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)

**Important information about our data migration:** /rsgrps rental PIs have their allocations until July 2021 before their initial /xdisk allocation expires. Afterwards, they will request /xdisk like all other PIs. Data that groups want to keep should be migrated to alternate storage before that time.
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

Our shared storage for HPC data changed significantly when we purchased the new and very fast Qumulo all flash array. In summary:

- Each user's home allocation was increased from 15GB to 50GB
- The /extra filesystem was discontinued and data migrated for existing users
- The /rsgrps rental filesystem was discontinued and data migrated for existing users
- The /groups filesystem is new and provides 500GB permanent space for each PI
- The use of /xdisk has been changed in several ways as detailed below.
- Local storage on nodes is available as /tmp

Details

Permanent Storage

<table>
<thead>
<tr>
<th>Storage Location</th>
<th>Previous Policies</th>
<th>01 Mar 2020 Changes</th>
<th>29 Apr 2020 Changes</th>
<th>Beginning Spring through Fall 2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>/home/uxx/netid</td>
<td>Users were allocated 15 GB of permanent storage space</td>
<td>/home directory allocations were expanded to 50 GB of permanent storage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/extra/netid</td>
<td>Users were allocated 200 GB of permanent storage space</td>
<td>/extra was discontinued in preparation for the move to the new storage array. Users with data in /extra prior to the discontinuation retained access until the next migration date.</td>
<td>Remaining /extra directories were migrated to users' PI /xdisk allocations and are accessible under /xdisk/PI/mig2020/extra /netid.</td>
<td></td>
</tr>
<tr>
<td>/groups/PI</td>
<td>/groups/PI is new and is provided at the PI level. Research groups are allocated 500 GB of permanent storage.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Additional Storage

<table>
<thead>
<tr>
<th>Storage Location</th>
<th>Previous Policies</th>
<th>01 Mar 2020 Changes</th>
<th>29 Apr 2020 Changes</th>
<th>Beginning Spring through Fall 2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>/xdisk/netid</td>
<td>Users may request additional storage of 2000GB to 1TB. The time limit is 45 days with one additional 45 day extension. This allocation is deleted after it expires.</td>
<td></td>
<td>Temporary allocations are now requested at the PI-level and will no longer be available on a per-user basis. Up to 20 TB may be requested for up to 150 days, renewable once for up to an additional 150 days, 300 days in total.</td>
<td></td>
</tr>
<tr>
<td>/rsgrps/netid</td>
<td>Research groups who have purchased space currently have access to their rental storage under /rsgrps/netid. Storage was rented by the terabyte (TB)</td>
<td></td>
<td>Individual /rsgrps directories will be migrated over to Qumulo. These will be stored in an xdisk allocation assigned to the relevant PI. Rental spaces with more than 20 TB will have their xdisk allocation increased to accommodate their rental capacity.</td>
<td></td>
</tr>
</tbody>
</table>

Storage Migration Timeline
Checking your Quota

Command Line

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

```
[netid@login3 ~]$ uquota
```

<table>
<thead>
<tr>
<th>Directory</th>
<th>used</th>
<th>soft limit</th>
<th>hard limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>/home</td>
<td>14.8G</td>
<td>50.0G</td>
<td>50.0G</td>
</tr>
<tr>
<td>/xdisk/PI</td>
<td>2.9G</td>
<td>200.0G</td>
<td>200.0G</td>
</tr>
</tbody>
</table>

User Portal

Users can check their storage usage in their online HPC portal under the storage tab.
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. A PI can request up to 150 days initially, either via the command line or through our web portal (no paperwork necessary!). The number of days requested may be extended at any time for up to an additional 150 days (300 days in total). 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. xdisk allocations can be renewed once to extend the time period. 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.

! **xdisk allocations are not backed up.** It is the user’s responsibility to save files stored in xdisk to alternate storage locations for backup and archive purposes. See Tier 2 Storage for more information on options for backing up your data.

Requesting xdisk Space

HPC Portal – New!

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/

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><strong>Display xdisk help</strong></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><strong>View current information</strong></td>
<td>Check current allocation size, location, and expiration date.</td>
<td>$ xdisk -c query</td>
<td>$ xdisk -c query XDISK on host: ocelote-head1.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>
</tbody>
</table>

It is the user’s responsibility to save files stored in xdisk to alternate storage locations for backup and archive purposes. See Tier 2 Storage for more information on options for backing up your data.
<table>
<thead>
<tr>
<th>Create an xdisk</th>
<th>Grants a user an xdisk allocation. Unspecified arguments (i.e. size/days) will automatically be set to the system defaults. Max Size: 20 TB Max Days: 150 days, can be extended for an additional 150 days after partition created.</th>
<th>$ xdisk -c create -m [size in gb] -d [days]</th>
<th>$ 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;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extend the xdisk time</td>
<td>Before an xdisk allocation has expired, it may be extended up to a maximum of 150 additional days.</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, “-“</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: ocelote-head1.hpc.arizona.edu Your resize to 400GB was successful $ xdisk -c size -m -100 XDISK on host: ocelote-head1.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 rights to manage xdisk to a user from their group. To add a delegate PI needs to click on Manage Delegates link on the home page of the portal:

To manage xdisk right a delegate first needs to switch to the PI account on the portal. To do that click on the “Switch User” link and enter your PI’s NetID to switch the account. After that a delegate can manage xdisk using the buttons 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 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>

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 is created, either the PI may manually create a directory and grant their group member ownership or, if the root directory is group writable, the user may create one themselves. Instructions for the PI:

$ mkdir /xdisk/PI/netid
$ chown netid /xdisk/PI/netid

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.

Q. Where did my stuff go after the migration?
For users with /xdisk and /extra allocations prior to the migration, these can now be found under:
General Research Data Storage (Tier 2)

Google Drive

HPC officially supports the usage of University-Affiliated Google Drive accounts which provides free, unlimited storage to users.

- Transferring Data Between HPC and Google Drive
- Requesting a University Google Drive Account
  - Students
  - Faculty/Staff
- Setting up a Shared Research Space
  - Instructions for Group Members

Transferring Data Between HPC and Google Drive

Information on how to transfer data between HPC and Google Drive is available from our Transferring Files section in our user guide.

Google Drive Limits

An individual user's Drive is subject to a daily quota imposed by Google:

1. Users may upload up to 750 GB to their accounts per day.
2. The maximum size of an individual file is 5 TB.
3. If a single file exceeds the 750 GB daily limit, the file will transfer successfully without interruption. Subsequent files will not transfer until the quota is reset.

Requesting a University Google Drive Account

Students

Students are automatically provided with a Google Drive account.

Faculty/Staff

Faculty and staff are not automatically provided with Google Drive accounts, but can receive one following the instructions below:

1. Go to: https://account.arizona.edu/welcome and log in.
2. Select Manage your Accounts

3. Select UA Google Apps for Education

Setting up a Shared Research Space

Researchers sharing Google Drive directories with collaborators who do not have UArizona affiliation should be aware:

1. File ownership cannot be transferred from one domain to another. For example, a UArizona researcher cannot transfer file ownership to a collaborator outside the institution.
2. Researchers accessing files or directories shared with them from different institutions cannot move these from Shared With Me to My Drive. This means those researchers may not be able to transfer files to/from that directory using software such as Globus.

Research groups can create a shared lab directory in their Google Drive so they can share data with their UArizona collaborators. To set up a shared directory, we recommend:

Instructions for PIs

1. If you have not created a UArizona Google Drive account, you will need to do so by following the steps under Requesting a University Google Drive Account above.
2. Create a folder in your Google Drive account by going to https://drive.google.com/, clicking New + in the upper left corner, selecting Folder, and giving your folder a descriptive name.

3. Your new research group folder will then be in My Drive. To add members of your research group, right click on the folder, go to Share, and enter the UArizona email addresses of your collaborators. This comes with the option to adjust the permissions for each user.
4. You and your group members may now add files to your Drive!

**Instructions for Group Members**

When your PI shares a research group folder with you, it will be located in the **Shared With Me** section. To make this directory accessible to a file transfer program (e.g. Globus – see: Transferring Files), you will need to move the directory to **My Drive**. To do this, go to **Shared With Me**, locate the relevant directory, right click it, select **Move To**, and choose **My 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></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>
<tr>
<td>-------------</td>
<td>--------------------------</td>
</tr>
</tbody>
</table>