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>
</tbody>
</table>
Checking your Quota

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

```
Check your storage quota

(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
Users can check their storage usage in their online HPC portal under the storage tab.

```
Important /xdisk Basics

- xdisks are temporary storage allocations available to each research group. No storage on HPC is designed to archive or persist data.
- Only PIs may request, alter, or delete an allocation from the command line. Group members may be given management rights allowing them to manage their group’s xdisk through our web portal.
- The maximum lifespan of an xdisk is 300 days. Allocations cannot be extended past the 300 day maximum.
- Groups may only have one active xdisk at a time.
- When an xdisk expires, the contents are deleted.
- Once an xdisk is deleted or expires, a new one may be immediately requested.
- xdisks are not backed up.
- Users must remove their data from HPC before their xdisk expires. We strongly recommend starting the backup process early.
```

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@email.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`. PIs may request a new xdisk allocation immediately after their previous one has expired.
**Requesting xdisk Space**

**User Portal**

PIs 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.

**Command Line**

xdisk is a locally written utility for PI's to create, delete, resize, and expire (renew) xdisk allocations.

<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 [query</td>
<td>create</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 rights**

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.

Managing Group Spaces

With the introduction of communal storage, i.e. /xdisk and /groups, users will need to work with their PIs to manage these spaces. Below offers a quick guide on how to do so and answers some common questions. In each case, PI is a stand-in for the UA netid specific to each group's PI and netid is the UA netid associated with an individual group member. When following directions, make sure to make the appropriate substitutions for your individual case.

/groups/PI

/groups/PI automatically grants read/write/execute privileges to all group members and does not come with any pre-generated subdirectories. It should be relatively straightforward to use and offers 500 GB of permanent storage space to be used by the whole group.

/xdisk/PI

/xdisk is a bit more complicated because it is not automatically created with full group read/write/execute privileges. For information on requesting an allocation, see the above guide Requesting xdisk Space.

Q. Who owns our group's /xdisk?
A group's PI owns the /xdisk allocation. By default, your PI has exclusive read/write/execute privileges for the root folder /xdisk/PI.

Q. Can a PI make their /xdisk accessible to their whole group?
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 Is, 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>

Q. Where can group members store their files?
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
Q. A group member’s directory isn’t in our /xdisk, how can we add it?

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 hpc-consult@list.arizona.edu to ask one of our consultants to change the file ownership.

Q. Do we need to request an individual allocation within the /xdisk for each user in our group?

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.

---

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

Research Technologies in partnership with UITS is implementing an AWS 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>TreeBASE</td>
<td><a href="https://www.treebase.org/">https://www.treebase.org/</a></td>
</tr>
</tbody>
</table>