arkansas Research Platform

SAU Campus
February 20

Day 2

supported by UAMS High Performance Computing AHPCC Arkansas High Performance Computing Center

ARKANSAS RESEARCH PLATFORM ARP

Day 2 Schedule

9:00 am	**SSH access**
	SSH clients (Windows/MacOS/Linux)
	Logging into Pinnacle and Grace
	Network Accessibility
	Pinnacle's 2-step login on the DMZ network
9:30 am	**Intro to Command Line Interface (CLI)**
	Bash shell
	Environment
	Basic linux commands
	File systems
	/home and /scratch
10:00 am	**Queueing System**
	Queues
	Running a sample job
	Interactive job
	Batch job
	Job arrays
10:30 am	**Software modules**
	Modules
	Conda environments
	Conda – GRACE deep dive
11:00 am	**Moving and Storing Data, part 2**
	Basic object storage principles; What is ROSS?
	Request access to ROSS – identities and secret keys
	ecs-sync
12:00 pm	**LUNCH**
1:00 pm	**Individual project help**
	Questions, Tell me more, and Bring your own project (optional for attendees)

Secure Shell Client

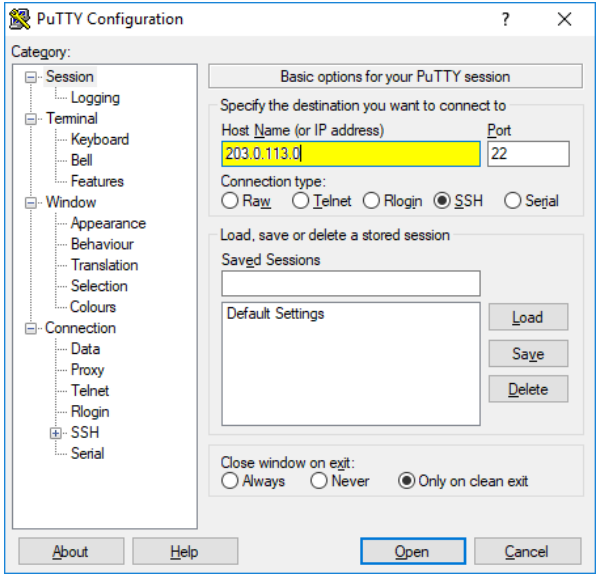
Secure Shell (SSH) is a protocol for exchanging information between two nodes on a network: an SSH server and an SSH client. SSH client requests a connection from and SSH server. Client runs only when a connection is established. Server runs continuously waiting for connection requests.

An SSH server is running on each of the login nodes on the ARP clusters (as well on each of the compute nodes).

Linux, Mac
SSH client is built in! (server too)

Windows
Windows Subsystem for Linux (WSL): Open PowerShell and run “wsl --install”
OR
1. please open this page in a browser:
   http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html
   (first link in google search with for keyword “putty”)

2. and right click and save to the desktop:
   putty.exe, pscp.exe and plink.exe



http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html

Secure Shell – log into Pinnacle and Grace

ssh user@hostname - log into a remote system

Linux

1. open a terminal and run
2. run:

   *ssh <username>@**hpc-portal2.hpc.uark.edu*** *Pinnacle*

   *ssh <username>@**login.hpc.uams.edu*** *Grace*
3. enter your supplied password when prompted.

Windows

1. double click on ***putty.exe*** on the desktop
2. specify

   hpc-portal2.hpc.uark.edu *Pinnacle*

   login.hpc.uams.edu *Grace*

   as “Hostname” and click on “Open” button.
3. enter your username (w/o ‘@uark.edu’ or ‘@uams.edu’)
4. enter your supplied password

ssh <username>@hpc-portal2.hpc.uark.edu *Pinnacle*

ssh <username>@login.hpc.uams.edu *Grace*

Network Accessibility – Pinnacle & Grace

Pinnacle

Network Location	Access Type		Server DNS	Availability
UARK Campus/UARK VPN	SSH	1	pinnacle.uark.edu	Open
UARK Campus/UARK VPN	HTTPS	2	pinnacle-portal.uark.edu	Open
UARK Campus/UARK VPN	SSH/HTTPS	3	hpc-portal.uark.edu	Open
UARK Campus/UARK VPN	SSH/HTTPS	4	hpc-portal2.hpc.uark.edu	Open
Other UASYS Campus	SSH	1	pinnacle.uark.edu	Open
Other UASYS Campus	HTTPS	2	pinnacle-portal.uark.edu	Open
Other UASYS Campus	SSH/HTTPS	3	hpc-portal.uark.edu	Open
Other UASYS Campus	SSH/HTTPS	4	hpc-portal2.hpc.uark.edu	Open
Commodity Internet	SSH	1	pinnacle.uark.edu	**Closed**
Commodity Internet	HTTPS	2	pinnacle-portal.uark.edu	**Closed**
Commodity Internet	SSH/HTTPS	3	hpc-portal.uark.edu	**Closed**
Commodity Internet	SSH/HTTPS	4	hpc-portal2.hpc.uark.edu	Open (2-step login)

Grace

Network Location	Access Type		Server DNS	Availability
UAMS Campus/UAMS VPN	SSH	1	login.hpc.uams.edu	Open
UAMS Campus/UAMS VPN	HTTPS	2	portal.hpc.uams.edu	Open
Other UASYS Campus	SSH	1	login.hpc.uams.edu	**Closed**
Other UASYS Campus	HTTPS	2	portal.hpc.uams.edu	Open
Commodity Internet	SSH	1	portal.hpc.uams.edu	**Closed**
Commodity Internet	HTTPS	2	portal.hpc.uams.edu	Open

Pinnacle – Demilitarized Zone (DMZ) access

hpc-portal2.hpc.uark.edu - login node for both SSH and HTTPS (OOD portal) accessible from anywhere on the internet. Access is granted using a 2-step process:

1. register your source IP and add it to the list of hosts allowed to access the Pinnacle SSH login nodes and Pinnacle OOD portal. Ssh into port 2022 of hpc-portal2.hpc.uark.edu.

```
ssh <username>@hpc-portal2.hpc.uark.edu -p 2022
```

```
[pawel@frontdesk ~]$ ssh testuser@hpc-portal2.hpc.uark.edu -p 2022
Password:
Last login: Tue Dec  7 16:58:10 2021 from 75.205.5.185
               Arkansas High Performance Computing Center
                        off-campus access portal
================================================================================
   https://hpc-portal2.hpc.uark.edu    <-  Open Ondemand Web Portal
   ssh hpc-portal2.hpc.uark.edu        <-  Standard AHPCC login

   Access allowed from 75.205.5.185 until Wed Dec  8 06:02:43 CST 2021
================================================================================
Connection to hpc-portal2.hpc.uark.edu closed.
[pawel@frontdesk ~]$
```

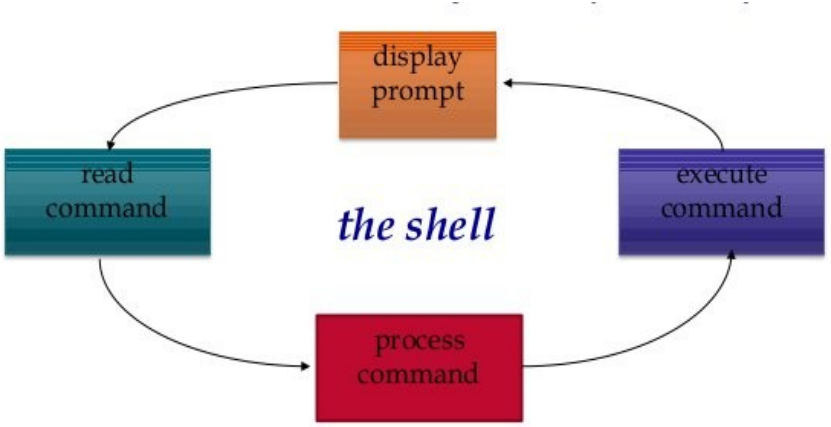
2. Log in as normal (without specifying port 2022):

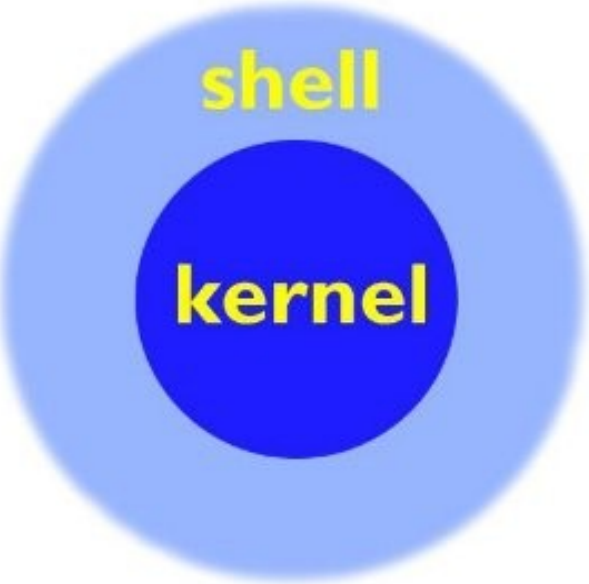
```
ssh <username>@hpc-portal2.hpc.uark.edu
https://hpc-portal2.hpc.uark.edu
```

Bash Shell

Kernel – core part of the operating system. Contains all the drivers necessary to communicate with the underlying hardware. Manages the CPUs, memory, I/O operations form all peripheral devices (disk, keyboard, mouse, printers, scanners etc.)

Shell – provides the Command Line Interface (CLI) for interaction with the kernel and maintains the environment, a set of variables. Bash is the most commonly used shell on Linux.





Kernel receives commands from the shell and performs them. There may be multiple shells interacting with the kernel at one time (multiple users logged in at once, each with mulitple shells).

Bash Shell - Environment

env – shows all of the environmental variables that are defined in your shell

- SHELL – the shell you are running
- HOSTNAME - name of the host
- PWD - present working directory
- USER - user name
- HOME - path of the user's home direcotry
- PATH – list of directories to search when user types in a command
- PS1 – configures prompt

. . .

[pwolinsk@login1 ~]

[username@host working_directory]

(~ is a shortcut notation for user's HOME directory)

```
[pwolinsk@login1 ~]$ echo $PS1
[\u@\h \W]\$
[pwolinsk@login1 ~]$ echo $SHELL
/bin/bash
[pwolinsk@login1 ~]$
[pwolinsk@login1 ~]$ export PS1="[hello \u@\h \W]\$"
[hello pwolinsk@login1 ~]$
```

Bash Shell - First Linux Commands

echo - display text string or contents of an environmental variable

who – list users currently logged on

whoami

hostname

date

pwd – show present working directory

man - show a manual page (help page) about any command *(in man type "/" followed by a search string, "n" to find next occurence)*

apropos - show any manual pages which reference a given string

history - show listing of your commands

nano – basic text editor

vi – powerful text editor

```
[pwolinsk@login1 ~]$ echo $PS1
[\u@\h \W]\$
[pwolinsk@login1 ~]$ echo $SHELL
/bin/bash
[pwolinsk@login1 ~]$
[pwolinsk@login1 ~]$ export PS1="[hello \u@\h \W]\$"
[hello pwolinsk@login1 ~]$
```

File Systems – Pinnacle & Grace

df - disk free **mount –** show mounted files systems and mount points

```
pinnacle-l1:pwolinsk:~$ df -h
Filesystem                                                                                  Size  Used Avail Use% Mounted on
devtmpfs                                                                                    7.8G     0  7.8G   0% /dev
tmpfs                                                                                       7.8G   24K  7.8G   1% /dev/shm
tmpfs                                                                                       7.8G  153M  7.7G   2% /run
/dev/mapper/centos-root                                                                     8.0G  6.7G  1.4G  84% /
/dev/vda1                                                                                  1014M  173M  842M  18% /boot
172.17.27.1@o2ib,172.17.27.21@o2ib:172.17.27.2@o2ib,172.17.27.22@o2ib:/scrfs  2.3P  1.2P  1.2P  50% /scrfs
172.17.27.13@o2ib:/scr1                                                                      11T  464G  9.7T   5% /scr1
172.17.27.14@o2ib:/scr2                                                                      11T  964G  9.3T  10% /scr2
tmpfs                                                                                      1.6G     0  1.6G   0%
/run/user/269030
pinnacle-l1:pwolinsk:~$
```

```
[pwolinsk@xbt002 ~]$ df -h
Filesystem                  Size  Used Avail Use% Mounted on
/dev/mapper/sysvg-root       88G   30G   59G  34% /
devtmpfs                     63G     0   63G   0% /dev
tmpfs                        63G     0   63G   0% /dev/shm
tmpfs                        63G  490M   63G   1% /run
tmpfs                        63G     0   63G   0% /sys/fs/cgroup
/dev/mapper/uservg-scratch  465G   33M  465G   1% /local-scratch
/dev/sda2                   997M  168M  829M  17% /boot
/dev/sda3                   500M   16K  500M   1% /boot/efi
uams-gs                     1.9P  1.4P  468T  75% /home
tmpfs                        13G     0   13G   0% /run/user/0
tmpfs                        13G     0   13G   0% /run/user/1692600016
[pwolinsk@xbt002 ~]$
```

File Systems – Pinnacle & Grace

ls – list files **ls -l** – long listing **ls -al** - show all files (including hidden files, starting with ".")

```
pinnacle-l1:guest01:~$ ls -l

drwxr-xr-x 2 guest01 guest01 32768 Aug 13  2009 mpiexamples
-rw-r--r-- 1 guest01 guest01     0 Mar 17 13:17 newfile
drwxrwxr-x 2 guest01 guest01 32768 Mar 17 14:55 tt
- ---------    ------- ------- ----- -----------  -----
|      |             |     |      |          |         |
|      |             |     |      |          |         |
| user/group/others user group  size modification  filename
|     permissions                 in       date
|                                bytes
file/directory/link
 r read   w write  x execute
```

```
pinnacle-l1:guest01:~$ ls -al
drwx------  6 guest01 guest01 32768 Mar 17 14:55 .
drwxr-xr-x 24 root    root        0 Mar 17 15:59 ..
-rw-------  1 guest01 guest01  4604 Mar 17 13:28 .bash_history
-rw-r--r--  1 guest01 guest01    24 Aug  5  2009 .bash_logout
-rwxr-xr-x  1 guest01 guest01   132 Mar 17 13:20 .bash_profile
lrwxrwxrwx  1 guest01 guest01    13 Jan 27  2011 .bashrc -> .bash_profile
...
drwxr-xr-x  2 guest01 guest01 32768 Aug 13  2009 mpiexamples
-rwx------  1 guest01 guest01     0 Mar 17 13:17 newfile
drwxrwxr-x  2 guest01 guest01 32768 Mar 17 14:55 tt
...
```

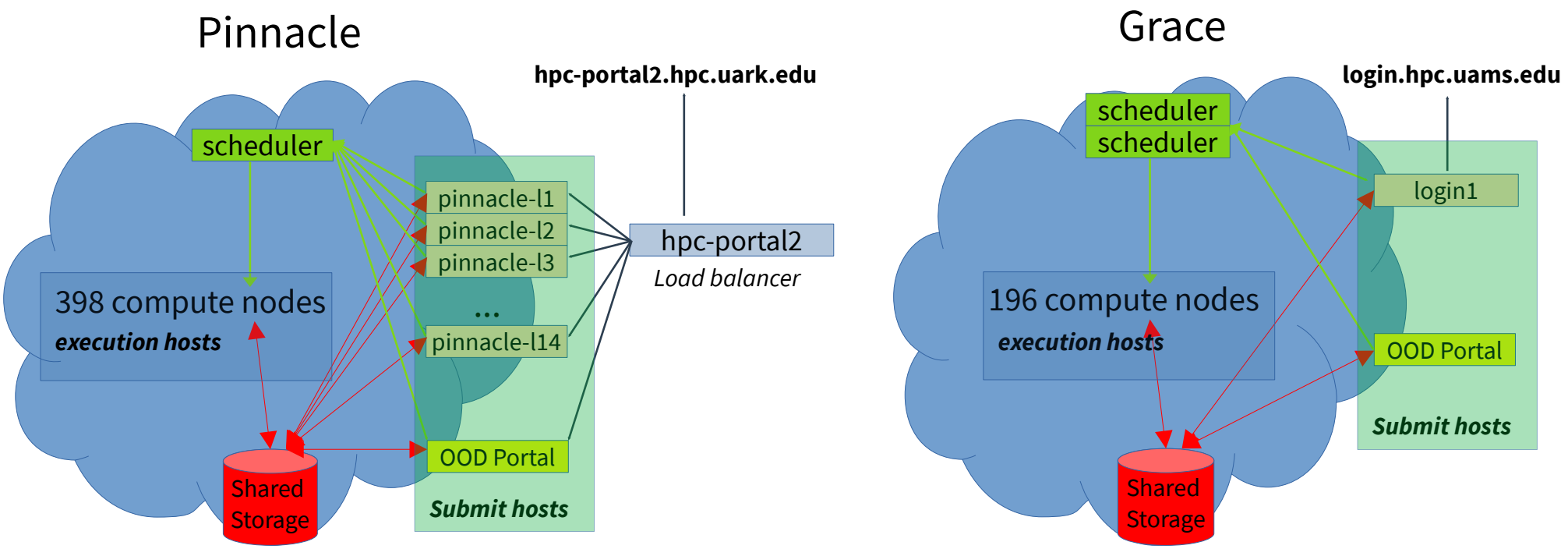
File Systems – /home and /scratch

***/home*/$USER** – file storage for each user **/scratch/<jobid>** - temporary directory for executing jobs

At the beginning of a job copy your input files to /scratch/<jobid> and at the end of job copy output files to back to /home/$USER

```
c1601:pwolinsk:~$ df
Filesystem                                                                      1K-blocks           Used      Available Use% Mounted on
devtmpfs                                                                         98129128              0       98129128   0% /dev
tmpfs                                                                            98146236             64       98146172   1% /dev/shm
tmpfs                                                                            98146236          52388       98093848   1% /run
tmpfs                                                                            98146236              0       98146236   0%
/sys/fs/cgroup
/dev/mapper/xcatvg-root                                                         229595616       16333344      213262272   8% /
/dev/sda1                                                                          517868         226032         291836  44% /boot
172.17.27.13@o2ib:/scr1                                                      11468687248      486872436    10403223304   5% /scr1
172.17.27.14@o2ib:/scr2                                                      11471001300     1011719572     9880567544  10% /scr2
172.17.27.1@o2ib,172.17.27.21@o2ib:172.17.27.2@o2ib,172.17.27.22@o2ib:/scrfs 2455426280448 1214971101936 1215600540940  50% /scrfs
tmpfs                                                                            19629248              0       19629248   0%
/run/user/1006
tmpfs                                                                            19629248              0       19629248   0%
/run/user/269030
c1601:pwolinsk:~$ pwd
/home/pwolinsk
c1601:pwolinsk:~$ ls -l /scratch |head
total 0
lrwxrwxrwx 1 root root 13 Feb 16  2022 1003741 -> /scr1/1003741
lrwxrwxrwx 1 root root 13 Feb 16  2022 1010024 -> /scr2/1010024
lrwxrwxrwx 1 root root 13 Feb 16  2022 1010044 -> /scr2/1010044
lrwxrwxrwx 1 root root 13 Feb 17  2022 1010047 -> /scr1/1010047
lrwxrwxrwx 1 root root 13 Feb 17  2022 1012608 -> /scr2/1012608
lrwxrwxrwx 1 root root 13 Feb 17  2022 1013345 -> /scr1/1013345
lrwxrwxrwx 1 root root 13 Feb 18  2022 1020146 -> /scr2/1020146
lrwxrwxrwx 1 root root 13 Feb 21  2022 1025910 -> /scr2/1025910
lrwxrwxrwx 1 root root 13 Feb 21  2022 1025911 -> /scr1/1025911
c1601:pwolinsk:~$
```

ARP Clusters – Queueing System



scheduler – coordinates all of the nodes in the queue (SLURM queueing system)
submit hosts – nodes from which users submit jobs to the queues
login hosts – nodes accessible from the external network
compute nodes – nodes that assigned to queues and run jobs

Queueing System

sinfo – show node/queue assigment **squeue** - jobs in the queues

Compute nodes are divided into queues (or partitions).

```
pinnacle-l1:pwolinsk:~$ sinfo
PARTITION  AVAIL  TIMELIMIT  NODES  STATE NODELIST
comp01        up    1:00:00      3  drain c[1402-1404]
comp01        up    1:00:00     42  alloc c[1407-1420,1501-1520,1603-1610]
comp01        up    1:00:00      2   idle c[1405-1406]
comp01        up    1:00:00      1   down c1401
comp06        up    6:00:00      1  drain c1404
comp06        up    6:00:00     42  alloc c[1407-1420,1501-1520,1603-1610]
comp06        up    6:00:00      2   idle c[1405-1406]
comp72        up 3-00:00:00     42  alloc c[1407-1420,1501-1520,1603-1610]
gpu06         up    6:00:00      2  alloc c[1713,1715]
gpu06         up    6:00:00     17   idle c[1612-1615,1701-1712,1714]
gpu72         up 3-00:00:00      2  alloc c[1713,1715]
gpu72         up 3-00:00:00     17   idle c[1612-1615,1701-1712,1714]
condo         up   infinite      4  alloc c[1307,1310-1312]
condo         up   infinite     21   idle c[1308-1309,1313-1328,1611,1719-1720]
pubcondo06    up    6:00:00      4  alloc c[1307,1310-1312]
pubcondo06    up    6:00:00     21   idle c[1308-1309,1313-1328,1611,1719-1720]
pcon06*       up    6:00:00      4  alloc c[1307,1310-1312]
pcon06*       up    6:00:00     21   idle c[1308-1309,1313-1328,1611,1719-1720]
cloud72       up 3-00:00:00      2   idle c[1601-1602]
cloud72       up 3-00:00:00      1   down c1329
himem06       up    6:00:00      6   idle c[1301-1306]
himem72       up 3-00:00:00      6   idle c[1301-1306]
pinnacle-l1:pwolinsk:~$
```

Queueing System

sinfo – show node/queue assigment ***squeue*** - jobs in the queue

```
pinnacle-l1:pwolinsk:~$ squeue
             JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
              3633     condo OG000050      aja PD       0:00      1 (Resources)
              3634     condo OG000050      aja PD       0:00      1 (Priority)
              3635     condo OG000050      aja PD       0:00      1 (Priority)
              3636     condo OG000050      aja PD       0:00      1 (Priority)
              3637     condo OG000049      aja PD       0:00      1 (Priority)
              3638     condo OG000050      aja PD       0:00      1 (Priority)
              3639     condo OG000050      aja PD       0:00      1 (Priority)
              3640     condo OG000050      aja PD       0:00      1 (Priority)
              3641     condo OG000049      aja PD       0:00      1 (Priority)
              3642     condo OG000049      aja PD       0:00      1 (Priority)
              3643     condo OG000049      aja PD       0:00      1 (Priority)
              3644     condo OG000049      aja PD       0:00      1 (Priority)
              3645     condo OG000048      aja PD       0:00      1 (Priority)
              3646     condo OG000049      aja PD       0:00      1 (Priority)
              3650     condo OG000048      aja PD       0:00      1 (Priority)
              3651     condo OG000048      aja PD       0:00      1 (Priority)
              ...
              6892    comp72 coalHMM_ tkchafin PD       0:00      1 (Priority)
              3632     condo OG000051      aja  R    8:19:34      1 c1307
              6551    comp72 Pmc21-MD  kypatel  R    6:09:20      8 c[1407,1409,1504,1506,1513,1605-1606,1608]
              6749    comp72 coalHMM_ tkchafin  R 2-10:42:46      1 c1411
              6761    comp72 coalHMM_ tkchafin  R 2-09:00:12      1 c1417
              6775    comp72 coalHMM_ tkchafin  R 2-01:20:12      1 c1415
              6776    comp72 coalHMM_ tkchafin  R 2-00:41:24      1 c1503
              6777    comp72 coalHMM_ tkchafin  R 1-19:01:08      1 c1510
              6779    comp72 coalHMM_ tkchafin  R 1-14:22:11      1 c1603
              6780    comp72 coalHMM_ tkchafin  R 1-13:25:36      1 c1416
              7147     gpu72 BO_bi_60 rpsander  R 1-21:56:14      1 c1715
              7299     gpu72 BO_mo_80 rpsander  R   16:11:37      1 c1713
              ...
              7311     condo fit_inte yingyuan  R    3:16:02      1 c1310
              7313     condo fit_inte yingyuan  R    3:08:58      1 c1311
              7314     condo      afm yingyuan  R    2:03:28      1 c1312
pinnacle-l1:pwolinsk:~$
```

Queueing System – First job - command line

sbatch – submit a job to the queue ***squeue*** - jobs in the queue

Copy a pre-made script into your home directory, examine it and run it on the login node.

```
pinnacle-l6:pwolinsk:~$ mkdir EXAMPLES
pinnacle-l6:pwolinsk:~$ cd EXAMPLES/
pinnacle-l6:pwolinsk:~/EXAMPLES$ cp /scrfs/ahpcc/ARP-workshop/first.sh .
pinnacle-l6:pwolinsk:~/EXAMPLES$ cat first.sh
#!/bin/bash

echo "Script is executing on: "
hostname
echo "My present working directory is $PWD"
echo "Time and date"
date
pinnacle-l6:pwolinsk:~/EXAMPLES$
pinnacle-l6:pwolinsk:~/EXAMPLES$ ./first.sh
Script is executing on:
pinnacle-l6
My present working directory is /home/pwolinsk/EXAMPLES
Time and date
Tue Sep 27 19:04:21 CDT 2022
pinnacle-l6:pwolinsk:~/EXAMPLES$
```

Queueing System – First interactive job - command line

sbatch – submit a batch job to the queue ***srun*** – submit an interactive job to the queue ***squeue*** - jobs in the queue

Submit the same script to the queueing system interactively – with **srun**.

```
pinnacle-l6:pwolinsk:~/EXAMPLES$ srun -p comp01 --pty /bin/bash
c1405:pwolinsk:~/EXAMPLES$ ./first.sh
Script is executing on:
c1405
My present working directory is /home/pwolinsk/EXAMPLES
Time and date
Tue Sep 27 19:09:15 CDT 2022
c1405:pwolinsk:~/EXAMPLES$ squeue -u pwolinsk
             JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
           1315893    comp01     bash pwolinsk  R       0:17      1 c1405
c1405:pwolinsk:~/EXAMPLES$ exit
exit
pinnacle-l6:pwolinsk:~/EXAMPLES$
```

Queueing System – First batch job - command line

sbatch – submit a batch job to the queue ***srun*** – submit an interactive job to the queue ***squeue*** - jobs in the queue

Submit the same script to the queueing system as a batch job – with **sbatch**.

```
pinnacle-l6:pwolinsk:~/EXAMPLES$ sbatch -p comp01 first.sh
Submitted batch job 1315894
pinnacle-l6:pwolinsk:~/EXAMPLES$ squeue -u pwolinsk
             JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
pinnacle-l6:pwolinsk:~/EXAMPLES$ ls -ltr
total 8
-rwxr-xr-x 1 pwolinsk pwolinsk 124 Sep 27 19:03 first.sh
-rw-rw-r-- 1 pwolinsk pwolinsk 130 Sep 27 19:13 slurm-1315894.out
pinnacle-l6:pwolinsk:~/EXAMPLES$ cat slurm-1315894.out
Script is executing on:
c1405
My present working directory is /home/pwolinsk/EXAMPLES
Time and date
Tue Sep 27 19:13:02 CDT 2022
pinnacle-l6:pwolinsk:~/EXAMPLES$
```

Queueing System – Second batch job – job array

sbatch – submit a batch job to the queue ***srun*** – submit an interactive job to the queue ***squeue*** - jobs in the queue

sbatch command with the **--array** parameter will run the same job script a specified amount of times with $SLURM_JOB_ARRAY_ID variable storing a unique array element number.

```
pinnacle-l6:pwolinsk:~/EXAMPLES$ cp /scrfs/ahpcc/ARP-workshop/second.sh .
pinnacle-l6:pwolinsk:~/EXAMPLES$ cat second.sh
#!/bin/bash

echo "My array task id is: $SLURM_ARRAY_TASK_ID"
pinnacle-l6:pwolinsk:~/EXAMPLES$ sbatch -p comp01 --array [1-5] second.sh
Submitted batch job 1315912
pinnacle-l6:pwolinsk:~/EXAMPLES$ ls -ltr
total 188
-rwxr-xr-x 1 pwolinsk pwolinsk    124 Sep 27 19:03 first.sh
-rwxrwxr-x 1 pwolinsk pwolinsk     62 Sep 27 19:32 second.sh
-rw-rw-r-- 1 pwolinsk pwolinsk     23 Sep 27 19:33 slurm-1315912_5.out
-rw-rw-r-- 1 pwolinsk pwolinsk     23 Sep 27 19:33 slurm-1315912_4.out
-rw-rw-r-- 1 pwolinsk pwolinsk     23 Sep 27 19:33 slurm-1315912_3.out
-rw-rw-r-- 1 pwolinsk pwolinsk     23 Sep 27 19:33 slurm-1315912_2.out
-rw-rw-r-- 1 pwolinsk pwolinsk     23 Sep 27 19:33 slurm-1315912_1.out
pinnacle-l6:pwolinsk:~/EXAMPLES$ cat slurm-1315912_5.out
My array task id is: 5
pinnacle-l6:pwolinsk:~/EXAMPLES$ cat slurm-1315912_4.out
My array task id is: 4
pinnacle-l6:pwolinsk:~/EXAMPLES$
```

Queueing System – Third batch job – job array

sbatch – submit a batch job to the queue ***srun*** – submit an interactive job to the queue ***squeue*** - jobs in the queue

sbatch command with the **--array** parameter will run the same job script a specified amount of times with $SLURM_JOB_ARRAY_ID variable storing a unique array element number.

```
pinnacle-l6:pwolinsk:~/EXAMPLES$ cp /scrfs/ahpcc/ARP-workshop/1fqy.pdb .
pinnacle-l6:pwolinsk:~/EXAMPLES$ head 1fqy.pdb
HEADER    MEMBRANE PROTEIN                        07-SEP-00   1FQY
TITLE     STRUCTURE OF AQUAPORIN-1 AT 3.8 A RESOLUTION BY ELECTRON
TITLE    2 CRYSTALLOGRAPHY
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: AQUAPORIN-1;
COMPND   3 CHAIN: A;
COMPND   4 SYNONYM: AQP1
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: HOMO SAPIENS;
SOURCE   3 ORGANISM_COMMON: HUMAN;
pinnacle-l6:pwolinsk:~/EXAMPLES$ cp /scrfs/ahpcc/ARP-workshop/third.sh .
pinnacle-l6:pwolinsk:~/EXAMPLES$ cat third.sh
#!/bin/bash

head -$SLURM_ARRAY_TASK_ID 1fqy.pdb |tail -1
pinnacle-l6:pwolinsk:~/EXAMPLES$ sbatch -p comp01 --array [100,102] third.sh
Submitted batch job 1315918
pinnacle-l6:pwolinsk:~/EXAMPLES$ ls -ltr
total 176
-rwxr-xr-x 1 pwolinsk pwolinsk    124 Sep 27 19:03 first.sh
-rwxrwxr-x 1 pwolinsk pwolinsk     62 Sep 27 19:32 second.sh
-rw-r--r-- 1 pwolinsk pwolinsk 158112 Sep 27 19:38 1fqy.pdb
-rwxrwxr-x 1 pwolinsk pwolinsk     58 Sep 27 19:38 third.sh
-rw-rw-r-- 1 pwolinsk pwolinsk     81 Sep 27 19:39 slurm-1315918_102.out
-rw-rw-r-- 1 pwolinsk pwolinsk     81 Sep 27 19:39 slurm-1315918_100.out
pinnacle-l6:pwolinsk:~/EXAMPLES$ cat slurm-1315918_102.out
REMARK   3
pinnacle-l6:pwolinsk:~/EXAMPLES$ cat slurm-1315918_100.out
REMARK   3   SIDE-CHAIN BOND               (A**2) : NULL   ; NULL
pinnacle-l6:pwolinsk:~/EXAMPLES$
```

Software Modules

Hunders of applications, compilers and scripting languages, some with multiple versions, are installed on Grace and Pinnacle. They cannot all be available to the users at once because they interfere with each other.

The Environment Modules system is a tool to help users manage the shell environment and provide access to specific applications and versions on demand. It does this by manipulating the values of shell environmental variables. The most important of those are **$PATH** and **$LD_LIBRARY_PATH** .

```
pinnacle-l6:pwolinsk:~$ ls /share/apps/
abinit          cplex                ganglia     idl              mathematics          OpenFOAM                qt              synopsys
accumulo        CPMD                 gappa       iftop            Matjes               openmpi                 Qt              szip
acml            crabz                Gaussian    iftop-0.17       matlab               opt                     QTXNetwork      TauDEM
alamode         cuda                 gcc         igb-5.3.5.4      maui                 oracle_developer_studio R               tcl
amber           curl                 gd          imp              mauiversions         orca                    rclone          texinfo
amber16tools17  dar                  gdal        intel            maven                orca_3_0_3_linux_x86-64 relion          thirdorder
amdappsdk       ddscat               gdrcopy     intel18          meep                 osu-micro-bench         rh6             torque
ansys           dealii               geomview    intel2020-2      MKDAT                ovito                   RITHM           torqueversions
...
pinnacle-l6:pwolinsk:~$ ls /share/apps/bioinformatics/
abyss              bismark            DataAnalysis   GCTA          last          MUMmer          picard          sate-tools-linux  swarm
admixture          blas               datamash       Gctf          lib           muscle          pilon           sepp              SweeD
aegean             blasr              dDocent        gemma         libBigWig     mutect          pindel          sepp-tools        swig
agalma             blast              deepTools      GeneMark      lordec        MutMap          pitchfork       seqtk             tabix
allpaths-lg        blat               diamond        genemark_es   lorma         ncbi            plink           shapeit           taiyaki
angsd              BLINK              dock           genometools   lsc           necat           PopLDdecay      sickle            tassel
annovar            blobtools          edirect        gtdbtk        MACSE         newbler         popoolation2    simple            TopHat
astalavista        bonito             eigen          GTDBTK        mafft         newick-utils    pplacer         smrtlink          Tracer
...
```

Software Modules

module avail - list all available modules
module load – load a module
module list – list modules currently loaded
module remove – remove a loaded module

```
pinnacle-l6:pwolinsk:~$ module avail

------------------------------------------------------------ /share/apps/modulefiles ------------------------------------------------------------
   BayesAss3-SNPs/1.0                edirect/20200815             (D)    java/ibmjdk_1.8.0                pqs/4.0-4
   BerkeleyGW/2.1                    elmer/1.0                           java/openjdk_1.7.0               pqs/4.0-8
   BerkeleyGW/3.0.1           (D)    elpa/2021.11.002                    java/openjdk_1.8.0_20            pqs/4.1-0
   CAFE/4.2.1                        emboss/6.6.0                        java/openjdk_1.8.0               pqs/4.1-2-env-modules
   CERN/6.24                         epa-ng/0.3.6                        java/openjdk_11.0.10             pqs/4.1-2                  (D)
   EnTAP/0.8.1                       exabayes/1.5.1                      java/openjdk_14.0.1              prisms-plasticity/1.3.0
   LIGGGHTS/1                        exonerate/2.4.0                     java/openjdk_16.0.0              prodigal/2.6.3
   METABOLIC/4.0                     express/1.5.1                       java/sunjdk_1.7.0_80             proj/5.0.1
   MUMmer/3.23                       fastANI/1.32                        java/sunjdk_1.7.0                proj/8.1.0                 (D)
   MaSuRCA/3.4.2                     fastepistasis/2.07                  java/sunjdk_1.8.0_72             pvm/3.4.6
   MetaGeneMark/3.38                 fastqc/0.11.5                       java/sunjdk_1.8.0_92             python/anaconda-2.7.13
   PAML/4.6                          fasttree/2.1.10                     java/sunjdk_1.8.0          (D)   python/anaconda-2.7.15
   PAML/4.8                          fastx/0.0.14                        java/sunjdk_11.0.10              python/anaconda-2.7.16
   PAML/4.9e                  (D)    fftw/2.1.5                   (D)    jellyfish/2.2.6                  python/anaconda-3.6.0
   PASTA/1.8.3                       fftw/3.2.2                          jellyfish/2.3.0            (D)   python/anaconda-3.7.3
   PASTA/1.8.6                (D)    fftw/3.3.4-double-openmp-old        julia/1.7.0-rc1                  python/anaconda-3.8
   PBSuite/15.8.24                   fftw/3.3.4-double-openmp            kallisto/0.43.1                  python/anaconda-3.9
   PGI/18.4                          fftw/3.3.4-double                   ktrim/1.1.0                      python/anaconda-3.10
   PGI/18.10                         fftw/3.3.4-openmpi                  last/719                         python/intel-2.7.16
   PGI/19.4                          fftw/3.3.4                          libxc/5.2.3                      python/intel-3.6.8
   PGI/2016                          fftw/3.3.6                          libxsmm/1.17.0                   python/intel-3.7.7
   PGI/2016.5                        fftw/3.3.8                          mafft/7.304b                     python/rh27
   PGI/2017                          flash/1.2.11                        mafft/7.490                (D)   python/2.7-rh
   PGI/2019                   (D)    fluent/2020R2                       maker/2.31.8                     python/2.7.3
   PopLDdecay/3.4.1                  fluent/2022R1                (D)    maker/2.31.9               (D)   python/2.7.5
   R/MR3.3.1                         flye/2.3.3                          mathematica/12.3                 python/2.7.9
   R/MR3.4.2                         flye/2.6                            mathematica/13.0                 python/2.7.11
   R/MR4.0.2                         flye/2.8.3                   (D)    mathematica/13.1           (D)   python/2.7.13-anaconda
   R/3.3.2                           gadget/2.0.7-SF                     matlab/r2017a                    python/2.7.15-anaconda
   R/3.4.3-deprecated                gadget/2.03-SF               (D)    matlab/r2019a                    python/2.7.15b
...
```

Software Modules

module avail - list all available modules
module list – list modules currently loaded
module load – load a module
module remove – remove a loaded module

```
pinnacle-l6:pwolinsk:~$ which python
/usr/bin/python
pinnacle-l6:pwolinsk:~$ python
Python 2.7.5 (default, Nov 16 2020, 22:23:17)
[GCC 4.8.5 20150623 (Red Hat 4.8.5-44)] on linux2
Type "help", "copyright", "credits" or "license" for more information.
>>> quit()
pinnacle-l6:pwolinsk:~$ module avail python
------------------------------------------------------------ /share/apps/modulefiles -------------------------------------------------------------
   python/anaconda-2.7.13    python/anaconda-3.10    python/2.7.5              python/2.7.16-anaconda    python/3.7.3-anaconda
   python/anaconda-2.7.15    python/intel-2.7.16     python/2.7.9              python/2.7.16-intel       python/3.7.7-intel
   python/anaconda-2.7.16    python/intel-3.6.8      python/2.7.11             python/2.7.18-anaconda    python/3.8-anaconda
   python/anaconda-3.6.0     python/intel-3.7.7      python/2.7.13-anaconda    python/3.5.1              python/3.9-anaconda
   python/anaconda-3.7.3     python/rh27             python/2.7.15-anaconda    python/3.6.0-anaconda     python/3.10-anaconda  (D)
   python/anaconda-3.8       python/2.7-rh           python/2.7.15b            python/3.6.0
   python/anaconda-3.9       python/2.7.3            python/2.7.15             python/3.6.8-intel

  Where:
   D:  Default Module
pinnacle-l6:pwolinsk:~$ module load python/3.10-anaconda
pinnacle-l6:pwolinsk:~$ module list
Currently Loaded Modules:
  1) os/el7   2) python/3.10-anaconda
pinnacle-l6:pwolinsk:~$ which python
/share/apps/python/anaconda-3.10/bin/python
pinnacle-l6:pwolinsk:~$ python
Python 3.10.4 | packaged by conda-forge | (main, Mar 24 2022, 17:38:57) [GCC 10.3.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> quit()
pinnacle-l6:pwolinsk:~$ module remove python
pinnacle-l6:pwolinsk:~$ which python
/usr/bin/python
pinnacle-l6:pwolinsk:~$
```

Conda (Python Modules)

Package, dependency and environment management for Python (and other languages)
On Pinnacle, load the Python 3.10 module and activate the base environment.

```
pinnacle-l5:pwolinsk:~$ module load python/3.10-anaconda
pinnacle-l5:pwolinsk:~$ module list

Currently Loaded Modules:
  1) os/el7   2) python/3.10-anaconda


pinnacle-l5:pwolinsk:~$ which python
/share/apps/python/anaconda-3.10/bin/python
pinnacle-l5:pwolinsk:~$ source /share/apps/bin/conda-3.10.sh
(base) pinnacle-l5:pwolinsk:~$ conda env list
# conda environments:
#
BoltzTraP-env            /home/pwolinsk/.conda/envs/BoltzTraP-env
bwa-env                  /home/pwolinsk/.conda/envs/bwa-env
circlator                /home/pwolinsk/.conda/envs/circlator
delly-env                /home/pwolinsk/.conda/envs/delly-env
...
htseq-3.10               /share/apps/python/anaconda-3.10/envs/htseq-3.10
mpi4py-mvapich-3.10      /share/apps/python/anaconda-3.10/envs/mpi4py-mvapich-3.10
mpi4py-openmpi-3.10      /share/apps/python/anaconda-3.10/envs/mpi4py-openmpi-3.10
phonopy-3.10             /share/apps/python/anaconda-3.10/envs/phonopy-3.10
pymor-3.10               /share/apps/python/anaconda-3.10/envs/pymor-3.10
qiime2-2022.2-3.10       /share/apps/python/anaconda-3.10/envs/qiime2-2022.2-3.10

(base) pinnacle-l5:pwolinsk:~$
```

Conda (Python Modules) – existing environments

Package, dependency and environment management for Python (and other languages)

conda env list – list installed environments **conda activate** <*env_name*> - activate env **conda deactivate** - deactivate

```
(base) pinnacle-l5:pwolinsk:~$ conda env list
bwa-env                  /home/pwolinsk/.conda/envs/bwa-env
circlator                /home/pwolinsk/.conda/envs/circlator
delly-env                /home/pwolinsk/.conda/envs/delly-env
...
htseq-3.10               /share/apps/python/anaconda-3.10/envs/htseq-3.10
mpi4py-mvapich-3.10      /share/apps/python/anaconda-3.10/envs/mpi4py-mvapich-3.10
mpi4py-openmpi-3.10      /share/apps/python/anaconda-3.10/envs/mpi4py-openmpi-3.10
...
(base) pinnacle-l5:pwolinsk:~$ which htseq-qa
/usr/bin/which: no htseq-qa in (/share/apps/python/anaconda-3.10/bin:/share/apps/python/anaconda-3.10/condabin:/share/
apps/python/anaconda-3.10/condabin:/share/apps/python/anaconda-3.10/bin:/usr/lib64/qt-3.3/bin:/usr/local/bin:/usr/bin:/
usr/local/sbin:/usr/sbin:/opt/ibutils/bin:/scrfs/ahpcc/libvirt/bin:/share/apps/bin:/home/pwolinsk/bin:/home/pwolinsk/
Leyh/t)
(base) pinnacle-l5:pwolinsk:~$ conda activate htseq-3.10
(htseq-3.10) pinnacle-l5:pwolinsk:~$ which htseq-qa
/share/apps/python/anaconda-3.10/envs/htseq-3.10/bin/htseq-qa
(htseq-3.10) pinnacle-l5:pwolinsk:~$  ls /share/apps/python/anaconda-3.10/envs/htseq-3.10
bin  compiler_compat  conda-meta  etc  include  lib  man  sbin  share  ssl  x86_64-conda_cos7-linux-gnu  x86_64-conda-
linux-gnu
(htseq-3.10) pinnacle-l5:pwolinsk:~$ conda deactivate
(base) pinnacle-l5:pwolinsk:~$ which htseq-qa
/usr/bin/which: no htseq-qa in (/share/apps/python/anaconda-3.10/bin:/share/apps/python/anaconda-3.10/condabin:/share/
apps/python/anaconda-3.10/condabin:/share/apps/python/anaconda-3.10/bin:/usr/lib64/qt-3.3/bin:/usr/local/bin:/usr/bin:/
usr/local/sbin:/usr/sbin:/opt/ibutils/bin:/scrfs/ahpcc/libvirt/bin:/share/apps/bin:/home/pwolinsk/bin:/home/pwolinsk/
Leyh/t)
(base) pinnacle-l5:pwolinsk:~$
```

Conda (Python Modules) – install your own package

Package, dependency and environment management for Python (and other languages)

conda config –add channels <*channel*> – add channel **conda search <*package*>**

```
(base) pinnacle-l5:pwolinsk:~$ conda config --add channels defaults
(base) pinnacle-l5:pwolinsk:~$ conda config --add channels bioconda
(base) pinnacle-l5:pwolinsk:~$ conda config --add channels conda-forge
(base) pinnacle-l5:pwolinsk:~$ cat .condarc
channels:
  - conda-forge
  - bioconda
  - defaults
(base) pinnacle-l5:pwolinsk:~$ conda search tensorflow
Loading channels: done
# Name                       Version           Build  Channel
tensorflow                     0.7.1          py27_0  conda-forge
tensorflow                     0.7.1          py34_0  conda-forge
...
tensorflow                     2.9.1 mkl_py37h58a621a_0  pkgs/main
tensorflow                     2.9.1 mkl_py38h96f9fba_0  pkgs/main
...
tensorflow                    2.10.0 cpu_py39h4655687_0  conda-forge
tensorflow                    2.10.0 cuda112py310he87a039_0  conda-forge
tensorflow                    2.10.0 cuda112py37h01c6645_0  conda-forge
tensorflow                    2.10.0 cuda112py38hded6998_0  conda-forge
tensorflow                    2.10.0 cuda112py39h01bd6f0_0  conda-forge
...
```

Conda (Python Modules) – install your own packages

Package, dependency and environment management for Python (and other languages)

conda install -c *<channel>* package=*<version>*=*<build>*

```
(base) pinnacle-l5:pwolinsk:~$ conda install -c conda-forge tensorflow=2.10.0=cpu_py310hd1aba9c_0
Collecting package metadata (current_repodata.json): done
Solving environment: -
The environment is inconsistent, please check the package plan carefully
The following packages are causing the inconsistency:

  - conda-forge/noarch::ipywidgets==7.7.1=pyhd8ed1ab_0
  - conda-forge/noarch::jupyterlab==3.4.4=pyhd8ed1ab_0
  - conda-forge/noarch::notebook-shim==0.1.0=pyhd8ed1ab_0
  - conda-forge/noarch::jupyterlab_server==2.15.0=pyhd8ed1ab_0
  - conda-forge/noarch::nbclassic==0.4.3=pyhd8ed1ab_0
  - conda-forge/noarch::nbconvert-core==6.5.0=pyhd8ed1ab_0
  - conda-forge/noarch::notebook==6.4.12=pyha770c72_0
  - conda-forge/noarch::nbconvert==6.5.0=pyhd8ed1ab_0
  - conda-forge/noarch::nbconvert-pandoc==6.5.0=pyhd8ed1ab_0
  - conda-forge/noarch::jupyter_server==1.18.1=pyhd8ed1ab_0
  - conda-forge/noarch::widgetsnbextension==3.6.1=pyha770c72_0
done

## Package Plan ##

  environment location: /share/apps/python/anaconda-3.10

  added / updated specs:
    - tensorflow==2.10.0=cpu_py310hd1aba9c_0
...
```

Users can't install into base environment

Conda (Python Modules) – create a private environment

Package, dependency and environment management for Python (and other languages)

conda create -n <*environment*>

```
(base) pinnacle-l5:pwolinsk:~$ conda create -n tensorflow-pawel
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: /home/pwolinsk/.conda/envs/tensorflow-pawel

Proceed ([y]/n)? y

Preparing transaction: done
Verifying transaction: done
Executing transaction: done
#
# To activate this environment, use
#
#     $ conda activate tensorflow-pawel
#
# To deactivate an active environment, use
#
#     $ conda deactivate

Retrieving notices: ...working... done
(base) pinnacle-l5:pwolinsk:~$ conda activate tensorflow-pawel
(tensorflow-pawel) pinnacle-l5:pwolinsk:~$
```

Users can install into private environments

Conda (Python Modules) – install into a private environment

Package, dependency and environment management for Python (and other languages)

conda install -c *<channel>* package=*<version>*=*<build>*

```
(tensorflow-pawel) pinnacle-l5:pwolinsk:~$ conda install -c conda-forge tensorflow=2.10.0=cpu_py310hd1aba9c_0
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##
  environment location: /home/pwolinsk/.conda/envs/tensorflow-pawel
  added / updated specs:
    - tensorflow==2.10.0=cpu_py310hd1aba9c_0

The following packages will be downloaded:
    package                    |            build
    ---------------------------|-----------------
    _openmp_mutex-4.5          |            2_gnu          23 KB  conda-forge
    absl-py-1.2.0              |     pyhd8ed1ab_0          94 KB  conda-forge
    aiohttp-3.8.3              |  py310h5764c6d_0         449 KB  conda-forge
    aiosignal-1.2.0            |     pyhd8ed1ab_0          12 KB  conda-forge
 ...
  wheel              conda-forge/noarch::wheel-0.37.1-pyhd8ed1ab_0
  wrapt              conda-forge/linux-64::wrapt-1.14.1-py310h5764c6d_0
  xz                 conda-forge/linux-64::xz-5.2.6-h166bdaf_0
  yarl               conda-forge/linux-64::yarl-1.7.2-py310h5764c6d_2
  zipp               conda-forge/noarch::zipp-3.8.1-pyhd8ed1ab_0
  zlib               conda-forge/linux-64::zlib-1.2.12-h166bdaf_3
Proceed ([y]/n)? y

Downloading and Extracting Packages
markdown-3.4.1       | 65 KB     |
################################################################################################################################ | 100%
```

Users can install into private environments

Conda (Python Modules) – install into a private environment

Package, dependency and environment management for Python (and other languages)

conda install -c *<channel>* package=*<version>*=*<build>*

```
libzlib-1.2.12       | 65 KB     |
########################################################################################################################## | 100%
grpc-cpp-1.47.1      | 5.1 MB    |
########################################################################################################################## | 100%
wrapt-1.14.1         | 51 KB     |
########################################################################################################################## | 100%
readline-8.1.2       | 291 KB    |
########################################################################################################################## | 100%
libedit-3.1.20191231 | 121 KB    |
########################################################################################################################## | 100%
zlib-1.2.12          | 92 KB     |
########################################################################################################################## | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
Retrieving notices: ...working... done
(tensorflow-pawel) pinnacle-l5:pwolinsk:~$ ls .conda/envs/tensorflow-pawel/
bin/                        etc/                        man/                        ssl/
compiler_compat/            include/                    sbin/                       x86_64-conda_cos6-linux-gnu/
conda-meta/                 lib/                        share/                      x86_64-conda-linux-gnu/
(tensorflow-pawel) pinnacle-l5:pwolinsk:~$ ls .conda/envs/tensorflow-pawel/bin/
2to3                    genbrk                  h5dump                  kpasswd          pip3
tensorboard
...
```

Users can install into private environments