

UT Southwestern

Medical Center

Lyda Hill Department of Bioinformatics

BioHPC

Cluster Storage Overview

01-20-2021

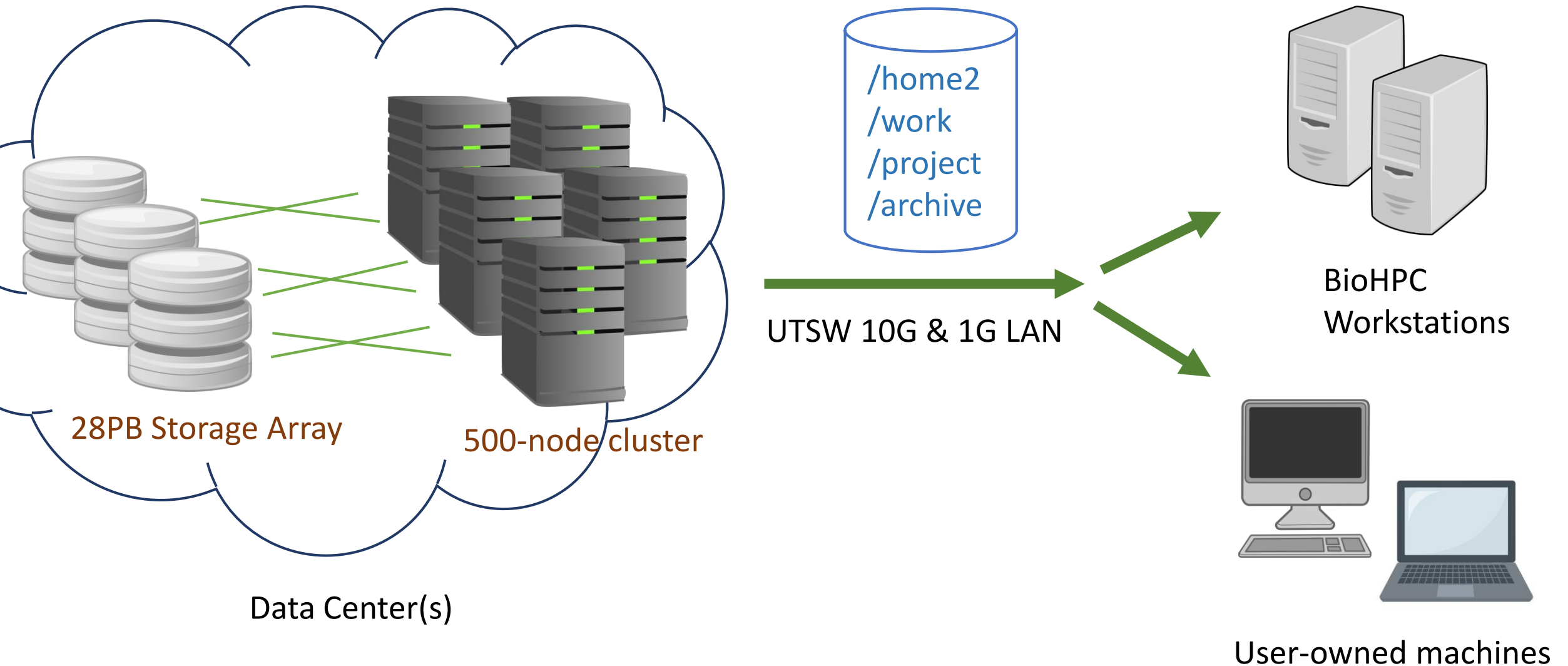
Storage Resources & Intended Use

Tools for Data Access

Data Sharing Methods

Backup Policies & Best Practices

Centralized, Network Storage



/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

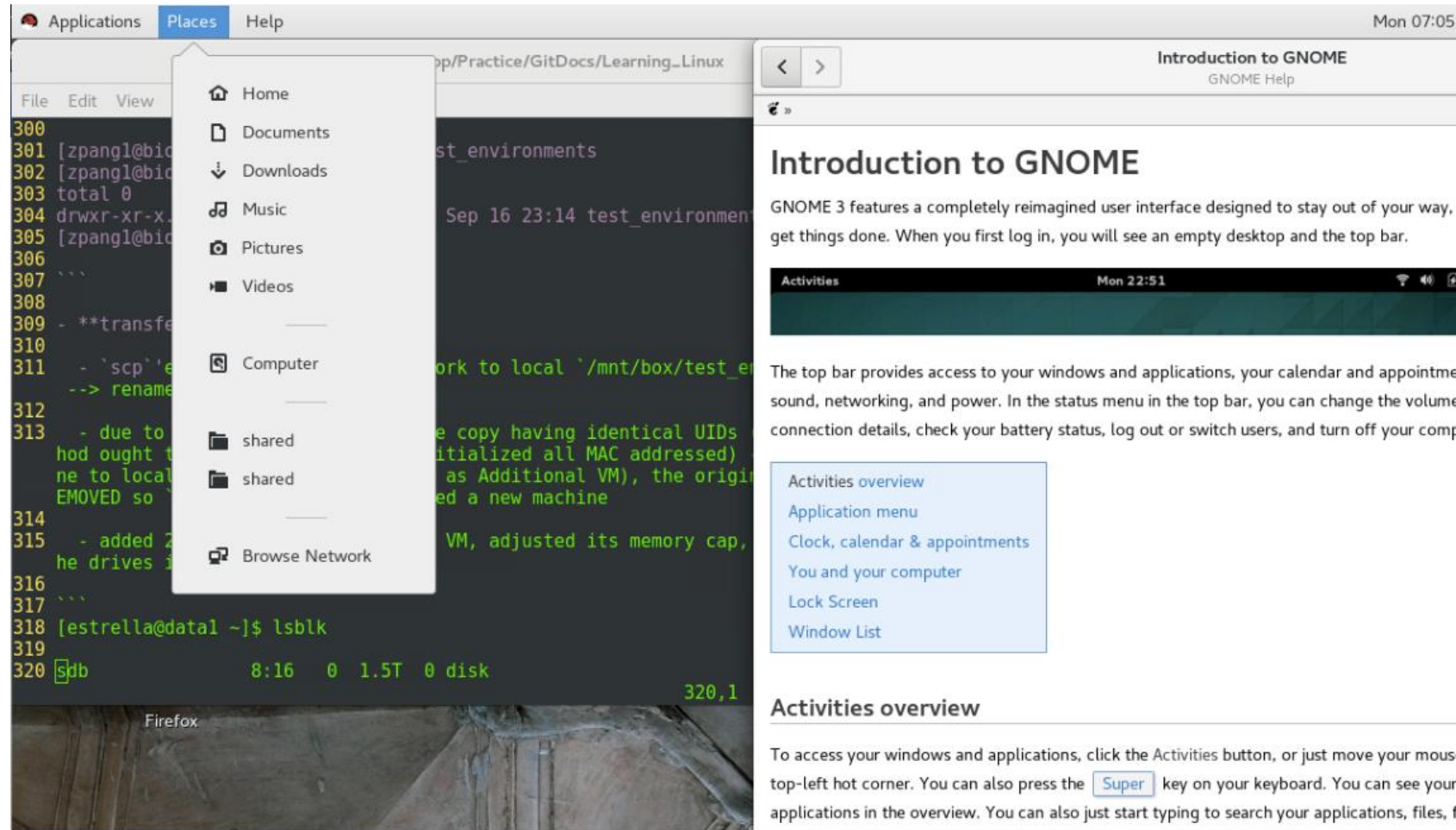
/archive/dept./lab/shared

/archive/dept./lab/you

- ▶ **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)

Linux GUI

- ▶ Remote Desktop
- ▶ Workstation / Thin Client



Mapping BioHPC Storage Directories to your local environment:

- ▶ Via our load-balanced Samba filesharing server

- ▶ **PC:**

Computer -> Map Network Drive

| | | |
|------------|------------------------------------------------------------------|----------------|
| Folder is: | <code>\\lamella.biohpc.swmed.edu\username</code> | <i>home2</i> |
| | <code>\\lamella.biohpc.swmed.edu\project\department\group</code> | <i>project</i> |
| | <code>\\lamella.biohpc.swmed.edu\work\department\group</code> | <i>work</i> |

- ▶ **MAC:**

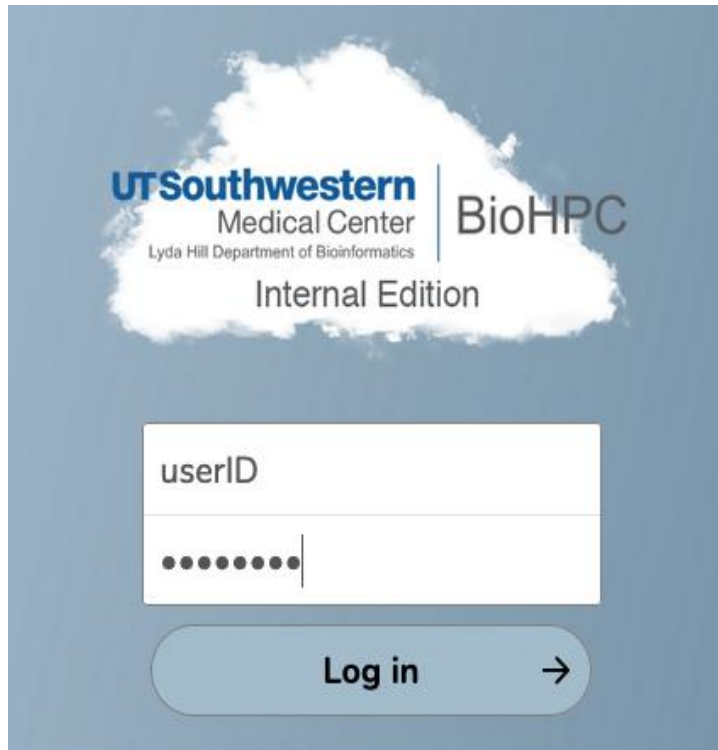
Finder -> Go -> Connect to Server

| | | |
|------------|----------------------------------------------------------------------|----------------|
| Folder is: | <code>smb://lamella.biohpc.swmed.edu/username</code> | <i>home2</i> |
| | <code>smb://lamella.biohpc.swmed.edu/project/department/group</code> | <i>project</i> |
| | <code>smb://lamella.biohpc.swmed.edu/work/department/group</code> | <i>work</i> |

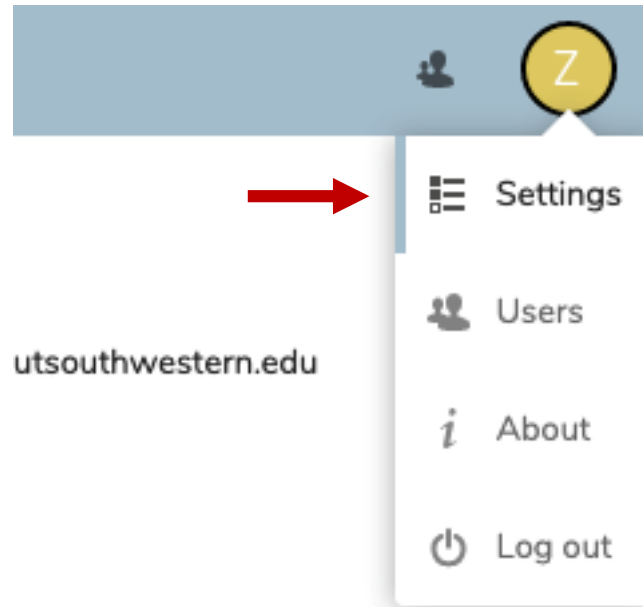
- ▶ **Web Access**
- ▶ **lamella.biohpc.swmed.edu**
 - ▶ Storage Gateway/Filessharing
 - ▶ **/home2** mounted by default (“Home”)
 - ▶ **/project, /work, /archive** require additional configurations



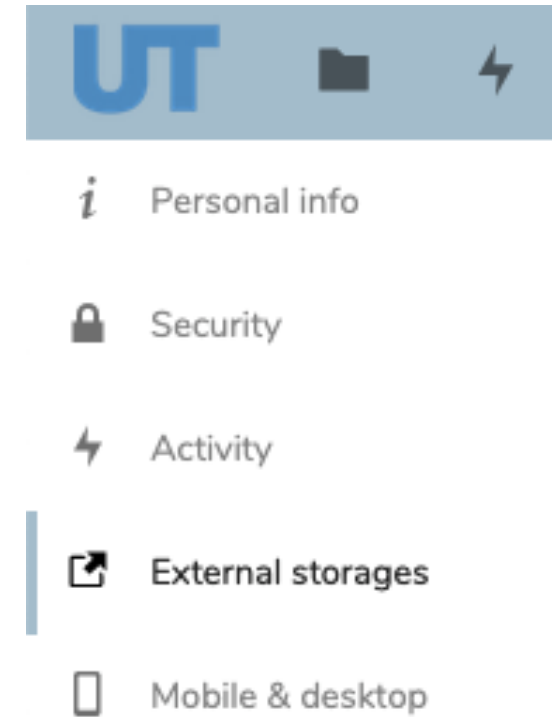
► External Storage Mount:



1.



2.



3.

► 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 | ... |
| Username and password | dept / lab / uid |
| | uid |
| | |

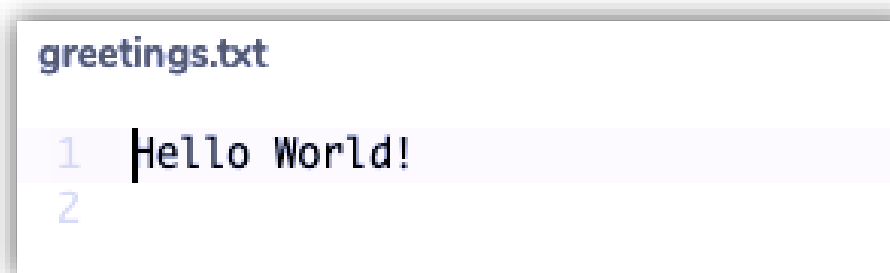
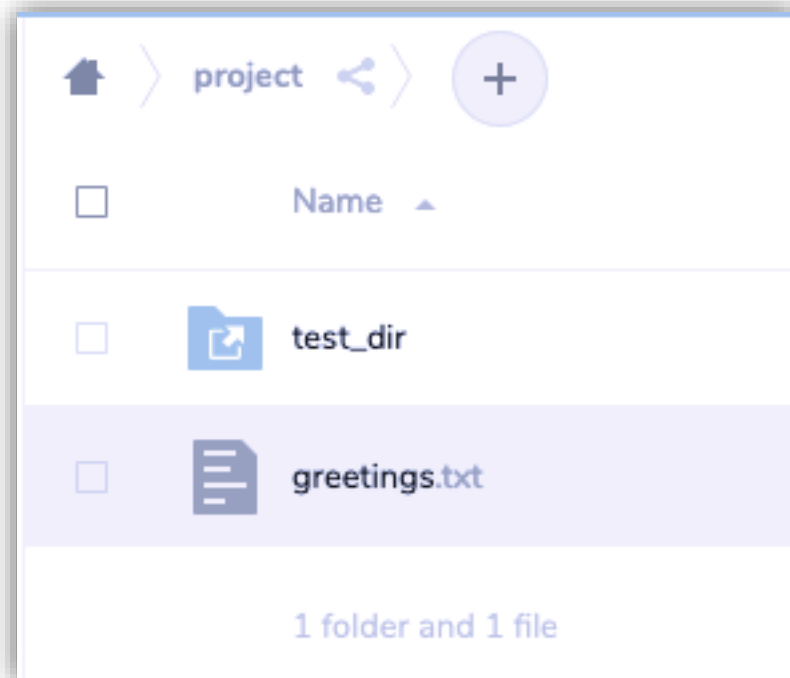
5.



6.

► External Storage Mount:

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



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

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

```
[zpang1@Nucleus005 ~]$ pwd
[/home2/zpang1
[zpang1@Nucleus005 ~]$ ls
CellProfilerLocal.cfg  Downloads                ldap002_iptabl
t      Videos
cloud_ssl              downtime-script.sh      line_chart.png
VirtualBox VMs
[Desktop                jupyter_notebooks      mlnx_ofed_eula
[Documents              ldap002_iptables_0.md  mlnx_ofed_eula
[zpang1@Nucleus005 ~]$ ls /project
[apps                   backup_home2            BICF
[apps_database          backup-images           biohpcadmin
[apps_new               backup_services         bioinformatics
[backup_archive         backup_work             biophysics
[zpang1@Nucleus005 ~]$
[zpang1@Nucleus005 ~]$ ls /work/biohpcadmin/zpang
[lists  VirtualBox_VMs  werk_notes
```

Among BioHPC Users

- ▶ ``../shared/..`` directories

Non-BioHPC Collaborators

- ▶ Cannot access cluster components (Thunder, Lamella)

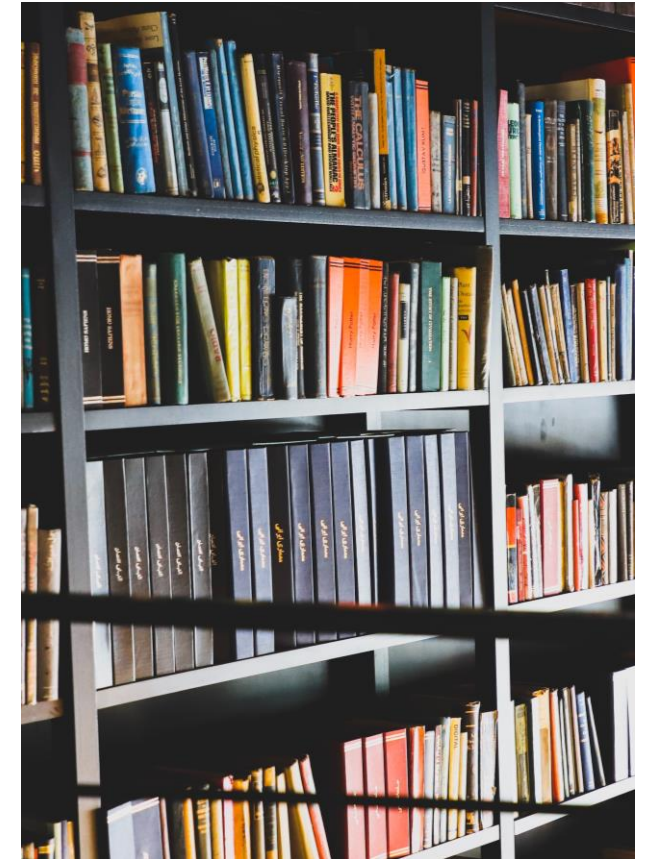
External Collaborators

- ▶ 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``








| 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

| | | Size | Modified |
|----------------------------------------------------------------------------------------|-----|---------|--------------|
|  | ... | Pending | a month ago |
|  | ... | 1.4 MB | 2 months ago |
|  | ... | Pending | 6 months ago |
| 1.  | ... | Pending | 22 days ago |
|  | ... | Pending | 2 months ago |




My_Scripts

★ < 1 KB, 22 days ago

Tags

Activity Comments Sharing

Name... →

2. 

Share link

Add to a project
Connect items to a project to make them easier

New share link

Sharing with Non-BioHPC Collaborators: < 100GB

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

The screenshot shows a file sharing interface for a file named 'My_Scripts'. The file size is '< 1 KB' and it was uploaded '22 days ago'. There are icons for 'Activity', 'Comments', and 'Sharing'. A red circle highlights the 'Copy link to clipboard' icon in the sharing options. A dropdown menu is open, showing various permissions: 'Read only' (selected), 'Allow upload and editing', 'File drop (upload only)', 'Hide download', 'Password protect' (checked), and 'Set expiration date' (checked). The expiration date is set to '22-10-2019'. There is also a 'Choose a password for' field and a 'Delete share link' option.

*Copy link to clipboard

▶ Share link

▶ Add to a project
Connect items

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

Sharing with Non-BioHPC Collaborators: < 100GB

The image shows a file sharing interface. At the top, it says "My_Scripts shared by Zengxing Pang (z pang1)". There is a "Download all files" button and a menu icon. A red circle highlights a plus sign icon with the text "*Upload Files" next to it. Below this is a list of files:

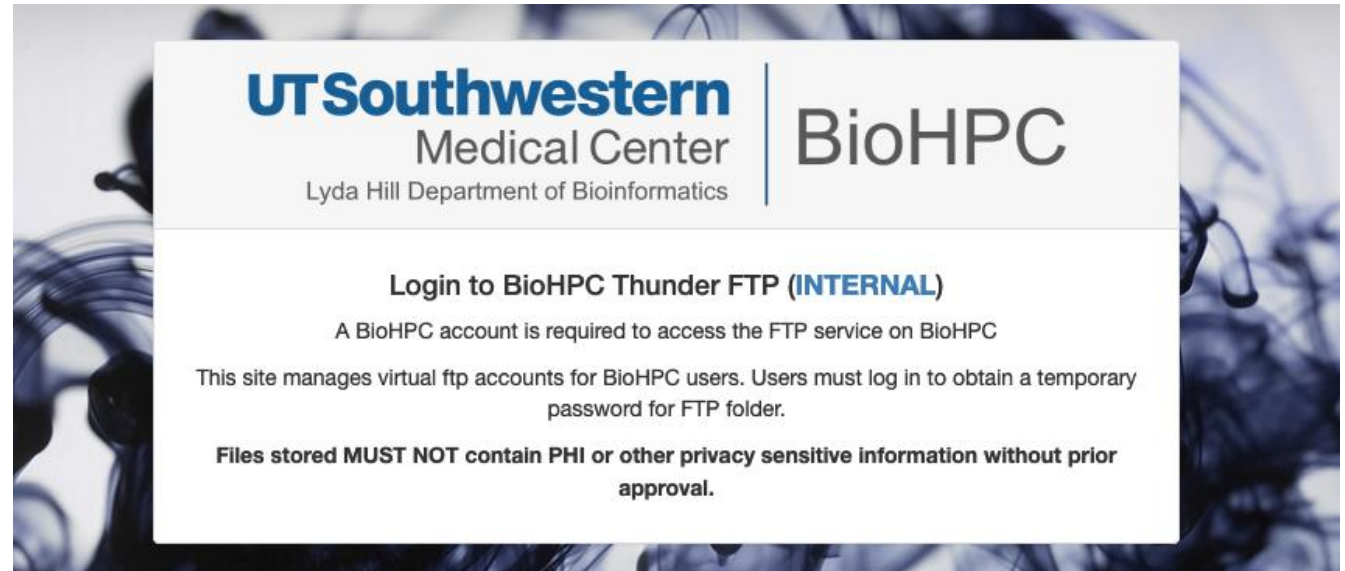
- Name ▾
- <> system_report.sh
- BLOWFISH
- blowfish.h2
- <> blowfish.h
- <> blowfish.cpp
- <> 14_DiskUse.sh
- <> 13th_Bash.sh
- <> 12th_Bash.sh

On the right, there is a preview of a password-protected share page. The URL is `biohpc.swmed.edu/index.php/s/MPJLd2XCFmPPDLS/authenticate/showShare`. The page features the UT Southwestern Medical Center BioHPC logo and the text "Internal Edition". A password field is shown with a "This share is password-protected" message above it.

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 accounts

Share with UTSW collaborator: +

Add new Guest

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

Guests: Create and manage guest FTP accounts

Share with UTSW collaborator: +

Email address *

Note

Directory Name *

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 /z pang1 /another_test | Extend Delete Resend |

Sharing with Non-BioHPC Collaborators: > 100GB

BioHPC user:

z pang1 has shared his BioHPC ftp folder with you with a message:

UTSW employee: z pang1 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

Host: Username: Password: Port:

Status: Connection established, waiting for welcome message...
Status: Insecure server, it does not support FTP over TLS.
Status: Logged in
Status: Retrieving directory listing...
Status: Calculating timezone offset of server...
Status: Timezone offset of server is 0 seconds.
Status: Directory listing of "/" successful

Local site: Remote site:

Local site directory listing:

- .ssh
- .vim
- .vnc
- .vscode
- Applications
- Desktop
 - Oct_16_19
 - Temps
 - Work

Remote site directory listing:

| Filename | Filesize | Filetype | Last modified |
|---------------------------|-------------|--------------|-----------------|
| .. | | | |
| limit_rate.png | 35,833 | png-file | 10/15/2019 1... |
| snow capped mountains.jpg | 14,089,0... | jpg-file | 10/15/2019 1... |
| test_statement.md | 6 | Visual Stu.. | 09/26/2019 1... |
| test_statement_back.md | 6 | Visual Stu.. | 10/15/2019 1... |

Local site file listing:

| Filename | Filesize | Filetype | Last modified |
|-----------|----------|-----------|---------------|
| .. | | | |
| Artifacts | | Directory | 10/01/2019 |
| Meetings | | Directory | 02/27/2019 |
| Pre_Docs | | Directory | 10/01/2019 |
| RedTape | | Directory | 03/02/2019 |
| .DS_Store | 10,244 | File | 10/08/2019 |

Sharing with External Collaborators:

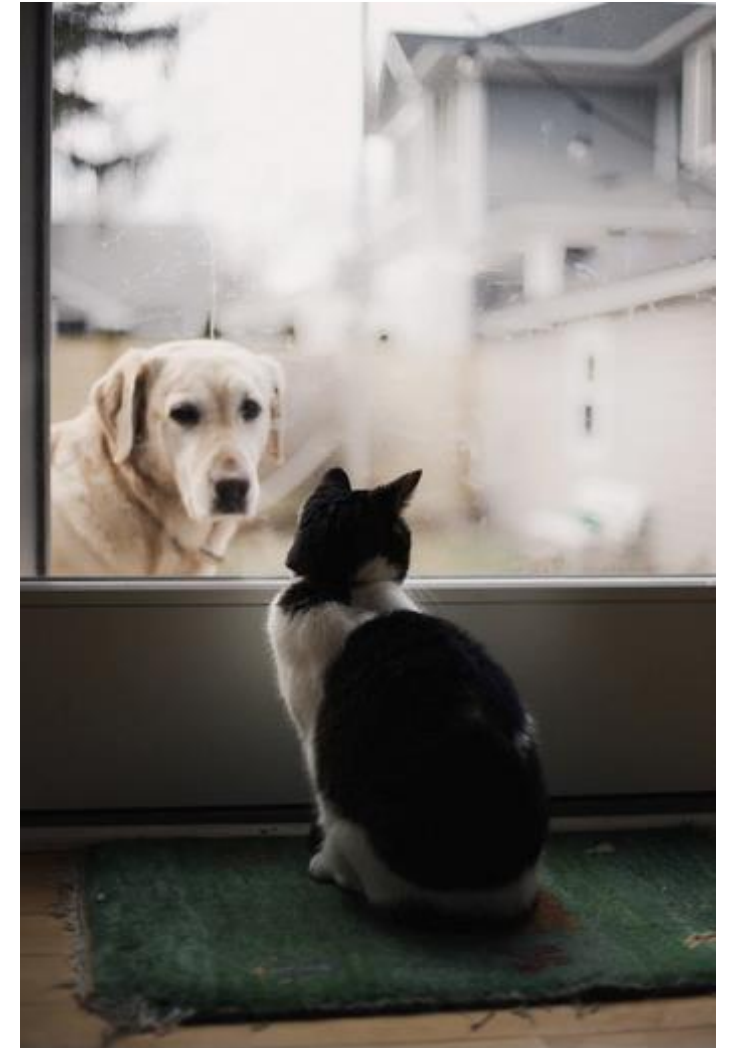
< 50GB -> BioHPC External File Exchange

cloud.biohpc.swmed.edu

> 50GB -> Flash FTP Service

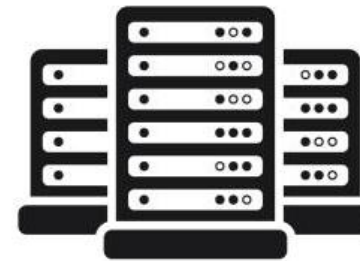
flash.biohpc.swmed.edu

- ▶ 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