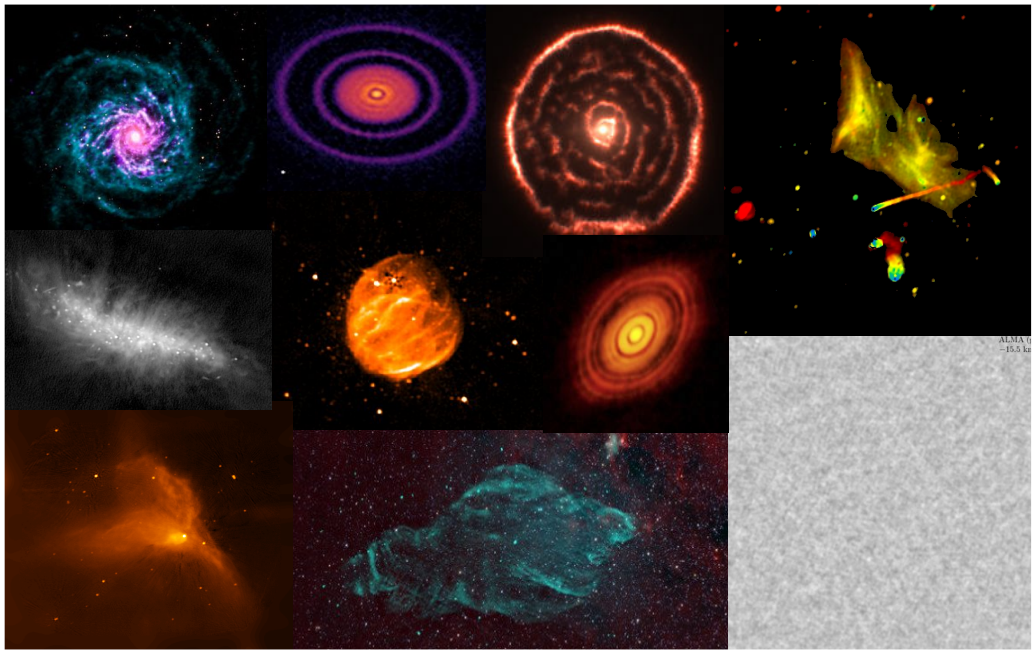
CASA and its evolution

CASA
Common Astronomy Software Applications



Urvashi Rau
National Radio Astronomy Observatory, Socorro, NM, USA

NM Update Meeting (14 October 2021)

CASA : Common Astronomy Software Applications

Data Analysis Tools for VLA / ALMA : casadocs.readthedocs.io/en/stable , casa.nrao.edu (1-3 releases / yr)

– **Data Format** : Measurement Sets

– **Data manipulation** : Import/Export, Meta-data summary, Selections, partitioning, averaging, etc.

– **Visualization** : Data plotting, Image viewing, Interactive GUIs

– **Analysis** : General purpose statistics and transformations for visibility data and images

– **Simulation** : Construct realistic datasets using sky and instrument models

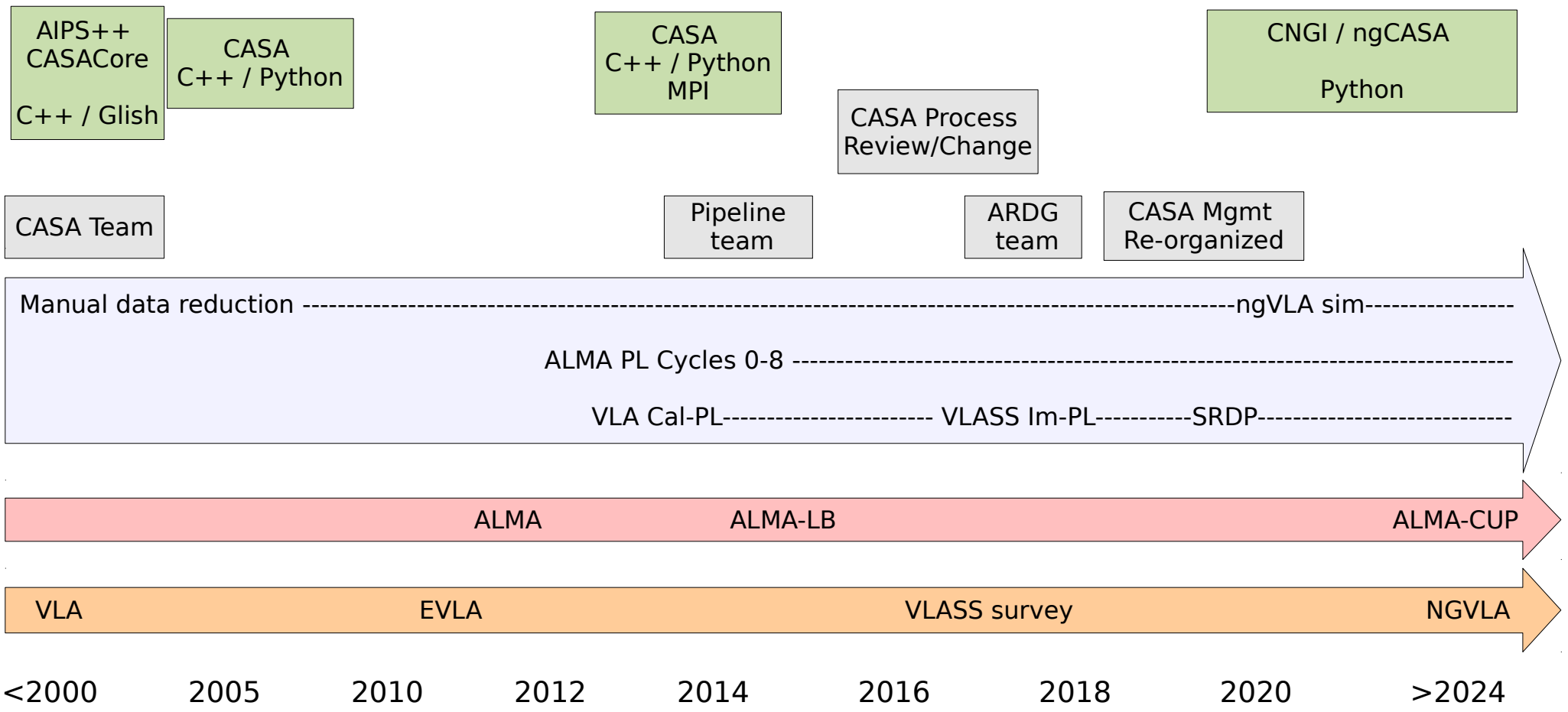
– **Flagging** : Identifying and discarding data corrupted by RFI (or other reasons)
  – Manual flags, meta-data based flags, and automatic outlier detectors

– **Calibration** : Estimate and undo direction-independent instrumental effects
  – Solve for antenna gains and bandpasses, delay, pol leakage, etc... and apply to data

– **Imaging** : Reconstruct images using principles of interferometry + undo other instrumental effects
  – Continuum images, Spectral line cubes, Mosaics, Wide Band and Wide Field corrections

New feature/algorithm development, bugs and their fixes, performance improvements..... happen continuously

CASA @ NRAO : Evolution



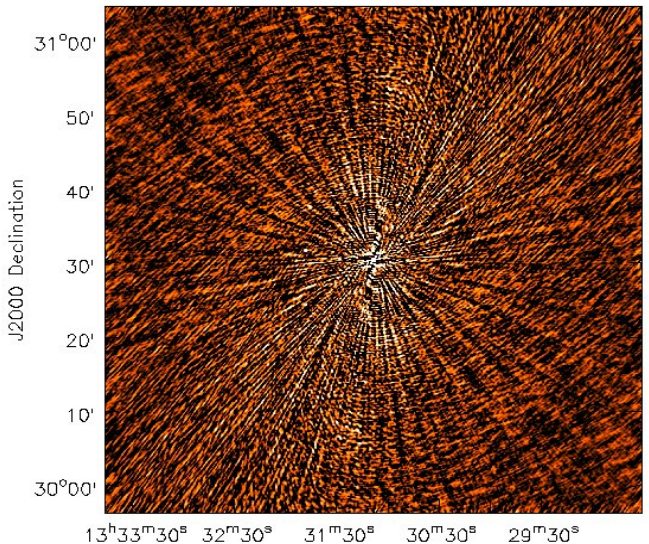
Creating algorithms to maximize the instrument's potential

Data : VLA L-Band data on the standard calibrator 3C286

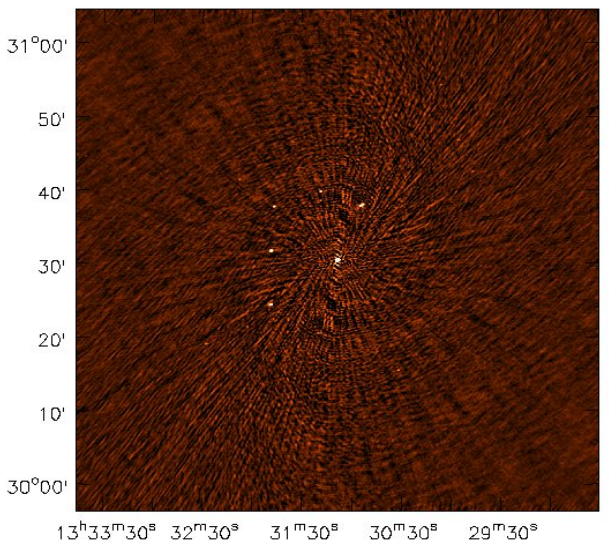
VLA → EVLA : Narrow-band to Wide-band

=> Sky and Instrument vary with frequency

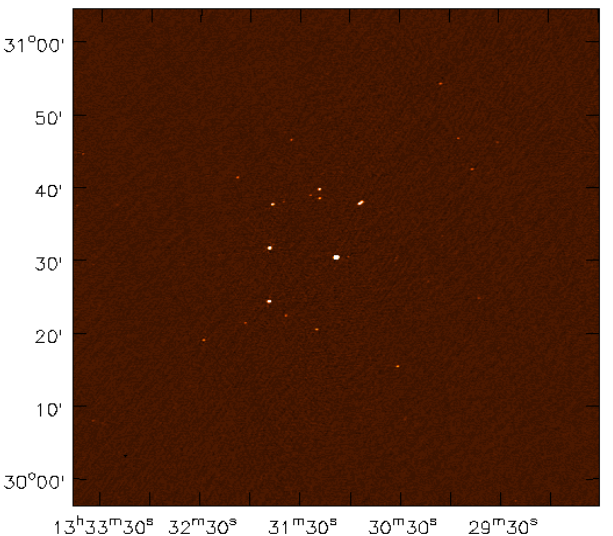
=> Needed new algorithms to achieve the full wideband sensitivity offered by the telescope



Dyn range = 500



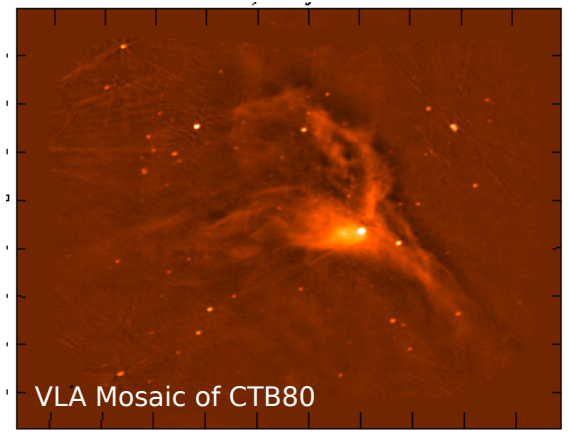
Dyn range = 10K

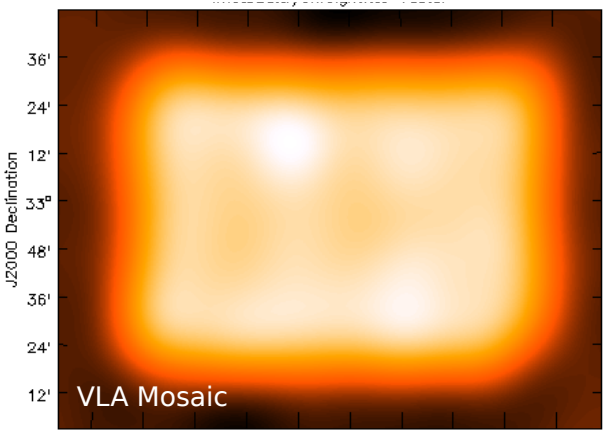


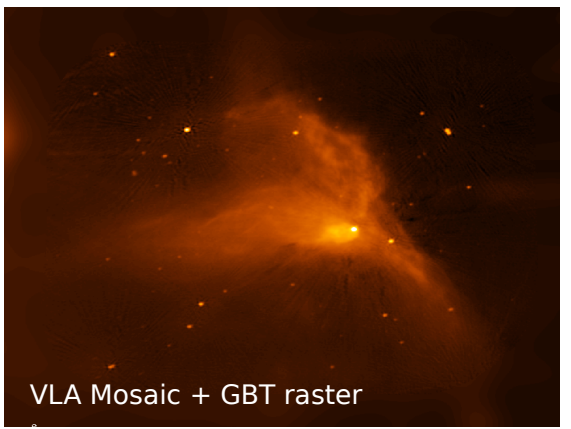
Dyn range = 100K

Wide-Band and Wide-Field Instrumental Corrections

Data : VLA L-band mosaic with 200 pointings. About 800 GB of RSRO data. Data observed ~10yrs ago.



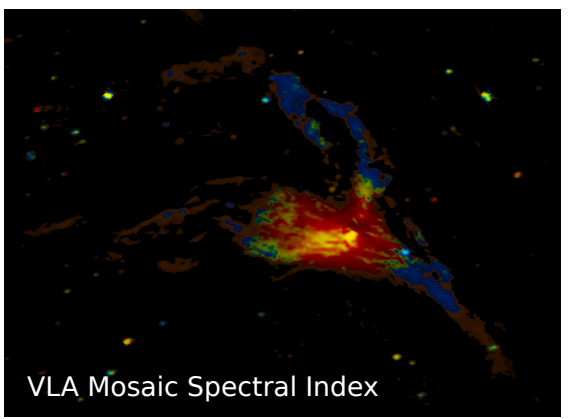




Algorithm and software development

– Parallelization of continuum imaging
– Wideband Multi-term imaging with multi-scale support
– W-term corrections during imaging
– Frequency-dependent antenna models and their use during imaging
– Joint Single Dish and Interferometer image reconstruction
– Gridding on a GPU

(Next : Polarization)



Support for Pipeline operations – in recent years

Pipelines : Automated data reduction (assisted by Data Analysts)
Provide Science Ready Data Products to Astronomers all over the world

Need : CASA code stability and numerical reproduceability
Need : Coordinated CASA packaging and release planning

CASA for ALMA : Work with the NAASC-based PL-working groups

– Algorithms : Image weighting, PSF shaping, Auto-masking, etc…

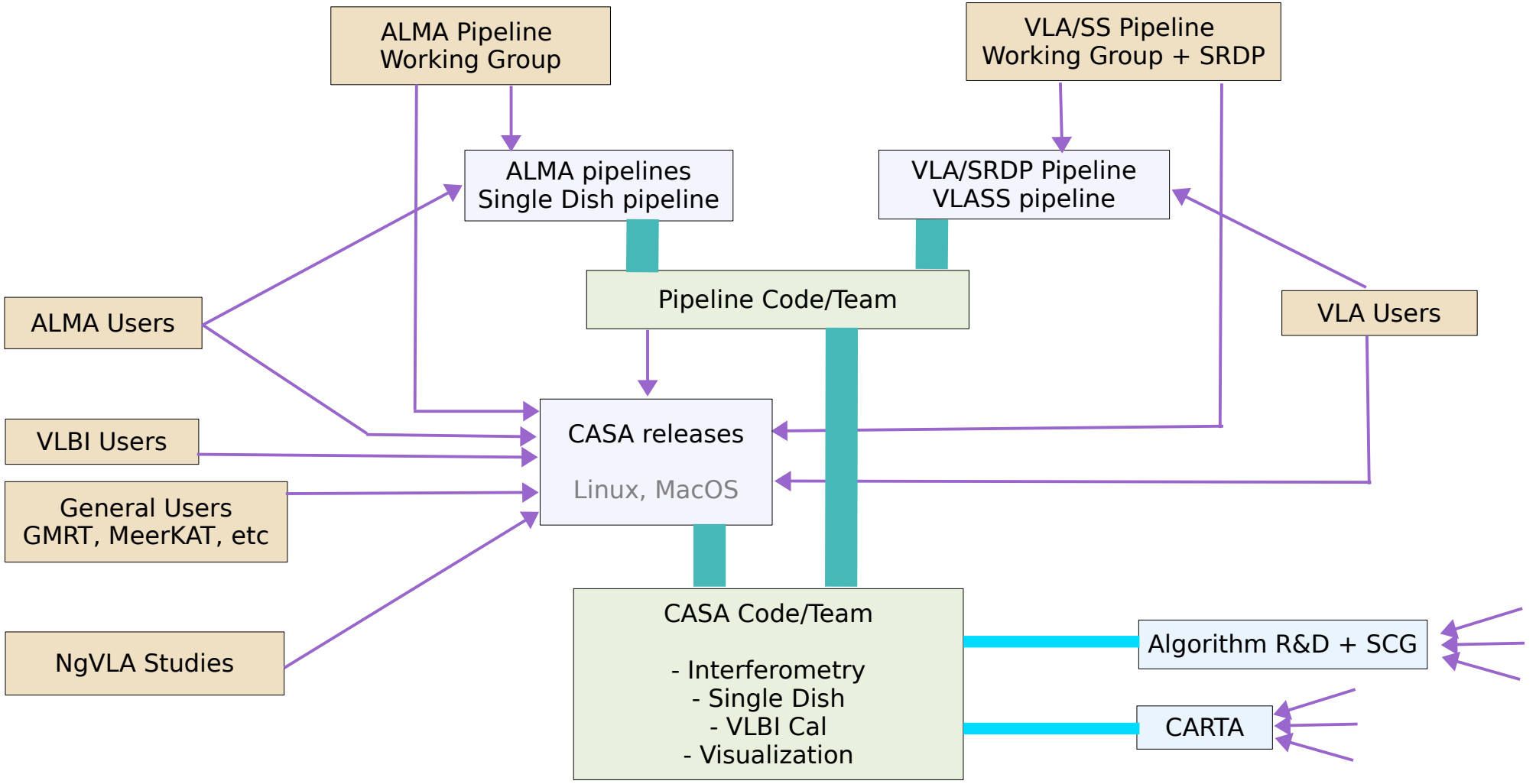
– Performance : Parallelization of Spectral Cube imaging

CASA for VLASS : Work with VLASS working group

– Algorithms : Wideband OTF (On-The-Fly) Mosaics and Pointing Corrections

– Performance : (ARDG) GPU development driven (largely by the compute load of VLASS)

CASA @ NRAO : Current Operations & Stakeholders



CASA development staff (and friends)

NRAO - CASA
Kumar Golap
Takahiro Tsutsumi
Jan-Willem Steeb
Ben Bean
Dave Mehringer
Wei Xiong
George Moellenbrock (VLBI)
Andrew McNichols
Neal Schweighart
Akeem Wells
Bob Garwood
Ville Suoranta
Darrell Scheibel (Visualization)
Pam Harris
Jorge Lopez
Josh Hoskins
Bjorn Emonts (User Liaison)

Ryan Raba (Group Lead + Infrastructure Dev Lead)
Urvashi Rau (Dep. Group Lead + Science Dev Lead)
Jennifer D. Meyer (Proj.Sci. + Validation Lead)

ESO – CASA
Sandra Castro (Verification Lead)
Federico Montesino
Enrique Garcia
Dirk Petry

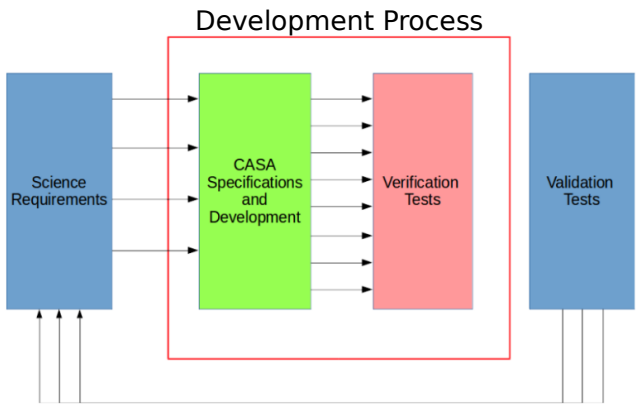
NAOJ – CASA
Takeshi Nakazato (Single Dish)
Wataru Kawasaki
Renaud Miel
Suminori Nishie

Pipeline
Joe Masters
Brian Kent
László Szücs
Rui Xue
Dirk Muders
Vincent Geers
Jose Sabater
Kana Sugimoto
Akira Yoshino
Yohei Hayashi
Hajime Ezawa

ASIAA - CARTA
Kuo-Song Wang
Anthony Moraghan
Heng-Tai Jan
Shou-Chien Hsu
Tien-Hau Chang
Ming-Yi Lin
Mark Chiang

JIVE - VLBI
Mark Kettenis
Des Small
Ilse van Bemmel

NRAO - ARDG
Sanjay Bhatnagar
Martin Pokorny
Mingyu Hsieh
Srikrishna Sekhar
Preshanth Jagannathan



Near Future : Algorithms & Infrastructure

Algorithms

- Full Pol and Wide Band Mosaics : Get improved antenna models and use them in new ways (with ARDG)
- ASP (W-ASP) deconvolution : Improve quality of images with multi-scale emission (with ARDG)
- Automatic flagging : New algorithm to target RFI that existing autoflag methods do not catch
- Joint single-dish and interferometer reconstruction : Numeric/usability improvements
- *GPU gridding (with ARDG)*

Infrastructure

- Visualization tools : New GUIs and modern frameworks
- CASA setup/config : Simplify
- Casa data repo and interface packaging : Refactor
- *CNGI (CASA Next Generation Infrastructure) project and introduction into CASA6*

GPU gridding

Gridding : Data from the telescope (observed visibilities) must be resampled onto a regular grid.

Compute Cost : This step dominates the data analysis and imaging sequence.
Algorithms that apply more accurate instrumental corrections increase cost further

Implementation on a GPU :

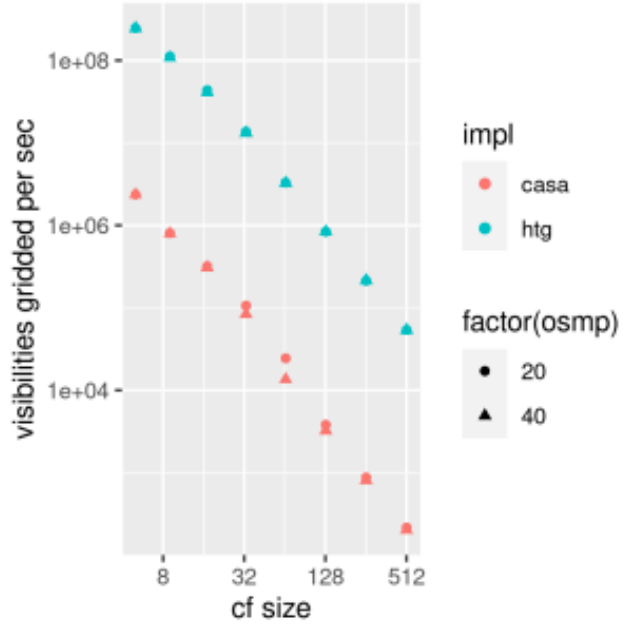
R&D prototype by Martin Pokorny and Sanjay Bhatnagar (ARDG)

Integration into production CASA is ongoing : Kumar Golap et al.

Target : VLASS operations

Results : 200 pointing mosaic at L-band using the VLA
- Parallel CASA : ~7 days
- GPU gridder : ~2.5 hrs

More information : ngVLA Computing Memo #4



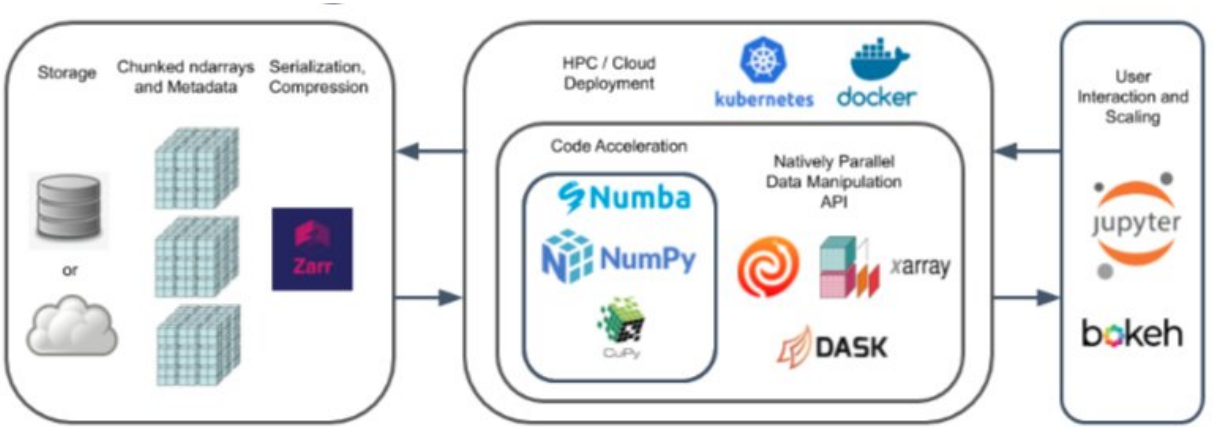
CNGI : CASA Next Generation Infrastructure

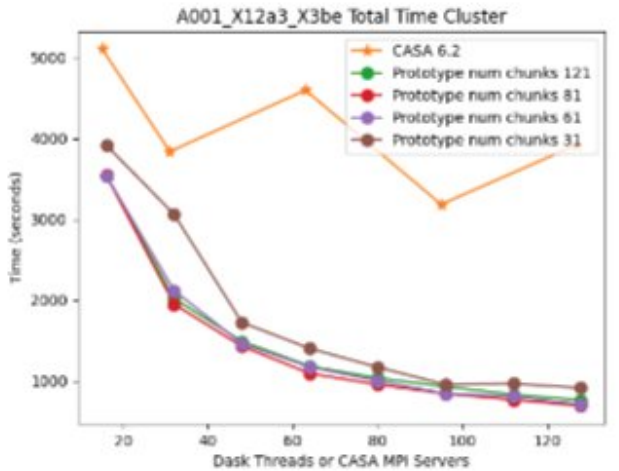
CNGI project : Study new technologies for replacing casacore and parallelization framework of CASA

- **Trade study** : Resulted in a choice of Python-based Xarray and Dask technologies

- **Prototype** : New data/image format, data manipulation, imaging gridders (*Full-pol WB Mosaics*)

- **Benchmarks** : Two large ALMA/VLA datasets, 2 to 6 times speedup (infrastructure choices)





- CNGI demo package : https://cngi-prototype.readthedocs.io/en/stable/
- SiRIUS simulator package : https://sirius-sim.readthedocs.io

(Work led by Ryan Raba & Jan-Willem Steeb)

Future : Transition from CASA6 to ngCASA

ngCASA – Data analysis software for the ngVLA and (upgraded) ALMA

- Likely (but not guaranteed) to be based on CNGI technology

- Code development to start at least 2+ years from now.

- Next 1-2 years

  – Define science and operational use cases (of ngCASA) and numeric metrics
  – Derive algorithm/feature specifications

CASA 6 – Will continue to be used for observatory operations and pipelines.

– Continue CASA6 development of new obs modes

– Try to avoid large code refactors in CASA6

– Introduce some CNGI tech as and where practical (for new feature development)

Questions ?