

ALMA Pipeline Status Update  
Sept 4, 2013  
Prepared for ANASAC Face to Face

The ALMA pipeline has progressed very rapidly over the past year. The ALMA Regional Centers have developed procedures to perform parallel reductions of Cycle 1 data, comparing the results of the human pipeline (cycle-0 data reduction style) and the automated pipeline. This period will serve the dual role of improving pipeline performance and developing confidence in the automated reductions. Once sufficient confidence in the pipeline calibration has been attained, parallel reduction will cease and the pipeline will be the primary data reduction path for remaining Cycle-1 data.

A strong “Tiger” team, led by Todd Hunter at the NAASC, working with the pipeline team has succeeded in translating the expertise gained from Cycle 0 data reduction into robust, automated heuristics. We have anticipated many of the changes expected in Cycle 1 data, but testing has been limited by the dearth of available data. For Cycle 1, we anticipate that there will be some required user intervention in the operation of the pipeline, particularly in providing updated flux estimates for the flux calibration targets, although extraordinary flagging may also occasionally require manual intervention. Both of these are accommodated naturally as part of the pipeline and accounted for in parallel reduction procedures.

The single dish pipeline has been tested at NAOJ and has been reported to be ready for use. The pipeline team has modified the single dish pipeline to more closely follow the structure of the interferometric pipeline for a common look and feel. Substantial improvement has been made in the performance of the single dish pipeline. Cycle 1 is the first observing cycle offering users single dish capabilities, and consequently there is still significant discussion about best practices for data acquisition and reduction. The pipeline team will need to track these developments.

For Cycle 1, the pipeline will produce a complete set of flags, appropriate calibration tables, reference images of the phase and bandpass calibrators for diagnostic use, and a WebLog summarizing the execution of the pipeline. A simple script to restore a measurement set to the fully flagged and calibrated state will be available. The WebLog is a self-contained set of HTML files which summarizes the pipeline execution sequence and allows exploration of the various steps. The weblog design was reviewed by a group of pipeline experts (both ALMA and VLA) in May of this year. We anticipate a repeat of this review early Q1 2014 based on the feedback from Cycle 1, and will solicit community feedback through a broader-scope user test prior to the beginning of Cycle-2.

Scientific support for the pipeline development team now follows the same pattern as other subsystems. Liz Humphreys has been named as the subsystem scientist and heads a working group consisting of: Eric Villard (Deputy), Remy Indebetouw, and Hiroko Shinnaga. In addition, Todd Hunter has an *ex officio* position on this

team in his capacity as the lead of the Tiger Team. There is a larger group of testers working on testing the pipeline in various capacities scattered worldwide, with the Data Content Managers at the JAO and the Data Analysts at the NAASC forming the majority of the personnel. The VLA pipeline has recently been adjusted to make use of the pipeline framework developed by the ALMA team, and is currently undergoing early testing. The pipelines are now sharing infrastructure and some heuristics. Both sets of heuristics are now stored with the CASA codebase and we are developing the procedures to allow the pipeline to be fully delivered as part of the standard CASA download.

Based on testing with commissioning and testing data, we expect that the interferometric pipeline is ready to support the calibration of most Cycle 1 observations (Band 9 observations will likely still require human intervention). The next six months of the pipeline effort are focused primarily on responding to Cycle 1 operations (some infrastructure improvements, archive ingest of pipeline products, and WebLog improvements) and preparing for ALMA Cycle 2. This latter work includes the addition of polarization calibration and support for the more sophisticated spectral window setups that will be made available to Cycle 2 proposers. We will also begin introducing science target imaging to the pipeline; an imaging tiger team is scheduled to begin work later this month. Another issue confronting the pipeline for Cycle 2 is the definition of a robust set of QA2 parameters by which the success of pipeline executions can be judged. The pipeline working group is preparing recommendations for the team to implement, but it is likely that a significant amount of iteration will be required before these metrics are considered reliable.