GLAST Large Area Telescope:

Service Challenge Update and Collaboration Compute Needs

Richard Dubois
Stanford Linear Accelerator Center
richard@slac.stanford.edu

Service Challenge Reminder

- **Coordinated Simulations for Science Group and ISOC preparation and testing**
- **Steering committee started meeting one year ago(!)**
  - **Representatives from Sci Groups, ISOC, Data Handling**
- **Activities so far**
  - **Through Spring '07, mostly developing machinery etc for Science Group datasets**
  - **Creation of a Level 0 science data set for ISOC development of L1 processing (along with data)**
- **Started cooking with gas this summer**
  - **More later on this :-)**

Connection to Science Groups

- **Several types of datasets have been identified for the Science Groups use:**
  - **Backgrounds Analysis & IRFs**
    - **"AllGammas"**
    - **"AllMuons"**
    - **Backgrounds**
    - **IRF Test Suite**
  - **Sky Analysis**
    - **1 orbit-yr obssim runs**
    - **55 orbit-day Gleam runs**
      - **survey**
      - **L&EO**
    - **1 orbit-year Gleam run**
      - **The definitive dataset for Science preparation!**
      - **Potentially huge backgrounds run needed**
      - **Targeting use of Lyon, Italian farms**
  - **Plus smaller scale specialty runs as needed**

Connection to ISOC

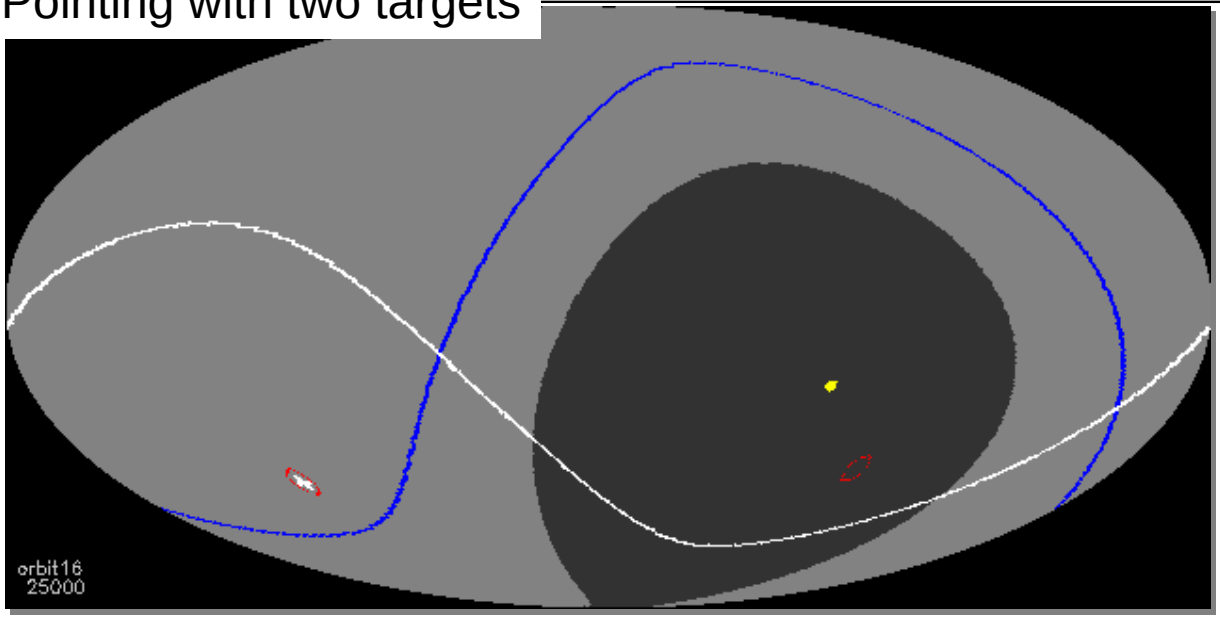
- ISOC testing has concentrated on Mission ETE tests
  - Used for control room type functions
    - Shift log (eLog), L1 pipeline, Data Catalogue, Monitoring
  - Used simulations to prep for ETE & provide realistic science data, extended running

- Simulate L0 science data
  - In support of OktoberTest - 16 orbits with different configurations, conditions
    - Downlink simulations for instrument readiness tests, such as calibrations, failed sensors etc.
  - Full simulation of sky+background (compute intensive)

- 55 day orbit run
  - Extended run to test ASP
  - Time trending of instrument quantities

Service Challenge Eye Candy

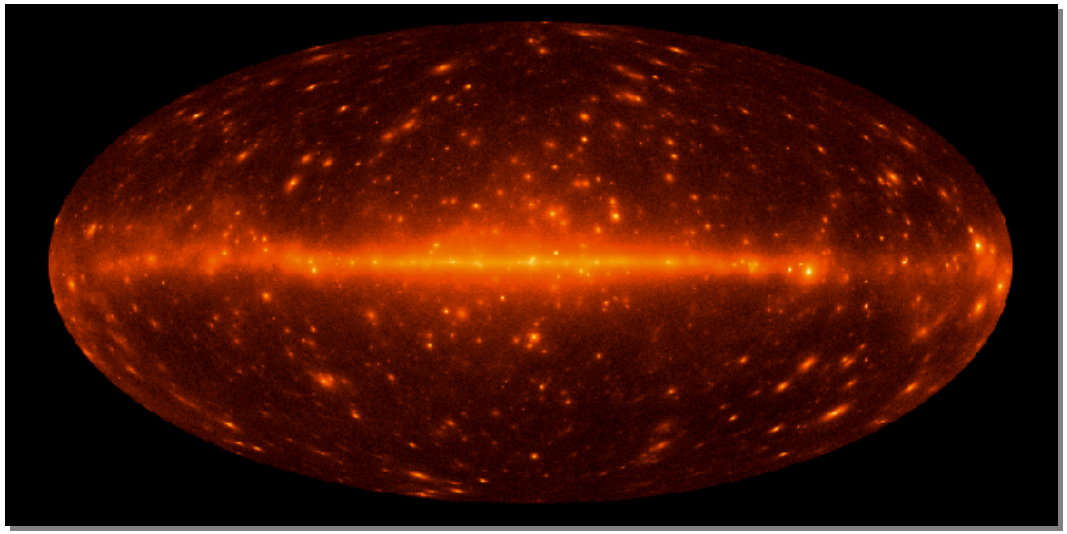
Pointing with two targets



Offline Sim of Onboard GRB Filter

GRB Trigger Time	0.02089589834
FirstRA	151.2563276
FirstDEC	-38.92002236
First Estimated Error	0.5743404438
nPhot w/ [0,100) MeV	20
nPhot w/ [100,1000) MeV	2
nPhot w/ [1,10) GeV	0
nPhot w/ > 10 GeV	0
Trigger window size	40
EnergyCut	-1

Alert notice!

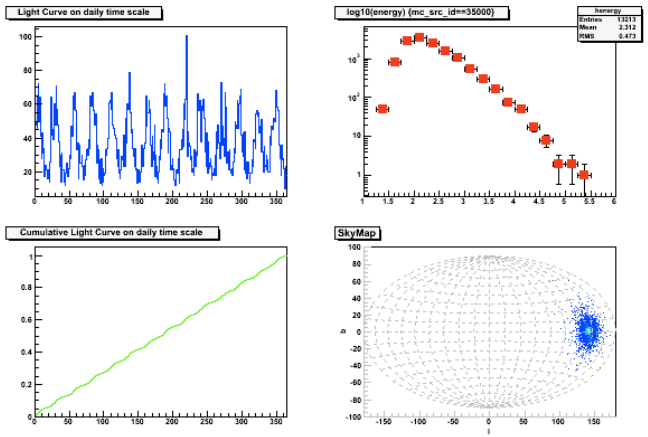


R.Dubois

Report of the 1yr sky model

Total Number of mc_src_id processed:1884
In GMT(UTC) date: 2007/06/21 - 07:45:00

LSI +61 303



Proposed Performance Update Timeline

Steve, Bill, Julie

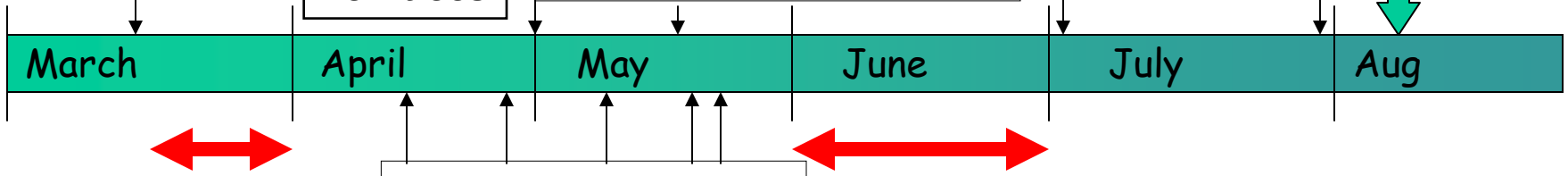
New merit variables complete and tested by 14 March:
- irreducible tags - Tracy
- multi-gamma event variables - Leon
- DOCA to screw holes - Eric

Pass 5 analysis, new trees

Reprocess for new* CTB variables, start 1 May. **Interim P5 IRFS by mid-May**

Iterate and check P5 analysis. FINAL by 1 Aug

Pre-launch analysis frozen



2-3B bkgd events, plus AG and diffuse

Improvements&**TESTING**:
- Beam test G4 tuning
- CAL per channel peds/gains/zero-suppression thresholds
- CAL recon robustness tweaks
- Spacecraft and other geometry updates
- OBFs

20-50B bkgd events plus AG and diffuse

? any infrastructure or package upgrade plans?

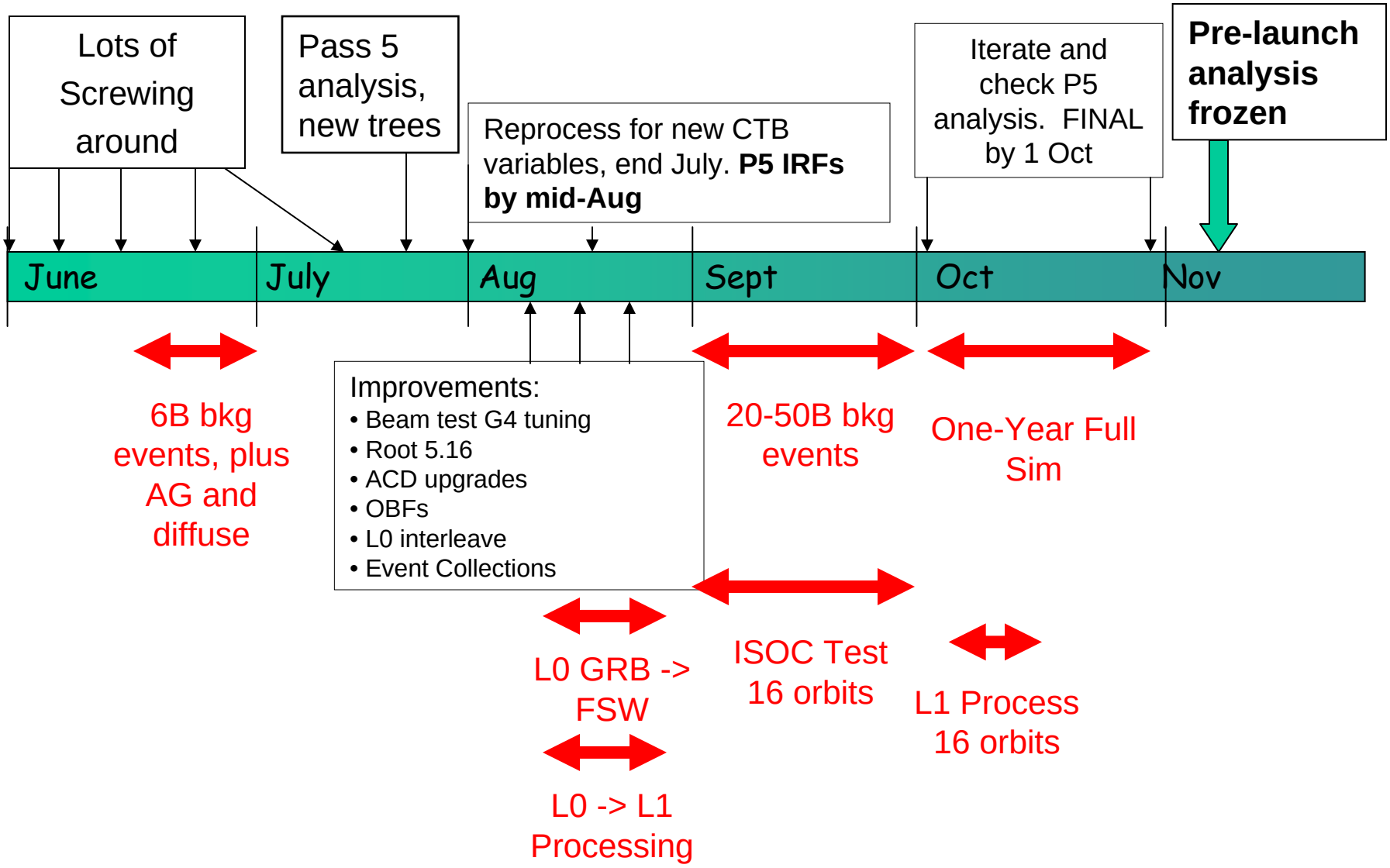
prior to each large run, we will have a short trial set of runs (bkgd, AG, Day) and a validation mini-review of a pre-defined set of plots.

Science analyses and performance page use:

DC2 IRFs	Pass4 IRFs ?	Interim Pass5 (middle IRFs)	Launch IRFs

* note: current CTB variables are from pass4 version 8 patch 2, used for handoff review and SWG presentation

Current Timeline



Simulations Plans for the Rest of '07

- L&EO
  - 56 days
  - 2 weeks pointed; 4 weeks survey with ARRs
- OktoberTest 16 orbits
  - 16 one-orbit runs with various LAT configurations, calibrations and failure modes
- One orbit-year
  - The Big One - a month in the making
- Level 0 interleaves for SO
  - Multi day simulations to exercise science aspects of L1/ASP
- Obssim equivalents of L&EO and 1-Year
  - And sundry others (they're cheap and easy)

Known Liens on Resources in 2007

- Existing Data & MC
  - Beamtest, I&T, MC => 85 TB remaining in hoard
- Beam Test
  - Perhaps another 5 TB to complete
- SC MC runs
  - ~30 TB for Gleam runs - 55 day + 1 yr
- OktoberTest runs
  - < 3 TB
- TVAC
  - 10 TB
- Leaves head room for another big simulation effort before launch

Needs for Post-Launch

Folder /MC-Tasks/ServiceChallenge/backgndSC-GR-v11r2/ru

Name	Type	Files	Events	Size	Created	Links
merit	Group	21599	31,681,038	22.5 GB	09-May-2007 10:50:37	Files
mc	Group	21599	31,682,566	1.7 TB	09-May-2007 10:50:37	Files
digi	Group	21599	31,682,566	62.9 GB	09-May-2007 10:50:37	Files
recon	Group	21599	31,682,566	259.8 GB	09-May-2007 10:50:37	Files
relation	Group	21599	31,682,566	51.7 GB	09-May-2007 10:50:38	Files

Total “data” ~ 450 GB/day => ~160 TB

Folder /ServiceChallenge/Interleave55d-GR-v11r12/runs

Name	Type	Files	Events	Size	Created	Links
merit	Group	7920	1,813,577,206	1.2 TB	03-Jul-2007 07:30:02	Files
mc	Group	7920	38,976,220	562.8 GB	03-Jul-2007 07:30:04	Files
digi	Group	7920	38,976,220	52.2 GB	03-Jul-2007 07:30:04	Files
recon	Group	7920	38,976,220	284.0 GB	03-Jul-2007 07:30:04	Files
relation	Group	7920	38,976,220	19.0 GB	03-Jul-2007 07:30:04	Files

MC ~ 25 TB per yr of simulation

Estimates are for ONE copy of the data.

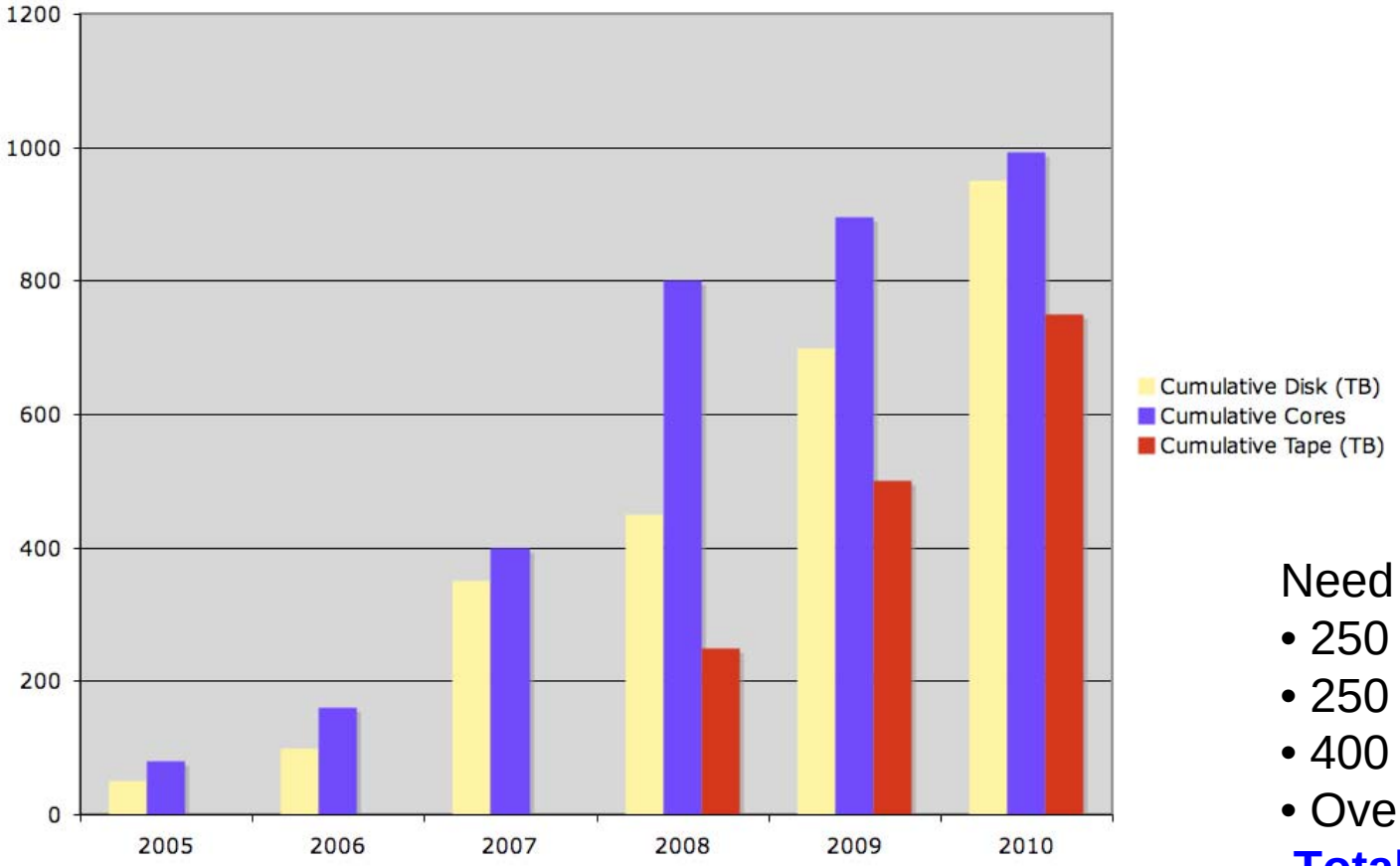
We will have to get smarter by the time of the first major reprocessing

Request 250 TB/yr to contain data & MC needs

Buy 50 more octal-core boxes for data/MC load - then maintenance

We still estimate ~100 cores to reconstruct a downlink in 1 hour, but it is looking like we will need to devote more to this in practice.

Computing Resource Projections



Need for/in '08:
- 250 TB disk: $288k
- 250 TB tape: $100k
- 400 cores: $200k
- Overhead: $76k

Total: $664k

Current pricing: $1.15k/TB disk; $0.40k/TB tape; $0.5k/core

The LAT Computing "GRID"

- SLAC
  - Pipeline2 configured originally to run at SLAC
  - 400 LAT cores now in batch farm
  - Plan to double this next year
- Lyon
  - P2 ported to Lyon
    - Beta test about to begin with SC runs
    - Plan to transfer all files to SLAC to minimise disk commitment at Lyon
    - Expect 200-300 CPUs for bursts and 50 TB of disk as needed
- Italy
  - Need to re-evaluate in light of competition at CNAF from LHC needs

Future Storage Options: Orbit Data

- 10% Solution: aka removing obvious background
  - 60 TB/yr
  - Assumes we can safely cut 90% of the background - we don't have the cuts determined yet - need input from C&A group

- Full disk in '07; 10% in '08 and onwards (add 25 TB contingency to 150)
  - ~175 TB for data
  - Plus disk/tape for MC
  - We'll see how disk usage goes with some experience under our belts in '08

Best Plan is to optimize purchase of (250 TB, 400 cores)

- No recon on disk
  - ~30-40 TB/yr
  - suffer latency in retrieval when needed
    - 1 TB retrieved in 1 day currently
    - Affects reprocessing merit; event displays; calibrations
    - Is there some need we'll find for fast turnaround that would make us regret not having things right there? Don't know yet.
- Wrinkles
  - Price of disk will continue to drop. How soon and how much? Another 20% last year
  - SCS has told us to start worrying about tape costs starting in '08

SAS Manpower Needs

- Already working closely with the collaboration
- Could not be done without this group-wide effort
- Quite a bit of ebb and flow in who does what

- Historical contributions outside SAS/SLAC:
  - Italy:
    - Core tools, G4 support – 2,5 FTE
    - TKR alg development – 1 FTE
    - SciTools & ASP development – 2,5 FTE
    - BT release management - 0,5 FTE
    - MC Pipeline operations - 0,25 FTE
  - France:
    - Core tools: 0,5 FTE
    - CAL sim/recon/calibs – ~5 FTE
    - Pipeline @ Lyon - 0,25 FTE
    - Source Catalogue – 3 FTE
  - US:
    - Core tools – 2 FTE (UW, Goddard)
    - CAL sim/calibs – 2 FTE (NRL)
    - ACD – 1 FTE (Goddard, SVAC)
    - TKR – 1 FTE (UCSC)
    - Sci Tools - 3 FTE (GSFC)
    - Sundry contributions

- Responsibilities outside SAS/SLAC:
  - Builds manager – Goddard
  - Code build tool (CMT) – LLR/France
  - Build tool gui & Event display – Udine/Italy
  - Source Catalogue – CEA/France
- Responsibilities inside SAS/SLAC:
  - Overall management
  - Instrument code release manager
  - SciTools release manager (with funding from GSSC)
  - Data handling
  - Documentation
  - System tests manager
  - Calibrations infrastructure
  - Overall Reconstruction; TKR
- Shortfall:
  - ~ 4 FTE for infrastructure work
  - GSFC working on a hire!
  - Hardest to get university assistance here

What to Take Away

- **ServiceChallenge + ETEs**
  - Being used to hone the tools, complete development and test end-to-end operations
- **Prudent approach is to:**
  - Acquire ~800 cores available to GLAST
    - Lesson learned from 5-ring circus of DC2, BT, EM
  - Keep full event details on disk in '08 : ~175 TB
    - Would shortchange CPU if push comes to shove
  - GLAST will do better science the more compute power it has access to
    - Have not hit the plateau yet!
    - Extend Pipeline 2 to France and Italy
- **SAS manpower is barebones in core software**
  - Expect to get 1 new FTE at GSFC via Steve
  - Short 3 with no prospects of getting them from the rest of the collaboration