Data Handling Status

Pipeline

- Oracle Upgrade/Tuning/Stress testing [done]

Low priority items

- Add prioritization of processes within pipeline
- Improve performance of web front-end
- Improve robustness of monitoring (ping can say OK even when server is stuck)

xrootd

- All new file servers (except wain021) now in production
- New version of xroot server/redirector/client ready to go
    - Supports read/write access restrictions
        - Write access will be restricted to glast/glastraw/glastmc accounts
        - Read access will be restricted to users listed (with SLAC username) in glast.stanford.edu database
    - Fixes a number of problem with xrootd found over last few months
        - In particular duplication of files when disks become (almost) full
- Plan to submit CCB request to put this into production (yesterday)
    - Aim to have it in production by early next week.

(Note this will restrict access to data in xrootd to glast users. Most of our NFS disks remain accessible via anonymous FTP. If we care we need someone to look into how to fix this)
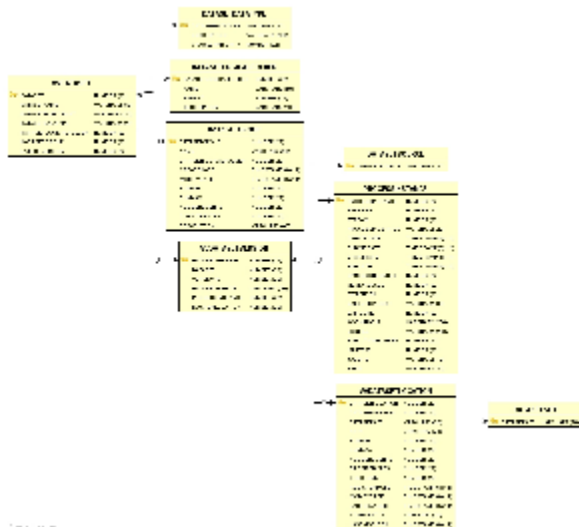
Low(er) priority items

- Turn on automatic archiving of data to tape
- Automatic taring of small files

Run Quality Database

- Designed to allow shiftees to flag runs as "good/bad/whatever"
- We have rough specification from Anders/Warren
- Karen has set up database tables, and initial web interface
    - Expect to get feedback from Anders etc in next couple of days
    - Expect details to change as we get early experience with data, so system has been designed to allow for flexibility
- Need to provide Warren with simple interface for marking runs as "ready to be reviewed" from L1Proc
- Need to link this in to data processing page for easy access by shifters
- Need to make this meta-data available to the data catalog

Data Catalog/Crawler

- Starting right before OpsSim2 L1 started registering multiple versions of datasets in the data catalog
    - The versions are designated only by a naming convention (vNNNN) which the catalog does not understand
    - Tools like the skimmer do not understand this convention and as a result skim multiple copies of the same file
- To fix this we have designed some extensions to the database scheme to support multiple versions of files
    - We will introduce some "views" to maintain backwards compatibility with existing tools
        - The views will only make most recent version visible
    - Will require changes to data catalog stored procedures, dataset registration, data crawler, L1Proc
        - This work is in progress

- Minor improvements to data crawler
  - Expect to submit CCB request this week
- Improvements to data catalog web interface
  - Delete old obsolete data
  - Speed up display of data using materialized views
  - Clearer indication of errors
  - Links to fire up WIRED on arbitrary datasets
  - Improved selection of subsets of data
- Automate loading of data into "Astro Server"
- Update skimmer to V5 backend
  - Make it possible for experts to test V6 backend
  - Skimming SVAC, GCR tuples

DataQualityMonitoring

- Many improvements to web application made by Max in last few weeks
  - Biggest outstanding problem is the memory usage with lots of runs selected
  - Production database needs to be cleaned before real data arrives

Report Application

- Pretty much done, the plots are being generated. There are some missing panels due to missing variables in the TelemetryTrending database.

Data Processing

- Top priority for Max
- It would be nice to improve performance of the main query
- The GCN notices query is to be re-written as the tables are changing
- There are a dozen JIRAs (mostly minor) to be implemented
- Add interface to Run Quality database

Miscellaneous

- Shift signup database/web application
  - Usable now, Charlotte adding
    - Ability to give up, transfer shifts
    - Ability for shift coordinate to reassign shifts
    - Support for Infrastructure, Flare Advocate Shifts
      - (Any other groups planning to use this need to let us know how they want their shifts set up)
- Continue robustness of Web Server infrastructure
- Improved monitoring of all aspects of infrastructure (interim system in place, being migrated to nagios)
- Improve "How to Fix" documentation
- Add "restart" functionality and improve usability of ServerMonitoring application
- Confluence/JIRA upgrades