

Where is my Cryo-EM Data?

You will first need a SLAC unix computer account to be access your data.

There are numerous ways to access your data. Each have their pros and cons which will be discussed. All data is kept on SLAC's global file system and is available from all centrally managed hosts.

All experimental data is stored under `/sdf/group/cryoem/exp/YYYYMM/YYYYMMDD-AA##` where Y is year, M is month, D is day, and AA## is the 4-character alphanumeric proposal number. Example: if your experiment name is 20180520-CS01, then your data will be located at `/sdf/group/cryoem/exp/201805/20180520-CS01`.

Permissions to access the data can be controlled via the elogbook <https://cryoem-logbook.slac.stanford.edu/lgbk/experiments>. One should add the SLAC unix ID whomever should have access to the data to the list of 'collaborators' for the experiment.

RSYNC/SCP

Rsync can be used over ssh to bulk transfer/synchronise data across different locations. **The data transfer nodes (`dtn01.slac.stanford.edu` and `dtn02.slac.stanford.edu` have been deprecated as of January 2024. Please use `centos7.slac.stanford.edu` for this purpose until further notice.**

Globus

One can create a free Globus account at <https://globusonline.org>. From there, one can download a local client ([Globus Connect Personal](#)) so that files may be copied to say your laptop, or staged to another Globus Endpoint that your institution may run.

The Globus 4 endpoint at `slac#cryoem` is no longer supported. (Globus shut down all version 4 endpoints worldwide on December 18th 2023).

There is a new Globus v5 instance running but is currently (as of Feb 2024) only available to S3DF users. Email your microscope support contact for an invitation to join S3DF (the new cluster), once you accept and complete your registration at coact.slac.stanford.edu, you should be able to access the **slac#s3df_globus5** endpoint from globus.org. You will be asked to choose an idP (identity provider), you should be able to log in using your SLAC unix credentials to either option. The Cryo-EM data is mounted on that endpoint at:

`/fs/ddn/sdf/group/cryoem/exp/`

SAMBA

If you are onsite at SLAC, you can access the data via samba/cifs. You should connect to `zslaccfs` in order to browse the global directory. From there you can access the cryo-EM disks under `cryoem`.

You should login with your SLAC Windows Account to use this.

1. On your Linux machine, open a terminal window.
2. Install the necessary software with the command `sudo apt-get install -y samba samba-common python-glade2 system-config-samba`.
3. Type your sudo password and hit Enter.
4. Allow the installation to complete.
5. Open a new file browser window.
6. At the bottom of the left navigation pane, click "Other Locations"
7. At the bottom of the window, in the "Connect to Server" field, Type `smb://zslaccfs/`
8. Open the cryoem directory.
9. Log in with your SLAC Windows account username and password, leave the "Workgroup" field as the default value or if using Ubuntu the domain should be "slac".
10. You can now browse the CryoEM disks in your Linux file browser.

Alternatively, you can mount it via command line.

1. Install cifs-utils. `sudo apt-get install cifs-utils`
2. Create a directory where you want to mount it. `sudo mkdir /mnt/slac`
3. `sudo mount -t cifs -o username=SLAC_USERNAME,vers=1.0,domain=slac,uid=id -u` //zslaccfs.slac.stanford.edu/cryoem/ /mnt/slac`

Note that the `uid` is important so that you as user can have the correct permissions on your local desktop.

Alternatively, you can mount it via command line.

1. Install cifs-utils. `sudo apt-get install cifs-utils`
2. Create a directory where you want to mount it. `sudo mkdir /mnt/slac`

```
3. sudo mount -t cifs -o username=SLAC_USERNAME,vers=1.0,domain=slac,uid=`id -u` //zslaccfs.slac.stanford.edu/cryoem/ /mnt/slac
```

Note that the `uid` is important so that you as user can have the correct permissions on your local desktop.

SSHFS

You can also use the FUSE based SSHFS to mount the filesystem via ssh. It is recommended to ssh into `dtn02.slac.stanford.edu` in order to use this method.

Using SSHFS, the filesystem will be mounted locally so that you may browse the directory as you would with SAMBA etc.

BBCP

BBCP provides multi-stream parallel transfers of data that allows transfers speeds to be significantly higher than that of scp. Details TBD.

Summary

		Globus	SAMBA	SSHFS	RSYNC/SCP	BBCP
Software Install		Clients available, Servers should already have it installed	Implementation already baked into MAC OS and Windows. Simple install for Linux.	Requires FUSE library. Not available for Windows (I think)	Command line tools usually already installed in most Linux and MAC OSes. GUI's available for SCP.	Command line tools only
Graphical Interface		Yes (Web)	Yes (OS)	Yes (OS)	Potentially	No
Command line interface		Yes	Yes (standard OS)	Yes (standard OS)	Yes (standard OS)	Yes (standard OS)
Performance		Fast	Fast	Slow	Slow	Fast
Access		Anywhere	At SLAC Only	Anywhere	Anywhere	Anywhere
Credentials		Globus ID + SLAC Unix	SLAC Windows	SLAC Unix	SLAC Unix	SLAC Unix
Ease of Use		Easy	Easy	Medium	Difficult	Difficult