UT Southwestern Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

# Cluster Storage Overview

*01-20-2021*

# Topics



***Storage Resources & Intended Use***

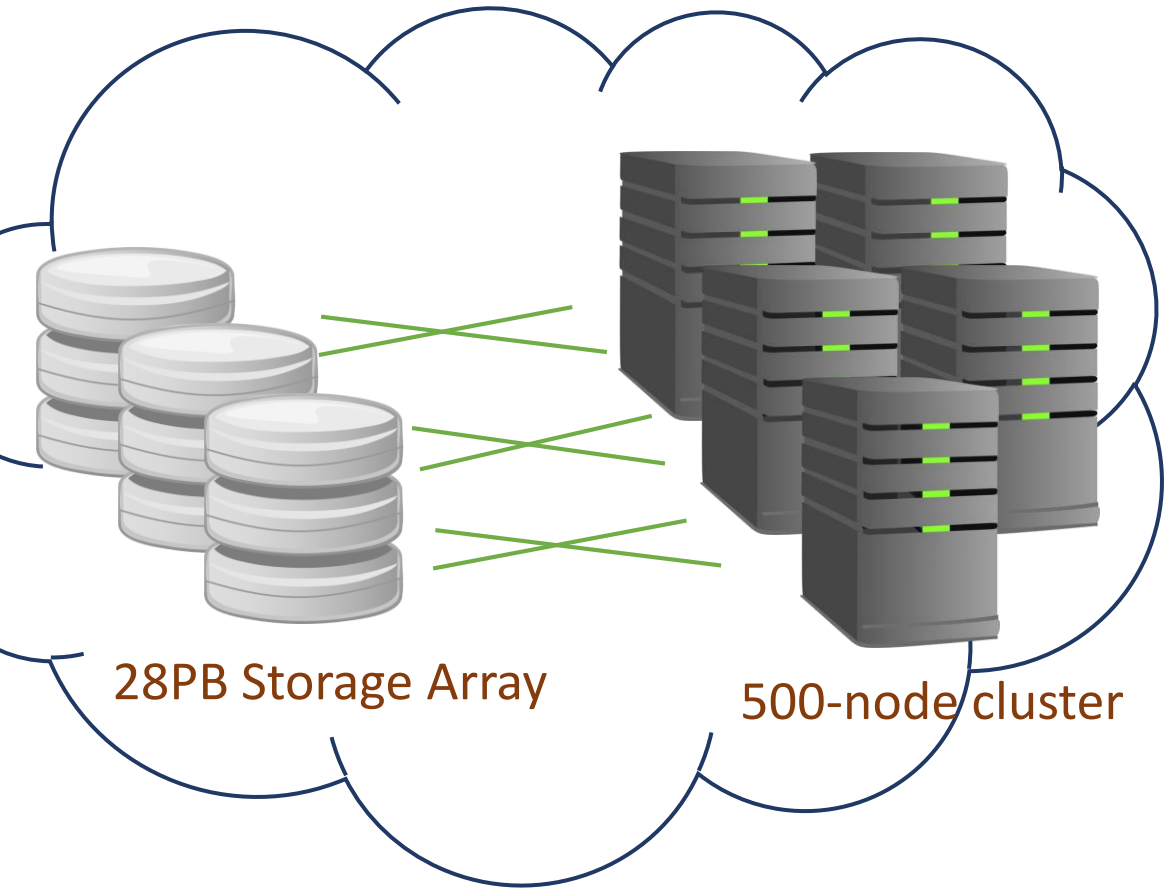
***Tools for Data Access***

***Data Sharing Methods***

***Backup Policies & Best Practices***

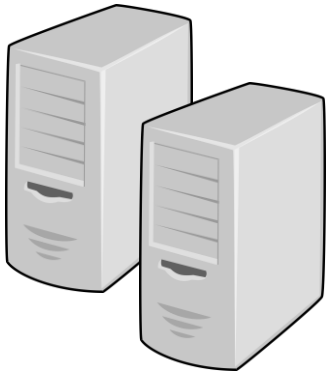
# Centralized, Network Storage






Data Center(s)

UTSW 10G & 1G LAN



BioHPC Workstations

User-owned machines

# Data Directories



## /home2

- **50 GB** per **user**

User environment, small documents & light weight scripting

*/home2/you*

## /work

- **5 TB** per **user**

Data processing

*/work/dept./you*

## /project

- **5 TB** per **lab**

Data processing
Large files, sharing

*/project/dept./shared*

*/project/dept./lab/shared*

*/project/dept./lab/you*

## /archive

- **5 TB** per **lab**

Long term storage, offload location

*/archive/dept./shared*

*/archive/dept./lab/shared*

*/archive/dept./lab/you*

# Access Methods



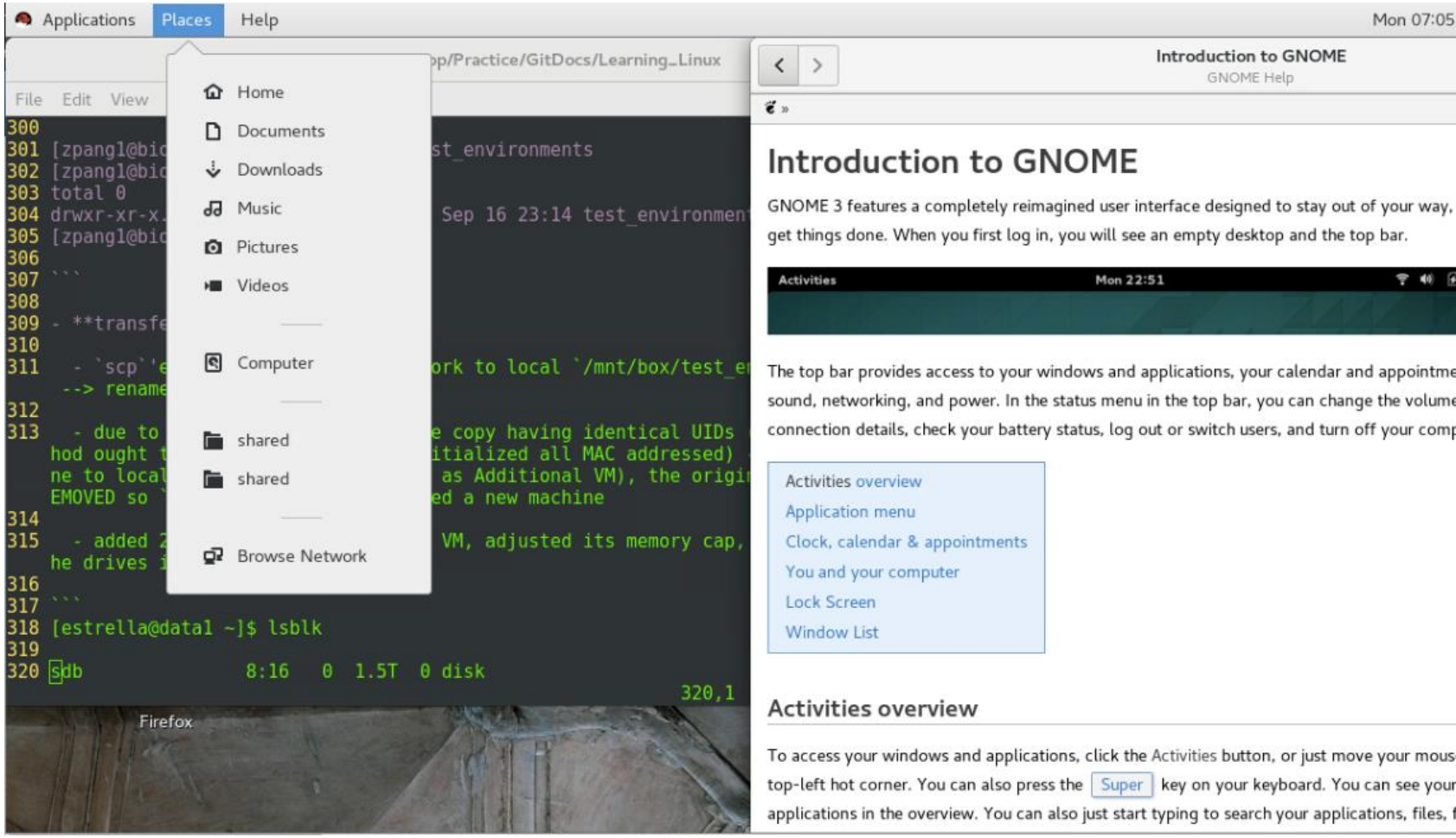
- **Direct browsing & editing**: Workstations/Thin Clients, Remote Desktop, Network mounting on local machines
- **Web browsing:** Lamella web interface; convenient for upload, download, & sharing internally
- **Specialized**: FTP servers for larger file transfers; Command Line Interface via SSH sessions
- **Sharing**: Lamella & Thunder FTP (internal), External Fileshare Server & Flash FTP (external)

# Access via Linux Environment



## 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

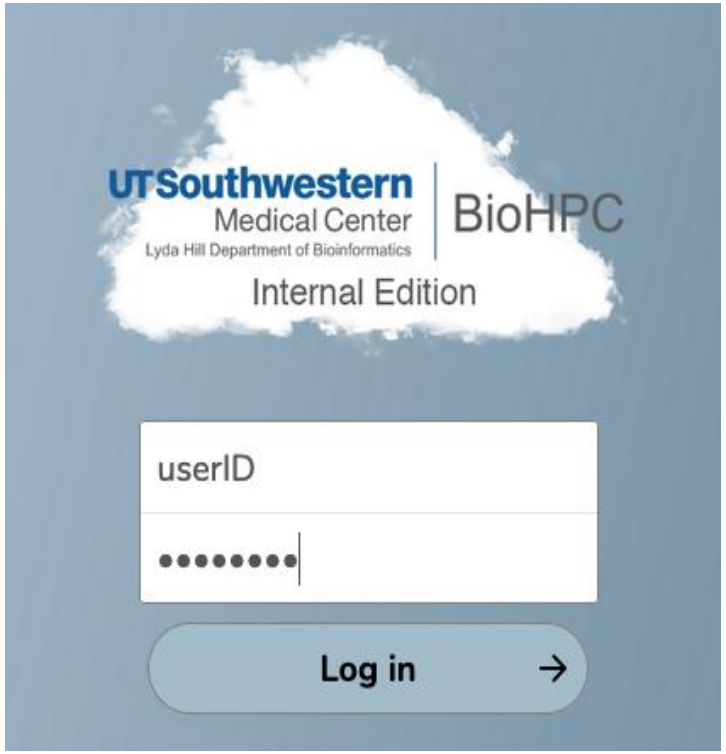


- **Web Access**
- **lamella.biohpc.swmed.edu**
  - Storage Gateway/Filesharing
  - **/home2** mounted by default ("Home")
  - **/project, /work, /archive** require additional configurations

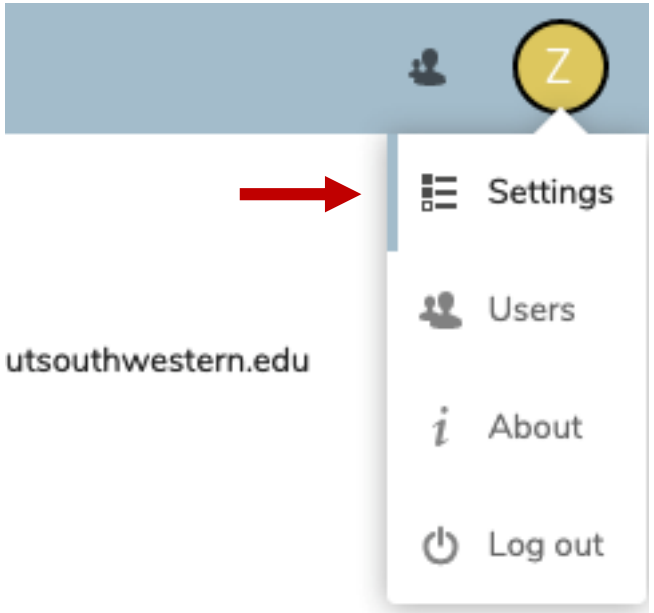
# Access via Web Interface

UTSouthwestern Medical Center | BioHPC
Lyda Hill Department of Bioinformatics

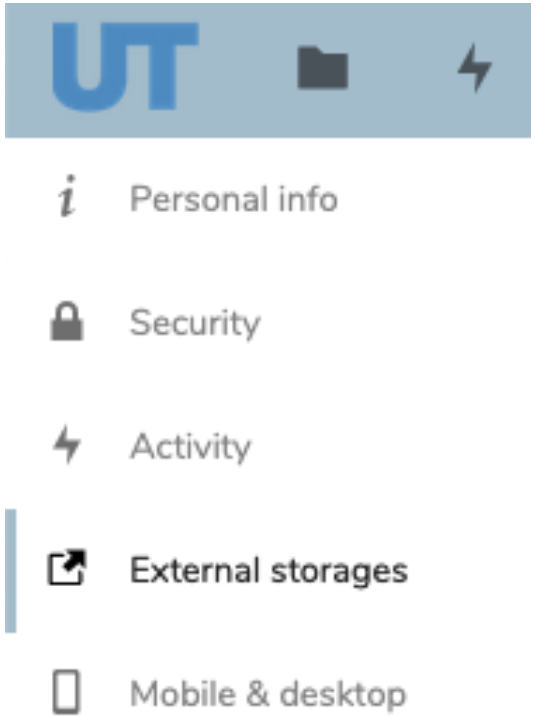
- **External Storage Mount:**




1.




2.

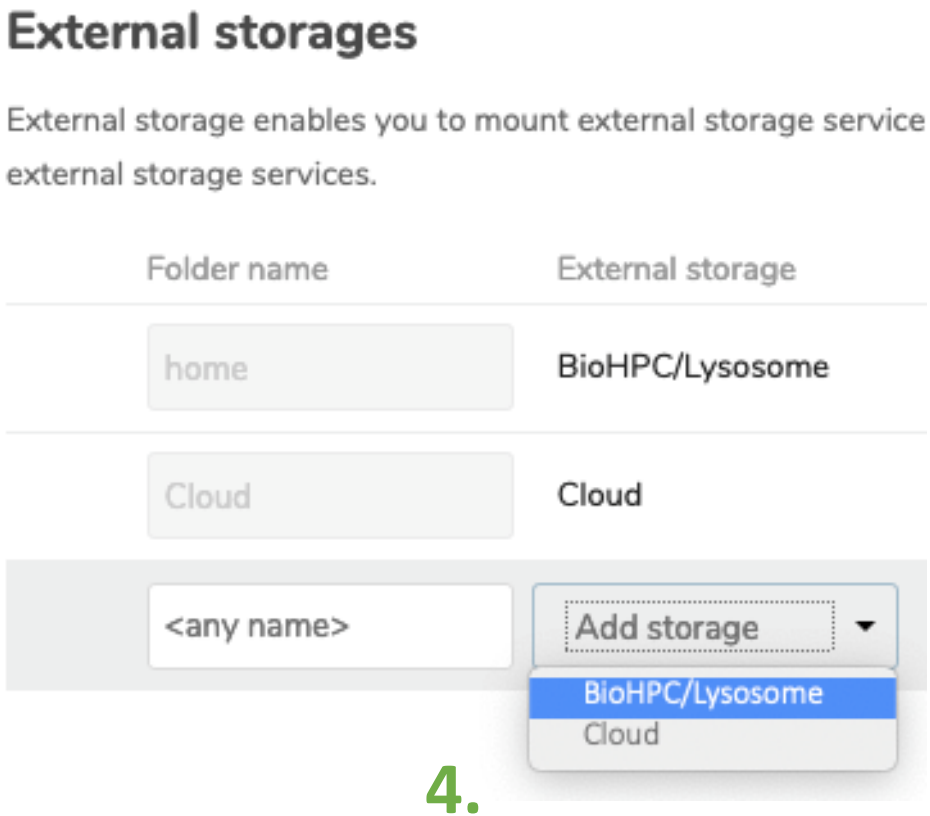


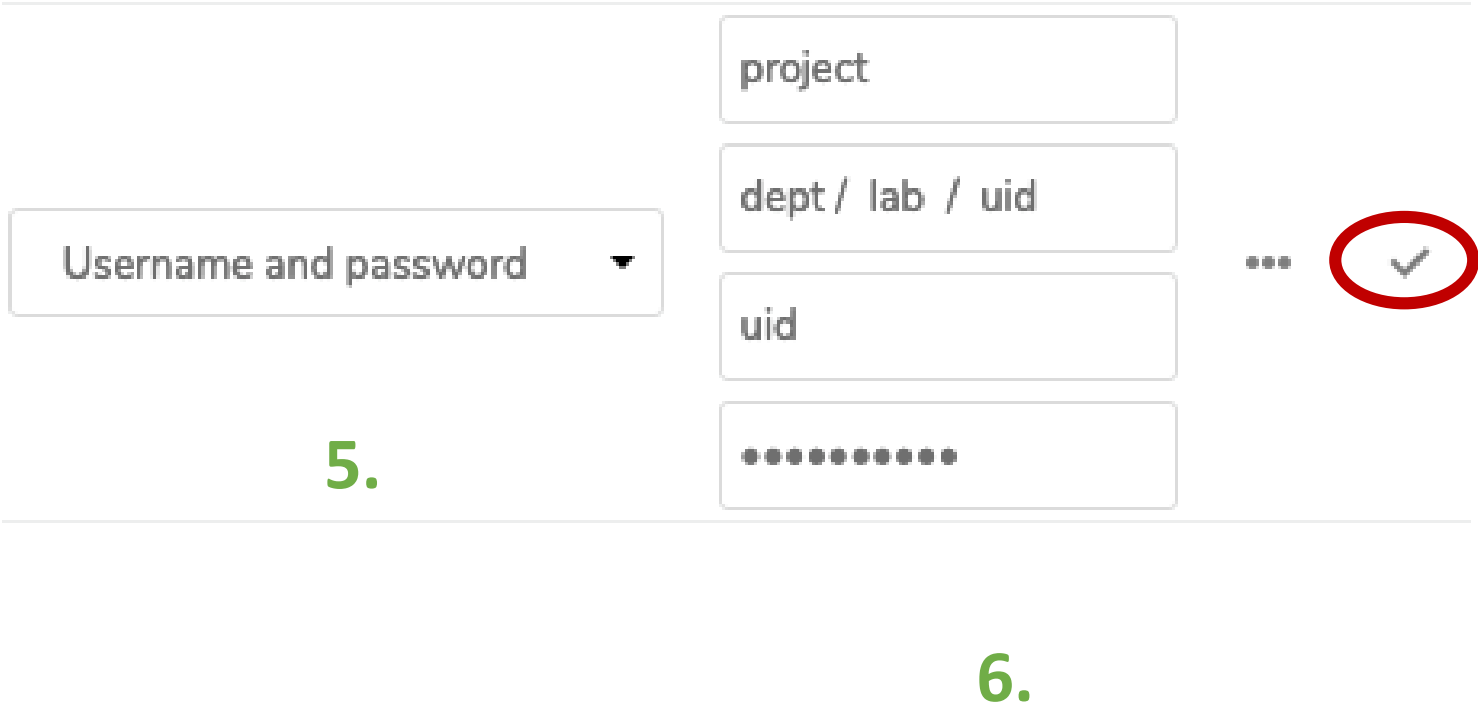

3.

# Access via Web Interface



- **External Storage Mount:**



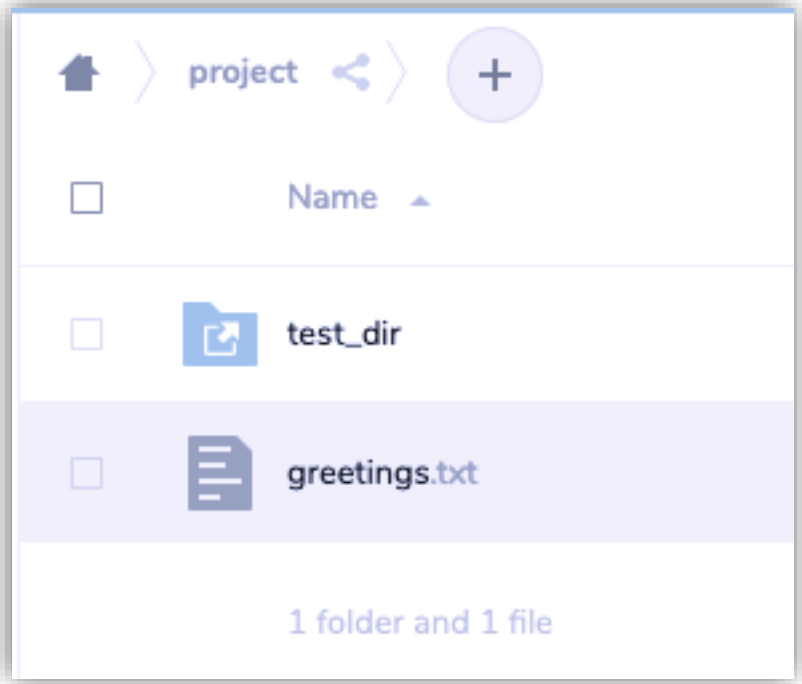



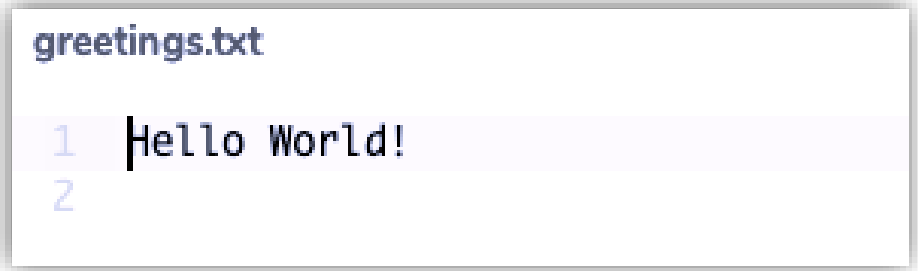
# Access via Web Interface



- **External Storage Mount:**

```
[zpang1@Nucleus006 ~]$ ls /project/biohpcadmin/zpang1/
greetings.txt  test_dir
[zpang1@Nucleus006 ~]$
[zpang1@Nucleus006 ~]$
[zpang1@Nucleus006 ~]$ cat /project/biohpcadmin/zpang1/greetings.txt
Hello World!
```



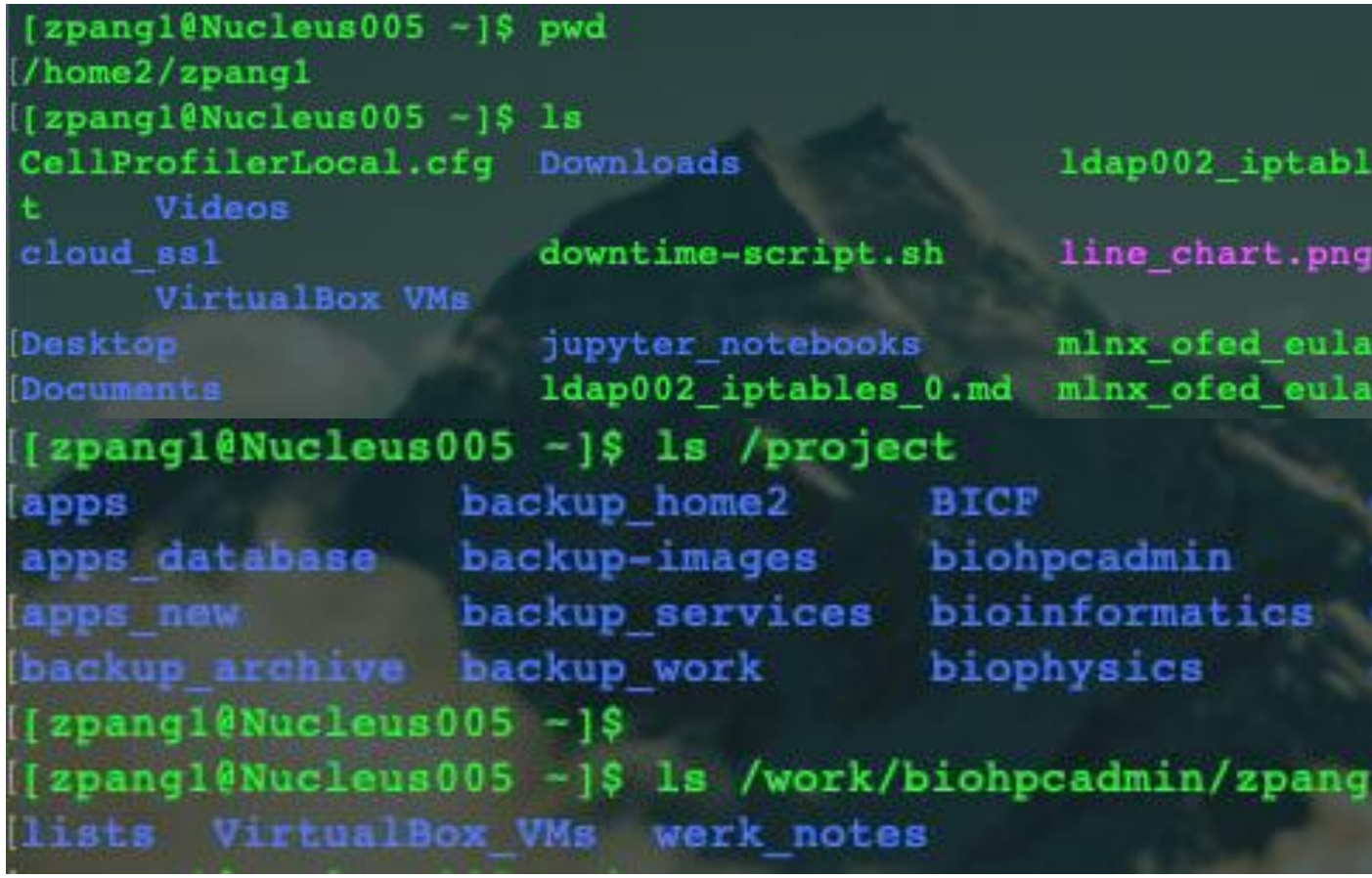



# Access via Command Line



```
ssh <your uid>@nucleus
```

- Nucleus.biohpc.swmed.edu
- Active BioHPC account
- SSH client/Terminal Emulator



## Among BioHPC Users

- `../shared/..` directories

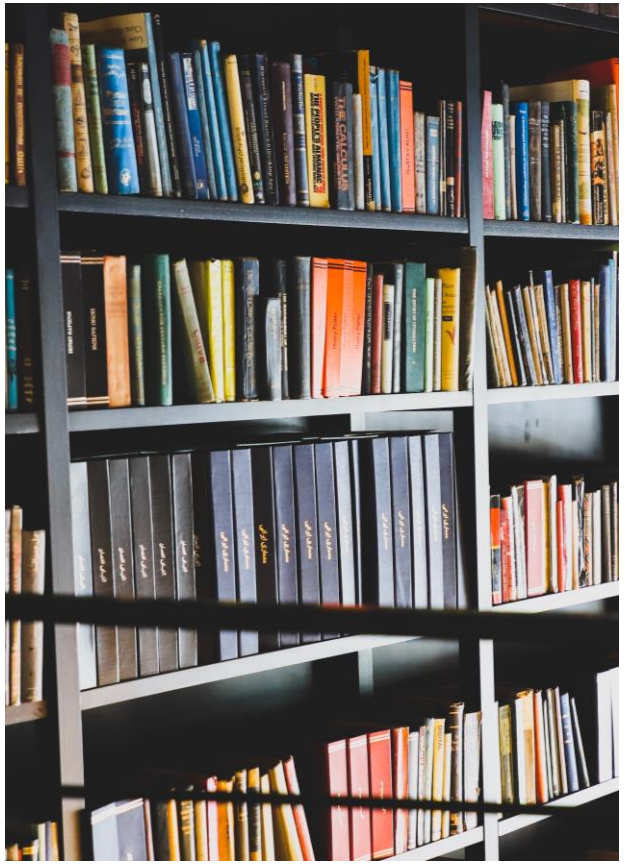
## Non-BioHPC Collaborators

- Cannot access
  cluster components (Thunder, Lamella)

## External Collaborators

- Blocked from UTSW internal network (Flash, Cloud)

# 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



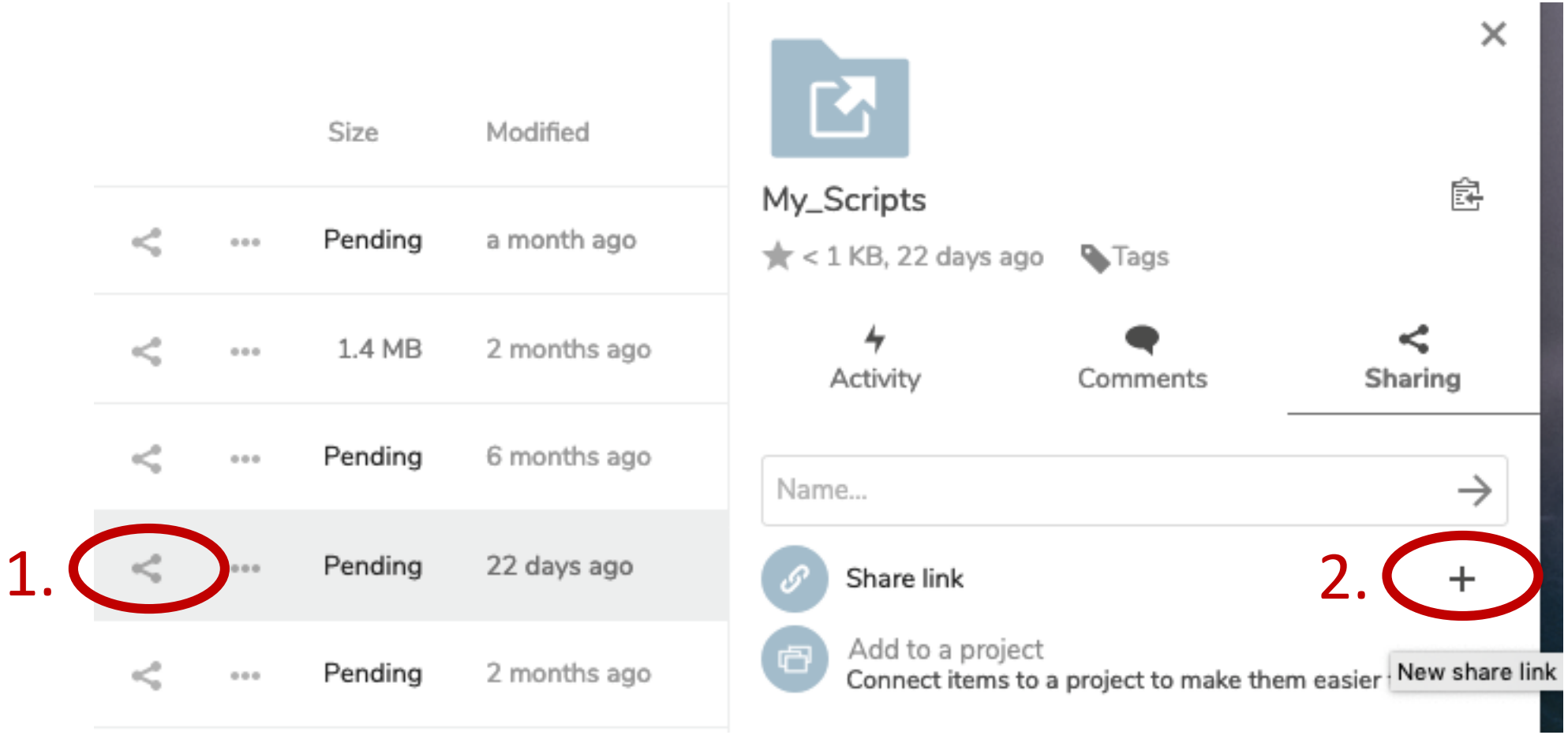
| File Sharing Size | Internal Shares | External Shares |
|---|---|---|
| < 50-100GB | **Lamella** | **External File Exchange** (cloud.biohpc.swmed.edu) |
| > 100GBs – TBs | **Thunder FTP** | **Flash FTP** |

# Sharing with Non-BioHPC Collaborators: < 100GB

**lamella.biohpc.swmed.edu**

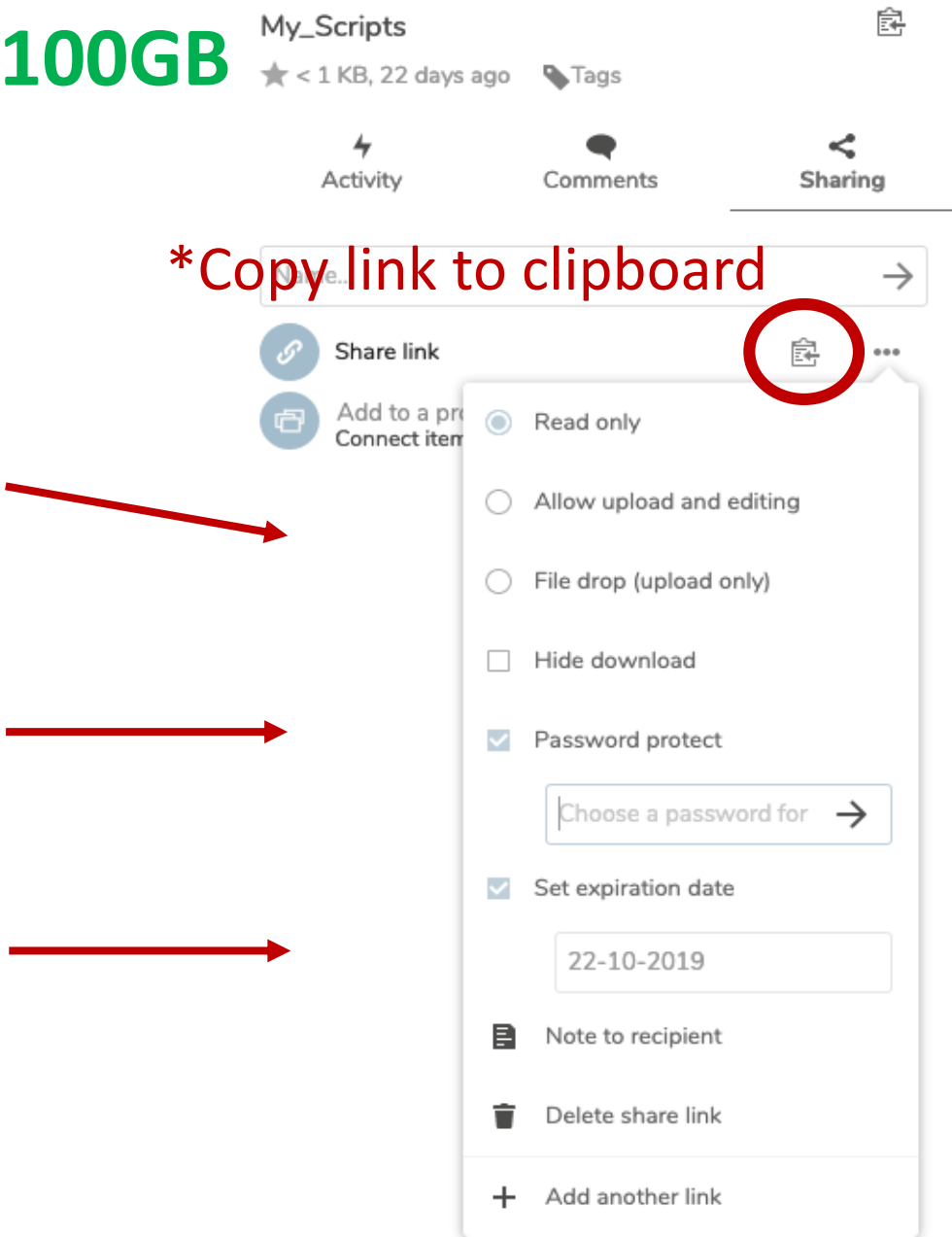
- `home` (/home2/<you>) -> shareable!
- Mapped `/project` & `/work` -> manually enable sharing
- Look for the icon
- Instruct collaborator to access via web URL

# Sharing with Non-BioHPC Collaborators: < 100GB



# Sharing with Non-BioHPC Collaborators: < 100GB

- Grant appropriate permissions!
  - Stick to principle of least privilege
- Password protect!
- Set Expiration date!

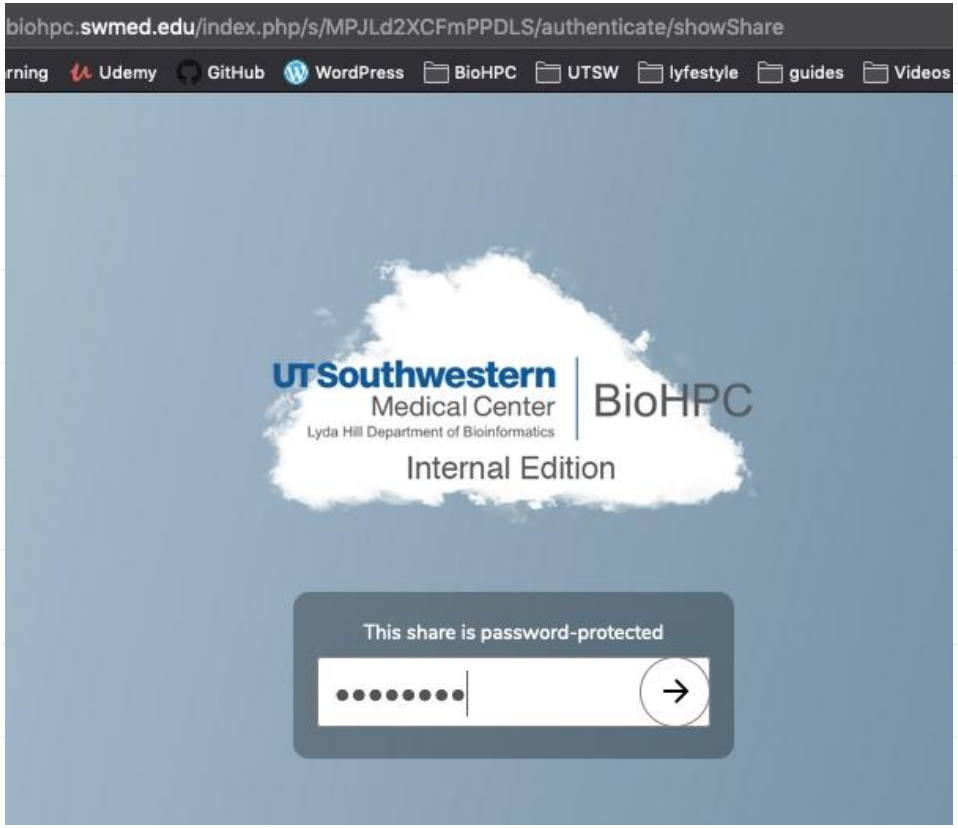


# Sharing with Non-BioHPC Collaborators: < 100GB

My_Scripts shared by Zengxing Pang (zpang1) — Download all files

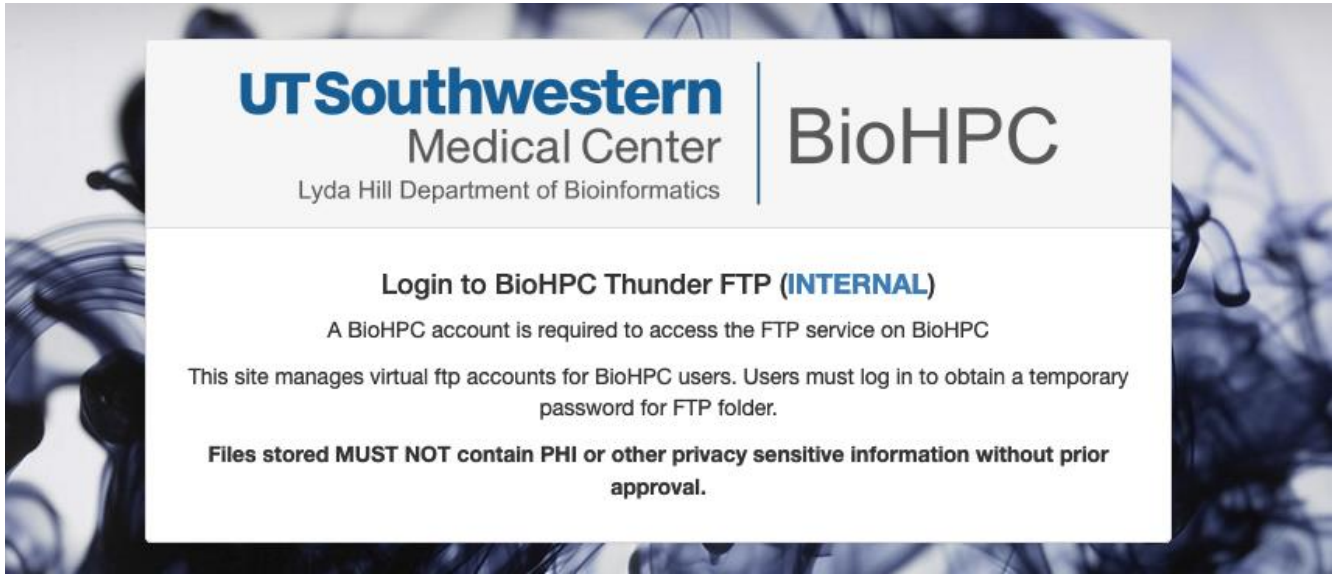
*Upload Files

| Name |
| --- |
| system_report.sh |
| BLOWFISH |
| blowfish.h2 |
| blowfish.h |
| blowfish.cpp |
| 14_DiskUse.sh |
| 13th_Bash.sh |
| 12th_Bash.sh |



# 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)

# Sharing with Non-BioHPC Collaborators: > 100GB

**Guests: Create and manage guest FTP acounts**

Share with UTSW collaborator: +

Add new Guest

You currently have not allowed any guest accounts to access your ftp folde
form.

**Guests: Create and manage guest FTP acounts**

Share with UTSW collaborator: +

Email address *

zengxing.pang@utsouthwestern.edu

Note

Note to Self

Directory Name*

another_test

Email Content

Email Content

Cancel Submit

guest created successfully

| Guestname | Expiration | Email | Note | directory | Actions |
|---|---|---|---|---|---|
| guest94202Szhi | Oct. 19, 2019, 7:53 a.m. | zengxing.pang@utsouthwestern.edu | | /project/thunder_ftp /zpang1 /another_test | Extend Delete Resend |

# 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:

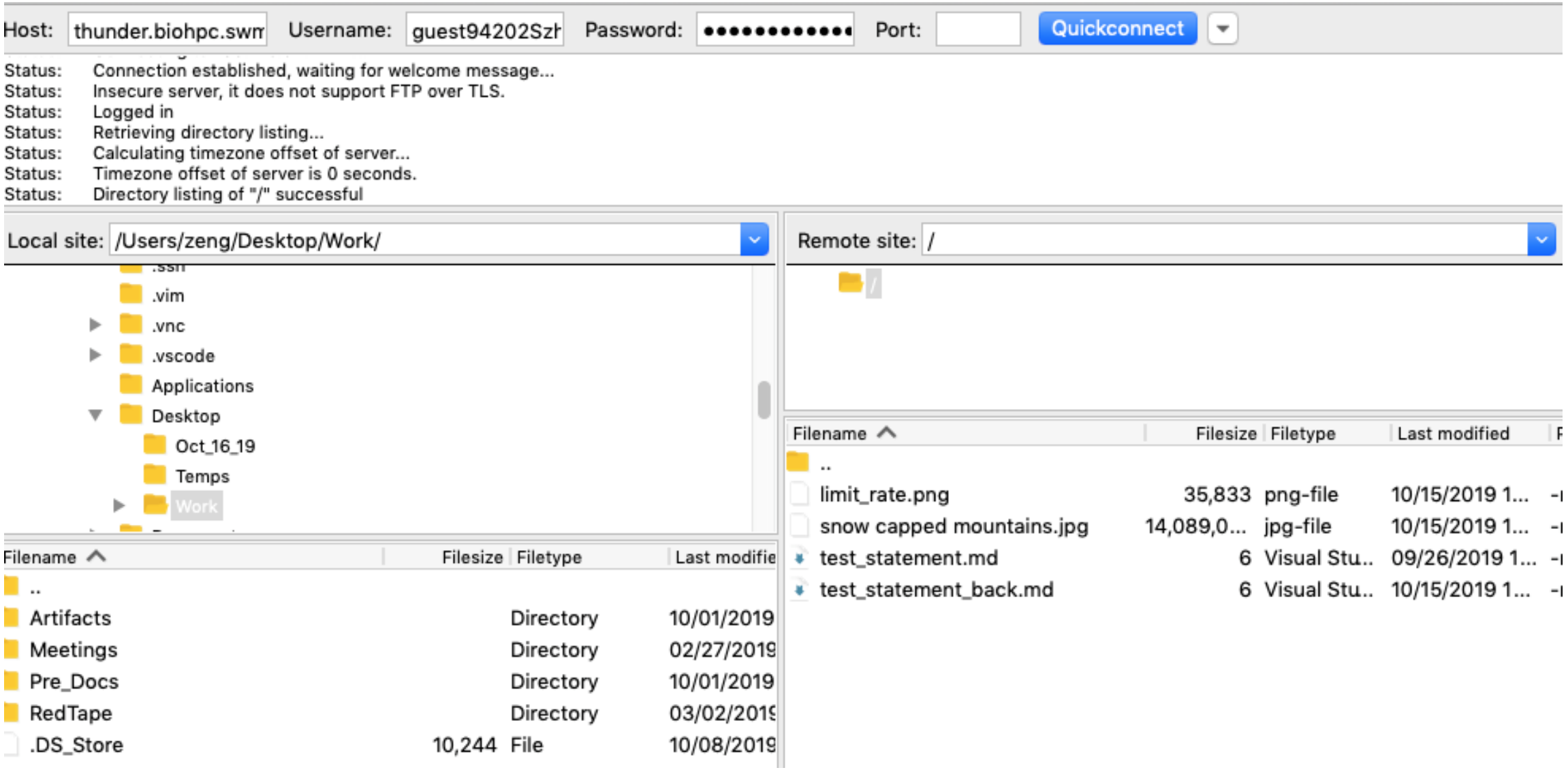
Thanks for using BioHPC, your temporary account and password are as follows, please save them, this account automatically expires in three days:

USERNAME: **guest94202Szhi**

PASSWORD: **zf5Eh7b3HtFV**

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

# Sharing with Non-BioHPC Collaborators: > 100GB



# Sharing with External Collaborators:

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

cloud.biohpc.swmed.edu

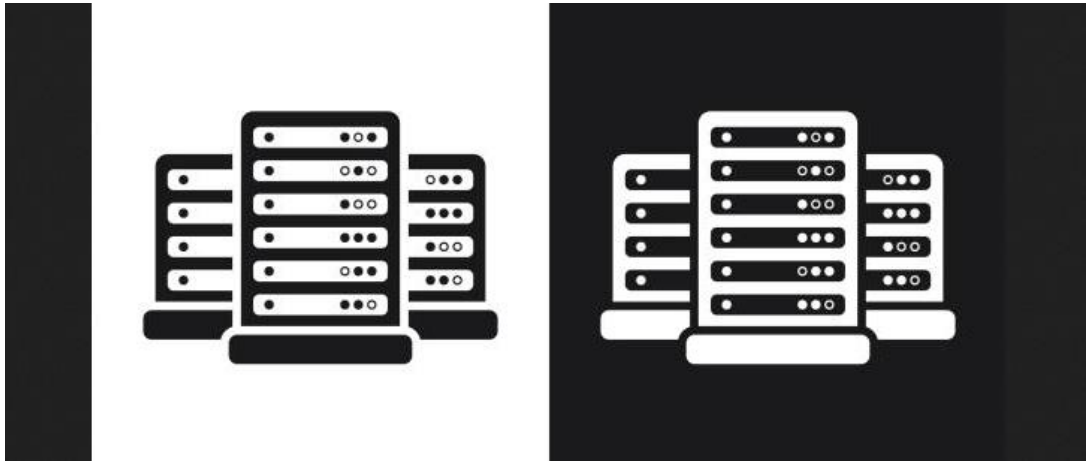
**> 50GB -> Flash FTP Service**

flash.biohpc.swmed.edu

- **External Equivalents of Lamella & Thunder (internal)**
  - **Identical operations**

# 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



# Storage Expansion – Spring 2021

- **+16 PB** with new IBM ES5000 Storage system
- **/clinical**: 1 PB of encrypted storage for PHI data
- Off-site replication of data, 24/7 uptime -- even during cluster maintenance

UT Southwestern Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

# Thank You!

*biohpc-help@utsouthwestern.edu*

*01-20-2021*