Data Challenge Planning Overview

• Purpose
• Boundary conditions
• Plan:
  – scope and progression of data challenges
  – schedule summary
  – implementation

This plan was generated in meetings with Richard Dubois, Seth Digel, Bill Atwood, and Steve Ritz. It has been discussed by the Analysis Group and the SSAC.
Purposes of the Data Challenges

• “End-to-end” testing of analysis software.
  – define the ends
  – define the tests (what is success?)
• Familiarize team with data content, formats, tools and realistic details of analysis issues (both instrumental and astrophysical).
• Develop additional methods for analyzing GLAST data, encouraging alternatives that fit within the existing framework.
• Provide feedback to the SAS group on what works and what is missing from the data formats and tools.
• Uncover systematic effects in reconstruction and analysis.

Support readiness by launch time to do all first-year science.
Still much to do on GLEAM to be ready:
- geometry review
- underlying physics review
- embed onboard filter
- other infrastructure (compilers, package versions, etc.
- Check, check, check everything!!
- finish implementing remaining gamma source fluxes

Still much to do on instrument analysis to be ready:
- background rejection
- performance evaluation and parameterization
- See Bill Atwood’s talk

Still much to do on science tools to be ready
- See Seth Digel’s talk tomorrow

Lots of other work demanding attention:
- Reviews…
- EM support
- Other calibration planning and development
- Construction, Integration, and Test planning and execution
Approach

• Use the September collaboration meeting as the start of the Data Challenges.
• Walk before running: design a progression of studies.
  – DC1. Modest goals to work out problems. Mainly use existing recon tools to explore where improvements could be made. Contains most essential features of a data challenge (see following slides).
  – DC2. More ambitious science goals. Encourage further development within the existing structure, based on lessons from DC1.
  – DC3. Support for flight science production.
Preparatory Work

- Preparation complete by September collaboration meeting:
  - complete geometry review
  - preparatory work meeting: THIS MEETING
  - complete embedding, verification of filter, evaluate: July
  - Simulation hard freeze date – TBA (needed soon!).
  - first new background rejection/performance analysis July & August
  - sky model fluxes implementation complete September 1

  - At Collaboration meeting:
    - first, PRELIMINARY instrument response functions presented
    - DC1 kickoff (see following slide)
    - small workshop on using tools for team
Data Challenge Progression

• **DC1**
  – modest goals:
    • 1 simulated day all-sky survey simulation (3M bkgd+gamma events to ground, => 400M generated events)
    • find flaring AGN, a GRB
    • single-day point source sensitivity. daily quicklook analysis development.
    • recognize simple hardware problem(s)
    • a few physics surprises
    • exercise:
      – exposure, orbit/attitude handling, data processing pipeline components, analysis tools
    • use existing recon, bkgd rejection and instrument response to show the problem areas that need improvement. secondary goal (not required) is to prototype improvements
  – baseline schedule:
    • Sept-Oct startup problems resolution.
    • Nov-Dec high-level tools beta testing. Finalize instrument response functions.
    • Dec 15 high-level tools release, workshop.
    • mid-January: interim reports (vrvs or face-to-face)
    • Feb 2004 closeout, and plan for DC2 (see following slide).
    • Then, break for I&T prep. Use the time for fixing problems learned in DC1, software advances, etc.
DC Progression

• DC2
  – more sophisticated goals:
    • 1 simulated month all-sky survey simulation (100M bkgd+gamma events post-filter. Method TBD) PLUS 1 simulated year of gammas
    • find AGN, bursts, pulsars
    • produce a toy 1-month catalog
    • detailed point source sensitivity and localization analyses
    • recognize more subtle hardware problems
    • a few more physics surprises
    • exercise:
      – exposure, data processing pipeline, analysis tools, quicklook. benchmark processing times, data volume, etc. connect to SSC.
    • use updated recon, bkgd rejection and instrument response to show the problem areas that need work. encourage improvements
  – tentative schedule:
    • freeze software version July 04. start generation in September
    • start DC2 October 2004 (beam test analysis ~complete)
    • Dec 15 2004 midterm reports milestone
    • Feb 2005 closeout, and plan for DC3 (see following slide).
DC Progression

• “DC3” – Flight Data Challenge!
  – physics groups will have been working on detailed analyses, based on experience with DC2, during the previous year.
  – main goal is realism to support running experiment analysis:
    • 1 full simulated year of data (methods TBD)
      – exercise everything: format data as it comes into the IOC. also confirm data storage, backup, processing speed. will be the reference sample for 1st year data analysis.
    • connect to SSC
    • demonstrate point source sensitivity and localization
    • recognize a few very subtle hardware problems. recognize a few realistic daily hardware problems -> feed to IOC and FSW.
    • physics surprises
    • use updated recon, bkgd rejection and instrument response. this will be our initial science performance. by this time, physics analysis groups should be up and running.
  – schedule:
    • freeze software version Feb 06. This is the first year flight version! start generation in May.
    • physics groups working
    • launch.
Notes:
1. The LAT schedule is being rebaselined.
2. The date of the beam test is not yet fixed. It should be complete, and the data analyzed, before the simulation freeze to incorporate the tuning into DC3.
Implementation Suggestions

• A small organizing committee (~3-5 people) should be put into place by the PI for DC1. Each person on the committee should have a clearly defined set of responsibilities for coordinating and overseeing the components of the work.

• The organizing committee should have a well-defined, prioritized list of plots and other results to be produced by data challenge participants.

• Review lessons from DC1, and re-evaluate scope and schedule for DC2 (and DC3) as appropriate.
Needed results from this meeting

- Checklist on remaining work.
  - dates, people responsible
- Sanity check on schedule.
- Fill in the holes! Especially need help checking, debugging infrastructure; thinking things through.
- Many other interesting SAS issues to discuss; however, for the data challenge to be possible, we must stay on topic at this meeting.

- Eye on the ball.