Storage

Where Should I Store My Data?

1. Data undergoing active analyses should be stored in HPC's local High Performance Storage (Tier 1).

2. Research data not requiring immediate access should be stored in General Research Data Storage (Tier 2).
   For example:
   a. Large datasets where only subsets are actively being analyzed
   b. Results no longer requiring immediate access
   c. Backups (highly encouraged)

3. Data no longer involved in ongoing analyses that need long-term archiving should be stored in Long term Research Storage (Tier 3).

HPC High Performance Storage (Tier 1)

Data stored on HPC are not backed up! All data on this storage should be backed up elsewhere by UA researchers, preferably in three places and two formats.

We strongly recommend that you do some regular housekeeping of your allocated space. Millions of files are hard to keep organized and even more difficult to migrate. Archiving or using a tool like tar will help keep our disk arrays efficient and potentially free up more space for you to use.

Summary

Every user has access to individual and group storage on the system where they can store data for active analyses as summarized below:
<table>
<thead>
<tr>
<th>Path</th>
<th>Description</th>
<th>Quota</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>/home/uxx/netid</td>
<td>An individual storage allocation provided for every HPC user</td>
<td>50GB</td>
<td></td>
</tr>
<tr>
<td>/groups/pi_netid</td>
<td>A communal storage allocation provided for every research group</td>
<td>500GB</td>
<td></td>
</tr>
<tr>
<td>/xdisk/pi_netid</td>
<td>Temporary communal storage provided for every group.</td>
<td>200GB-20TB</td>
<td>Up to 300 days</td>
</tr>
<tr>
<td>/tmp</td>
<td>Local storage available on individual compute nodes.</td>
<td>&lt; 800GB to 1.4TB</td>
<td>Only accessible for the duration of a job's run.</td>
</tr>
</tbody>
</table>

### Checking your Storage Quota and Usage

#### Command Line

To check your storage usage, use the command `uquota`. For example:

```bash
(puma) [netid@junonia ~]$ uquota
used  soft limit  hard limit
/groups/pi_netid      6.6G       500.0G      500.0G
/home                   37.1G      50.0G       50.0G
/xdisk/pi_netid        12.9G       9.8T       9.8T
```

#### User Portal

You can check your storage allocation through our online user portal by navigating to the Storage tab and clicking **Check Disk Quotas**:

![Check Disk Quotas](image)

### xdisk

#### Overview
Overview

What is xdisk?

xdisk is a temporary storage allocation available to all PIs and offers up to 20 TB of usable space for their group for up to 300 days. A PI can request an allocation either via the command line or through our web portal (no paperwork necessary!). Once an xdisk allocation is created, it is immediately available for use.

Because xdisk allocations are temporary, they will expire as soon as their time limit is reached. Warnings will be sent to every group member at their netid@arizona.edu addresses beginning one week before the expiration. It is the group's responsibility to renew xdisk allocations or copy files to an alternate storage location prior to the expiration date. Once an xdisk allocation expires, everything in it is permanently deleted. PI's may request a new xdisk allocation immediately after their previous one has expired.

Requesting an xdisk Space

Requesting an xdisk Space

Faculty members (PIs) or their designated xdisk delegates are able to request, alter, extend, and delete an xdisk allocation from the web portal under the storage tab: https://portal.hpc.arizona.edu/portal/

To request a new allocation, select Manage XDISK, fill in the form, and submit.

Check Disk Quotas
Manage XDISK
Delete XDISK

CLI Commands
xdisk CLI

xdisk is a locally written utility for PI's to create, delete, resize, and expire (renew) xdisk allocations. This functionality is usable by PIs only.

<table>
<thead>
<tr>
<th>Xdisk Function</th>
<th>Information</th>
<th>Command</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display xdisk help</td>
<td>Commands given in brackets are optional. If left blank, you will get system defaults.</td>
<td>$ xdisk -c help</td>
<td>$ xdisk -c help /usr/bin/xdisk -c [query</td>
</tr>
<tr>
<td>View current information</td>
<td>Check current allocation size, location, and expiration date.</td>
<td>$ xdisk -c query</td>
<td>$ xdisk -c query XDISK on host: ericidle.hpc.arizona.edu Current xdisk allocation for &lt;netid&gt;: Disk location: /xdisk/&lt;netid&gt; Allocated size: 200GB Creation date: 3/10/2020 Expiration date: 6/8/2020 Max days: 45 Max size: 1000GB</td>
</tr>
<tr>
<td>Create an xdisk</td>
<td>Grants an xdisk allocation. Max Size: 20000 GB Max Days: 300</td>
<td>$ xdisk -c create -m [size in gb] -d [days]</td>
<td>$ xdisk -c create -m 300 -d 30 Your create request of 300 GB for 30 days was successful Your space is in /xdisk/&lt;netid&gt;</td>
</tr>
<tr>
<td>Extend the xdisk time</td>
<td>Prior to its expiration, if your xdisk's time is under the 300 days, you may increase it until the 300 day limit is reached.</td>
<td>$ xdisk -c expire -d [days]</td>
<td>$ xdisk -c expire -d 15 Your extension of 15 days was successfully processed</td>
</tr>
<tr>
<td>Resize an xdisk allocation</td>
<td>You may resize your allocation by specifying the increase/decrease in gb. To reduce the size, use a negative sign, &quot;-&quot;</td>
<td>$ xdisk -c size -m [size in gb]</td>
<td>$ # Assuming an initial xdisk allocation size of 200 gb $ xdisk -c size -m 200 XDISK on host: ericidle.hpc.arizona.edu Your resize to 400GB was successful $ xdisk -c size -m -100 XDISK on host: ericidle.hpc.arizona.edu Your resize to 300GB was successful</td>
</tr>
<tr>
<td>Delete an xdisk allocation</td>
<td>Permanently deletes your current xdisk allocation. Be sure to remove any important data before deleting.</td>
<td>$ xdisk -c delete</td>
<td>$ xdisk -c delete Your delete request has been processed</td>
</tr>
</tbody>
</table>

Delegating xdisk Management
Delegating xdisk Management Rights

When a user is added as a delegate, it allows them to manage group storage on their PI's behalf through the user portal. It will still not be possible for the delegate to manage storage via the CLI interface.

PIs can delegate xdisk management rights. To add a group member as a delegate, the PI needs to click on Manage Delegates link on the home page of the portal:

Once a group member has been added, they can manage their group's xdisk through the web portal. To do this, they should log into our web portal, click the Switch User link, and enter their PI's NetID. They can then manage their group's space under the Storage tab.

FAQs

A group's PI owns the /xdisk allocation. By default, your PI has exclusive read/write/execute privileges for the root folder /xdisk/PI.
If they so choose, a PI may allow their group members access to their /xdisk by running one of the following commands:

<table>
<thead>
<tr>
<th>Command</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ chmod g+r /xdisk/PI</td>
<td>Group members can see the contents of the directory with ls, but may not access it or make modifications (e.g. add, delete, or edit files/directories)</td>
</tr>
<tr>
<td>$ chmod g+rx /xdisk/PI</td>
<td>Group members can access the directory and see files but cannot make modifications (e.g. add, delete, or edit files/directories)</td>
</tr>
<tr>
<td>$ chmod g+rwx /xdisk/PI</td>
<td>Group members are granted full read/write/execute privileges.</td>
</tr>
</tbody>
</table>

When an /xdisk allocation is created, a subdirectory is automatically generated for and owned by each individual group member. If the directory /xdisk/PI does not have group privileges, group members may not access the root directory, but may access their individual spaces by:

$ cd /xdisk/PI/netid

If a user joins the group after the xdisk was created and /xdisk/PI is not writeable for group members, contact our consultants and they can create one.

Typically when an /xdisk allocation is created, it will automatically generate a directory for each group member. In the unlikely event that it doesn't or, more commonly, a group member is added after the allocation has been created, either the PI may manually create a directory or, if the root directory is group writable, the user may create one themselves. If the group’s /xdisk does not have full group permissions, the PI may run:

$ mkdir /xdisk/PI/netid

then can reach out to our hpc consultants to request an ownership change.

No, the full /xdisk allocation is available for every member of the group. It’s up to group members to communicate with one another on how they want to utilize the space.

A group’s PI owns the /xdisk allocation. By default, your PI has exclusive read/write/execute privileges for the root folder /xdisk/PI.

xdisk is a temporary storage space available to your research group. When it’s close to its expiration date, notifications will be sent to all members of your group. For detailed information on xdisk allocations, see: Storage.

xdisks are managed by your group's PI by default. This means if you want to request an xdisk or modify an existing allocation (e.g., extending the time limit or increasing the storage quota), you will need to consult your PI. Your PI may either perform these actions directly or, if they want to delegate xdisk management to a group member, they may do so by following the instructions under Delegating xdisk Management Rights.

To modify your allocation's time limit or storage quota, your PI can either do so through the Web Portal under the Storage tab, or via the command line. If your PI would like to delegate management rights to a group member, they may follow the instructions under Delegating xdisk Management Rights. Once a group member has received management rights, they may manage the allocation through our web portal.

If you're getting errors using xdisk commands in a terminal session, check that you are on a login node. If you are on the bastion host (hostname: gatekeeper), are in an interactive session, or are on the filexfer node, you won't be able to check or modify your xdisk. When you are on a login node, your terminal prompt should show the hostname junonia or wentletrap. You can also check your hostname using the command:

$ hostname
If you're trying to extend your group's allocation but are seeing something like:

```
(puma) [netid@junonia ~]$ xdisk -c expire -d 1
invalid request_days: 1
```

for every value you enter, your xdisk has likely reached its maximum time limit. To check, go to portal.hpc.arizona.edu, click **Manage XDISK**, and look at the box next to **Duration**. If you see 300, your allocation cannot be extended further.

If your allocation is at it's limit, you will need to back up your data to external storage (e.g., a local machine, lab server, or cloud service). Once your xdisk has expired (either by expiring or through manual deletion), you can immediately create a new allocation and restore your data. Detailed xdisk information can be found on our Storage page. You may also want to look at our page on Transferring Data.

No, once an xdisk has reached its time limit it will expire. It's a good idea to start preparing for this early by making frequent backups and paying attention to xdisk expiration emails.

Once an xdisk expires, all the associated data are deleted. Deleted data are non-retrievable since HPC is not backed up. It's advised to keep frequent backups of your data on different platforms, for example a local hard drive or a cloud-based service like Google Drive, or (even better) both!

---

**General Research Data Storage (Tier 2)**

Google Drive Storage Notice

Free unlimited Google Drive storage will be going away for usage greater than 15GB, and UITS would like to migrate users off by April 2023. As a result, we will be transitioning to Amazon Web Services (AWS) S3 as a Tier 2 option. We will continue to support researchers using Google Drive during this transition.

Research Technologies in partnership with UITS is implementing an Amazon Web Services (AWS) S3 rental storage solution. This service provides researchers with an S3 account which is managed by AWS Intelligent Tiering. After 90 days of nonuse, data will be moved to Glacier. After 90 additional days, it will be moved to Deep Glacier. There will be no charge for data stored at either Glacier level, nor for any transfer charges. The data can be retrieved at any time, although it will take a while.

For information on setting up and using an S3 account, see: Tier 2 Storage

For information on Google Drive, see: Google Drive

**Long term Research Storage (Tier 3)**

HPC does not actively support long term research storage. Individual groups are responsible for managing and archiving their data. Some options for data archival include:

<table>
<thead>
<tr>
<th>Archival Service</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWS Glacier</td>
<td><a href="https://aws.amazon.com/glacier/">https://aws.amazon.com/glacier/</a></td>
</tr>
<tr>
<td>Dryad</td>
<td><a href="https://datadryad.org/">https://datadryad.org/</a></td>
</tr>
<tr>
<td>Zenodo</td>
<td><a href="https://zenodo.org/">https://zenodo.org/</a></td>
</tr>
<tr>
<td>Figshare</td>
<td><a href="https://arizona.figshare.com/">https://arizona.figshare.com/</a></td>
</tr>
<tr>
<td></td>
<td><a href="https://figshare.com/">https://figshare.com/</a></td>
</tr>
<tr>
<td>TreeBASE</td>
<td><a href="https://www.treebase.org/">https://www.treebase.org/</a></td>
</tr>
</tbody>
</table>