Data Storage

Local and Cluster-Wide data storage locations

- Home Directories
- DFS
- Node-Local Storage

Home Directories

User home directories are regularly backed up, but limited in space. Individual quotas start at 20GB, but can be increased on request.

They are NFS mounted from the server node nfs-0, and appear on every user-accessible node as "/nfspool-0/home/username". Your home directory can be referenced as "/export/home/username", "/export/nfs0home/username", or the shortcut "~".

On logging in, you should see a message reporting your current home directory usage. If it is getting full, either delete/move data, or ask for more space. When it reaches 100%, it will be difficult to log in correctly.

DFS

We have several Distributed File Systems on Greenplanet. They are generally accessible from every node, and are each made up of multiple servers.

Data

/DFS-L/DATA (1.6PB)

DFS-L is the current main Lustre (2.12) file system for data. The /DFS-L/DATA subdirectory contains directories for each research group, which contain sub-directories for each user in that group.

There is no quota for this file system, but usage is tracked for billing purposes. Each group gets one TB free, with additional data storage at $2/TB/month. [Data usage is aggregated with the deprecated /DFS-B]

/DFS-B/DATA (109TB)

/DFS-B is an older file system using BeeGFS (7.2). It uses some of the same physical servers as /DFS-L, so space used on one lowers capacity on the other. This is why usage is aggregated with /DFS-L for billing.

At the time, that it was initially installed, BeeGFS had superior metadata performance to Lustre (i.e. it could handle lots of small files quickly). Since then, Lustre performance on these workloads has improved enough to not need multiple file systems competing for resources. /DFS-B will not be accessible on the newer Rocky Linux 8 side of the cluster, so migrating from /DFS-B/DATA to /DFS-L/DATA is recommended.

/D2/DATA (0TB)

/D2/DATA is currently being built but will be based on the next stable Lustre with long-term support (2.17 or 2.18, TBD). There will likely be a long period of coexistence of /DFS-L and /D2 for data migration,

Scratch

Scratch file systems are intended for short or medium term data storage. It is free of charge, but old and unused data can be purged if space becomes tight.

Scratch data deletion policy (as of 24 February 2026)

Temporary files and directories created by known slurm submission scripts (e.g. /XXX-L/SCRATCH/$group/$user/$SLURM_JOB_ID) will be deleted if they meet all of these criteria:

- Not an actively running/pending job
- Not open on login or compute nodes
- Have a most-recent access time older than 365 days ago

Other files that we can't identify as obvious trash will have these deletion criteria:

- 1) Not open on login or compute nodes
- 2) Have a most-recent access time older than 1080 days ago

/XXX-L/SCRATCH (143TB)

A smaller Lustre (2.15) file system built on older, but highly redundant hardware. Currently very full, and may be decommissioned as the disks age out.

/X2/SCRATCH (263TB)

Another small Lustre (2.15) file system, built on less old but still redundant hardware. Currently the default scratch area used by the Slurm submission scripts we provide for multi-node or large jobs.

Node-Local Storage

Each compute and login node has a local storage area, accessible as /scratch and /work. On most nodes, this is a small disk (200GB-1TB), but some diskless nodes use space on /XXX-L.

Single-node slurm jobs can create temporary workspace her under the user's directory (e.g. /scratch/username/slurm_job_id).

These locations can only be accessed from the node they are on, and are subject to automatic cleanup when jobs end. If jobs end due to hardware or queuing system failures, temporary files may recoverable.