

Chandra Source Catalog Review

CXCDS Data Processing Implementation

Janet DePonte Evans
Kenny Glotfelty
Chandra X-ray Center
February 9, 2006

Data Processing Infrastructure

- *Chandra SDP* is fully automated, and is controlled by our *Automated Processing (AP)* software
- *AP* is an application that manages pipeline processing
 - Based on the *STScI OPUS* application with many *Chandra*-specific enhancements
 - Executes pipeline stages: retrieves data, runs pipeline, store results
 - *Darch (Data Archiver/Retriever)*: Archive/Retrieves data to/from cache and/or archive
 - *OST (Observation status tracker)*: Kicks off a pipeline when all the required data products are available

Data Processing Infrastructure cont.

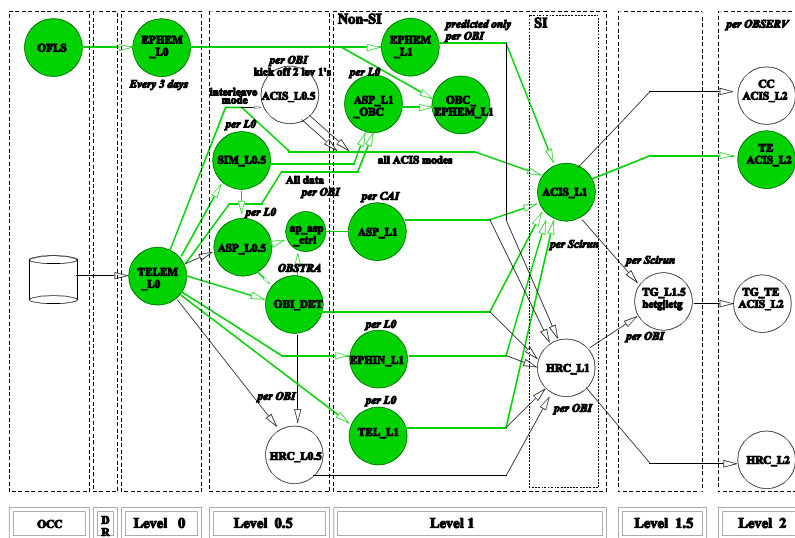
- *Processing* is split into several *levels* based on functionality
 - Level 0 – 2: telemetry decom, instrument calibrations, data filtering
 - Level 3: Chandra Source Catalog production
- Within each level, there are multiple pipelines that process specific types of data
 - Each pipeline comprises several (or several dozen) programs that perform specific tasks
- *Processing* is inherently non-linear, with numerous parallel threads that need to be completed before further steps can proceed

February 8 – 10, 2006

Chandra Source Catalog Review

Page 3

SDP Pipeline Diagram

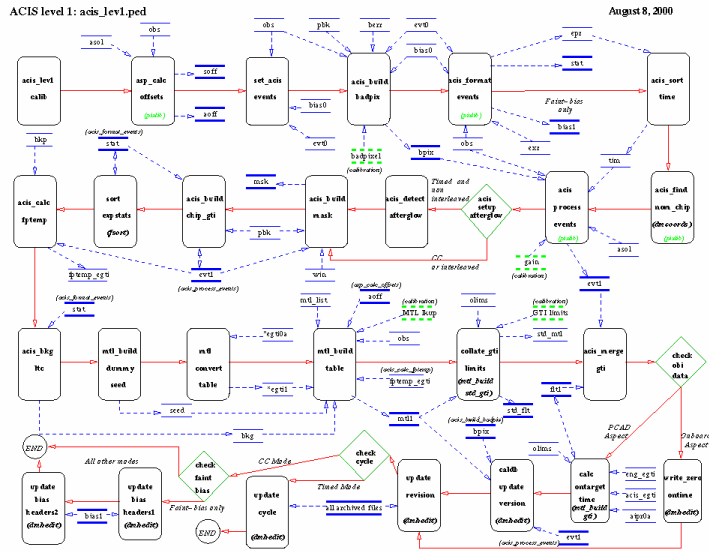


February 8 – 10, 2006

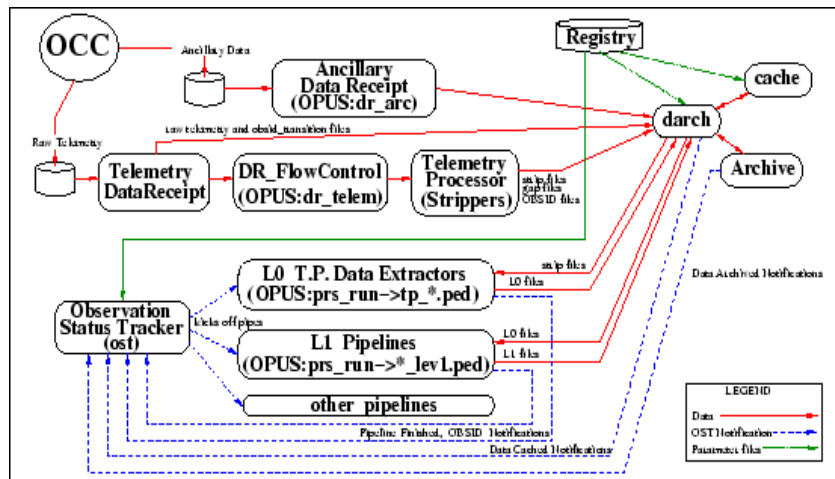
Chandra Source Catalog Review

Page 4

ACIS Level 1 Pipeline Diagram



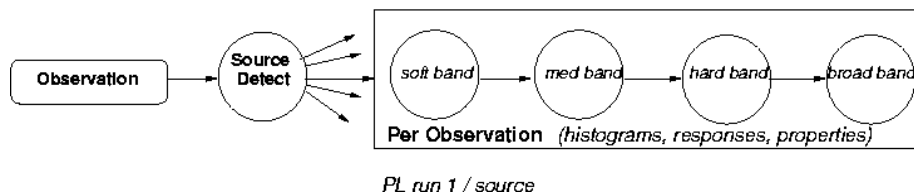
AP Configuration for SDP



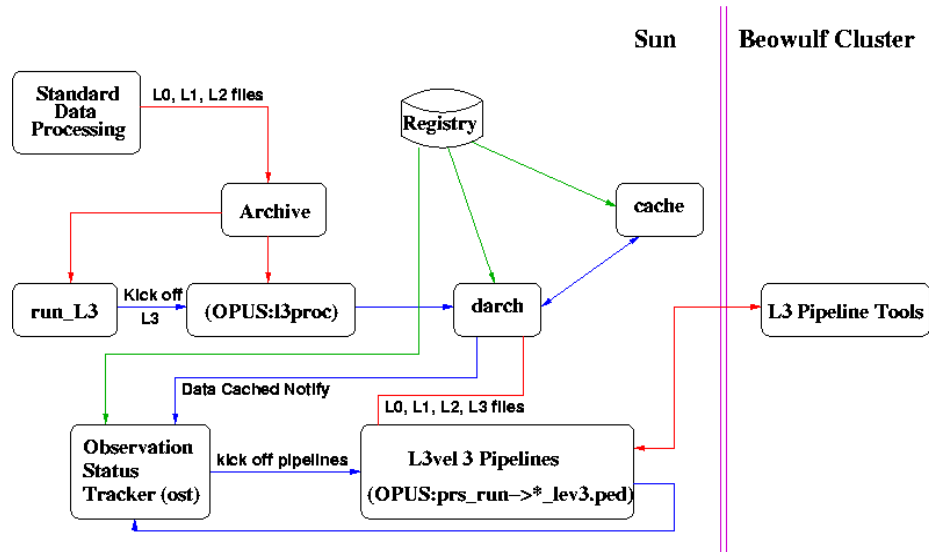
Supporting Chandra Source Catalog

- AP components have been configured to also manage Chandra Source Catalog pipelines
 - Configured on Sun and 15 node Beowulf/Linux cluster
- 2 pipelines developed to support L3
 - Detect:
 - applies calibrations, builds maps, source detection, merge
 - Per Source:
 - Creates histograms and response maps for each source & computes source properties

L3 Pipeline Diagram



AP Configuration for CSC



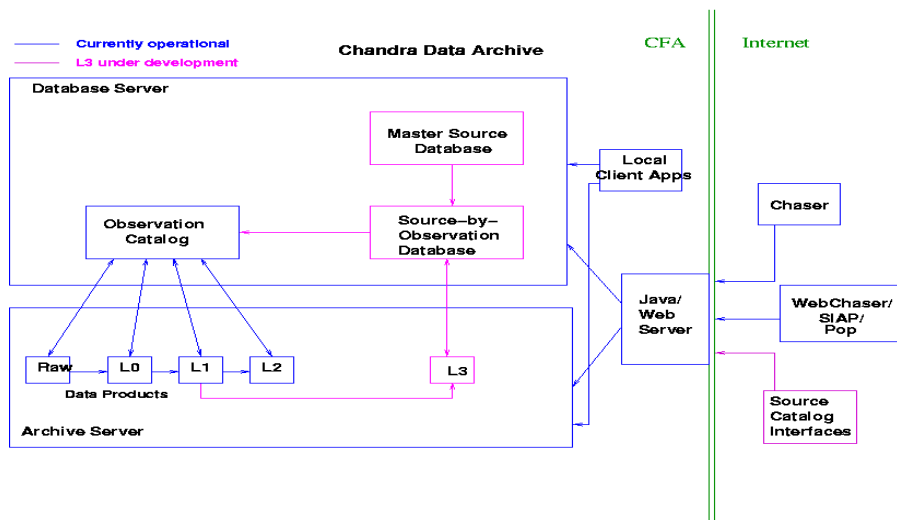
Chandra Data Archive

- Data is managed by 2 types of servers
 - Archive server stores data products and all other data files
 - Database server which stores file metadata
- Applications have 2 access interfaces to the archive
 - Internal access is through a Local Client Application
 - Utilized by *AP*
 - External access through a Java/Web Server which serves as a gateway or as an application server
 - Utilized by Chaser/Webchaser/VO

Chandra Data Archive cont.

- Maintains data history of all products
- Tracks data product version information
- Accounts for proprietary time and triggers public release of data

Chandra Data Archive Architecture Overview



Development Status

- Chandra Source Catalog production has been mapped to the extensible and scalable design of the data system
- New tools developed for L3 will also be released in CIAO for data analysis use. To date they include 6 new tools and 1 major enhancement:
 - merge_src, roi, dither_region, streakloss, dmimgfilt, glvary, acis_streak_map
 - wavdetect upgrades for better false source statistics
- L3 pipeline for point-like detection and properties has been developed

Development Status cont.

- AP configured on Sun/Linux to support L3 processing
 - Currently evaluating speed improvements in cache server
- Archive structures and pipeline interfaces for L3 data products are about to be integrated in the system
- Design of the source catalog and its association with the Observing Catalog and data products are in progress
- Have processed ~20 Obsids from “List of 100”
 - science review and feedback
 - system robustness study
 - timing evaluations