Requirements and constraints for GLAST Mirrorsite

T. Kozanecki, J.P. LeFevre
December 1, 2002

Document version: 1.0
Document status: Draft
Keywords: GLAST, database, mirror site

1 Introduction

GLAST will make its reconstructed data available through two sites, the LAT Instrument Operations Center (LIOC) at SLAC (Stanford) and the Science Support Center (SSC) at GSFC/NASA. In addition plans exist to make the data available at European mirror sites, e.g. one site in France and possibly one in Italy. All data sites will become accessible to guest investigators.

Software accessing the databases will run on servers of SSC through a web interface. Users will extract the events of interest from the database on the SSC's server and download the files to the user's server. Users will download from the SSC's web interface a suite of tools, which will run on the user's server. The supported systems are Windows and UNIX. It is expected that the access to data and tools is the same at SSC and SLAC, and at the mirror sites.

This note formulates the present view of the data storage and access, mainly with the aim to clarify the architecture and get feedback from groups, which have experience with providing access to large databases.

2 Event size and data volume

The expected size of reconstructed events is 200 Bytes/event. Most events will be proton events; about 10% of the events are gammas and are of interest for the final science analysis. The expected number of events and required storage are summarized in the following table.

Time period	Events	Storage	Photons	Storage
1 day	3 M	0.6 GB	0.3 M	0.06 GB
1 year	1000 M	200 GB	100 M	20 GB
10 years	10,000 M	2 TB	1000 M	200 GB

The mirror sites will provide photon data to guest investigators and team members.

Members of the instrument team may, however, for diagnostic purposes, want to have in addition access to (a subset of) the full data sample, or to the raw data. These data are accessible to team members at the central distribution sites (LIOC, SSC). It is therefore at present not planned to make them available at the mirror site.

3 Data format and version management

3.1 Data format

The data will be stored in FITS format, following HEASARC standard. FITS files are flat data files, and the format is well known to the astrophhysics community. Tools for convertion to ROOT files should be made available.

The file size has not yet been considered. It will be determined by limitations of the hardware and peripheral system (e.g. at CCIN2P3 the file size is limited to 200MB), and by the optimization of the selection process.

The grouping of the data into pixels ('tesselization') according to position coordinates, or according to time of recording, was considered; no conclusion has been reached. Whereas the position ordering is useful for most searches, it requires frequent updates of the database, or the use of an index related to the position. The time ordering simplifies the extension of the database when new data arrive. But as a consequence most searches have to traverse all data. One could consider a mixture: keep the data in time order, but set up an index, which corresponds to a pixel in position (granularity ~ fraction of the resolution of the instrument). Some database systems allow geographical indexing (PostgreSQL, PostGiS, Oracle Spatial).

3.2 Version management

A change history must be kept, which documents any change to the data available at the mirror site. This change history keeps track of the addition of data or any upgrade of data.

The data may be upgraded occasionally, for example because reconstruction programs, calibrations or selection cuts were improved. To reflect such changes, each event carries identifiers for the versions of calibration and reconstruction. In case that the event samples are updated, only one version of the data will be made available at the mirror site.

4 Architecture

The architecture of the system has not yet been detailed, however the rough layout could be as shown in Figure 1:

1. Web interface for user requests: This interface is also responsible for authentication and authorization (accounting is not foreseen).
2. Storage of the data and search engine, depending on the data organization, supported by an additional database for meta data.
3. Output of the data for user download from an ftp server.

4.1 General requirements

1. The three domains described above should be as independent of each other as is reasonable. This will ease future upgrades.
2. Only freely available software components should be used.
3. The data processing chain must be independent from the data access chain.
4. Most of the content of the database should be publicly visible.

5. However, actual access to the data may be restricted : the characteristics of the dataset can be displayed, but the corresponding data cannot be downloaded. Access to data and the type of data available for download should be controlled via password. The access policy is defined by the GLAST team.

4.2 Requirements on the user interface

1. A set of parameters is presented to select a subset of the data in the database (region in sky, data quality, energy band, etc). Default values or menus may be presented as examples or to provide possible choices.
2. An estimate of the size of the selected data subset should be provided.
3. Data can be selected, restricting increasingly the choices.
4. The selection can be revised at any point. At each stage the client can go back and revise his choices.
5. The last choice is remembered at logoff and is recorded for later reuse.
6. Once the list of selection criteria is ready, the transfer of the corresponding data from the storage system can be requested.
7. The history of changes must be accessible (see Section 3.2).

4.3 Requirements on data download

1. The files with the user selected data are identified by a unique name, which is defined by the download system.
2. In addition to the data the used selection criteria are downloaded.
3. The data should be available on the ftp server for at least several days.

4.4 Requirements on the data search

The requirements on access speed and number of requests handled per time interval are documented in Ref. 2. They depend on the amount of data to be digested. (extract few typical numbers?)

4.5 Resource considerations

The system must make cost-effective use of resources and must be easy to maintain (0.3 FTE ?). The following table lists some numbers for the h/w of a disk system for 1TB storage.

Device	**kEuro (~ k$)**	**Comment**
6 disks 160 GigaByte	11	B. Schaefer $300/disk
Chassis for disks	0.6	
Server PC	3.8	
Total	15.4	

Today's storage capacity: 1 disk => 160 GigaByte (may become 400 GB in 1-2 years)

5 Possible implementations

For the three architectural domains the following implementations were recommended or considered.

5.1 User interface on web

Software tools:

Java Server Pages (JSP, used by [1], scalable, clear separation between GUI and processing)
PHP (for small projects ok, but difficult to maintain, not very scalable)
Tomcat (used by [1], interface between GUI and the database, e.g. MySQL)

5.2 Storage and fast selection of data

Store meta date in MySQL, PostgreSQL (allows geographical index, PostGiS) or Oracle (Oracle Spatial, allows geographical index)

Fast selection: Storage on parallel PCs, e.g. the Beowulf system, see Figure 2, .

a) store data in parallel
b) store data time ordered
c) store data ordered by location

Dedicated libraries for parallel machines: MPI/PVM.

5.3 Data transfer to user

Transfer data to the ftp server, from where they can be downloaded. (control who is authorized to download what?)

6 Communication between central repository and mirror site

The following questions have to be addressed:

1. How are the sites kept in sync?
2. Which tools to use for data transfer and code transfer
3. How is the update of data and code managed
4. Security questions

For these questions we expect to profit from the tools and methods, which are developed for the various GRID projects.
The Linux program 'rsync' for synchronizing different sites should be considered.

7 References

(to be updated and re-ordered)

1. EROS database Gateway, http://edbg.in2p3.fr:8080/index.html
2. Science Tools Databases requirements, accessible from Ref. 3
3. Review of Standard Analysis Environment for LAT Data September 2002 http://www-glast.slac.stanford.edu/ScienceTools/reviews/sept02/