Storage Options at the University of Utah

CHPC Group Space:

CHPC currently allows CHPC PIs with sponsored research projects to buy into file storage at a price determined by the cost of the storage. This RAID-based network file system can be Samba/CIFS mounted by resources outside of CHPC or directly mounted on resources that CHPC manages. A more detailed description of this storage offering is available at https://www.chpc.utah.edu/documentation/white_papers/CHPCStoragev3.pdf. The current pricing is $150/TB for the lifetime of the hardware that is purchased with a 5 year warranty. CHPC purchases the hardware for this storage in bulk (currently 400TB at a time) and then sells it to individual groups in TB quantities. Therefore, depending on the amount of group storage space you are interested in purchasing, CHPC may have the storage to meet your needs on hand; if not it would take approximately 4-6 weeks to get the new space purchased and made operational.

For non-research applications, we use a total cost of operation pricing model, which currently is the cost mentioned above, plus the cost of operation including CHPC staff and datacenter charges, which adds ~225% to the cost mentioned above.

Archive backups of group level storage to tape quarterly is available by request for the cost of the backup tapes. We recommend that groups purchase a quantity of tapes to allow for two copies, so that backups can be alternated between the two sets. Contact CHPC for current pricing of this backup solution. Alternatives to this backup solution are discussed below, and CHPC may deprecate support for the tape back-up solution since it is becoming increasingly people-intensive and does not scale with current growth in storage needs on campus.

CHPC Archive Storage:

CHPC has an archive storage solution based around object storage. This solution uses the open-source Ceph software (http://ceph.com/), a distributed object store suite developed at UC Santa Cruz. We have an initial raw capacity of 1.15PB, with a cost of $80/TB raw space. In order to calculate the cost per TB of usable space you must consider the replication configuration. Initially, we are offering an 6+3 erasure coding configuration which results in a price of $120/TB (research subsidized price; total cost of operation price for non-research usage to be determined) of usable capacity for the 5-year lifetime of the hardware. As we currently do with our group space, we will operate this space in a condominium model by reselling this space in TB chunks.

One of the key features of the archive system is that users manage the archive directly, unlike the tape archive option. Users can move data in and out of the archive storage as needed -- they can archive milestone moments in their research, store an additional copy of crucial instrument data, or retrieve data as needed. This archive storage solution will be accessible via applications that use Amazon’s S3 API. GUI tools such as transmit (for Ma, https://www.panic.com/transmit/) as well as command-line tools such as s3cmd and rclone (https://rclone.org/, https://www.chpc.utah.edu/documentation/software/rclone.php) can be used to move the data. In addition, Globus (https://www.globus.org, https://www.chpc.utah.edu/documentation/software/globus.php) can be used to access this
space; however note that the globus Ceph plugin is a new tool that is still in development and should be treated as such.

**Storage outside of CHPC – University of Utah Google Drive Storage:**

An alternative to the CHPC administered storage options is the use of Google Drive via the University of Utah and the Google Apps for Education program, which currently provides unlimited storage. Details about this storage can be found at [http://gcloud.utah.edu/](http://gcloud.utah.edu/) -- please note the table on this site discussing the types of data for which it is applicable to use this offering.

At [https://www.chpc.utah.edu/documentation/software/rclone.php](https://www.chpc.utah.edu/documentation/software/rclone.php) CHPC provides a description of how to use the rclone tool to do data transfers and synchronizing of files between your local storage and google drive (as well as to the CHPC provided archive storage option mentioned above).

**Additional information about Backup Options**

In the above section on CHPC group space, our tape backup option is discussed. Along with this option, individual users can use the Google drive storage to manage their own backup of data in this group space. A third alternative is to purchase not only CHPC group space but also CHPC archive space, and manage your own backup between these two spaces.