Science Ready Data Products and the ngVLA

Jeffrey S. Kern
National Radio Astronomy Observatory
Charlottesville VA 22903

Abstract

With angular resolution comparable to ALMA and future ELTs, the ngVLA will be the centimeter wavelength instrument contributing to multi-wavelength astronomy throughout the next decades. To maximize the impact of the ngVLA it is important that it be a facility available to all astronomers, not only to those who consider themselves radio astronomers. Building on the successes of the ALMA and VLA pipelines NRAO has undertaken a multi-year project to provide science ready products to the users of our telescopes and archive. The SRDP initiative addresses one of the technical risk areas of the ngVLA, and demonstrates the Observatory's ability to deliver this critical functionality.

The SRDP project is currently refining and reviewing the technical and scientific requirements for the existing NRAO telescopes. Capabilities will be made available to the community as they are developed and mature. As the project progresses, a rich archive of science quality radio images from both ALMA and the VLA will be provided to the astronomical community. The current status, concepts, and plans for the SRDP project are presented, as are the implications for the ngVLA facility.

ngVLA Data Processing

The next generation VLA will enable science at unprecedented resolution and sensitivity throughout the centimeter to millimeter regime. Although the science is revolutionary, the technology is merely an update to current state of the art electronics and digital processing capabilities. Table 1 (below) summarizes several of the key science use case requirements and the implied post-correlator data rates.

In defining the requirements for the SRDP project we are guided by the following set of use cases. These are still under development and may change as we refine and review the requirements. The table at bottom right summarized current status of the primary use cases by telescope.

SRDP Precursors at NRAO

The VLA Sky Survey is a three-epoch survey of the sky visible from the VLA. Totalling nearly 40,000 square degrees this 5-Band (2-4 GHz) survey will achieve 120 Jy RMS in each epoch, and a cumulative sensitivity of 60 mJy. When complete this high resolution (1.5 arcsec) survey is expected to have detected over 5 million radio sources. The survey also represents the first science ready data products automatically generated for the Jansky VLA. Using a slightly modified version of the standard VLA calibration pipeline, each VLA survey is automatically calibrated upon completion of observation. After a quality assurance step, the calibration products are stored and the imaging processing is begun.

The SRDP Project is currently seeking a full time Project Scientist to lead the heuristics development and validation team. Details and application instructions are available at: http://jobs.sdvsrle.com/nrao/jobs/okskyfem

The ALMA pipeline has been used throughout Cycle 4 for the automated calibration and imaging of ALMA data. Over 95% of Cycle 4 standard mode observations were successfully calibrated by the pipeline. Approximately two thirds with no human intervention. NRAO and the ALMA project continue to refine the heuristics in order to decrease the fraction of observations that require human intervention.

The imaging pipeline has been used throughout Cycle 4 to produce reference images, suitable for establishing the quality of the observations. Improving the quality of the images, and providing interfaces to allow user specified re-imaging are high priority deliverables from the NRAO SRDP Project.

The Pipeline Weblog, initially developed for the ALMA pipeline, is a hierarchical representation of the pipeline processing that provides both a summary view for the novice, and a detailed view for the expert. The SRDP project will expand the use of the weblog in the archive.

SRDP Data Products

<table>
<thead>
<tr>
<th>Use Case</th>
<th>ALMA</th>
<th>VLA</th>
<th>VLASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Calibration</td>
<td>Science Ready</td>
<td>Reference</td>
<td>Science Ready</td>
</tr>
<tr>
<td>Restore / Recalibration</td>
<td>None</td>
<td>None</td>
<td>N/A</td>
</tr>
<tr>
<td>Standard Imaging</td>
<td>Reference</td>
<td>None</td>
<td>Quick Look: Science Ready Single Epoch: Preliminary</td>
</tr>
<tr>
<td>Optimized Imaging</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Large Project</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

Table 1: Key parameters and the associated data rates for selected Key Science Goals. Although these data rates are cost prohibitive in the current epoch ($5100/yr) per PB historical trends in storage predict that the cost per petabyte will be ($1M) per year but tractable.