STATUS OF ALMA PIPELINE

Formally: APHS (ALMA Pipeline Heuristics System):
Initiated in 2004

See:
   http://almasw.hq.eso.org/almasw/bin/view/PIPELINE/PipelineHeuristics
   http://almasw.hq.eso.org/almasw/bin/view/PIPELINE/DesignReviews

Goals:
   Help provide good ALMA standard product, timeline by Prop call (late 2012?)
   End-to-end processing, Capture expert knowledge
   Use and improved by ARCS, transfer parts to non-experts, to casapy
   Begin limited use in ES in Nov 2011?

Effort:
   Yearly design review, workshops, tests
   First, VLA, PdB data. Now using Science Verification data

   BUT

PRESENT CAPABILITES and USE

Attempts to reduce SV data suggests:

Too slow for SV and ES reductions (general problem, esp 3840 ch)
Some very inefficient algorithms
Strange method of calibration and editing---use 1 ch data
All or nothing approach, no interaction
Use of casa tools and python and native html displays. No tasks
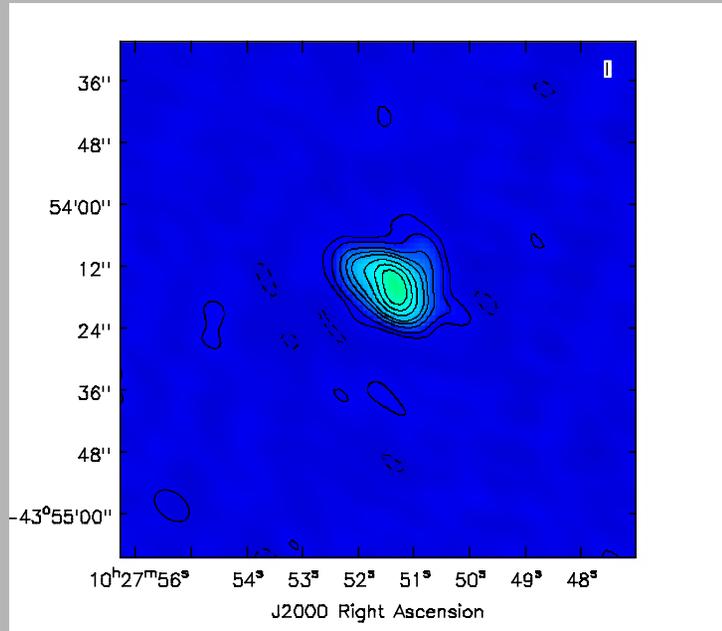Not modular enough

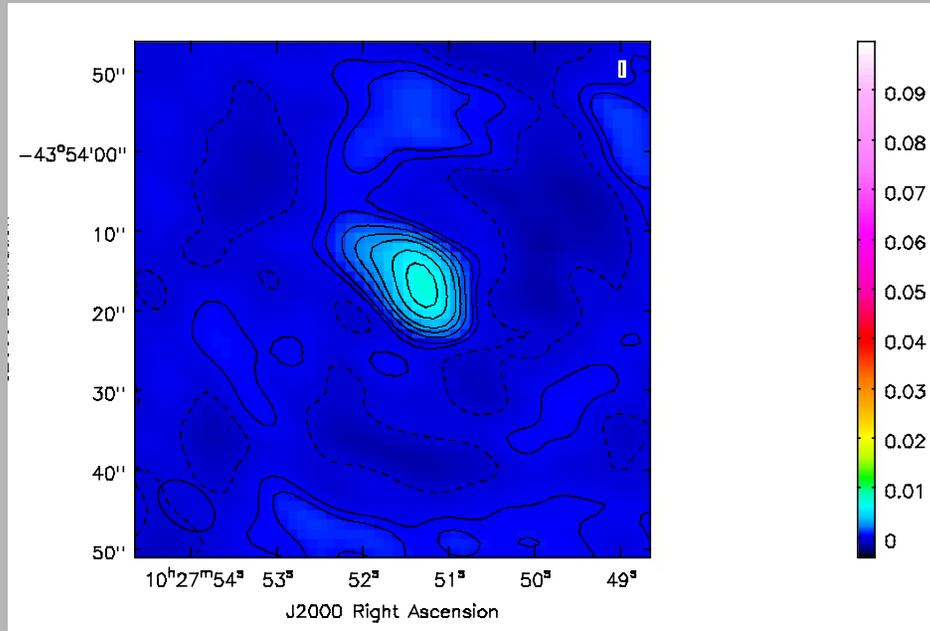Not yet ready – Restructuring of project needed

Did use for ngc3256 data set
   (demonstration of results at end on request)

NGC3295 Comparison

SV reductions   Pipeline Reductions

Continuum image of spw 2 (no line)
~3 hours on this laptop to get to clean images (cont+sl)
 only found strong line in spw 0
Main problem: Only knows how to flag outliers
   No heuristics for typical ALMA corruptions.

MODIFIED PIPELINE EFFORT

Organizational Changes
   Combined with Casa group and EVLA pipeline effort (last month)
   Jeff Kern as Head of this effort
   External group; Pipeline Algorithm and Users Group (PAUG)
      to aid in heuristics and testing (Ed F as chair-homonid, Baltasar, Chris Wilson)

Goals over the next six months
   --**Reorganize and Redesign present Pipeline code/methods to reach goals**
      Are the results useable and trusted?
      Understood and modifiable by ARCS and PI's
      Modular so that parts can be redone and exported
      As sympatico with casapy as possible
      Clear reporting of results (QA2, later)

 Time-Frame
   ~Nov 2011: Improved infrastructure coming well-along
   ~Dec 2011: Beginning to help with ES reductions
   ~Feb 2012: Develop sophisticated heuristics
               egs: single-dish + 7m + 12m software, polarization

Suggested Overall Reduction Steps

1. ASDM to MS with required apriori calibrations, editing, eg, task importalma
***1A. Some input information about the scientific goals. (PdB)
2. Flag any 'gross' outliers (data and weights) in all sources (easy for ALMA)
3. Determine bandpass (with necessary editing) and apply to all
4. Determine gain calibration (with necessary editing) + Abs flux Calibration
5. Apply calibration and general checking of integrity
   Cal Data: More edits? Interference? May have to repeat some of above
   Cal Images: Each spw, general statistics, problems?

6. Check target data (outliers?) and determine line channels
7. Make clean images of continuum and sp cube channels
8. Self-calibration possible? What data to use, further editing? Apply?
9. Remake images, subtract continuum
   Reasonable stopping point for pipeline?

*** Example of useful input to guide reduction steps:
   a. Continuum only and expected dynamic range?
   b. Expected Continuum/line ratio? -> time variable BP?
   c. Emission angular size?
   d. What else?

Possible Stages for each Step

Use Reduction step 3. Bandpass as an example)

1. Decide on heuristics and generate casapy code.
   What data to use? bandpass, flux cal, phase cal?
   One bandpass or time variable? Flag edge channels

2. Execute the code in recent casapy-stable as much as possible?

3. Review results:
   'messages' during execution
   analysis of antenna-based phase cal and bandpass cal
      flagging needed in time/freq for outliers or low sensitivity
      Now done by plotms and plotcal.
   HARD TO DO: Develop sophisticated heuristic
   redo step if needed

4. Report results
   Add basic plots statistics to QA2 for this step
   Similar report as from EVLA pipeline, hand-reductions

Pipeline Summary

Pipeline infra-structure is being modified
Make more congenial to casapy
Make more modular
Reflect current reduction practices
Interaction with User/algorithm group. ARC's, JAO
   to test and improve pipeline.

Time-Frame (probably too optimistic)
   ~Nov 2011: Improved infrastructure coming well-along
   ~Dec 2011: Beginning to help with ES reductions
   ~Feb 2012: Develop sophisticated heuristics
               egs: single-dish + 7m + 12m software, polarization