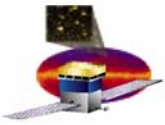


## LAT Engineering Meeting

### Findings and Recommendations of the Ad Hoc Committee on End-to- End (E2E) Testing

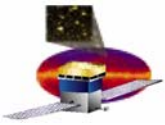
Michael Huffer, [mehsys@slac.stanford.edu](mailto:mehsys@slac.stanford.edu)  
Stanford Linear Accelerator Center



# Outline

---

- Committee
  - Charge and Scope
  - Composition
  - History
  - Findings (underlying assumptions)
- Tests
  - Power sequencing, booting, and configuration
  - L1 Trigger
  - Event transport
  - L3 Trigger (filter)
  - Housekeeping and monitoring
- Recommendations



# Charge

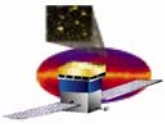
---

***To ensure success, it is necessary for the GLAST LAT Project to have detailed and carefully-designed end-to-end tests and procedures for the LAT. Plans already exist for the subsystems to verify the Level III and lower requirements, and for LAT scientific verification and calibration, however end-to-end tests cover a broader system-level scope. The charge to this group is to ensure that the LAT Comprehensive Performance Test (as indicated in MD-1312-01) adequately verifies the detector through spacecraft science data interface and to define the objectives and requirements for LAT through IOC ground end to end tests. The resulting plan should include enough detail for the I&T and SE groups to produce from it a detailed and complete LAT test procedures document.***

***Because of the inherently broad nature of this plan, there will be some overlap with the content of the existing test plans. It is not necessary to spend significant effort determining which test should be in which plan, or to determine whether a test is a functional test or a performance test; rather it is much more important to determine that there are no missing elements, and to provide that comprehensive analysis in a single document. Where appropriate, it should be shown how the individual subsystem tests tie together at the LAT level. To avoid duplication, existing tests in other documents should be referenced explicitly.***

***In addition to providing a component of the formal documentation of the Level II requirements verification, the end-to-end test plan is the document that will be consulted when someone asks, for example, "How do we know the data system doesn't hang, or corrupt the data, under all expected operating conditions?"***

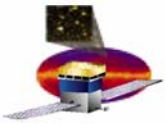
***We would like a preliminary outline of the report by XXXX (TBR), along with any feedback on the charge, and a first draft by XXXX (TBR). At that point, we will assess how best to proceed, based on your results. Our intention is to have the plan complete and carefully reviewed by members of the team and a few outside experts, and to place the plan under project control, by the end of February 2004 (TBR).***



# Scope

---

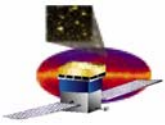
- Committee's focus (with Peter's concurrence) limited to *Data Handling Systems*
  - E2E defined: "from detector digital outputs to instrument's output telemetry"
- Why?
  - Charge of testing entire instrument much too broad to provide any meaningful input to the experiment
  - Matches best our interpretation of charge
  - Matches well with expertise of committee
  - Testing outside this scope is already well defined and enumerated by both individual subsystem test plans and calibration tests within the SVAC
  - T & DF depends on correct operation of mechanical/thermal/detector systems (Data handling tests *implicitly* test these systems)
- The scope of tests recommended by the committee are *explicitly* meant to be necessary, but in no way asserted to be *sufficient*



# Membership

---

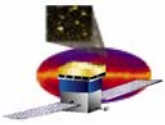
- Subsystems
  - Bill Atwood, UCSC
  - Neil Johnson, NRL
- Science
  - Steve Ritz, GSFC
- Systems Engineering
  - Tom Leisgang, SLAC
- Integration and Test
  - Eduardo do Couto e Silva, SLAC
- T & DF
  - Jim Russell, SLAC
  - Mike Huffer, SLAC



# History and Methodology

---

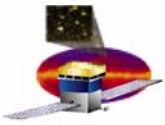
- Ten (2 hour) meetings from September 03, through March 04
- Much of the work accomplished off-line through “homework” assignments
- Established web site for archival purposes
- Work divided in three phases:
  - Acquisition of necessary background information
  - Analysis of data and findings
  - Consolidation of these findings into a report



# Assumptions

---

- Engineering model tests already performed and leading up to integration provide an adequate base for implementing the committee's recommendations
- Subsystems adequately tested at the unit level and further...
  - These tests and calibration procedures are available to recommended tests
- Limited performance tests can be derived by their full performance equivalents

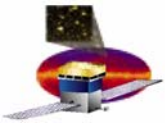


# Tests (general considerations)

---

- A test which does not pass *stops* the integration process
  - Minimize the number of tests
    - These are *functional* tests, *not* science performance tests
  - Specify clear and precise pass criteria
  - Require minimum (or no) off-line analysis
  - Minimize statistics necessary to pass judgment
  - Allow room for iteration (durations should be reasonable)
- Each of the recommended tests were characterized by their:
  - description
  - rationale
  - pass criteria
  - schedule phasing (when and how often)
  - accumulation of necessary amount of statistics
- Durations, phasing, and statistics are meant to be *suggestive*

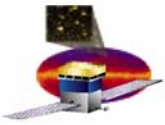




# Power Sequencing

---

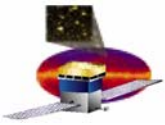
- Robust and reliable operation is a necessary prerequisite to *all* tests
  - Short-circuiting these requirements is ultimately short sighted
- Its important that both Primary and Redundant systems be tested equally well
- There is an inherent risk in power cycling the instrument
  - Careful consideration must be paid to how often this process is done
- Power-up tests are tied to SBC booting and fall into two categories:
  - Cold boot:
    - No power applied to the LAT
  - Warm boot:
    - SIU and LAT feeds powered on, but not necessarily remainder of LAT
- Both cold and warm boots can be exercised on the testbed
  - useful to iron out procedural and operational problems, however...
  - does not eliminate need to conduct these tests on the flight article
- Success criteria for power-up tests is determined by...
  - successful operation of primary communication channels (C/R and 1553)
  - monitoring appropriate quantities on housekeeping stream (voltages, currents, etc)



# Configuration

---

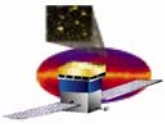
- Involves correct loading of all registers not already configured during power sequencing (mostly all ~100K detector Front-End registers)
- Robust and reliable operation is a necessary prerequisite to all subsequent tests
  - Short-circuiting these requirements is ultimately short-sighted
- Its important that both Primary and Redundant systems be tested equally well
- Success determined by reading back and matching against expectations
- Test may be commissioned, but not fully exercised, using the test-bed
  - No Front-End Register emulation
  - Mitigated to a large degree by subsystem unit testing
    - Underscores importance of unit testing by subsystems...



# Trigger and Unit testing

---

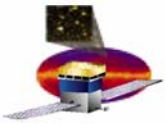
- LAT is differentiated from other space-based experiments because it is a triggered instrument
  - Science requirements depend on the correct and understandable operation of the trigger
- Trigger is distributed across the instrument and responsibility for implementation crosses subsystem boundaries...
  - The tower (Calorimeter, Tracker, and Electronics)
  - The Veto system (ACD)
  - Signal Transport and decision (Electronics)
  - Trigger software (FSW?)
  - Filter (FSW)
  - Simulation (Off-line)
- Committee could not identify a specific administrative entity (or individuals) responsible for:
  - Definition and testing of trigger *unit* requirements
  - Coordination of those trigger activities which do cross subsystem boundaries
  - Definition and implementation of a trigger *system* test plan
- Unit testing (if it existed) would need to assure:
  - Correct operation of the GEM
  - Understanding of the relative timing offsets and jitters of the trigger request lines with respect to a well understood T0



# Trigger Testing

---

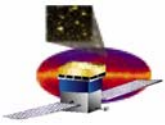
- Defined as the integration of multiple towers and the ACD
- Two categories of trigger tests:
  - **Timing and control signals:** Ensure that fast trigger signals delivered both to and from the subsystems arrive at the expected relative times with adequate margin to assure stable system
    - Enough cosmic ray data to understand and characterize timing at the 1 % level
    - Testbed can used to iron out procedure and operation
    - “One time tests (after instrument complete)”
      - Once during thermo-vac to verify behavior over full temperature range
      - Once on orbit during initial check-out
  - **Data:** Ensure the trigger’s contribution to the event is correct
    - Depend on good understanding and characterization of timing & control signals
    - Performed at nominal settings
    - On flight article need high rates for good statistics
    - Testbed also used to precisely understand “edge conditions”
    - May overlap with general data taking tests (see below)



## Some wrinkles...

---

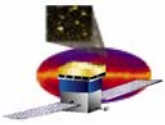
- CNO and CAL (high) signals are hard to find in cosmics...
  - Need some good ideas...
    - Testbed requires understanding modeling assumptions
    - Best alternative seemed to be solicited charge inject (calibration strobe), but...
      - Requires establishing common (calibration) timing across towers and ACD
      - Understanding limitations due to differences between signal shape of injected charge versus “real particles”
- Understanding false triggering is very important!
  - Look for, and use signatures in acquired data:
    - Appearance of galloping trigger rate
    - Similarity of adjacent (in time) events
    - Similarity in trigger patterns in adjacent events
    - Non-Poisson distribution in arrival of cosmic data



# Event Tests (general considerations)

---

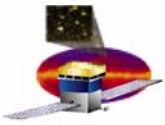
- A calibrated detector is assumed
  - Calibration results will have an impact on analysis of test results in terms of reference histograms, comparisons, trending, etc...
- The Event Filter is *not* required
  - Good. Decouples filter delivery from tests
  - In some cases (tests) data volume exceeds available output bandwidth
  - These cases require simple (real-time) event selection
    - Effort involved in developing this software is not onerous
- “Nominal” rate testing and Cosmic Ray (CR) test discussed in more detail
  - Why?
    - Nominal rate testing is both necessary and *hard*
    - CR testing is used as the base for almost all other tests by varying the:
      - LAT environment
      - configuration
    - Exception is tests which deliberately introduce errors (important!). For these type of tests
      - Testbed offers most flexibility
      - Dataflow system has built-in mechanisms (use!)



# Nominal rate testing

---

- The non-vetoed orbit average trigger rate is routinely 1 to 5 KHZ , with ...
  - peak rates a factor of 2 to 4 higher, consequently the committee prefers...
  - the term: “nominal rate” (routine) to “high rate” testing (anomalous)
- These are the tests where we answer the key charge to the committee: “How do we know the T & DF system doesn’t hang or corrupt event data under all expected operating conditions?”
- Not easy! Requires:
  - T & DF system with sufficient bandwidth and functionality to support operation in this regime
  - Varying event rate, distribution, and size without affecting physics
- Not so difficult in the test-bed, but how to extrapolate to flight article?

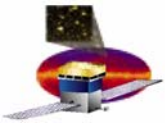


# How?

---

- Initial suggestion:
  - Lower tracker thresholds and look for cosmic induced muons
  - Demonstrate that the signature of these events is independent of trigger rate
    - Expected event rate
    - Angular distribution
    - CAL deposition patterns
    - Noise occupancy
- Unfortunately not an effective test
  - Trigger timing varies with threshold
  - Ditto for tracker noise occupancy

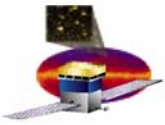




# Two simple alternatives

---

1. Fix nominal thresholds, trigger normally, but...
  - Artificially turn on layers and use expected noise to generate rate
    - Will introduce small change in accepted angular and energy distributions
    - Can be corrected with an *analysis* not a *trigger* selection
2. Fix nominal thresholds, trigger normally, but...
  - Use the so-called “solicited condition” as a trigger
    - Simply an “on demand” trigger under program (software) control
    - Allows for the introduction of a configurable background
      - rate (up to 100 KHZ is realizable)
      - time distribution (for example Poisson)
  - The committee recommends two (2) as the standard nominal rate tool
    - Less invasive
    - More flexibility
  - However, the two techniques are independent and both could be developed and played off against one another in order to validate each technique



# Two other methods considered...

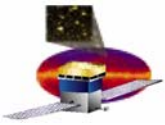
---

## 1. Airplane test

- Aircraft altitudes provide a real particle flux, over all angles, at nominal rate
- While attractive (for these reasons alone), after due consideration, this method was dismissed:
  - Not sufficient time in air to exercise completely the proposed test suite
  - No practical way to iterate
  - Alternatives *do* exist which satisfy the scope of proposed testing (see below)

## 2. Van de Graff (VDG) test

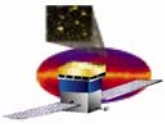
- Not subject to the objections raised for the airplane test
- Offers several attractive features
  - Triggering on events which do *not* pass through the LAT
  - Tests are simple, meaningful, and subject to unambiguous, quantitative analysis
- The committee does recommend VDG tests and proposes two options:
  1. Use the VDG directly by increasing its flux by a factor of ~ 100
  2. Use the VDG flux demonstrated during EM testing and supplement the rate using the solicited condition
- Of the two options, the first is preferable, but the second is an acceptable alternative



# Baseline CR test

---

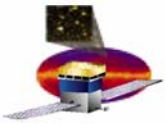
- Record cosmic rays
  - Typical runs should record 1 million events (~1 hour of data taking)
  - Pass criteria:
    - No data system hangs
    - No transport errors (parity errors, timeouts, etc...)
  - Measure, monitor, and record basic data quality indicators (at least)...
    - Total trigger rate, per tower and per layer trigger rate ( tracker and calorimeter)
    - For each detector system: hit and noise occupancy rates
    - For each detector system: reconstructed angular and energy distribution
    - TOT distributions
    - Deadtime distributions (by time and subsystem)
    - Veto and CNO metrics
  - Perform test at least once per week and more frequently when practical, up to once per day, for trending and history analysis
  - Longer runs (~ 1 day) should be done periodically, at least once per month
  - Data quality should match expectations to within statistical precision



# The Level 3 Filter

---

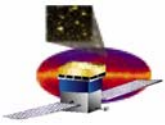
- Depends on the correct operation of:
  - Initialization
  - Triggering
  - Event transport
- Functionality and robustness of the filter algorithms can be both studied and tested in detail both off-line and in the testbed
- Principal issue is the impact of the filter on the overall operation of the T & DF system.
  - Its in the deadtime path
  - The testbed is the most suitable tool to address this issue
  - The VDG with the filter inserted could be used as a sanity check on the integration and correct operation of the filter in the test article.



# Housekeeping and Monitoring

---

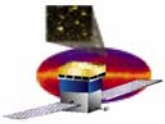
- Depends on the correct operation of:
  - Initialization
  - Triggering
  - Event transport
- No additional data taking appears necessary to verify housekeeping and monitoring functions
- Verification is accomplished by (successful) analysis of the test data taken for both trigger and event transport testing, both outside and inside thermo-vac



# Recommendations

---

- A complete suite of tests should be defined to test every aspect of the T &DF system including:
  - Initialization (power on, booting and configuration)
  - Triggering (both L1 and L3)
  - Event transport
  - Housekeeping and monitoring
- Careful testing over the full range of expected event rates and data volumes should be performed
- The Airplane test while appealing, is unnecessary
- Van de Graff testing, as described, should be included in the test plan



## More...

---

- An explicit administrative entity should be called out and staffed appropriately to ensure the success of the trigger. Using as an example the LAT subsystem pattern, this entity would be responsible for:
  - Definition and implementation of *Unit* testing
  - Definition and implementation of a *System* test plan
  - Coordination of trigger activities which cross subsystem boundaries
  - Liaison with I & T in commissioning and integrating the trigger
- This entity should fall within the control of the electronics subsystem