User Software Overview

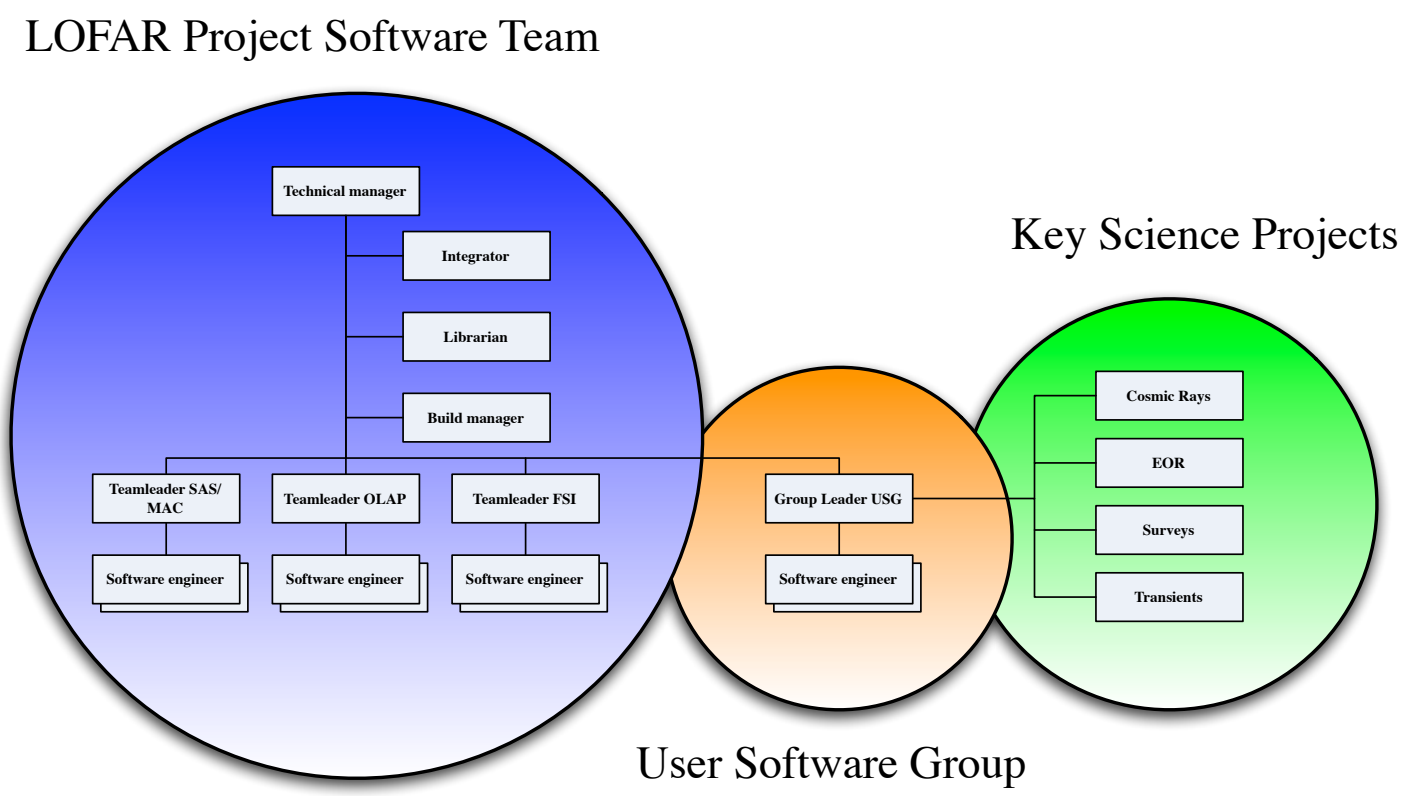
LOFAR DCLA Project Meeting
26 June 2007

Michael Wise

Outline

- Group Overview
- Science Pipelines
  - Designs and module breakdowns
  - Implementation and roll-out
  - KSP status summary
- Infrastructure and Tools
  - Data formats
  - Core libraries
  - Offline analysis tools
- Support Activities
- Demos

Activities



- Coordination *KSP developments, LST, E-LOFAR consortium, partner organizations*
  - Development *Requirements, algorithms, infrastructure, tools, and testing*
    - Support *Development, testing, documentation, code repository*

Science Pipelines

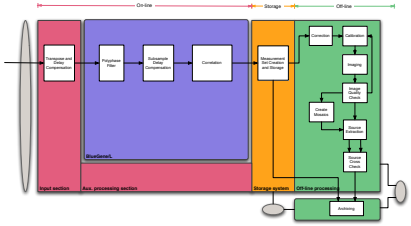
Science Pipelines

- Based on KSP requirements
- Mix of LOFAR project and USG/KSP development
- Includes online and offline components
- Pipeline and interactive components

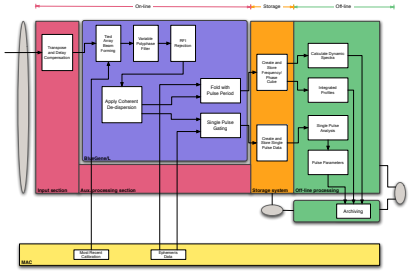
⇒ *Preliminary integrated software plan*
9 observing modes
~ 60-70 FTE-yrs of effort
Phased roll-out of capability

Pipeline Designs

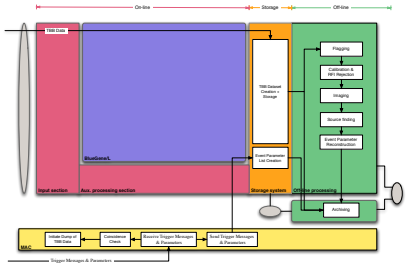
Survey mode



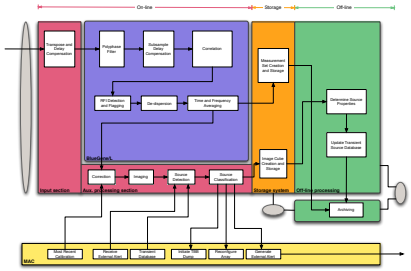
Known pulsars mode



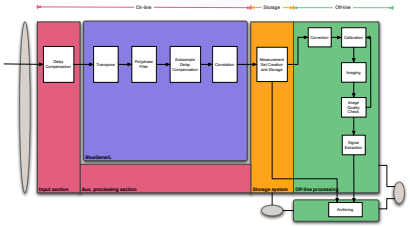
VHECR mode



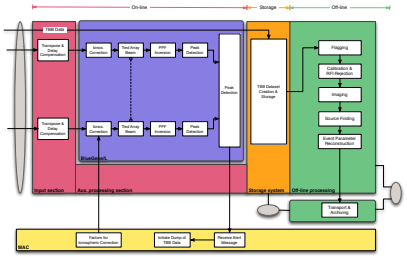
Transient detection mode



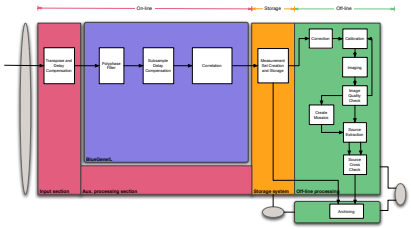
EoR mode



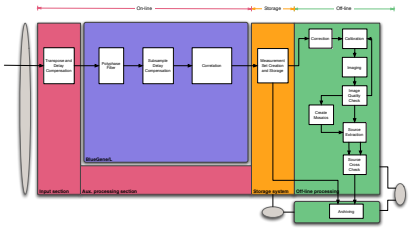
UHEP mode



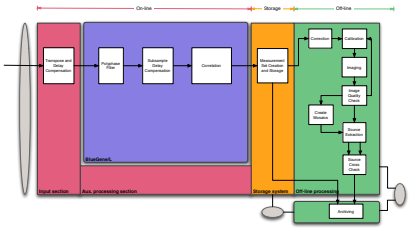
Pulsar Survey mode



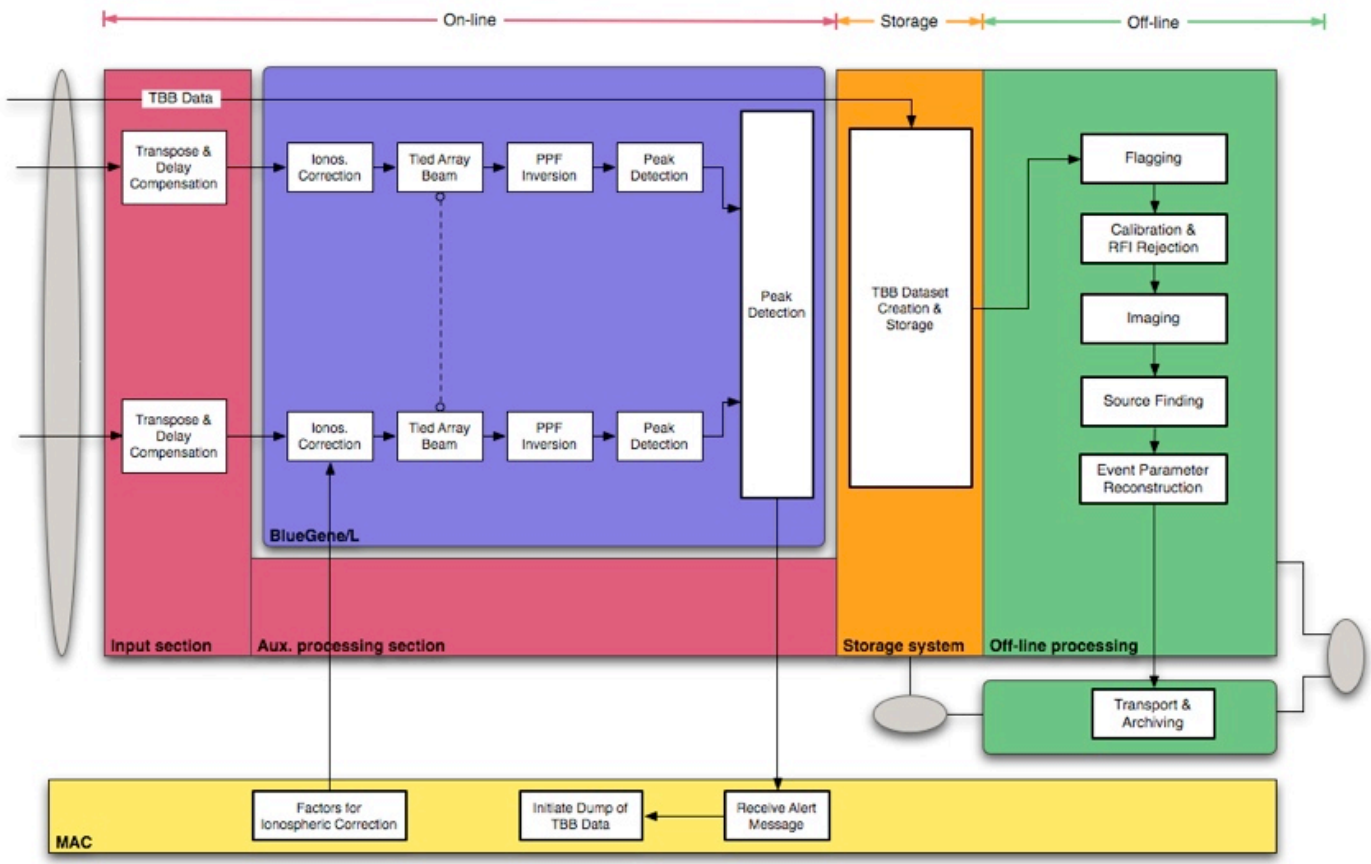
HECR mode



TS mode

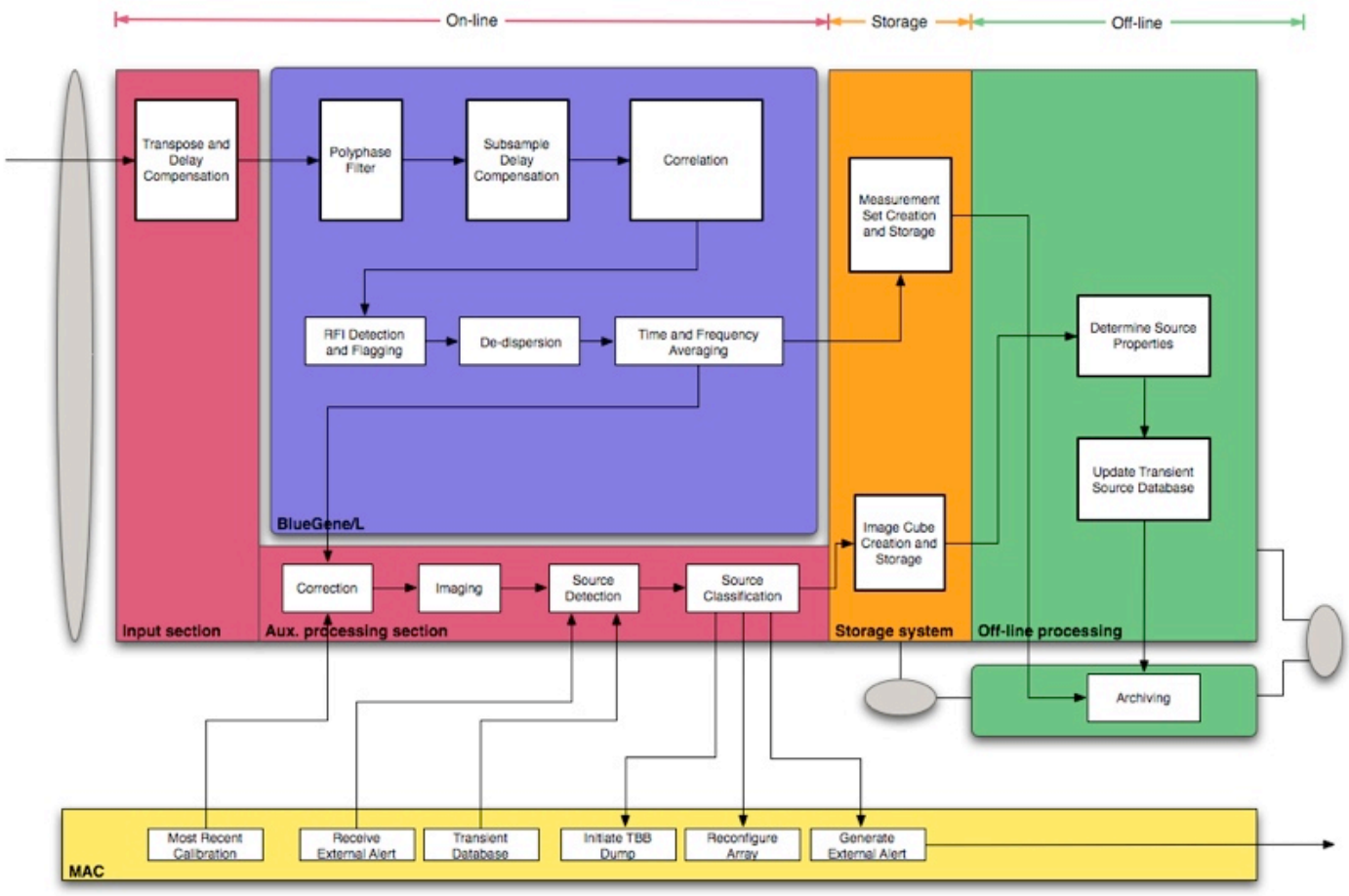


Pipeline Designs



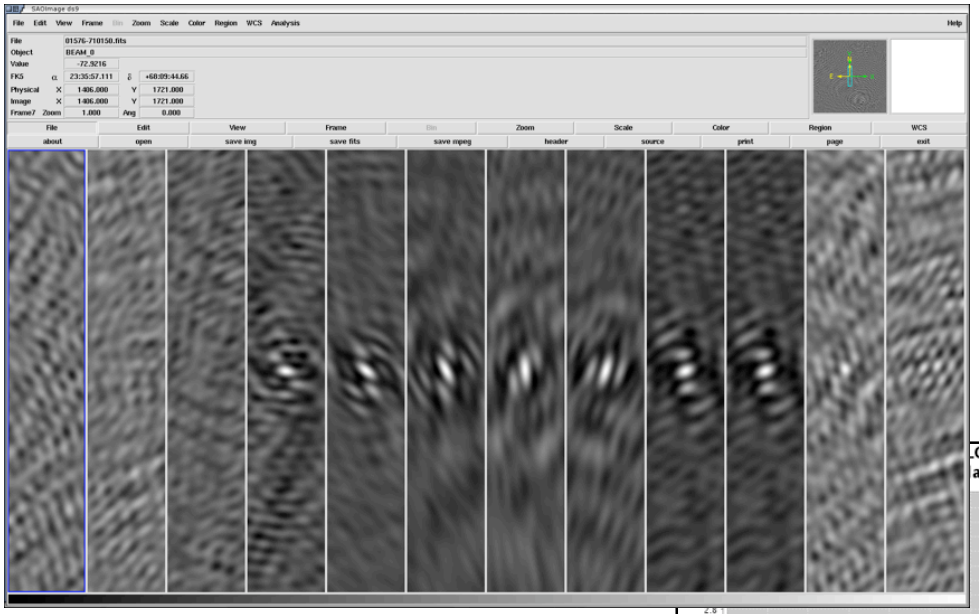
UHEP mode

Pipeline Designs



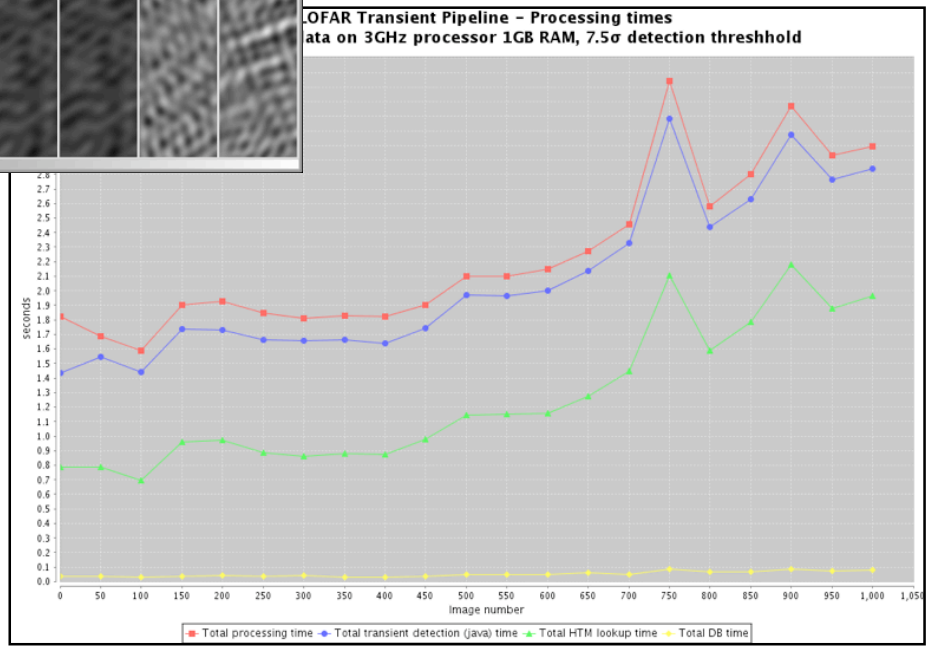
Transient detection mode

Transient Pipeline



- *Real-time pipeline*
- *Rapid detection*
- *Source classification*
- *Trigger generation*
- *Transient database*

Detection and storage performance tests using simulated datasets



Survey Pipeline

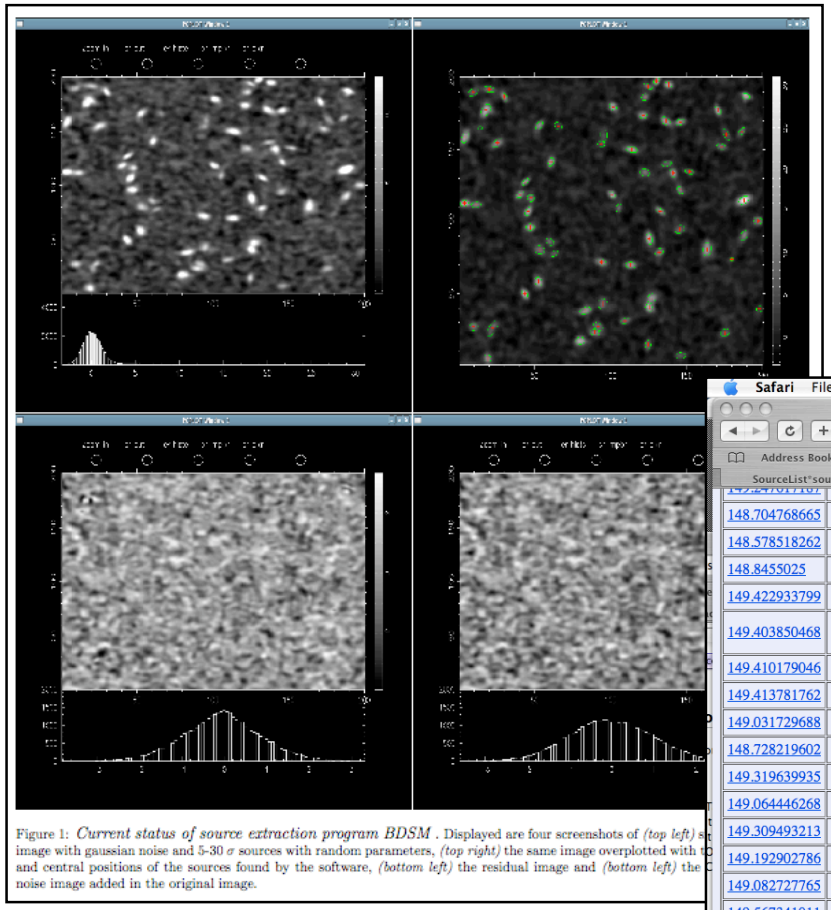
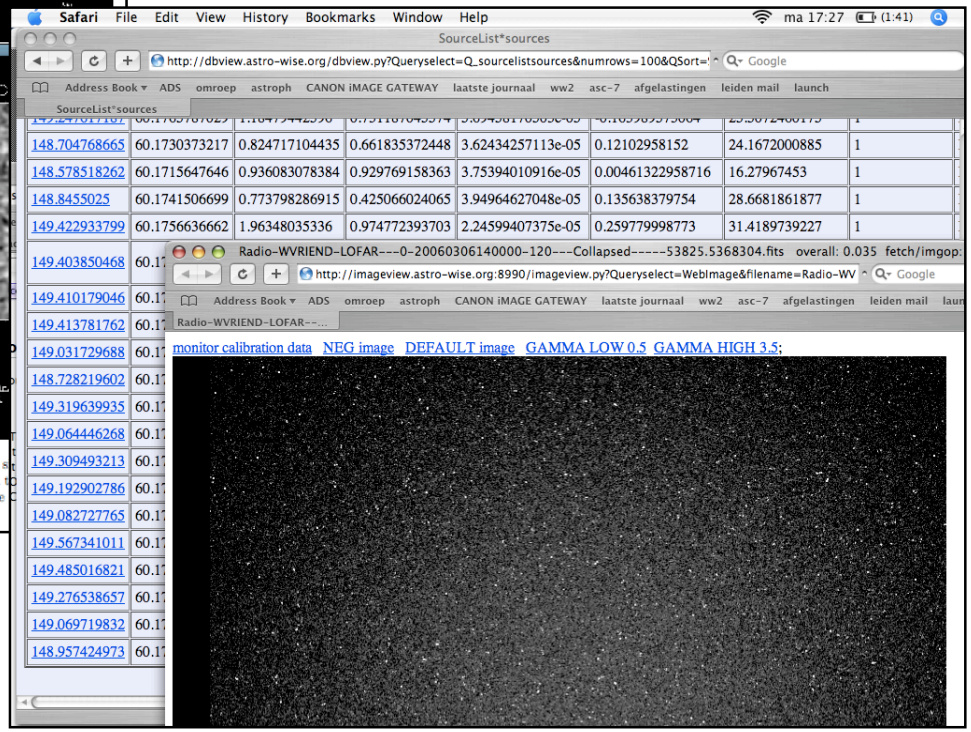


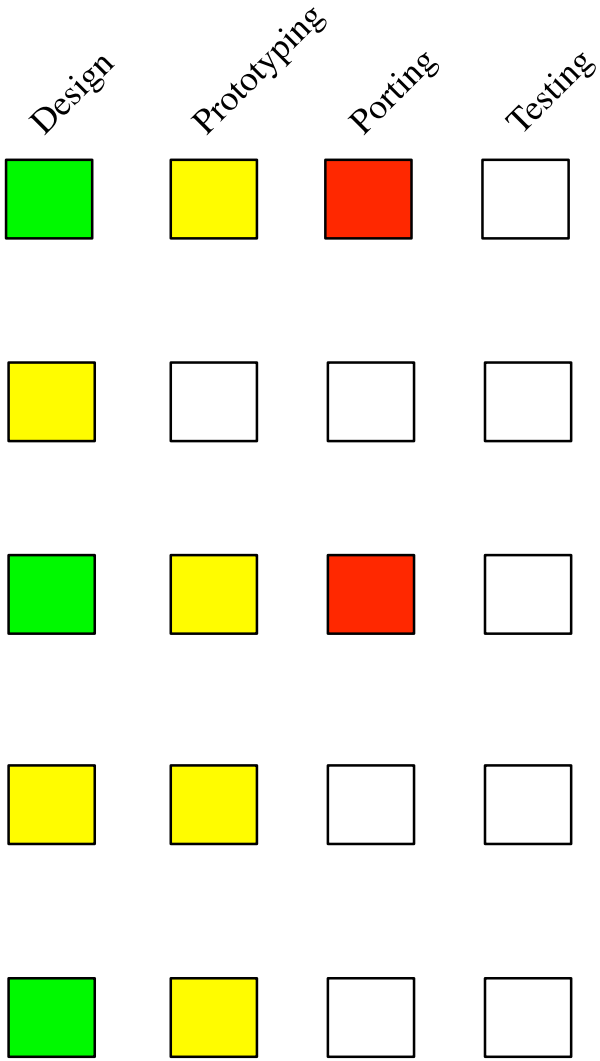
Figure 1: *Current status of source extraction program BDSM* . Displayed are four screenshots of *(top left)* s image with gaussian noise and 5-30 σ sources with random parameters, *(top right)* the same image overplotted with t and central positions of the sources found by the software, *(bottom left)* the residual image and *(bottom left)* the noise image added in the original image.

- *Quality control*
- *Mosaicing*
- *Source detection*
- *Catalogs*



KSP Status Summary

- Cosmic Rays
  - Prototype pipeline near completion
  - Porting to USG framework underway
- EoR
  - Preliminary pipeline design
  - Component module inventory needed
  - No developer available
- Pulsars
  - Designs and prototypes exist
  - Data products and formats need definition
  - Work begun on de-dispersion module
- Surveys
  - Prototype source detection pipeline exists
  - Remaining modules require specification
- Transients
  - Prototype pipeline near completion
  - Performance testing about to begin
  - Classifier module needs better design



Infrastructure and Tools

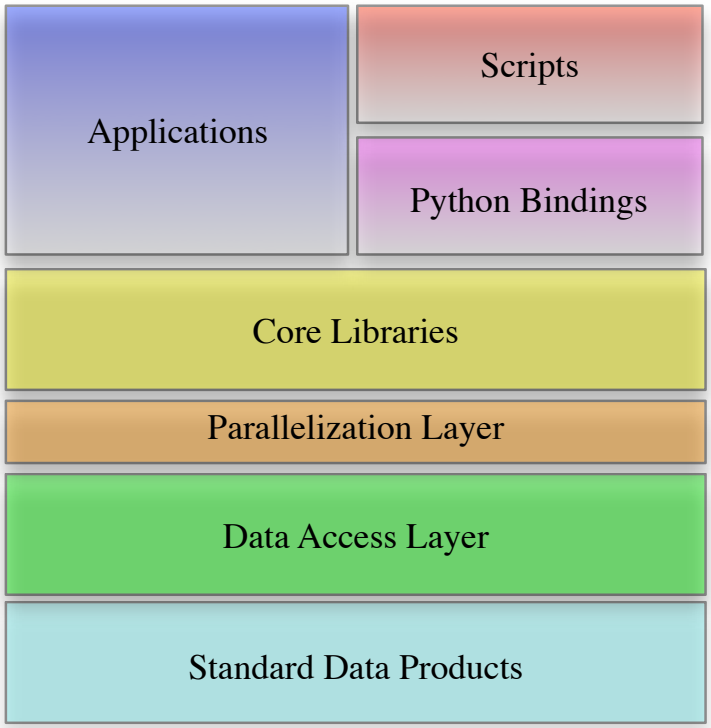
System Design

Requirements

- Support HPC (*Large files, fine and coarse parallelism, GRID, etc.*)
- Compatibility with other major packages (*CASA, IRAF, ROOT, CIAO, etc.*)
- Support multiple platforms
- Support custom analysis
- Rapid development
- Supportable

Design

- Python scripting layer
- Native APIs will be in C/C++
- Python bindings for all libraries
- Support (MPI+PVM)
- Use or adapt existing software
- Open-source software



Standard Data Products

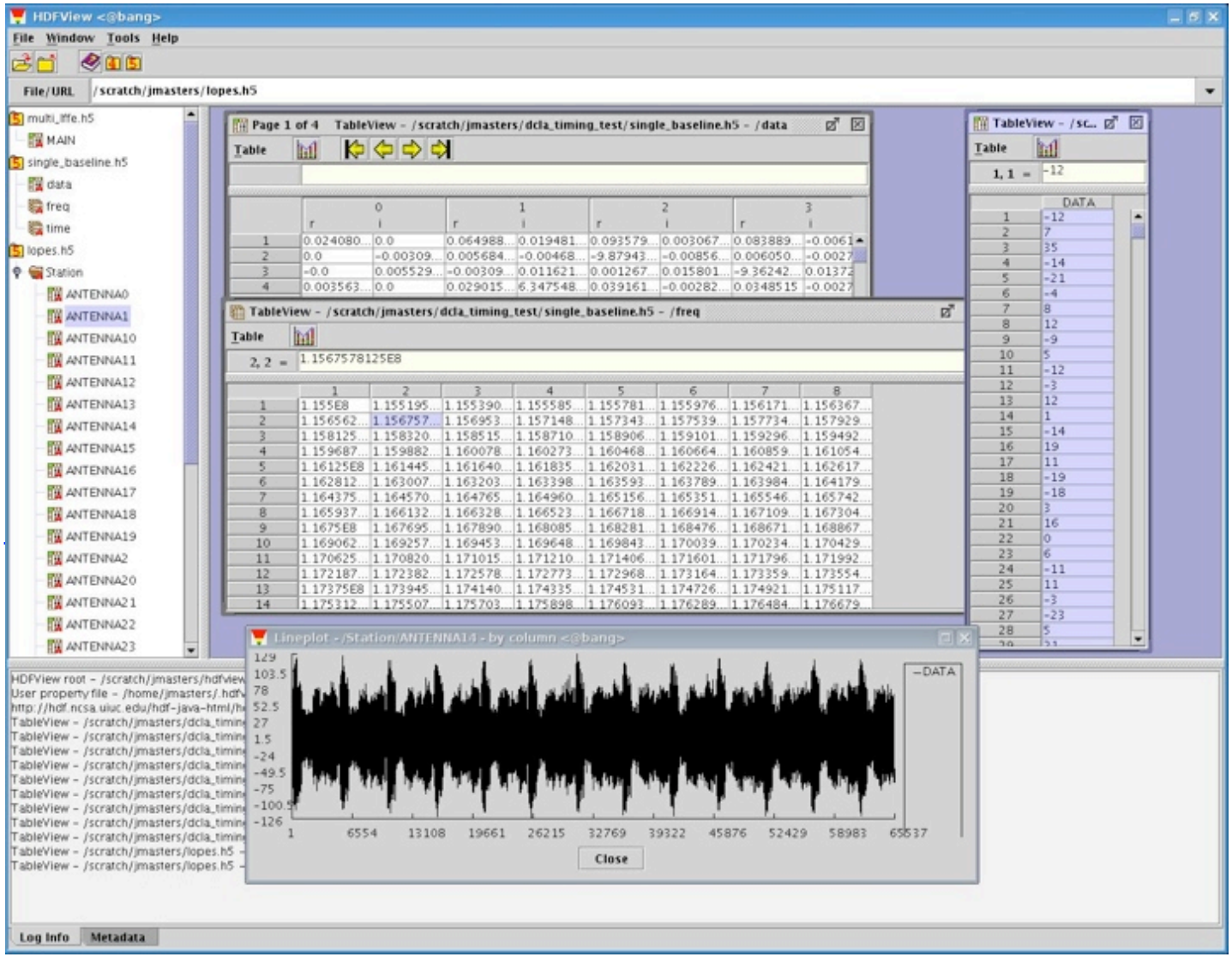
Data	Formats	ICD	DAL I/O	Availability
Time series	HDF5	1.0	R+W	Now
Beam-formed	HDF5	0.5	R+W	Soon
UV data	MS/HDF5	0.2	R/R+W	Q3 07
Image cubes	FITS/HDF5	0.2	--/R+W	Q4 07

Supporting several formats

- AIPS++ measurement sets, tables *(CASACORE)*
- FITS images, tables *(CFITSIO)*
- HDF5 tables, image cubes *(HDF5IO)*
- Raw telemetry formats *(TBB, beam-formed, etc..)*
- LOPES, ROOT, etc...

Need formats for metadata and calibration products

Standard Data Products



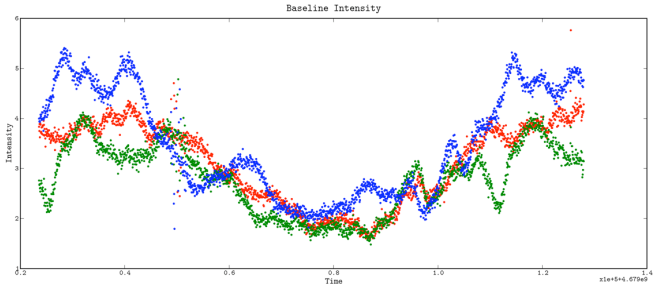
Core Libraries and Tools

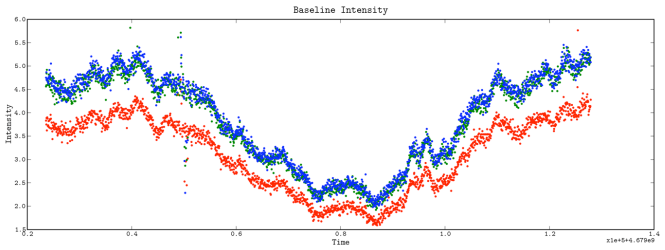
- DAL under development
  - Beta version of HDF5 I/O ~Now
  - Full I/O capabilities ~Q3 07
  - Slicing/Collapsing ~Q4 07
  - Streaming, Parallel I/O ~Q1/Q2 08

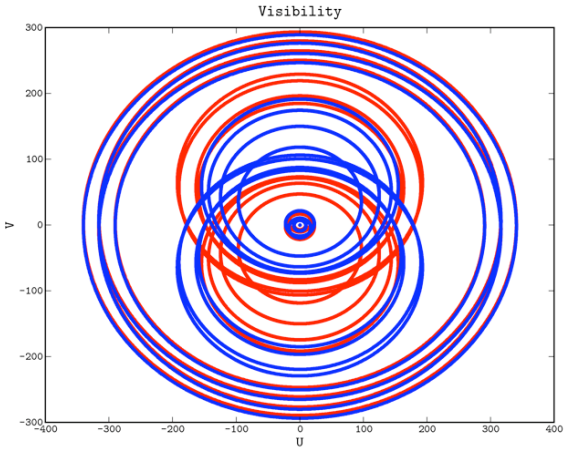
⇒ ***Used for CS1 TBB time series data this summer***

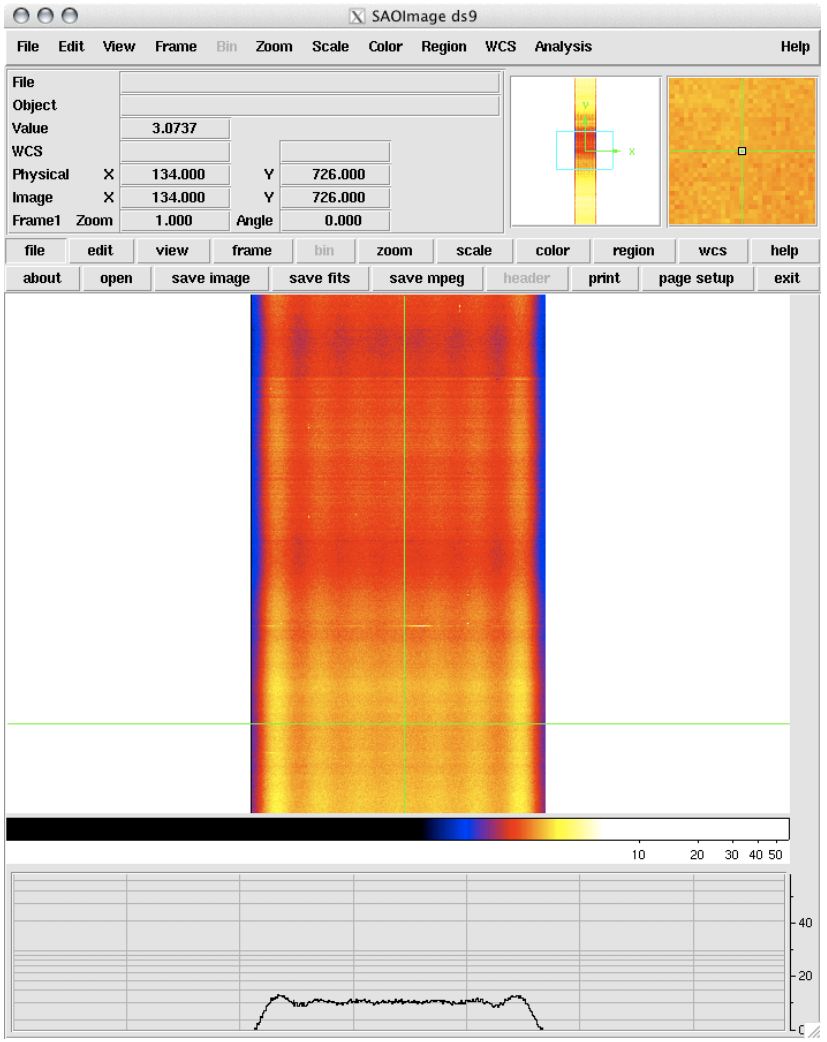
- Additional core libraries
  - Parallelization Layer (OpenMP-aware)
  - Data visualization (DVL), General math library (GML)
  - Need better requirements (especially for DVL)
  - Beta version(s) ~Q4 07/Q1 08
- Tools
  - Reprocessing (metadata inspection, flagging, visualization)
  - Analysis (source detection+characterization, mosaicing, visualization)
  - Currently compiling specs. (inputs welcomed!)

Visualization Tools

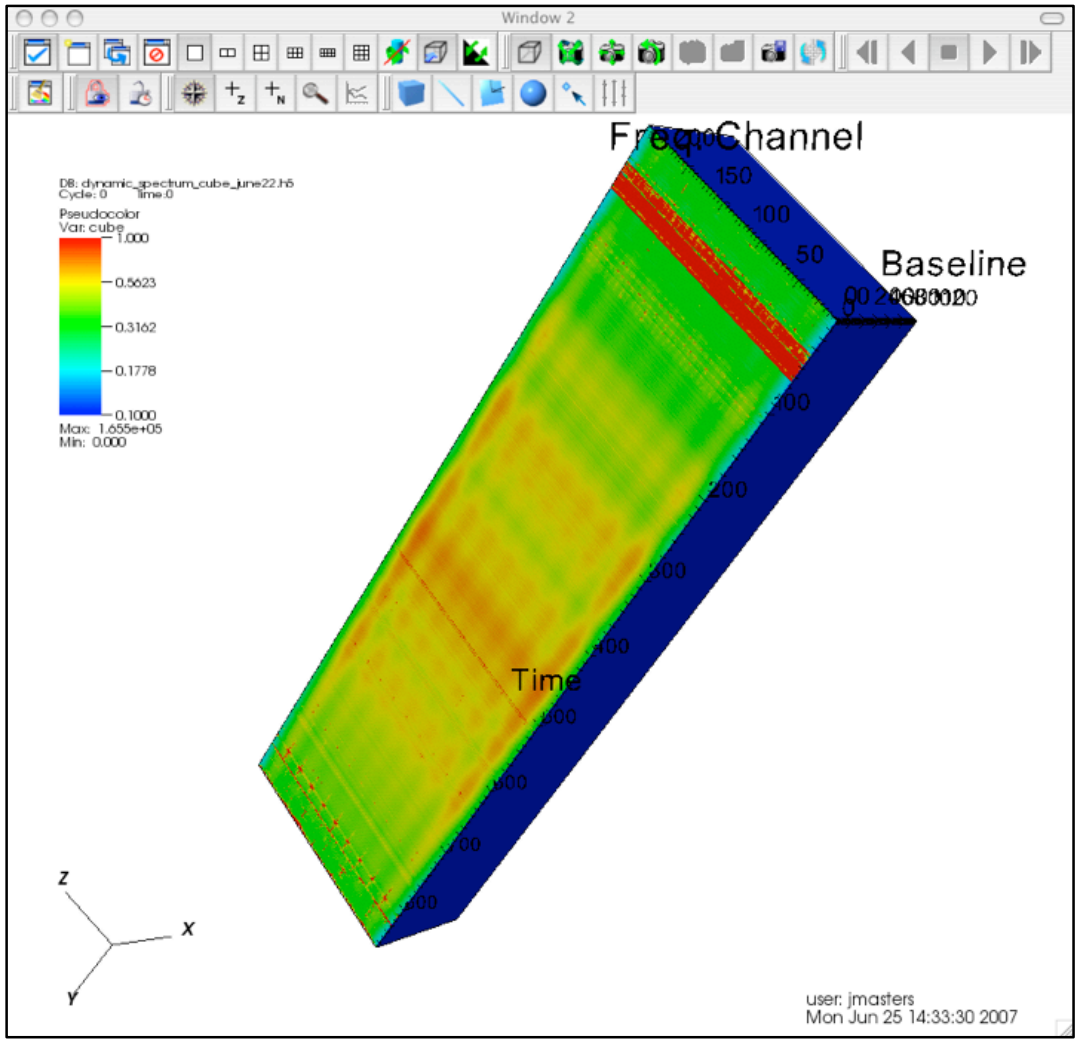




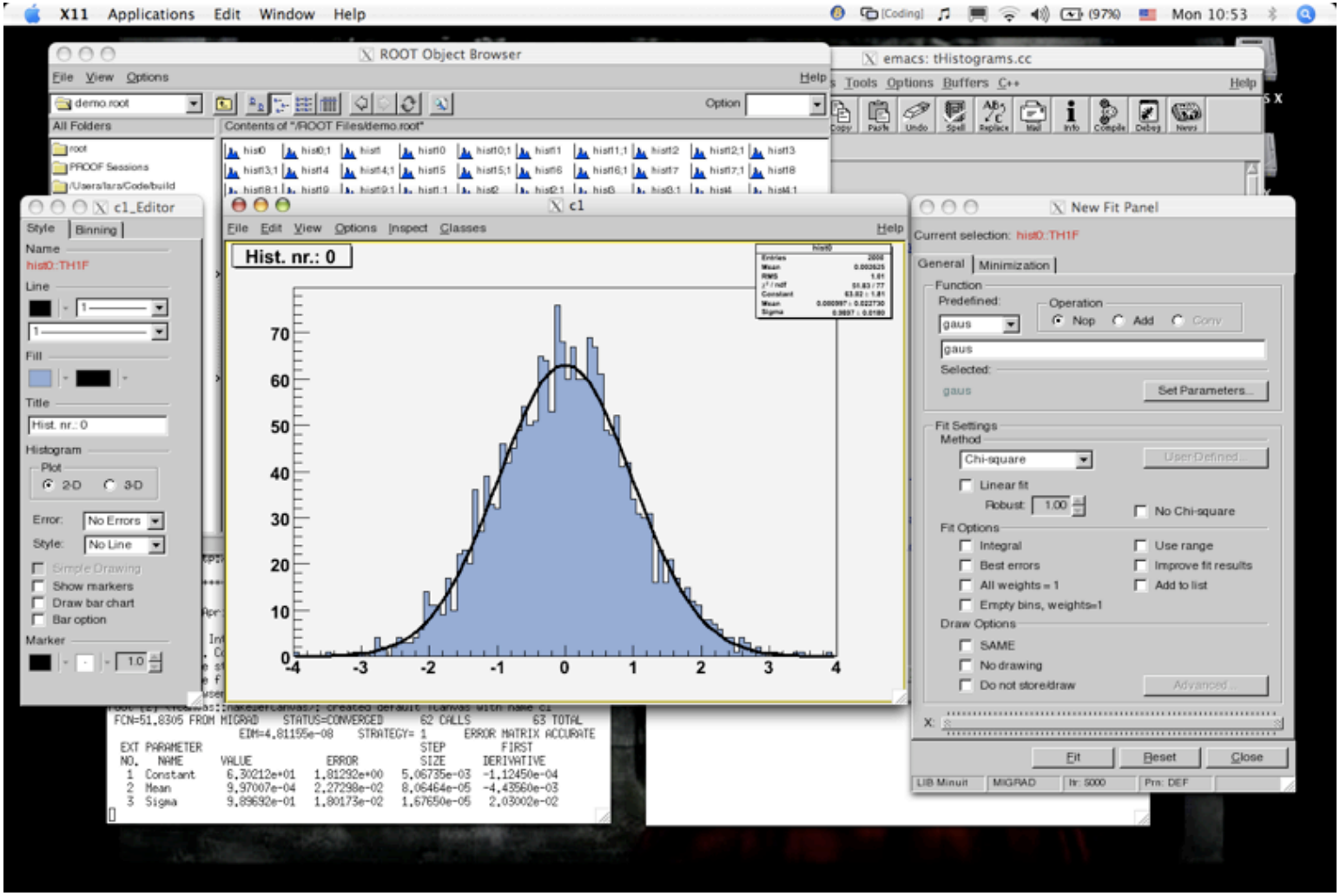




Visualization Tools



Access External Packages

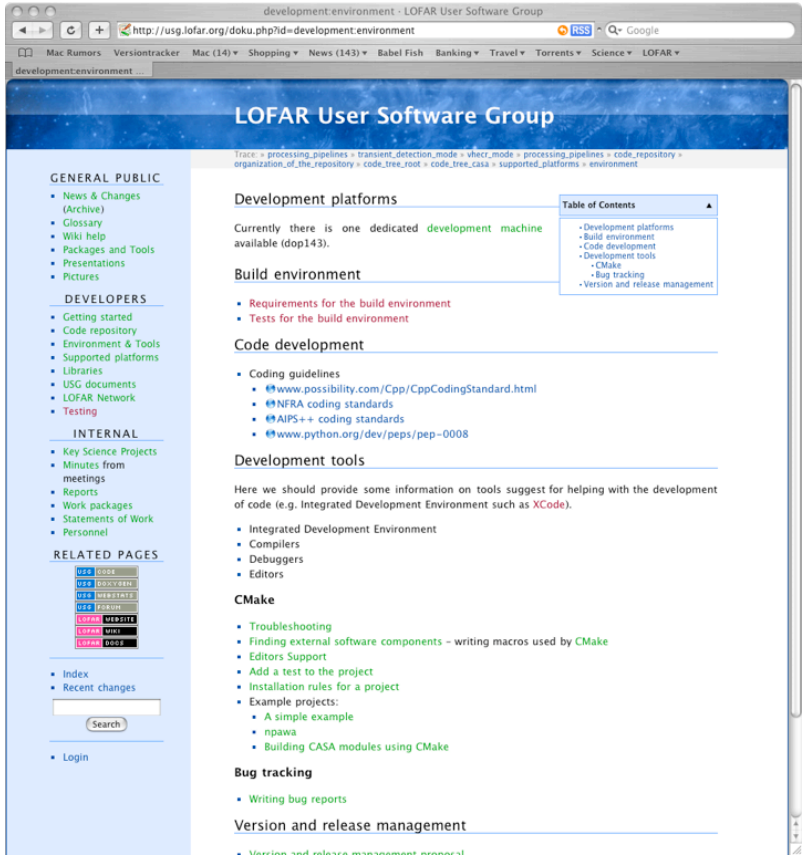


Support Activities

Support Activities

- Established USG web server
  - Online now
  - Some developer's documentation available
- Development and porting platforms
  - USG development machine available (SUSE 10.1 +VMWare)
  - 3 Porting machines for Debian, RedHat, and Fedora available
- Software repository available
  - Preliminary build environment defined
  - Preliminary code tree defined
  - Available to USG/KSP developers now
- Coming soon
  - User contributed code area
  - Bug tracking, Discussion forum
  - Build and testing framework

USG Web Server



http://usg.lofar.org

Next Steps

- Focus on real-time pipelines
- Finalize standard data products
- Define metadata formats and interfaces
- Refine library and tool requirements
- Produce development timelines
- Preliminary release schedule

Demos

John Swinbank ⇒ *Transient detection pipeline*

Joe Masters ⇒ *Data Access Library*