LAT Data Server

- **Serve what?**
  - **Glibly stated – all the 'science' output files produced by the pipeline**
  - **From I&T, generic MCs and Level 1 operations**
    - **Real data – instrument analysis (eg runs taken with different hardware configurations, etc) – I&T, ISOC**
    - **MC & Level 1 – astronomy analyses**
      - **Include pointing, livetime histories?**
      - **Event selection by location on sky, time, energy mainly**
    - **Perhaps connected to FRED**
      - **We envisage a pool of servers running Gleam with remote internet access. Data server technology presumably needed for finding the desired events.**
  - **To whom?**
    - **LATers anywhere**
      - **Not just at SLAC**
      - **Though data transfer latencies for SLAC users will be minimal, since data will just be moving around local nfs disks.**
  - **Prototype Server for DC1**
    - **Written by Navid (see his talk)**
    - **Allowed for event selection on Merit ntuples by user supplied TCut on all allowed variables; or via web form for standard cut variables**
    - **Found all the events that pass and created output file for user to fetch**
    - **Could find "Tree" events by (run, event) ID**
    - **Was told by hand what directories to search**

Issues

- Pipeline database is processing-centric and only contains file information
  - Events are known to be in particular files by run ID, but nothing else is know about them
  - Do we need to recast or mirror the db to enhance access for the server?
- Do we need to index the events in some fashion to get good performance?
  - And only for 'astronomical' analysis events?
  - What is good performance? As fast as reasonable? As we can afford? Spec'ed out by SSC for L1 server; what about instrument analysis?
  - Are the promises of random access to Root files + PROOF just hot air? Can we do parallel processing?
- C++ API?
  - For local to SLAC nfs access to files
  - eg pipelinedataSets.cxx class for access to Root files
  - Any possibility of remote access (à la xrootd)?
- Output targeted for ScienceTools in FITS
- Access to files generated not-at-SLAC
  - Quite related to pipeline issues of how to maintain the bookkeeping for MC (at least) generated on LAT CPU farms
- What about Level 2 products??