
Introduction to BioHPC

The meeting will begin soon.
Please mute your microphone.

[email] biohpc-help@utsouthwestern.edu
[register] portal.biohpc.swmed.edu/accounts/register
[portal] portal.biohpc.swmed.edu

Welcome! A Quick Note Before Getting Started...

- Most future BioHPC training sessions will be hybrid!
- Choose to join us online, or In-Person
- **Classroom Location: G9.102**
- Users are encouraged to attend in-person.



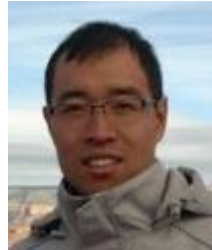
The BioHPC Team



Liqiang Wang
Director



Paniz Karbasi
*Ph.D. Electrical
and computer
engineering*



Peng Lian Ph.D.
*Biomedical
Engineering,
Bioinformatics*



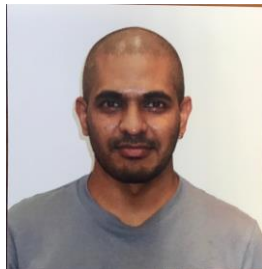
**Xiaochu Lou -
Ph.D.**
Biochemistry



Devin O'Kelly
*Ph.D., Biomedical
and Molecular
Imaging*



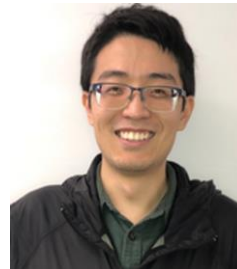
Hung Le
*Ph.D.,
Computational
Chemistry*



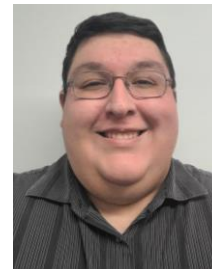
**Suresh
Pannerselvam**
*Ph.D., Medical
Sciences*



Kenny Huynh
*B.S., Computer
Science*



Xueyan Li
*M.S.,
Computational
Geophysics*



Felix Perez
*M.S., Computer
Science*



Joel Sandoval
*B.S., Biomedical
Engineering*

BioHPC institutional resources are available to member departments

- **Center for Alzheimer's and Neurodegenerative Diseases**
- **Cecil H. and Ida Green Center Reproductive Biology Sciences**
- **Children's Research Institute**
- **Department of Cell Biology**
- **Department of Biophysics**
- **Department of Biochemistry**
- **Department of Immunology**
- **Department of Ophthalmology**
- **Department of Pathology**
- **Department of Radiology**
- **Department of Urology**
- **Department of Population and Data Sciences**
- **Department of Pediatrics**
- **Department of Psychiatry**
- **Department of Internal Medicine**
- **Department of Anesthesiology and Pain Management**
- **Green Center for Systems Biology**
- **Hamon Center for Therapeutic Oncology Research**
- **Harold C. Simmons Comprehensive Cancer Center**
- **Lyda Hill Department of Bioinformatics**
- **McDermott Center for Human Growth and Development**
- **Peter O'Donnell Jr. Brain Institute**
 - **And sub-departments**
- **Texas Institute for Brain Injury and Repair**
- **Touchstone Diabetes Center**

Regularly updated list available at:

<https://portal.biohpc.swmed.edu/content/about/>

If you don't see your department in our membership list, come talk with us biohpc-help@utsouthwestern.edu

BioHPC Business Model

- The UT Southwestern BioHPC team provides and maintains high-performance computing, storage and client systems for the UTSW research community.
- The business model can be summarized as **“By the users - For the users”**
 - BioHPC team works directly with Department chair or administrator on resource requirements
 - Department chair or administrator internally coordinates with PIs on the total expected amount of compute resources and storage.
- Questions? Please reach out to us: biohpc-help@utsouthwestern.edu

What is HPC, and why do we need it?

High-performance computing (HPC) is the use of aggregated computing power to solve large-scale or computationally intensive problems in science and engineering.



The power to solve problems that personal computers can't

CHALLENGES

Huge, Diverse, "Big Data" Datasets

Complex Algorithms

Difficult / Inefficient Software

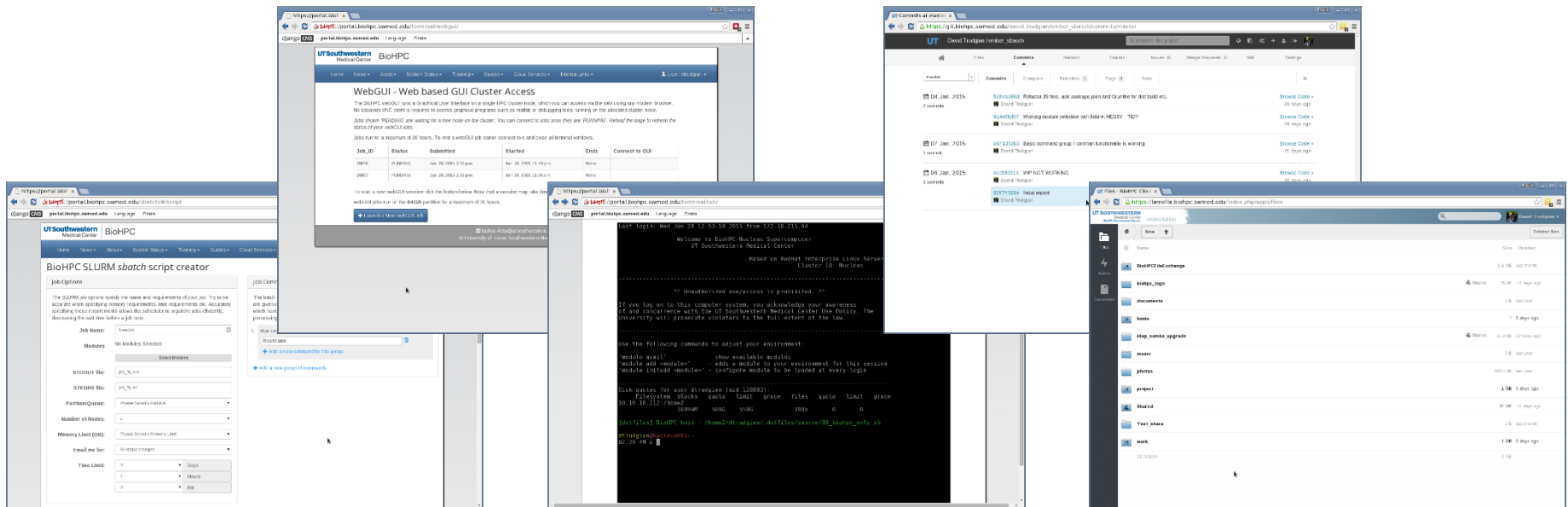
Quick Start - Usage Scenario

- **You have a complex analysis that takes a long time to run on your desktop PC,** and you would like to use your desktop for other things like doing research, making figures and writing manuscripts.
- Everything is on your desktop's hard drive: scripts, data, libraries for running code, etc.
- You would like to securely store your data and code on BioHPC and use its computational power to run your analysis.

What does BioHPC provide? – Cloud Services

A big focus at BioHPC is **easy access** to our systems.

Our **cloud services** provide web-based access to resources, with only a browser.

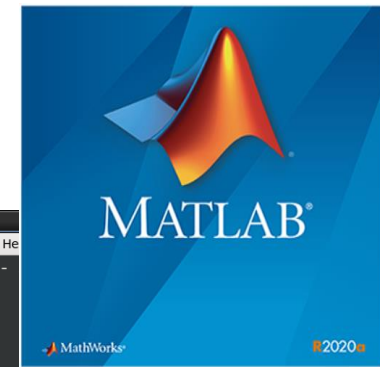


All cloud services accessible via portal.biohpc.swmed.edu

What does BioHPC provide? – Software access

A wide range of software packages are available on Nucleus as *modules*.

Module	Usage	url	Keywords	Versions
3D modeling				
mricon	0	http://people.cas.sc.edu/rorden/mricon/index.html	imaging	10-2016
Alignment				
bustools	1	https://github.com/BUSTools/bustools	RNA-Seq, Alignment, analysis	0.39.3
mauve	0	http://darlinglab.org/mauve/mauve.html	genome, alignment	2.4.0
star	4968	https://github.com/STAR-Fusion/STAR-Fusion	RNA-Seq, Alignment, analysis	2.7.2b; 2.4.2a; 2.5.2b; 2
Alignment				
star	4968	https://github.com/STAR-Fusion/STAR-Fusion	RNA-Seq, Alignment, analysis	2.7.2b; 2.4.2a; 2.5.2b; 2
Assembler				
spades	2180	http://cab.spbu.ru/files/release3.13.0/manual.html	genome, assembly	gcc3.13.0
Bioinformatics				
bfast	0	https://sourceforge.net/projects/bfast	search, sequencing	0.7.0a
bismark	104	http://www.bioinformatics.babraham.ac.uk/projects/bismark/	cytosine, methylation, mapping	0.21.0; 0.14.5;
Biophysics				
nmrpipe	0	https://www.ibbr.umd.edu/nmrpipe/index.html	nmr, spectroscopy, pipeline	20181101
Biostatistics				
mzmine	12	http://mzmine.github.io/	LC-MS, biostatistics, spectroscopy	2.40.1
Biostatistics				
R	8418	http://www.r-project.org	Biology, Biostatistic	3.3.2-gccmk1(default); 2.15.3.1.0-intel; 3.2.1-intel; 3.4.1-gccmk1; 3.4.1-gccmk1; 3.5.1-gccmk1;
C				



```
File Edit View Search Terminal He
-----
abyss/1.3.6
acml/gcc/64/5.3.1
acml/gcc/fma4/5.3.1
acml/gcc/mp/64/5.3.1
acml/gcc/mp/fma4/5.3.1
acml/gcc-int64/64/5.3.1
acml/gcc-int64/fma4/5.3.1
acml/gcc-int64/mp/64/5.3.1
acml/open64/64/5.3.1
acml/open64/fma4/5.3.1
acml/open64/mp/64/5.3.1
acml/open64/mp/fma4/5.3.1
acml/open64-int64/64/5.3.1
acml/open64-int64/fma4/5.3.1
acml/open64-int64/mp/64/5.3.1
almost/2.1
amber/12
annovar/lastest
autoBuster/1.10.0
automake/1.13
bcftools/1.1
intel-cluster-runtime/mic/5.3
intel-tbb-oss/ia32/41_20130613oss
intel-tbb-oss/intel64/41_20130613oss
iozone/3_414
java/oracle/jdk1.7.0_51
lapack/gcc/64/3.4.2
lapack/open64/64/3.4.2
matlab/2013a
matlab/2013b
matlab/2013b_mdcs
matlab/2014a
matlab/2014b
mmdb/1.23.2.2
mpfr/3.1.2
mpiBLAST/1.6.0
mpich/ge/gcc/64/3.0.4
mpich/ge/open64/64/3.0.4
mpich/intel/3.0.4
mpiexec/0.84_432
```

You can ask biohpc-help@utsouthwestern.edu for additions/upgrades etc.

UTSW Network

Your PC

SSH

Browser

Drive Mount

Portal

Lamella

BioHPC

nucleus005

On-Demand
(Cloud)
Services

SLURM Scheduler

Nucleus (Compute)

CPU Nodes

GPU Nodes

Storage

home2

work

project

archive

Internal
cloud
storage

The World

Outside UTSW

BioHPC Cloud

External cloud

Legend

Node/Computer

Storage

Interface

Interaction

Nucleus is our supercomputing cluster (our “compute”)

CPU Nodes

Light Nodes:

32 GB 276 nodes

Heavy Nodes:

128 GB 24 nodes

256 GB 89 nodes

256 GBv1 54 nodes

384 GB 18 nodes

Nucleus (Compute)

CPU Nodes

GPU Nodes

GPU Nodes

Tesla K20/K40 8 nodes

Tesla P4 16 nodes

Tesla P40 16 nodes

2x Tesla P100 12 nodes

Tesla V100 32 nodes

4x Tesla V100 12 nodes

Ampere A100 16 nodes

4x Ampere A100 16 nodes

Coming soon

256 GB ~300 nodes

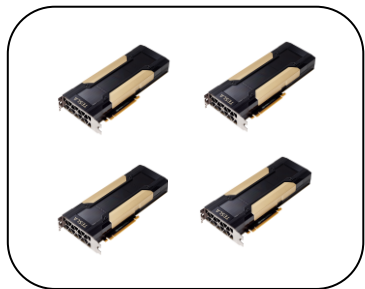
We are installing 100 new nodes.

<https://portal.biohpc.swmed.edu/content/about/systems/>

RHEL 7.7, GNOME 3, Bash

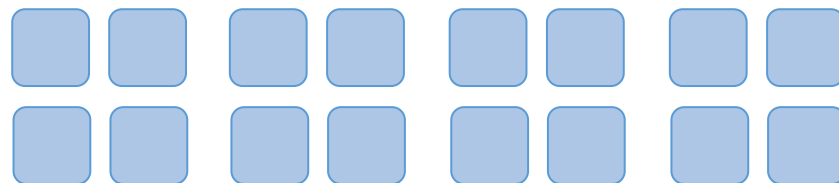
How many nodes can I use at once (if they are available)?

4 GPU nodes



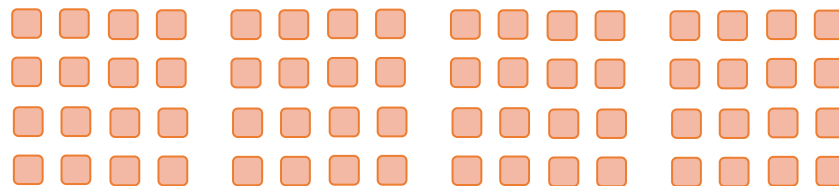
AND

16
heavy nodes
(> 32 GB ea.)



OR

64
light nodes
(32 GB ea.)



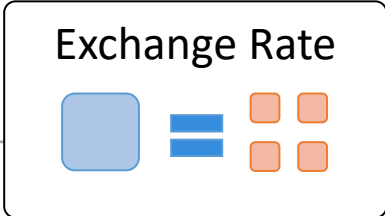
Exchange Rate



One (1)
Heavy

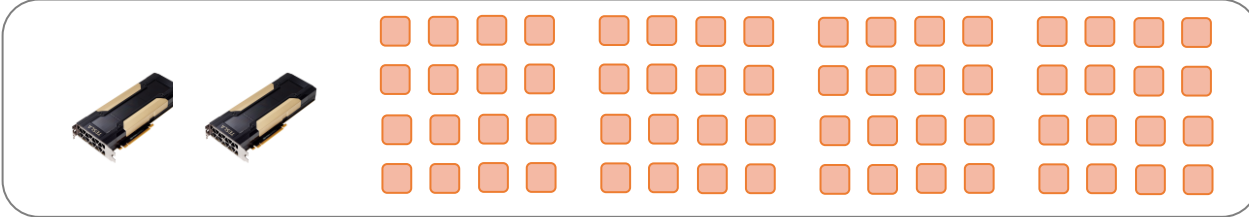
Four (4)
Light

Possible total allocations

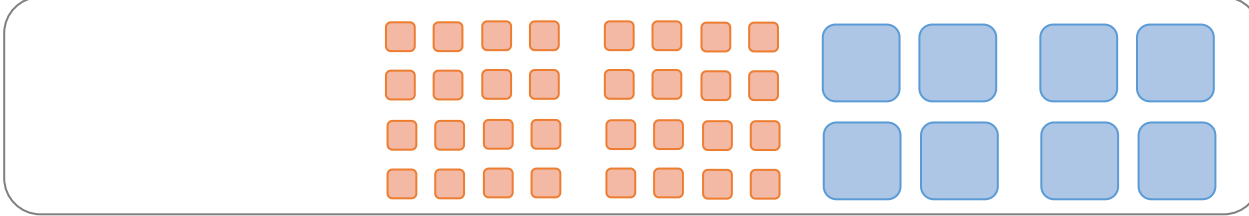


2 GPU + 64 light nodes

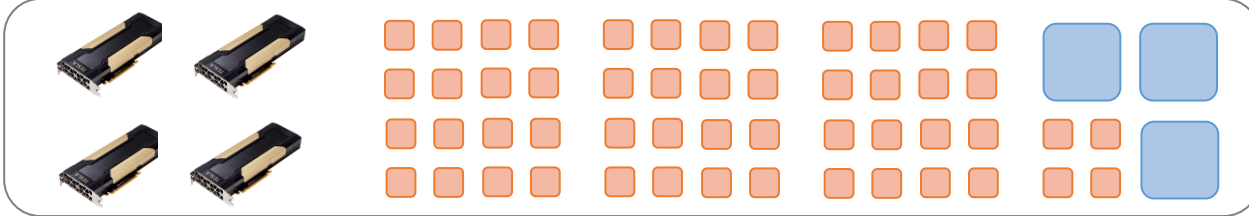
1024 CPU cores
2048 GB RAM



0 GPU + 8 heavy + 32 light

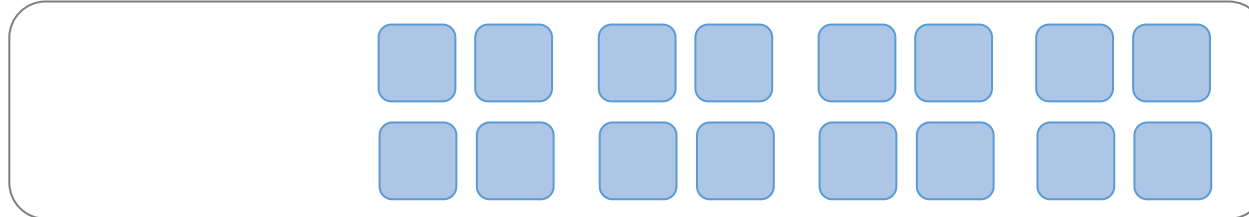


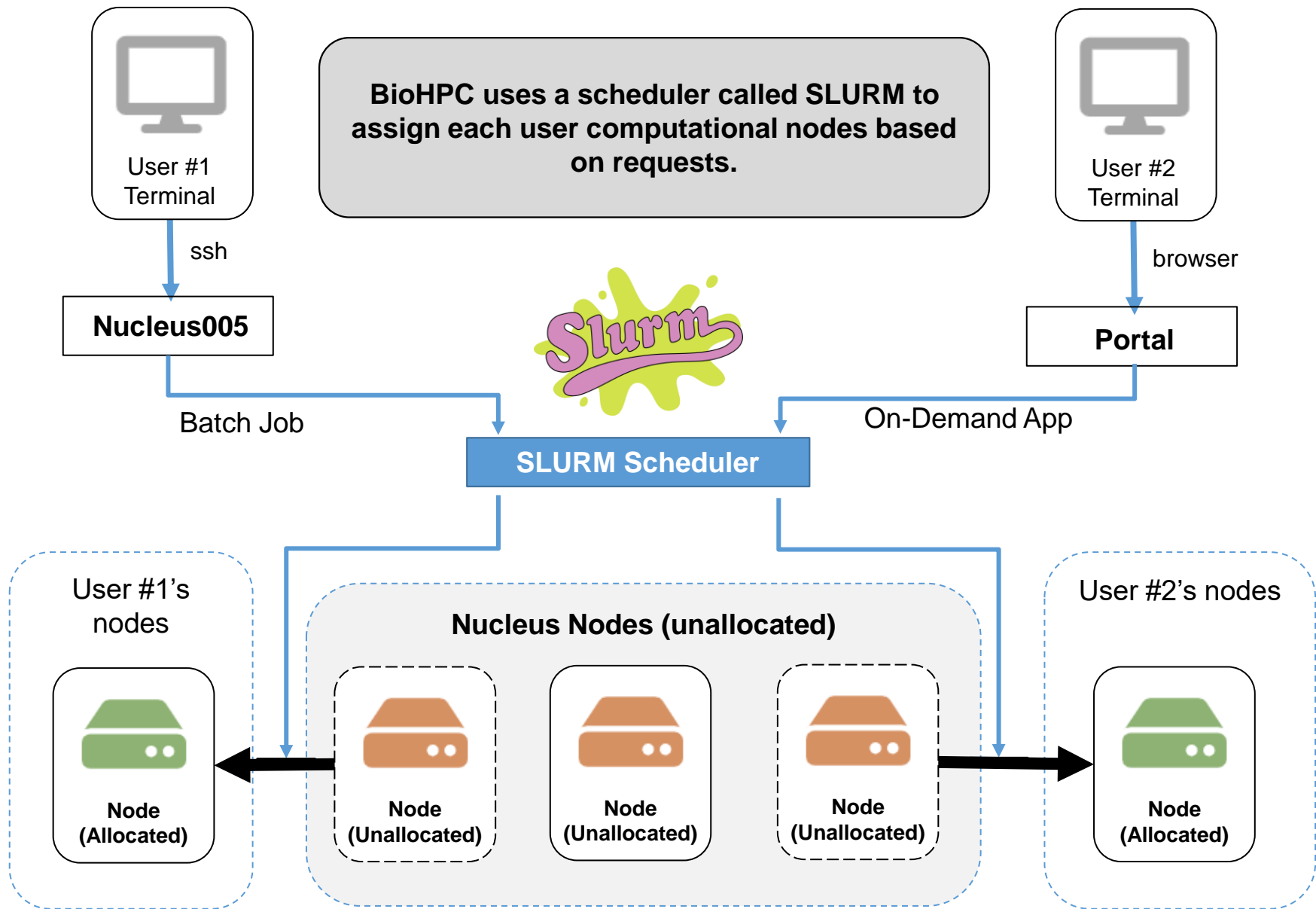
4 GPU + 3 heavy + 52 light



0 GPU + 16 heavy

448 CPU cores
4096 GB RAM





BioHPC - Storage

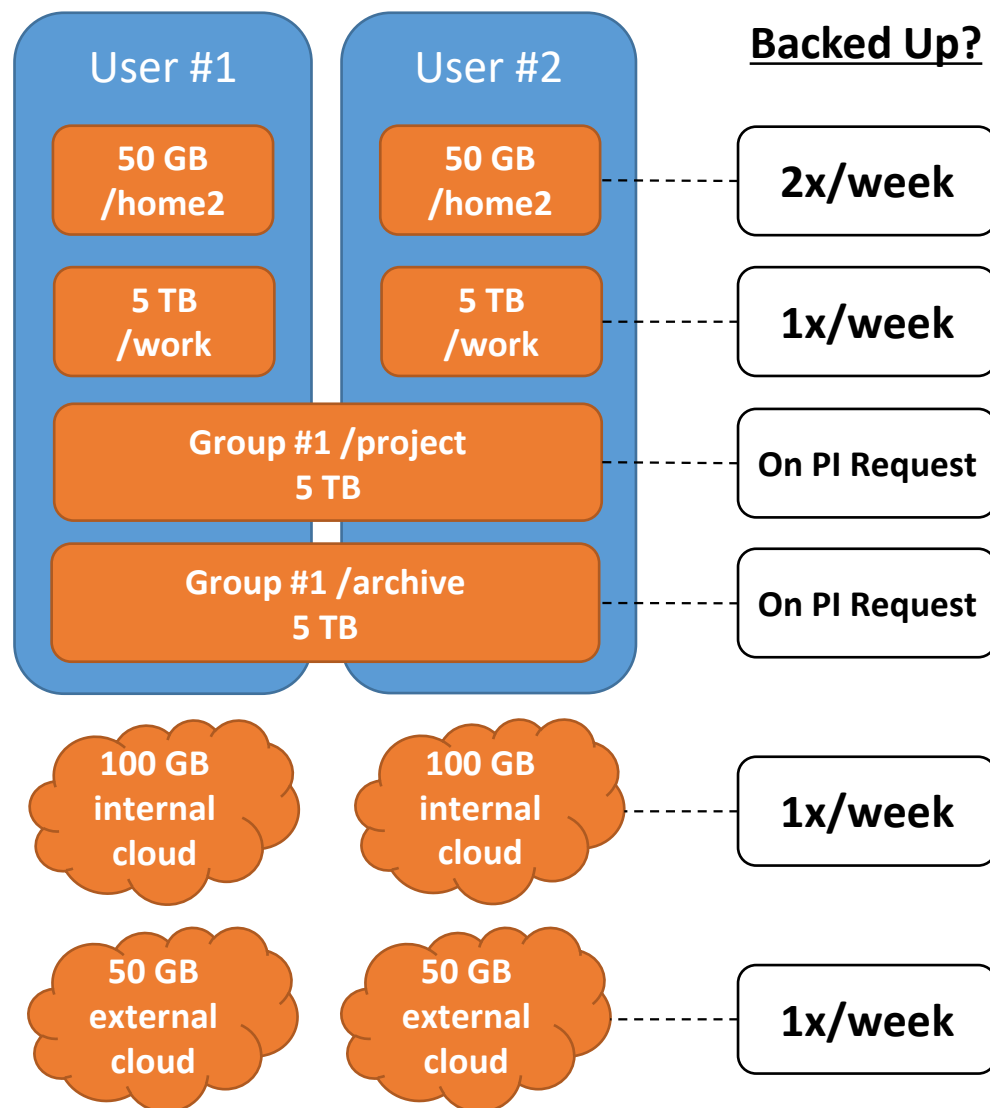
Every user has a **user quota**.
Every group has a **group quota**.
Group quotas are shared among members of that group.

Every user gets:






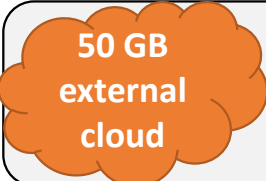
- **50 GB** in **/home2**
- **5 TB** in **/work**
- **100 GB** internal cloud storage
- **50 GB** external cloud storage

Every group gets:

- **5 TB** in **/project**
- **5 TB** in **/archive** – **7.5 TB effective**
 - **2/3x “data discount”**



Info and backup policy: <https://portal.biohpc.swmed.edu/content/guides/storage-cheat-sheet/>

Scripts, Configurations		<u>/home2/<username></u> Small files which are frequently accessed
“Hot” data		<u>/work/<department>/<username></u> Data which is read and written frequently; active analyses
“Warm” data		<u>/project/<department>/<group>/<username></u> Data which is read and written occasionally; raw data
“Cold” data		<u>/archive/<department>/<group>/<username></u> Data which is rarely read; archival/retained data
Inside UTSW		Dropbox-like interface, for sharing <u>inside</u> UTSW network
Outside UTSW		Dropbox-like interface, for sharing <u>outside</u> UTSW network

Exceptions for O'Donnell Brain Institute (OBI)

If you are a member of OBI's sub-departments, your paths are:

The same!

/home2

/home2/<username>

Small files which are frequently accessed

/work

/work/OBI/<department>/<username>

Data which is read and written frequently; active analyses

/project

/project/OBI/<department>/<group>/<username>

Data which is read and written occasionally; raw data

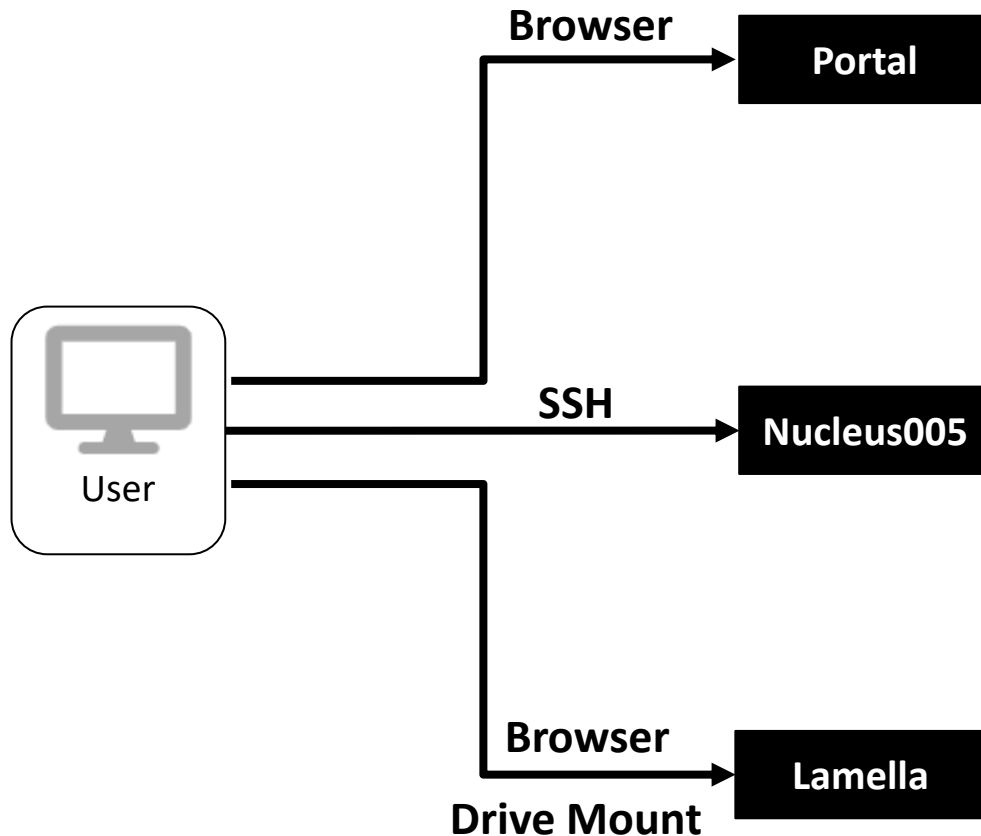
/archive

/archive/OBI/<department>/<group>/<username>

Data which is rarely read; archival/retained data

Whenever you see <department> in general instructions, replace it with OBI/<department>

Three basic ways to interact with BioHPC



portal.biohpc.swmed.edu

Web Services, Guides,
Training, Information

nucleus.biohpc.swmed.edu

The login node. Schedule
jobs, hop to another node

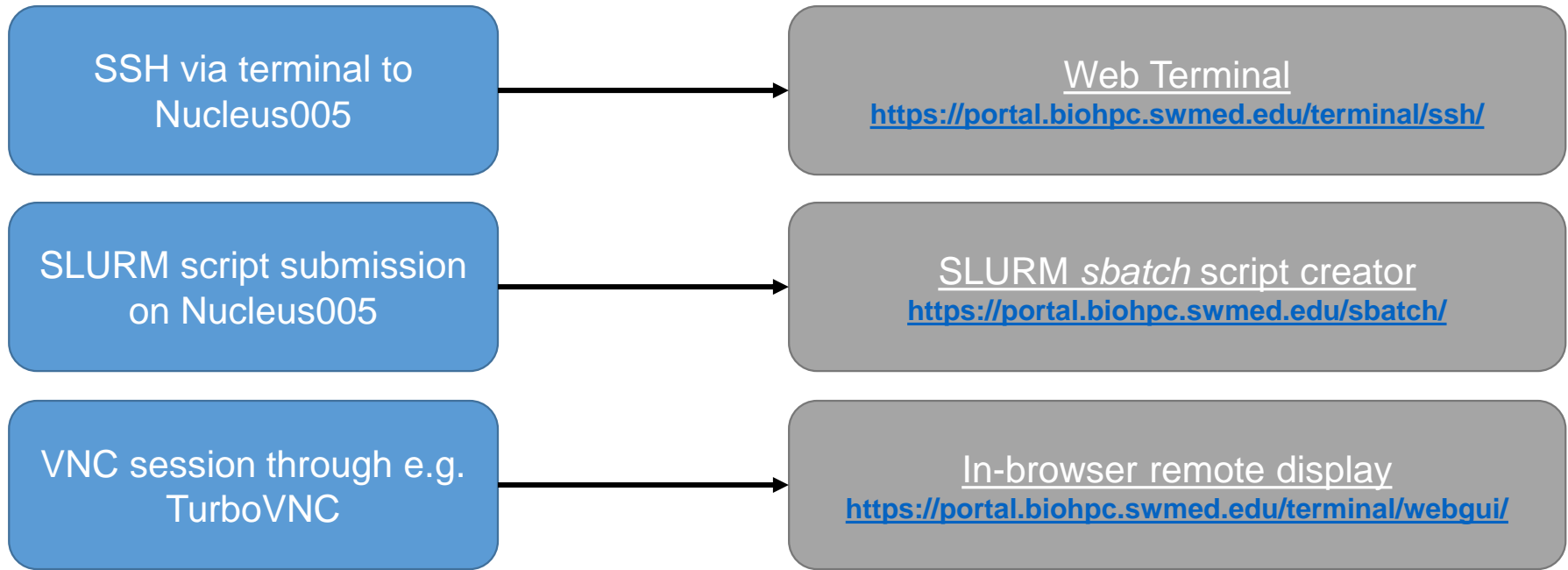
lamella.biohpc.swmed.edu

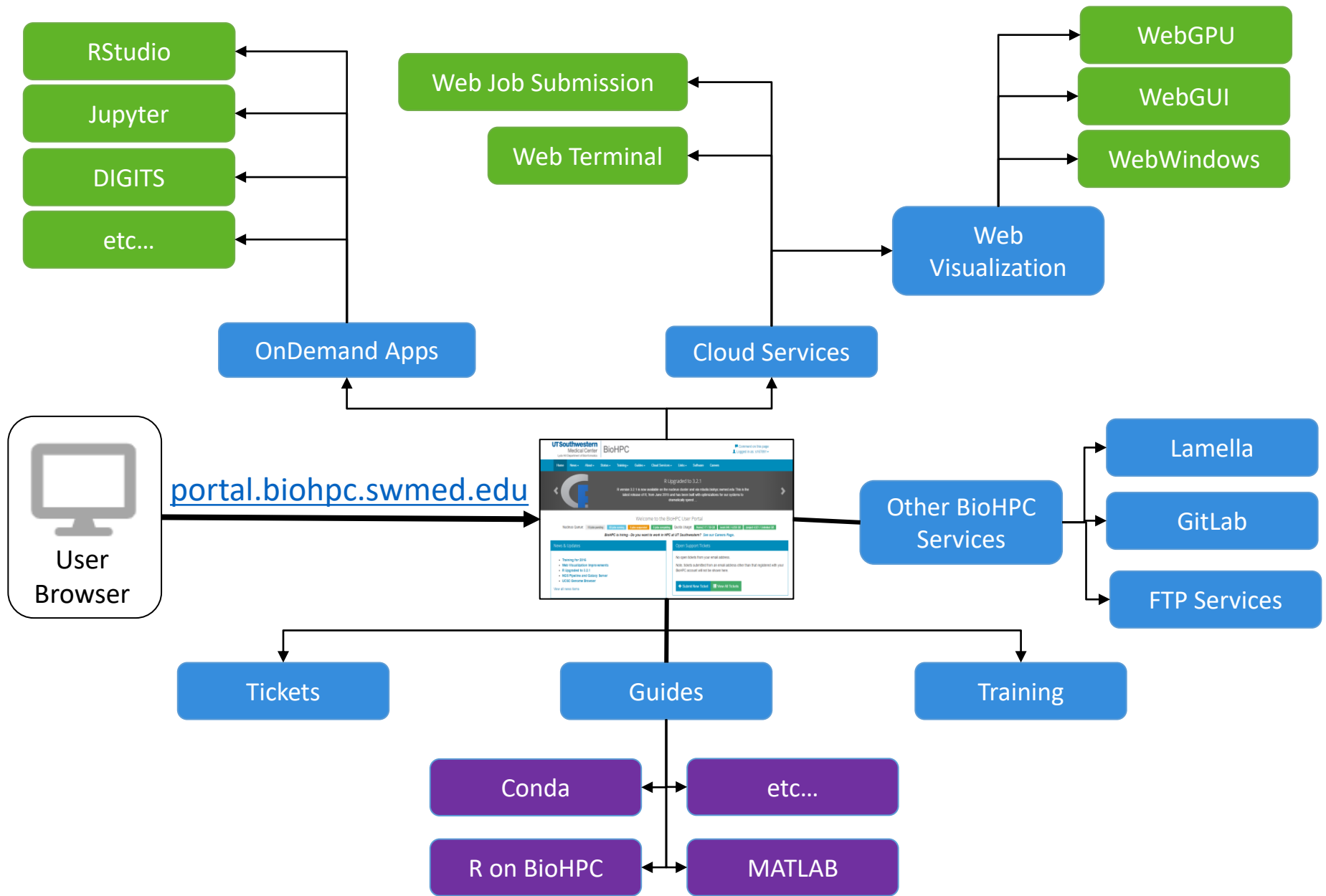
Storage

Accessing BioHPC – Portal

Our goal is for BioHPC to be available to users of all technical backgrounds, and to streamline the process of becoming more comfortable and familiar with HPC by lowering the barriers to access.

Many basic tasks have a Portal-based alternative.





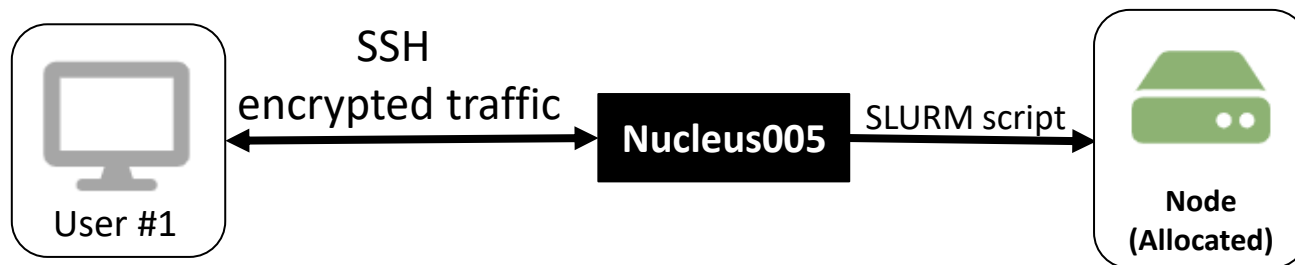
Accessing BioHPC – Secure Shell (SSH)

SSH provides a command line that you can use to perform tasks on a remote computer.

Users may connect to the Nucleus login node (**nucleus005**) via Secure Shell (SSH) sessions. This is preferred to other methods (e.g. WebGUI) as it does not use up a node.

```
> ssh <username>@nucleus.biohpc.swmed.edu
```

Nucleus005 is the main hub, and is for scheduling jobs to run on **other nodes**.
You should never run analyses or intensive computation on Nucleus005



Accessing BioHPC - SSH Clients

Linux

Most Linux distributions have OpenSSH installed by default.

Simply open a terminal window.

Mac

The Mac Terminal should already be able to SSH by default.

Windows

Either:

Install an SSH client

- PuTTY
- Windows Subsystem for Linux (WSL)
- Git Bash (provided by Git)
- ConEmu
- MobaXterms

Or use Command Prompt (cmd) or Powershell

BioHPC Portal

Navigate to Web Terminal.

Portal > Cloud Services > Web Terminal

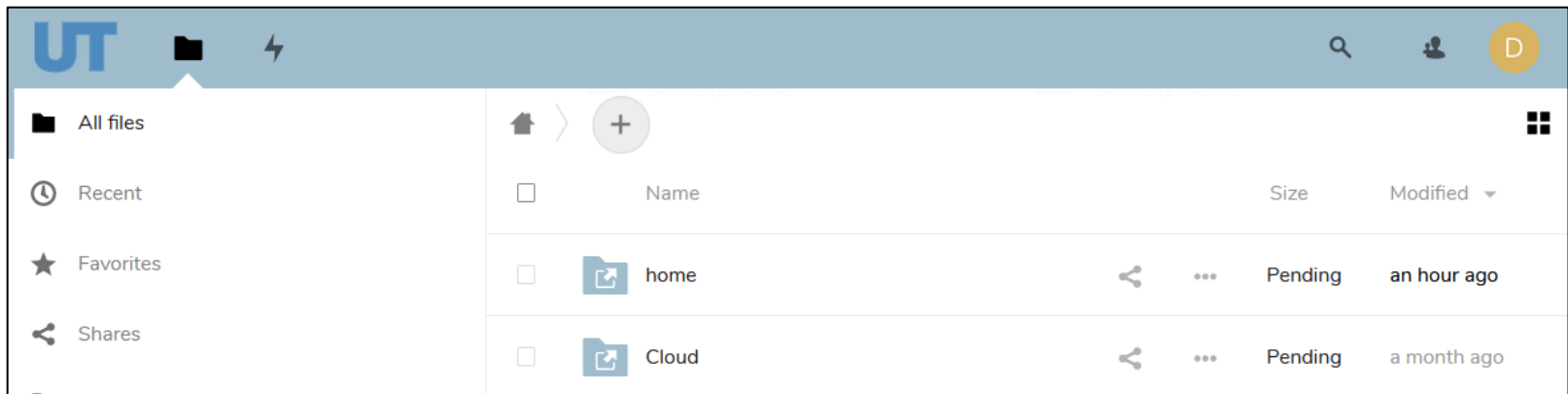
<https://portal.biohpc.swmed.edu/terminal/ssh/>

Accessing BioHPC - Lamella

Lamella is BioHPC's storage gateway. By default, it is configured to provide cloud access to your internal cloud, external cloud, and home2 storage.

Lamella translates data transfer protocols so that you can access BioHPC storage from non-BioHPC computers.

There is a browser-based interface, similar to Dropbox or other cloud storage, at <http://lamella.biohpc.swmed.edu>



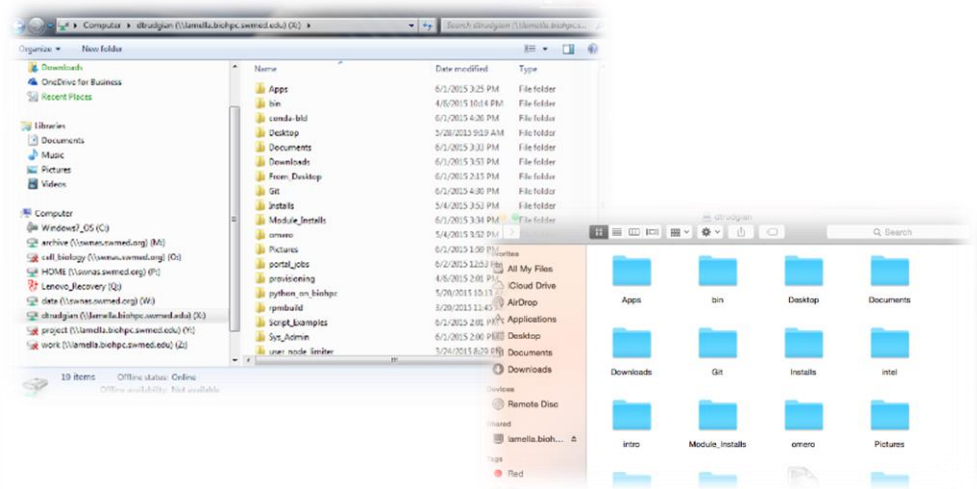
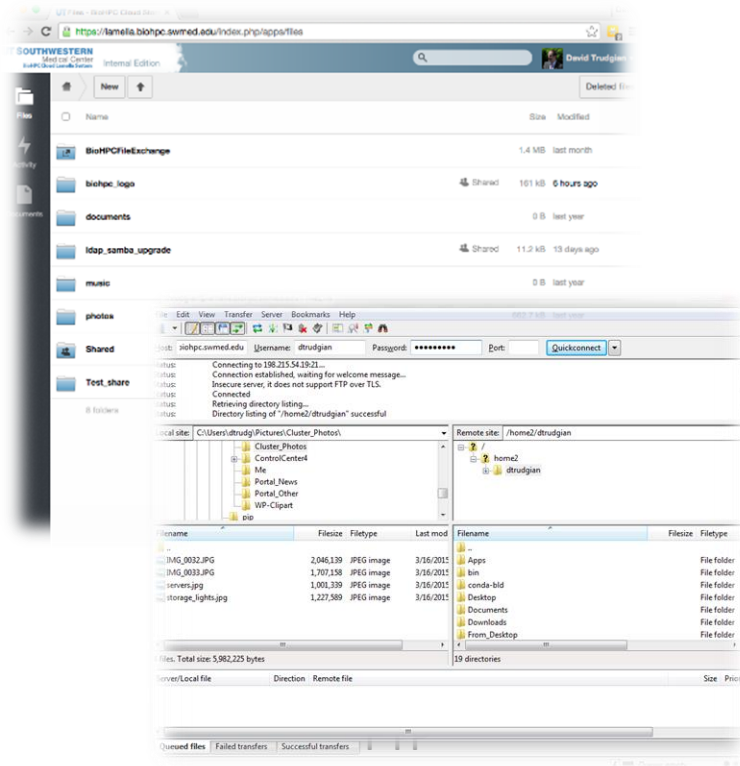
Cloud storage guide: <https://portal.biohpc.swmed.edu/content/guides/biohpc-cloud-storage/>

Accessing BioHPC - Lamella

Multiple ways to access your files:

Web Interface (Lamella)
FTP Clients (Thunder/Flash)

Windows / Mac network drive mounts
(SMB /WebDav)



lamella.biohpc.swmed.edu

Accessing BioHPC – Thin Clients and Workstations

- BioHPC has specially-configured desktop clients available for order.
- Can submit jobs directly to the cluster (**like Nucleus005**)
- Direct access to the cluster storage systems.
- Same software stack and modules as compute nodes.
- Provides you with a graphical desktop (like a web visualization session)
- **Customizable (to some extent)**, persistent, can run some software we disallow on compute nodes (e.g. Virtualbox)
- Thin clients are useful for light workloads.
- Workstations are designed for intensive development or local computation.



Okay, sounds great....

But how do I use any of this?

Quick Start - Usage Scenario

You have a complex analysis that takes a long time to run on your desktop PC, and you would like to use your desktop for other things like doing research, making figures and writing manuscripts.

We will walk through a sketch of the process via the following basic walkthrough:

1. Move data and code from your desktop to BioHPC, via **Lamella mounts**
2. Submitting a SLURM batch script for your code, via **Web Job Submission**
3. Checking on our script via the **Web Terminal**
4. Visualizing our results via a **Web Visualization** session and **loading modules**
5. Running OnDemand apps

Hands on BioHPC – 1. Manage Files with Lamella / Cloud Storage Gateway

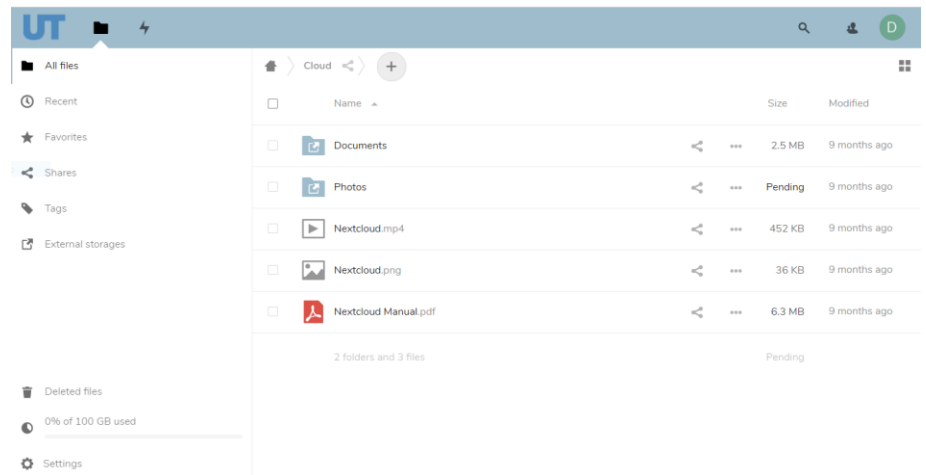
Cloud storage gateway – web-based.

<https://lamella.biohpc.swmed.edu>

- 100GB separate space +
- Mount /home /project /work
- Internal

<https://cloud.biohpc.swmed.edu>

- 50GB space
- External file transfer
- Accessible from Internet



Hands on BioHPC – 1. Setting up Lamella to access project and workspace

<https://lamella.biohpc.swmed.edu>

Log-in credentials, save in session uses the BioHPC login credentials and only saved in the user session, giving increased security. The drawbacks are that sharing is disabled, as lamella has no access to the cluster storage credentials.

Username and password mechanism requires a manually-defined username and password. Remember to click the gear icon and enable sharing.

The screenshot shows the 'External Storage' configuration page. It features a table with columns for 'Folder name', 'External storage', 'Authentication', and 'Configuration'. There are two rows for 'project' and 'work' folders, both using 'BioHPC Lysosome' external storage and 'Username and password' authentication. The 'Configuration' column contains fields for folder name, external storage name, and password. Annotations include orange arrows pointing to the 'Log-in credentials, save in session' option, the 'project' and 'work' rows, and the 'Add storage' button. A red box highlights the 'Enable sharing' checkbox in the configuration dropdown menu. Text at the bottom explains the folder naming conventions for home, private project space, and lab shared project space.

Folder name	External storage	Authentication	Configuration	username	password
home	BioHPC Lysosome	Log-in credentials, save in session			
Cloud	Cloud	Log-in credentials, save in session			
project	BioHPC Lysosome	Username and password	project	biohpcadmin/ydu	ydu
work	BioHPC Lysosome	Username and password	work	biohpcadmin/ydu	ydu
Folder name	Add storage				

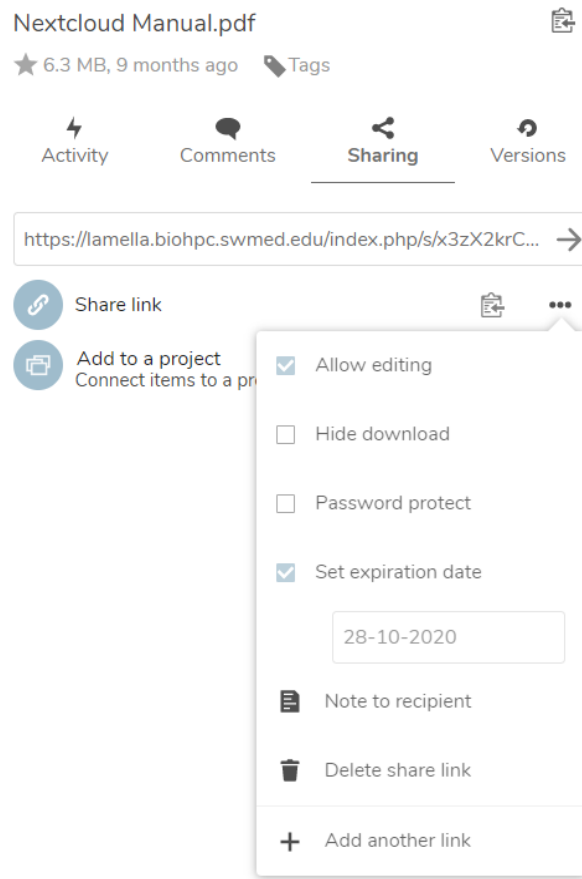
BioHPC Endosome/Lysosome

project
work
home

For home leave blank
For private project space:
department/lab/user
For lab shared project space:
department/lab/shared

Hands on BioHPC – 1. Manage Files with Lamella / Cloud Storage Gateway

File Sharing



Lamella cloud storage <https://lamella.biohpc.swmed.edu> : sharing with user **inside** UTSW
File Exchange <https://cloud.biohpc.swmed.edu> : sharing with user **outside** UTSW

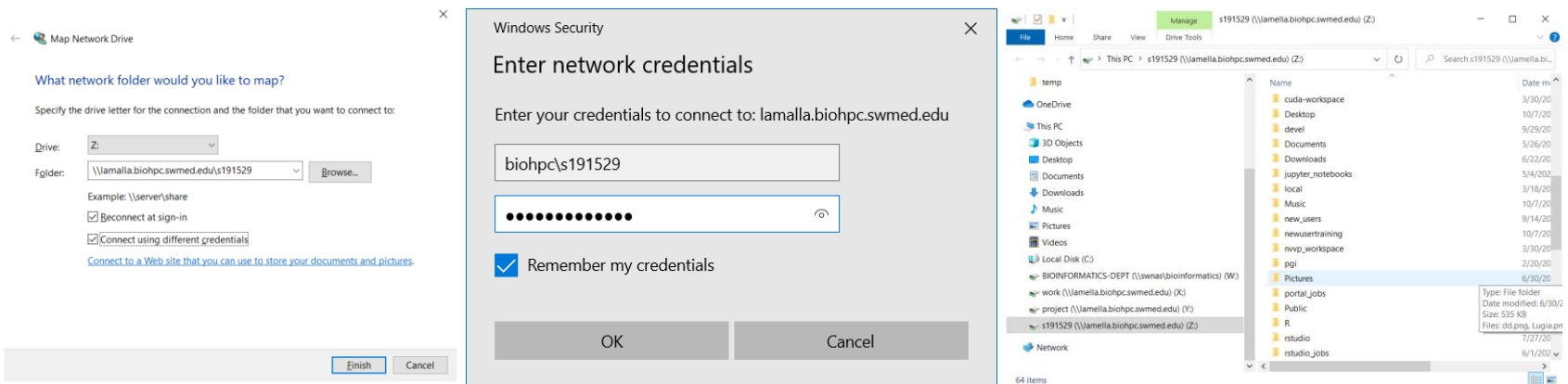
Hands on BioHPC – 1. Accessing BioHPC Storage Directly from Windows

Computer -> Map Network Drive

Folder is: \\lamella.biohpc.swmed.edu\username (home dir)
\\lamella.biohpc.swmed.edu\project
\\lamella.biohpc.swmed.edu\work

Check 'Connect using different credentials'

Enter your BioHPC username and password when prompted.

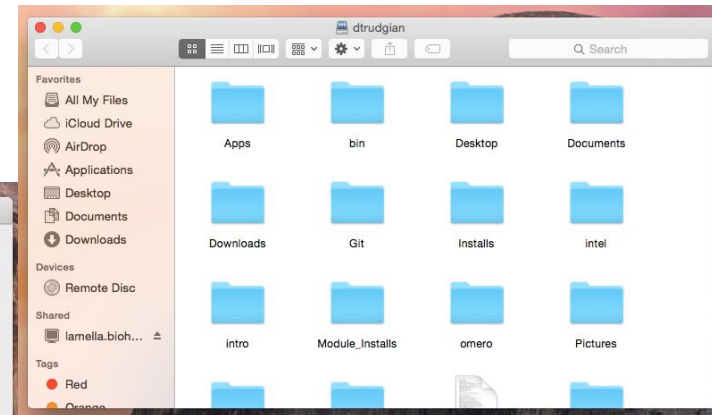
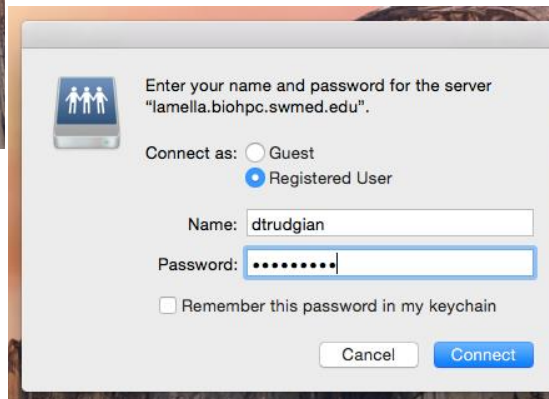
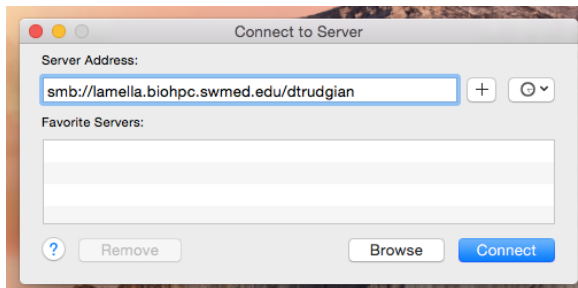


Hands on BioHPC – 1. Accessing BioHPC Storage Directly from Mac OSX

Finder -> Go -> Connect to Server

Folder is: `smb://lamella.biohpc.swmed.edu/username` (*home dir*)
 `smb://lamella.biohpc.swmed.edu/project`
 `smb://lamella.biohpc.swmed.edu/work`

Enter your BioHPC username and password when prompted.



Hands on BioHPC – 2. Web Job Script Generator

<https://portal.biohpc.swmed.edu> -> Cloud Services -> Web Job Submission

UT Southwestern Medical Center BioHPC

Home News About System Status Training Guides Cloud Services Internal Links User: dtrudgian

BioHPC SLURM *sbatch* script creator

Job Options

The SLURM job options specify the name and requirements of your job. Try to be accurate when specifying memory requirements, time requirements etc. Accurately specifying these requirements allows the scheduler to organize jobs efficiently, decreasing the wait time before a job runs.

Job Name: MyJob

Modules: No Modules Selected

STDOUT file: job_%j.out

STEDRR file: job_%j.err

Partition/Queue: 64GB - 64GB Nodes

Number of Nodes: 1

Memory Limit (GB): Please Select a Memory Limit

Email me for: All status changes

Time Limit: 0 Days, 2 Hours, 0 Min

Job Commands

The batch script contains one or more commands. Each command is given in a group they will be processed in parallel. Each group of commands will be processed on each node.

- Run commands in parallel

hostname

+ Add a new command to group

+ Add a new group of commands

SLURM *sbatch* Script

This is the script created from the options you have chosen on the form. You can edit the script directly, but this will disable the ability to modify it further using the form. When your job is ready click the button at the bottom of this page to submit it to the cluster. You can also copy and paste the script into a file, as a template for future use.

```
#!/bin/bash
#
# CREATED USING THE BIOHPC PORTAL on Thu Jan 29 2015 14:56:28 GMT-0600 (CST)
#
# This file is batch script used to run commands on the BioHPC cluster.
# The script is submitted to the cluster using the SLURM 'sbatch' command.
# Lines starting with # are comments, and will not be run.
# Lines starting with #SBATCH specify options for the scheduler.
# Lines that do not start with # or #SBATCH are commands that will run.
#
# Name for the job that will be visible in the job queue and accounting tools.
#SBATCH --job-name MyJob
#
# Name of the SLURM partition that this job should run on.
#SBATCH -p 64GB # partition (queue)
# Number of nodes required to run this job
#SBATCH -N 1
#
# Time limit for the job in the format Days-H:M:S
# A job that reaches its time limit will be cancelled.
# Specify an accurate time limit for efficient scheduling so your job runs promptly.
#SBATCH -t 0-2:0:0
#
# The standard output and errors from commands will be written to these files.
# %j in the filename will be replaced with the job number when it is submitted.
#SBATCH -o job_%j.out
#SBATCH -e job_%j.err
#
# Send an email when the job status changes, to the specified address.
#SBATCH --mail-type ALL
#SBATCH --mail-user david.trudgian@UTSouthwestern.edu

module load

# COMMAND GROUP 1
hostname

# END OF SCRIPT
```

Submit Job To Cluster

Hands on BioHPC – 3. Web Terminal

<https://portal.biohpc.swmed.edu> -> Cloud Services -> Web Terminal

Nucleus Login Shell

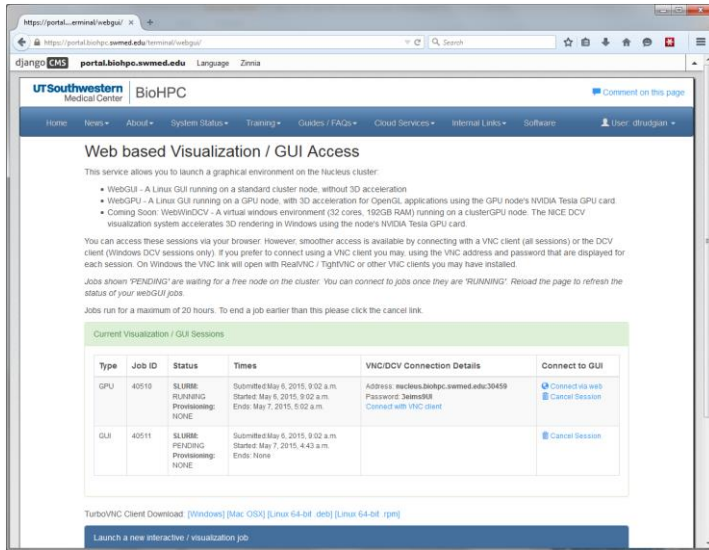
This terminal provides a login session to the BioHPC Nucleus cluster. Please login using your BioHPC password, your username is automatically detected.

For safety, the terminal is only active when your mouse is over it. Remember to log-out of your session (CTRL-D or exit) when you are finished!

Password:

Hands on BioHPC – 4. Web Visualization: Graphical Interactive Session via Web Portal/VNC

<https://portal.biohpc.swmed.edu> -> Cloud Services -> Web Visualization



The screenshot shows the BioHPC web portal with the following content:

Web based Visualization / GUI Access

This service allows you to launch a graphical environment on the Nucleus cluster.

- WebGUI - A Linux GUI running on a standard cluster node, without 3D acceleration
- WebGPU - A Linux GUI running on a GPU node, with 3D acceleration for OpenGL applications using the GPU node's NVIDIA Tesla GPU card
- Coming Soon: WebProDCV - A virtual windows environment (32 cores, 16GB RAM) running on a cluster/GPU node. The NICE DCV visualization system accelerates 3D rendering in Windows using the node's NVIDIA Tesla GPU card.

You can access these sessions via your browser. However, smoother access is available by connecting with a VNC client (all sessions) or the DCV client (Windows DCV sessions only). If you prefer to connect using a VNC client you may, using the VNC address and password that are displayed for each session. On Windows the VNC link will open with RealVNC / TightVNC or other VNC clients you may have installed.

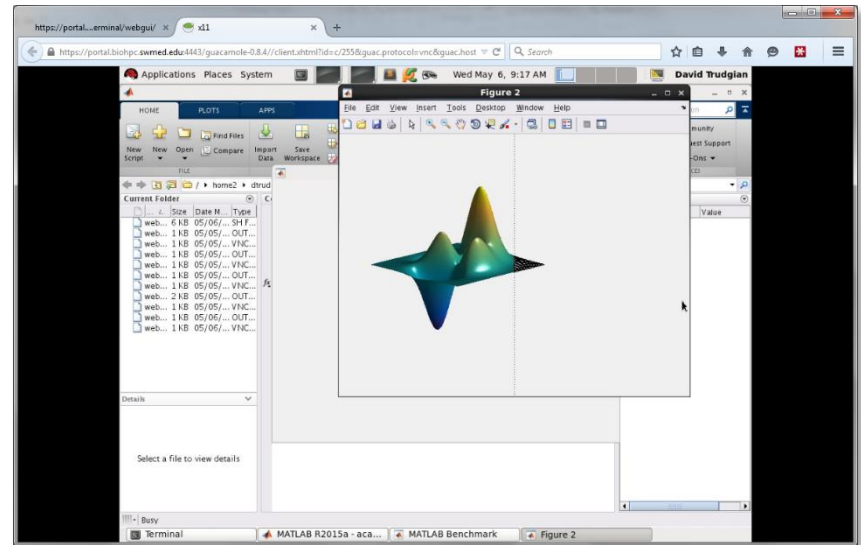
Jobs known "PENDING" are waiting for a free node on the cluster. You can connect to jobs once they are "RUNNING". Reload the page to refresh the status of your webGUI jobs.

Jobs run for a maximum of 20 hours. To end a job earlier than this please click the cancel link.

Type	Job ID	Status	Times	VNC/DCV Connection Details	Connect to GUI
GPU	40510	SLURM: RUNNING Provisioning: NONE	Submitted: May 6, 2015, 9:02 a.m. Started: May 8, 2015, 9:02 a.m. Ends: May 7, 2015, 8:02 a.m.	Address: nucleus.biohpc.swmed.edu:30459 Password: 3e1na9R8 Connect with VNC client	Connect via web Cancel Session
GUI	40511	SLURM: PENDING Provisioning: NONE	Submitted: May 6, 2015, 9:02 a.m. Started: May 7, 2015, 4:43 a.m. Ends: None		Cancel Session

TurboVNC Client Download [Windows] [Mac OS X] [Linux 64-bit deb] [Linux 64-bit rpm]

[Launch a new interactive / visualization job](#)



The screenshot shows a Windows desktop environment with a 3D visualization of a protein structure. The window title is "Figure 2". The desktop background is "PLOT1". The taskbar shows "MATLAB R2015a - oca..." and "MATLAB Benchmark". The system tray shows the date and time: "Wed May 6, 9:17 AM".

Connects to GUI running on a cluster node. WebGPU sessions have access to GPU card for 3D rendering.

Hands on BioHPC – 4. (continued) Software Modules

```

dtrudgian@Nucleus005:~
03:16 PM $ module avail

----- /cm/local/modulefiles -----
cluster-tools/6.1      ipmitool/1.8.12      null                use.own
cmd                    module-git            openldap            version
dot                    module-info           openmpi/gcc/64/1.6.5-mlnx-ofed
freeipmi/1.2.6         mvapich2/gcc/64/1.9-mlnx-ofed  shared

----- /cm/shared/modulefiles -----
abyss/1.3.6            fftw2/openmpi/open64/64/double/2.1.5  NAMD/2.9/ibverbs-smp-CUDA
acml/gcc/64/5.3.1      fftw2/openmpi/open64/64/float/2.1.5    NAMD/2.9/multicore
acml/gcc/fma4/5.3.1    fftw3/openmpi/gcc/64/3.3.3            NAMD/2.9/multicore-CUDA
acml/gcc/mp/64/5.3.1   fftw3/openmpi/intel/3.3.3             netcdf/gcc/64/4.3.0
acml/gcc/mp/fma4/5.3.1  fftw3/openmpi/open64/64/3.3.3         netcdf/intel/4.3.0
acml/gcc-int64/64/5.3.1  fftw3/shared/3.3.4                   netcdf/open64/64/4.3.0
acml/gcc-int64/fma4/5.3.1  gcc/4.8.1                             netperf/2.6.0
  
```

Command	Summary
<code>module list</code>	Show loaded modules
<code>module avail</code>	Show available modules
<code>module load <module_name></code>	Load module into environment
<code>module unload <module_name></code>	Unload module from environment
<code>module help <module_name></code>	Help text for a specific module
<code>module -h</code>	General module command help

***BioHPC may install additional modules upon request.**

Hands on BioHPC – 5. OnDemand apps

BioHPC OnDemand ▾

- OnDemand DIGITS
- OnDemand Jupyter
- OnDemand JupyterLab
- OnDemand BisQue
- OnDemand RStudio
- OnDemand Applications
- OnDemand CryoSPARC
- OnDemand CLARA

BioHPC has made convenient interfaces for accessing some commonly used software packages.

OnDemand apps consume 1 light (32 GB) node each.

Current OnDemand Applications

matlab

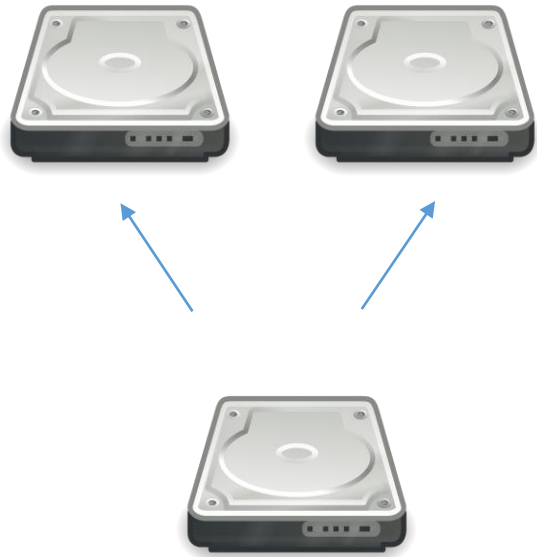
ilastik

chimera

reion

fastqc

What is BioHPC not intended for?



Rigorous file version backups
BioHPC is primarily a *computational resource*, not a *storage provider*.

Non-Research Computing
BioHPC is publicly funded; every user is expected to be responsible with taxpayer dollars.

Things to Remember

1. If you ever have any BioHPC-related questions or need assistance, contact us at ***biohpc-help@utsouthwestern.edu***
2. Be conservative when requesting resources – Do you really need a heavy node when a light node will do?
3. Make reasonable attempts to use the resources efficiently.
 - Use one node for multiple tasks if you can, rather than reserving several nodes.
 - Cancel or close any jobs/session you no longer need.
 - Try to optimize your code.
4. Keep notes in case you need our help troubleshooting.
5. **Never run any intensive code on Nucleus005, AKA login node, AKA web terminal AKA nucleus.biohpc.swmed.edu.**

Persistent, improper use of BioHPC resources is grounds for having your account frozen.

Getting Effective Help

Email the ticket system: biohpc-help@utsouthwestern.edu

What is the problem?

Provide any error messages, logs, and diagnostic output you may have.

When did it happen?

What time? Cluster node or client workstation? What job ID? What folders were you in and accessing?

How did you run it?

What did you run? Which parameters and settings did you use?

Any unusual circumstances?

Have you compiled your own software? Do you customize startup scripts?

Can we look at your scripts and data?

Depending on the problem and its complexity, we may need your permission to access your storage allocation. Please let us know if there is any data that is sensitive.

BioHPC can also provide more detailed assistance to help *facilitate* research.

- **Collaborations**

- Need help beyond a support ticket?
- The BioHPC team can provide in-depth consultations to assist with particular research projects.

- **Liaison**

- Lab with need for *dedicated*, domain-specific computational support
- Co-hired by BioHPC and the lab

Next Steps

- Please register immediately if you have not done so at (UTSW VPN required):
 - <https://portal.biohpc.swmed.edu/accounts/register/>
- New user registrations will be manually reviewed and processed (**a week or two**)
 - Your PI should be prepared to approve your activation.
- You will receive an activation notice when registration is processed.
- Spend some time experimenting with our systems and browse our guides.
- Check the training schedule and attend relevant sessions on BioHPC Portal
 - <https://portal.biohpc.swmed.edu/content/training/calendar-2023/>
- QUESTIONS: biohpc-help@utsouthwestern.edu

Thank you!

Please acknowledge our contribution:

This research was supported in part by the computational resources provided by the BioHPC supercomputing facility located in the Lyda Hill Department of Bioinformatics, UT Southwestern Medical Center.

If your PI wants to include info in a grant proposal on BioHPC as a resource, or request grant funds to contribute to their Department's HPC costs, please contact our Department Administrator, **Rebekah Craig** (**Rebekah.Craig@utsouthwestern.edu**) for advice on grant language.