NRAO Support for Algorithm Development

Ed Fomalont, Project Scientist

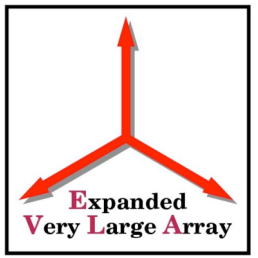
Nicole Radziwill, AD

NRAO End to End Operations

Observatory-Wide Support

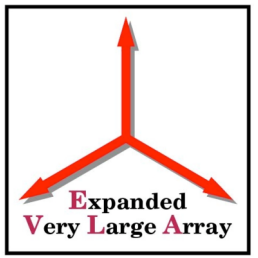
- EVLA imaging will have unprecedented dynamic range, sensitivity and spectral resolution
  - Such images are not obtained ‘automatically’. Imaging based on VLA exp
  - Need high performance computing, algorithms, convenient user-systems
- Myers presentation gave highlights of current and anticipated algorithm effort.
  - NRAO has a lot of talented people available to achieve the EVLA goals
  - Recent effort concentrated on making software packages robust
  - A lack of observatory-wide coordination in meeting these goals
  - Insufficient documentation, publicizing of NRAO algorithm effort
- E2E operations has the charge to insure these high quality EVLA images by: increasing observatory-wide support and coordination, funding additional resources, and filling gaps in overall effort.

E2E Goals for Algorithm R&D



- Promote **vitality in communications and collaborations** among scientific and technical staff
- **Separate efforts** of algorithm R&D and production software implementation to promote wide involvement in problem solving
- Develop our **core capabilities** in the related areas of algorithm R&D for science and High Power Computing (HPC)
- **Invest** in R&D equipment that can be used by a wide audience throughout NRAO and outside, especially for hard EVLA problems
- **Partner** with others in astronomy, mathematicians, computer scientists, scientific computing specialists, image processing specialists, and students more frequently and effectively
- **Improve our internal organization** of these efforts to better support all of the above

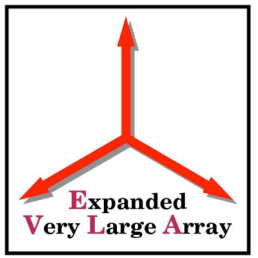
E2E Organization in Support of Algorithm R&D



Ed Fomalont – coordinating strategic scientific efforts across Observatory

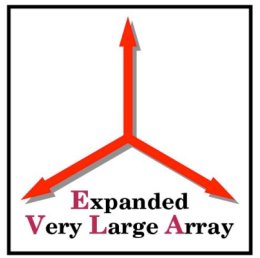
- Steve Myers – technical coordination for interferometric and "hybrid" efforts
- Crystal Brogan - ALMA; Michael Rupen - EVLA
- Cotton, Greisen, Owen, Perley, Uson, vanMoorsel, Whysong
- **Bhatnager, Golap, Moellenbrock, Urvashi, King**
- Dana Balser – technical coordination for single dish efforts
- Scott Ransom – technical coordination for pulsar work
- Ron DuPlain – technical coordination for HPC work
- Nicole Radziwill – strategic alliances with HPC organizations, external funding

Present E2E Efforts



- **Stronger CV-AOC Link:** CASA group to CV, CV developers to AOC
- **Algorithm R&D Week** workshops:
  - AstroGPU (November 9-10, 2007) - Princeton
  - Supercomputing 07 (November 11-14, 2007) - Reno
  - EVLA algorithm workshop (November 15-17, 2007) - Socorro
  - 2008 workshop plans: mosaicing, GBT feed array, HPC
- **Free up CASA developers' time**: Bhatnagar, Golap, Moellenbrock
  - Additional FTE support for user-oriented issues.
- Aggressively pursuing **external funding** & collaborations (e.g. NSF OCI petascale applications with Pittsburgh Supercomputer Center
- **Establish a memo series** to better communicate issues, progress and results; promote publishing journals (e.g. A&A, IEEE, SIAM)

Present E2E Investment

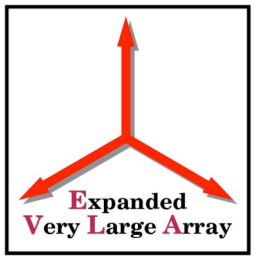


- **Algorithm R&D support will be ~$180K/yr**
  - From NRAO E2E Operations, not EVLA construction/contingency
  - $100K/yr: Hire support staff (1.5 FTE) to free developer time for R&D from core group of people at NRAO. Next week 0.5 FTE arrives
  - Support workshops, collaborations, travel (~$50K/yr)
  - Equipment investments (~$30K/yr)
- **External Grant Proposals**
  - Proposal submitted to NSF OCI (Petascale Applications) July 2007

    $2M/5 yr for 0.3 supercomputing expert P/T at PSC, 2 graduate students,

    1+ postdoc through UVA Computational Science Center. Decision Feb 2008
- **Collaboration and visits from other institutions:**
  - ALMA, LWA, LOFAR, ATA, eMERLIN, SKA, VLBA: colloquia, visits for one week to several months.

Algorithm Memos



All memos will be collected soon in the Algorithm R&D Memo Series

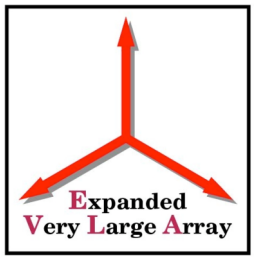
Casa-based investigations (http://www.aoc.nrao.edu/evla/memolist.shtml)

- EVLA Memo 62: Polarization effects in imaging
- EVLA Memo 67: W-projection wide-field imaging
- EVLA Memo 84, 100: Pointing self-calibration
- EVLA Memo 101: Multi-frequency synthesis

Obit-based (http://www.cv.nrao.edu/~bcotton/Obit.html)

- Image pixelization and dynamic range
- Automatic CLEAN windowing
- Beam squint and Stokes V on the VLA
- Several ionospheric-effect papers
- http://www.atn.csiro.au/projects/sky/Memoseries.html (SKA)
- http://www.aoc.nrao.edu/aips/aipsmemo.html (AIPS)

Overview of EVLA Needs (1)



KEY ISSUES

- Processing is inefficient at present
- Larger data volumes will require faster processing
- More sophisticated algorithms required to achieve higher dynamic ranges

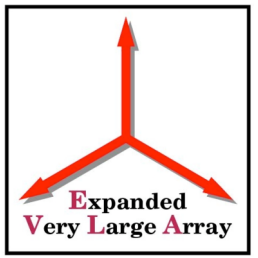
Present inefficient processing

- Can barely keep up with harder VLA problems
- Speed-up of factor of 10 to 50 would help significantly with no other advance
- Clusters/multi-processing, develop internal expertise, get external guidance

Large data volumes

- EVLA volume (25 Mbyte/s in 2011) from 10 to 100 times that of VLA
- Fast efficient I/O, data archiving and format, central data access
- Convenient platform for users of big data sets

Overview of EVLA Needs (2)



More sophisticated algorithms

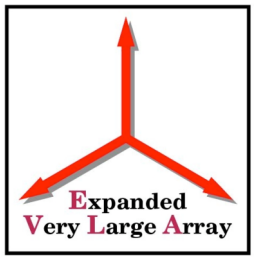
Existing algorithms, techniques OK to about 1000:1 dynamic range (DR)

Increase efficiency of algorithms in cluster/multi-procesing environments

Develop algorithms to go from 10^3 to 10^5 DR (~easiest to hardest, but linked)

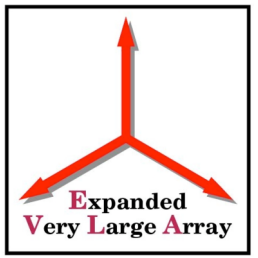
- Incorporate known and parallactic primary beam effects
- Recognize and remove low level interference (gross interference easier to remove)
- Extended source deconvolution techniques
- Wide-field imaging:
- slowly changing pointing errors (algorithm or reference pointing)
- Non-isoplanicity (>1.0 GHz minor, <1.0 GHz major proble, but not key EVLA science)
- Wide-bandwidth imaging: hybrid to full bandwidth implementation
- Band-pass closure, time variations, continuum subtraction, strong line contamination
- Unexpected limitations (eg. baseline closure errors, system non-linearities, OTF map)
- Polarization changes over band-pass/beam, understand corruption of I

R&D Timescales



- **2007.8 to 2008.4 (before WIDAR)**
  - CASA beta releases, ramping up user support, CASA infrastructure support
  - Continued EVLA debugging using AIPS with a lot of hard, investigatory work
  - Continued algorithm development of stationary (parallactic angle) primary beam effects
  - Begin HPC investigations with November 2007 workshop; substantial progress can be made
  - Complete SDM filler into CASA from WIDAR correlator
- **2008.4 to 2010 (debugging WIDAR, first science)**
  - WIDAR correlator tests: AIPS → CASA transition to debugging, checking
  - Commissioning basic correlator modes with good science, but easier problems
  - High frequency continuum, multi-line spectral imaging - medium-hard problems
  - Testing and software for special correlator modes
  - HPC implementation; I/O strategy set and beginning implementation
  - Algorithms must be easily available for astronomical use
- **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.

Future Considerations



- **Possible software trade-offs:**
  - Limited resources are available; must prioritize the software development for maximum science output and EVLA output volume and modes
  - Will reevaluate resources and goals every six months (in November and May) to ensure that we are on track
- **Computing Efficiency (CE) and I/O speed:**
  - Good algorithmic development will depend on CE and I/O. Cluster/parallel processing investigation and development crucial
  - Evolve into processing platform for EVLA users, especially hard problems
- **How much automated processing should a user rely on?**
  - Good pipeline for initial calibration, editing, band-pass. Can probably use ALMA infrastructure. Pipelining of hard projects difficult.