MeerKAT Software Developments

Jasper Horrell, ThunderKAT Workshop, Arniston – April 2011





Control and Monitoring (CAM)

CAM system basics in place and running KAT-7:

- Standardised low-level CAM interface to devices (katcp).
- Proxy and systems layer (python) talks katcp to devices.
- Python-based low-level user interface allows full access to the system via proxies.
- User scripting interface (also python) built on top of low-level python interface enables standard operations (commissioning, imaging runs etc).
- Fairly rich Flex-based GUI for now.
- New signal displays testing out HTML5-based matplotlib backend.

CAM system still under major development. Lots to add: subarrays, task execution, scheduling, ops planning, proposal management etc, etc. Development will continue using KAT-7 as testbed and targetting MeerKAT.



CAM team of 6 people, managed by Lize van den Heever.

Science Processing (SP)

- During 2010 focus re-established on "science processing" with a dedicated team fairly early days. 6 people currently plus some part-time inputs. Hiring another soon, we hope.
- Currently cutting our teeth on KAT-7 data capture, archive and analysis software, early imaging, scripting framework and general operation scripts.
- Imaging: SP focus on data processing and archive post correlation. Define, design and build the facility systems to support this over next few years.
- Non-imaging: Current SP focus on raw data capture, offline (software) correlation, data analysis, VLBI data capture and processing. A small SETI project performed and thinking about pulsars and transients.
- Interface to the approved large survey project teams re data products, formats, pipeline development, etc. Define boundaries between what is "facility" and what is "outside".
- SIMULATIONS (new team?), prototypes, involvement in early science projects, etc.





Observation Planning/Tracking



KAT-7 Early Fringes (2009)





KAT-7 Horizon Masks





KAT-7 Images (2010)



KA Africa SOUARE KILOMETRE ARRAY

Pic A – 4-dish, single pol



SKA AFRICA

Pic A

Clean image of Pic A at 1820 MHz





Hot off the press!

- Successful baseline calibration data runs using new KAT-7 correlator (5 antennas for now, full pol).
- Fringe stopping tests to follow
- Imaging tests to follow...



Baseline Calibration Fringes – KAT-7 Correlator



Imaging – Next Steps

- Imaging on KAT-7 using new correlator:
 - 5-dish, warm feeds (now).
 - 7-dish, cold feeds (July-Aug 2011)
- Imaging data reduction (important to have more than one package for testing):
 - Ludwig python (exploratory)
 - CASA
 - MeqTrees
- Spectral line imaging (Q3 2011 correlator dependent)
- (Engineering) imaging pipelines development (starting Q3-Q4 2011) next slide



Imaging Pipelines

- (Very) non-trivial !!
- One pipeline to rule them all? almost certainly not!
- A few pipelines to cover the common cases? perhaps.
- Science quality pipeline output? with time and experience... perhaps.



Pipeline "algorithm"

An initial approach:

- Start with CASA and MeqTrees KAT-7 reductions.
- Script and apply to new observations.
- Evaluate and improve.
- Iterate
- ...
- Think MeerKAT size and science. Interact with science teams. Simulate and reduce. Expect quality and performance bottlenecks. Re-evaluate technology choices.
- Apply to emerging MeerKAT.
- Rework and refine
- Iterate



BIG JOINT EFFORT REQUIRED!!

• ...

KAT-7: Archive (Dec 2010)

- Started archiving visibility data from KAT-7 in Dec 2009
- 1948 data captures saved to archive
- Approx 351 GB visibility data
- Approx 353 GB monitoring data!!
- Approx 17 GB logs
- Trivia: Fornax A (247 observations) and Pic A (242) are most observed targets
- This for Fringe Finder (4-input, single ROACH board correlator), excludes KAT-7 correlator data to date.



KAT-7: Archive Future

- Archive will need more and more attention as data requirements evolve. A rework currently underway.
- Already seeing the need to support multiple data sources and formats – Fringe Finder, KAT-7 and holography
- Currently reviewing and testing Apache OODT components to assess viability for MeerKAT data catalog and archiving
- ASTRON and NRAO also have systems of interest....
- Early days....
- The threat of raw data!!! ☺



KAT-7 Data Archive

111	Fringe	Finder

Search Tall Stats EL Find

Fringer Finder Archive

Observer	Jasper	
Description		
Target	Centaurus A	Ffducklus Counch
File Name		FFArchive Search
Start Date		
End Date		

Contact Us

Portions of the look and feel of this site are based on content from mozilla.org | Licensing.

Fringe Finder Search Tall Stats EL Find Fringer Finder Archive Your search matched 4 files. Refine Search New Search DQ DQ .650 1270673254.h5 (1.0 GB) Observed by jasper. Data capture started at 2010-04-07 22:47:34 CAT. Description: Overnight run 4-ant VV imaging of CenA with longer (3 min) cal duration Targets: 3C273, Centaurus A Details EQ EQ .NSC 1270670306.h5 (136.0 MB) Observed by jasper. Data capture started at 2010-04-07 21:58:26 CAT. Description: short test run 4-ant VV imaging of CenA with longer (3 min) cal duration Targets: 3C273, Centaurus A Details DQ DD .h5 1270664341.h5 (267.0 MB) Observed by jasper. Data capture started at 2010-04-07 20:19:01 CAT. Description: longer baseline cal scan run with 4 ants Targets: 3C123, Taurus A, 3C161, J1819-6345, J1939-6342, Orion A, Hydra A, 3C273, Virgo A, Centaurus A, Pictoris A, 30408-6545, 30522-3627 Details EQ EX .NSC 1270598639.h5 (74.0 MB) Observed by jasper. Data capture started at 2010-04-07 02:04:00 CAT. Description: 4 antenna runs with fixed cable swaps Targets: Hydra A, 3C273, Virgo A, Centaurus A, J0408-6545, J1819-6345, J1939-6342 Details Refine Search New Search

Contact Us

Portions of the look and feel of this site are based on content from mozilla.org | Licensing.



KAT-7: Voltage Recording

Nov 2010:

- First site test of raw data capture (full band 400 MHz for a few secs plus GPUbased digital down-converter for "indefinite" period)
- First VLBI data capture run (KAT-7 HartRAO first fringe test)
- First SETI data capture for Project Dorothy
- 300 GB recorded



The Beast - "Big Gulp?"

- 2 x 10 GbE NICs
- 2 x GTX 460 2GB GPUs
- 2 X 6 Core 3.0 GHz Xeons
- 16 x 2 TB HDD
- 24 GB RAM







VLBI Observations and Processing

- Joint HartRAO (26 m) and KAT-7 (12 m, warm feed) antenna L-band observation on 3C273 run on 23 Nov 2010. Data recorded to disks. Shipped to Cape Town for processing.
- Hart recorded 16 sub-bands of 16 MHz in dual pol (circular). The 1696 1712 MHz band was used in fringe test. Data extracted from the Mark5A format using some DiFX utilities and extra C code.
- KAT recorded full 400 MHz band between 1614 and 2014 MHz (single linear polarization). Extracted the same 16 MHz band as Hart data using a software digital downconverter developed for the purpose.
- Fringe stopping software written in python and applied to Hart data (matched to centre frequency and updated over time in small blocks).
- Correlation in python to produce lag plot. Spike shows correlation power where the signals are correctly time aligned. Success!!!
- Re-checked the detection using the DiFX software correlator success, eventually!



VLBI Lag Plot (python correlator)





VLBI Lag Plot (DiFX)



9 x 1 sec integration



9 sec integration



Raw Data: Next Steps

- Development of dual pol full-band raw data capture for all 7 dishes underway (2 10 mins) plus system to record sub-band from single antenna / beam for 3 8 hours. Expect operational in 2nd half of 2011. This will enable:
 - Software correlation test (for comparison with ROACH correlator).
 - Software beamforming test (for comparison with ROACH beamformer).
 - "Routine" VLBI with KAT-7 beamformer (to disk in VDIF format) by end 2011 using the new hardware/software.
 - (Engineering ?) exploration of RFI, fast transients, pulsar timing and search, SETI (no expected dates yet).
- New 10 Gbps link should enable e-VLBI. Expect first tests using DiFX Q1 2012.



MeerKAT to provide:

- Calibrated visibilities
- Certain standard pipelines (spectral line image cubes, continuum images)
- Storage for visibility data (project disk quotas?)
- A 10 PB archive on site plus 3 PB in CT, plus likely European mirror
- Archive access and facilities to reprocess from archive
- Flexible architecture for "black belt" users plus space in the Array Processor Data Centre for additional user-supplied equipment. Care is needed re power promises – can be large cost.
- Support for some re-use of existing mature packages where possible (support CASA, MeqTrees).



SP: In place now

- Basic data capture framework to HDF5 file archive format and MS export writing capability
- Simple archive for commissioning purposes
- SCAPE commissioning package
- First imaging software
- Simulation and data reduction tests in Meqtrees (Oleg, Brad, etc)
- Prototype raw data system based on ROACH sampling, followed by GPUs preprocessing and disk storage. Allows for raw data recording for commissioning, VLBI, software correlation, SETI
- Emerging NRAO and ASTRON collaborations
- Some great early results on the KAT-7 system
- A lot of work to be done over the next few years (we won't be bored).



SP: Roadmap (likely)



Conclusions

- Major joint effort needed over coming years to ensure that the science teams hit the ground running
- Also need to support commissioning and smaller PI proposals
- Technology within reach for MeerKAT scale imaging
- Fine tuning (e.g. imaging pipeline) will take time
- Technology challenges in pulsar search area
- KAT-7/8 will continue to be an excellent engineering and science test bed
- Plan re-use software where possible, but expect significant custom developments
- Lots to do. Technical team will need to grow a bit. Project teams can help (a lot).

Questions...

Credits (full SP Team):

Thomas Bennett Mattieu de Villiers Jasper Horrell Tshaks Namalili Simon Ratcliffe Ludwig Schwardt

jasper @ ska.ac.za

http://www.ska.ac.za

KAT-7 – our playground







