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
    - for doing digi, recon
    - creating reports
    - Preparing for calibrations
  - Generic MC
    - DC2, background runs etc etc
  - ISOC
    - Flight operations
    - What about environmental testing, at Spectrum Astro, KSC?
  - Should there be a “User Facility” for Joe Schmoe to run his random MCs?

Initial Pipeline Requirements

- **Planned to be the backbone of ISOC**
  - Highly configurable in terms of what it can run
    - Any old scripts, not just tailored to GlastRelease
  - Flexible in terms of conditions for initiating a process
  - Steve Culp will show the envisaged uses in the ISOC

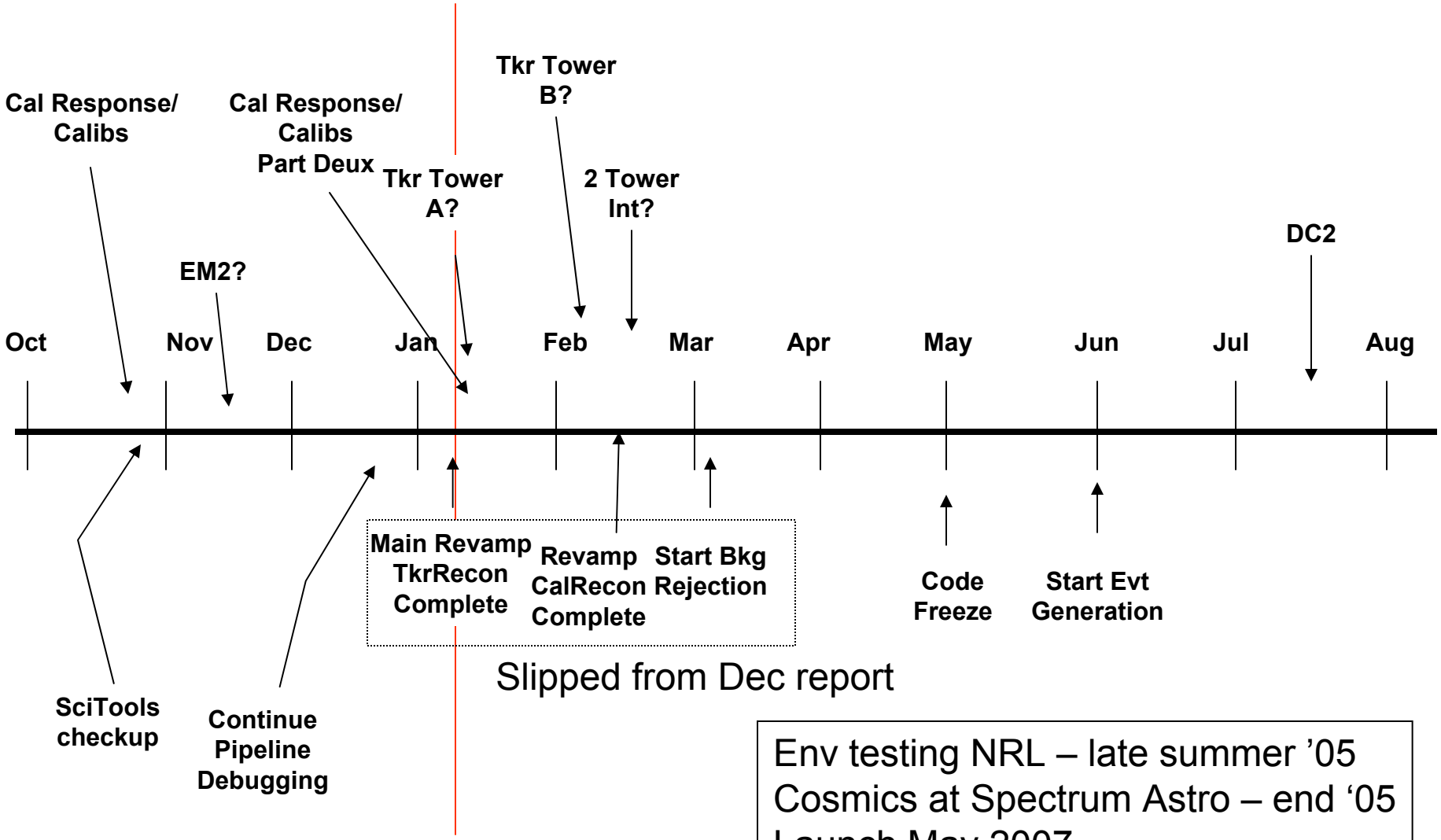
- **High level reqs (Flight Ops):**
  - automatically process Level 0 data through reconstruction (Level 1)
  - provide near real-time feedback to ISOC (evolved 'system tests')
  - facilitate the verification and generation of new calibration constants
  - re-process existing data
  - produce bulk Monte Carlo simulations
  - backup all data that passes through

http://confluence.slac.stanford.edu/display/GRITS/3.2+Functional+Spec+for+Processing+Pipeline

- **Further discussion on design issues at:**

http://confluence.slac.stanford.edu/display/Gino/Gino+Design+and+Feature+Requests

Timeline



Slipped from Dec report

Env testing NRL – late summer ’05
Cosmics at Spectrum Astro – end ‘05
Launch May 2007

Newly Identified Functions

- Gino as server
  - We should be able to hold conversations with it to check aliveness etc; handle log files; finer control of scheduling db checks etc
- Splitting input files
  - Will need to route input data (after digi?) to the famous 75 CPUs to turn downlink around in an hour; and reassemble somehow afterwards
- Concatenate/prune MC runs
  - A post-processing step to collect all good runs, and concatenate and prune them
- Better identification of underlying apps run in Gino (see MC talk)
- Archiving strategies
  - How to recognize when all processing on a run is complete and to grab all produced files (even ones the pipeline is not told about)
- Write protect output files
- What about remote MC generation? eg Lyon, Perugia
  - Berrie did ½ our backgrounds generation for DC1 at Lyon
  - Offering to do more for DC2