Gemini FITS Data Storage Infrastructure: Data Curation Policies

Paul Hirst 2009-Oct-21

**Filesystem**

The files will be stored on a standard linux filesystem (e.g. ext3, XFS etc), on storage systems that are tolerant to at least single drive failures. Filesystems will be sized so that the summit systems store at least the previous 12 months of raw data from the telescope at which they are located and base facility systems store at least 24 months of raw data from the telescope at the site at which they are located.

**UNIX file permissions**

These apply to direct logins to the storage servers and to systems NFS mounting the storage volumes.

Files will not have world read or write permission. Files will have group read permission and all Gemini staff user accounts that require read access to the data will be added to that group. Files will not have group write permission. Files will be owned by a username that staff use to modify them when necessary (this could be dataproc).

**NFS exports**

The filesystems will be NFS exported only to Gemini workstations on which the login database is Gemini controlled and on which the UID-username mapping is consistent with other Gemini machines. The filesystems will not be NFS exported to laptop machines. NFS export to non operations machines is discouraged - HTTPS / wget will be the preferred access mechanism for general science staff workstations, though there will be cases (e.g. DAS workstations) where NFS will probably be more appropriate and this will be accommodated.

**HTTPS access**

Access to the data will be provided via HTTPS. This will be abstracted such that the user need not know the absolute path to the datafile on disk, instead simply specifying the filename or obs-id. Access will be controlled by a username / password scheme. – Though this needs to be thought out well with respect to not revealing sensitive passwords in wget command lines for example, or requiring typing a username/password each time. This will probably implemented using cookies (i.e. user authenticates once, which provides them a cookie which the server recognizes to grant subsequent access).

**Tracking of data files.**

Each storage server will run a database to keep track of the files it handles. The database will provide tracking information regarding where files are and were located, including tape and disk storage, and the dates and times of critical file handling operations. Example use cases include querying for the location on tapes and disk of a given file, querying when a file was last modified, etc.

**Archiving data to tape**

Data will be archived to tape by scripts provided as part of this system, using dedicated tape drives in the server machines also provided as part of this system. These scripts will be designed to be invoked automatically, for example via cron. The system will track which files are on which tapes. Files may be on multiple tapes as well as disks. The system will understand the concept that a tape may become unreadable and will provide means to recover from that situation, by recovering the data from other tapes and/or disks and re-writing the failed tape.

The system will initially use either LTO-3 or LTO-4 tape systems. This decision will be made based on the price-performance balance at time or purchase.

All data will be written to two separate tapes, using different tape drives. In normal operations, data from each telescope will be written to tape by the data storage systems located at the respective summit and base facility.

Of these two tapes, one tape from each site (Hawaii and Chile) will be retained locally, the other tape will be sent to the other Gemini site for storage. Both tapes will be verified as being readable before being placed in secure storage.

If a file has been updated (changed) since it was last backed up, the new version will also be backed up.

**Disk space Curation**

The system will track free space on the disks and will automatically (with configurable operator confirmation) delete files from disk which are known to be stored in at least two other locations.

**Integrity Checks**

The system will store checksums of data files it is tracking, and will provide a means to generate a list of files on disk that do not have the expected md5sum.

**GSA**

The system will store the same checksum as used for files in the Gemini Science Archive. This will allow a simple query to determine which version of a file tracked on the system corresponds to the version in the GSA (if any).