GLADE file spaces

The Globally Accessible Data Environment – a centralized file service known as GLADE – uses high-performance GPFS shared file system technology to give users a common view of their data across the HPC, analysis, and visualization resources that CISL manages.

File space	User quota (Data volume)	User quota (File count)	Backup	Purge policy	Descriptions/notices
Home: /glade/u/home/*username*	50 GB	N/A	Yes	Not purged	User home directory Access: POSIX
Scratch: (Cheyenne) /glade/scratch/*username*	10 TB	N/A	No	120 days	Temporary computational space Access: POSIX
Scratch: (Derecho) /glade/derecho/scratch/*username*	30 TB	10 M	No	180 days	Derecho's scratch file system also includes a limit on a users' total number of files
Work: /glade/work/*username*	1 TB	N/A	No	Not purged	User work space Access: POSIX
Project: /glade/p/*entity/project_code*	N/A	N/A	No	Not purged	**Deprecated; will be migrated to Campaign Storage in 2023**
Collections: /glade/collections	N/A	N/A	No	Not purged	Curated collections (CMIP, RDA, others) Access: POSIX
Campaign Storage: /glade/campaign	N/A	N/A	No	Not purged	Project space allocations (via allocation request)
GLADE system status report					

Page contents

Overview

CISL backs up the GLADE **home** file space several times a week and also creates snapshots to enable users to recover deleted files quickly and easily. Data can remain in each of these spaces in accordance with the policies detailed below. The policies are subject to change; any changes necessary will be announced in advance.

CISL does not provide backups of other spaces. You are responsible for the safe storage of any data that must be preserved.

Best practice: Check your space usage regularly with **gladequota** as described below, and remove data that you no longer need.

You can conserve GLADE space by storing large files, such as tar files, rather than numerous small, individual files. This is because the system allocates a minimum amount of space for each file (currently configured to 16 KB), no matter how small the file is. Thus 16KB is the smallest amount of space the system can allocate to a file, including empty files, directories and symlinks.

Home space

Each user has a ***/glade/u/home/username*** space with a quota of 50 GB* for managing scripts, source code, and small data sets. It is backed up weekly, with backups retained for several months. CISL also creates snapshots of the space to enable users to recover deleted files quickly and easily.

Backup policy

- Your /glade/u/home directory is backed up several times a week while your account is open. Each backup is kept for several weeks. The frequency and scheduling of backups and the length of time they are kept may change with no prior notice. If you are unable to find files that you would like to restore in the snapshots of your home directory, contact the NCAR Research Computing help desk to request restoration of the files if they are in a backup.
- Core dump files are not backed up. Core dump file names typically follow this format: **core.xxxxx** (where the extension can include from one to five digits).
- CISL does not purge or scrub your home directory, and deletes files only as stated in the following data retention policy.

Data retention policy

- When your account is closed, files will remain in your home directory for 30 days, after which CISL backs up the final contents and removes them from the file system. This backup is retained for six months after account termination. However, note that your project and group memberships are also terminated as described below, which can limit your ability to access files based on shared group permissions.
- If one or more of your project allocations expires but your account is not closed, files are retained in your home directory.
- Core dump files are not archived.

Work space

Your **/glade/work/*username*** space is best suited for actively working with data sets over time periods greater than what is permitted in the scratch space.

The default quota for these spaces is 1 TB.

Purge policy

- This space is not purged or scrubbed. CISL deletes files only as stated in the following data retention policy.

Data retention policy

- **When your user account is closed, files are retained for 30 days before being deleted**.
- Files are not recoverable from backups, as there are none.
- If one or more of your project allocations expires but your account is not closed, your work directory files are retained.

Scratch file space

Each user has a **/glade/derecho/scratch/*username*** space by default, with an individual quota of 30 TB. The scratch file space is intended to support output from large-scale capability runs as well as computing and analysis workflows across CISL-managed resources. It is a **temporary** space for data to be analyzed and removed within a short amount of time.

If you will need to occupy more than your quota of scratch space at some point, contact the NCAR Research Computing help desk to request a temporary increase. Include a paragraph justifying your need for additional space when making your request.

Purge policy

Individual files are removed from the scratch space automatically if they have not been accessed (for example: modified, read, or copied) in more than 120 days. A file's **access time** (atime) is updated at most once per day for purposes of I/O efficiency. To check a file's atime, run **ls -ul filename**.

Users may not run “touch” commands or similar commands for the purpose of altering their files' timestamps to circumvent this purge policy. CISL staff will reduce the scratch quotas of users who violate this policy; running jobs may be killed as a result.

In addition:

- CISL routinely monitors scratch space usage to ensure that it remains below the 90% mark and to determine if a reduction in the retention period is necessary.
- We will announce in advance any changes to the retention period.

Best practice: To help us avoid the need to shorten the retention period, please use this space conscientiously.

Delete files that you no longer need as soon as you're done with them rather than leave large amounts of data sitting untouched for the full 120 days. If you need to retain data on disk for more than 120 days, consider using your **/glade/work** space or Campaign Storage space.

Data retention policy

- When your account is closed, files are purged automatically as they become 120 days old.
- If one or more of your project allocations expires but your account is not closed, your scratch files are removed as stated in the purge policy.
- Files are not recoverable from backups, as there are none.

Campaign Storage / project space

Dedicated project spaces on the **/glade/campaign** file system are available through our allocations process to support longer-term disk needs that are not easily accommodated by the scratch or work spaces. Allocations for project spaces are made to collaborative groups of users through the University/CHAP or NCAR allocations processes. The allocations are based on project needs and resource availability. Requests are reviewed according to the various allocation schedules.

If you have a user account and project space but lack the directory permissions you need for that space, contact the NCAR Research Computing help desk to request changes. Identify the directories and the permissions you are requesting.

Access reports

CISL generates weekly usage reports throughout /glade to help users manage their data. The reports provide a summary of when files were last accessed, how much space is used, and details for the top 25 users. The files are named **access_report.txt** and can be found at the top-level of shared file spaces, for example:

- /glade/campaign/group_name/ (or similar, depending on the project)
- /glade/p/lab_name/group_name/ (or similar, depending on the project)
- /glade/p/univ/project_code/
- /glade/p/uwyo/project_code/

Data retention policy

- Users' files are not deleted from project space after their accounts become inactive.
- Files are not recoverable from backups, as there are none.
- CISL reserves the right to reclaim space from expired projects.

As disk space is a limited resource shared by the entire community, permanent storage of infrequently accessed data should be avoided. These spaces should be used as efficiently as possible so that other projects can take advantage of the resource. The Quasar tape storage system is more appropriate for long-term storage of infrequently accessed data.

Checking space usage

Knowing your quotas and usage is important

All files that you own in your home, work, or scratch spaces are counted against your GLADE quota, regardless of the directory in which they are stored. If you write files to another user's home or scratch space, for example, they still count against your own individual user quota for that space.

If you reach your disk quotas for the GLADE file spaces (see **gladequota** below), you may encounter problems until you remove files to make more space available. For example, you may not be able to log in, the system may appear hung, you may not be able to access some of your directories or files, your batch jobs may fail, and commands may not work as expected.

If you cannot log in or execute commands, contact the NCAR Research Computing help desk. You can check your space usage as shown below.

gladequota command

This command will generate a report showing your quota and usage information:

```
gladequota
```

* Output from the **gladequota** command will show the home space quota as 100 GB instead of 50 GB. This is because the system stores dual copies of users' data for increased data integrity and safety. In some circumstances, queries of storage utilization from **du** and **ls** will also report a duplicated data footprint in your home directory for the same reason.

General data retention policies and procedures

Project data

When a sponsored project approaches expiry, there are several steps in the process that affect the accessibility of associated data:

- 30 days before expiration, the project PIs will receive an email reminding them of the pending expiration. The project team should assess remaining files and disposition appropriately in preparation for group deactivation.
- 90 days after project expiration, the UNIX group associated with the project is removed. At this point users with accounts remaining on the system will likely no longer have access permissions to the projects' data, as the primary group no longer exists. It is therefore *imperative* that any remaining project data be relocated and ownership permissions assessed prior to this group deactivation.
- Finally, files are removed as scheduled on the timeline described above for the relevant file system.

Restoring access to project data

CISL has limited ability to modify access to project data after the 90-day post-expiry window. Such modifications require the approval of the original project owner. CISL has no ability to restore data after the purge or removal policies stated above have taken effect.

User accounts

User accounts are deactivated when they are no longer associated with an active project. When a user account is deactivated, several steps in the process affect the accessibility of the users' data:

- 30 days after a user account is deactivated, a final home directory backup is performed and the home directory is removed.

- The user’s work directory is removed. No backup is performed.
- Finally, additional scratch files are removed as scheduled on the timeline described above for the relevant file system.

Restoring access to collaborators' data

A typical request for data access comes not from the departing user, but from remaining collaborators. Colleagues occasionally request access to a departed users' files, sometimes many months after the account is terminated, often when they realize the original owner set permissions that limit their access.

While CISL has a limited ability to help in these situations, there are also legal limits to what we can do. For example, CISL cannot share files beyond the clear intent of the original owner as inferred from the UNIX file permissions. If a collaborator would like access to a file that was previously group- or world-readable, we may be able to help. If the original file was restricted to user-only read, however, we cannot override those intentions. The only exceptions to this policy are in compliance with broader UCAR IT records or investigation policies as described in UCAR's 1-7 Information Security Policy.