GLAST Large Area Telescope:

Exploring the γ-ray Sky

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

http://www-glast.slac.stanford.edu/software

Über Outline

- Introduction to GLAST & C++ world
- Reconstruction events in a pair conversion telescope
- Astronomy analysis with GLAST
- Data Handling

Outline

- Introduction to GLAST
- The Instrument
  - Pair conversion telescope
- Code Development Environment
- Users: code installation, documentation
- Overview of C++ World: Gaudi, GEANT4 etc

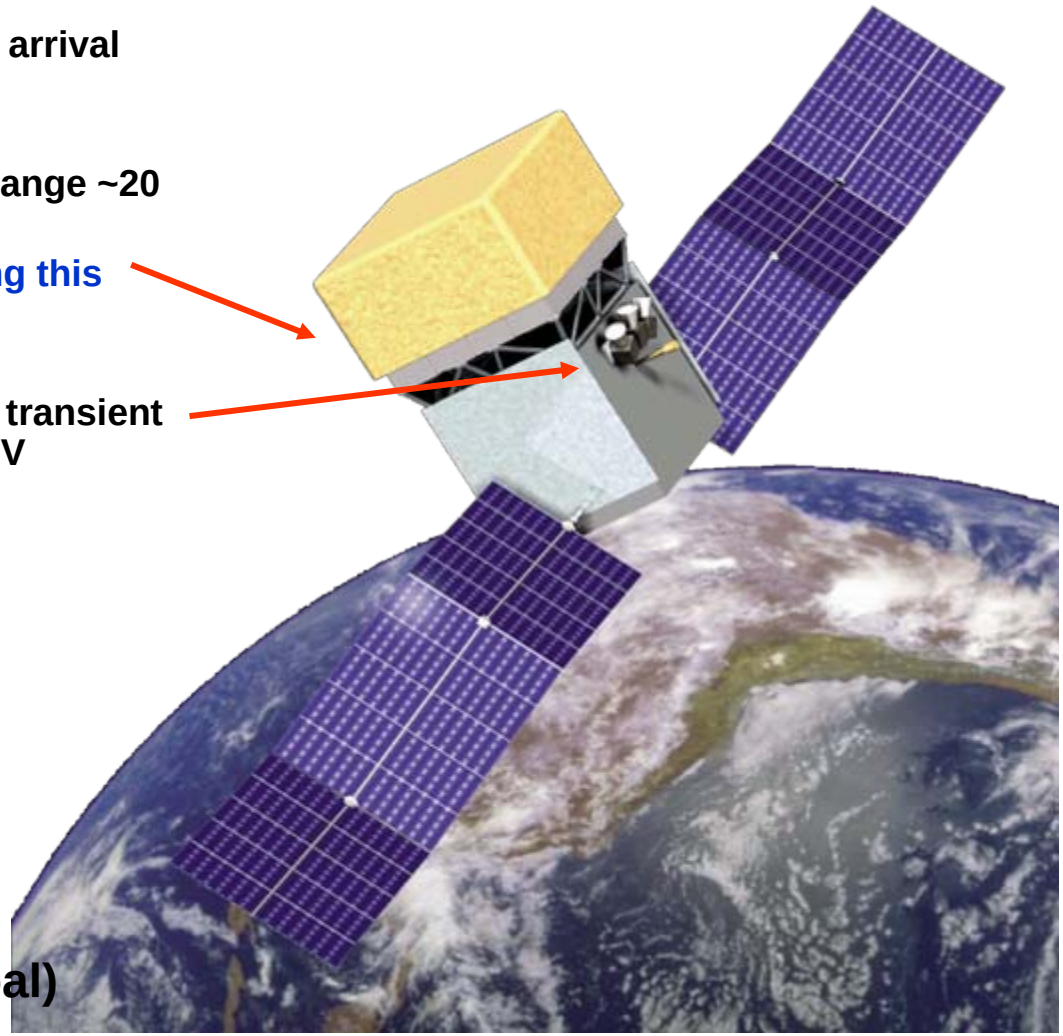
GLAST Mission

GLAST measures the direction, energy and arrival time of celestial gamma rays

-LAT measures gamma-rays in the energy range ~20 MeV - >300 GeV

- **There is no telescope now covering this range!!**

- GBM provides correlative observations of transient events in the energy range ~20 keV – 20 MeV



Launch: August 2007
Florida

Orbit: 565 km,
28.5° inclination

Lifetime: 5 years
(minimum; 10 yrs goal)

GLAST Participation

France Germany Italy Japan Sweden USA

NASA - DoE Partnership on LAT

LAT is being built by an international team

- Stanford University (SLAC & HEPL, Physics)
- Goddard Space Flight Center
- Naval Research Laboratory
- University of California, Santa Cruz
- University of Washington
- Ohio State University
- CEA/Saclay & IN2P3 (France)
- INFN & ASI (Italy)
- Hiroshima University, ISAS, RIKEN (Japan)
- Royal Inst. of Technology & Stockholm Univ. (Sweden)

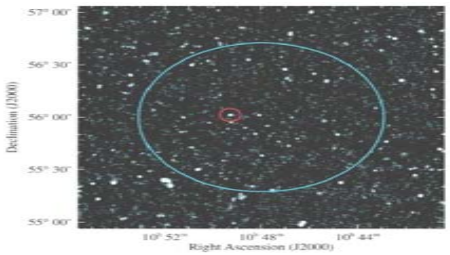
LAT managed by SLAC
PI – Peter Michelson

GBM is being built by US and Germany

- MPE, Garching (Germany)
- Marshall Space Flight Center

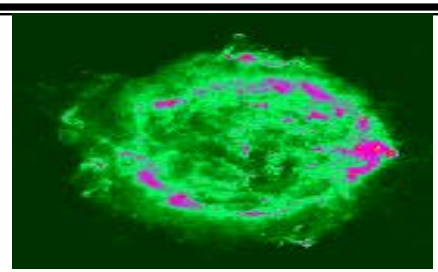
Spacecraft and integration - Spectrum Astro

GLAST science - the sky above 20 MeV

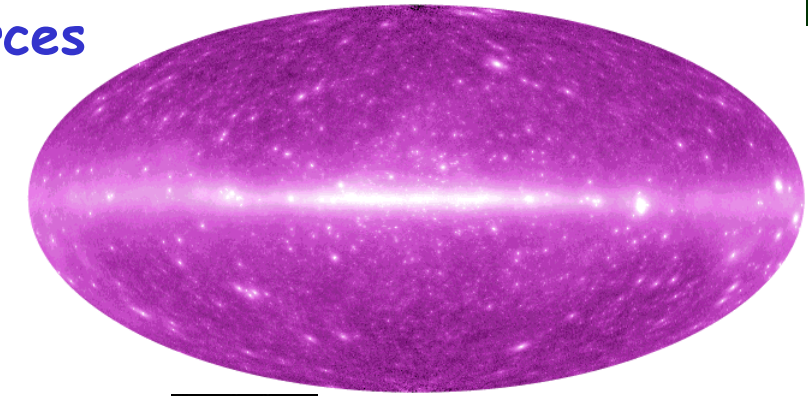


Unidentified sources

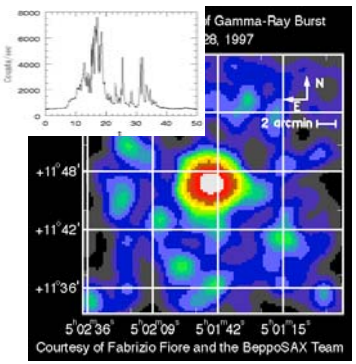
Active Galactic Nuclei

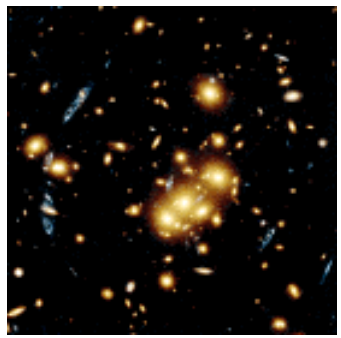


Cosmic ray acceleration



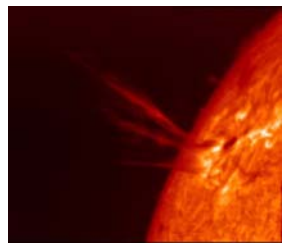
Pulsars





Gamma Ray Bursts

Dark matter



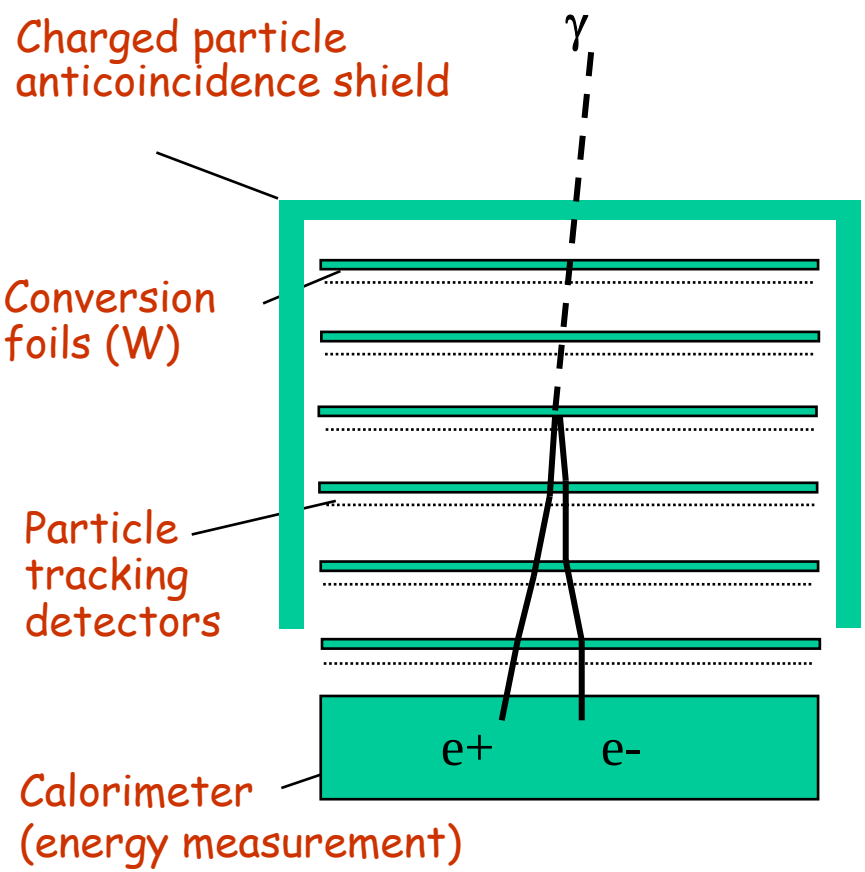
Solar flares

0.01 GeV 0.1 GeV 1 GeV 10 GeV 100 GeV 1 TeV

R.Dubois

γ detection – pair conversion telescope

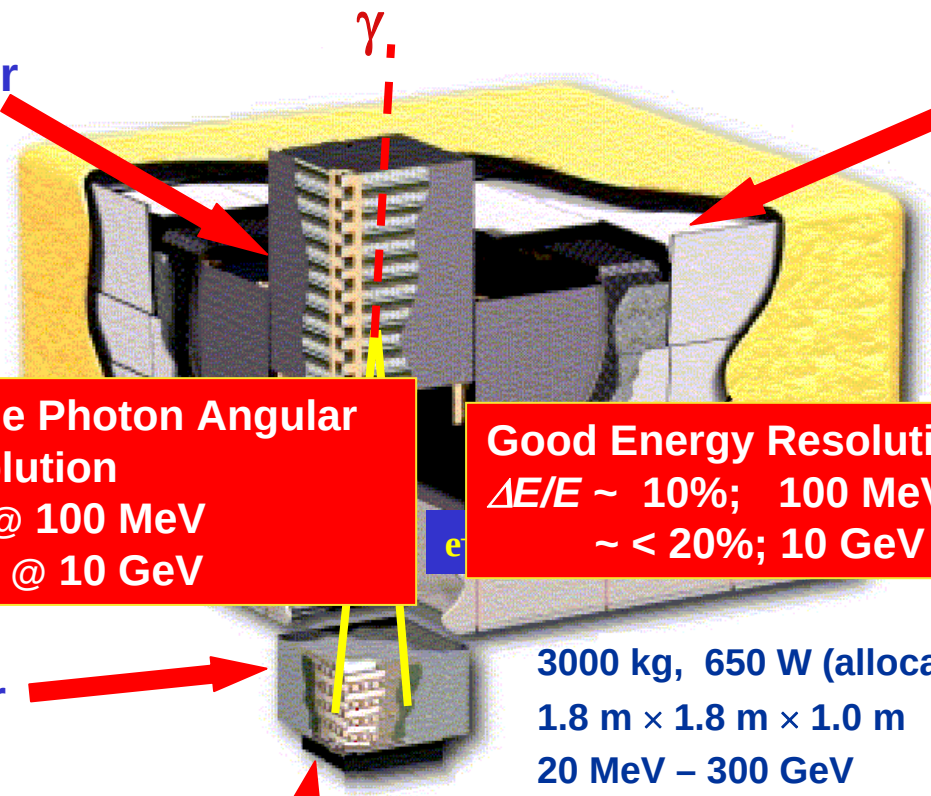
Pair production is the dominant photon interaction in our energy range



GLAST Concept

- Low profile for wide f.o.v.
- Segmented anti-detector to minimize self-veto at high E.
- Finely segmented calorimeter for enhanced background rejection and shower leakage correction.
- High-efficiency, precise track detectors located close to the conversions foils to minimize multiple-scattering errors.
- Modular, redundant design.
- No consumables.
- Low power consumption (650 W)

GLAST Large Area Telescope (LAT)



Si Tracker Tower

pitch = 228 μm
5.52 10^4 channels
12 layers × 3% X_0
+ 4 layers × 18% X_0
+ 2 layers

Single Photon Angular Resolution
3.5° @ 100 MeV
0.15° @ 10 GeV

ACD

Segmented scintillator tiles
0.9997 efficiency
⇒ minimize self-veto

Good Energy Resolution
ΔE/E **~ 10%; 100 MeV – 10 GeV**
~ < 20%; 10 GeV – 300 GeV

Thermal
rs)

3000 kg, 650 W (allocation)
1.8 m × 1.8 m × 1.0 m
20 MeV – 300 GeV

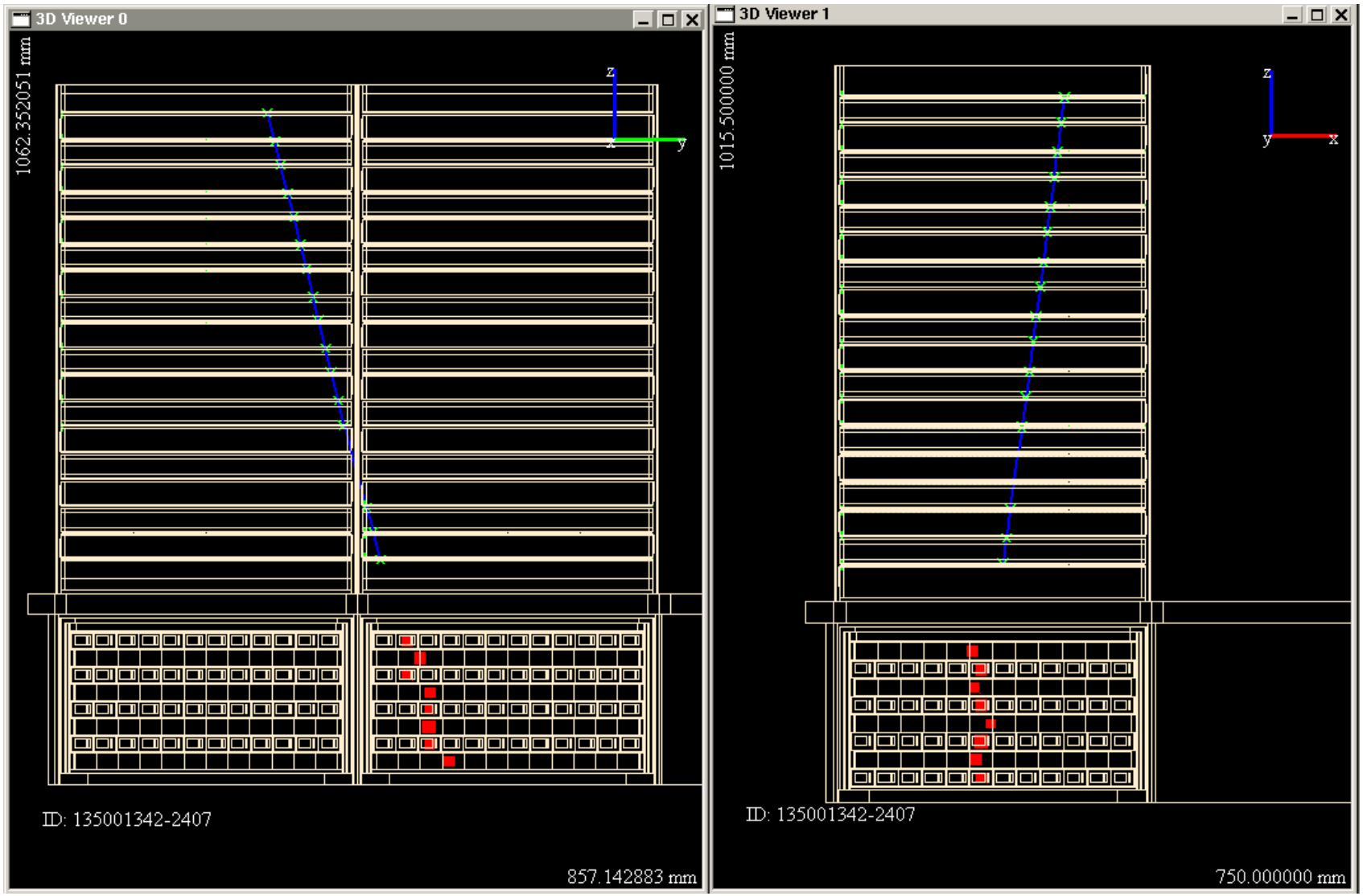
CsI Calorimeter

Hodoscopic array
8.4 X_0 8 × 12 bars
2.0 × 2.7 × 33.6 cm
⇒ cosmic-ray rejection
⇒ shower leakage correction

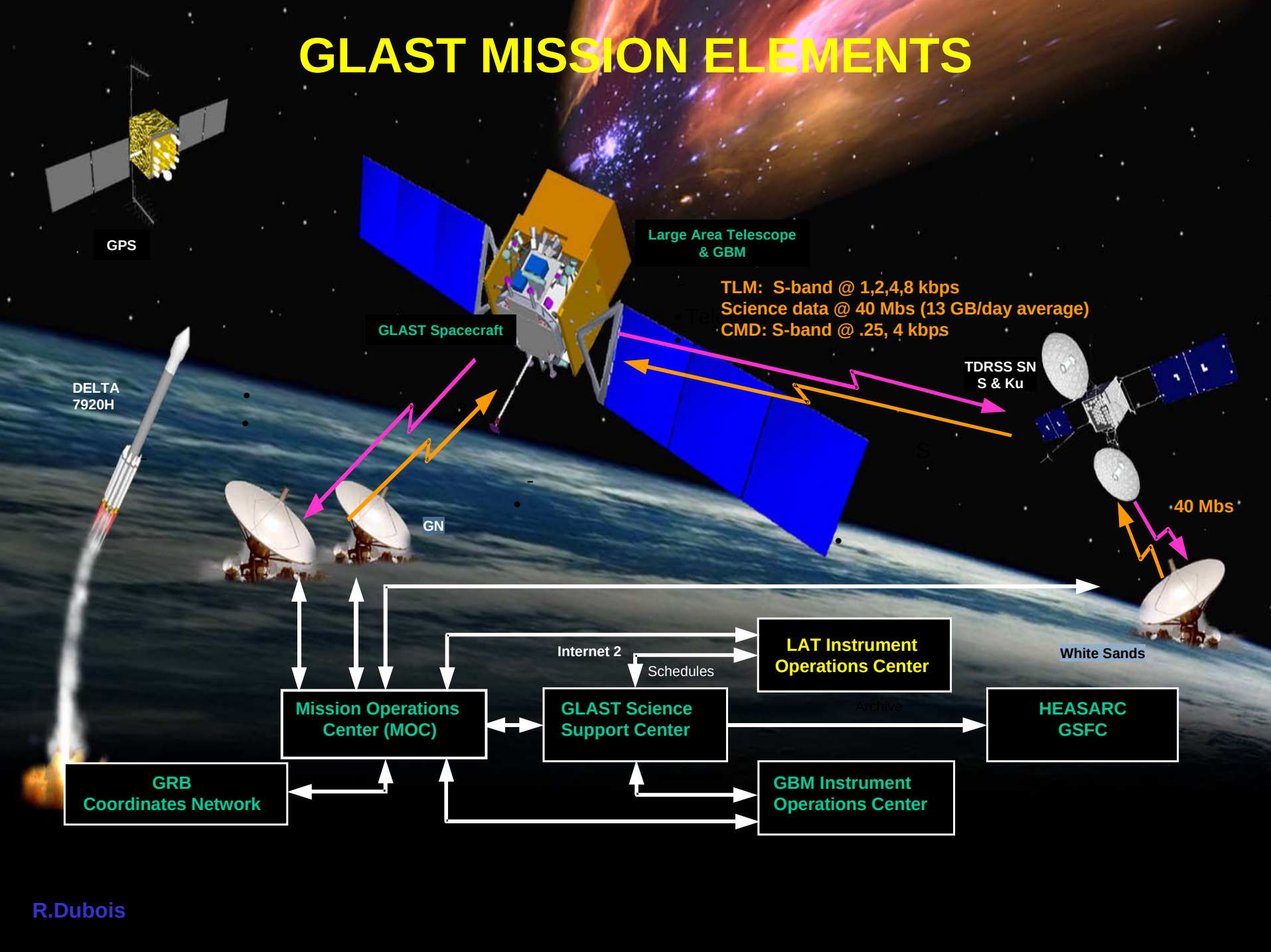
Data acquisition

16 identical towers
300 Hz average downlink

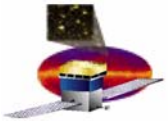
Cosmic Ray Muon for Two-Towers



GLAST MISSION ELEMENTS



R.Dubois




GLAST Software

SAS
LAT sim/recon
ScienceTools
Infrastructure
System Tests
Simulation
Architects
Analysis Tools
Pipeline
ACD
CAL
TKR
GEANT4
Likelihood
Pulsars
GRBs
Obs Sim
List the rest
Code Distribution
System Tests
Issues Tracker
SLAC Linux environment
Sundry Utilities
Event Display
Release Manager
Data Server
Documentation
SLAC Windows Environment
Caliibrations
Code Management Tool
Recon
Event Interpretation
ACD
CAL
TKR
User Interface
I/O
Release Manager
Architect
Flight Int Support
Release Manager
I/O

~25 FTEs

Software Development Approach

- Enable distributed development via cvs repository @ SLAC
- Extensive use of electronic communications
  - Web conferencing (VRVS), Instant Messaging (icq)
- CMT tool permits equal development on Windows and Linux
  - 'requirements' file generates MS Project or gnu Makefiles from single source
  - Superior development environment on Windows; compute cycles on linux
- documentation and coding reviews enforce coding rules
- "Continuous integration"
  - Eliminate surprises for incoming code releases
  - Build code when packages are tagged; alert owners to failures in build or running of unit tests. Results tracked in database.
  - Developing comprehensive system tests in multiple source configurations. Track results in database; web viewable.

Documentation: User Workbook

Workbook for Offline Users

HOME

Site Map

GLAST Links | SAS Software | Get Connected | Installing GLAST S/W (End-user | Developer) | GLEAM | Running GLAST Applications (FRED | MRvcmt | ROOT | Science Tools) | Advanced

ROOT: 1: Overview & Setup 2: Outputs 3: View Ntuple 4: RootTreeAnalysis 5: Accessing Data

ROOT 3 View Ntuple: View Summary Ntuple | Create Histogram | Remove this navbar

Print Version

Use Case I: Summary Ntuple

This section provides detailed procedures to open and view a summary ntuple, create TCuts, and create an ASCII file containing ntuple contents.

Open and View a Summary Ntuple

1. To download an example summary ntuple ROOT file, go to:

   ftp://ftp-glast.slac.stanford.edu/glast.u07/mcenery/systests/GlastRelease/v6r2p8/AllGamma/linux/

   Download the *AllGamma_Merit.root* file and save it in *yourWork* directory.

   Troubleshooting Tip: Make sure that, if you have not set up a permanent environment for ROOT analysis, your temporary environment is set up correctly. (Refer to Set Root Environment Variables: Linux or Windows.)

2. Start up ROOT then, in sequence, enter the following commands:

TFile f("AllGamma_Merit.root", "READ")	open the Summary Ntuple file
f.ls()	view its contents
*TTree *MeritTuple = (TTree*)f.Get("MeritTuple")*	load the summar ntuple TTree
MeritTuple->StartViewer()	start the TreeViewer

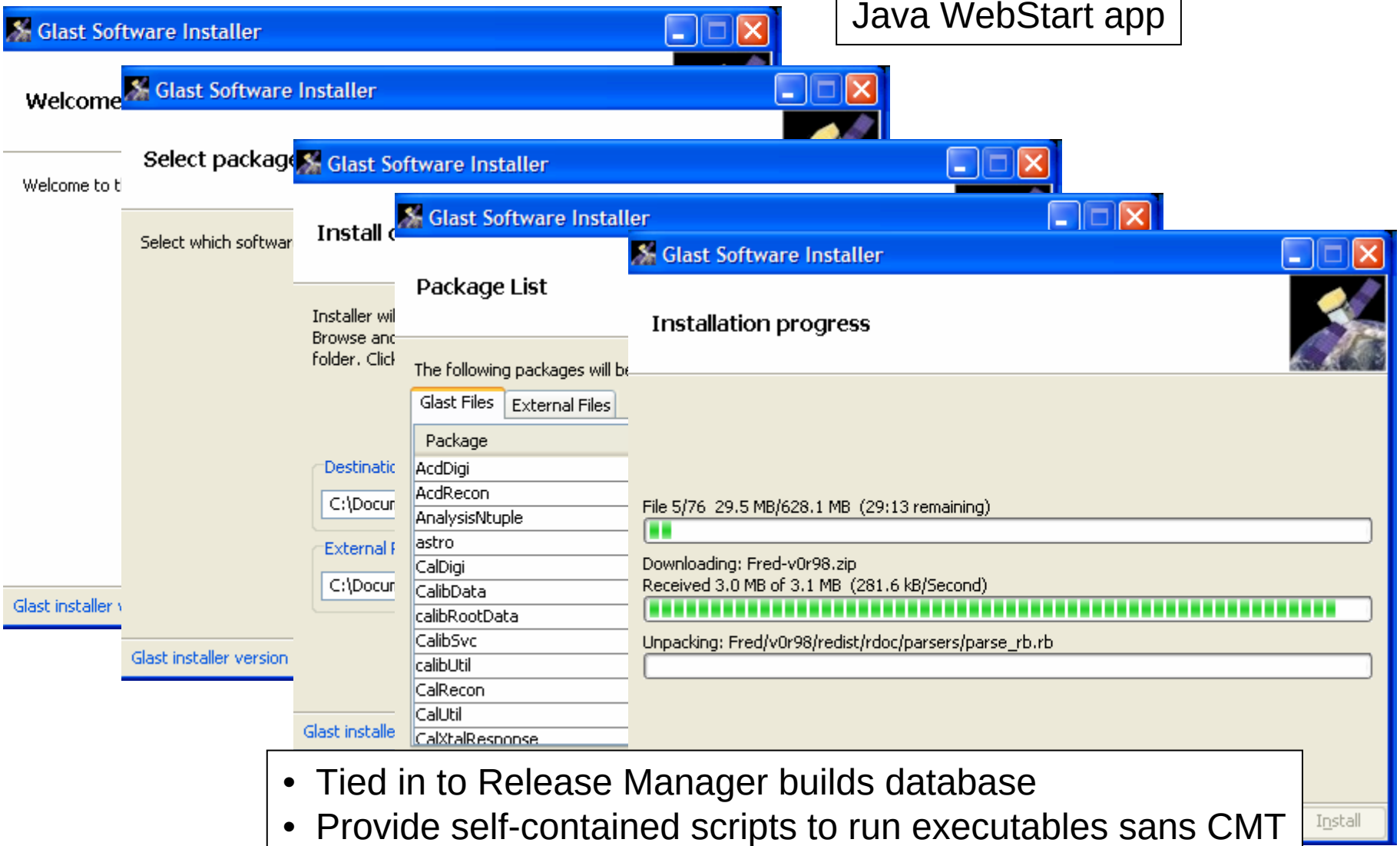
   Your ROOT session should look similar to the following:

Follow on lead from SLD, BABAR, but …

- work with Tech Writer
  - skilled at extracting information from us wackos
  - worries about layout, organization
  - can write good
- we're struggling with apparent conflict of web navigation vs "printed book". Pursuing the former.

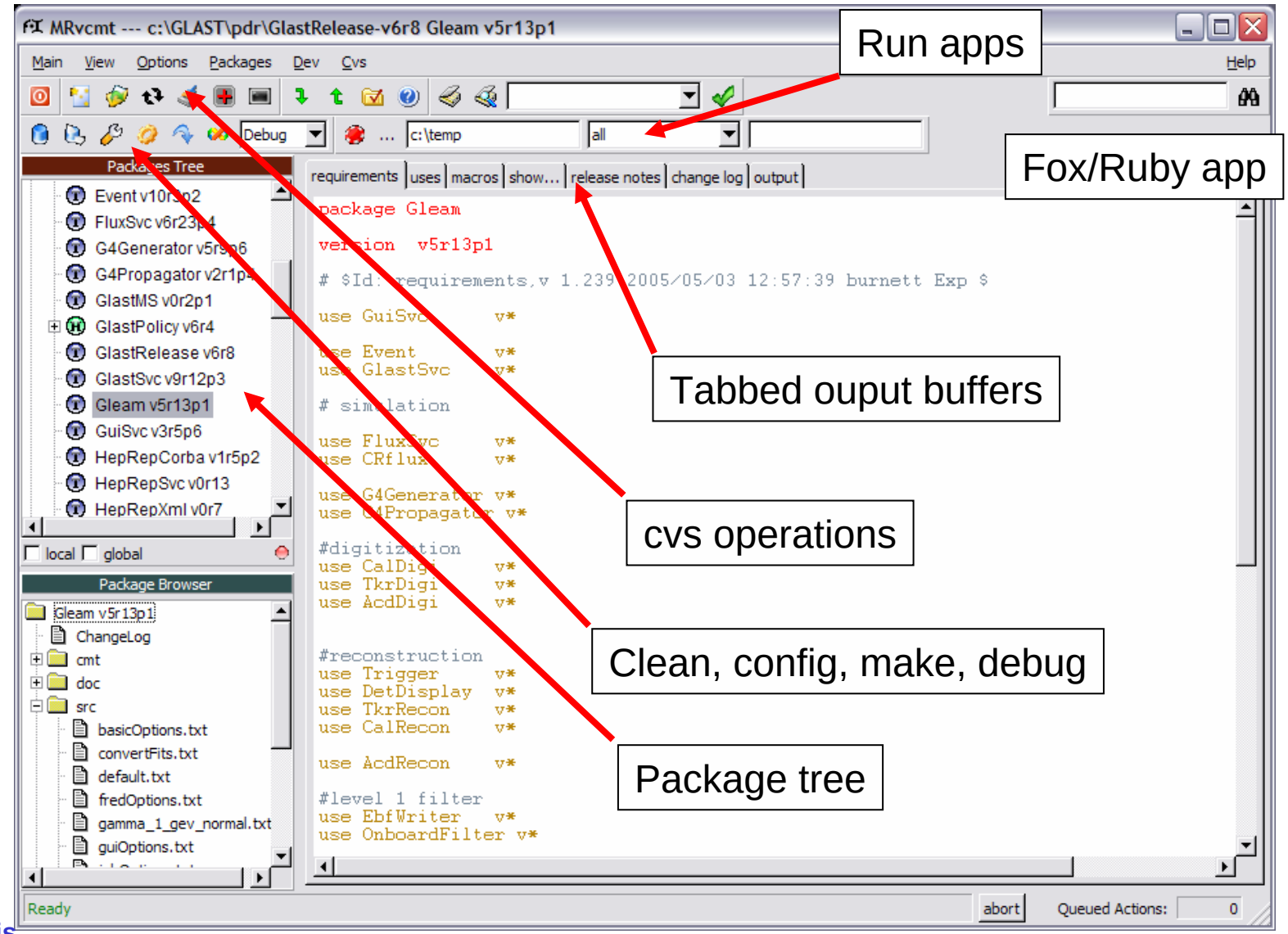
Code Distribution

Java WebStart app



- Tied in to Release Manager builds database
- Provide self-contained scripts to run executables sans CMT

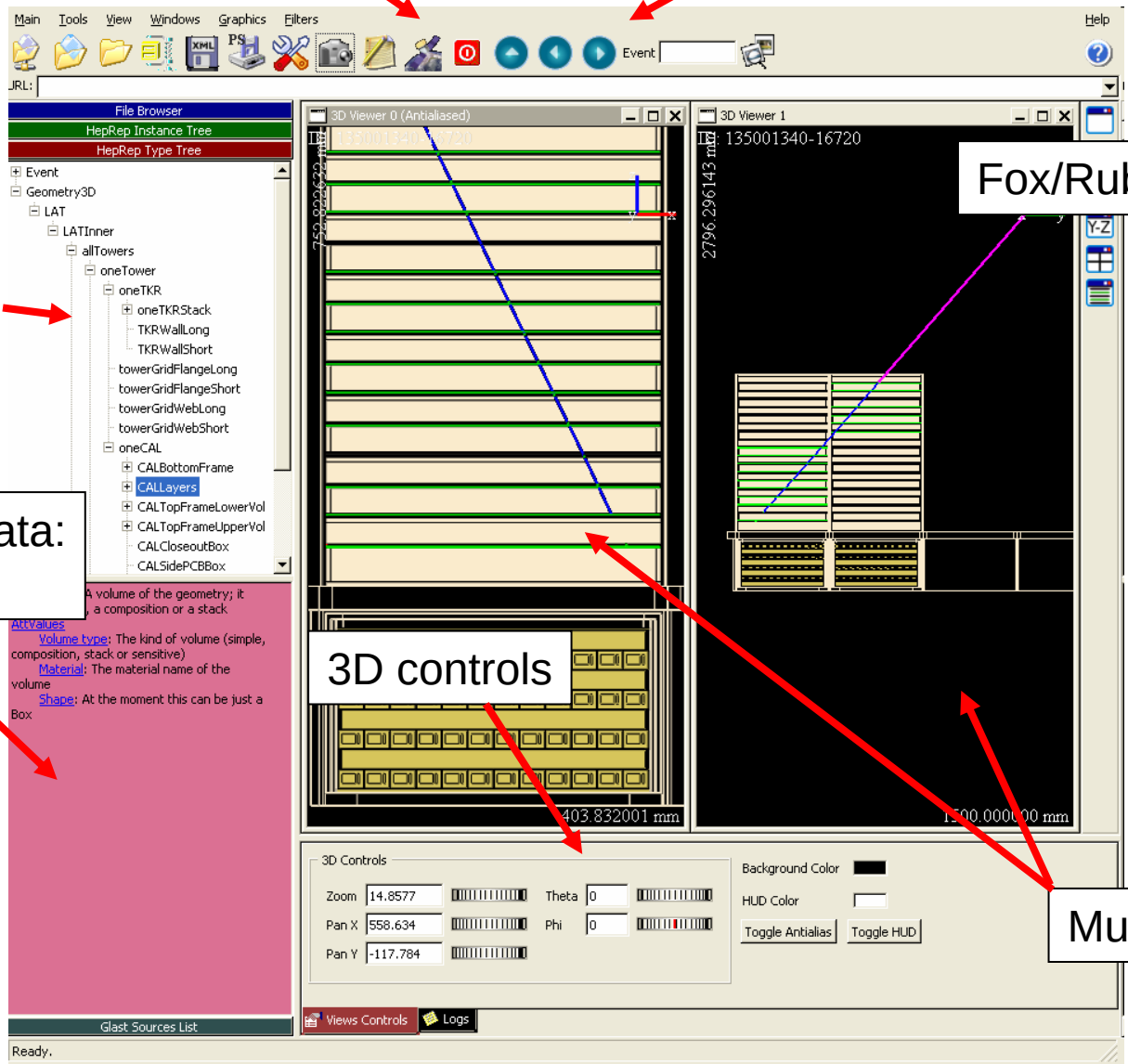
MRvcmt – gui for code development



FRED

GLAST plugin
GlastRelease config

Event control



Fox/Ruby/C++ app

Graphics tree

Graphics metadata:
HepRep

3D controls

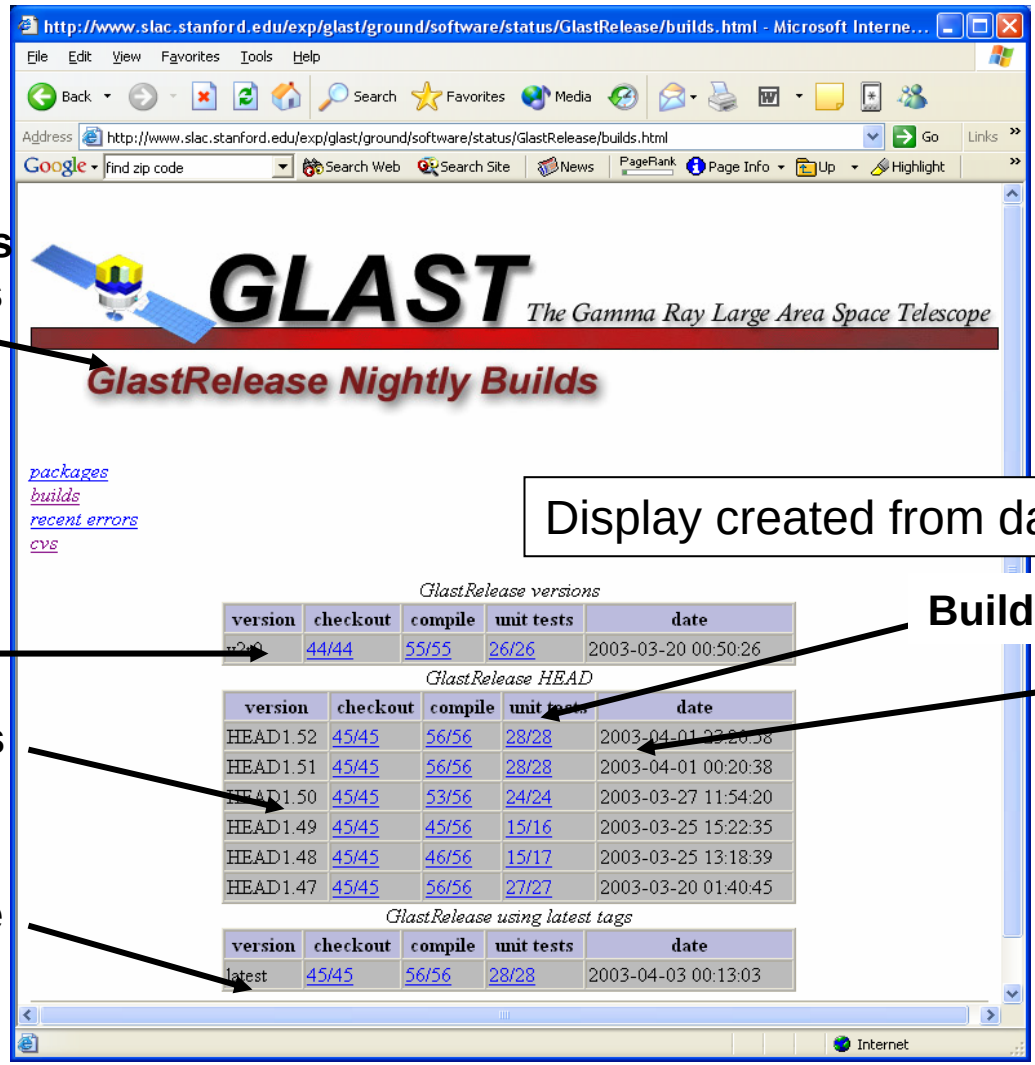
Multiple views

Issues Tracker; CCB; wiki

- User JIRA web issues tracker
  - Commercial product but affordable
  - Handles bugs, features, improvements
  - Full user/group management
  - “roadmaps” for version evolution/project management
- Change Control Board
  - Code used in pipeline – sim/recon; executive scripts; pipeline itself
  - Require documentation of all changes – preferably backed up by JIRA issues
  - Demonstration that fixes work; system tests on sim/recon
  - Using wiki tool to record actions
  - 4-person board – adjudicated by email so far
- Wiki
  - Commercial product (Atlassian – same parent as JIRA)
  - Simple web editing independent of user OS
  - Space management; same groups and users as JIRA

Code Builds

Performing builds for Science Tools also



Display created from database query

Past release

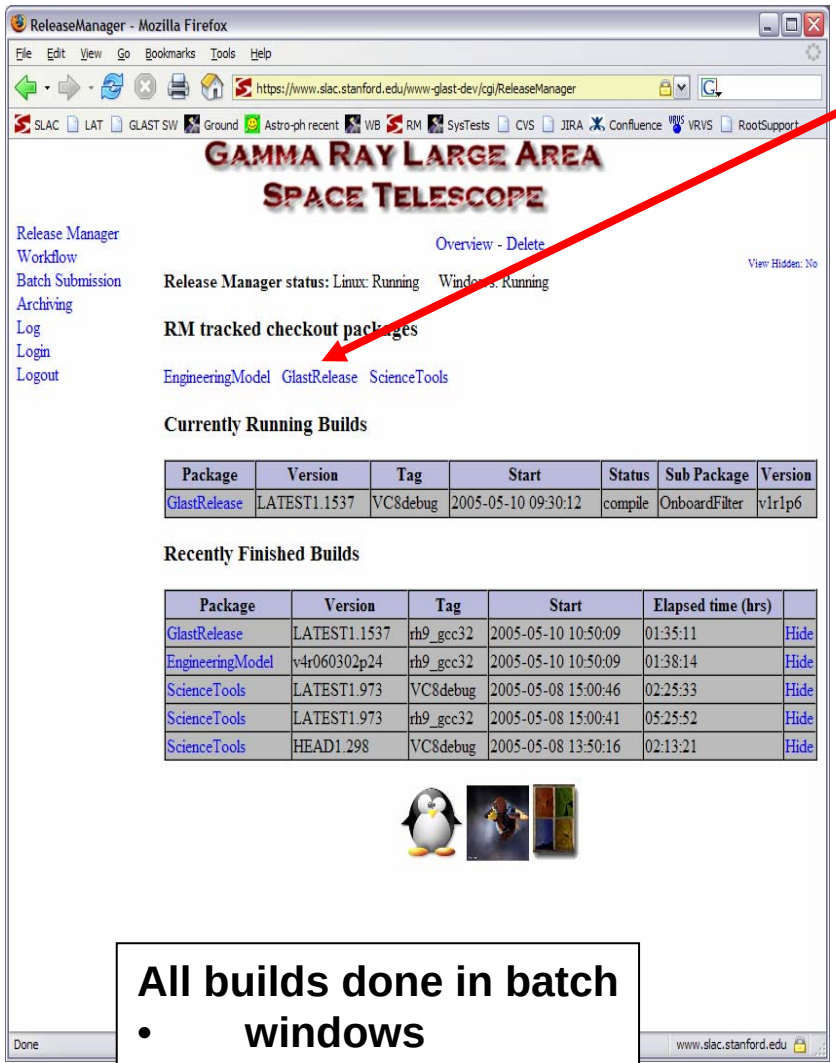
Release in progress

Future release

Build status

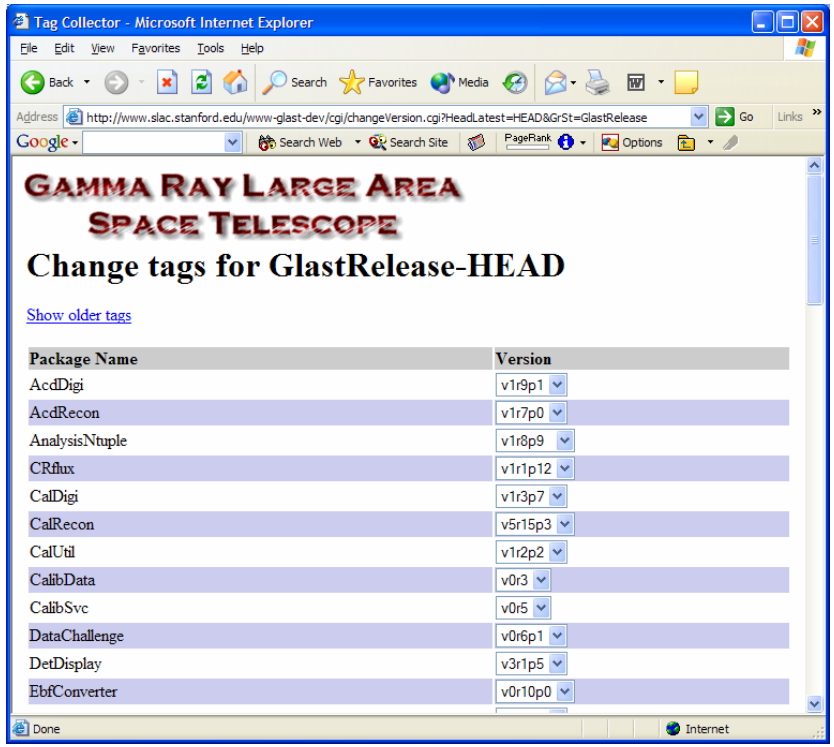
Unit test status

More Code Builds



Multiple packages being tracked

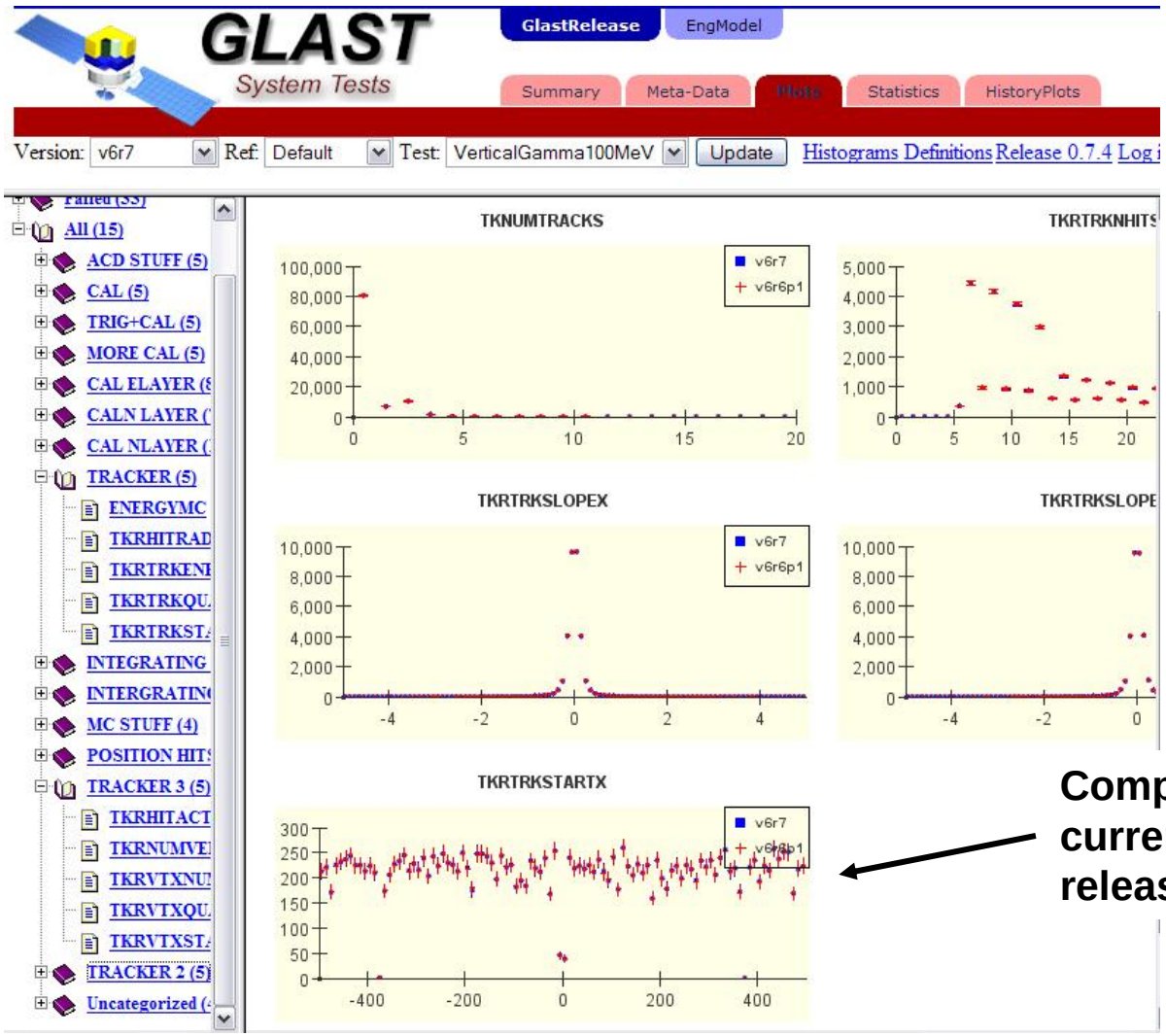
Web tag collector



All builds done in batch

- windows
- linux

System Tests



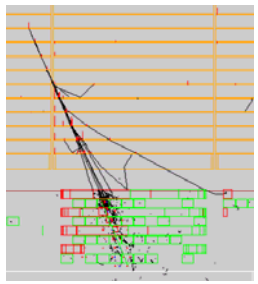
Comparison of current to previous release.

Sim/Recon Toolkit

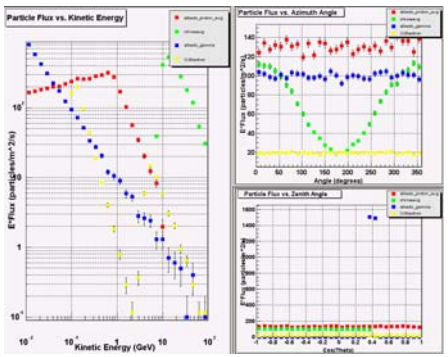
Package	Description	Provider	Status
ACD, CAL, TKR Recon	Data reconstruction	LAT	90% done In use
ACD, CAL, TKR Sim	Instrument sim	LAT	95% done In use
GEANT4	Particle transport sim	G4 worldwide collaboration	In use
xml	Parameters	World standard	In use
Root 4.02.00	C++ object I/O	HEP standard	In use
Gaudi	Code skeleton	CERN standard	In use
doxygen	Code doc tool	World standard	In use
Visual C++/gnu	Development envs	World standards	In use
CMT	Code mgmt tool	HEP standard	In use
ViewCvs	cvs web viewer	World standard	In use
cvs	File version mgmt	World standard	In use

Instrument Simulation and Reconstruction

3 GeV gamma interaction



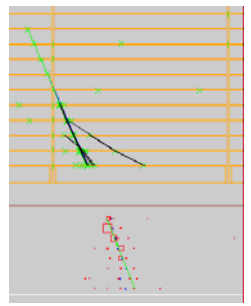
Source Fluxes



Particle Transport

"Raw" Data

Instrument data

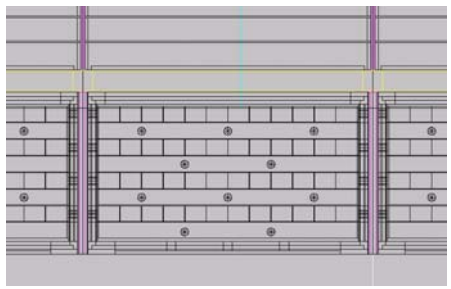


3 GeV gamma recon

Recon

Geometry

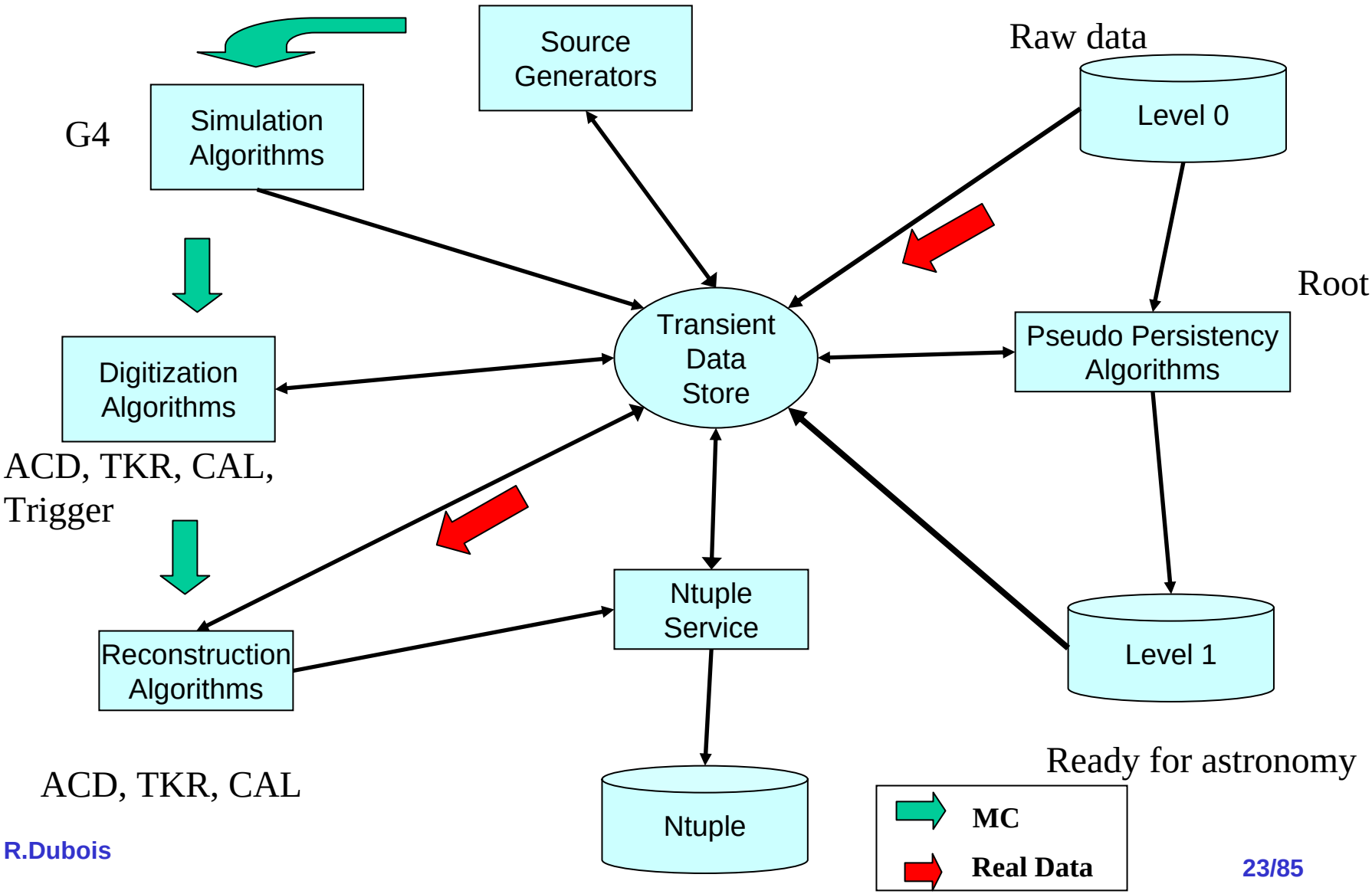
Background Rejection - Particle ID



CAL Detail

Full geometry in xml with
C++ interface
G4 discovers instrument
from the xml

Data flow in the Gaudi framework



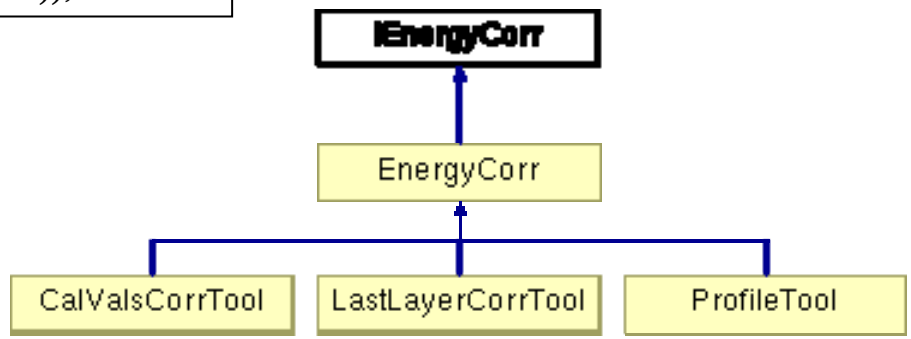
Example of Using Gaudi Tools

```
IEnergyCorr* m_lastLayerTool;
sc = toolSvc()->retrieveTool(m_lastLayerToolName, m_lastLayerTool);

m_lastLayerTool->setTrackSlope(slope);
m_lastLayerTool->doEnergyCorr((*it)->getEnergySum(),(*it));
```

Retrieve tool by name via base class

Refer to base class functions. Does not know which concrete tool it is.



jobOptions

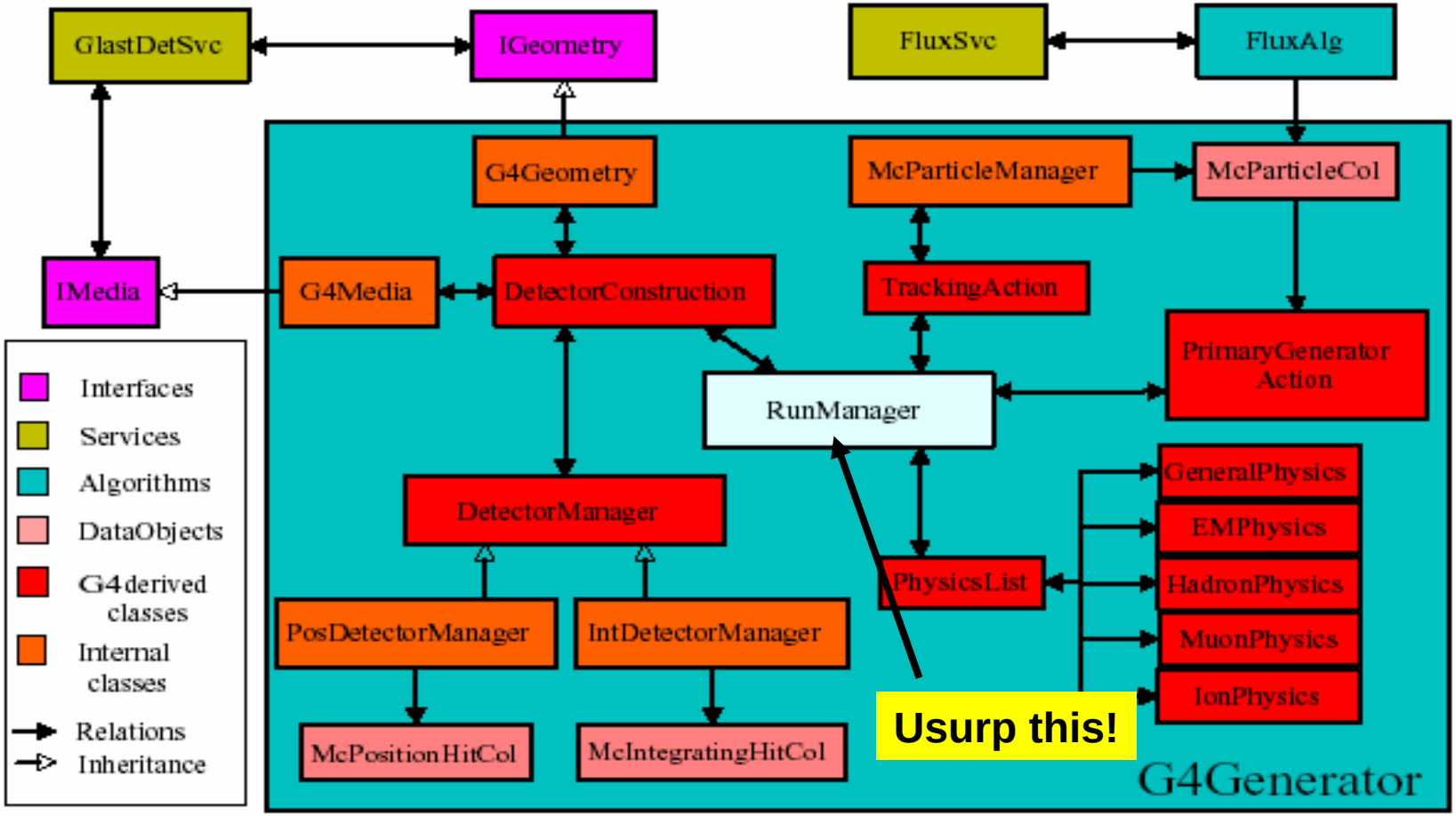
Parameters:

CalClustersAlg.callNumber	this parameter is used to distinguish multiple calls to **CalClustersAlg** (for example, before and after TkrRecon). The default value is 0 .
CalClustersAlg.clusterToolName	name of tool performing clustering. Default is **SingleClusterTool**
CalClustersAlg.lastLayerToolName	name of tool performing last layer energy correction
CalClustersAlg.profileToolName	name of tool performing profile fitting energy correction

Concrete classes that customize behaviour

Tools id'ed by name in ascii config file ("jobOptions")

Gaudi Interface to Geant4



http://www-glast.slac.stanford.edu/software/core/documentation/reviews/G4Generator/g4greview.pdf

Data Challenges

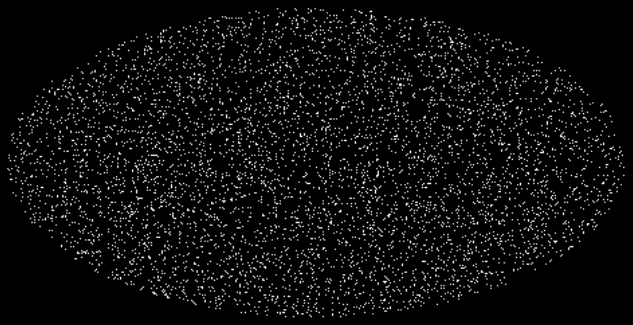
- Ground software is amalgam of HEP instrument software and Astro FTOOLS
- Adopt HEP's "Data Challenges" to create a series of end-to-end studies: create a progression of ever more demanding studies
- DC1. Modest goals. Contains most essential features of a data challenge.
  - 1 simulated day all-sky survey simulation
  - find GRBs
  - recognize simple hardware problem(s)
  - a few physics surprises
  - Exercise all the components
- DC2, start beginning of CY06. More ambitious goals. Encourage further development, based on lessons from DC1. One simulated month.
- DC3, in CY07. Support for flight science production.

DC Components

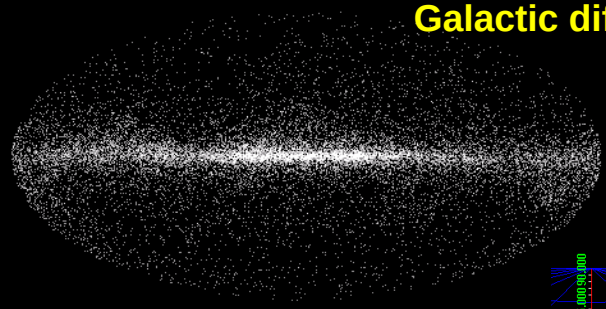
- **Focal point for many threads**
  - **Orbit, rocking, celestial coordinates, pointing history**
  - **Plausible model of the sky**
  - **Background rejection and event selection**
  - **Instrument Response Functions**
  - **Data formats for input to high level tools**
  - **First look at major science tools – Likelihood, Observation Simulator**
  - **Generation of datasets**
  - **Populate and exercise data servers at SSC & LAT**
  - **Code distribution on windows and linux**
- **Involve new users from across the collaboration**
- **Teamwork!**

The Simulated DC1 Sky

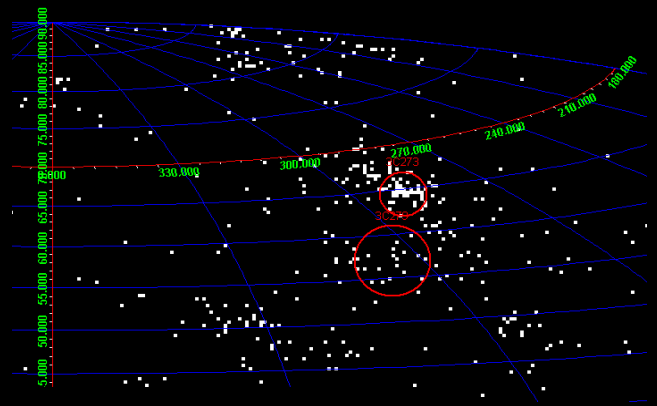
Extragalactic diffuse

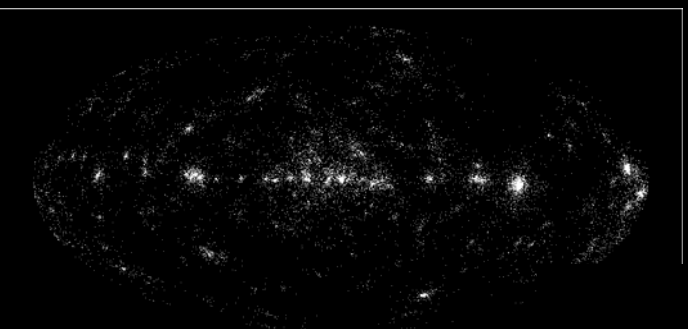


Galactic diffuse



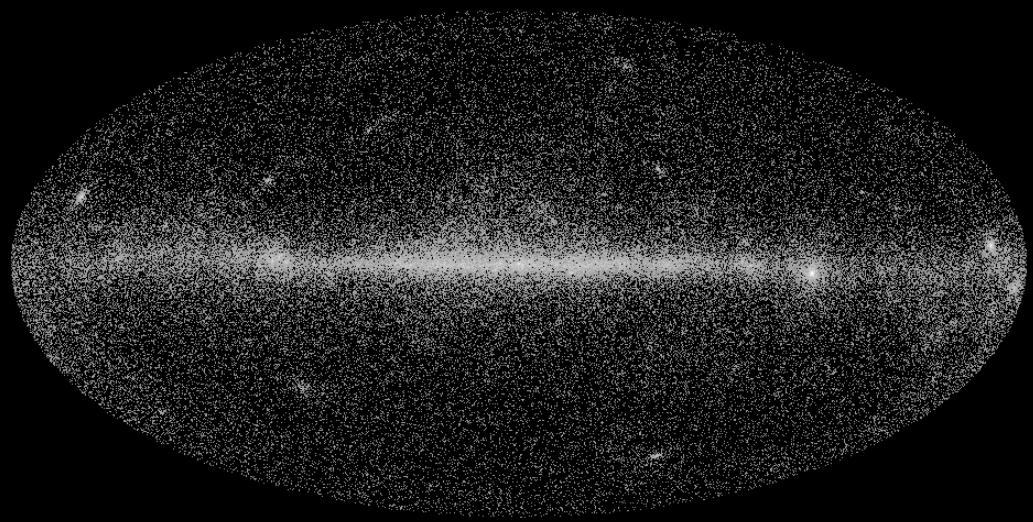
Fiddling 3C273/279





EGRET 3EG

Our Sky



GLAST Large Area Telescope:

Reconstruction

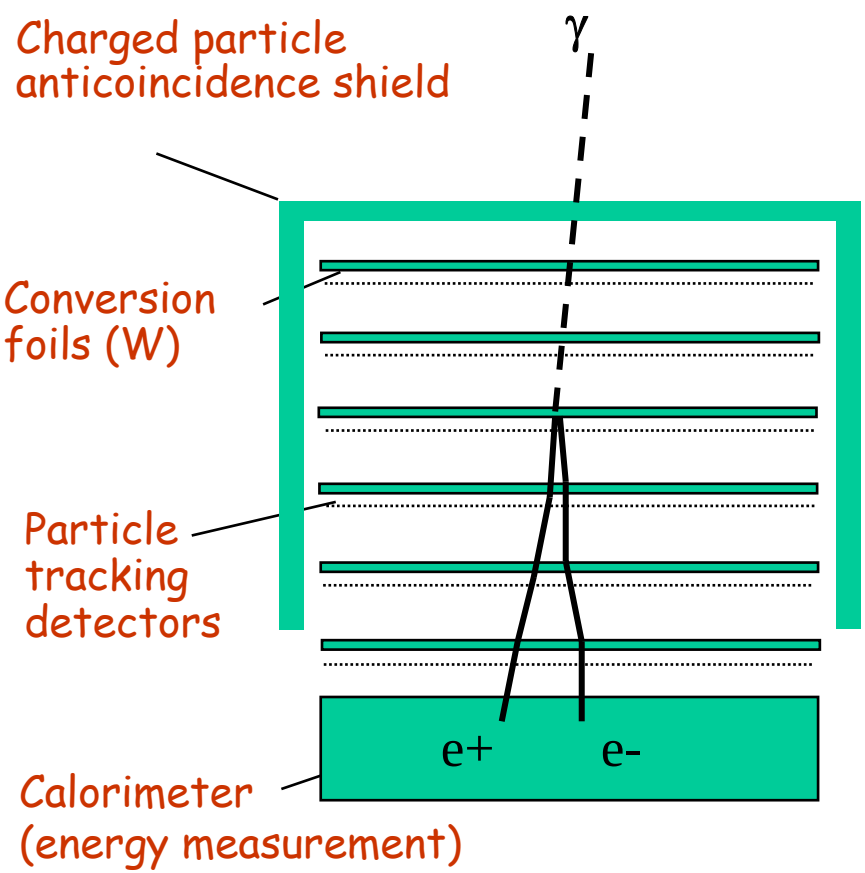
Tracy Usher
Stanford Linear Accelerator Center
usher@slac.stanford.edu

http://www-glast.slac.stanford.edu/software

GLAST Reconstruction

Anatomy of a "Typical" Event

Pair production is the dominant photon interaction in our energy range



- **Reconstruction Goals:**
  - **Incident Gamma Direction and Energy**
  - **Reject Backgrounds**
- **Incident Gamma converts in the tracker**
  - **In particular, conversion occurs in one of the converter foils – ie at a well defined location**
- **Resulting electron-positron pair range out of tracker (TKR)…**
  - **No magnetic field, tracks are "straight lines"**
  - **Resulting two tracks "point" back to incident Gamma**
- **And into the CsI Calorimeter (CAL)**
  - **Measures total energy of electron-positron pair**
  - **= Gamma energy**
- **Surrounding Anti-Coincidence Detector (ACD) vetoes any wayward charged particles**

GLAST Reconstruction

What makes it challenging…

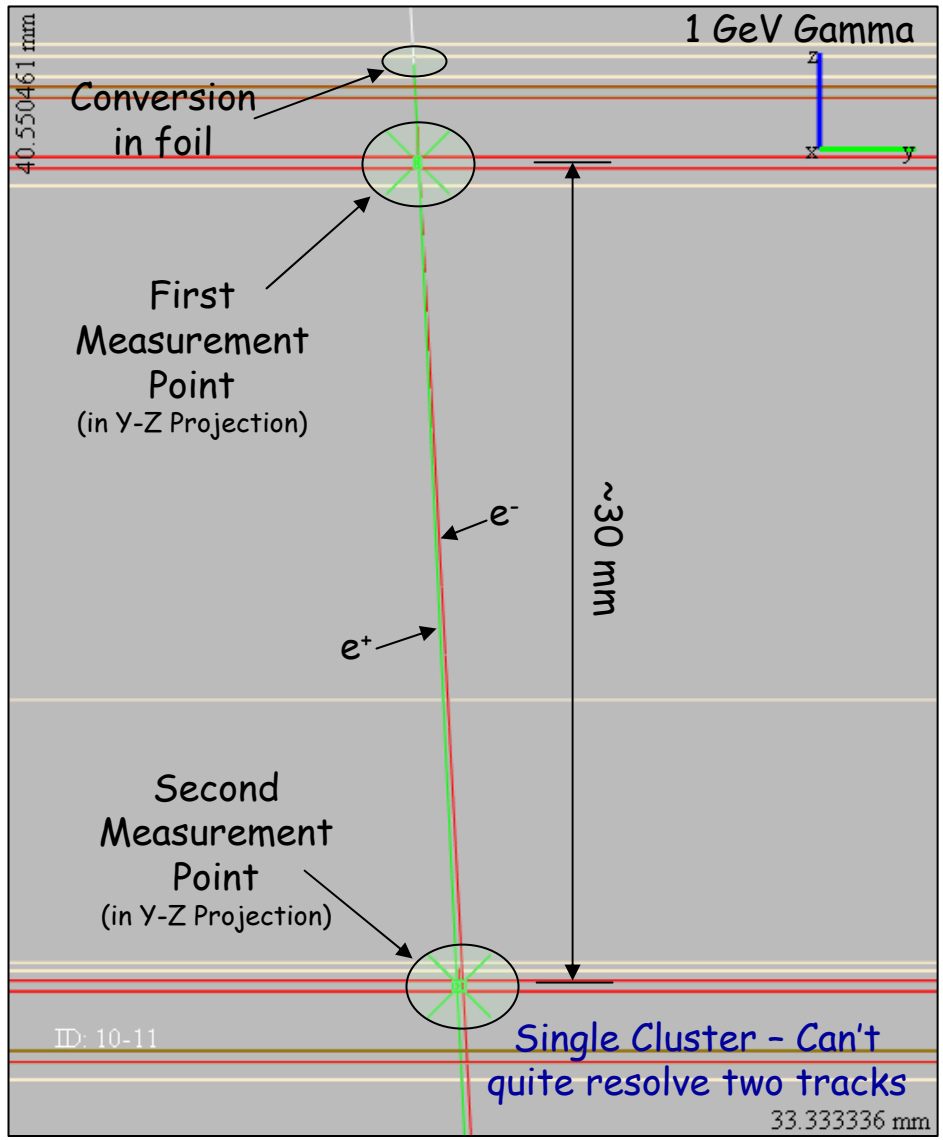
- **Track Opening Angle ~0**
  - **Resolve**
    ~ 2 * 228 um / 30 mm = ~15 mr

Strip Pitch

~ Tray Spacing

< ~50 MeV photons to resolve tracks without "help"

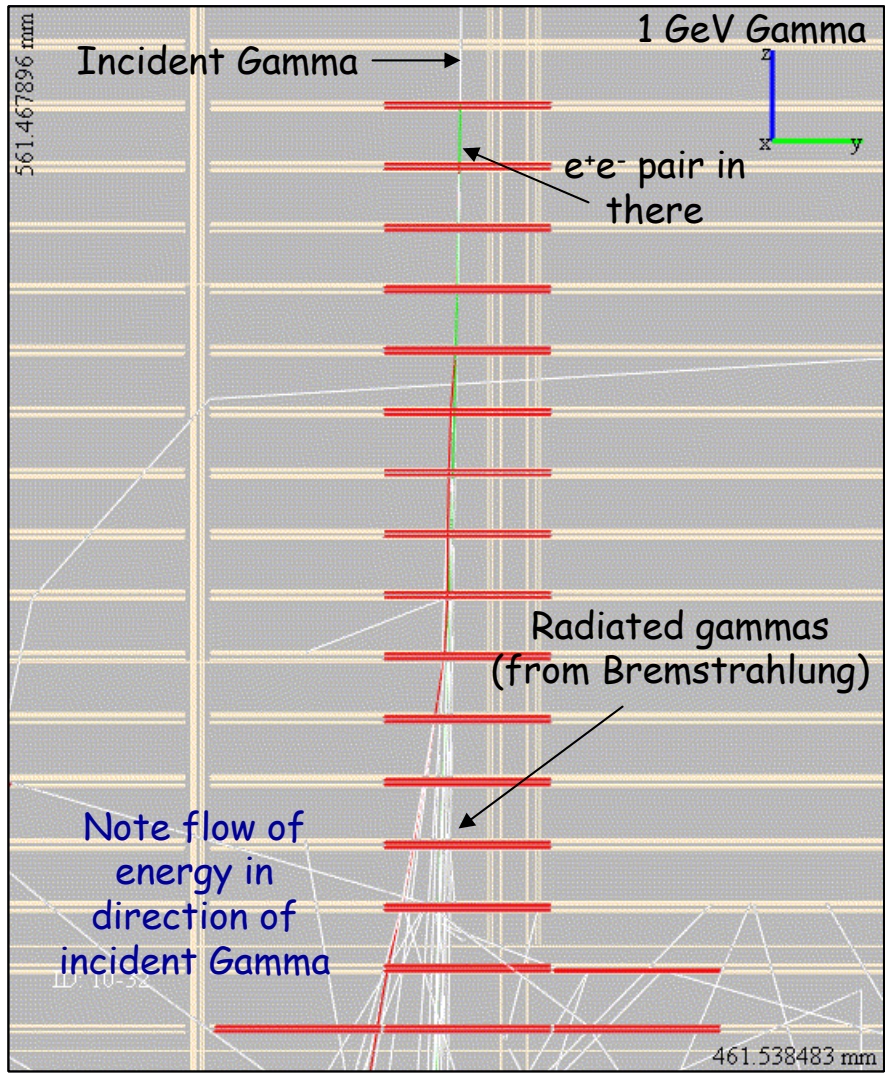
- **Looking for "v"s may not be the correct strategy for gamma direction reconstruction**
  - **Well… see next slides…**



GLAST Reconstruction

What makes it challenging…

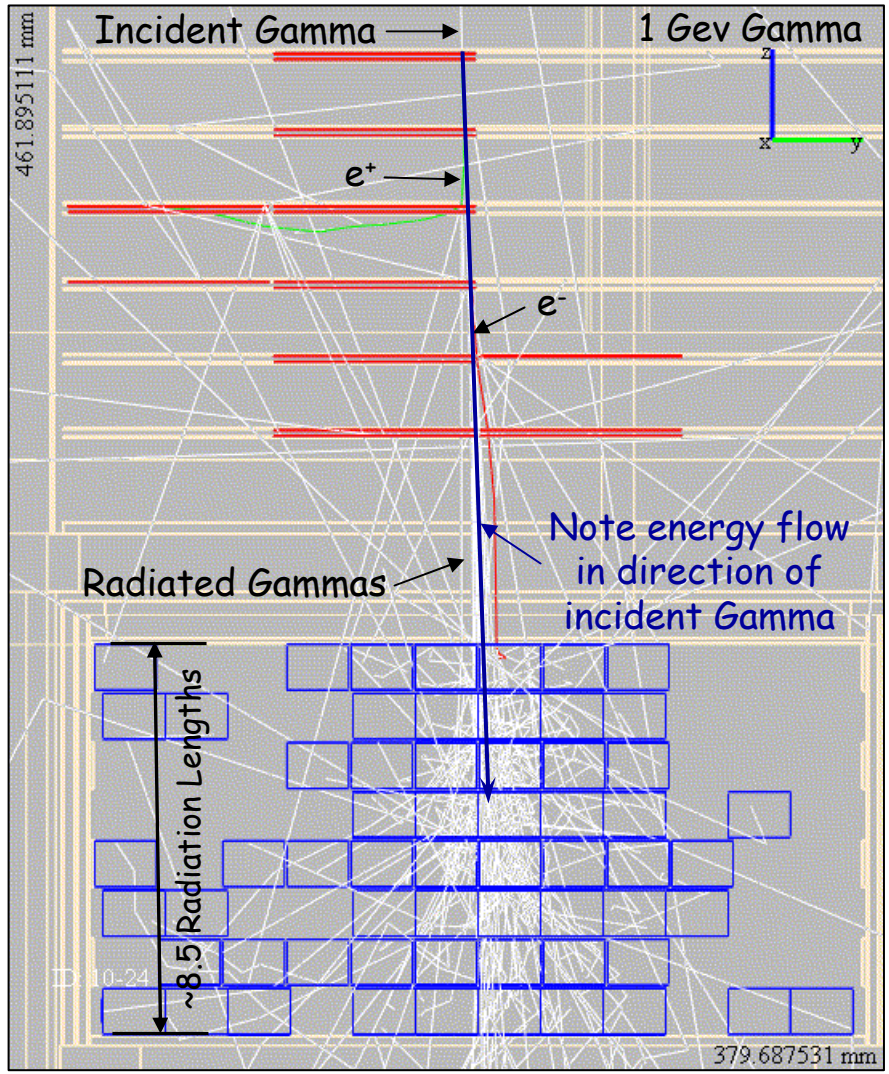
- **Tracker has a lot of material**
  - **Actual tracker is ~ .3 rl**
    - **Could live with this…**
  - **Converter foils are ~ 1.1 rl**
    - **Love them: convert gamma**
    - **Hate them: tracking electrons**
  - **Total ~ 1.4 rl**
    - **For particles traversing active area of tracker**
    - **Does not include walls between towers, etc.**
- **Issues to deal with**
  - **Gammas can (and do) convert outside the foils**
  - **e^+e^- pair interact with tracker**
    - **Multiple scatter**
    - **Primary e^+ or e^- can stop in the tracker**
    - **e^+ and e^- radiate energy**
    - **etc.**



GLAST Reconstruction

What makes it challenging…

- **Calorimeter Issues**
  - **Measure Event Energy – Not Track Energy(ies)**
    - **Don't have resolution to separate**
    - **Large fraction of measured energy from Brems**
    - **Implications for determining gamma direction when you do have two track events…**
  - **Measure Fraction of Event Energy**
    - **Energy "loss"**
      - **in tracker**
      - **Leaking out of Calorimeter**
    - **Significant contribution at**
      - **lower energies (e.g. < 1 GeV)**
      - **for conversions starting higher in the tracker**
    - **Must augment total energy determination with contribution from tracker**



GLAST Reconstruction

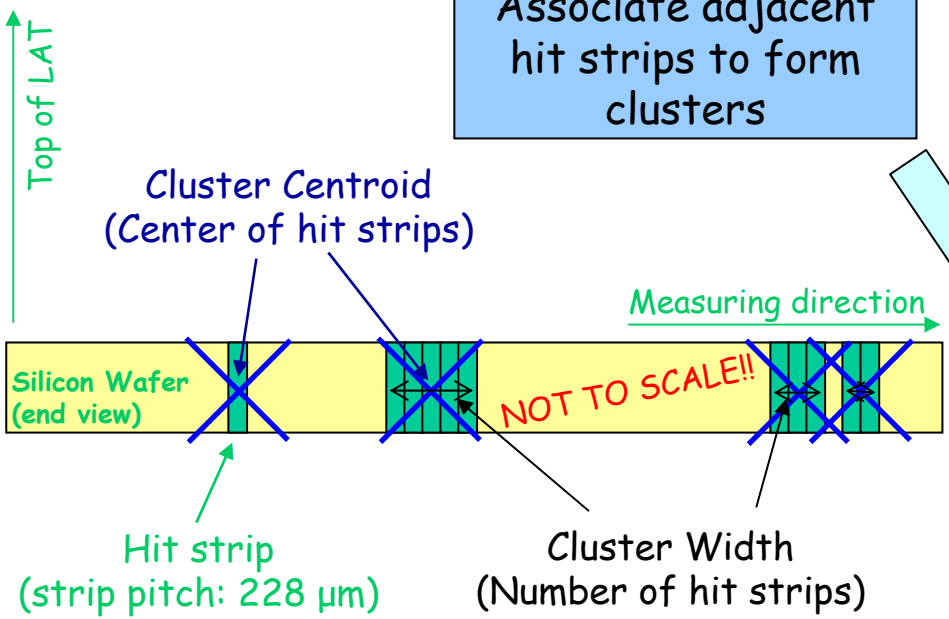
- Summary: Slightly more complicated than first thought
- But still follow the “Standard” HEP Approach
  - Tracking
    - Change Goal slightly
      - Still look for two tracks
        - » Multiple Scattering separates them
      - But emphasize the “longest, straightest” (highest energy) track
    - Algorithms to assign energy to tracks in final fits
    - Provide enough information to reject “bad” events
  - Calorimetry
    - Look for total event energy
    - Algorithms to correct for
      - Losses in the tracker
      - Leakage
      - Etc.
  - Both: Algorithms to help reject background

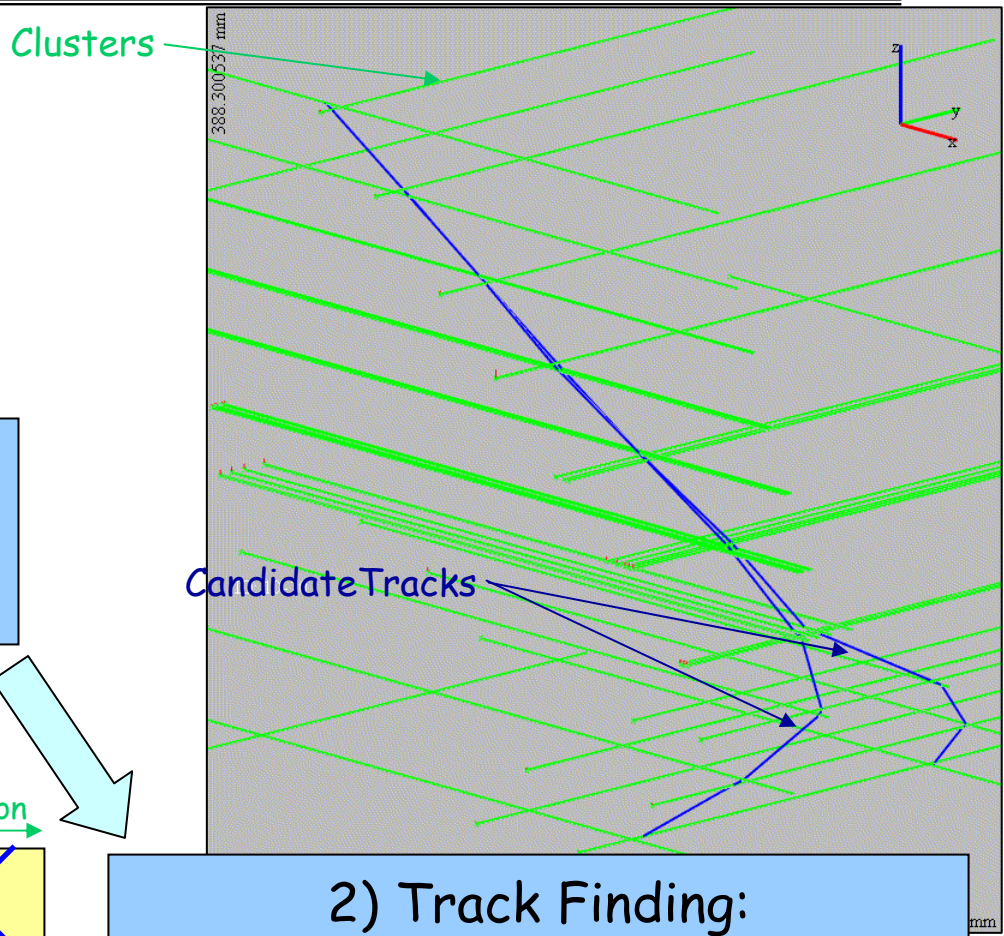
Tracking Overview

"Standard" HEP Approach

Raw Data
1) Hit Strips
2) "ToT"

1) Clustering:
Associate adjacent hit strips to form clusters





2) Track Finding:
Assume start point, direction and energy
Follow track and attach clusters allowing for deviations due to multiple scattering, etc.

Tracking Overview

"Standard" HEP Approach

Track Finding
(see previous page)

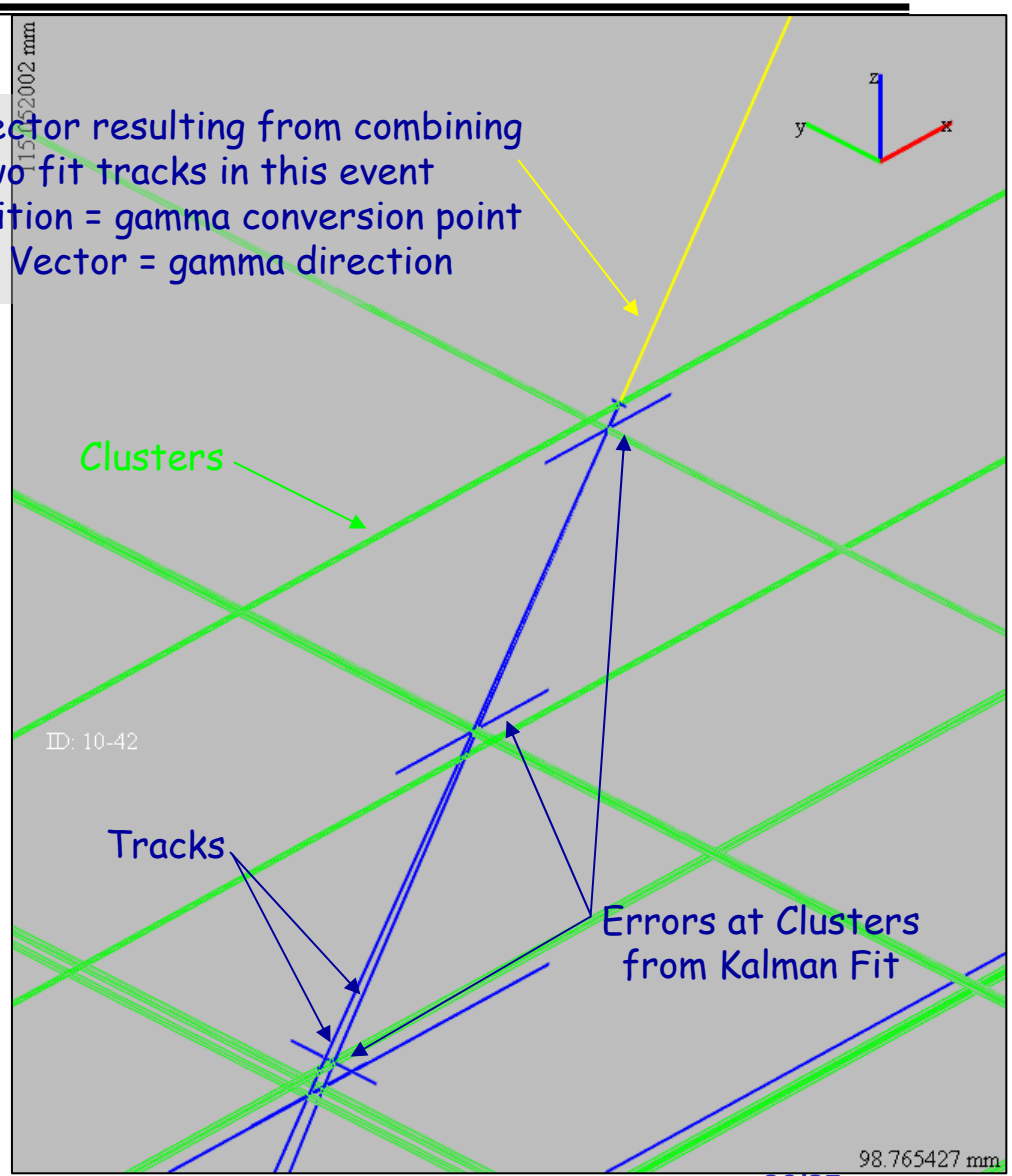
"Vertex" vector resulting from combining the two fit tracks in this event
Vertex Position = gamma conversion point
Vertex Vector = gamma direction

3) Track Fit:

Kalman Filter Fit to get final Track Parameters

4) "Vertexing":

Combine Tracks to determine conversion point and direction of the incident Gamma
Note: a "vertex" can consist of only one track



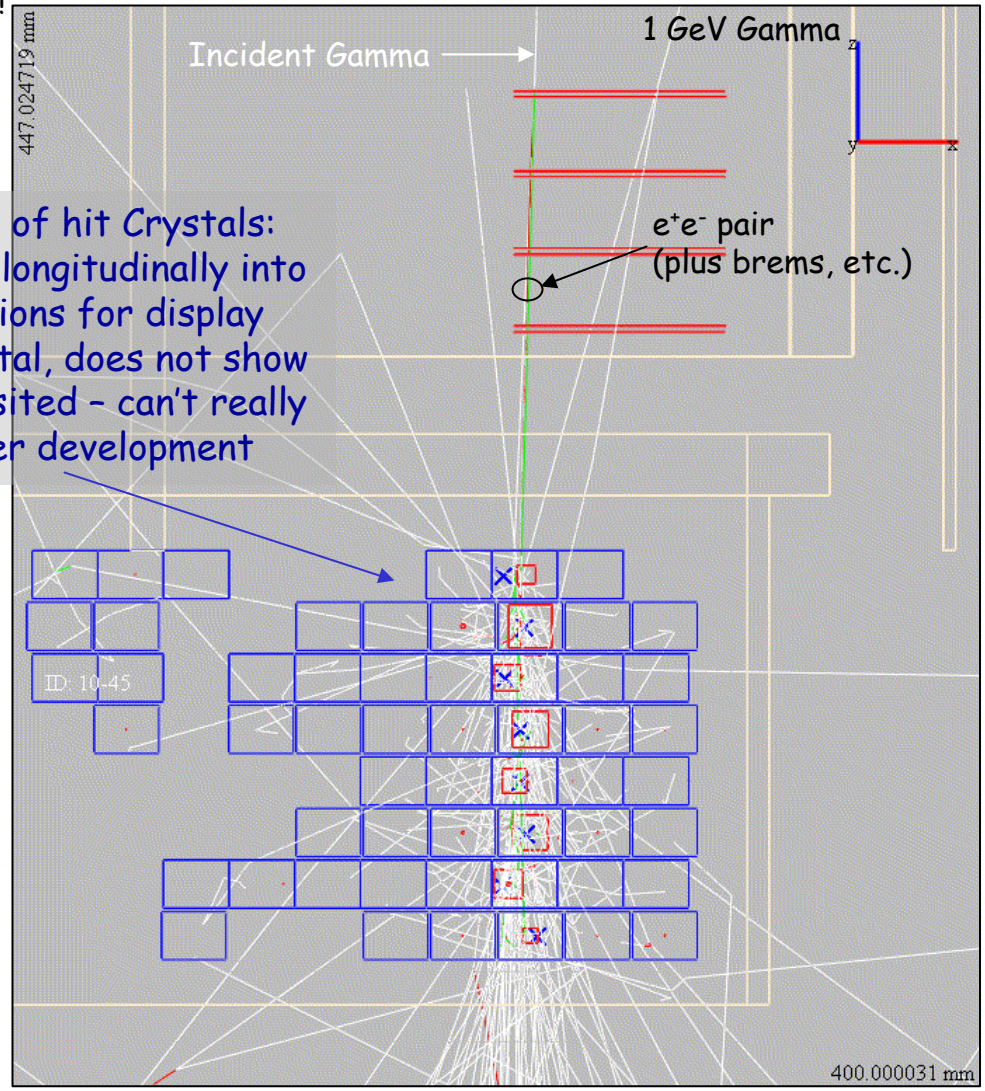
Calorimetry Overview

"Raw" Data
Corrected Energy, "Position" per Crystal (Xtal)

Corrections/Calibrations applied before this step!

Clustering:
Associate hit Xtals together to form clusters
Gamma Energy: All hit Xtals are from incident Gamma – Energy is sum over all hit Xtals
Backgrounds: Isolated sets of hit Xtals from own "cluster"

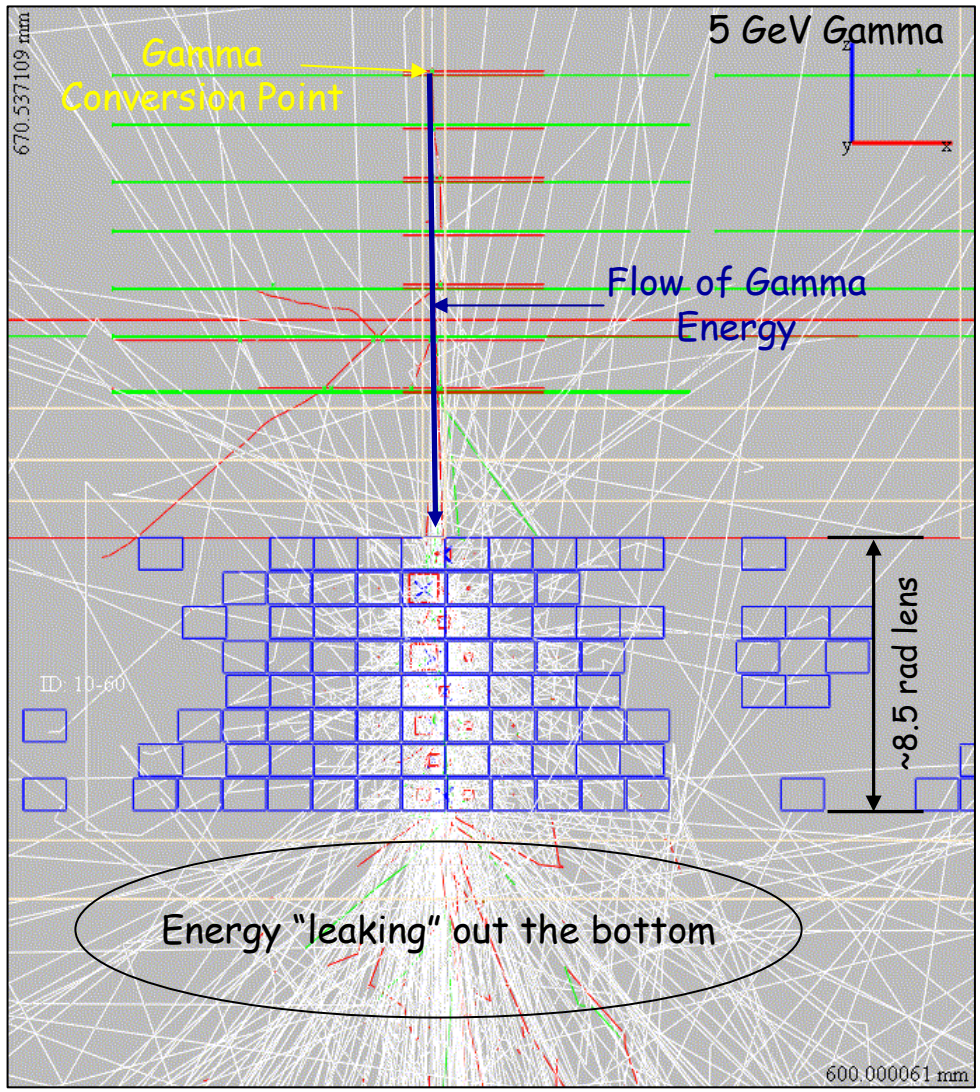
MC Display of hit Crystals:
Divides Xtal longitudinally into eight sections for display
Shows Hit Xtal, does not show energy deposited – can't really see shower development



Calorimetry Overview

Energy Correction Algorithms:
1) Shower Profile
2) Inefficiencies due to Geometry
3) Leakage
4) etc.

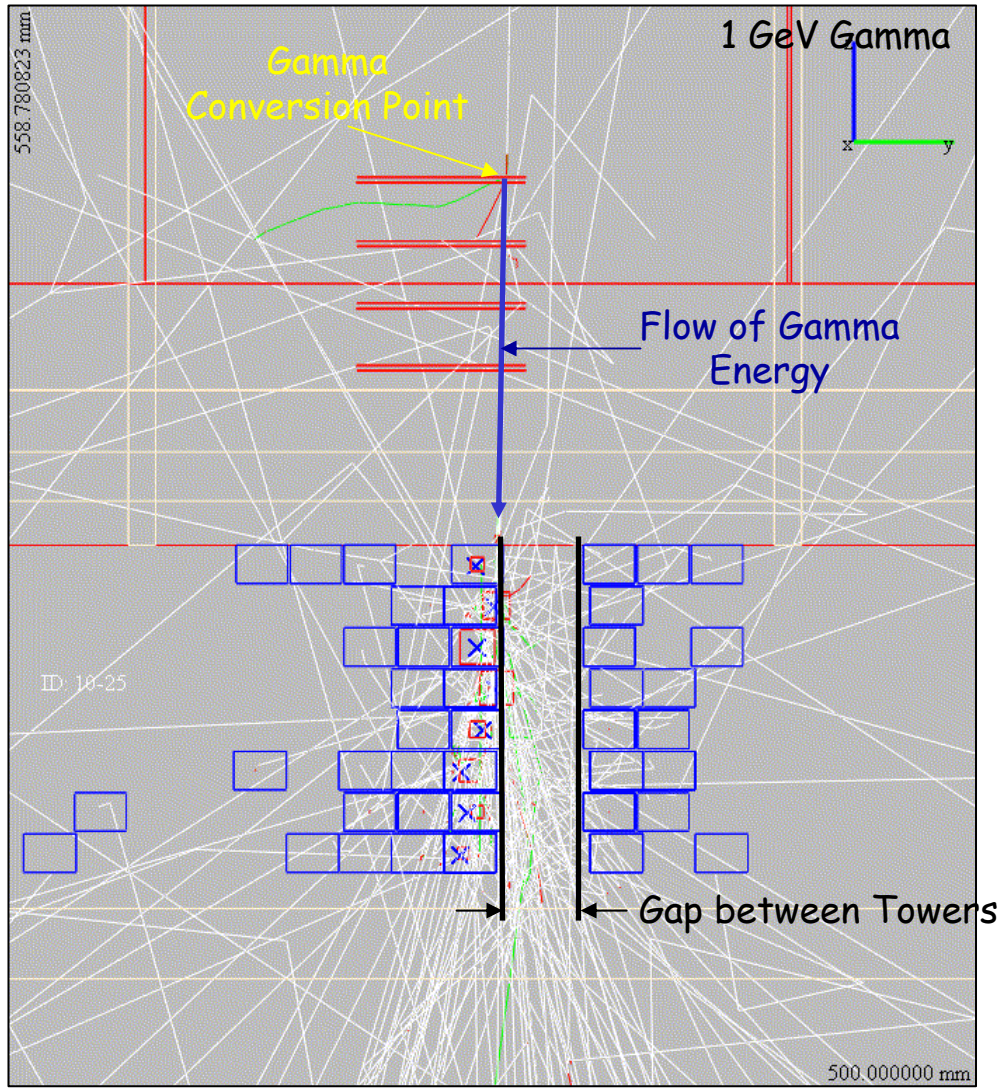
(see next slides)



Calorimetry Overview

Energy Correction Algorithms:
Inefficiencies due to Geometry

Given the direction of energy flow (from the reconstructed Gamma direction), can apply geometric corrections to account for energy "lost", e.g. between towers

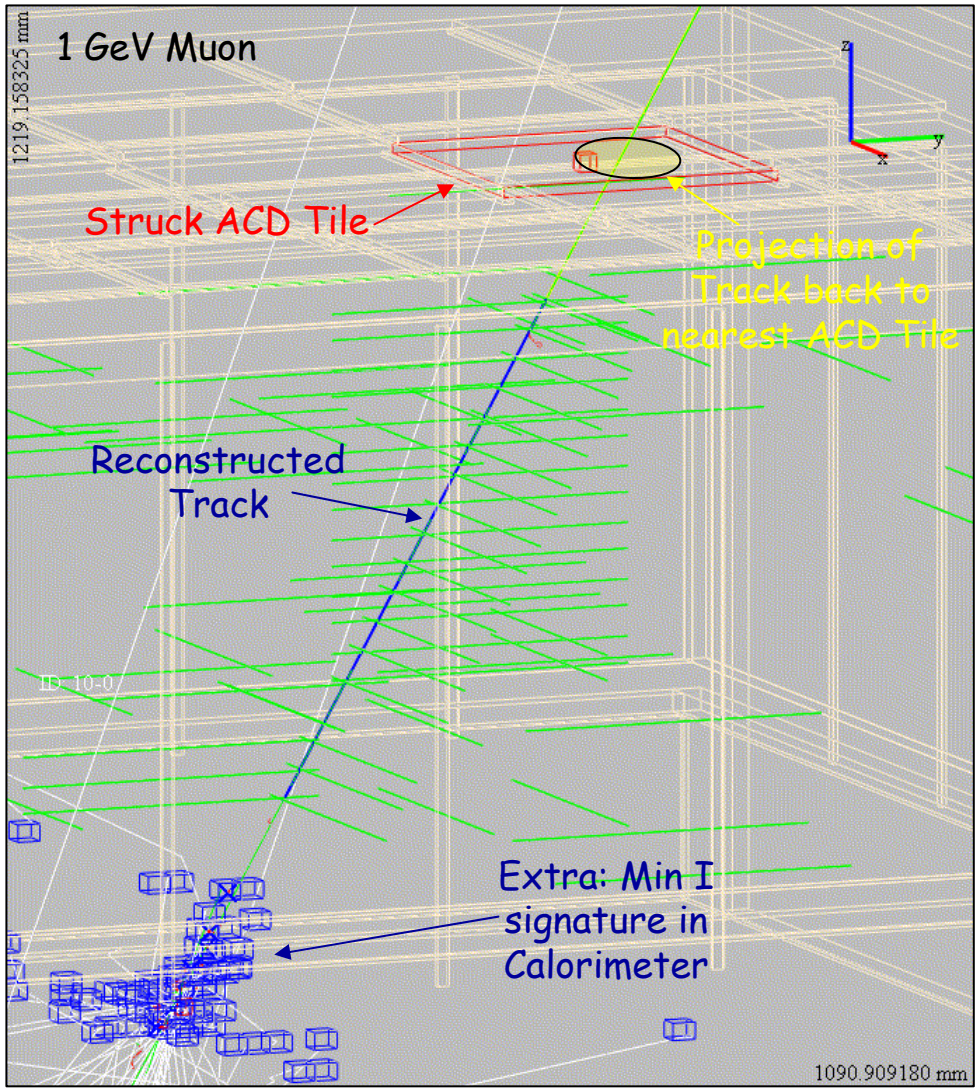


Putting It All Together

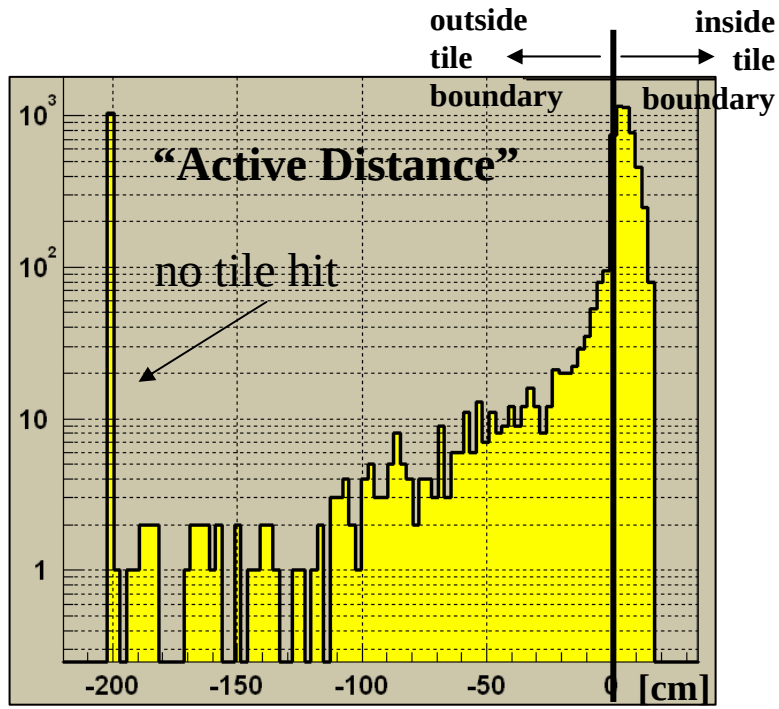
- Chicken…
  - Track Finding/following needs starting values:
    - Initial Position
    - Initial Direction
    - Initial Energy
  - This from the Calorimeter…
- Or Egg?
  - Energy Correction algorithms need gamma direction
  - This from the tracking…
- Solution: Iterative Reconstruction
  - First Pass
    - Calorimeter Reconstruction
      - Through Clustering
        - » Total Energy
        - » Cluster Centroid
        - » Cluster Axis
    - Tracker Reconstruction
      - Track Finding/Following
      - Track Fit and Vertexing
        - » Good enough for 2nd pass
  - Second Pass
    - Calorimeter Reconstruction
      - Energy Correction Algorithms
    - Tracker Reconstruction
      - Track Fit and Vertexing
        - » Use “improved” energy

Background Rejection

Example: Charged Particles in Tracker



- Project Track to plane of struck tile
- Calculate distance to nearest edge
- Sign
  - Positive if track projection inside the tile
  - Negative if track projection outside the tile
- Reject if inside the tile



End of Reconstruction

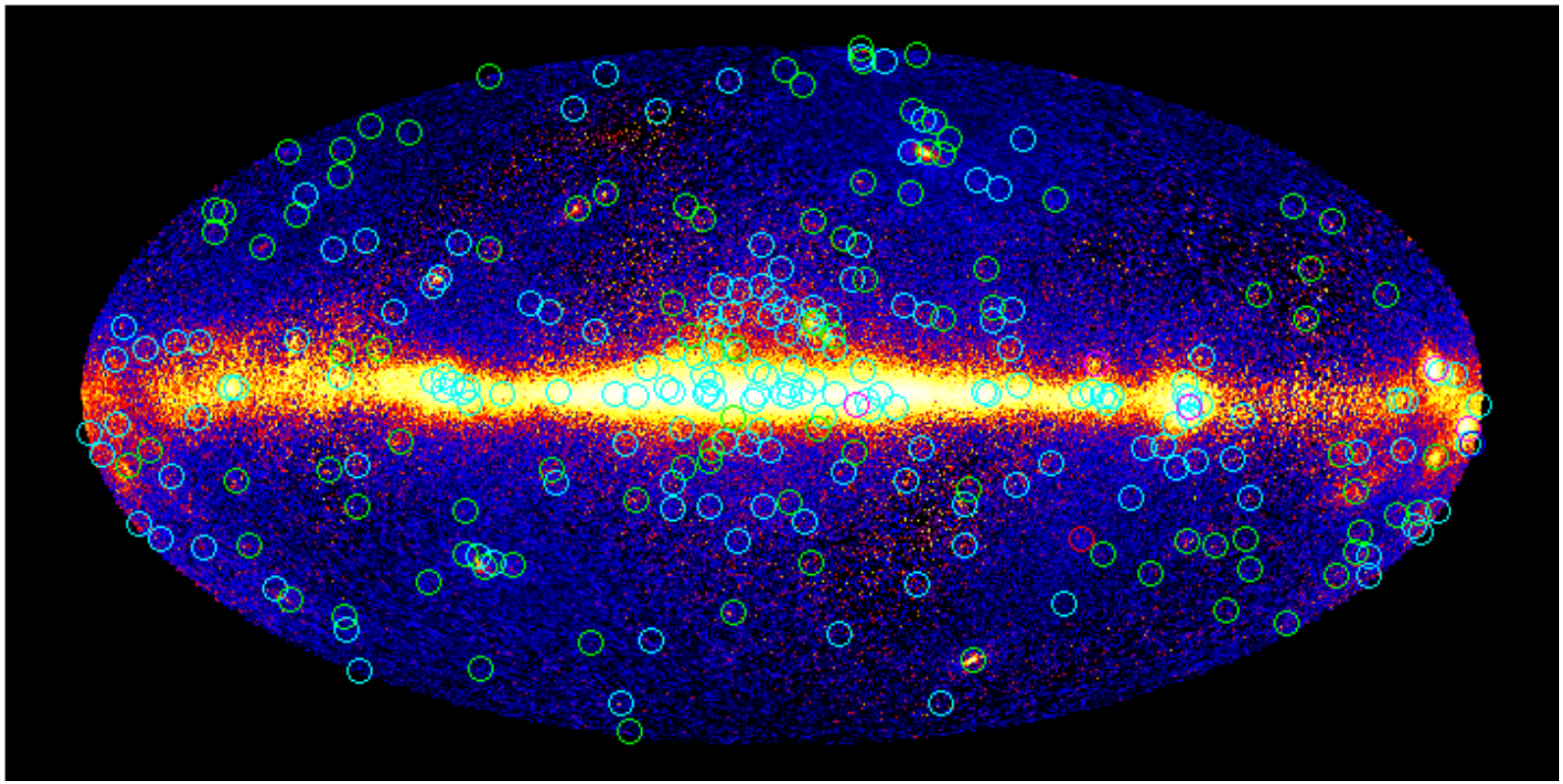
- **Two Levels of Output at end of Reconstruction:**
  - **Root Trees**
    - **Basically, all the output of all steps of reconstruction**
    - **Enough information to read back in and continue reconstruction from that point**
    - **Detailed offline analysis for reconstruction algorithm improvements**
    - **Main component of System Tests**
  - **Output Ntuple with two branches**
    - **A detailed branch which contains enough information for checking of reconstruction performance**
    - **The analysis branch which is passed on to the next stage…**

LAT Science Tools for Gamma-Ray Astronomy

James Chiang
GLAST Science Support Center
jchiang@slac.stanford.edu

The Gamma-Ray Sky

- **EGRET All-Sky Map and 3rd EGRET Catalog:**

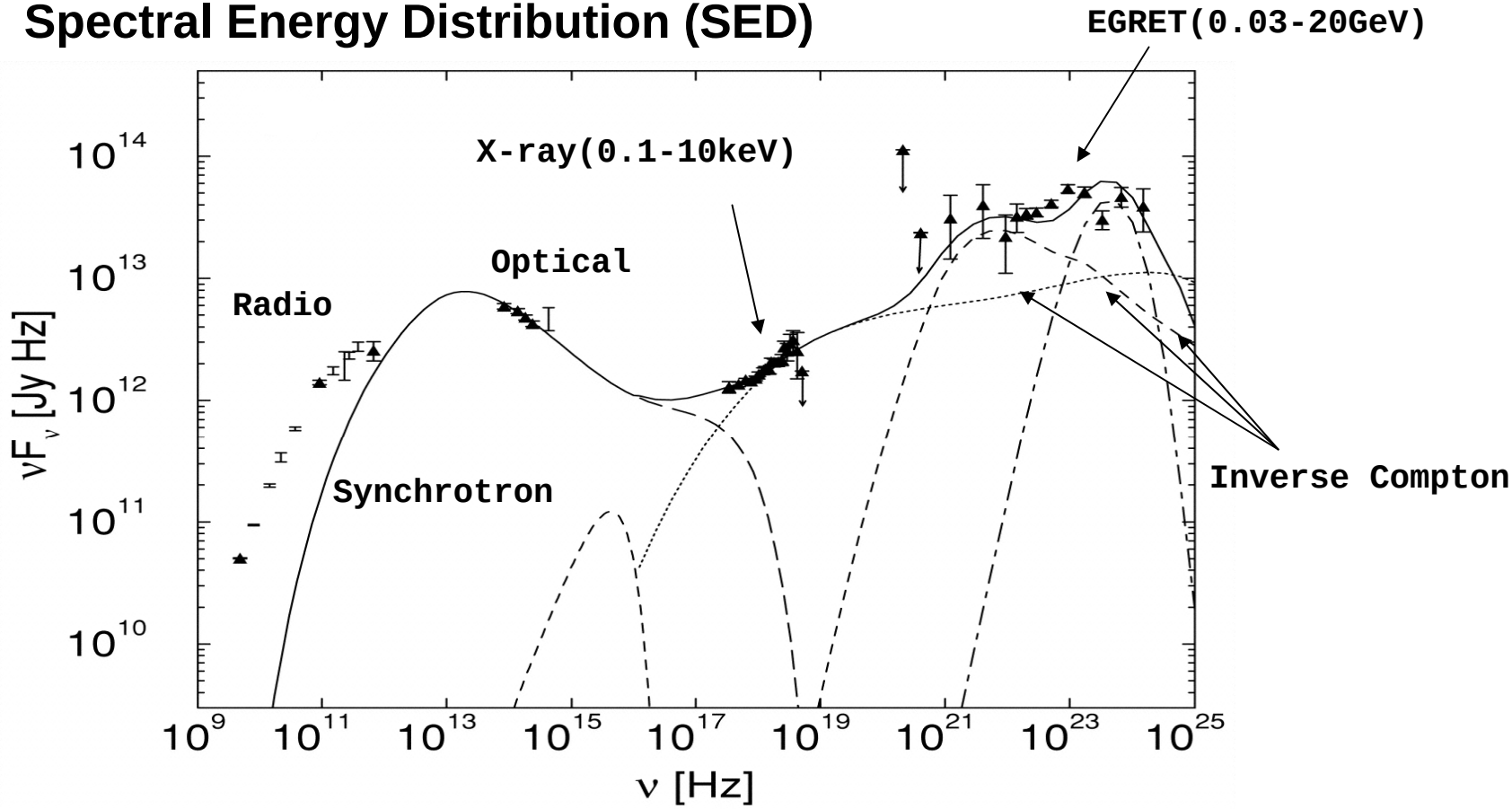


	3EG	GLAST		3EG	GLAST
AGNs	94(67)	~3000	Unids	170	O(10^3)?
Pulsars	5	~ 10s	Sol. Flare	1	?
galaxies	1(LMC)	>1?			

Dark Matter, SNRs, etc.

Example Source Class: Blazars

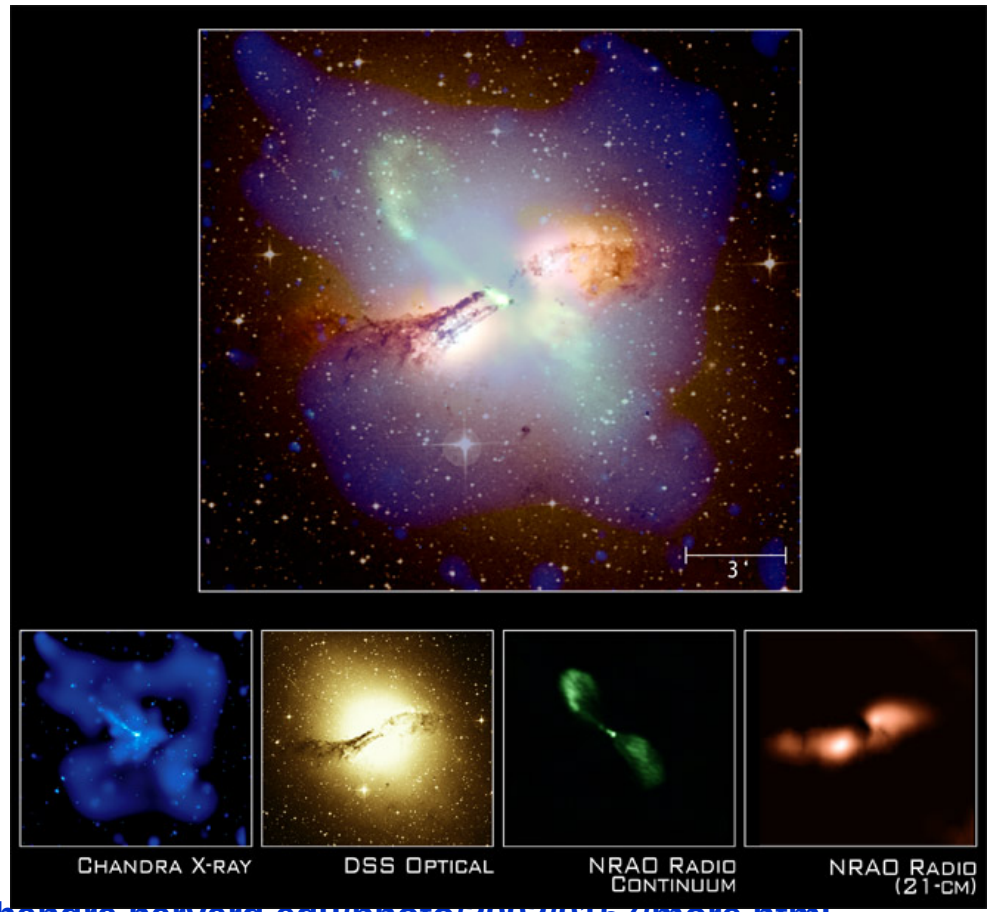
- **Spectral Energy Distribution (SED)**



3C 279 (Hartman et al., 2001, ApJ, 553, 683)

Blazars (cont.)

- **Radio morphology and its evolution implies a relativistic outflow (jet):**

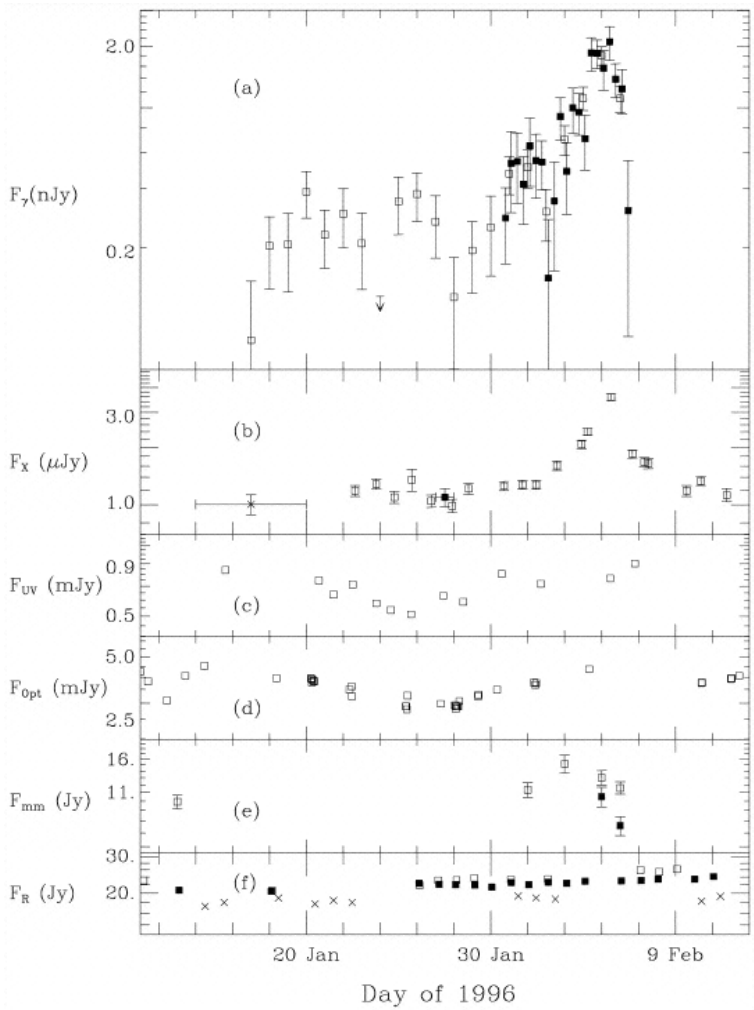


- http://chandra.harvard.edu/photo/2002/0157/more.html

Multi-wavelength Observations

- Light curves and rapid variability across wavebands, e.g., 3C279 in 1996 (Wehrle et al 1998) require coordinated monitoring efforts with other missions and ground-based teams
- After 1st year, all data becomes public immediately

∴ It must be straight-forward to analyze LAT data by investigators outside of the collaboration.

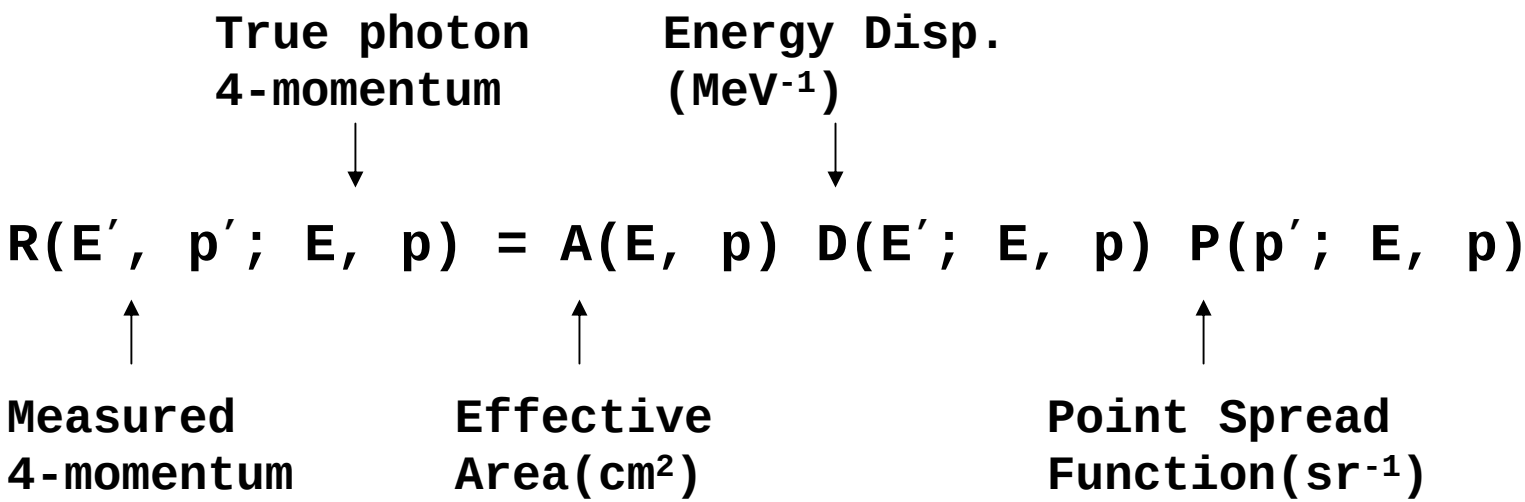


Data Analysis for High Energy Astronomy

- **Framework driven by**
  - **Desire for uniformity between missions**
  - **Guest observer support (HEASARC)**
  - **Aggregate nature of the data: events are (almost) never analyzed individually**
- **Standardized software and data formats**
  - **FITS files for images and tabular data**
  - **FTOOLs for examining and manipulating contents**
    - **Can be mission-specific**
    - **User interfaces – "parameter interface layer", ballistic operation**
    - **Often scripted (Tcl, Perl), some GUI use**
  - **High level analysis applications: Xspec (from Xanadu suite), Sherpa, ISIS (from CIAO), etc.**
  - **Unix-based tradition; GLAST pushing for Windows support**

Instrument Response Functions (IRFs)

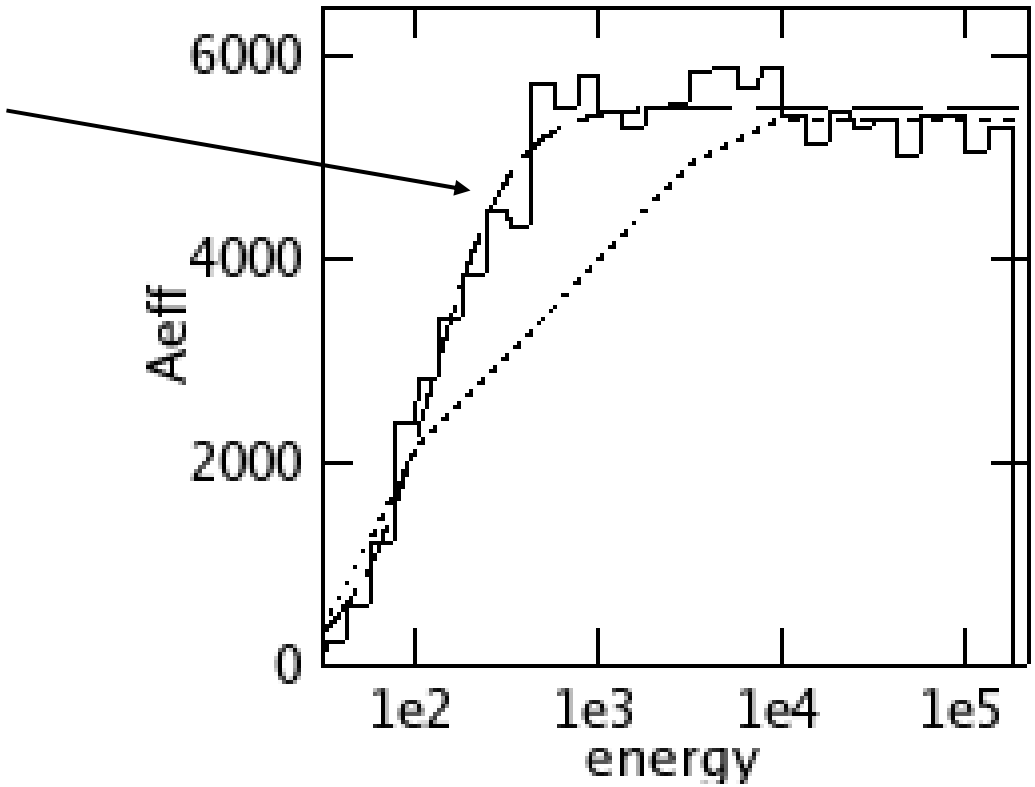
- The *linchpin* between the event reconstruction and Science Tools
- The IRFs are a statistical description of the performance of the LAT for measuring photon properties, e.g., a transition matrix.
- They are derived from real calibration runs using a photon source (e.g., real data + recon) and/or from Monte Carlo simulations using GlastRelease (GEANT 4 + recon).
- The total response, R, is usually factored into three components:



Example IRF Generation: DC1

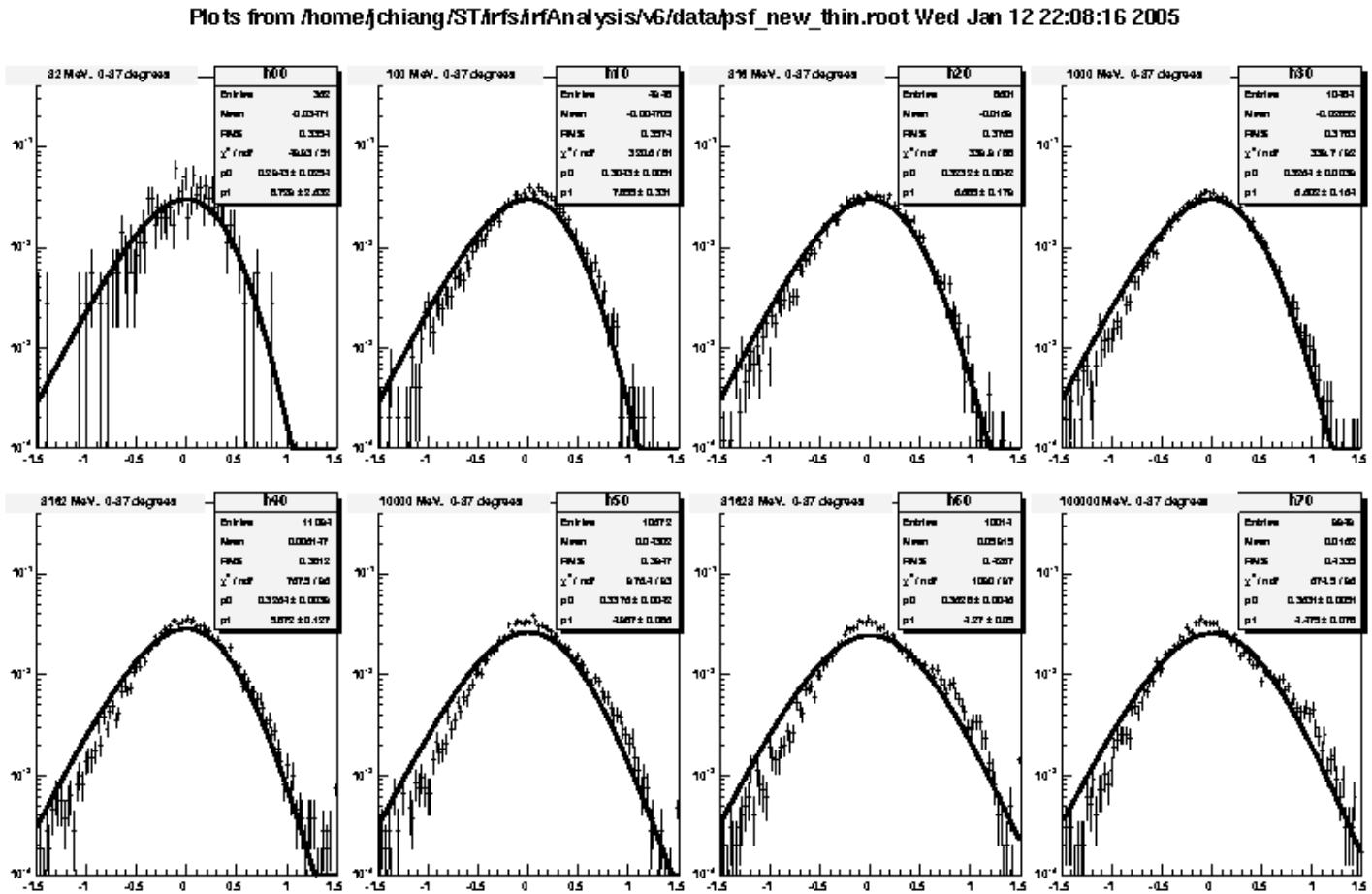
- ~5 M "AllGamma" events are generated covering 2π sr and spanning energies 20 MeV to 200 GeV.
- Effective Area -- detector "cross-section" as a function of energy:

LAT performance is strongly impacted by cuts on particle background

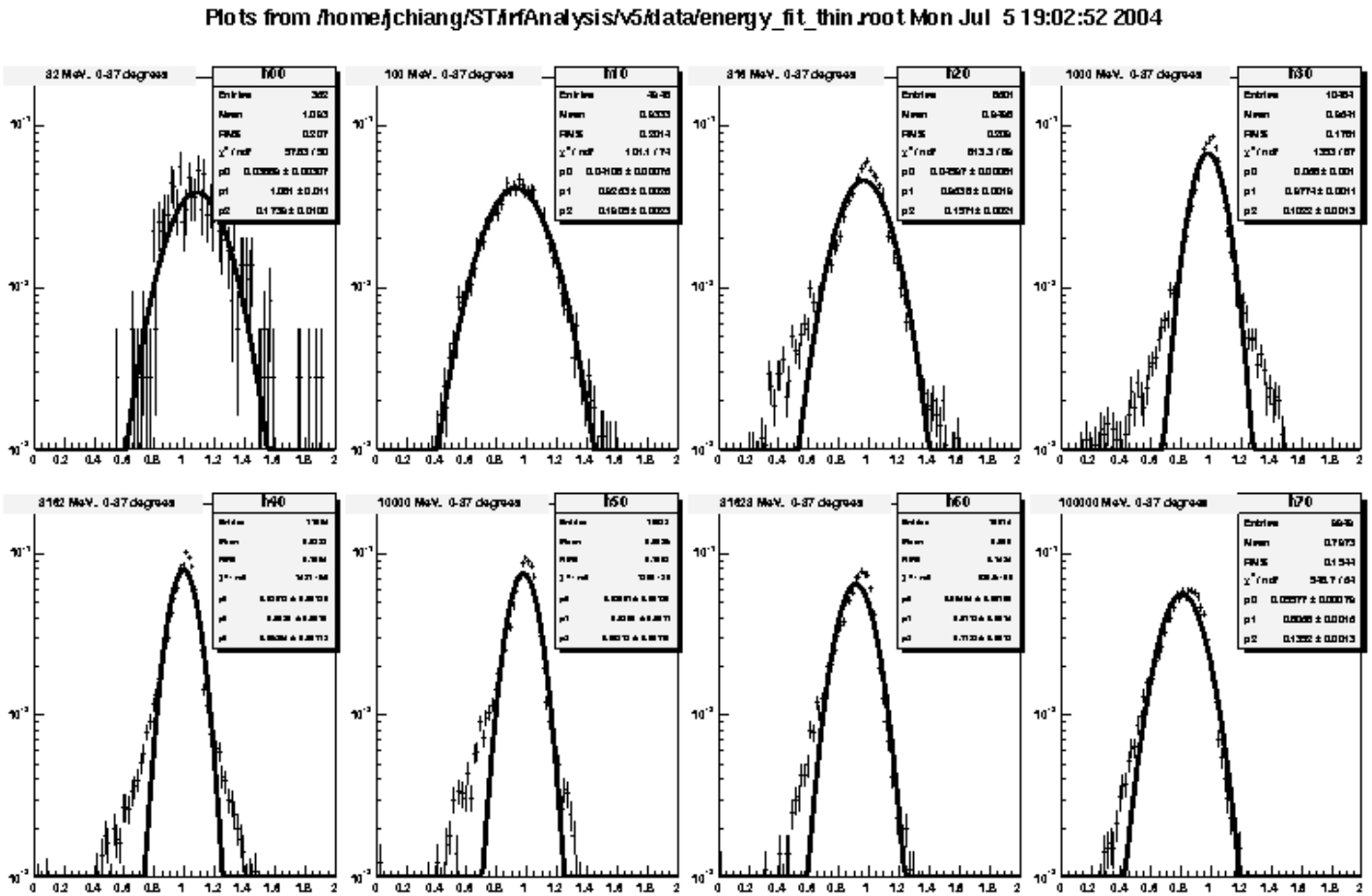


Point Spread Functions (PSFs)

- Data are partitioned into broad energy and inclination bins
- Angular deviations are scaled by ~E^{-1} to account for multiple scattering



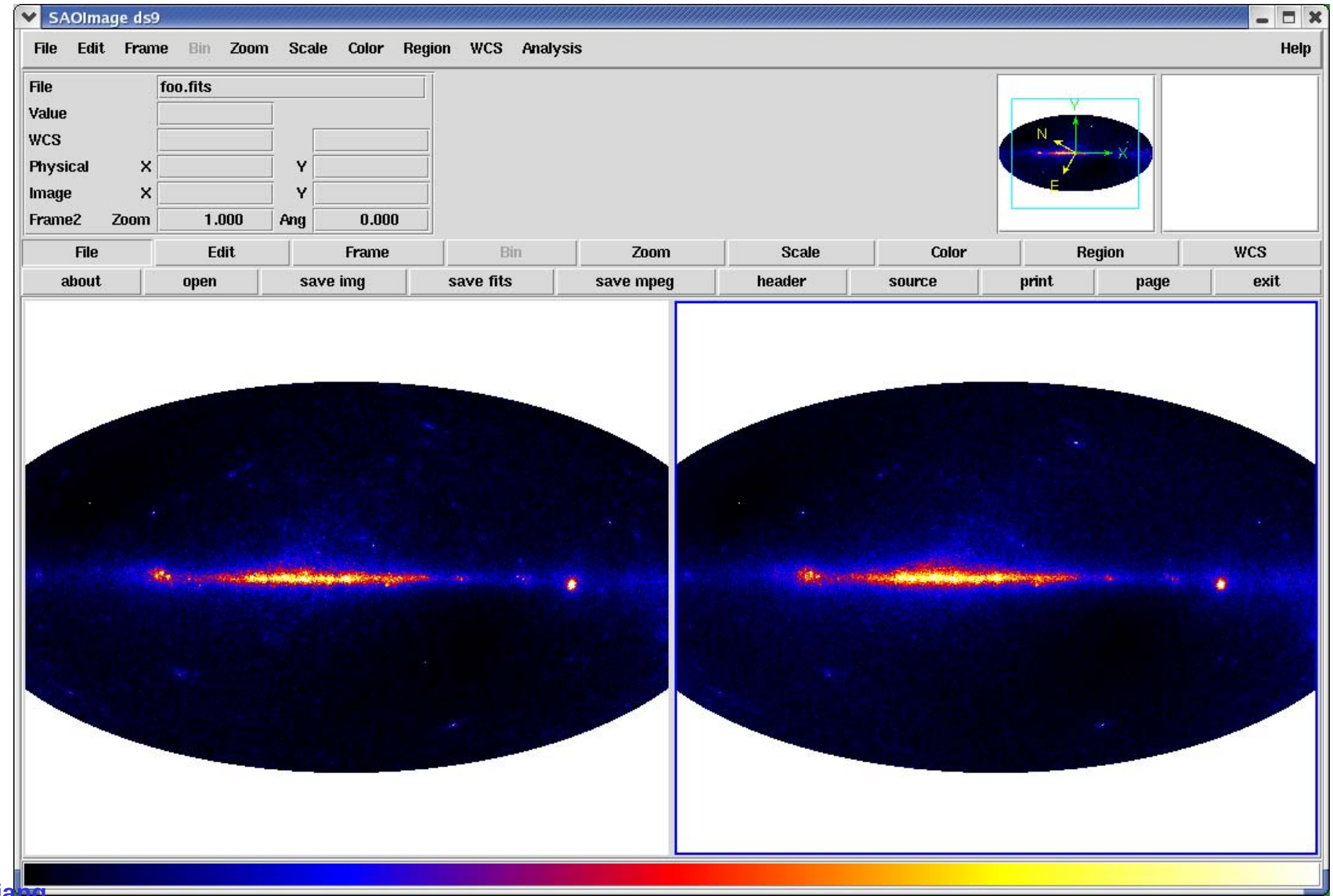
Energy Dispersion



Using the IRFs for Simulation

- **Assuming perfect knowledge of incident photons, the distributions of measured quantities should (ideally) be identical for all three ways of obtaining them:**
  - **Real observations + Gleam reconstruction**
  - **Gleam simulation + recon**
  - **IRFs**
- **For Science Tools development and testing, we have developed a high level observation simulator that reads in the same sky model as Gleam, but uses the IRFs to produce simulated events:**
  - **Source flux (photons $cm^{-2}s^{-1}$) x A (cm^2) = rate of detected events**
  - **True photon 4-momentum & P & D $\Rightarrow$ Apparent photon 4-momentum, i.e., smeared by instrument response**

IRF Simulation vs Gleam

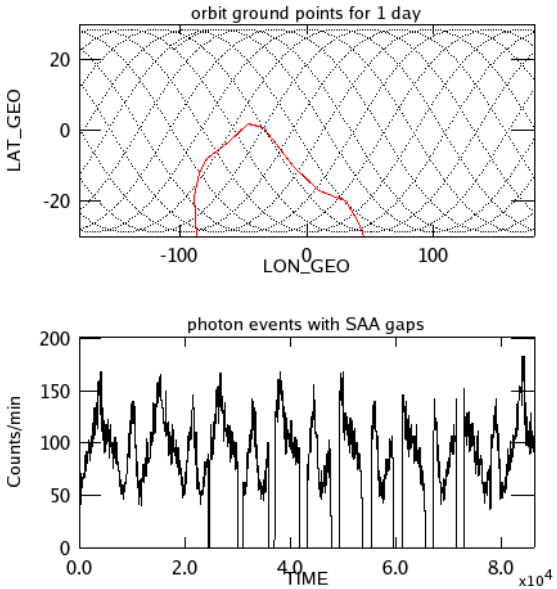


Typical Steps in a LAT Analysis

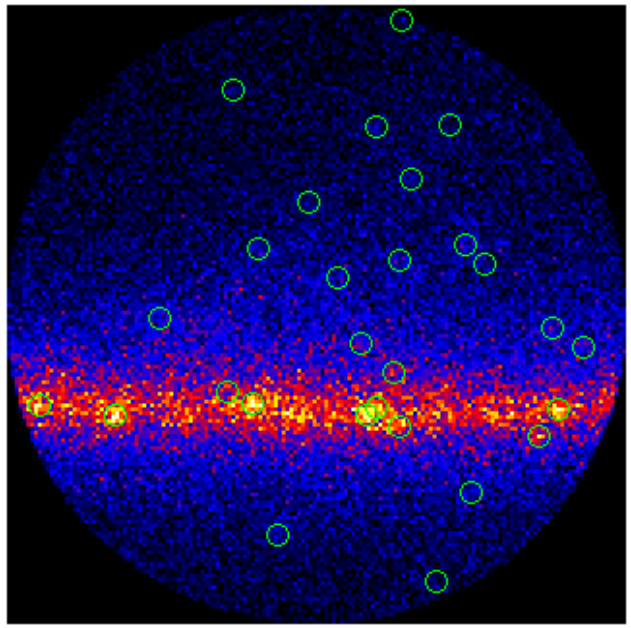
- Acquire data
  - download from GSSC server
- Preliminary visualization
  - counts and exposure maps
- Analysis-specific data selections
  - GTIs, ROI, event type
- Source identification
  - Source detection and identification: image processing techniques, wavelet analyses, etc.. (should be fast).
- Source characterization
  - Maximum Likelihood estimate (MLE) of source properties—flux, spectrum, position (computationally expensive).
  - Multi-wavelength spectral fitting (using Xspec).

Data Selection

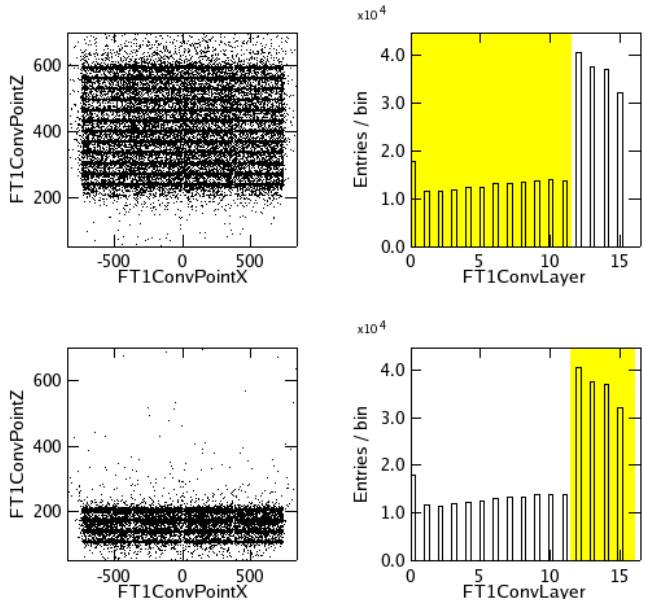
South Atlantic Anomaly (SAA) passages handled by "good time intervals" (GTIs)…



event data are partitioned into "regions-of-interest" (ROIs)…

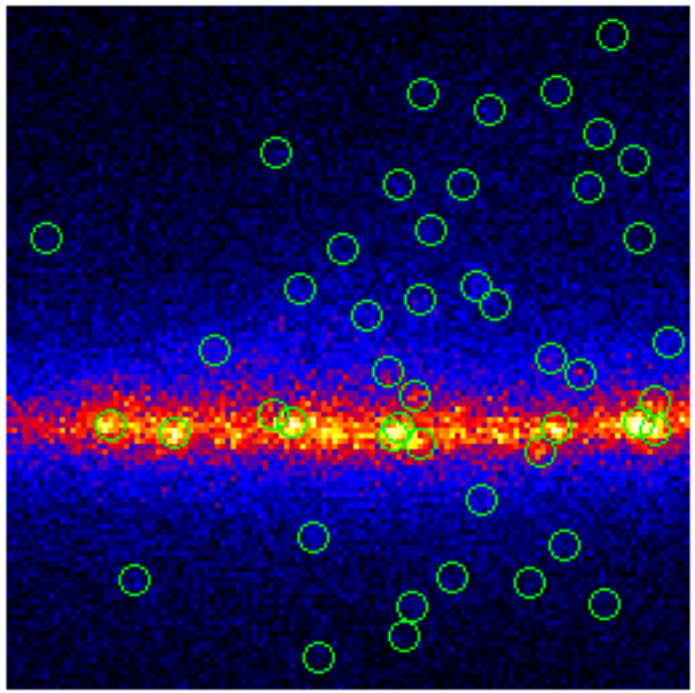


and by event type, eg. "front" vs "back" (depends on IRF granularity)

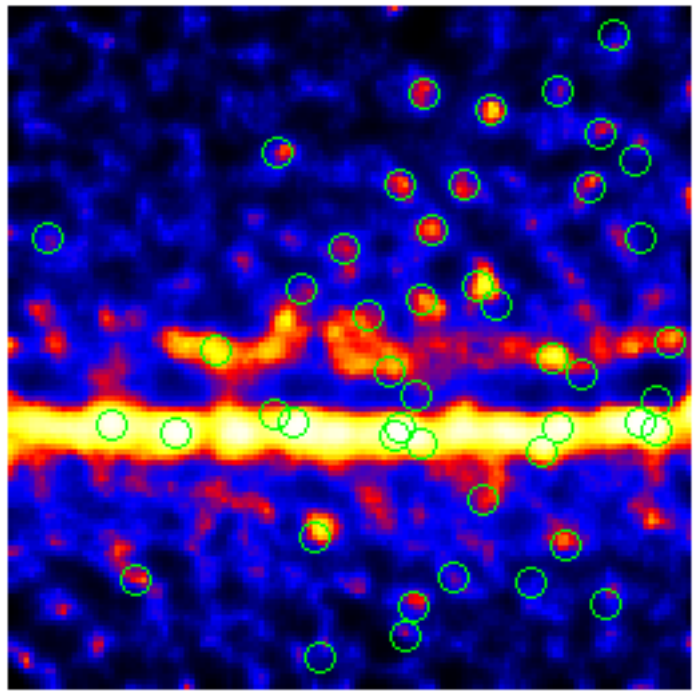


Source Identification

- De-noising and deconvolution (wavelets, etc.)
- Source finder (preferably automated)



Input counts map
1 week simulation time



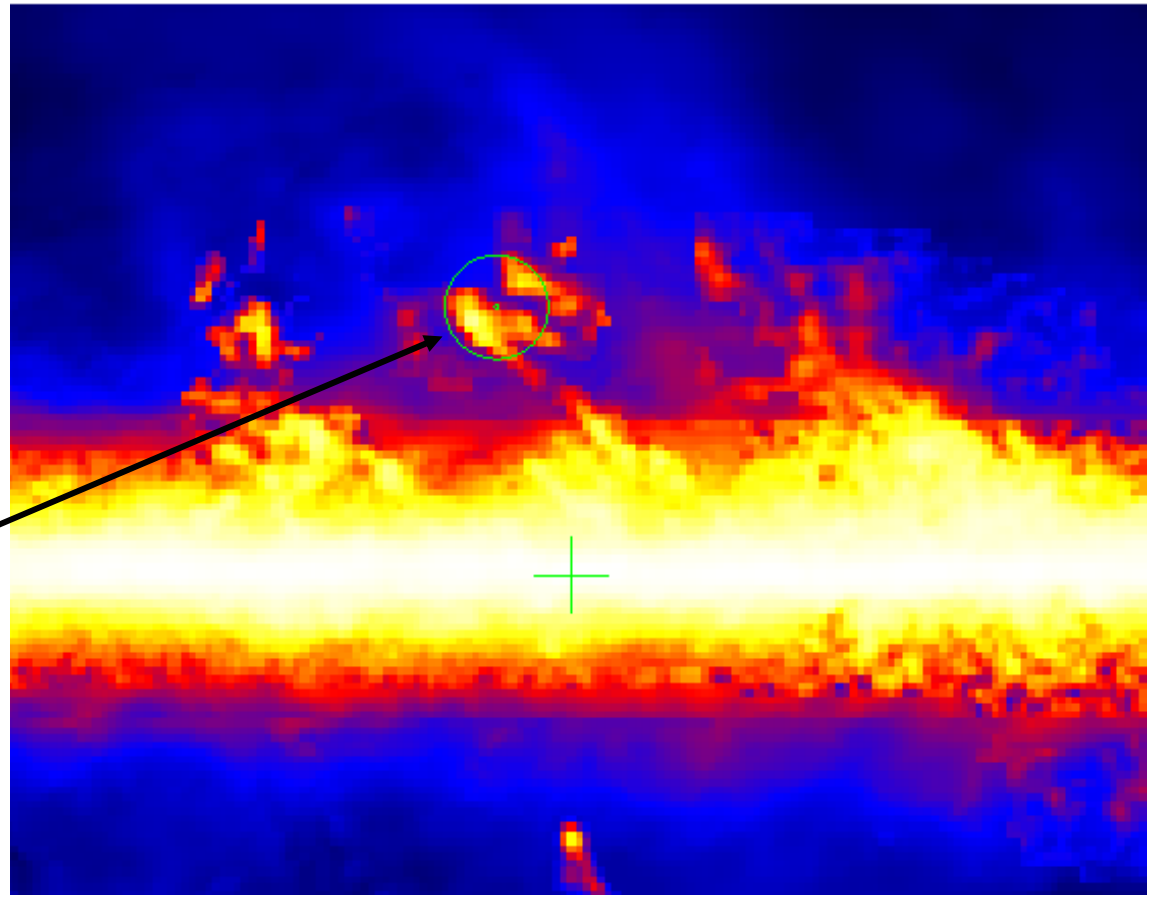
Deconvolved map using
EM algorithm ⇒ MLE

Galactic Diffuse and Source Confusion

Emission results from cosmic ray interactions with interstellar gas.

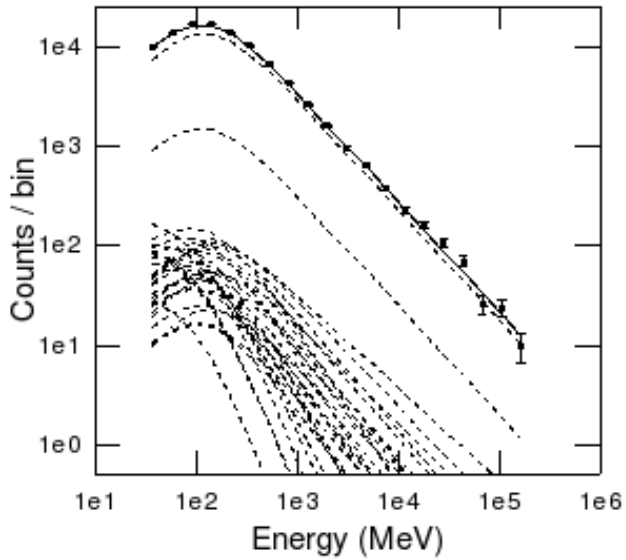
Models rely on HI & CO observations for the gas distribution.

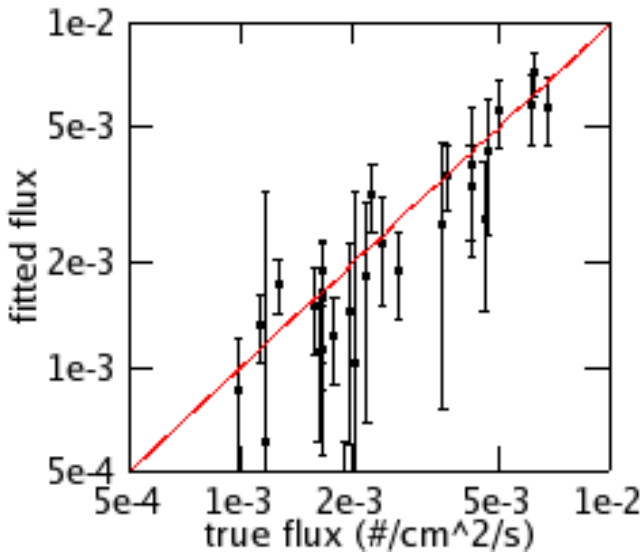
These observations reveal structures on angular scales similar to the PSF:
~3.5^{o} @ 100 MeV
~0.1^{o} @ 10 GeV

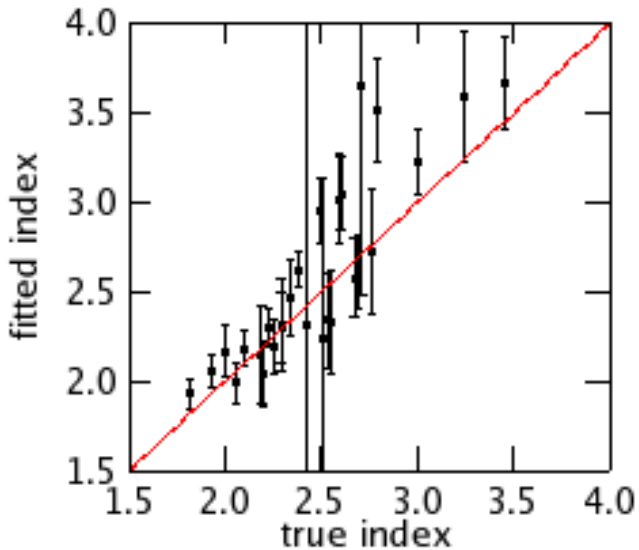


Source Characterization

- Maximum likelihood for ascertaining source parameters
  - flux, spectral index, source position
  - $>$ 50 parameter fits for a single ROI are common



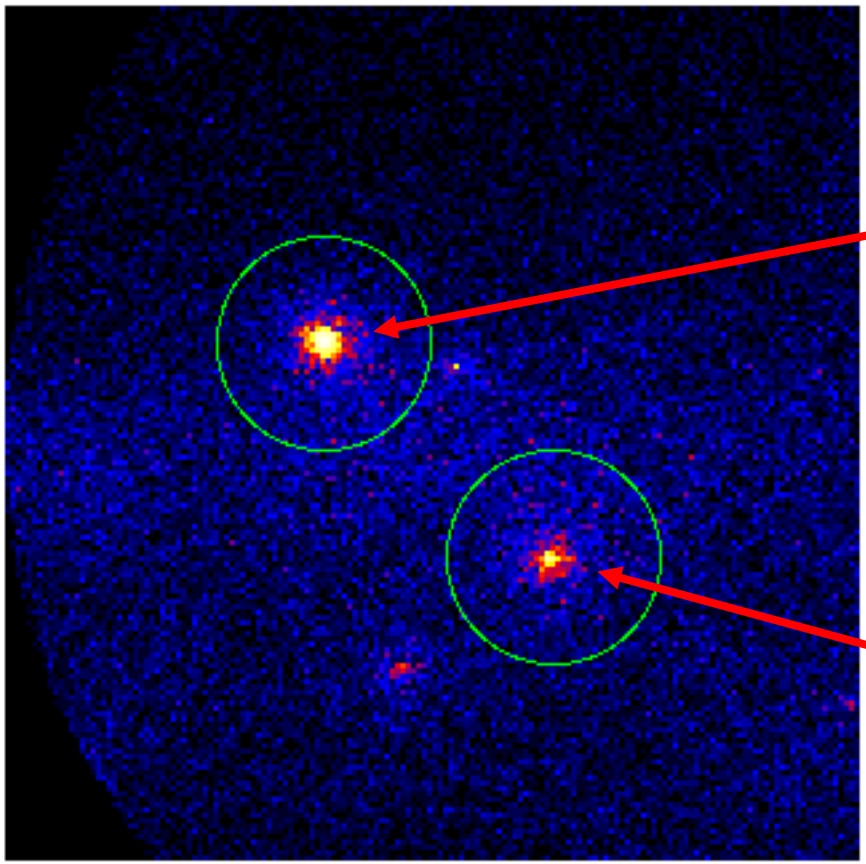


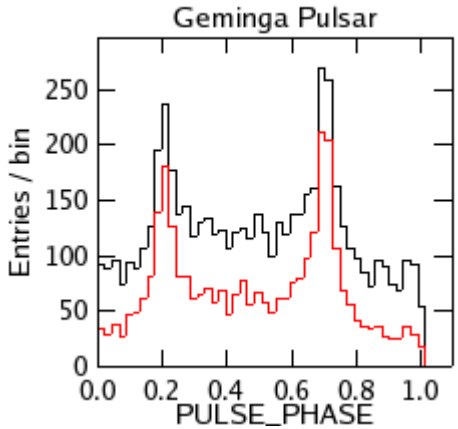


Types of Gamma-Ray Sources

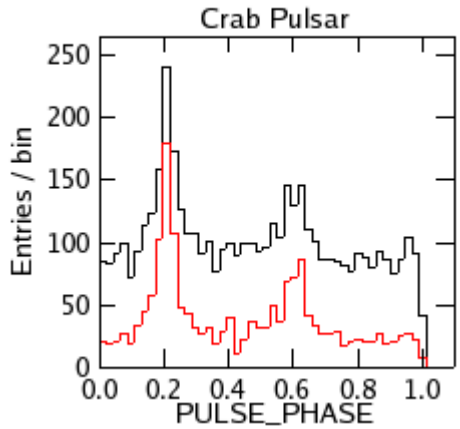
- **Pulsars**
  - Rapidly rotating neutron stars ($P \sim 10^{-3}$-10s, $(dP/dt)_{Crab} \sim 10^{-13}$ s/s) with $B \sim 10^{9}$-10^{12} G
  - Ephemerides from radio observations
- **Blazars**
  - Variability over a wide range of time scales (hours to months)
  - Multi-wavelength monitoring is crucial
- **Gamma-ray Bursts**
  - Very short time scales, < 10s of seconds
- **Diffuse/extended emission**
  - Milkyway galaxy, LMC, supernova remnants
  - Extragalactic diffuse may comprise unresolved discrete sources such as blazars
- **New physics:**
  - Dark Matter sources

Pulsars



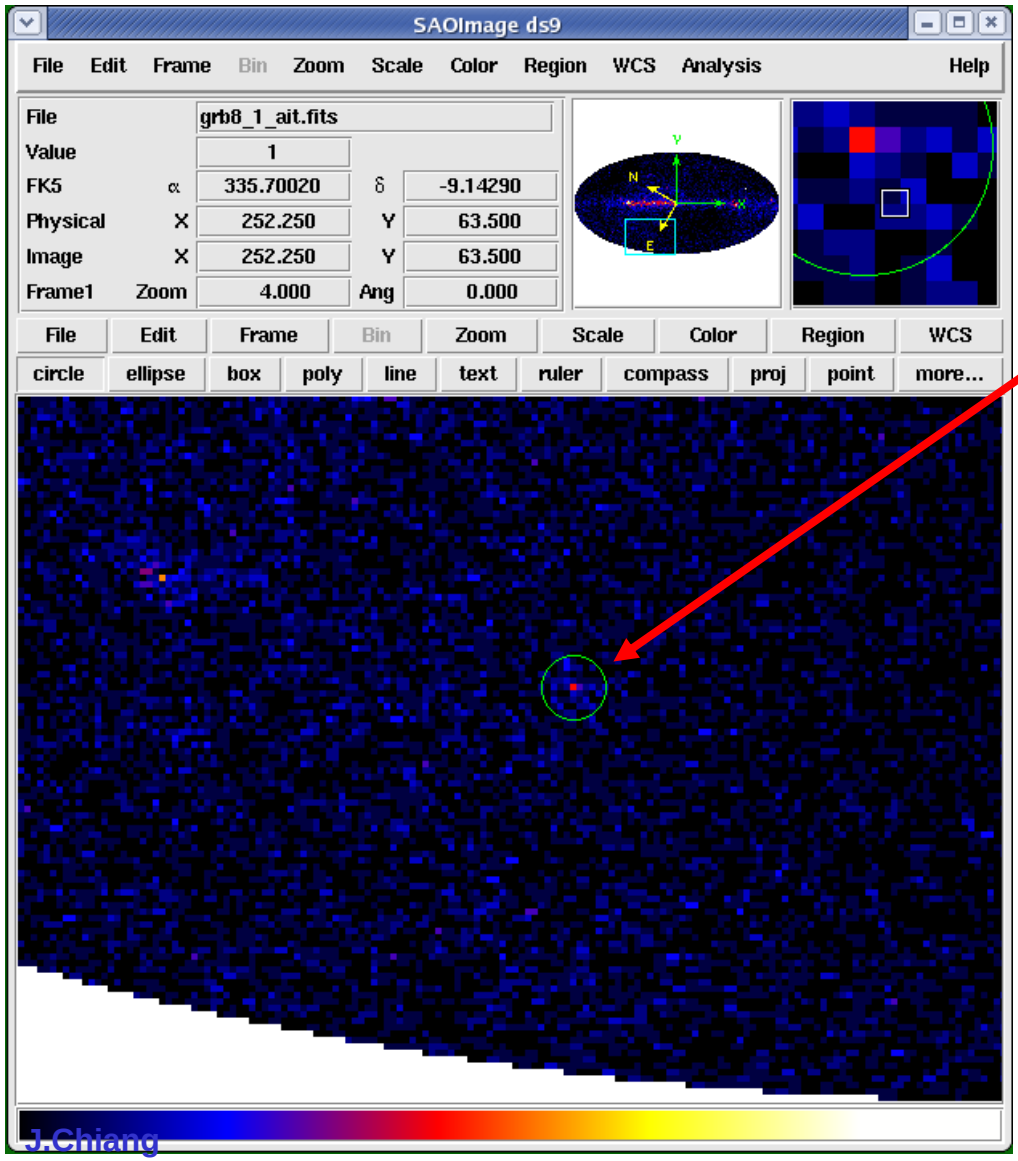


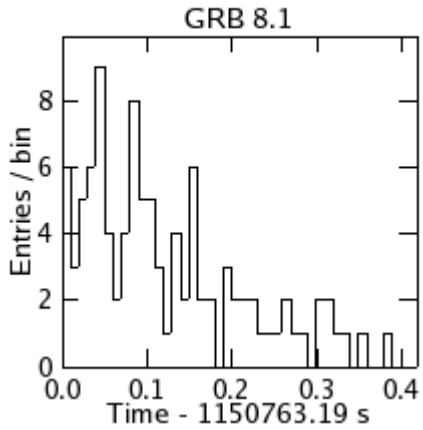
P=237ms

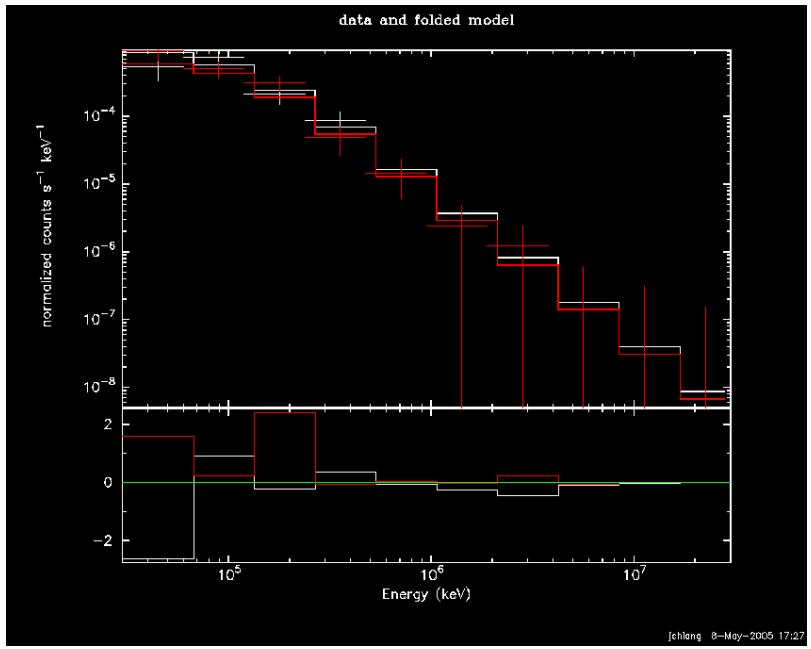


P=33ms

Gamma-Ray Bursts







J.Chiang

For more on ScienceTools…

http://www-glast.slac.stanford.edu/ScienceTools/

http://www.slac.stanford.edu/exp/glast/ground/software/status/documentation/ScienceTools/latest/Likelihood/latest/tutorial.html

http://glast.gsfc.nasa.gov/cgi-bin/ssc/LAT/STCDataQuery.cgi

GLAST Large Area Telescope:

Exploring the γ-ray Sky

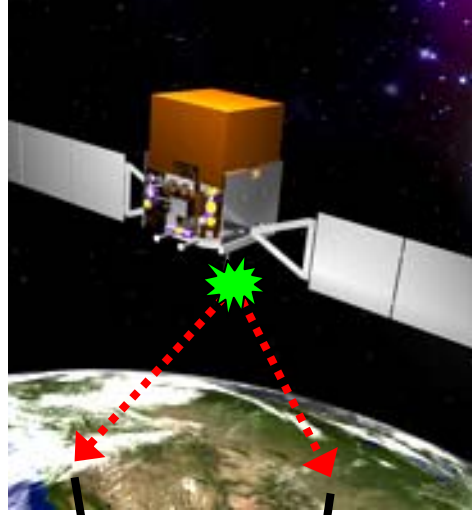
Daniel Flath
Stanford Linear Accelerator Center
dflath@slac.stanford.edu

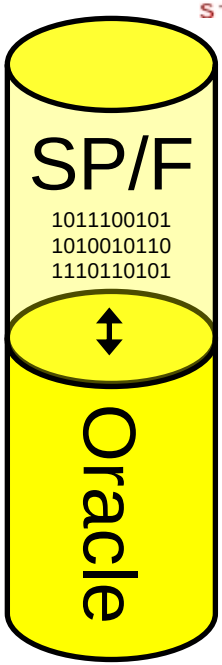
http://www-glast.slac.stanford.edu/software

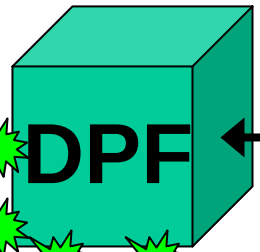
Data Handling: Outline

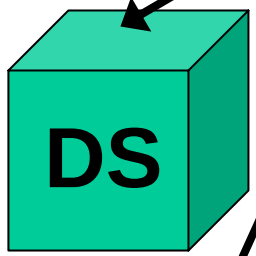
- Automation: Pipeline
  - Guaranteed 1 hour turnaround
  - 24x7x365x10
  - Parallel processing of data – Workflow
  - Web based monitoring
- System Tests
  - Monitoring data and software
  - Web based
- Data Server
  - Public data server, for public
  - Glast data server, for collaboration
    - Why? (pull region of sky from many orbits)
    - Users don't (want to know) much about recon/simulation
- Technologies (to be) used
  - 3rd Generation Web Application Containers
    - Web applications that work like desktop applications
  - Mix of commercial products and Open Source projects
    - JIRA
      - Bug tracker
      - Project management tool
    - Confluence
      - Documentation repository
      - "Super WIKI"
  - Commercial tools
    - Themselves based on Open Source projects
    - Examples of 3rd Generation web applications


STANFORD LINEAR ACCELERATOR CENTER

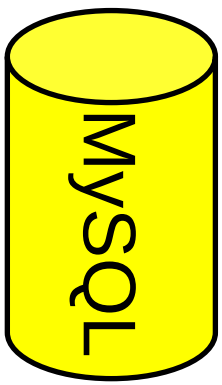


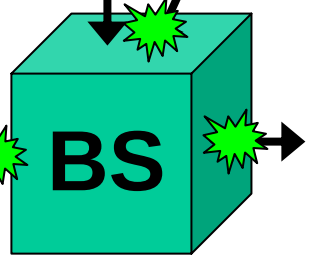

SP/F
1011100101
1010010110
1110110101
Oracle


DPF

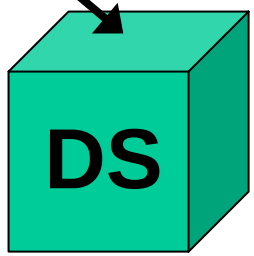

DS

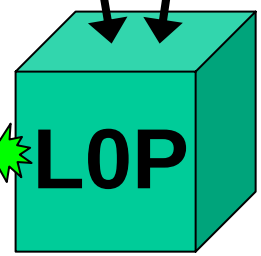

NASA
GSFC


MySQL


BS

LSF


DS

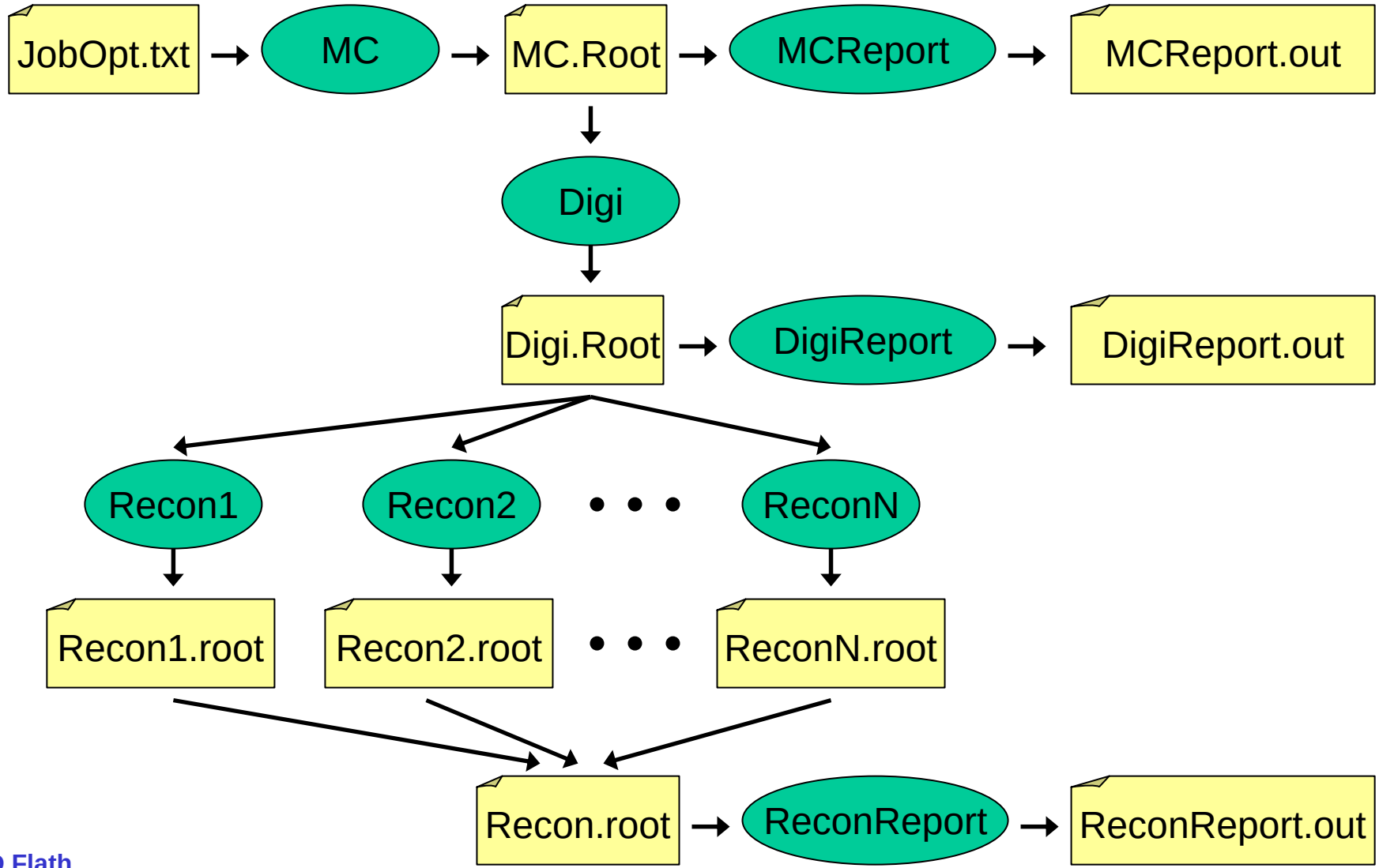

L0P

Pipeline Intro

- **What is the pipeline?**
  - Envisaged as tool to provide a tree of processing on a given input dataset
  - Full bookkeeping to track what happened
  - Archive all files touched

- **Used by whom?**
  - Online
    - for sweeping integration data out of the clean room and to tape
    - populate eLogbook
  - SVAC (Science Verification and Calibrations)
    - for doing digi, recon
    - creating reports
    - Preparing for calibrations
  - Generic MC
    - DC2, background runs etc etc
  - ISOC (Instrument Science Operations Center)
    - Flight operations
    - What about environmental testing, at Spectrum Astro, KSC?

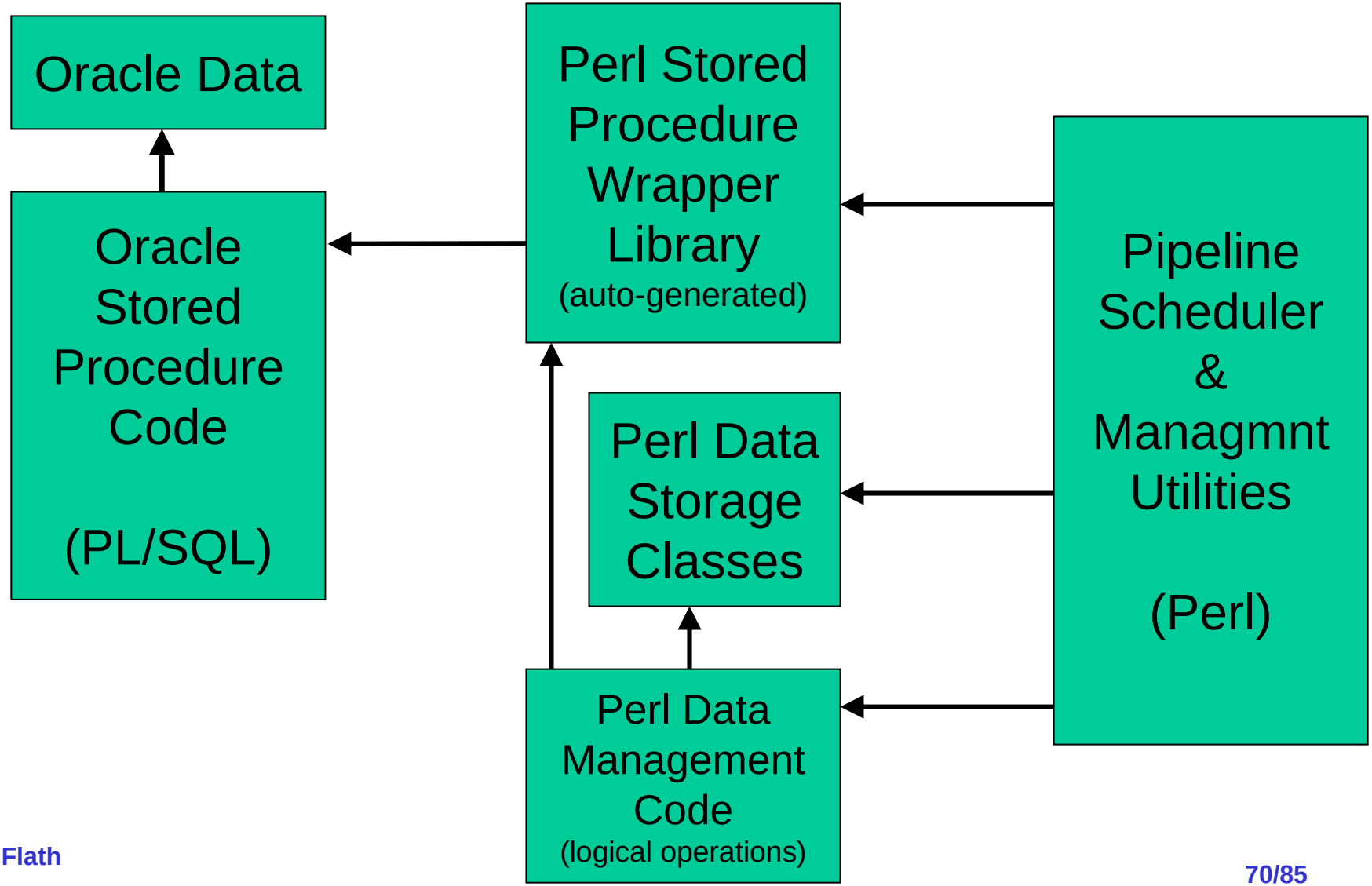
Sample Processing Chain



Current Pipeline: Major Components & Tech Used

- **RDBMS** (relational database management system)
  - Oracle
  - Contains all processing and data product history and relationships
- Data Exchange Layer
  - Oracle PL/SQL
  - Compiled SQL queries provide read/write access to tables
- DB Access Layer
  - Perl::DBI
  - Auto-Generated subroutines wrapping every public stored function and procedure
  - Provides simple, seamless DB interface to Perl Utilities
  - Also Perl classes representing each record type
- Scheduler, utilities
  - Perl
  - Higher level code to manage data and processing
  - Little dependency on actual table structure gives developer freedom to write maintainable, extensible code

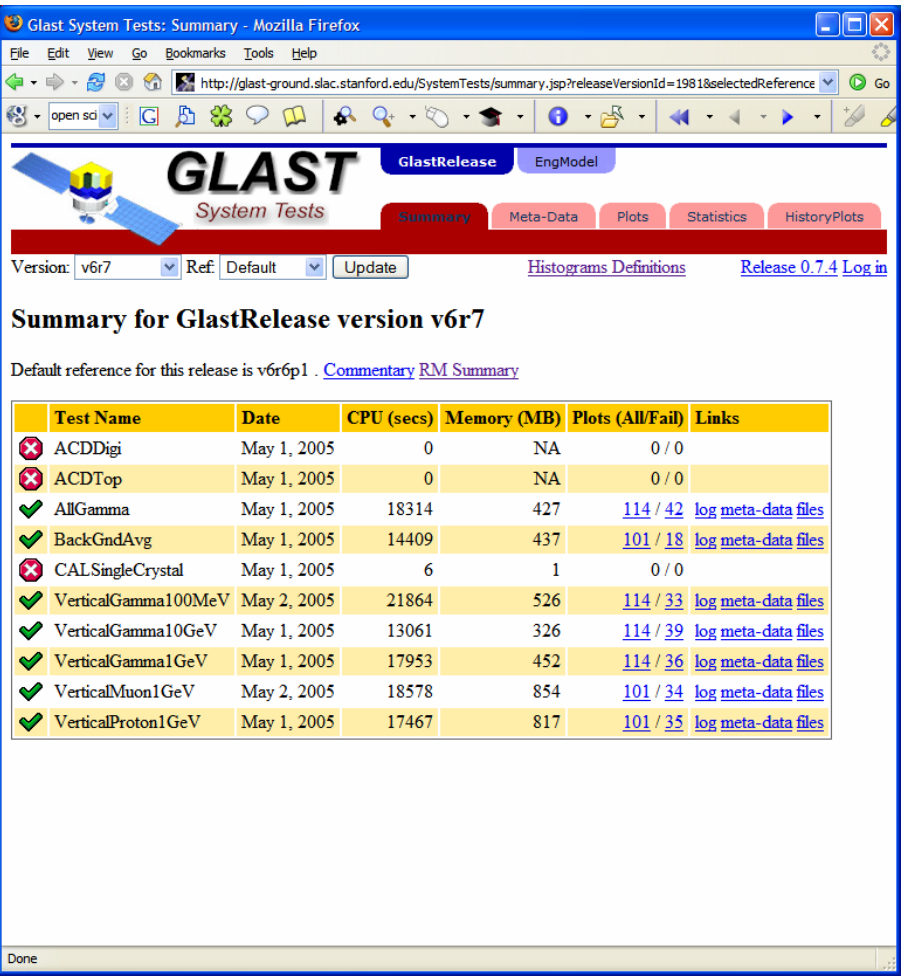
Pipeline Component Dependencies

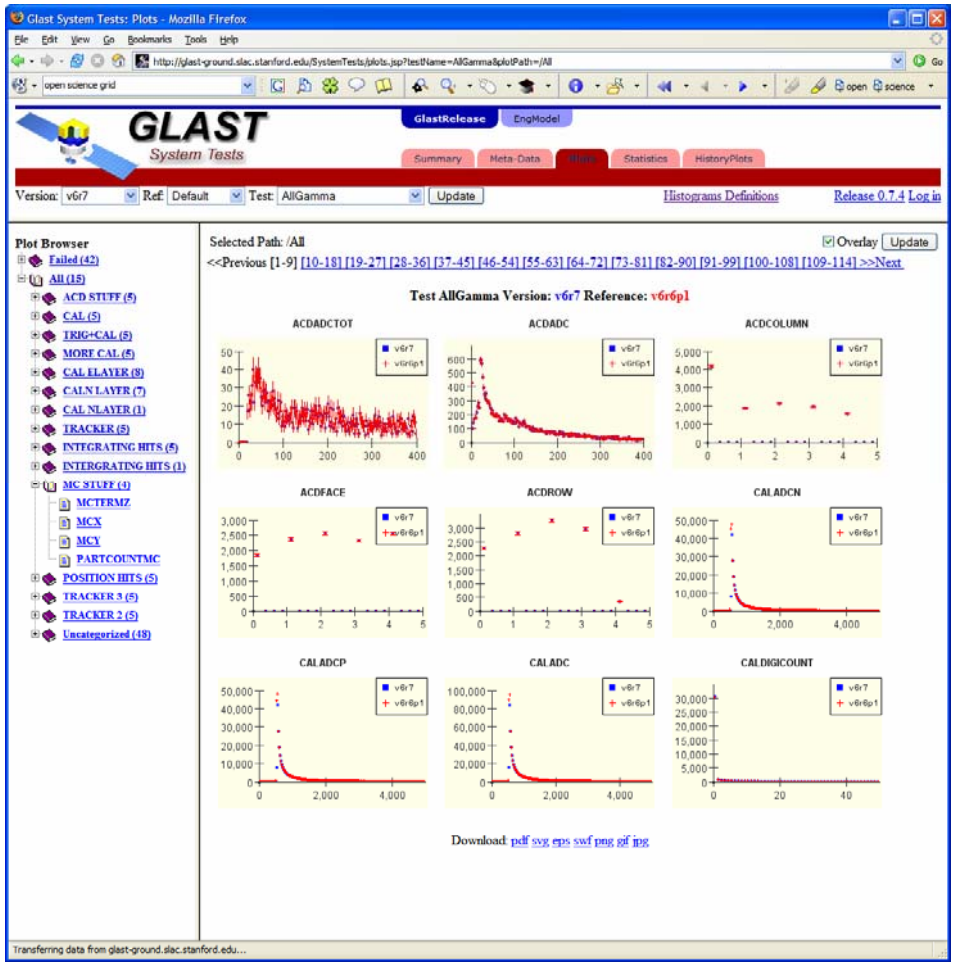


System Tests

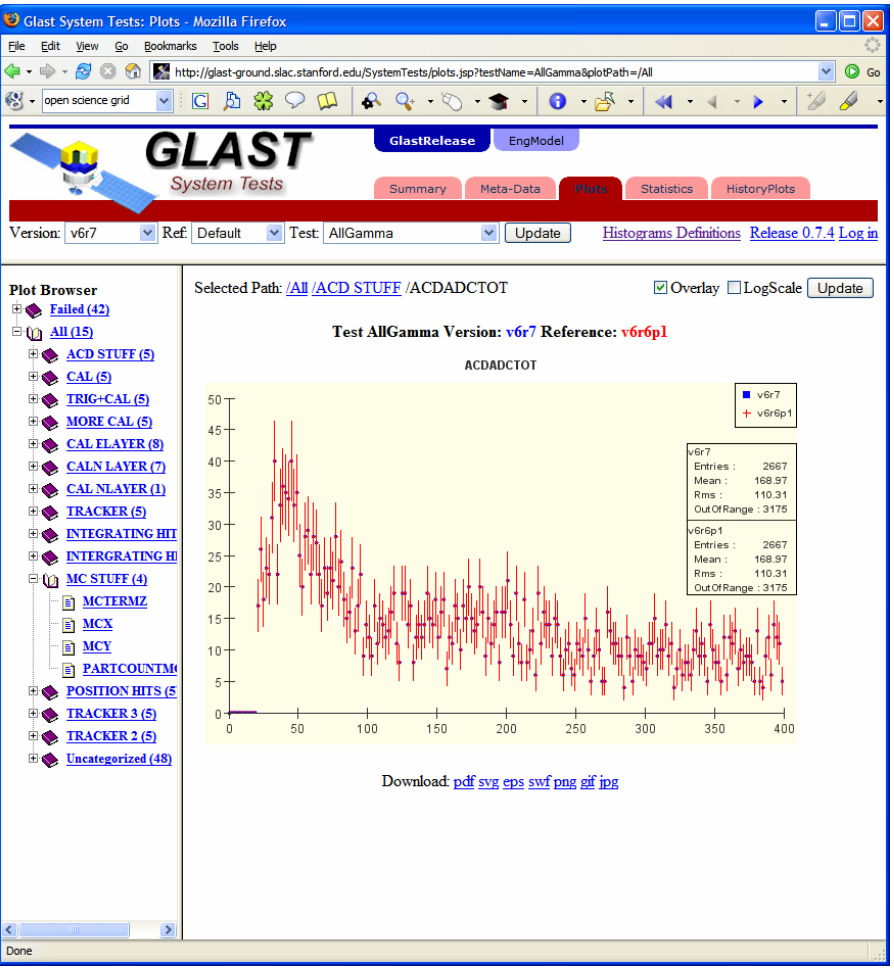
- Goals
  - Provides mechanism for validating:
    - Software releases (now)
    - Data quality (after launch)
  - Run (automatically) after each software release
    - Compares plots to references and flags problems
- Web based access to system tests results from any platform
  - No software install needed
  - Accesses data from combination of
    - Oracle database tables
    - Root files
  - Implemented using JAIDA, xrootd, JSP, Tomcat

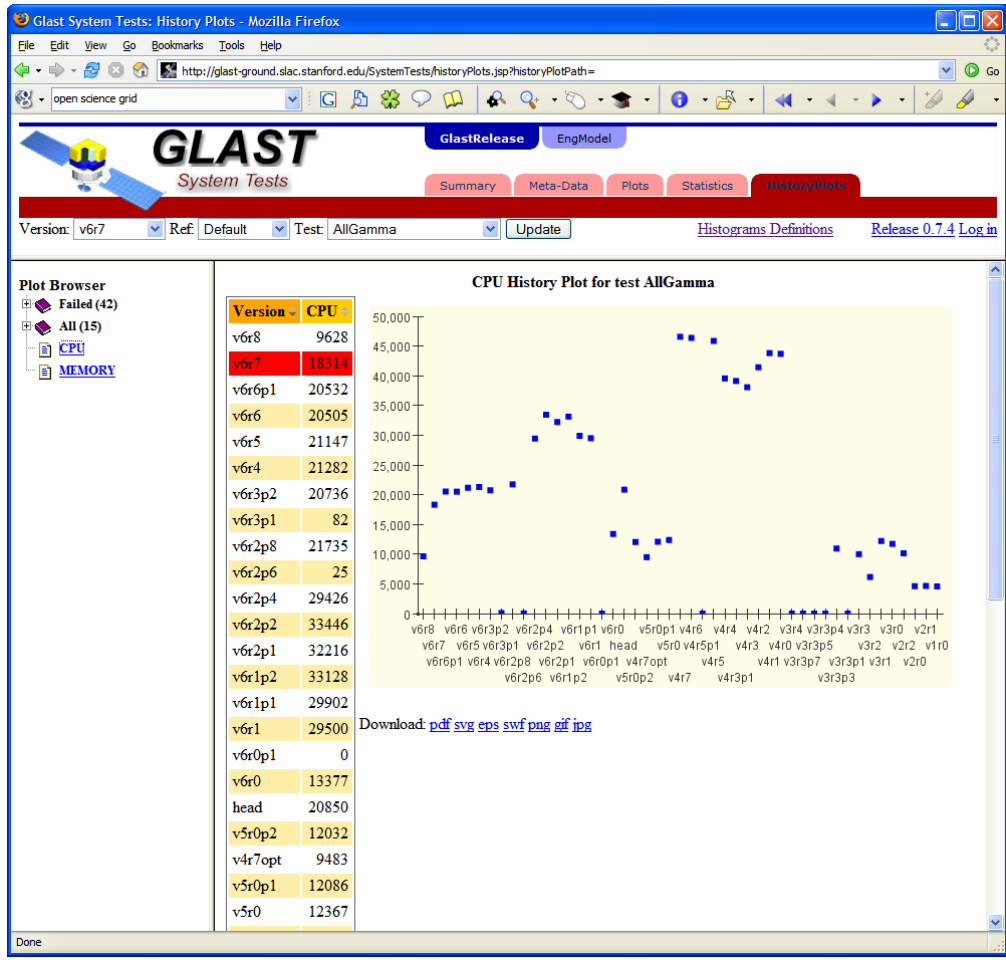
System Tests





System Tests





Data Server

- Glast will run two data servers
  - One for the public at Goddard Space Flight Center
  - One at SLAC for Glast collaborators
- Glast Physicists will access data via Data Server
  - Pulls events associated with
    - Particular region of the sky
      - Satellite doesn't stay still so this is spread throughout data.
    - Energy range
    - Time Period
  - Removes need for users to know how/where data is stored
    - For most astrophysics measurements physicists only need to know about photon direction and efficiency, details of reconstruction/simulation are largely irrelevant
  - Should be able to download data in various formats
    - List of run/events
    - Tuples (FITS, root, possibly with choice of number of columns)
    - Full root trees
  - Should be able to browse events
    - with web based event display (WIRED)
  - Should be able to store personal favorite searches
    - Should be able to download incremental updates to data
- Expect to get 100M events/year for 10 years
  - Small compared to Babar, but we want fast turnaround

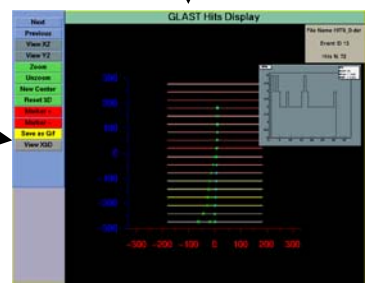
Data Server

Web Form
Region in Sky:
Time Range:
Energy Range:
Gammas/Events:

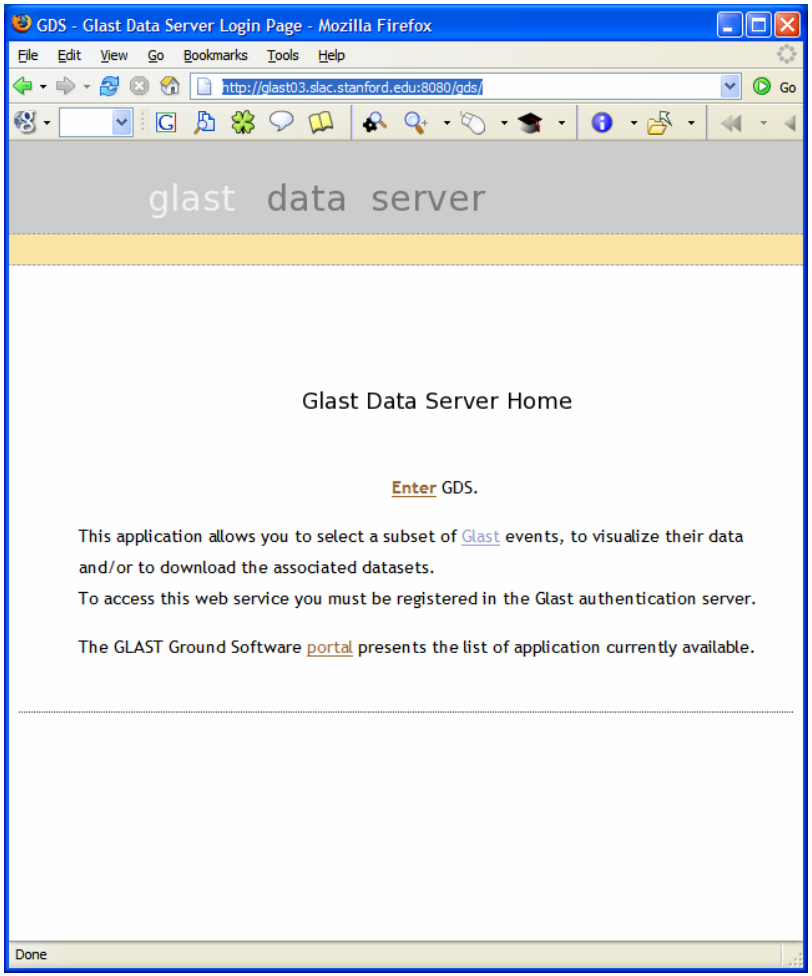
You selected 39383844 events
Change Criteria
Add “TCut”
Browse Events
Download: Event Selection
Compressed Tuple
Full Merit Tuple Full Root Tree

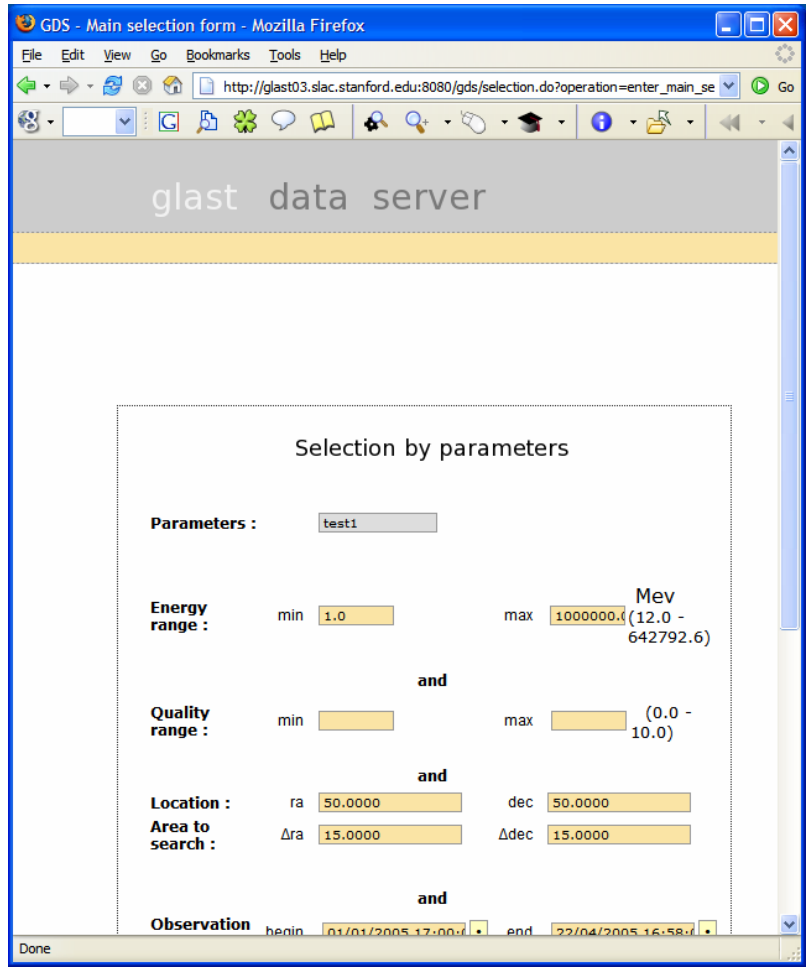
In memory meta-data
Binned by sky position, time, energy

Root Event Store

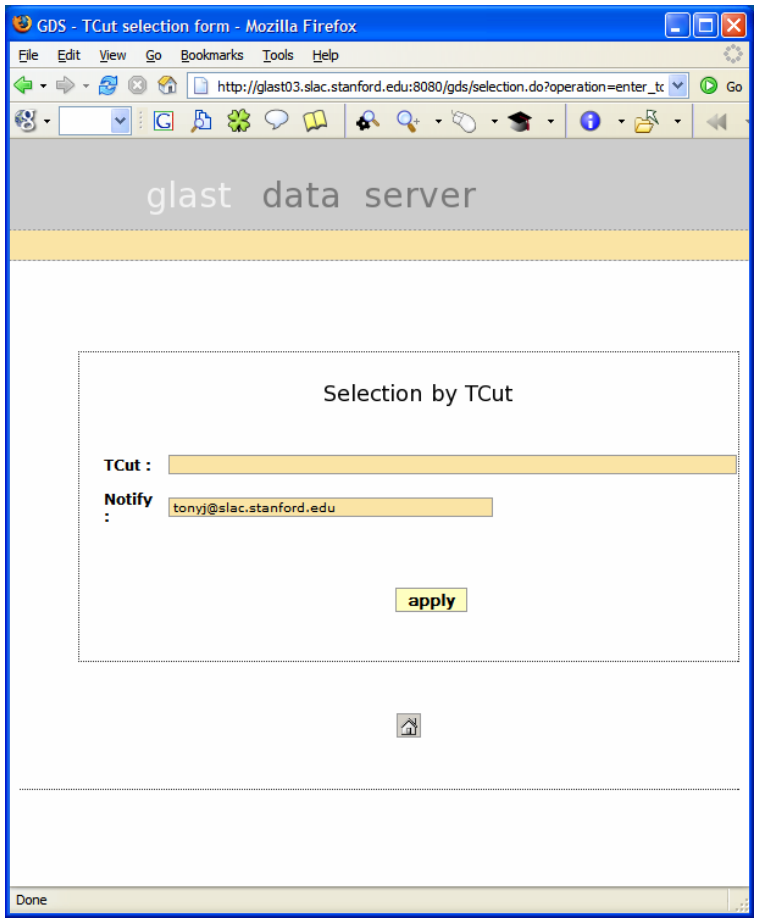


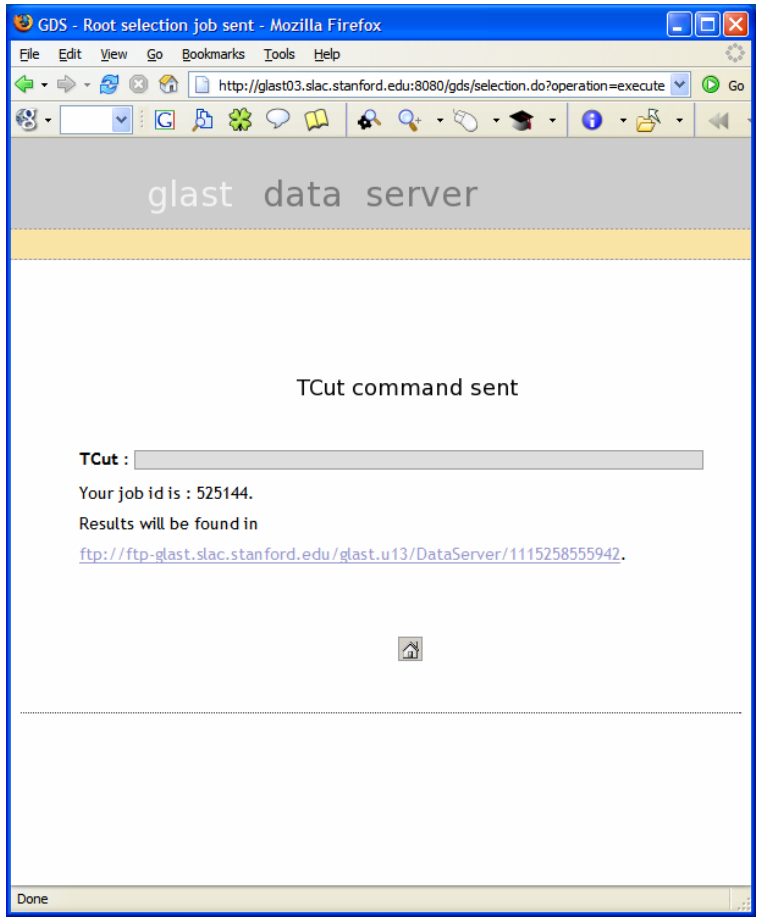
Data Server


GDS - Glast Data Server Login Page - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://glast03.slac.stanford.edu:8080/gds/
Go
glast data server
Glast Data Server Home
Enter GDS.
This application allows you to select a subset of Glast events, to visualize their data and/or to download the associated datasets.
To access this web service you must be registered in the Glast authentication server.
The GLAST Ground Software portal presents the list of application currently available.
Done


GDS - Main selection form - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://glast03.slac.stanford.edu:8080/gds/selection.do?operation=enter_main_se
Go
glast data server
Selection by parameters
Parameters :
test1
Energy range :
min
1.0
max
1000000.(
Mev
(12.0 - 642792.6)
and
Quality range :
min
max
(0.0 - 10.0)
and
Location :
ra
50.0000
dec
50.0000
Area to search :
Δra
15.0000
Δdec
15.0000
and
Observation
begin
01/01/2005 17:00:(
end
22/04/2005 16:58:(
Done

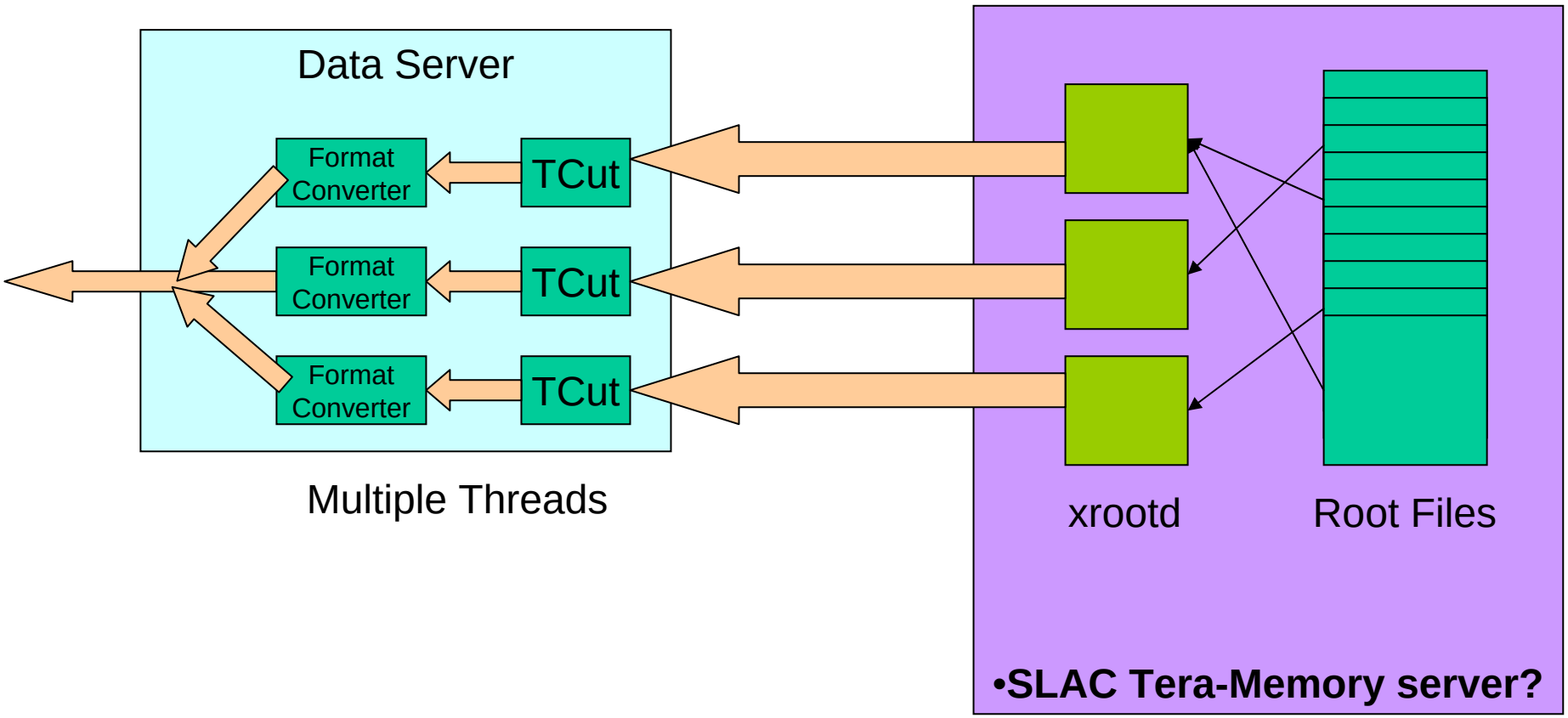
Data Server


GDS - TCut selection form - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://glast03.slac.stanford.edu:8080/gds/selection.do?operation=enter_tc
Go
glast data server
Selection by TCut
TCut :
Notify :
tonyj@slac.stanford.edu
apply
Done


GDS - Root selection job sent - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://glast03.slac.stanford.edu:8080/gds/selection.do?operation=execute
Go
glast data server
TCut command sent
TCut :
Your job id is : 525144.
Results will be found in
ftp://ftp-glast.slac.stanford.edu/glast.u13/DataServer/1115258555942.
Done

Future Data Server

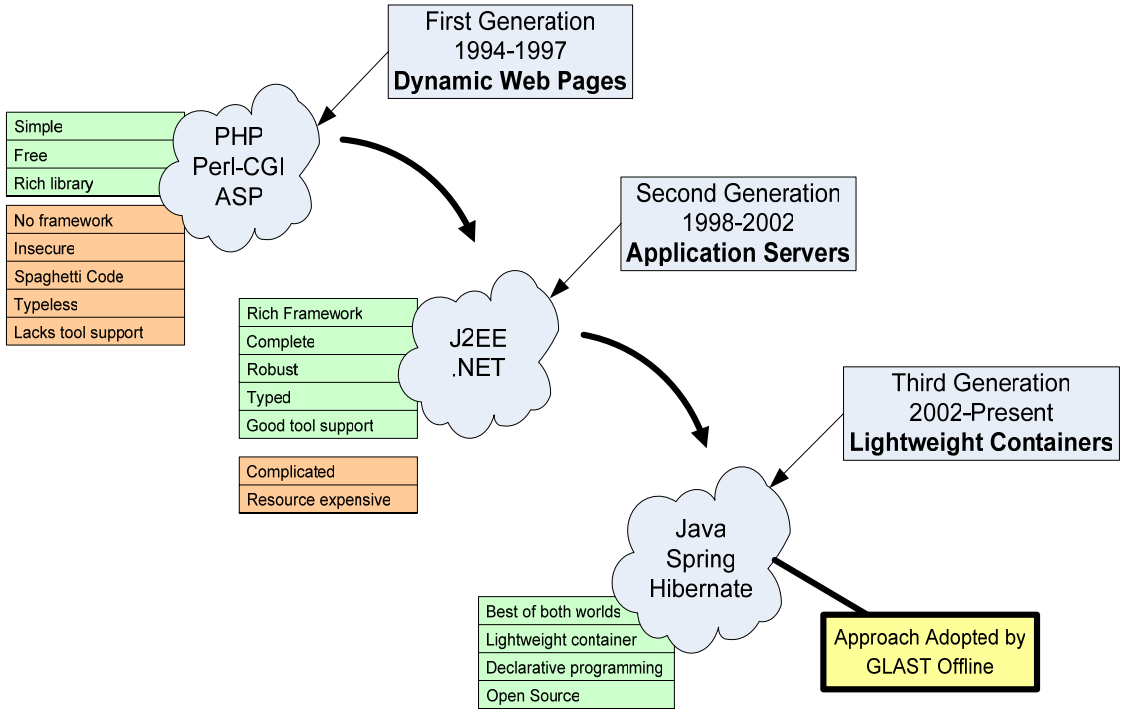
- **Future plans for Data Server:**
  - **Instead of delivering data via FTP, use real-time streaming**



Back to the Future: Adopt 3rd Generation Web Applications

- **Program as little as possible, declare as much as possible**
  - **State your requirements and intentions in configuration files**
  - **Write code only for your specific problem domain, leave the rest to the** container

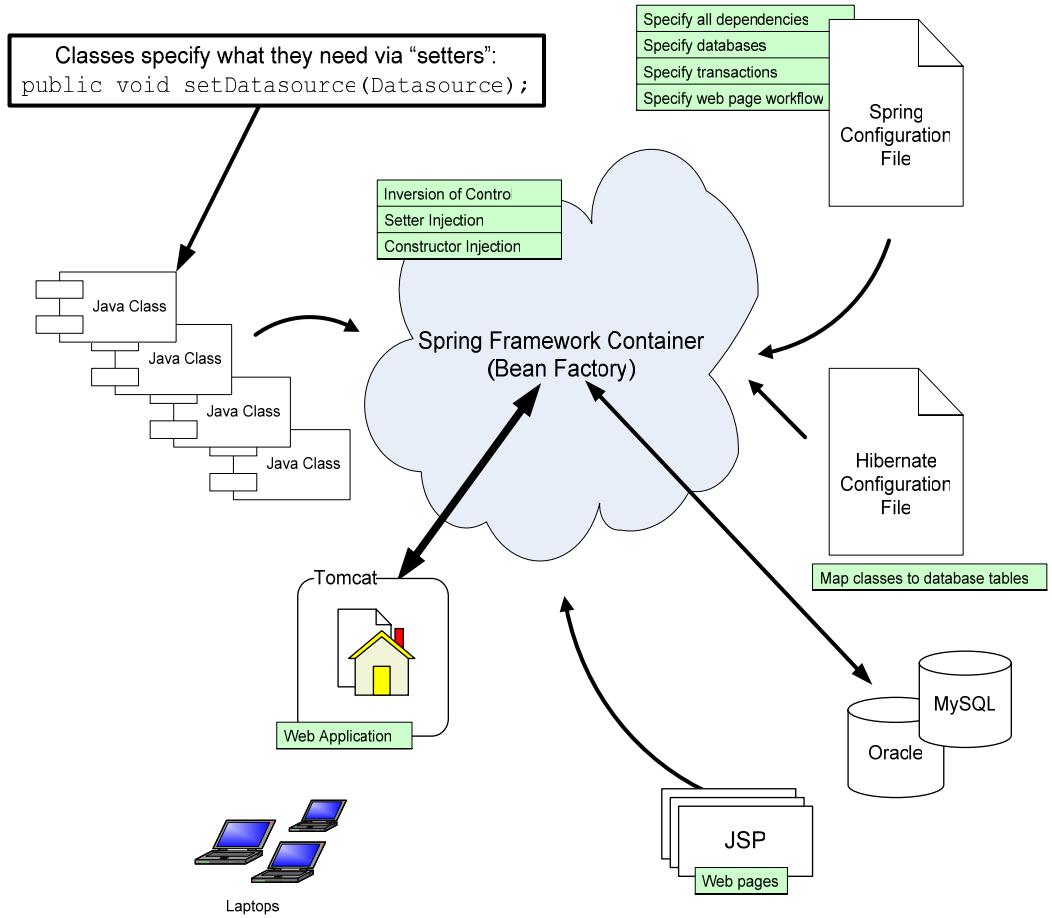
Brief History of Web Application Platforms



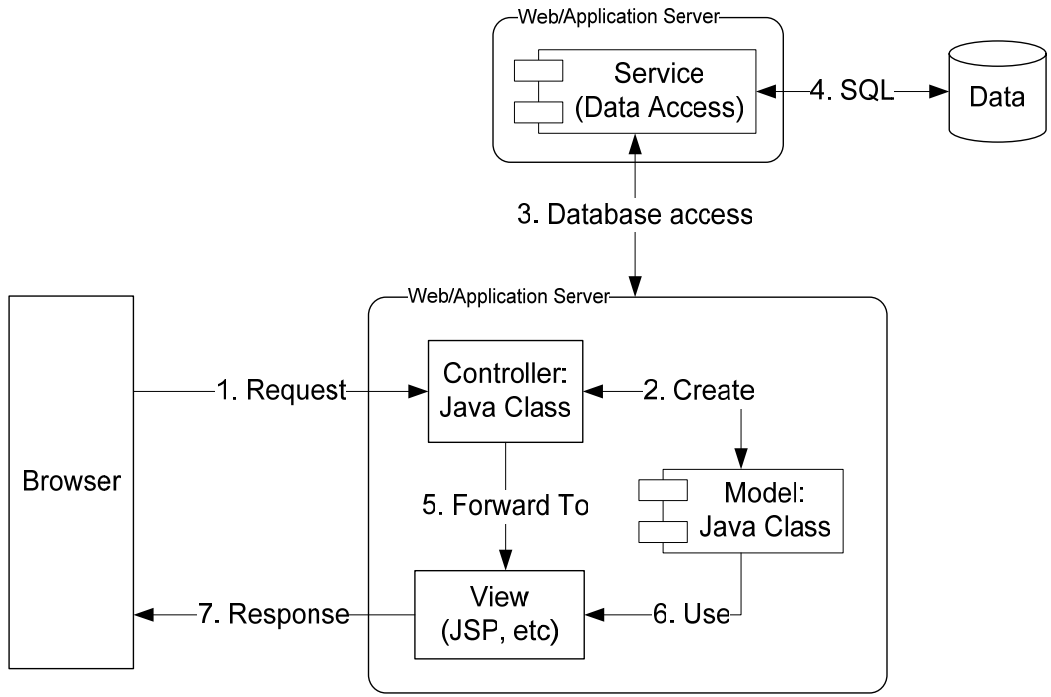
References

http://www.springframework.org
http://www-106.ibm.com/developerworks/webservices/library/co-tmline/notes
http://www-106.ibm.com/developerworks/java/library/j-what-are-ejbs/part1/
http://perl.apache.org/about/history.html

Lightweight Container

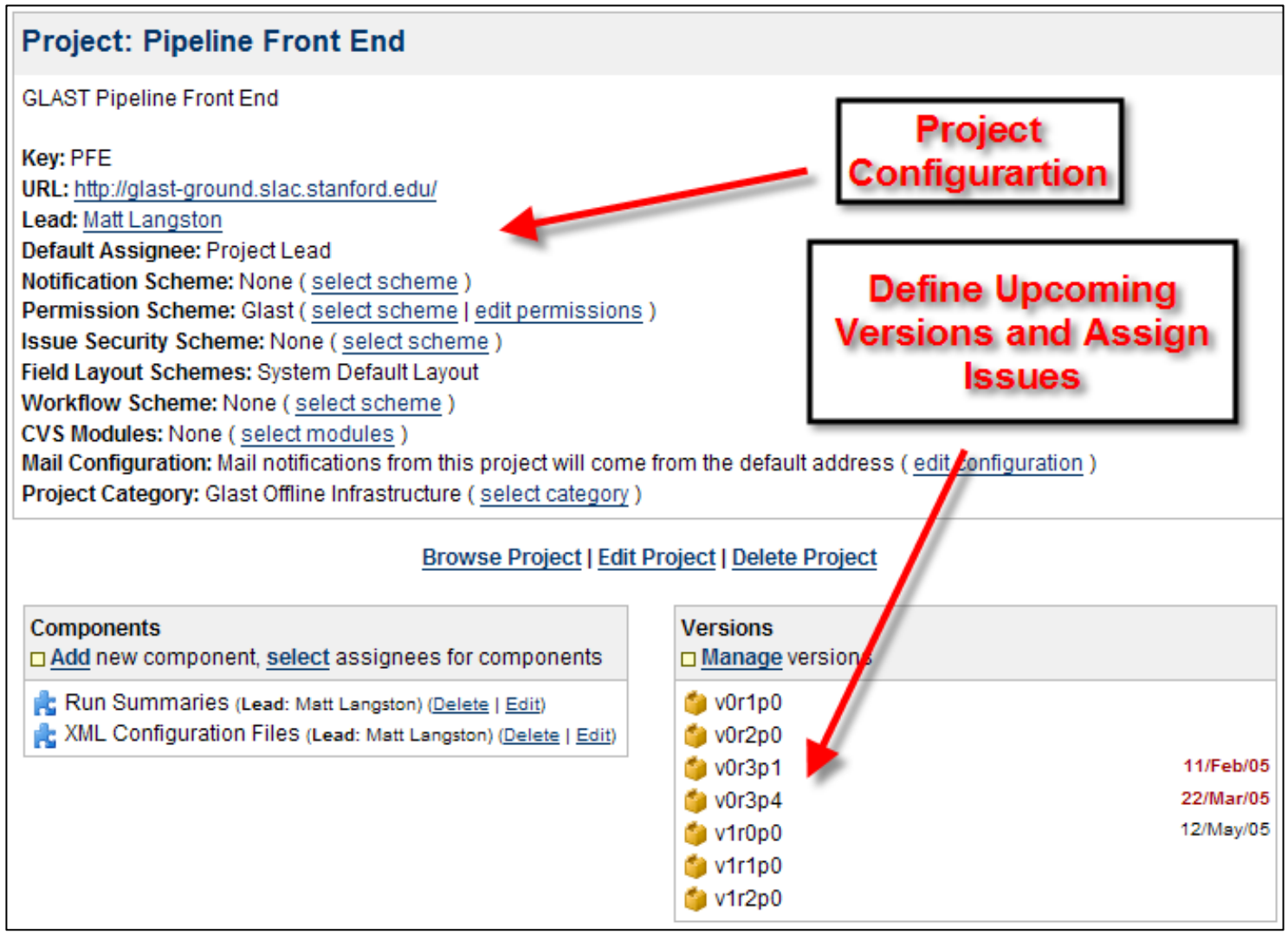


MVC type 2
Model **V**iew **C**ontroller

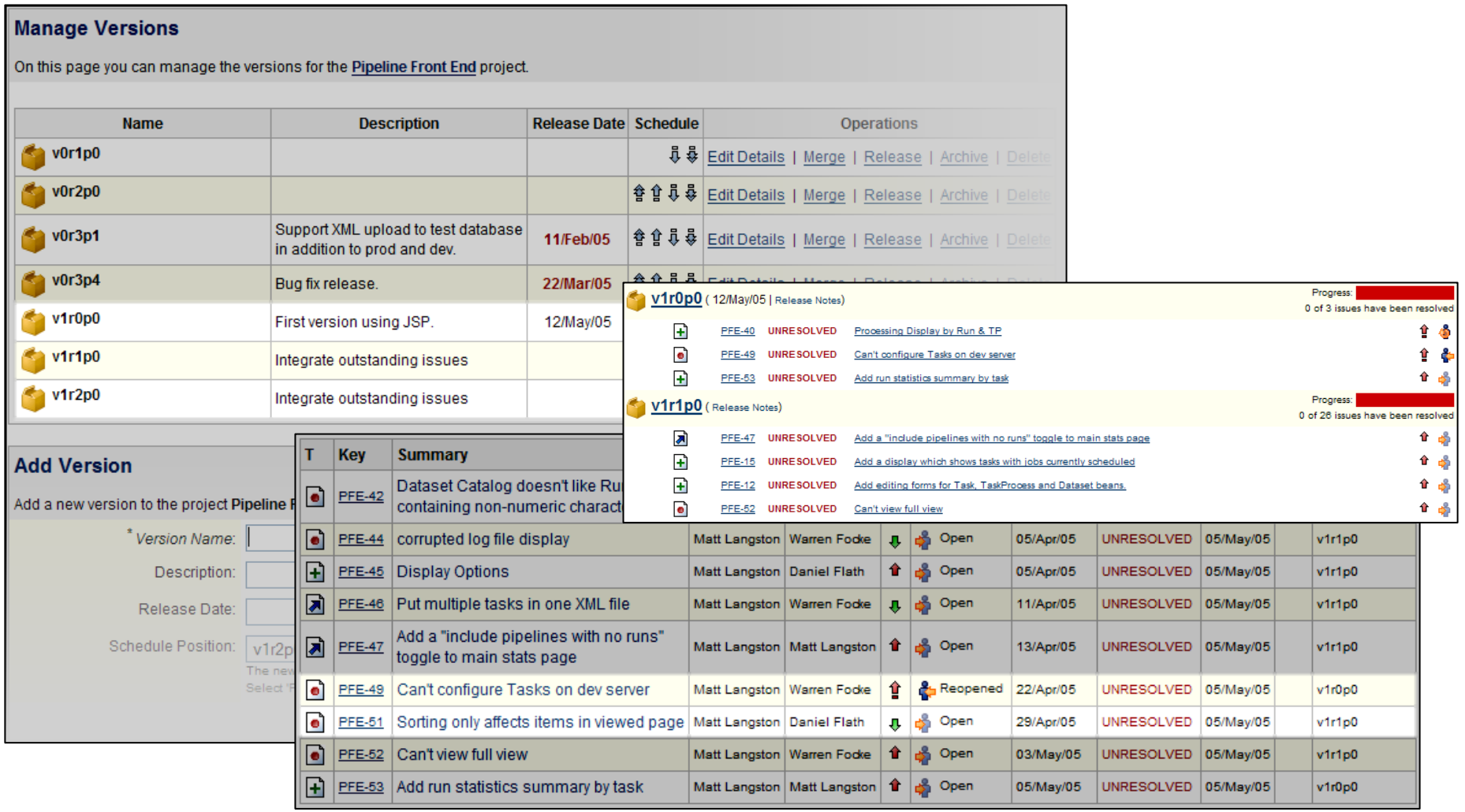


Guaranteed workflow (page order)
Guaranteed security (force login, HTTPS, Role membership, etc.)
Coarse-grained security (directory, page, page fragment
Fine-grained security (member functions, arguments, element in returned array)
Auto-populate web forms
Data validation (date ranges, run ranges, etc.)

JIRA Web Application



JIRA as a Project Management Tool



Manage Versions

On this page you can manage the versions for the Pipeline Front End project.

Name	Description	Release Date	Schedule	Operations
v0r1p0				Edit Details | Merge | Release | Archive | Delete
v0r2p0				Edit Details | Merge | Release | Archive | Delete
v0r3p1	Support XML upload to test database in addition to prod and dev.	11/Feb/05		Edit Details | Merge | Release | Archive | Delete
v0r3p4	Bug fix release.	22/Mar/05		
v1r0p0	First version using JSP.	12/May/05		
v1r1p0	Integrate outstanding issues			
v1r2p0	Integrate outstanding issues			

Add Version

Add a new version to the project Pipeline P

*Version Name:

Description:

Release Date:

Schedule Position: v1r2p

The new

Select "F

v1r0p0 (12/May/05 | Release Notes) Progress: 0 of 3 issues have been resolved

PFE-40 UNRESOLVED Processing Display by Run & TP

PFE-49 UNRESOLVED Can't configure Tasks on dev server

PFE-53 UNRESOLVED Add run statistics summary by task

v1r1p0 (Release Notes) Progress: 0 of 26 issues have been resolved

PFE-47 UNRESOLVED Add a "include pipelines with no runs" toggle to main stats page

PFE-15 UNRESOLVED Add a display which shows tasks with jobs currently scheduled

PFE-12 UNRESOLVED Add editing forms for Task, TaskProcess and Dataset beans.

PFE-52 UNRESOLVED Can't view full view

T	Key	Summary									
	PFE-42	Dataset Catalog doesn't like Ru containing non-numeric charact									
	PFE-44	corrupted log file display	Matt Langston	Warren Focke		Open	05/Apr/05	UNRESOLVED	05/May/05		v1r1p0
	PFE-45	Display Options	Matt Langston	Daniel Flath		Open	05/Apr/05	UNRESOLVED	05/May/05		v1r1p0
	PFE-46	Put multiple tasks in one XML file	Matt Langston	Warren Focke		Open	11/Apr/05	UNRESOLVED	05/May/05		v1r1p0
	PFE-47	Add a "include pipelines with no runs" toggle to main stats page	Matt Langston	Matt Langston		Open	13/Apr/05	UNRESOLVED	05/May/05		v1r1p0
	PFE-49	Can't configure Tasks on dev server	Matt Langston	Warren Focke		Reopened	22/Apr/05	UNRESOLVED	05/May/05		v1r0p0
	PFE-51	Sorting only affects items in viewed page	Matt Langston	Daniel Flath		Open	29/Apr/05	UNRESOLVED	05/May/05		v1r1p0
	PFE-52	Can't view full view	Matt Langston	Warren Focke		Open	03/May/05	UNRESOLVED	05/May/05		v1r1p0
	PFE-53	Add run statistics summary by task	Matt Langston	Matt Langston		Open	05/May/05	UNRESOLVED	05/May/05		v1r0p0

Overall Summary

- **GLAST offline software represents a confluence of HEP and Astro communities**
  - **Looks like HEP for instrument simulation**
    - **C++; Gaudi; Geant4; Root; Kalman filter tracking etc**
  - **Looks like a telescope for analysis of the sky**
    - **FTOOLS, FITS etc**
  - **We have a small group**
    - **Trying to automate as much as we can**
    - **Trying for good gui tools; good user doc**
    - **Release Manager, System Tests**
    - **processing pipeline**
      - **Backbone of Science Ops Center**
  - **Modest data volumes**
    - **Keep it all on disk**
    - **Provide smart data servers for the collaboration**