UTSouthwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Storage Overview of BioHPC

[email]   biohpc-help@utsouthwestern.edu
[lamella]   https://lamella.biohpc.swmed.edu/
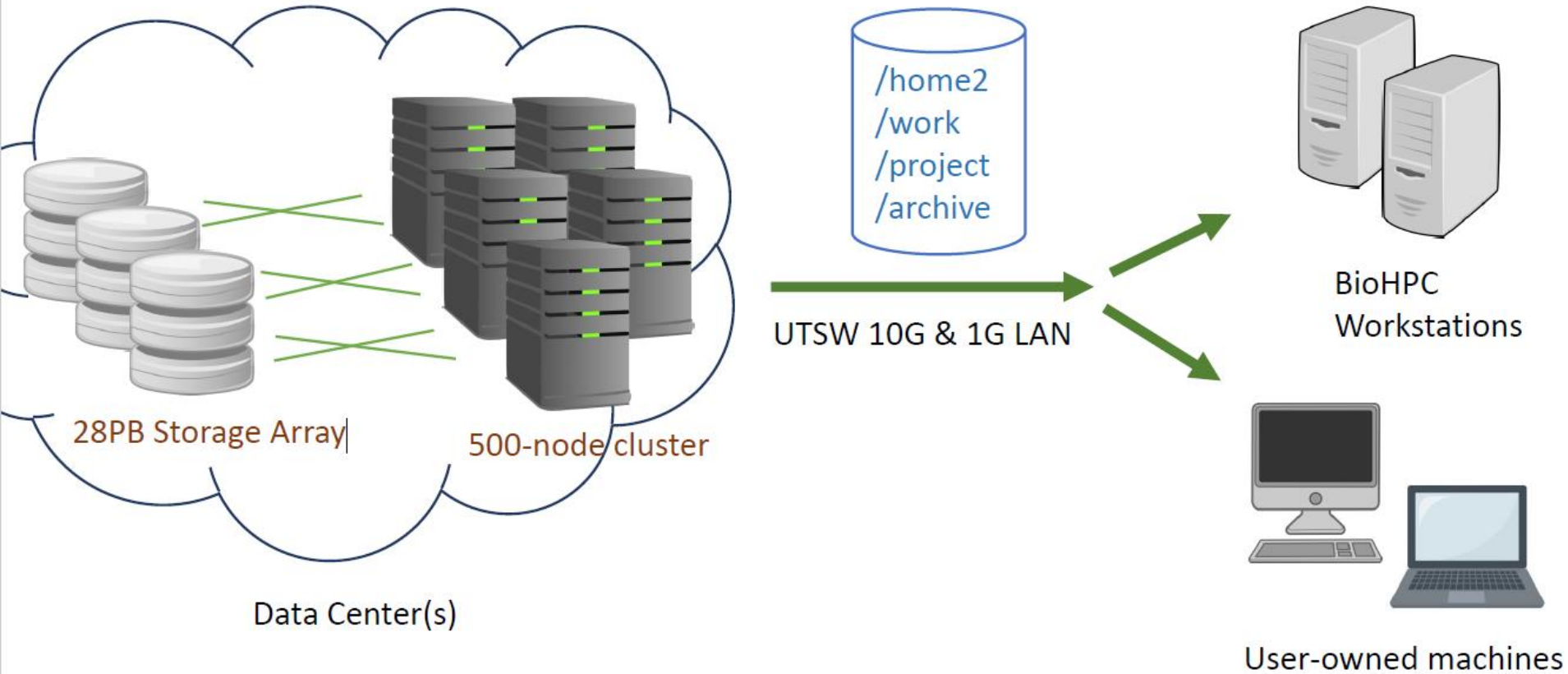[cloud]   https://cloud.biohpc.swmed.edu/

05-11-2022

Topics

Storage Resources & Intended Use

Tools for Data Access

Data Sharing Methods

Backup Policies & Best Practices

Centralized, Network Storage

28PB Storage Array

500-node cluster

Data Center(s)

/home2
/work
/project
/archive

UTSW 10G & 1G LAN

BioHPC Workstations

User-owned machines

Data Directories

/home2

- **50 GB** per <u>user</u>

User environment, small documents & light weight scripting

/home2/you

/work

- **5 TB** per <u>user</u>

Data processing

/work/dept./you

/project

- **5 TB** per <u>lab</u>

Data processing
Large files, sharing

/project/dept./shared

/project/dept./lab/shared

/project/dept./lab/you

/archive

- **5 TB** per <u>lab</u>

Long term storage, offload location

/archive/dept./shared

/archive/dept./lab/shared

/archive/dept./lab/you

Note: Exceptions for some OBI (O'Donnell Brain Institute) users, replacing **dept.** with **OBI/dept.**
2) Quota changes need get approval from Department.
3) Quota usage report: https://portal.biohpc.swmed.edu/content/, ssh nucleus005

UTSouthwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Access Methods

❖ **Direct browsing & editing**: Workstations/Thin Clients, Remote Desktop, Network mounting on local machines

❖ **Web browsing:** Lamella web interface; convenient for upload, download, & sharing internally

❖ **Specialized**: FTP servers for larger file transfers; Command Line Interface via SSH sessions

❖ **Sharing**: Lamella & Thunder FTP (internal), External FileshareServer <cloud.biohpc> & Flash FTP (external)

UTSouthwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Access via Linux Environment/Desktop

Linux GUI

- ▶ Remote Desktop

- ▶ Workstation / Thin Client
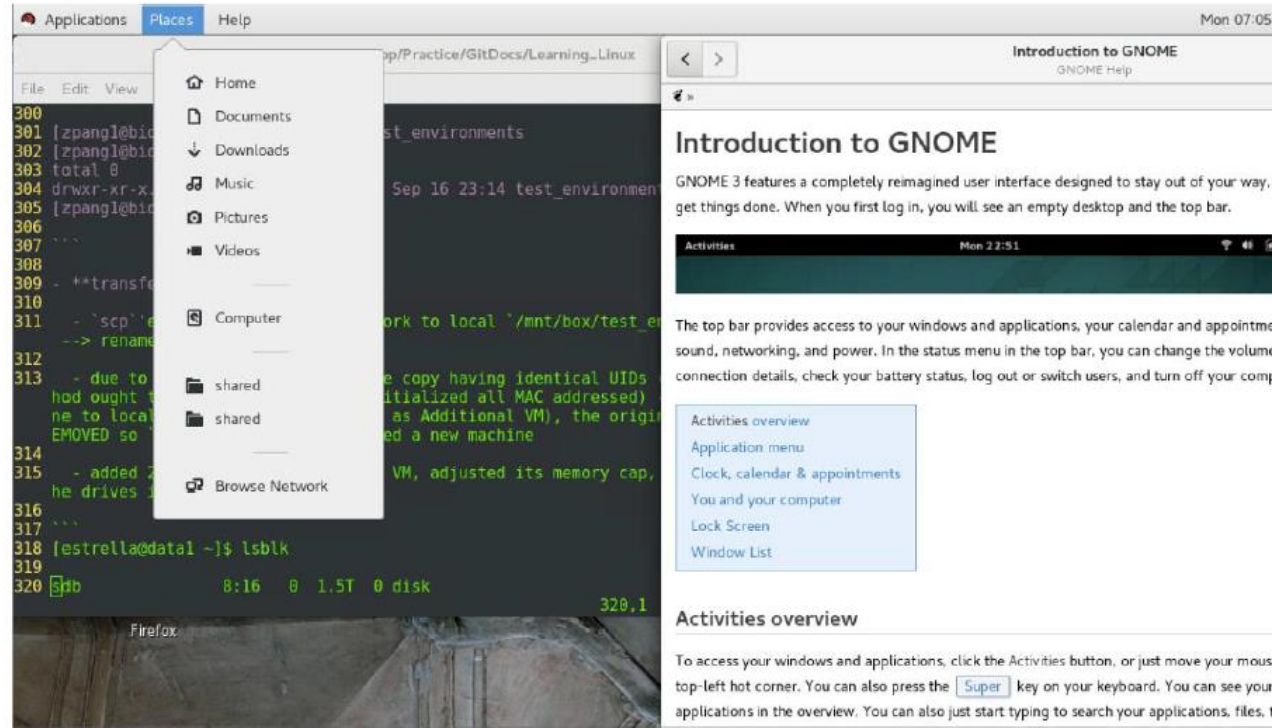
Access via Network Drives

Mapping BioHPC Storage Directories to your local environment:

▸ Via our load-balanced Samba filesharing server

    ▸ **PC:**

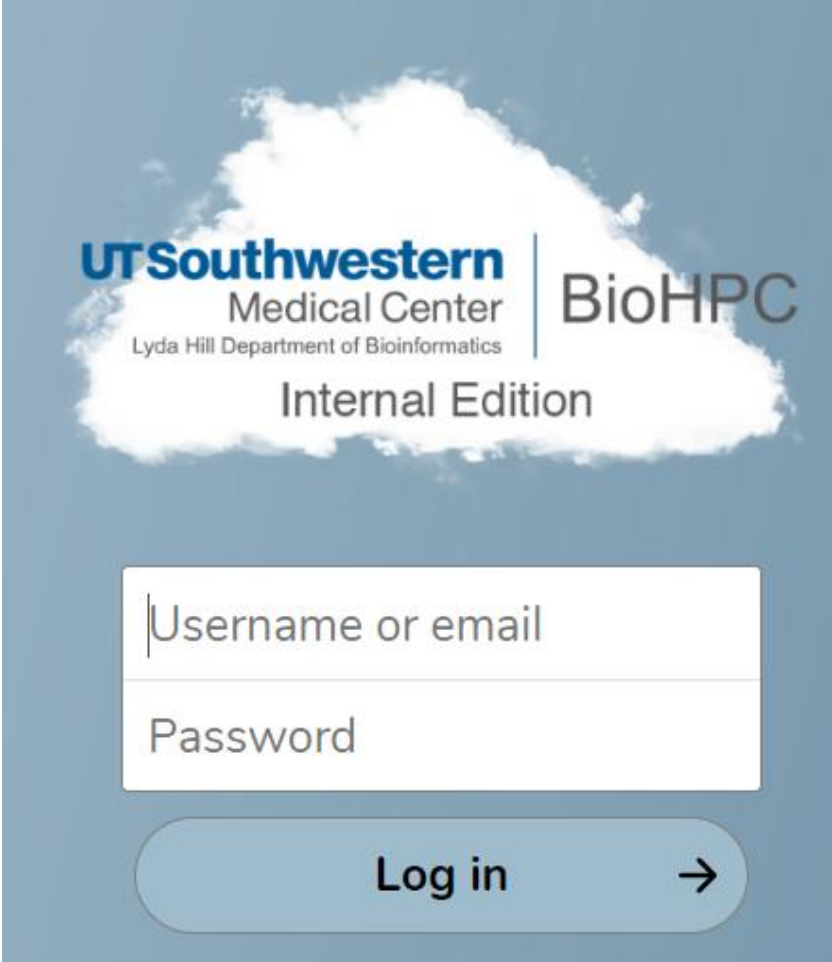| Computer -> Map Network Drive	
Folder is: \\lamella.biohpc.swmed.edu\username	home2
\\lamella.biohpc.swmed.edu\project\department\group	project
\\lamella.biohpc.swmed.edu\work\department\group	work

    ▸ **MAC:**

Finder -> Go -> Connect to Server	
Folder is: smb://lamella.biohpc.swmed.edu/username	home2
smb://lamella.biohpc.swmed.edu/project/department/group	project
smb://lamella.biohpc.swmed.edu/work/department/group	work

Access via Web Interface

❖ **Web Access**

- ## Lamella.biohpc.swmed.edu

  - Storage Gateway/Filesharing

  - **/home2** mounted by default ("Home")

  - **/project, /work, /archive** require additional configurations

Access via Web Interface

▸ **External Storage Mount:**
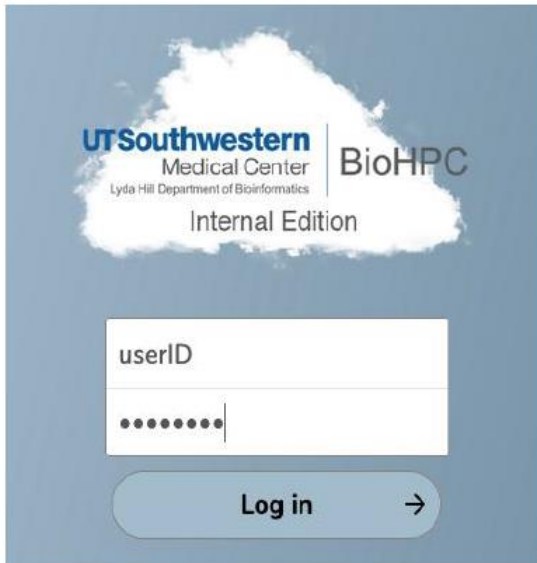
1.
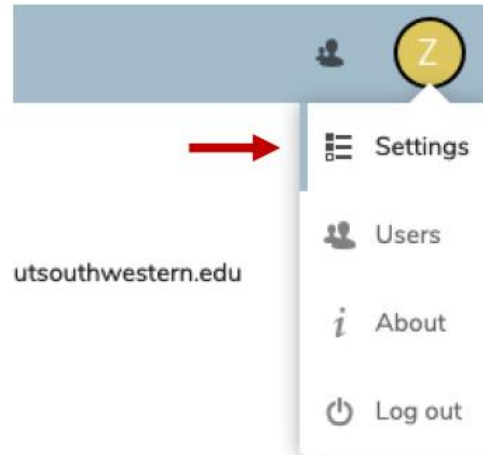
2.
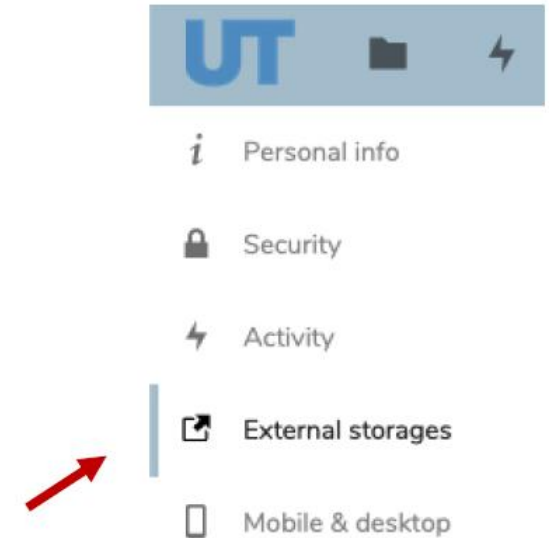
3.

Access via Web Interface

► **External Storage Mount:**

External storages

External storage enables you to mount external storage service
external storage services.

Folder name	External storage
home	BioHPC/Lysosome
Cloud	Cloud
<any name>	Add storage ▼
	BioHPC/Lysosome
	Cloud

4.

project

dept / lab / uid

uid

••••••••••

Username and password ▼   •••   ✓

5.

6.

Username and password ▼   •••

☑ Enable previews

☑ Enable sharing

7. (optional: select enable sharing is required to share files with others)

UTSouthwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Access via Web Interface

[s179389@Nucleus006 s179389]$ cat /project/biohpcadmin/s179389/greeting.txt
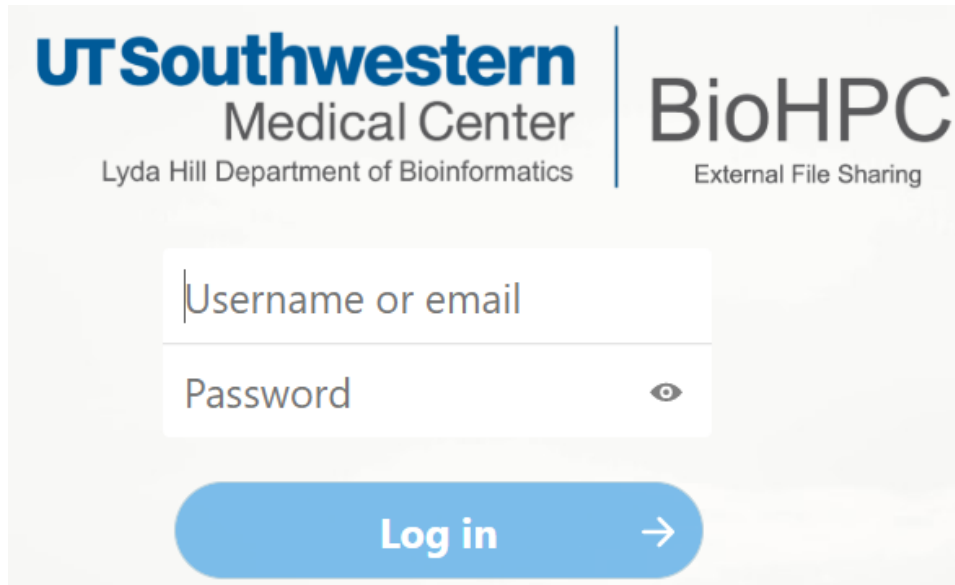Hello World and Good Morning!

🏠 〉 project ⌣ 〉 ⊕

☐ Name

☐ 📄 greeting.txt

https://lamella.biohpc.swmed.edu/index.php/
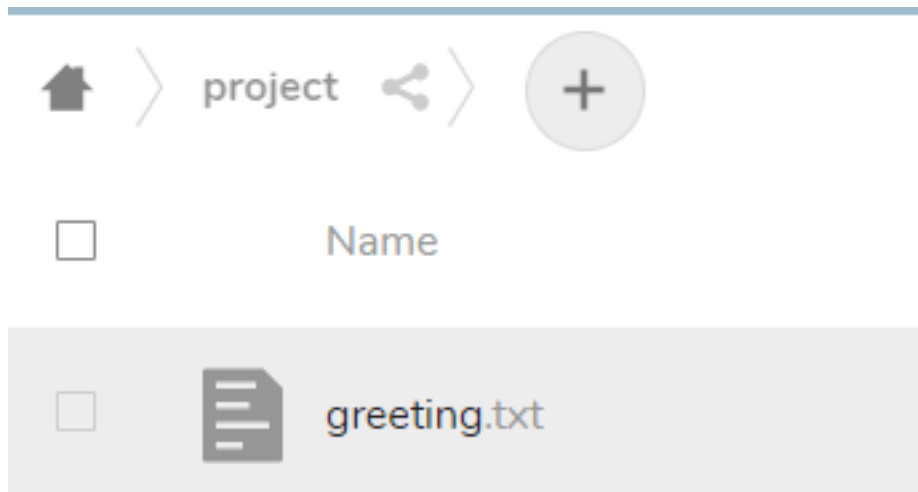
greeting.txt

1 Hello World and Good Morning!
2

UTSouthwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Access via Web Interface

❖ **Web Access**

- Cloud.biohpc.swmed.edu

  - External Filesharing

  - Interface similar to Lamella

  - separate cloud storage space

  - Can not share /home2, /project, /work, /archive directly

Access via Command Line

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

- Activate BioHPC account

- SSH client/Terminal Emulator

- BioHPC web terminal

- BioHPC workstation

- nucleus.biohpc.swmed.edu

- BioHPC webGUI/webGPU session

- BioHPC remote desktop

```
[s179389@Nucleus006 s179389]$ pwd
/project/biohpcadmin/s179389
[s179389@Nucleus006 s179389]$ cat greeting.txt
Hello World and Good Morning!
[s179389@Nucleus006 s179389]$ ls -ll greeting.txt
-rwxr--r-- 1 s179389 biohpc_admin 31 May  9 16:12 greeting.txt
[s179389@Nucleus006 s179389]$ ls /project
AIRC            backup_home2     biohpcadmin     CRI        flash_ftp_backup
apps            backup_services  bioinformatics  cryoem     GCRB
apps_database   backup_work      biophysics      DAPM       greencenter
apps_new        BICF             CAND            DPDS       hackathon
backup_archive  biochemistry     cellbiology     flash_ftp  HCTOR
[s179389@Nucleus006 s179389]$
```

Filesharing

Among BioHPC Users

- `` `../shared/..` `` directories

Non-BioHPC Collaborators

- Cannot access
  cluster components (Thunder, Lamella)

External Collaborators

- Blocked from UTSW internal network (Flash, Cloud)

© Bob Greenfield

UT Southwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Filesharing

Sharing Among BioHPC Users

▸ Sharing with other departments (made upon request)
  ▸ `/archive/shared`   &   `/project/shared`

▸ Sharing within your department:
  ▸ `/archive/<dept>/shared`   &   `/project/<dept>/shared`

▸ Sharing within your lab:
  ▸ `/archive/<dept>/<lab>/shared`   &   `/project/<dept>/<lab>/shared`

Filesharing

File Sharing Size	Internal Shares	External Shares
< 50-100GB	**Lamella**	**External File Exchange** (cloud.biohpc.swmed.edu)
> 100GBs – TBs	**Thunder FTP**	**Flash FTP**

UTSouthwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Filesharing

Sharing with Non-BioHPC Collaborators: < 100GB

lamella.biohpc.swmed.edu

- ▸ `home` (/home2/<you>) -> shareable!

- ▸ Mapped `/project` & `/work` -> manually enable sharing

- ▸ Look for the [<] icon

- ▸ Instruct collaborator to access via web URL

Filesharing

Sharing with Non-BioHPC Collaborators: < 100GB

Filesharing

Sharing with Non-BioHPC Collaborators: < 100GB

- Grant appropriate permissions!
  - Stick to principle of least privilege

- Password protect!

- Set Expiration date!

My_Scripts
< 1 KB, 22 days ago   Tags

Activity   Comments   Sharing

*Copy link to clipboard

Share link

Add to a pr...
Connect item

- Read only
- Allow upload and editing
- File drop (upload only)
- Hide download
- Password protect
  - Choose a password for →
- Set expiration date
  - 22-10-2019
- Note to recipient
- Delete share link
- Add another link

UTSouthwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Filesharing

Sharing with Non-BioHPC Collaborators: > 100GB
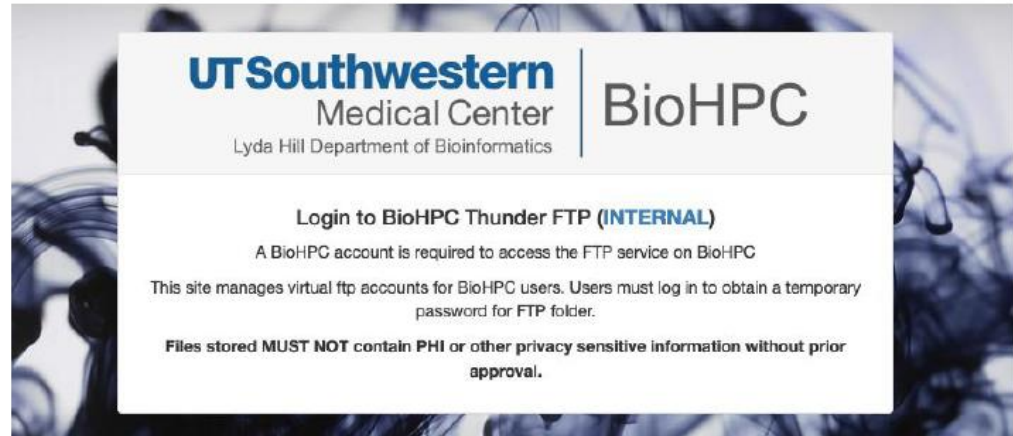
thunder.biohpc.swmed.edu

▸ `/project/thunder_ftp/<uid>`
  -> share folder

▸ Upload limit = **/project quota***

▸ If size exceeds quota -> Contact us!

▸ Access via **web** –or– **FileZilla** (if collaborator wishes to upload)

UT Southwestern Medical Center | **BioHPC**
Lyda Hill Department of Bioinformatics

Login to BioHPC Thunder FTP (INTERNAL)

A BioHPC account is required to access the FTP service on BioHPC

This site manages virtual ftp accounts for BioHPC users. Users must log in to obtain a temporary password for FTP folder.

Files stored MUST NOT contain PHI or other privacy sensitive information without prior approval.

UT Southwestern Medical Center | **BioHPC**
Lyda Hill Department of Bioinformatics

Filesharing

Sharing with Non-BioHPC Collaborators: > 100GB

UT Southwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Filesharing

Sharing with Non-BioHPC Collaborators: > 100GB

BioHPC user:

zpang1 has shared his BioHPC ftp folder with you with a message:

UTSW employee: zpang1 has shared his ftp folder with you, please click on below link to view your ftp credentials

please click the link below to view your temporary message:

https://thunder.biohpc.swmed.edu/manage/view_guest_account
/?q=9297aa00f02f82c464c47b2d5e16673b86146c691b4f442facf666e0aeeb8c4a153db883fd05b9a6f1ad04c716c946c5

Guest Account Session:

Thanks for using BioHPC, your temporary account and password are as follows, please save them, this account automatically expires in three days:

USERNAME: guest94202Szhi

PASSWORD: zf5Eh7b3HtFV

URL: ftp://guest94202Szhi:zf5Eh7b3HtFV@thunder.biohpc.swmed.edu

UTSouthwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Filesharing

Sharing with Non-BioHPC Collaborators: > 100GB

Filesharing

Sharing with <u>External</u> Collaborators:

< 50GB -> **BioHPC External File Exchange**

cloud.biohpc.swmed.edu

> 50GB -> **Flash FTP Service**

flash.biohpc.swmed.edu

▸ **External Equivalents of Lamella & Thunder (internal)**

   ▸ **Identical operations**
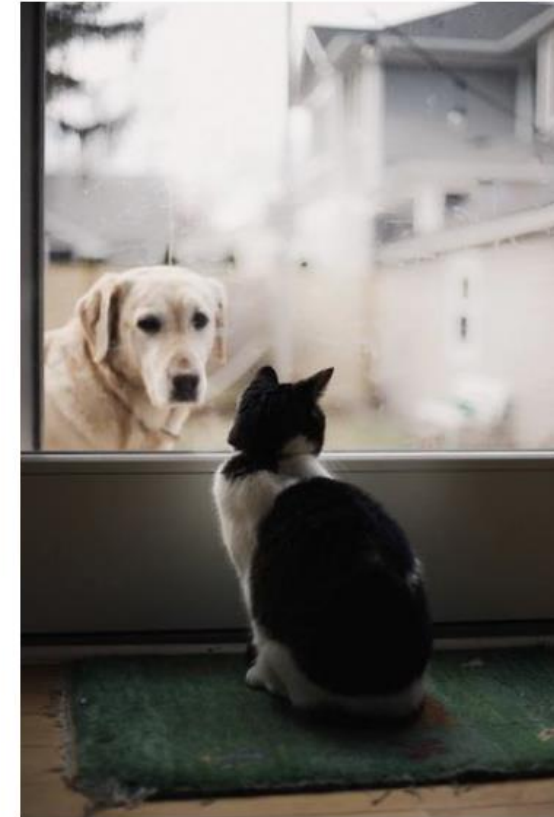
Backup

Mirror – Not backup!

- No incremental backups by default, only present states are synced

- Used for system restore points

- **/home2 & /work** mirrored weekly

- **/project & /archive** OPTIONAL for incremental backup, upon PI's request

Data Hall Migration

- Most of BioHPC nodes, storage will be moved to new location as school planed.

- All BioHPC storage backup weekly before data hall migration finished 4th quarter 2022.

- New Storage ~20 PB in plan.

UTSouthwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

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.

UT Southwestern
Medical Center
Lyda Hill Department of Bioinformatics
BioHPC