Interactive Remote Datareduction Framework

JIVE contribution to WP 3.4

Mark Kettenis, Software Project Scientist

JIVE

Joint Institute for VLBI ERIC

JIVE

- Promote and advance the use of VLBI for astronomy
  - Central correlation
  - User services
  - Network support
  - Innovation
  - EC liaison/representation
- Founded in 1993
  - Base budget from partners in 9 countries:
    - China, France, Germany, Italy, Spain, Sweden, United Kingdom, South Africa, the Netherlands
  - Hosted by ASTRON in Dwingeloo, NL
  - A European Research Infrastructure Consortium as of December 2014
- Large number of external projects
  - Many people with temp positions in R&D/Science
- 24 people, 8 nationalities

EVN: European VLBI Network

- Consortium of radio telescopes
  - Involving 15 different organizations around the world: Europe, China, Puerto Rico, South Africa, Russia, Korea
  - More than 30 telescopes
  - And 12 more antennas for “Global observations” with NRAO
- Covering range of frequencies
  - Workhorse frequencies 18cm, 6cm,
  - Also available: SX, 5cm, 1.2cm
  - And at specific stations 90cm, 21cm, UHF, 50cm, 2cm, 7mm
- Reaching mas resolutions
  - From 15mas for 1.4 GHz EVN
  - To 1 mas at 5GHz with Asian, African or American baselines
- Sensitivity of 5μJy in 8hr at 1.4 GHz
  - Combination of Big Antennas and 1 Gbps bandwidth (and soon 2 or 4Gbps)
  - Big antennas also vital for spectroscopy (mJy sensitivity)
- Operational approximately 60 days/year
  - 3 sessions augmented with e-VLBI once a month

VLBI with the SKA

- VLBI is part of the Level 1 requirements
  - Coherent beamformer
- Official VLBI Working Group
- Multiple beams
  - For simultaneous observations of target and calibrators

New software needed

- VLBI (still) uses "classic" AIPS
- Radio Astronomy community uses CASA
- New developments use CASA/casacore.
  - Conversion of polarisation basis
  - Wide field imaging
  - Beam shape correction

Exploratory Data Reduction

- Semi-interactive using (Python) scripts
- Re-run scripts with new parameters and/or small variations
- Redundant re-calculation of intermediate data products
- Errors caused by attempts to avoid re-calculation

Framework for minimal recomputation

1. Turn CASA scripts into execution graphs
2. Compare execution graphs from CASA scripts
   - Using a Graph Partitioning approach
3. Generate sub-graph to be executed
   - Using intermediate data projects from object cache

- HILADO project delivered code to create execution graphs (UCAM) and Haskell code to compare execution graphs (JIVE)
- Use ZFS as backend for the object store

Jupyter

- A new name for IPython notebooks
- Web-based, enabling remote computation
  - Data sets getting too large to ship around
- Used at HPC centers, e.g. TACC
- Good match: CASA uses IPython!

Goals

- Integrate CASA with Jupyter
- Integrate the HILADO minimal recomputation framework with Jupyter
- Deliver Python and Haskell modules necessary to achieve this
- Deliver a description on how to set up a complete envrionment