

Data Server needs from DC2 Perspective

- Anticipate that many of the DC2 requirements will be a subset of the I&T, ISOC and Science Tools needs. Some differences:
 - Slight differences in usage (users possibly more likely to want to download the entire sky/dataset).
 - Timescales!! (needs to happen soon)

Data Challenge provides an opportunity to test and get feedback on the data servers.

Experience from DC1

- A Success!!
- GSSC server provided FITS data in a form suitable for analysis with the science tools. Well documented and easy to select data based on region of sky or time period.
- SLAC server provided access to root and fits data formats, flexible interface to apply cuts and subselections to the data.
- SLAC also had a manually created page to serve AllGamma and BackGndAvg merit tuples.
- Small niggle: A majority of post-kickoff DC1 users wanted to perform all sky source/GRB searches. Neither server easily provided this so many people bypassed the servers altogether and just grabbed the files directly.

Data Challenge Needs

- Two somewhat separate roles:
- Pre-kickoff
 - Support the development of classification/background rejection analyses.
- Post-kickoff
 - Provide access to the simulated sky data. At a minimum we need to serve FT1 and FT2 fits files.

Analysis/Reconstruction (1)

- MC AllGamma and BackGndAvg root tuples:
- Used for classification analyses, background studies, trigger rates etc.
 - Would be nice to see the mc merit tuples as they are created (so that we don't need to be on the small, elite mailing list of people notified about a new dataset).
 - Webpage/access dynamically generated from a DB.
 - Prepruned, preconcatanated tuples (for fast access).
 - Also allow user specified cuts, perhaps also allow selection of banks of variables (FT*, PT* etc).
 - Access to the log files (c.f. system tests, so that a user can find out how the data were generated).

Analysis/Reconstruction (2)

- Access to full mc, digi and recon root trees.
 - Used to get events, filtered by event number or cuts, for viewing in event viewer
 - Rerun recon on subset of data? (not so obvious that it would not be easier to start from scratch)
- How is this going to relate to the event display?
 - Pruned root files, then run Gleam and Fred?
 - Directly via a server (some resistance to this from offsite users).
 - An aside, is event reproducibility still an issue? We don't store the full MC tree.

Some general comments

- It would be good to merge (or move close together) several of the current SLAC services.
 - i.e. download site for pruned, concat merit tuples should be linked to the data peeler.
- The cut/filter interface could be made a little easier to use.
 - Perhaps have drop down list of standard cuts.
 - More documentation.
- Should be easy to navigate to more information about the GlastRelease version being used: systests, commentary, release diffs etc.

Post-kickoff

- Provide access to the simulated sky and spacecraft pointing data.
 - Need an easy, fast method to download entire dataset (or filtered by largish preset time range).
- SLAC server needs to provide access to all the data required for a science analysis (i.e. access to FT2 information).
 - One way to do this, would be to change the definition of the exposure tree so that it really does describe exposure, then provide a root file with the two trees or convert to FT1 and FT2 fits format.
- Event Display
 - For example, someone may wish to look at all events above 10 GeV from a GRB.