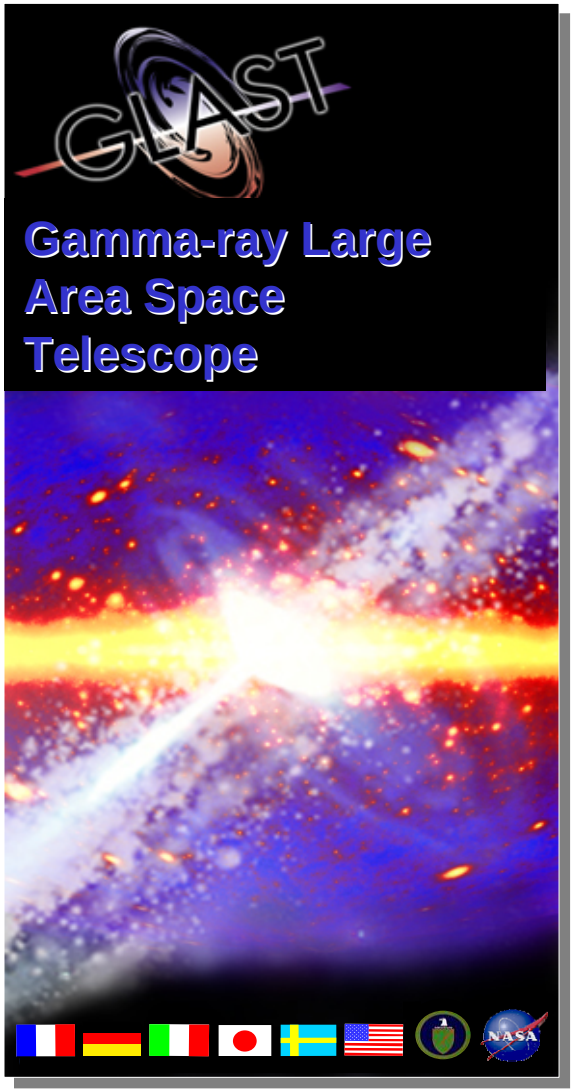


GLAST Large Area Telescope

LAT Reboot Resolution Team

January 04, 2007

Monthly Status

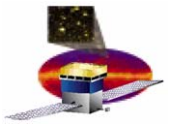
Jana Thayer
Erik Andrews

RRT Status

- Reboot summary data are maintained on the ISOC / FSW Website:
  - http://confluence.slac.stanford.edu/display/ISOC/FSW
- Update on fishbone analysis
- B0-8-0, incorporating team suggestions, to be uploaded after functional test is complete
- Diagnosing reboots during observatory test
  - Memory dump procedure defined
    - Pre-defined memory locations are dumped (~2 hours)
    - Turning procedure into a “blue sheet”
    - Incorporating changes to procedure due to inaccessibility of SSKI rack
  - FSW on call 24/7 to diagnose reboots
  - Additional dumps may be defined after some analysis

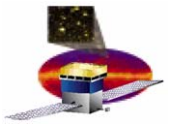
Fishbone Status

- The fishbone analysis has been an effective tool used to focus the attention of the team on the most likely causes
- Two major branches have been virtually eliminated
  - Hardware failures
  - Operations/environment
- Plan forward is to focus on the specific software and software/firmware/hardware interactions that are possible causes



Fishbone

Cause	Watchdog Status	Exception Status
1. Hardware failure induces reboot	Very unlikely	Very unlikely
1.1. Component failure	Very unlikely	Very unlikely
1.1.1. Isolated part failure	Eliminated	Eliminated
1.1.2. Design defect causes stresses that take out specific parts systematically	Very unlikely	Very unlikely
1.1.2.1. CPU board	Very unlikely	Very unlikely
1.1.2.1.1. Bridge chip	Very unlikely	Very unlikely
1.1.2.1.2. SDRAM	Very Unlikely	Very Unlikely
1.1.2.1.3. EEPROM (SUROM)	Eliminated	Eliminated
1.1.2.2. SIB	Very unlikely	Very unlikely
1.1.2.2.1. EEPROM	Eliminated	Eliminated
1.1.2.3. LCB	Very unlikely	Very unlikely
1.2. Environmentally induced	Eliminated	Eliminated
1.2.1. Vibration failure damages chips	Eliminated	Eliminated
1.3. Intermittent hardware failure	Eliminated	Eliminated
1.3.1. Intermittent double bit errors	Eliminated	Eliminated
1.3.1.1. EDACs hardware failure	Eliminated	Eliminated



Fishbone (Continued)

Cause	Watchdog Status	Exception Status
2. Software induced reboot	Possible	Possible
2.1. Operating system flaw	Possible	Possible
2.1.1. Priority inversion	Unlikely	Unlikely
2.1.2.OS does not provide memory protection	Possible	Possible
2.2. Application software bug	Possible	Possible
2.2.1. Memory overwrite	Possible	Possible
2.2.1.1. Generic overwrite	Possible	Possible
2.2.1.2. Watchdog timer overwritten	Eliminated	N/A
2.2.2. Interrupt locks	Unlikely	Unlikely
2.2.3. Task exception	Eliminated	Possible
2.2.4.Race Conditions	Unlikely	Unlikely
2.2.5. Shared memory conflicts	Very unlikely	Very unlikely
2.3. Primary Boot Sequence	Eliminated	Eliminated
2.3.1. Boot I/O	Eliminated	Eliminated

Fishbone (Continued)

Cause	Watchdog Status	Exception Status
3. Operations/environment	Very unlikely	Very unlikely
3.1. Over/undervoltages	Very unlikely	Very unlikely
3.2. Over or under temperature	Eliminated	Eliminated
3.3. Command sequence has unintended side effects	Eliminated	Eliminated
3.4. VSC induced	Very unlikely	Very unlikely
3.4.1. Erroneous times in timetones induces reboots	Very unlikely	Very unlikely
3.4.2. Clock/signal jitter or noise	Very unlikely	Very unlikely
3.5. SEU induced	Very unlikely	Very unlikely
3.5.1. SDRAM and SUROM	Very Unlikely	Very unlikely
3.5.2. Other hardware	Very unlikely	Very unlikely
3.6. EMI induced	Very unlikely	Unlikely
3.6.1. Noise from external source leaks in	Very unlikely	Unlikely
3.6.1.1. Clock/signal jitter or noise	Very unlikely	Unlikely
3.6.2. Noise from another box in the LAT	Very unlikely	Unlikely

Fishbone (Continued)

Cause	Watchdog Status	Exception Status
4. LAT software interacts with computer firmware/operating system feature	Possible	Possible
4.1. Feature documented in vendor errata sheets	Possible	Possible
4.1.1. Errata 13	Eliminated	Eliminated
4.1.2. Errata 15	Possible	Possible
4.1.3. Errata 20	Eliminated	Eliminated
4.1.4. Errata 24	Possible	Possible
4.2. Undocumented and previously unknown errata	Possible	Possible
4.3. 1553 Protocol & I/O	Very Unlikely	Very unlikely
4.4. Interrupt handlers (PPS, GRB)	Eliminated	Eliminated
5. EPU/SIU hardware design flaw	Unlikely	Possible
5.1. Marginal clock jitter/timing/noise	Eliminated	Eliminated
5.2. LCB FPGA error	Unlikely	Possible
5.2.1. LCB incorrectly writes memory	Unlikely	Possible
5.2.2. LCB doesn't release PCI and hangs CPU	Very unlikely	Very unlikely
5.3. EMI within the box	Very unlikely	Very unlikely

Plan forward

- **Possibilities remaining:**
  - Code (ISR or task?) is spinning with interrupts disabled (watchdog)
  - Corruption in exception vectors, interrupt dispatch code, ISR code, or task scheduling code
    - Memory overwrite (watchdog, exception)
    - Task exception (exception)
  - Hardware
    - Errata 15 and 24 (watchdog, exception)
    - LCB FPGA error: incorrectly writes memory (exception)
- **Plan forward:**
  - Apply workarounds for BAE errata (B0-8-0)
  - Implement FSW changes resulting from code review
  - In case of watchdog, use expanded LSW trace, coherent copy of interrupt stack, task stack, and dump of last ISR and subroutines called from ISR to determine whether problem is an interrupt, code corruption in an interrupt routine, or code corruption elsewhere
  - In case of an exception, task ID, PC, etc are recorded
  - Data to FES conversion tool being implemented in order to run LAT data taken prior to reboot through the testbed to try to reproduce the reboot

FSW Status

- B0-6-15+ - installed on LAT 11/21
  - Includes correction to LSW trace code (LSW V0-1-1)
    - Root cause for many exceptions has been identified and corrected
      - Signed/unsigned char issue
      - Memory overwrite
- B0-8-0 new build
  - 17 Jira's addressed
  - Incorporate changes related to RAD750 errata # 15 and # 24
  - Include enhancements recommended by RRT
    - Expanded LSW trace
      - Records task switches
      - Interrupt entry/exit
      - Interrupts enabled/disabled?
      - Value of watchdog timer
      - PC of incoming/outgoing task
    - VXW: write-through mode enabled
  - Build is being tested
  - Install on LAT after completion of functional test

Reboots during Observatory Test

- If a reboot occurs during Observatory Test:
  - Planned LAT testing is terminated
  - Memory dump procedure is run by LAT operator
    - Pre-defined memory locations are dumped and received
      - SIU dumps: via primary boot 1553 housekeeping*
      - EPU dumps: via diagnostic telemetry
    - Approx 3 Mbytes and two hours for this dump
    - Turning procedure into a “blue sheet” and incorporating changes due to inaccessibility of SSKI rack
  - FSW may define additional dumps after some analysis
  - FRB will be convened
  - FSW on call 24/7 to diagnose reboots

*If dropouts in 1553 housekeeping persist, our ability to diagnose SIU reboots will be impaired. If the problem cannot be fixed, can we request a dump of the housekeeping partition of the SSR to get the data dumps?

- Neil has circulated a document describing the process and timescale
  - LAT dump activity can be delayed until convenient in observatory test sequence – only impact is continuous primary boot telemetry.
  - LAT power configuration cannot be changed until the dumps are complete else data will be lost.
  - LAT reboot diagnosing activities may extend beyond planned shift boundaries and potentially impact Observatory test activities, else diagnostics will be lost.

GLAST Large Area Telescope

Monthly Mission Review

Backup

Stanford Linear Accelerator Center

Testbed and LAT

- **The testbed was created to ---**
  - **Simulate the LAT as closely as possible for the purposes of FSW development**
  - **Approximate system-level testing on the flight LAT prior to LAT integration**
  - **Pre-launch testing of the performance of the FSW/DAQ at the expected on-orbit trigger rates**
- **Although the testbed could be improved, the cost of such improvements and the inherent limitations of the testbed for the purpose of diagnosing reboots must be understood**
- **The root cause of the reboots has not yet been isolated to a failure of hardware/software or identified as a dependence/independence on the data or the environment**
  - **The answers will affect which potential improvements are most likely to reproduce the problem on the testbed**
  - **One cannot make an intelligent decision on how to deploy our limited resources for testbed improvement without first isolating root cause.**

Testbed/LAT differences

- Although DAQ side of LAT and testbed are logically identical, the FES does not simulate the TKR, CAL, ACD detectors, or the front-end electronics that read them out.
  - It simulates the interface to the DAQ system at cable-level
  - To do more would require a larger FPGA on the FES and substantial work to re-design the FES.
- The FES does not simulate the LAT's latency from the time a TACK is issued to the time that each bit of the readout data is presented to the DAQ
- There is no "connection" of the LATC configuration to the FES logic.
  - The FSW on the testbed can execute in an environment where all of the input data sources and output data sinks are available, but the timing of these sources and sinks differs substantially from the LAT
- The processors on the testbed differ from the LAT
- The SDRAM organization on the EM processor board is slightly different from the flight board (BAE erratum 24) - reboot seen with work around in place
- Minor differences in parts and logic between parts of DAQ system on testbed and LAT resulting from the fact that flight parts were used on the LAT, commercial parts on testbed.

Statistics

- It would require ~ 1 billion events *with the unique characteristic* that causes the reboot in order to have a good chance of reproducing the problem
  - The FES was not designed to handle this kind of data volume and would require an infrastructure re-design to be able to serve that amount of data to the DAQ
  - Where would we get this data?
    - It would take a lot of time to make a good MC simulation of the sea-level muons
    - This MC would not necessarily contain the unknown characteristic that would produce the ocrruption
    - Play the LAT data back through the FES
      - Does the event that caused the reboot actually make it to the ground?
      - If the unique characteristic that causes the memory corruption is a combination of the data content of the event and the way the event is broken into packets by flow control and the locations in memory where those packets end up being placed by the LCB, then this setup may still fail to reproduce the problem.

LAT data

- **Every event on the LAT is unique**
  - **The LAT is an analog detector with millions of channels each with their own noise and characteristic response.**
  - **Two identical events hitting the detector will produce subtly different outputs in the analog and digital outputs which will vary the likelihoods of noise-induced triggers**
- **The LAT's sense of time and time sequences is not reproducible**
  - **Each processor, GASU, 1553, VSC, and GPS have their own clock which runs at its own frequency and varies with time and temperature.**
  - **Two identical runs started at exactly the same time will always be different: the number of machine instructions executed between events will differ, the arrival of the SC attitude message can vary, the arrivals of packets at an EPU's circular buffer will differ, etc. As a result, identical events in identical runs can could arrive in EPU memory divided into different sequences of packets located at different memory addresses.**
- **The VSC network load varies**
  - **This affects the VSC's ability to keep up with the data stream sent to the SDI.**

FES data

- **The FES data does not complete specify the contents of an event's data stream.**
  - **GEM contribution not specified by FES**
  - **GEM event data can only identify the identify the relative timing of the trigger primitives whose assertion led to a TACK. The LAT has a finer granularity.**
    - **Converting LAT data to FES format made difficult by the fact that there are no data recorded on which to base a logical decision as to precisely which primitives should be asserted or the order in which they should be asserted or for how long they should be asserted.**
  - **Latency between TACK and read out of data not simulated**
    - **To upgrade the FES would require 6-12 months of effort and a lot of money**
    - **Not worth pursuing unless we get evidence that this is the cause of the problem and that the upgrade would have a significant chance of identifying and leading us to a solution**

Some improvements

- Creating data to FES conversion tool
  - Will have to make up the GEM contribution; that is a roadblock
- Could investigate getting rad750 EM variant from BAE
  - Buying one costs ~ $130K
  - Could we swap or borrow?
- Could get flight processors going, but it is not clear that this would lead us any closer to a solution
  - Requires putting on a modern PBC and flight build
  - Mounting to testbed requires the develoment of controls
    - ESD, contamination control, safety
    - EGSE cables attaching to the crates would need to be build, HiPOt tested, etc. prior to mating to the crates
    - Maintain power on time logs and have a cognizant person present whenever flight crates are powered
      - Limited resource pool of “cognizant” people
    - Would need administrative controls to restrict the usage of these processors and this would interfere with all other testing
      - FSW development and test, I&T/LICOS/system testing, PROC development, etc.

Top FSW JIRAs (Critical or Major Severity)

Priority	Key	Summary	Issue Type
Critical	FSW-292	Implement GRB detection algorithm	New Feature
Major	FSW-456	EMP and LCM do zlib compress with malloc/free, should use MBA_alloc/free	Improvement
Major	FSW-305	Summary/statistics telemetry stream needs to be created for on-board event processors	Improvement
Major	FSW-369	MSG needs to disable reports from within the MSG task	Bug
Major	FSW-576	Bug in CAL data compression algorithm	Bug
Major	FSW-623	CLONE -Documentation for several apids needs to be added to standard webpage	Improvement
Major	FSW-341	LPA Mode Change/Flush Behavior is Incorrect	Bug
Major	FSW-682	LTC estimation filter timescale is too short	Improvement
Major	FSW-680	Swap LHKPnxHP3DSIT and LHKPnxHP5DSIT to address miswiring of thermal sensor	Bug
Major	FSW-562	Make sure that PIG's power sequence is still correct	Improvement
Major	FSW-168	Add LIM mode status to regular housekeeping packet	New Feature

Top FSW JIRAs (Critical or Major Severity) (2)

Major	FSW-684	There need to be general no-op commands for each task.	New Feature (7/21 ENHANCEMENTS MTG)
Major	FSW-685	Expand LHKDIAGAPID argument range for LHKREQDIAGPKT	Improvement (7/21 ENHANCEMENTS MTG)
Major	FSW-686	Mnemonic LHKSMEM0MPTID should be LHKSMEMDMPTID	Improvement (7/21 ENHANCEMENTS MTG)
Major	FSW-687	LHKT0TEM28V0ST and LHKT0TEM28V1ST et al are missing conversion	Improvement (7/21 ENHANCEMENTS MTG)
Major	FSW-270	mnemonics in telemetry packet 720/0x2D0 do not begin with ?L?	Improvement
Major	FSW-698	Separate LTC master config files into fof, data	Improvement (7/21 ENHANCEMENTS MTG)
Major	FSW-699	Create report to identify configuration files in use	Improvement (7/21 ENHANCEMENTS MTG)
Major	FSW-701	Add flexibility to MSG level output based on destination	Improvement (7/21 ENHANCEMENTS MTG)
Major	FSW-702	EPU secondary boot indication	Improvement (7/21 ENHANCEMENTS MTG)
Major	FSW-704	Read, report and clear flag registers	Improvement (7/21 ENHANCEMENTS MTG)
Major	FSW-703	Ensure all registers are set	Improvement (7/21 ENHANCEMENTS MTG)
Major	FSW-705	Support chip reset commands (and possibly others)	Improvement (7/21 ENHANCEMENTS MTG)
Major	FSW-688	LMEMPAD re-use	Improvement (7/21 ENHANCEMENTS MTG)