South Africa: SKA Regional Centre

Dr. Rob Simmonds Associate Director IDIA



Overview



- MeerKAT RC / Tier 2 facilities
- SDP DELIV
 - Data delivery design for the SKA
- DOME collaboration: ASTRON, IDIA, SKA-SA, IBM
 - Data distribution and data processing for MeerKAT data
- SKA-SA activities
 - KAT 7 and MeerKAT
 - Archiving MeerKAT data
 - Pipelines / Machine learning
- CyberSKA portal and visualization activities

ARC / IDIA / Tier 2 systems

- Hardware deployed at UCT and NWU to act as MeerKAT Tier 2 system
- Currently 48 compute nodes (1500 cores) / 1.5PB storage
- Additional systems being deployed at UWC, UP and Wits will be used for distributed prototyping
- Exploring expansion of ARC to AVN / SKA partner countries



Software environment



- OpenStack virtualized environment with docker deployment support
- Mix of Swift object store, CEPH block and object store and BeeGFS parallel file-systems
- Interoperating between sites using OpenStack services
- Adding OCCI for federating beyond ARC
- Access services include cyberSKA and Jupyter notebooks
- CASA production and development environments in place
- Currently using keystone for AuthN and AuthZ
 - In process of moving AuthN to SAFIRE/eduGAIN
 - Exploring approaches to role based AuthZ

SDP DELIV

- SDP design consortium work package for data delivery
- Partners from UCT, CADC, SKA-SA/SAC, ASTRON, IAA and Oxford
- Initially not clear which parts of architecture should be part of SDP
- Some parts of what was initially planned have moved to Observatory Support in SDP





"DOME" Collaboration



- Project working on MeerKAT regional science processing activities
- Partners: ASTRON, IDIA, SKA-SA and IBM
- Two components:
 - Data transport for moving data to Tier 2 centres
 - Processing pipelines for executing at Tier 2 centres
- Data transport implement key parts of SDP DELIV data transfer components
 - Request, Prepare and Deliver
 - Query component not currently part of project, but will be added to system in collaboration with CADC
- Focusing on use cases
 - Radio Astronomy Calibration & Imaging

Pipelines



- Have Docker container with:
 - CASA, drive-casa, python libraries, Jupyter Notebook & Hub.
- Developing software framework for running pipelines in Notebook
- Proof of concept working
- Still needs to be combined with data transport system to provide automated archive to product execution

Current Notebook Dashboard on ARC



Pipeline Control & Data Exploration

Visualization

JUPyter ACF1G1 Last Checkpoint Last Tuesday at 5:06 PM (autosaved)	Control Panel Logout	Jupyter plotting with apercal Last Checkpoint. 01/27/2017 (subserved)	Control Panel Cogout
File Edit View Imart Gell Karnal Help	Pythan Z. O	File Edit View Inset Gell Kernel Help	Python 2. C
E + 9: 0 E + + H B C Mandoor - Califordiar		5 + > 2 6 + + x = C Cook 1 = Calliolar	
ACF1G1 In this Netebook I am going to try to calibrate WSRT Observations of you guessed it ACF1G1. First, I'm going to use this hack to source the silriad_start.ab file, so that I can access MIRIAD.		<pre>in []; from apercal import lib lib.ertup_logger(logfile='/home/frank/mplot.log', quiet=True) import pylab as pl setup logger with lib.setup_logper() logging to file. To see the log in a bash window use the following command: tail -n +1 -f /home/frank/mplot.log</pre>	
<pre>In []: import sys sys.path.append('/home/frank/spercel/') import discretes</pre>		Now import the plotting library	
import os command = ['hash', '-c', 'cource /home/frank/miriad/miriad_start.sh && env']		In [9]: mpiot.swpit.vis='/home/frank/acf/acfigi/working/call.wv',	
<pre>proc = subprocess.Popen(command, stdout = subprocess.PIPE) for lise in proc.stdout: (key, _, value) = lise.partition('s') os.esviron(key) = value.replace('\n',') proc.communicate()</pre>		Out(\$): 31() YT1 million (4403 EE: 4-1) 11() YT1 million (4403 EE: 4-1) S S	4 1- 3
<pre>in (); from spercal import lib lib.setup_logger('info', logfile='/heme/frank/anfigl-reduction.log') import pylab as pl tastplotlib laline from spercal import calibrate</pre>		kupitose biol 0.m Lingebae	
<pre>root - 13FO : Logging 44artadl root - 13FO : To see the log in a bash window use the following command: root - 13FO : tail = n -1 -f /homm/frank/afigi-reduction.log</pre>			
Setup logger with lib.setup_logger()		వరింది దరి నది వరి వరి వరింది కరి వరి Door	⁰ 27 ^m
<pre>in (0): ccal = calibrate.crosscal() scal = calibrate.wselfcal()</pre>		22 1 XY T caller 1400 UTs i- 4 II Y I'I'I caller 1400 CT	9 - 3-1 P
The file rames are: 11406328.50, T0.M5 11406339, 50, T0.M5 11406330, 50, T1.M5 11406330, 50, T1.M5 11406330, 50, T1.M5 11406330, 50, T3.M5 11406330, 50, T5.M5 11406330, 50, T6.M5 11406330, 50, T6.M5 11406330, 50, T6.M5 11406330, 50, T6.M5 11406331, 50, T0.M5			
In [9]: path = '/home/frank/act/actigl/working/'		fans Trae	
in [10]: call = calibrate.eource()		Im (0): mplot.uvplt7	

SKA SA - Programmes







MeerKAT First Light (FR II galaxy)





MeerKAT Processing and Archive



SKA1 Era Regional Centres





Galaxy Morphology

Arun Aniyan



Shapelet and Deep Learning approaches under development



FRII Type Galaxy



Bent Tail Galaxy

1 2 3 4

80 100 120 140

100 120

0 20 40 60



Pulsar Candidate Selection



Image: NASA











CYBERSK A Cyberinfrastructure pl	A atform to meet the needs	of data intensive rad	dio astronomy on route to the S	5KA		
Home Profile Settings myD	ashboard myGroups 🖂	Tools About Help		Search	Go Log out 🗵	
Cameron Kiddle	Contacts	EDIT =	Event calendar	EDIT - Recent Astro-ph E	Eprints EDIT -	
Subscribe to feed Bookmark this Edit page layout My applications My blog My bookmarked items			 Imaging Science Technical Me Weekly imaging science technic 10:00 - 11:00, 28 Jun 2011 CANARIE Site Visit (Tentative Semi-annual site visit by CANA 1:00 - 3:00, 29 Aug 2011 ADASS XXI Astronomical Data Analysis Sof Systems Conference 6 Nov 2011 - 10 Nov 2011 	eeting arXiv:0812.2984v1 Testing the cosmologica galaxies with the SKA arXiv:0812.0141v2 A generalised Measurem Zernike theorem for wide interferometry arXiv:0811.1070v1 The Directional Depende Technique for UHE Neutur arXiv:0811.0211v1 Pulsar searches and timi	evolution of magnetic fields in PDF ent Equation and van Cittert- e-field radio astronomical PDF ino Detection PDF ing with the SKA PDF	
My event calendar My files				Active Users	EDIT -	
My pages My publications My tasked items My activity Contact's activity Site activity DMS support of Astronomical Data This subgroup is for planning and developing specialized support Astronomical Data in the CyberSKA Data Management System. CyberSKA Sys Admins	EDIT - tronomical Data planning and red support n the CyberSKA Data 1.	Activity Contacts Samuel George bookmarked Detect Thresholds and Bias Correction in Intensity (4 hours ago) Russ Taylor updated a page titled 5 hours ago)	EDIT - Cameron Kidd Busy preparing http://ska2011 Location: calgary SubGroups (9 Cameron Kidd Busy preparing http://ska2011 Location: calgary my "Job is wait back later for r Location: castrop	Cameron Kiddle Busy preparing for SKA 2011 - http://ska2011.org/ update (12 days ago) Location: calgary, alberta, canada Samuel George my "Job is waiting for processing, please check back later for results" (5 days ago) Location: astrophysics group, university of		
	CyberSKA Sys Adm	ins	Mircea Andrecut has posted a new this discussion topic SKA 2011 Tr UofC participants I can take one person. (16 hours ago)	comment on ravel Plans for Ohe, uk	cambridge, cavendish laboratory cambridge cb3 Ohe, uk	
	Group for developers working/c		Arne Grimstrup has posted a new o this discussion topic SKA 2011 Tr UofC participants Arrive in Banff: July 3 Depart Banff:	comment on ravel Plans for July 8 I do not	ated Collaboration - Phase III days ago by Cameron Kiddle	
	Portal Support Group for portal sup	oport -	have a vehicle but was considering Banff Airport shuttle bus.	taking the More pages	More pages	



- Over 700 users
- Latest version enables federation of multiple portals
- iRods used for data management
- Provides access to data sharing, collaboration, visualisation and data search tools

Visualisation

- CyberSKA remote radio astronomy viewer
- Currently used with up to 360 GB data cubes
- Provides range of visual analytics algorithms
- Now developing CARTA viewer – NRAO collaboration
- Aim to scale analytics to multi-terabyte cubes
- Considering HDF5 formats for efficient execution









- SKA SA aim to support activities for SRC development
- SDP DELIV has developed design for delivering data to SRCs and is working on interface definition document
- MeerKAT RC framework being developed in multi-partner collaboration
- CyberSKA portal and CARTA viewer development is ongoing
- South Africa planning to have SRC in addition to SKA1 Mid Processing Centre

