Software Development at NRAO for Science Support

Nicole Radziwill, AD End to End Operations
Observatory Software Goals

• Deliver reliable, maintainable software systems for telescope operations, reusing concepts, designs and code wherever possible

• Broaden access to our instruments throughout the user community and provide users of NRAO research facilities with a consistent experience

• Achieve efficiency and effectiveness in software operations

Cooperation across NRAO occurs to meet these goals:

• First goal is primary responsibility of telescope software groups
• Second is primary responsibility of End to End Operations
• Third is everyone’s responsibility
Separation of Concerns

- **Telescope Computing Groups (e.g. EVLA Computing)** – getting the instrument to function well and be useful in an observing context
  - Monitor & Control
  - Observing
  - Scheduling

- **End to End Operations (E2E)** – providing a consistent experience to the end user, integrating work with other NRAO telescopes
  - Portal for online access & communication with scientists
  - Proposal Submission
  - Data Analysis/Pipelines (including coordination of algorithm R&D)
  - Operations Analytics
  - Archive Access & Infrastructure
  - Science information on main NRAO web
External Focus

• E2E role is also differentiated from telescope software groups by its external focus, which includes:
  – Coordination with NRAO Education & Public Outreach (science web, presentation of our capabilities as One Observatory)
  – Coordination with NRAO Computing & Information Services (for example, transitioning HPC facilities from R&D to operations)
  – Coordination with NRAO Science & Academic Affairs on policy issues (e.g. data archive and proposal policies)
  – Coordination with NRAO HR (broadening access to diverse groups)
  – Establishment/maintenance of external relationships (Pittsburgh Supercomputer Center, NCSA)
  – Pursuit of external, non-programmatic funding
High-Performance Computing (HPC)

• Making better use of HPC (e.g. clusters, GPUs, FPGAs) is critical to meeting the needs for data handling and data processing, especially for EVLA
  • We are working to ensure that NRAO software engineering staff can develop and expand their HPC skills
    – More involvement and participation in national HPC community
    – Need to leverage national facilities (NCSA, PSC, NMCAC etc)
  • Algorithm R&D proposal submitted to NSF in July
    – ½ FTE at supercomputer center dedicated to NRAO for 5 yrs
    – Access to 2 postdoc and 2 grad students each year for support
Major Subsystems

- High Level Architecture (HLA)
- Portal (user authentication/gateway)
- Proposal preparation and handling (PST/PHT)
- Observation preparation (OPT)
- Array Scheduling (OST)
- Archive access & infrastructure (AAT)
- *Activities & current status of Data Analysis to follow*
- *Algorithm R&D technical details discussed later*
E2E Contribution

- High Level Architecture (HLA)
- **Portal (user authentication/gateway)**
- **Proposal preparation and handling (PST/PHT)**
- Observation preparation (OPT)
- Array Scheduling (OST)
- **Archive access & infrastructure (AAT)**
- **Data Analysis & Algorithm Development**
  - *Primary EVLA subsystems will be discussed by B. Butler*
  - *E2E subsystems will be discussed by N. Radziwill*
EVLA Project Organization

EVLA Project Organization

Numbers refer to WBS level 2 tasks.
# E2E Resources

**Nicole Radziwill**, AD End to End Operations  
**Ed Fomalont**, Project Scientist, End to End Operations

## Algorithm Research & Development
- Sanjay Bhatnagar
- Kumar Golap
- George Moellenbrock
- Steve Myers (25%)
- Ed Fomalont
- Urvashi Rao-Venkata
- Abittur
- Juan Uson, Bill Cotton, Rob Reid, Brian Mason, Scott Ransom

## Data Management & High Performance Computing
- Gareth Hunt
- Darrell Schiebel
- Boyd Waters
- Wes Young
- John Benson
- Ron DuPlain
- **Joe Brandt, Amy Shelton** (25%)
- Eric Sessoms
- Libby Miller
- Pat Murphy, Stephan Witz (50% each in webmaster position)
- Tony Remijan

## Observatory-Wide Science Support
- McMullin/Replacement
- David King  
  Laura Glendenning  
- **Dana Balser** (50%)
- Jared Crossley (75%)
- Ashish Arte and Paul Schock at Open Sky Software in Austin, TX working on Interactive Services/PST/etc.  
- Darren Hoyt at Category 4 Solutions in Charlottesville, VA working on web and design issues

---

**Individuals in boldface are full-time NRAO employees; others are fixed-term or contract employees.**

*Individuals in italics participate in these activities in their research time and are not committed to specific deliverables.*
E2E Contributions
Interactive Services Portal  
(http://my.nrao.edu)

- Initial version of portal developed by EVLA Computing  
- Used in a number of proposal calls  
- Being used for development of Observation Preparation Tool, Observation Scheduling Tool, Operator’s Screen  
- Turned over to E2E/OpenSky at June 2007 proposal deadline; upgraded for October 1 proposal deadline, after which EVLA applications can be retrofit to updated portal
Portal Dashboard

The Dashboard is a management tool to help you keep track of news from the NRAO telescopes, latest developments, your data, your collaborators, and the status of dynamic scheduling.

You will also be able to keep track of the various requests for feedback that NRAO has open at any given time. This utility will be continually evolved over the next one to two years, and new items will be added based on feedback from the user community, so please be sure to tell us if there is something you would like to see on the Dashboard that would make working with NRAO easier.
Proposal Preparation (PST) & Handling (PHT)

• Originally developed by EVLA Computing and used in a number of proposal calls for VLA and GBT
• Turned over to E2E/OpenSky for June 2007 proposal deadline; upgraded for Oct 2007
• **Must still get EVLA hardware definitions supported (notably WIDAR)**
• Full proposal handling system & VLBA integration planned for February 2008 proposal deadline
• Results from dynamic scheduling required to transition from instrument-centric view (sources, resources, sessions) to science objectives view (ALMA approach)
Proposal Preparation & Handling

ALMA Tree Structure applied and used for all NRAO telescopes; First step in integration
Archive access (AAT)

- Searches and retrieves data from the disk archive
- Raw data and processed products made available
- Will support Virtual Observatory (VO) access on the processed products
- EVLA will have common science data model and already has archive storage software and hardware (NGAS) in common with ALMA, so archive access tool can and will be common
AAT - VLA/VLBA Prototype

• VLA and VLBA raw data currently accessible via web application, available and continually refined since October 2003
• Non-proprietary data openly available
• Proprietary data made available via portal login or staff-generated “key”
• Both simple (Project ID, for example), and complex (akin to VO cone search) searches supported
• More data downloaded via this mechanism than is taken real-time at the VLA (~3 GB/day)
• Screen shots of recent work to follow
### Image Search Results

**Show Query Parameters**

Displaying 5 Rows out of 5 Found

<table>
<thead>
<tr>
<th>Basic Info</th>
<th>UV Info</th>
<th>Image</th>
<th>Image Info</th>
<th>Links</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field: J005652.2+694046 (0000 image center) 4.860 GHz [C-band] Sky offset: 0.0 arcmin</td>
<td>Telescope: VLA Collection: VLAPipeline Config: B Polarization: FULL Observed: 1999-10-31</td>
<td><img src="image1" alt="" /></td>
<td>Strikes: 1 Beam: 4.390 arcsec for R = 0550.97 mas = 0.134 mJy</td>
<td></td>
</tr>
<tr>
<td>Field: J005551.7+694046 (0000 image center) 4.860 GHz [C-band] Sky offset: 0.0 arcmin</td>
<td>Telescope: VLA Collection: VLAPipeline Config: B Polarization: FULL Observed: 1999-12-04</td>
<td><img src="image2" alt="" /></td>
<td>Strikes: 1 Beam: 1.3136 arcsec for R = 0916.37 mas = 0.038 mJy</td>
<td></td>
</tr>
<tr>
<td>Field: J005655.7+694016</td>
<td>Telescope: VLA</td>
<td><img src="image3" alt="" /></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Extra Info**

- Project: A193388 Observer: Itso, Paul Source: M82 (11 observations at the VLA)
- Project: A293989 Observer: Stumpe, Richard A. Source: M82 (39 observations at the VLA)
Archive Progress & Plans

- Updated archive access tools for VLA/GBT/12m/140ft are being made available in Fall 2007
- Access to observing scripts and logs made available for VLA Fall 2007 and GBT Spring 2008
- E2E providing support for archive development required to support WIDAR in Summer 2008
- Interface between shared and non-shared EVLA subsystems is through archive
Data Analysis with CASA

- Exciting year, with the beta release coming up next month, and phase-in of user support for the beta test group
- “Train the trainer” and tutorials planned this winter and spring
- Details about the package and its current capabilities to be provided by Joe McMullin
Algorithm R&D

KEY ISSUES
- Processing is inefficient at present, 10-50x speedup beneficial by itself
- Larger data volumes will require faster processing
- More sophisticated algorithms required to achieve higher dynamic ranges

• 2007.8 to 2008.4 (before WIDAR)
  - CASA betas, ramping up user support, infrastructure support, continued EVLA debugging in AIPS
  - Continued algorithm development of stationary (parallactic angle) primary beam effects
  - Begin HPC investigations with November 2007 workshop; substantial progress can be made here

• 2008.4 to 2010 (debugging WIDAR, first science)
  - WIDAR correlator tests: AIPS → CASA transition to debugging, checking
  - Commissioning basic correlator modes with ‘simple’ science
  - Aggressively pursue wide-field/wide-bandwidth imaging, multi-line spectral imaging
  - HPC - I/O strategies determined and beginning implementation

• 2010 to 2013 (evolving into mature EVLA capabilities)
  - Shared-risk science support → algorithm development motivator
  - HPC and I/O further improvements to handle harder EVLA cases
  - Progress on 10^5 DR, wide-field, wide-bandwidth imaging, full spectral line support.
Backup
Algorithm R&D/HPC
Issues for EVLA

<table>
<thead>
<tr>
<th></th>
<th>Low Frequency</th>
<th>High Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Cases</strong></td>
<td>• Pointing Self-Cal</td>
<td>Ionosphere not an issue</td>
</tr>
<tr>
<td></td>
<td>• Non-isoplanatic calibration issues (remove ionospheric/tropospheric effects that limit dynamic range)</td>
<td>• Removing tropospheric effects</td>
</tr>
<tr>
<td></td>
<td>• RFI mitigation (detection, flagging, etc.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Continuum Imaging</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Wide-band high fidelity imaging</td>
<td>• Data volume/processing times</td>
</tr>
<tr>
<td></td>
<td><strong>Continuum Imaging (Full Polarization)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Wide-band high fidelity imaging</td>
<td>• Data volume/processing (4 data streams as compared to single polarization continuum imaging)</td>
</tr>
<tr>
<td></td>
<td>• Wide-field full polarization imaging (many options; primary need is to make the processing more efficient)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Wide-band wide-field full polarization imaging</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Spectral Line Imaging</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Data volume/processing</td>
<td>• Data volume/processing (more spectral lines, more channels)</td>
</tr>
<tr>
<td></td>
<td><strong>Spectral Line Imaging (Full Polarization)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Data volume/processing (4 data streams as compared to single polarization continuum imaging)</td>
<td>• Data volume/processing (more spectral lines, more channels, 4 data streams)</td>
</tr>
<tr>
<td><strong>All Cases</strong> (systematic improvement of imaging process)</td>
<td>• Multi-scale deconvolution/MEM</td>
<td></td>
</tr>
</tbody>
</table>

**Color Key:**
- Green – problem is well in hand, many options available
- Blue – problem more complex, but solvable with time
- Orange – problem much more complex, must focus resources
- Red – most complex problems, dependent on other solns