EVLA Project Book, Chapter 10

DATA MANAGEMENT

Tim Cornwell, Frazer Owen, and Gustaaf van Moorsel
Last changed 2001-July-18

Revision history
2001-July-18: Initial version

Summary
The DM deliverables will be built as part of the NRAO-wide End-to-End project (e2e). The e2e Project Book (E2ePB: http://www.nrao.edu/e2e/documents/e2eprojectbook.doc) is the principal definition. In this chapter, we briefly recapitulate the elements of e2e, and comment on EVLA-specific modifications.
The objective of the e2e project is to provide integrated end-to-end processing for observations made on NRAO telescopes. The ends are proposal submission (beginning) and scientific analysis (end). The software to be delivered must handle:

- Proposal submission and handling
- Observation preparation
- Real-time observing
- Scheduling of sets of observations
- Archiving of observations
- Pipeline processing including calibration and imaging

These are to be treated as a whole, so that information flows through the above steps without loss due to *e.g.* format changes, omissions, *etc.* For this to be possible, some changes in current operations may be necessary.

10.1 Conceptual Design

The conceptual architecture and design is discussed further in E2ePB. A straw man set of components is:

- Common Operational Model
- Proposal Submission Toolkit
- Observation Planning Toolkit
- Observation Description Toolkit
- Observing Toolkit
- Observation Scheduling Toolkit
- Archive Toolkit
- Pipeline Toolkit
- Calibration Toolkit

A more detailed description of each of these deliverables is given below, each in a separate section. Note that this breakdown is given mainly to focus discussion, and it may be revised as a result of subsequent work.

10.1.1 Common Operational Model

The Common Operational Model describes the operational model to be used for NRAO telescopes. The COM is an essential ingredient of the e2e package: it specified how NRAO telescopes are to be operated in the light of the e2e project.

10.1.2 Proposal Submission Toolkit

The proposal submission toolkit allows scientists to submit proposals to NRAO telescopes, ensuring that the information necessary to the review of the proposal is included and, as much as possible, is correct.
The PST does not require detailed description of the observation (a responsibility of the ODT), and it does not have capabilities for detailed planning of observations (a responsibility of the OPT). However, in many cases, it must be possible to go from the information in the proposal to a valid observing script. This script may require substantial subsequent modification.
An observer will use the proposal submission toolkit to construct, verify, and submit a proposal for one of more of the NRAO telescopes. This may be done collaboratively with other scientists at other locations.

10.1.3 Observation Description Toolkit

The observation description toolkit allows description of observations to be made with NRAO telescopes. The output from the ODT can be used either for an observation or as input to the Observation Planning Toolkit.
The ODT is a natural successor to existing tools like JObserve, SCHED, and the GBT Observe. The key innovations over these tools are

1. Only one tool serves all telescopes.
2. The tool will be able to drive the observation planning toolkit.
3. The ODT will provide a sufficiently high level description of the observation that the pipeline toolkit can derive suitable processing heuristics. Alternatively, the ODT will contain instructions to the pipeline.

The description will be as a script in a high level scripting language such as Glish or the simple GBT observation scripting language.
Classically, synthesis radio telescopes have been scheduled using static ASCII files obeying some well-defined format. Extra constructs for various looping features are typically added as extra definitions. Single dish telescopes have been more interactive, and some include provision for writing simple scripts in some command language.
The recent work of Rick Fisher for the GBT Observe program is one step in this direction (see *e.g.* http:/www.nrao.edu/~rfisher for more detailed examples). He defines a simple scripting language that is translated into Glish for execution when observing. The EVLA C&M team decided that one could go one step further and use Glish itself as the scheduling language.

10.1.4 Observation Planning Toolkit

The observation planning toolkit allows planning of observations using NRAO telescopes. It is a collection of AIPS++ tools that may be used by a scientist in conjunction with the Observation Description Toolkit to aid the planning of an observation. Factors relevant to the scientific success of an observation are calculated and displayed. The OPT must work in close conjunction with the ODT to allow a scientist to explore many options easily. The OPT includes an exposure time calculator as one of the simplest capabilities, but also includes simulation of observations using detailed models of physical effects such as calibration errors. Also included is a telescope simulator that can read observation descriptions and provide a MeasurementSet containing the simulated data.
The AIPS++ simulator is designed to simulate the data collected by a radio telescope. Modeling of the observation process is via the Measurement Equation formalism in AIPS++, and the data are stored in an AIPS++ MeasurementSet. Definition of the observing setup and data collection strategy is currently hard-coded into the interface for the simulator. This will be evolved to match the Glish used in the Observation definition.

10.1.5 Observation Scheduling Toolkit

The Observation Scheduling Toolkit allows NRAO telescope operators to schedule a telescope from Observation Descriptions. It includes capabilities for normal queue scheduling, and dynamic scheduling.
A telescope operations staff will receive observations descriptions, perhaps as scripts. The staff will then feed these to a queue for subsequent observation. The queue may be dynamically scheduled or fixed order.
The OST is not controllable by an observer but the current queue may be viewed by anyone.

10.1.6 Observing Toolkit

The Observing Toolkit allows an authorized observer to interact with observations on NRAO telescopes in Real-Time from any location on the net. Although the different telescopes may offer some form of this capability, the goal for this toolkit is to make the capabilities and interface the same for all NRAO telescopes.

10.1.7 Archive toolkit

The archive toolkit allows interactions with the archive, including submission, querying, distribution requests, *etc.* A telescope control system may submit data to be archived, or a request for some archived data. A scientist may submit a query to a catalog or a specific distribution request.
A number of NASA organizations have considerable expertise in archiving. In particular, STScI has a system that could be adapted to our use. We are therefore pursuing the possibility of outsourcing the development and installation of the archive to STScI. The physical archives will be located at NRAO sites, but the software would be that developed by STScI.

10.1.8 Pipeline Toolkit

The Pipeline Toolkit allows submission, monitoring, and interaction with a pipeline reduction system in AIPS++. A pipeline may run in a variety of contexts:

- In a user's desktop machine,
- In a user's personal network,
- In a dedicated telescope-based network,
- In a Grid.

The pipeline will run the same software, AIPS++, as that used for interactive data reduction. The pipeline is split conceptually into *mechanism* and *heuristics.* This task describes just the mechanism. The definition of telescope-specific heuristics is split into a separate task.

10.1.9 Pipeline Heuristics

The Pipeline heuristics are rules to help in the processing of observations on the pipeline. They are driven by the data itself, pipeline processing parameters, and meta-data. EVLA specific heuristics will be needed.

10.1.10Calibration Toolkit

The Calibration Toolkit handles calibration information for all NRAO telescopes. The primary goal is to consolidate all information on calibrators, flux and position reference systems, in one logical location, from which other catalogs may be derived as necessary.

10.2 Interfaces to other EVLA sub-systems

[TBD after discussion with other groups, especially monitor and control]

10.3 Construction plans

The e2e project uses resources from the EVLA project and other NRAO activities. For the EVLA work, two positions are to be funded for nine years.
[More details once WBS is complete]

10.4 References