

Disk Space and Data Storage

Category: Docs

Table of Contents



- [Centralized Data Stores](#)
 - [Quotas](#)
 - [Permissions](#)
 - [Snapshots](#)
- [Local Disk](#)
 - [Desktops](#)
 - [Laptops](#)
- [Non-SCI data stores](#)
 - [UIT](#)
 - [CHPC](#)

Data lies at the heart of almost everything we do in the computing world, and there are a myriad of attributes that are important in figuring where to store data and how to access it. To enable SCI users to best manage their data, SCI IT provides a number of different data storage locations. This document details those locations and their features.

If you need more disk space in a centralized data store (home directory, research project, etc), would like to set up a new shared directory, or need to recover lost data that you weren't able to find in your snapshots, please [contact SCI IT](#).

Centralized Data Stores

SCI IT provides centralized disk space that is fast, highly available, accessible from any SCI computer, and regularly backed up. There are two main types of centralized storage:

- **Home Directories**– the home directory you have when you

log into a SCI computer is meant for your personal storage, and may be used however you like. These are accessed via `/home/{sci,collab}/<username>`.

- **Research and Other Directories**– these directories are provided for specific purposes, such as for a specific research project or group, departmental web pages, etc. They are usually accessed via `/usr/sci` but other directories also exist.

Quotas

Each of these directories are limited to a set size (a quota) in order to manage costs and assure availability. For home directories, quotas are:

- Faculty: 1TB
- Staff/Postdoc/Grad: 500GB
- Undergrad: 100GB
- Collaborator: 10GB

Permissions

Access to files and directories, including shared directories such as those under `/uu/sci.utah.edu/projects`, are managed using standard [UNIX filesystem permissions](#). Each researcher or research group has their own methodology for how they manage access control– if you need access to a particular set of files/directories, please contact that person or group directly about being given access.

Snapshots

If you accidentally delete data in one of the centralized data stores, you may be able to recover it yourself by copying it out of a snapshot directory. Snapshots are automatically taken on an hourly, nightly, and weekly basis by the file servers that provide our centralized disk space. Snapshots can be accessed via the `.snapshot` directory, which is a special

subdirectory that exists in every directory in our centralized storage. Note that the `.snapshot` directory will not appear unless you explicitly reference it on the command line, as in the example below, where I remove a file and then get a backed up copy from the hourly snapshot taken at 5pm (17:00 in military time):

```
[@memphis] ../dev/desktop > ls -al
```

```
total 40K
```

```
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 17:58 ./
```

```
drwxr-xr-x 4 clake sci-it 1.0K Oct 23 17:51 ../
```

```
-rw-r--r- 1 clake sci-it 6.1K Jun 26 14:04 50unattended-upgrades
```

```
-rwxr-xr-x 1 clake sci-it 582 Sep 04 17:58 dns_check*
```

```
-rwxr-xr-x 1 clake sci-it 4.0K May 31 17:09 harden_sshd*
```

```
-rwxr-xr-x 1 clake sci-it 286 Jul 01 15:13 set_mtu*
```

```
-rwxr-xr-x 1 clake sci-it 181 Apr 30 17:06 Ubuntu_24.04_customization*
```

```
[@memphis] ../dev/desktop > rm dns_check
```

```
rm: remove regular file 'dns_check'? y
```

```
[@memphis] ../dev/desktop > ls -al
```

```
total 36K
```

```
drwxr-xr-x 2 clake sci-it 2.0K Oct 23 17:53 ./
```

```
drwxr-xr-x 4 clake sci-it 1.0K Oct 23 17:51 ../
```

```
-rw-r--r- 1 clake sci-it 6.1K Jun 26 14:04 50unattended-upgrades
```

```
-rwxr-xr-x 1 clake sci-it 4.0K May 31 17:09 harden_sshd*
```

```
-rwxr-xr-x 1 clake sci-it 286 Jul 01 15:13 set_mtu*
```

```
-rwxr-xr-x 1 clake sci-it 181 Apr 30 17:06 Ubuntu_24.04_customization*
```

```
[@memphis] ../dev/desktop > ls -lt .snapshot
```

```
total 132K
```

```
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 17:00 hourly_sci_17:00/
```

```
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 16:00 hourly_sci_16:00/
```

```
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 15:00 hourly_sci_15:00/
```

```
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 14:00 hourly_sci_14:00/
```

```
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 13:00 hourly_sci_13:00/
```

```
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 12:00 hourly_sci_12:00/
```

```
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 11:00 hourly_sci_11:00/
```

```
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 10:00 hourly_sci_10:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 09:00 hourly_sci_09:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 08:00 hourly_sci_08:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 07:00 hourly_sci_07:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 06:00 hourly_sci_06:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 05:00 hourly_sci_05:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 04:00 hourly_sci_04:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 03:00 hourly_sci_03:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 02:00 hourly_sci_02:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 01:00 hourly_sci_01:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 00:05 nightly_sci_10_23/
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 00:00 hourly_sci_00:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 22 23:00 hourly_sci_23:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 22 22:00 hourly_sci_22:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 22 21:00 hourly_sci_21:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 22 20:00 hourly_sci_20:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 22 19:00 hourly_sci_19:00/
drwxr-xr-x 2 clake sci-it 2.5K Oct 22 00:05 nightly_sci_10_22/
drwxr-xr-x 2 clake sci-it 2.5K Oct 21 00:05 nightly_sci_10_21/
drwxr-xr-x 2 clake sci-it 2.5K Oct 20 00:05 weekly_sci_2024_10_20/
drwxr-xr-x 2 clake sci-it 2.5K Oct 19 00:05 nightly_sci_10_19/
drwxr-xr-x 2 clake sci-it 2.5K Oct 18 00:05 nightly_sci_10_18/
drwxr-xr-x 2 clake sci-it 2.5K Oct 17 00:05 nightly_sci_10_17/
drwxr-xr-x 2 clake sci-it 2.5K Oct 13 00:05 weekly_sci_2024_10_13/
drwxr-xr-x 2 clake sci-it 2.0K Oct 06 00:05 weekly_sci_2024_10_06/
drwxr-xr-x 2 clake sci-it 2.0K Sep 29 00:05 weekly_sci_2024_09_29/
```

```
[@memphis] ../dev/desktop > cp .snapshot/hourly_sci_17:00/dns_check .
```

```
[@memphis] ../dev/desktop > ls -al
```

```
total 40K
```

```
drwxr-xr-x 2 clake sci-it 2.5K Oct 23 17:58 ./
drwxr-xr-x 4 clake sci-it 1.0K Oct 23 17:51 ../
-rw-r--r- 1 clake sci-it 6.1K Jun 26 14:04 50unattended-upgrades
-rwxr-xr-x 1 clake sci-it 582 Oct 23 17:58 dns_check*
-rwxr-xr-x 1 clake sci-it 4.0K May 31 17:09 harden_sshd*
-rwxr-xr-x 1 clake sci-it 286 Jul 1 15:13 set_mtu*
```

```
-rwxr-xr-x 1 clake sci-it 181 Apr 30 17:06 Ubuntu_24.04_customization*
```

```
[@memphis] ../dev/desktop >
```

Local Disk

Desktops

Practically all modern desktop and workstation computers come with locally installed hard disks that provide space well in excess of what is needed by the operating system. The standard build in SCI makes the extra space not needed by the operating system available for users to access and use as scratch space.

This local disk space is available on the local machine as `/scratch` or `/scratch_{storage_type}` (e.g. `/scratch_nvme` or `/scratch_ssd`).

The benefits of using this scratch space is that it there generally is a lot of space available, and when used locally on a machine, it is fast. However, this scratch space *is not backed up*, so in the event of a hardware failure or data deletion, SCI IT will be unable to recover data. As such, this is best used for data that can be recreated or recovered from elsewhere (i.e. kept in multiple places).

Laptops

Similarly to desktops, most modern laptops come with a large amount of available space on them. Given that *laptops are not backed up by SCI IT*, any important information on your laptop should be backed up by copying it to a centralized data directory or by some other means of backing up your laptop (e.g. 3rd-party backup software, manual copying to an external disk, etc). If you do backup your laptop outside of SCI, please be mindful about sensitive data and where/how that data may be copied.

Non-SCI data stores

The University of Utah contains many other groups and Departments that provide data storage services, in addition to those provided at SCI. Two of the most relevant ones are:

UIT

UIT has several options for storage, including personal and departmental storage via [Box](#) and [One Drive](#), departmental shares, [Google Workspace](#), etc. For more information, see the [UIT Service Catalog](#).

CHPC

When using resources in the Center for High Performance Computing, there are several options for data storage, including home directories, scratch space, etc. Additionally, they do provide storage in their Protected Environment if using sensitive data. For more information, see the [CHPC page on Storage Services](#).