The ALMA Science Archive: Implementation

Juande Santander-Vela, Matthias Bauhofer, Holger Meuss, Felix Stoehr, Alisdair Manning
European Southern Observatory, ALMA Archive Subsystem

The Atacama Large Millimeter Array (ALMA) radio interferometer has started Early Science observations, providing a copious stream of new, high-quality astronomical datasets of unprecedented sensitivity and resolution. We present here how the ALMA Science Archive (ASA) is being implemented, leveraging existing Virtual Observatory (VO) technologies and software packages, together with Web 2.0 techniques, to provide scientists with an easy to use, multi-parameter discovery tool for ALMA data, integrated with the VO.

ASA Architecture

The ASA implementation follows the requirements spelled out by S. Etoka et al. in 2007 (1), and more recently by F. Stoehr et al., (see 2 in this poster exhibition).

The highlights of the implemented architecture (see Figure 1) are:

- a harvesting process to pass data from the ALMA Fronted Archive (AFA) to the ASA tables;
- a simple database structure (see 2 for details) supporting aggregation of data by Fields, Spectral Windows, Projects, and product Dependency for elaborated data products;
- reuse of community-maintained software, such as the openCADC stack (3), and the VOView (4), to support the Table Access Protocol (5);
- and a Spring MVC-based web graphical interface that works as VO client of the web application services;

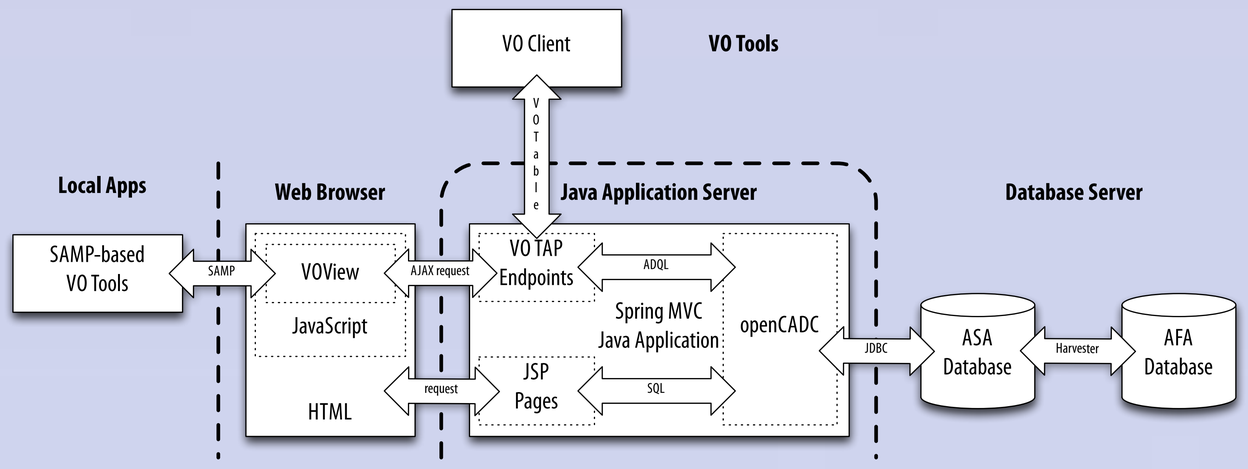

Figure 1: The layered view of our current ASA implementation shows the harvesting process feeding the ASA database from the ALMA Front-end Database, the Spring-based Java application providing GUI and endpoint services, the web-browser with embedded VOView, and different clients of the system.

Query Interface

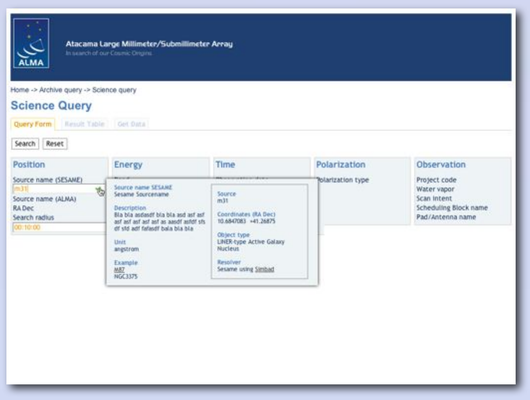

Figure 2: Above, the search interface showing IVOA characterisation-oriented axes, and dynamic name resolution via Sesame using JavaScript. To the right, the search results in VOView.

Figure 2 shows two tabs of the search interface. The initial tab (left) supports searching on Position, Energy, Time, Polarisation, and Observation provenance parameters, and uses JavaScript and AJAX to allow for in-page searching, and automatic coordinates' resolution for object names. The result tab shows data found in the VOView, and thus allows for interactive filtering, searching, column reordering, hiding, and sorting. Search operators such as `!` (NOT), `*` (wildcards) and `..` (numeric ranges) are supported.

We are enhancing the VOView so that it can interact with SAMP-based applications (see 6), allowing interactivity between local desktop applications and our query interface.

Harvesting

Until the ALMA pipeline heuristics are fully defined, only the raw data coming from observations will be available. ASDMs (Viallefond, see 7) are *harvested* from the ALMA Front-end Archive (AFA), where they were created and stored by the ALMA Correlator subsystem. For details on how the data reaches the AFA, see 8.

The most delicate part of the algorithm (see Figure 3), once non-observational data have been rejected, is the *folding*, or dimensionality reduction, of the data. During folding, data from all other ASDM tables are compiled at the Field level.

In particular:

- integration time from Scans is attributed to observed Fields;
- weather and water vapour information is averaged;
- spectral information is aggregated per Field via interval operations, computing characteristic channel numbers, bandwidth, etc.;
- Antennas and Station pairs are created, baseline vectors calculated, and then reprojected to estimate spatial resolution and field of view;
- and other quantities are estimated.

In total, more than 70 parameters are calculated per Field, and available to be both searched and retrieved.

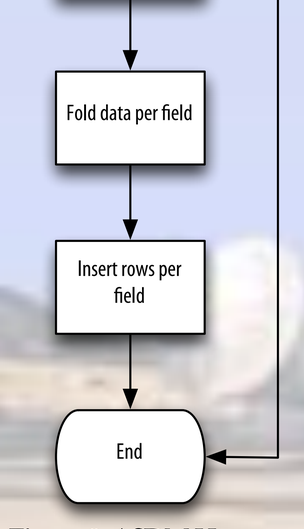

Figure 3: ASDM Harvester

References

1. S. Etoka, G. Fuller, and A. Wicenec, "ALMA Science Archive Requirements," Draft COMP-70.50.00.00- 004-A-GEN, Atacama Large Millimeter Array, April 2007.
2. F. Stoehr, et al., "The ALMA Science Archive: Design", ADASS XXI, **P147**
3. Patrick Dowler et al., openCADC: http://code.google.com/p/opencadc/
4. Dean Hinshaw, VOView: http://code.google.com/p/voview/
5. P. Dowler, G. Rixon, and D. Tody, "Table Access Protocol," IVOA Recommendation 1.0, March 2010.
6. M. B. Taylor, T. Boch, M. Fitzpatrick, A. Allan, L. Paioro, J. Taylor, and D. Tody, "Simple Application Messaging Protocol," IVOA Recommendation 1.11,, April 2009.
7. F. Viallefond, "The Alma Science Data Model," ADASS XV, vol. 351, ASP Conference Series, 2006
8. A. Manning, et al., "ALMA Archive: Data to the Scientists", ADASS XXI, **P089**