Ideas to analyze the 99-00 GLAST Beam Test Data

José A. Hernando, Wilko Kröger

January 24th 2000

This note suggests guidelines for the analysis of the GLAST test beam data taken during the winter 99-00 at SLAC. It covers the following subjects: data format, how to make the data available for everyone, a collection of public programs to perform analysis, and a list of some analysis topics to be addressed. We assume two times scales: a short one (few weeks), when very preliminary data need to be presented to some agencies, i.e NASA, and a medium term one (few months) where preliminary but reliable results should be obtained. The note is divided into three sections: 1) Data stripping, 2) Software infrastructure 3) Analysis topics. This note is indicative and should be consider a proposal for an open discussion.

1 Data Stripping

The raw data should be filtered, pre-processed and stored into a well designed accessible structure - that for old habits, we call DST-. The following is a list of tasks:

- Filtering of the data.

1. Synchronization of tracker, calorimeter, ACD and tagger data. (Some of the data don't necessarily need the tagger, only the synchronization between Cal/Tkr/ACD might be needed).
2. Select "good events" that should go into the DST's. Definition of the criteria for tracker, Cal, ACD, tagger if needed. Preliminary study to define this criteria needs to be performed.
3. What about house keeping data ?
4. Adding reconstructed objects (i.e.: tracks) to the DST's (see later in this note).

- Definition of the DST format. Richard presented at the last software meeting the structure of a Root-Tree in which the test beam data is been stored by Daniel Flath and people of his group. The ROOT-tree is itself a well designed storage structure for our data, with the possibility to add new branches.
- Processing of the raw data to produce the DST's. A group, with large computing capabilities, should be responsible for the processing of the DATA. The DST's should be easily accessible to everyone involved in the test beam.
- The information of the DST should be available via web. It should contain information obtained during processing (e.g.: no tagger data) and the run log information. The group leading the production should be responsible for its publication.
- There should be a possibility to re-run the data stripping, in case something went wrong or someone wants to apply different selection criteria.

2 Infrastructure

For Infrastructure we understand: a) a collection of software programs (to be created or modified from existing ones) in order to reconstruct and/or analyze the data; b) a collection of programs that will allow the user an easy access and manipulation of the data.

The a) part has two mayor topics:

1. **Monte Carlo Simulations**. We assume that the GLASTsim program (B. Atwood, T. Burnett) is used to generate Monte Carlo data for the GLAST test beam tower. Eduardo, Sawyer and Mario have been working defining the test-beam geometry within GLASTsim. The actual output for the GLASTsim raw data are *.irf* files. We suggest that there should be a program to convert the *.irf* files into the DST format used for the real data.

2. **Reconstruction/Analysis program**. The RECON program should read the DST's (real or MonteCarlo) and generate high level objects (i.e. tracks, calorimeter clusters). The reconstruction code should be flexible and modifiable for the user. One should realize that there are different analysis levels, some require only high level processed data (i.e PSF studies), and others (i.e tracking efficiencies) require to modify the reconstruction code. The output of the RECON program should be a well defined accessible structure. In particular, it should be possible to add the new objects to the initial DST. Instead of adding to the initial DST one also could generate a new DST that includes the reconstructed objects and a subset of the original DST data. We see two possible, not exclusive, ways to reconstruct the test beam data:

   - *Use GLASTsim.* Up to the moment, GLASTsim has been used to generate/reconstruct data. The reconstruction of GLASTsim uses *.irf* data files and it produces ASCII ntuples as output. In order to accommodate the above commented requirements, the following changes need to be implemented: a) read the DST's - the program discussed in the MonteCarlo section, could be bidirectional and convert *.irf files* $\leftrightarrow$ ROOT-tree DST's -; b) the possibility of the users to modify the reconstruction as they need (i.e. masking layers into the track reconstruction for tracking efficiency studies) and c) A more flexible output, similar to a ROOT-tree DST's; for example, the output should include the list of reconstructed tracks.

   - *Use a RECON stand-alone program.* This program will read real and simulated data DST's and perform a stand-alone reconstruction. The different reconstruction code could be provided by the different groups (ACD, Tracker, Calorimeter). This program should have the possibility to access and modify the reconstruction code for the different sub detectors or even exchange it with a different version. The output of the RECON program should be stored. One possibility is to add new branches with reconstructed data to the original ROOT-tree. Jose, Wilko and Dan have created an stand-alone RECON program, based on GLASTsim. This programs contains a simple framework to read the ROOT-tree files, hooks to incorporate reconstruction code for the different sub detectors, and some analysis code. It includes the tracker reconstruction and the GLASTsim 3D graphics display. This program could be used as initial structure for RECON. The calorimeter, and ACD code could be added.

We think that separating the RECON from GLASTsim for the test-beam could be of great convenience to develop the analysis, and a learning lesson for the future of GLASTsim.

This approach opens the question about a general frame work. It would be desirable to design a frame work that persists longer then the beam test analysis and would be part of the glast software. However, because of the short time frame for the analysis we doubt that this is feasible, but for the long term one could migrate/evolve towards a GLAST frame work.

The b) tasks requires a collection of programs to make the data easy accessible for everyone. These are some of the topics to be discuss:

1. A Framework to read-write the DST's (ROOT-tree). The adoption of a ROOT-tree as a main storage structure should not imply the use of ROOT as data analysis package. For that, we should be able to contemplate the possibility of converting the ROOT-tree DST's into HBOOK, ASCII ntuples or IDL structures, which is more convenient for the different groups or people who do analysis.

2. All the programs used to simulate/reconstruct the data: GLASTsim for simulation; GLASTsim or RECON for the reconstruction, should be available for everyone.

3. A "HOWTO" should be provided (maybe via web), of how to use the DST read-write Framework; as well as how to use GLASTsim or the RECON programs.

For each of the task mentioned in this section, we need to assign group of people responsible, and a time-scale.

To summarize the main tasks:

1. *A framework to read-write the DST's (suggested: Richad and Daniel ROOT-trees)*
2. *A simulation program (suggested: GLASTsim that produces a ROOT-tree).*
3. *A RECON program (that include the reconstruction code or tools of Tracker, Cal, ACD) and produces an organized output (suggested: a stand-alone program that reads the ROOT-tree and add branches of new data; and/or: modification of GLASTsim to read the DST and produce an flexible output).*

3 Analysis Topics

This is a partial list of analysis topics to be address:

- Tracker
    1. Tracker alignment.
    2. Tracker resolution, performed with electrons and clean hadrons.
    3. Tracker efficiency studies for different angles.
    4. Tracker efficiency as a function of threshold.
    5. Time-over-threshold (ToT) performance.
    6. Trigger efficiency of the tracker (self-triggering rate).
    7. Could the ToT be used to separate electrons from gammas?
    8. Performance of the tracking reconstruction.
    9. Data/MC comparison: distributions: hits distributions by plane, number of track reconstructed, missed hits, etc.

10. Tracker reconstruction performance for SuperGLAST layers.
11. Energy determination with tracker information.
12. PSF as a function of different energies (tagger) and angles.
13. PSF specific studies for the SuperGLAST layers.

- Calorimeter
- Calorimeter-Tracker
  1. PSF as a function of the measured energy.
  2. Energy correction.
- ACD
- GLASTsim tower
  1. Hadron rejection level.
  2. Data/MC comparison of hadron interactions.

Redundancy in some of the previous analysis should be encouraged.

Time Scale:

Some very preliminary results could be obtained within few weeks with the tools already developed by some groups: Berrie's ntuples, Robert's tracker efficiency program, etc. That should be in principle enough for the NASA inquires. A list of plots to be presented should be made and possible candidates to provide these plots found.

The medium term plan should contemplate the points covered in this note. A group should be assigned to ecah of the topics and a time-scale should be specified. We should be able to publish the test beam results by summer 2000.