

User Policies: CryoEM Data Processing and Storage

Processing server: “imagingsuitesupermicro” (10.20.1.13)

Storage server: “nanosci_imaging_storage” (10.20.1.11)

Both servers are managed by Imaging Facility and Amedee’s lab. Please contact facility staff (twang1@gc.cuny.edu) or Amedee (adesgeorges@gc.cuny.edu) for approval of user accounts and any other data processing/storage related requests. There are policies I expect everyone will follow.

1. In principle, only data collected in the facility can be put on both servers for storage and processing.
2. Temporary storage and processing of data collected outside ASRC are allowed, however with limitations. Such data will be asked to move out of the servers after the storage is 75% full for that server.
3. On processing server, all data and project directories must be in /data/your-username/. On storage server, all data must be in /home/Data/your-username/.
4. Anyone who plans to use both servers must have user account on the server, which requires approval from both the facility and Amedee. Either facility staff or Amedee can create user account for you. Generic user accounts, “exx”, “cryosparc_user”, are not allowed to submit jobs on processing server and jobs associate with generic user accounts will be forced to stopped by facility staff at anytime.
5. Connecting from outside ASRC needs VPN. ASRC employees must work with their initiative and ASRC IT to get VPN.
6. VPN is not available to non-ASRC employees (I will update when it becomes available), who need to have a teamviewer account and request approval from facility staff or Amedee to be added to “Lab: IS cryoEM” group in our corporate account (via ASRC IT).
7. Please always keep in mind that this computer is not cluster (no PBS system or similar, all Jobs submitted will start right away). There is no queue for submitted jobs. If the total requested number of CPU/GPU from all users at any given time is larger than the number of CPU/GPU of the computer, multiple jobs will share the resources, e.g. if two job share one CPU each job only use 50% of that CPU, or later job will kill the job that is already running on that GPU. The later is true for RELION jobs running on GPUs.
8. Check the GPU status and job status before submit jobs, only request to use those GPUs that are not occupied. For RELION jobs, each user should request to use two GPUs and leave the rest two to other users. (This requires that you should not leave the gpu option blank in RELION that will use all gpu resources on the computer). The GPU ID is 0,1,2,3 that can be put into two groups: 0,1 and 2,3.
9. Cryosparc jobs start with GPU 0 if you did not specify GPU in job queue. Please assign GPU 0,1 for Cryosparc jobs and GPU 2,3 for RELION jobs.
10. The default (each time you login) user directory is /home/your-username/ on both servers. Please do not save any data in this directory nor using this

directory as your project-working directory. Bulky data should be put in /data/your-user-name/ on “imagingsuitesupermicro”, which has about 20TB space, and in /home/Data/your-user-name/ on “nanosci_imaging_storage”, which has about 160TB space.

11. All users should be in asrcimaging group. Please check the permissions of your directory (I usually give 755 permission when I create user account), let me know if you are not satisfied and we can discuss to modify the permissions.
12. Please do not use “imagingsuitesupermicro” for long-term storage. I will keep all data if the directory (/data/) has more than 25% storage space, I will send notice to ask you to delete data either more than four weeks old or data not collected in facility (you will have three days to back up your data after receiving notice) once the storage space are more than 75% occupied.
13. For “nanosci_imaging_storage”, facility staff will send notice to users to move their data not collected in the facility to your own storage space when the storage space (/home/Data/) is more than 75% occupied.