Commissioning and Science Verification: MeerKAT

Complexity and Scale

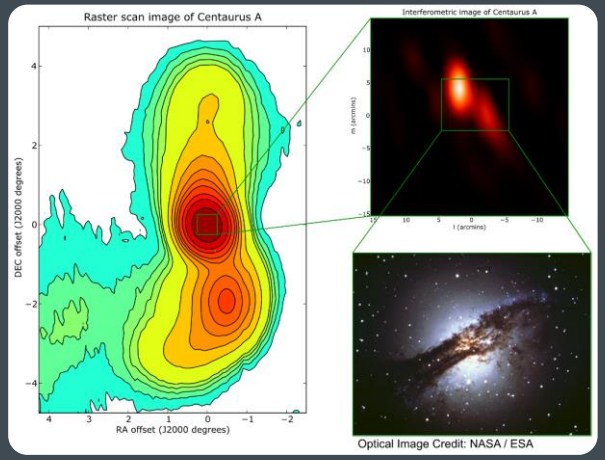
Simon Ratcliffe

64 Dishes Complete – Inaugurated 13 July 2018

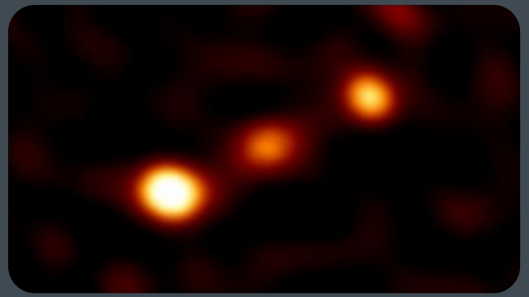
Continuous Commissioning

Two Dish Image

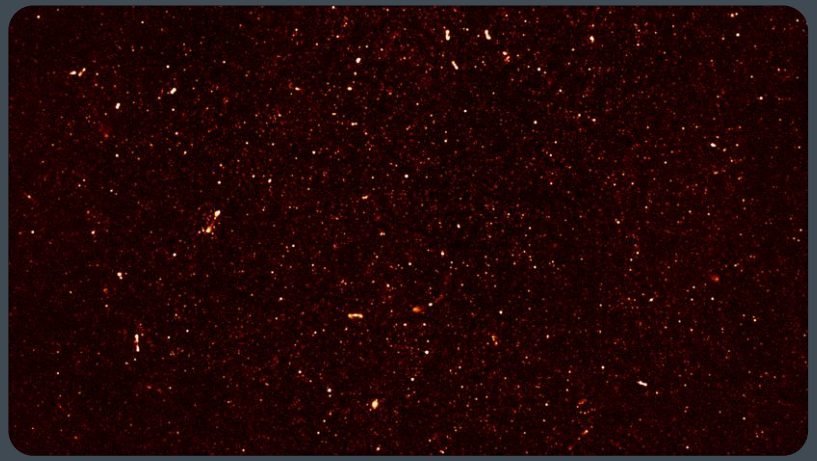
Single Dish Raster



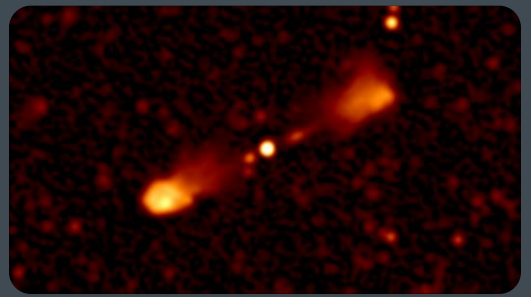
4 Dish MeerKAT



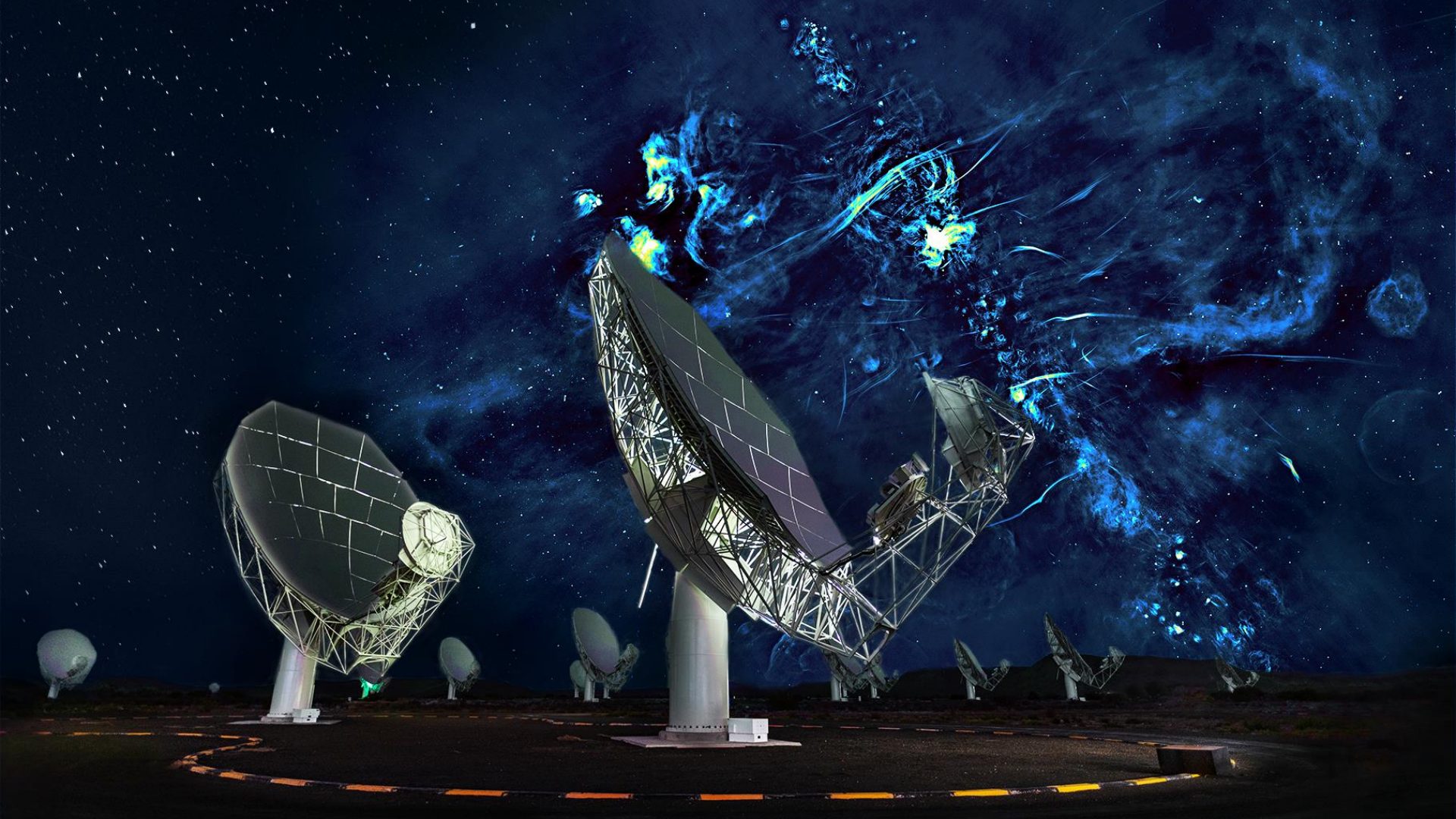
16 Dish MeerKAT



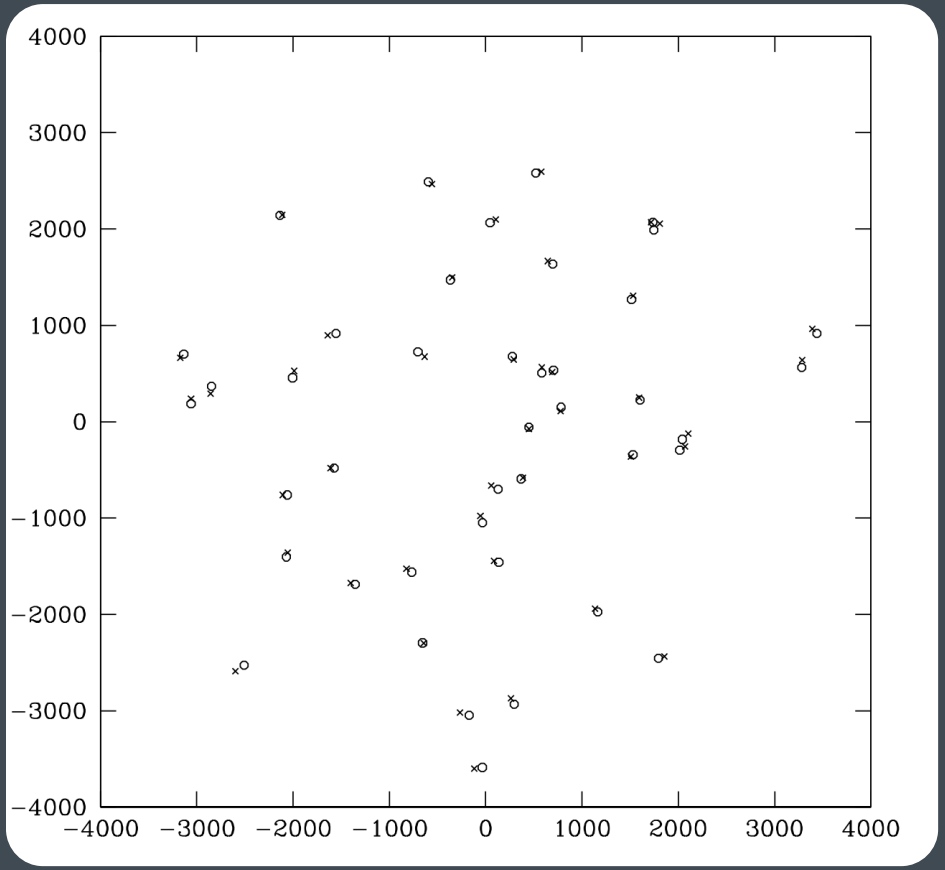
64 Dish MeerKAT



32 Dish MeerKAT



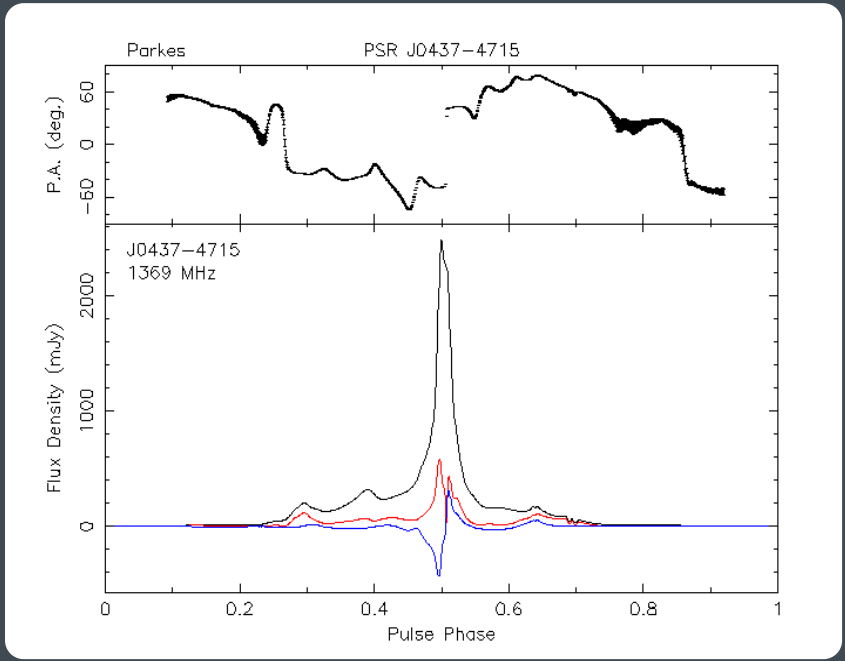
Getting better all the time…



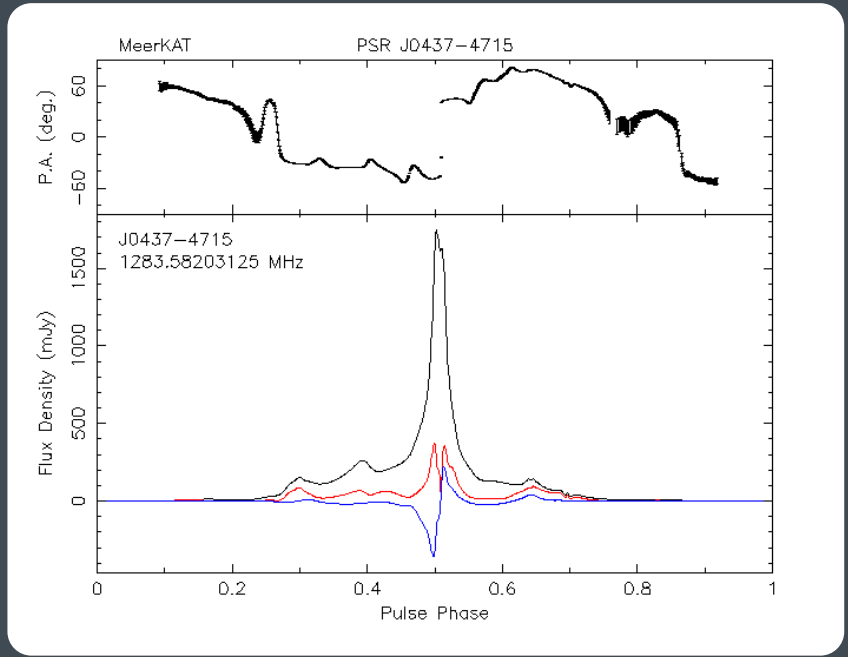
Mean offsets are
-0.08 arcsec in RA
+0.14 arcsec in DEC

J. Condon, B. Cotton, T. Mauch
Axes are arcsec offset from phase center.
Position differences x100 to show offsets.

Partnerships with USE very beneficial



300 Hours



4 Hours - MeerKAT

System Complexity is Starting to Hurt

Engineering agility, coupled with schedule pressure (WSJF) tends to favor producing small features more often.

For commissioning this introduces more context switching and pressure to understand the latest exciting feature before the old one is fully verified.

The system is getting more complex as it evolves - mostly software – which inevitably introduces more bugs

That can’t be right, we won’t have bugs because…

We have code reviews
We have unit and smoke tests
We integrate early and often
We have simulators
We have continuous builds
We have signed ICDs
We have smart people
We test on the sky
We have a modular system
We can rollback changes quickly
We have a robust Engineering Change Process
We have a ticketing system
~~We have a silver bullet~~

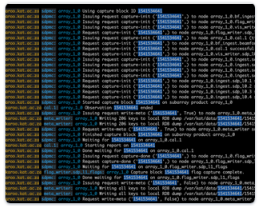
Complexity – An Aside

SR-1458

ThunderKAT OBS files missing for P-ID:SCI-XXXXXX-AA-01

Document link for Observation request UZ For

Attachments



Linked issues

relates to

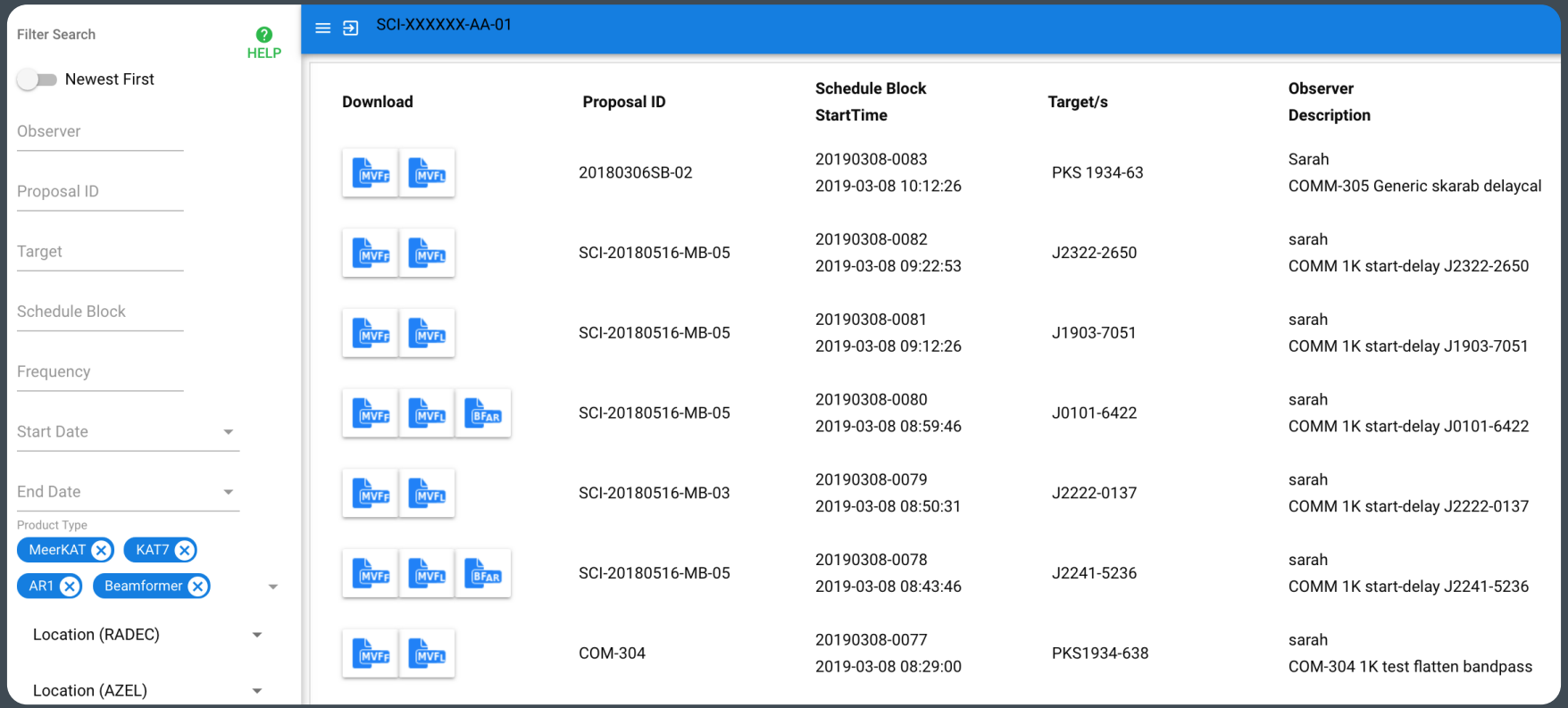
SR-1446 Observations not in archive - umbrella ticket TO DO

Activity Comments

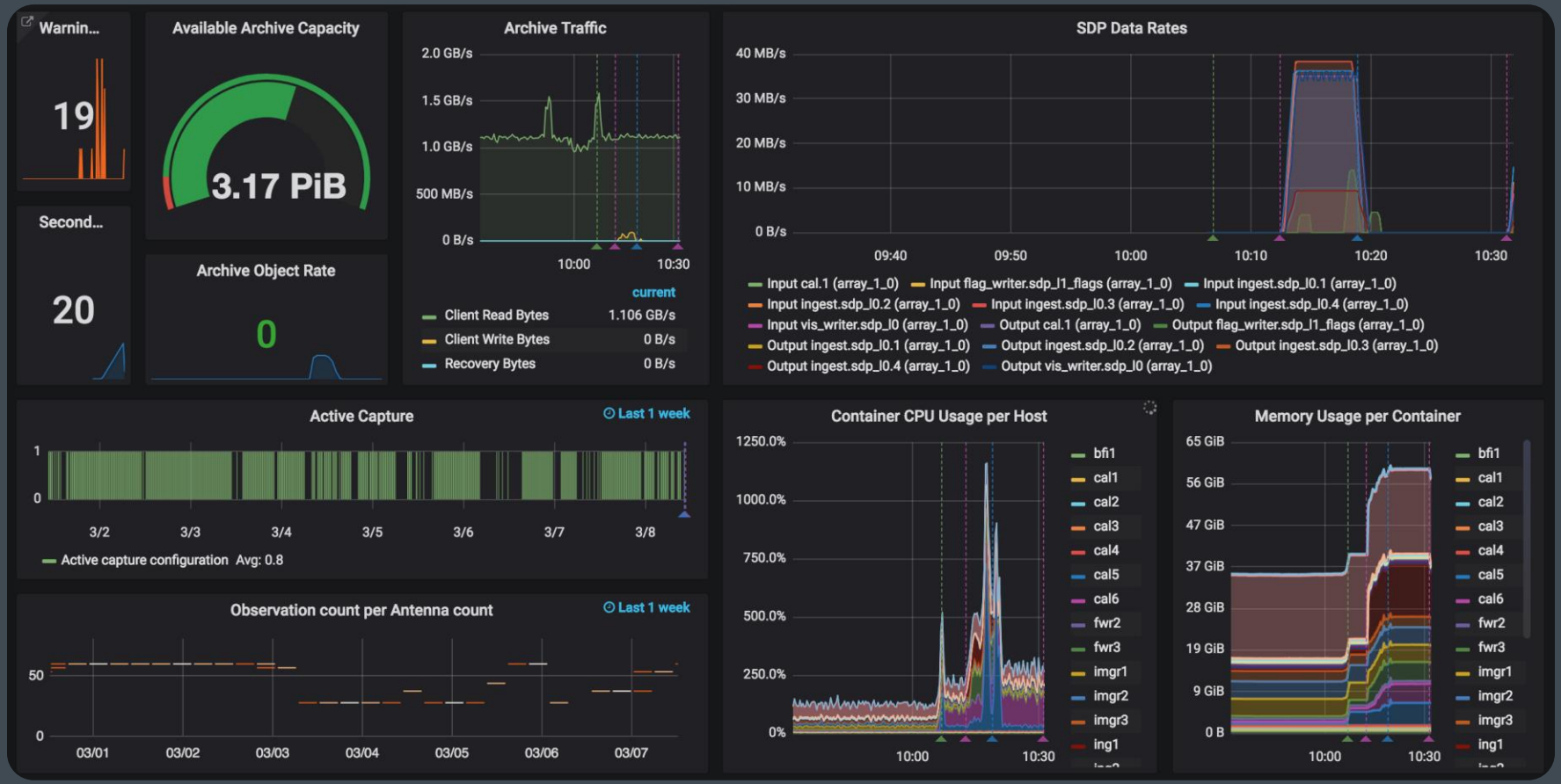
Complexity – An Aside

```
2019-03-05 16:24:22.230Z INFO     Quorum = 90% -> 26 of 28 receptor(s), i.e. 2 straggler(s) tolerated
2019-03-05 16:24:22.237Z INFO     Delay tracking = on
2019-03-05 16:24:22.237Z INFO     Switching off all noise diodes
2019-03-05 16:24:22.354Z INFO     Noise diode will not fire automatically while performing canned commands
2019-03-05 16:24:22.367Z INFO     Stopping / readying antennas
2019-03-05 16:24:23.089Z INFO     INIT
2019-03-05 16:24:23.089Z INFO     Capture block ID = 1551803062
2019-03-05 16:24:23.304Z INFO     -------------------------
2019-03-05 16:24:23.308Z INFO     New compound scan: 'interferometric_pointing'
2019-03-05 16:24:24.135Z INFO     slewing to target
2019-03-05 16:24:56.211Z INFO     Waiting for sensor 'lock' == True had 1 straggler(s): ['m044']
2019-03-05 16:24:56.212Z INFO     target reached
2019-03-05 16:24:56.217Z INFO     tracking target
2019-03-05 16:25:06.224Z INFO     target tracked for 10.0 seconds
2019-03-05 16:25:06.226Z INFO     Initiating 10-second track on target 'J0010-4153'
2019-03-05 16:25:06.226Z INFO     Offset of (0.000000, -1.500000) degrees
2019-03-05 16:25:06.812Z INFO     slewing to target
2019-03-05 16:25:13.941Z INFO     target reached
2019-03-05 16:25:13.949Z INFO     tracking target
2019-03-05 16:25:23.956Z INFO     target tracked for 10.0 seconds
2019-03-05 16:25:23.957Z INFO     Initiating 10-second track on target 'J0010-4153'
2019-03-05 16:25:23.958Z INFO     Offset of (0.000000, -0.750000) degrees
```

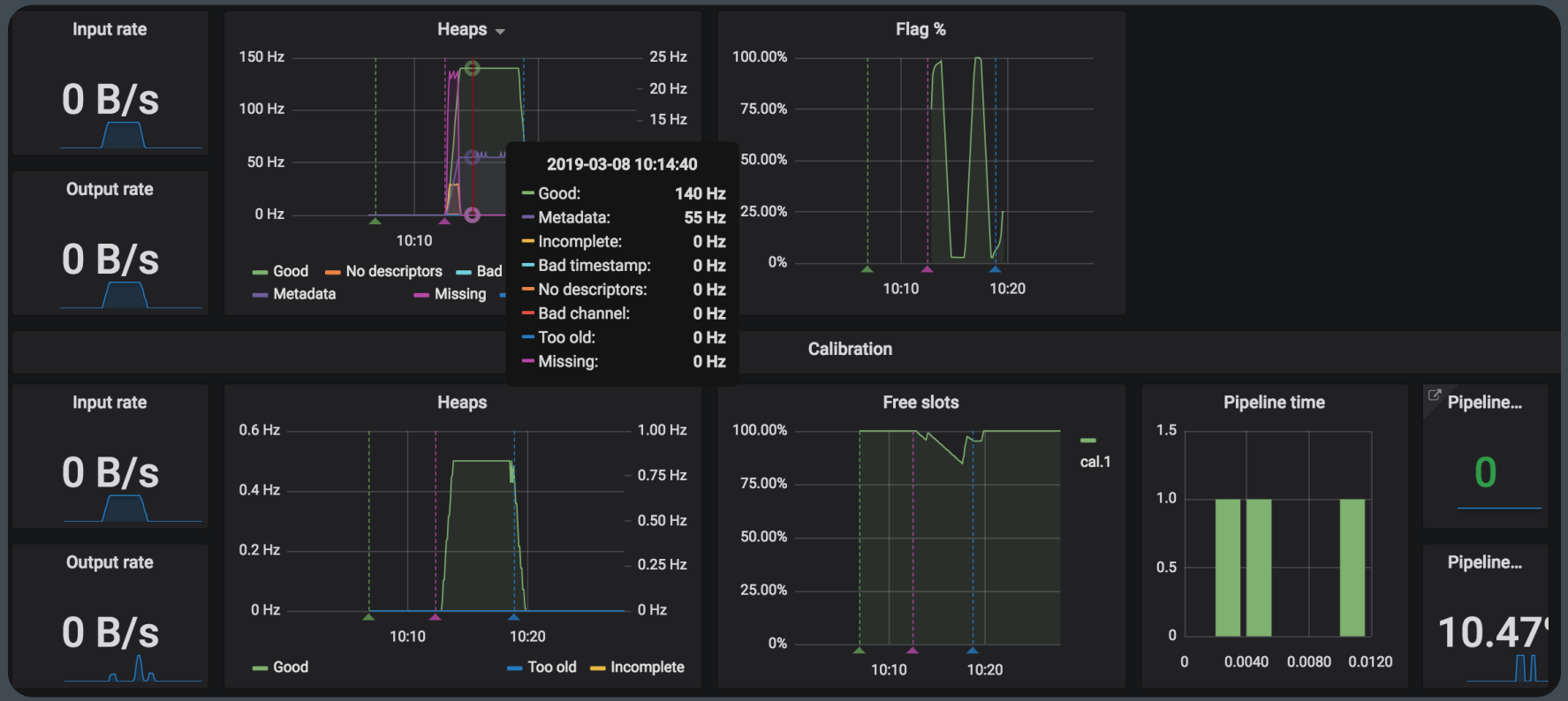
Complexity – An Aside



Complexity – An Aside

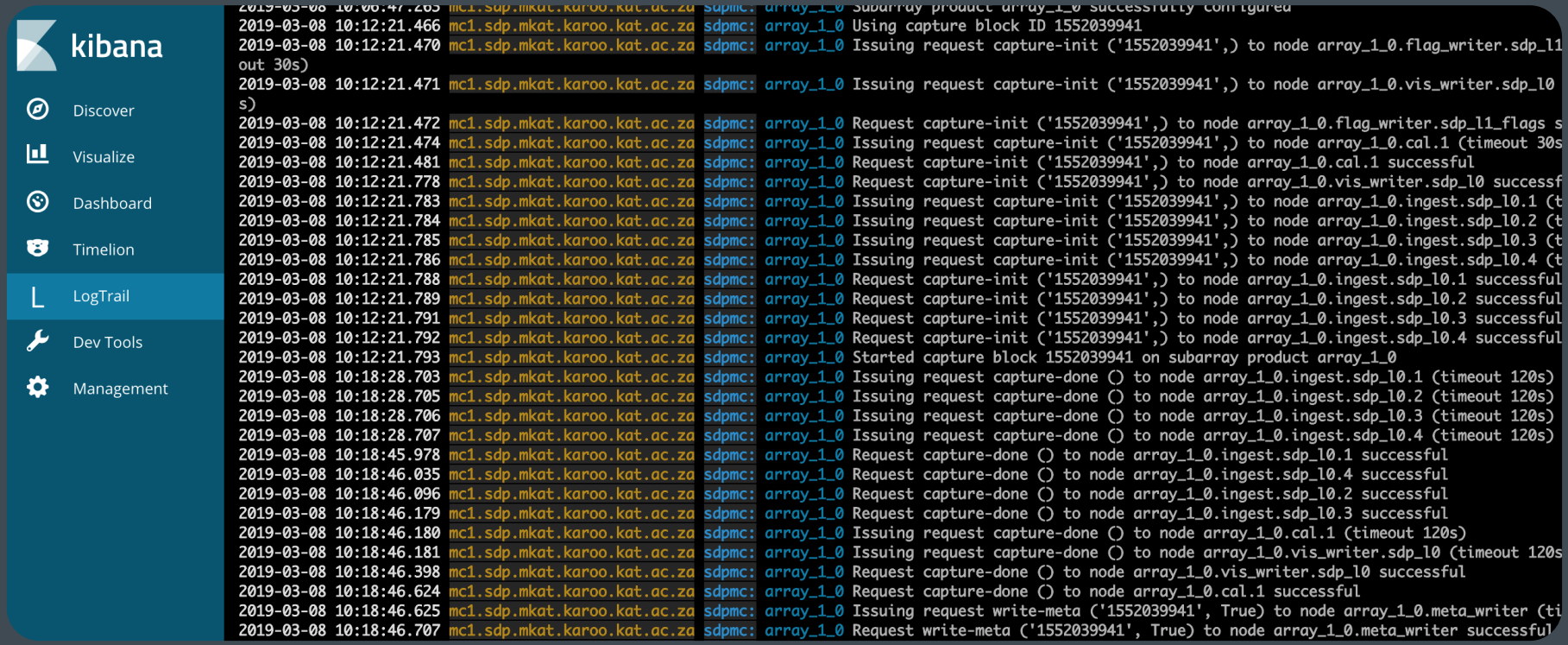


Complexity – An Aside



Complexity – An Aside

Complexity – An Aside



Complexity – An Aside

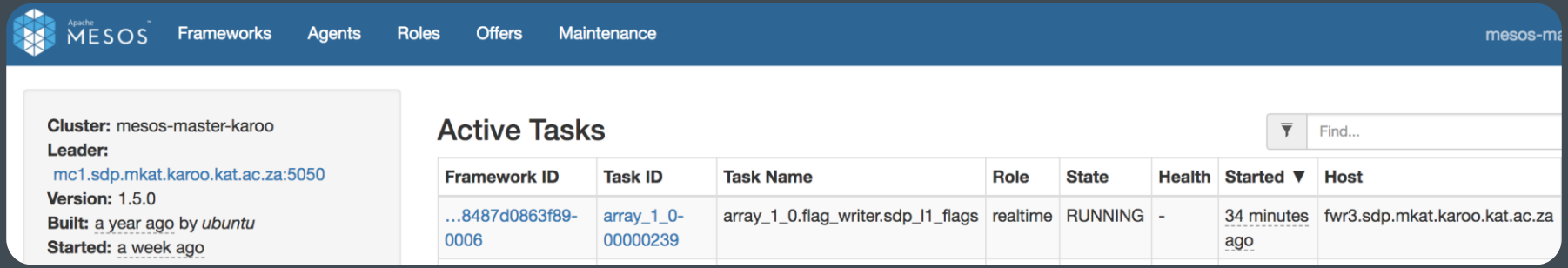
array_1_0

CAPTURING

Tasks | Config | Capture blocks

Name	State	Mesos State	Host
telstate	READY	TASK_RUNNING	cal6.sdp.mkat.karoo.kat.ac.za
cam2telstate	READY	TASK_RUNNING	cal6.sdp.mkat.karoo.kat.ac.za
meta_writer	READY	TASK_RUNNING	fwr3.sdp.mkat.karoo.kat.ac.za
ingest.sdp_l0.1	READY	TASK_RUNNING	ing1.sdp.mkat.karoo.kat.ac.za
ingest.sdp_l0.2	READY	TASK_RUNNING	ing1.sdp.mkat.karoo.kat.ac.za
ingest.sdp_l0.3	READY	TASK_RUNNING	ing1.sdp.mkat.karoo.kat.ac.za
ingest.sdp_l0.4	READY	TASK_RUNNING	ing1.sdp.mkat.karoo.kat.ac.za
cal.1	READY	TASK_RUNNING	cal5.sdp.mkat.karoo.kat.ac.za
vis_writer.sdp_l0	READY	TASK_RUNNING	fwr3.sdp.mkat.karoo.kat.ac.za
flag_writer.sdp_l1_flags	FINISHING	TASK_RUNNING	fwr3.sdp.mkat.karoo.kat.ac.za
timeplot.sdp_l0	READY	TASK_RUNNING	cal6.sdp.mkat.karoo.kat.ac.za

```
kat@cms3:~/tbennett$ s3cmd -v --config s3cfg_kat_c1 ls s3://1551803062-sdp-l0/
                     DIR   s3://1551803062-sdp-l0/correlator_data/
                     DIR   s3://1551803062-sdp-l0/flags/
                     DIR   s3://1551803062-sdp-l0/weights/
                     DIR   s3://1551803062-sdp-l0/weights_channel/
kat@cms3:~/tbennett$
```

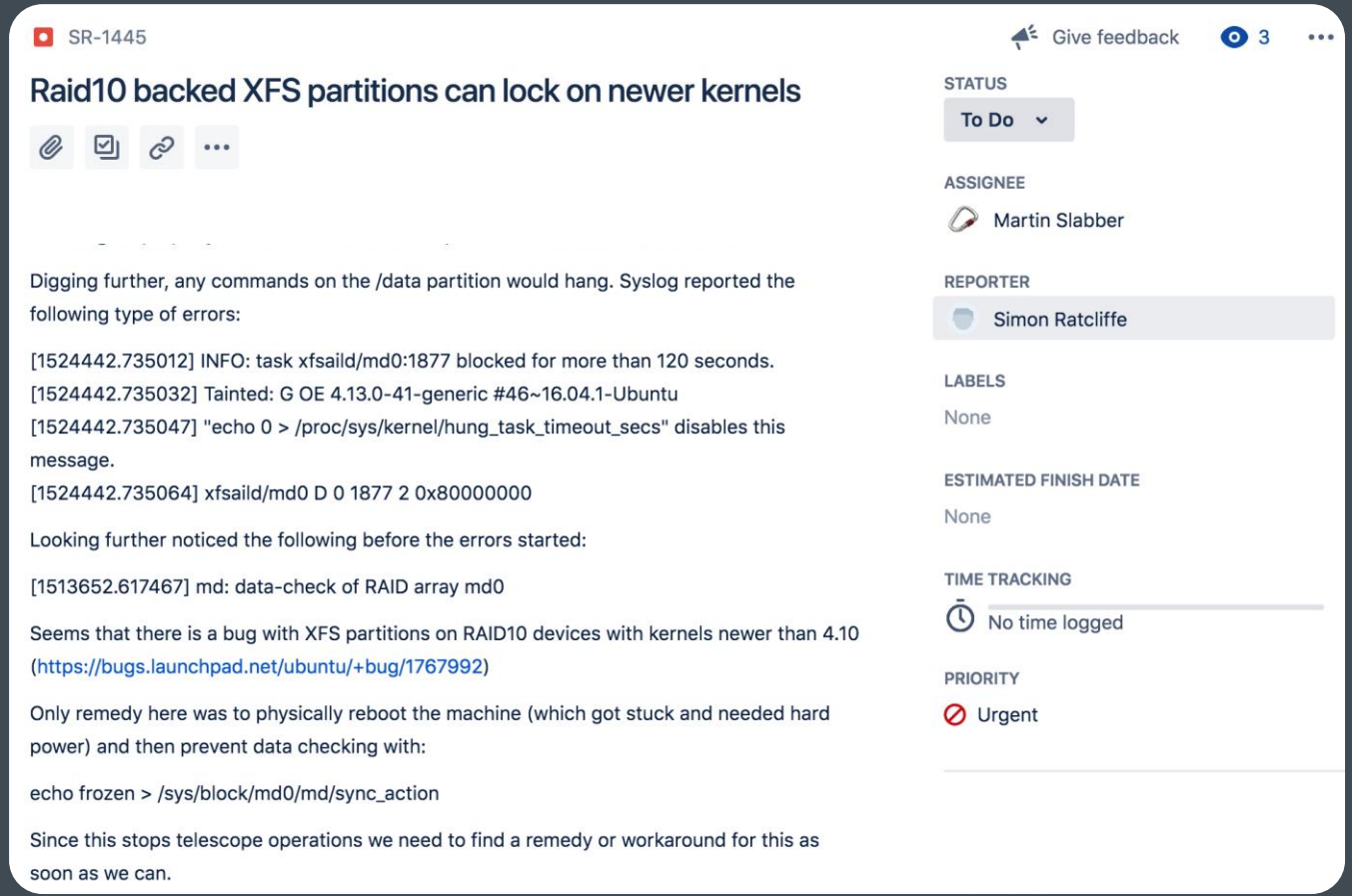



```
kat@fwr3:/$ telnet localhost 31173
Trying 127.0.0.1...
Connected to localhost.
Escape character is '^]'.
Python 3.6.7 (default, Oct 22 2018, 11:32:17)
[GCC 8.2.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
---
This console is running in an asyncio event loop.
It allows you to wait for coroutines using the 'await' syntax.
Try: await asyncio.sleep(1, result=3)
---
>>> loop
<_UnixSelectorEventLoop running=True closed=False debug=False>
>>> await asyncio.sleep(1, True)
```

Complexity – An Aside

```
Mar  3 00:57:01 fwr3 kernel: [2307278.036251] md: data-check of RAID array md0
Mar  3 01:03:48 fwr3 kernel: [2307685.087047] INFO: task xfsaild/md0:1881 blocked for more than 120 seconds.
Mar  3 01:03:48 fwr3 kernel: [2307685.087066]       Tainted: G           OE   4.13.0-41-generic #46~16.04.1-Ubuntu
Mar  3 01:03:48 fwr3 kernel: [2307685.087082] "echo 0 > /proc/sys/kernel/hung_task_timeout_secs" disables this message.
Mar  3 01:03:48 fwr3 kernel: [2307685.087098] xfsaild/md0     D    0  1881      2 0x80000000
Mar  3 01:03:48 fwr3 kernel: [2307685.087100] Call Trace:
Mar  3 01:03:48 fwr3 kernel: [2307685.087105]  __schedule+0x3d6/0x8b0
Mar  3 01:03:48 fwr3 kernel: [2307685.087107]  schedule+0x36/0x80
Mar  3 01:03:48 fwr3 kernel: [2307685.087110]  wait_barrier+0x14e/0x1b0 [raid10]
Mar  3 01:03:48 fwr3 kernel: [2307685.087112]  ? wait_woken+0x80/0x80
Mar  3 01:03:48 fwr3 kernel: [2307685.087114]  raid10_write_request+0x3d/0x840 [raid10]
Mar  3 01:03:48 fwr3 kernel: [2307685.087115]  ? r10bio_pool_alloc+0x24/0x30 [raid10]
Mar  3 01:03:48 fwr3 kernel: [2307685.087118]  ? mempool_alloc+0x73/0x180
Mar  3 01:03:48 fwr3 kernel: [2307685.087120]  ? md_write_start+0xc8/0x210
Mar  3 01:03:48 fwr3 kernel: [2307685.087122]  ? find_busiest_group+0x47/0x4d0
Mar  3 01:03:48 fwr3 kernel: [2307685.087124]  raid10_make_request+0xd0/0x130 [raid10]
Mar  3 01:03:48 fwr3 kernel: [2307685.087126]  md_handle_request+0xb5/0x130
Mar  3 01:03:48 fwr3 kernel: [2307685.087128]  md_make_request+0x6c/0x170
Mar  3 01:03:48 fwr3 kernel: [2307685.087130]  generic_make_request+0x12a/0x300
Mar  3 01:03:48 fwr3 kernel: [2307685.087131]  submit_bio+0x73/0x150
Mar  3 01:03:48 fwr3 kernel: [2307685.087132]  ? submit_bio+0x73/0x150
Mar  3 01:03:48 fwr3 kernel: [2307685.087162]  _xfs_buf_ioapply+0x2e7/0x4a0 [xfs]
Mar  3 01:03:48 fwr3 kernel: [2307685.087177]  ? xfs_buf_delwri_submit_buffers+0x14b/0x280 [xfs]
Mar  3 01:03:48 fwr3 kernel: [2307685.087189]  xfs_buf_submit+0x63/0x210 [xfs]
Mar  3 01:03:48 fwr3 kernel: [2307685.087201]  ? xfs_buf_submit+0x63/0x210 [xfs]
Mar  3 01:03:48 fwr3 kernel: [2307685.087213]  xfs_buf_delwri_submit_buffers+0x14b/0x280 [xfs]
Mar  3 01:03:48 fwr3 kernel: [2307685.087225]  ? xfs_buf_delwri_submit_nowait+0x10/0x20 [xfs]
Mar  3 01:03:48 fwr3 kernel: [2307685.087237]  xfs_buf_delwri_submit_nowait+0x10/0x20 [xfs]
Mar  3 01:03:48 fwr3 kernel: [2307685.087248]  ? xfs_buf_delwri_submit_nowait+0x10/0x20 [xfs]
Mar  3 01:03:48 fwr3 kernel: [2307685.087264]  xfsaild+0x286/0x750 [xfs]
```

Complexity – An Aside



The solution

?

Aside from complexity, scale is next…

Increased reliability and new modes challenge commissioning due to ever increasing amounts of data

Since inauguration we have collected 1000's of hours of data. At least 50% of this has not been looked at in any significant detail.

It is only getting worse – commensal just started (narrow + wide) and 32k is once again imminent.

Increased system performance also starts to show the warts in our existing toolchain. Even with a responsive developer, getting changes in 3rd party packages takes time.

The solution

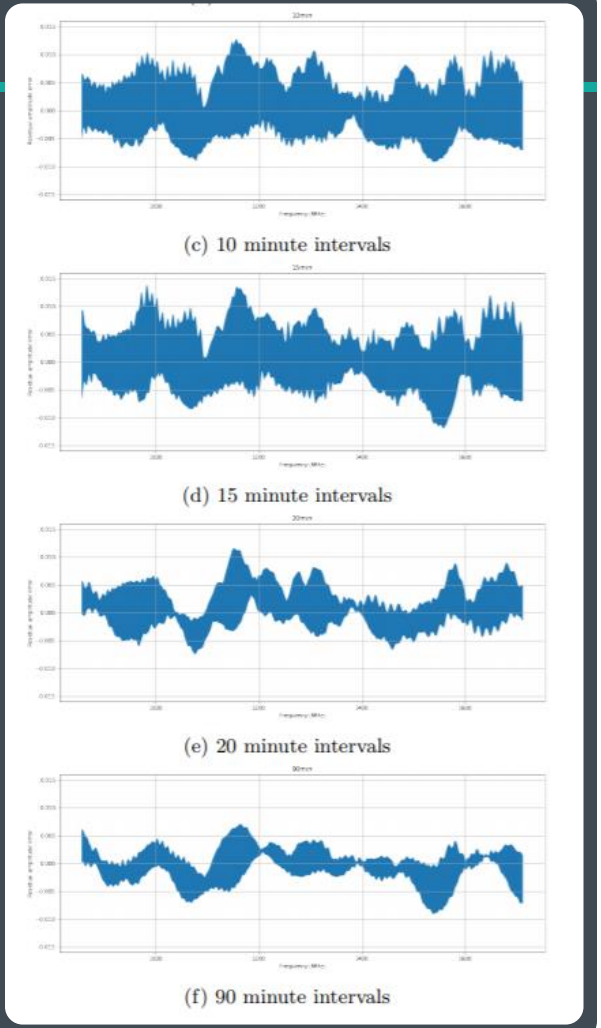
(this time we do have something)

Control your toolchain as far as practicable

Really helpful having science commissioning, science processing and RARG (Radio Astronomy Research Group) in a single unit.

Allows us to push in multiple directions simultaneously, looking for both performant and high quality solutions.

Particularly for new algorithmic or quality issues this approach has shown real utility.



(c) 10 minute intervals

(d) 15 minute intervals

(e) 20 minute intervals

(f) 90 minute intervals

PKS 0407 Substructure
Credit: B.Hugo et al

MeerKAT Data Products (this month)

Continuum image pipeline: baseline subtraction, self-cal solutions, and best effort images.

Spectral line pipeline: image cubes (up to 100 hrs joint)

Reduction for QA is done as standard for qualifying observations (currently 15mins on target for continuum / 3 hours for spectral)

Full res spectral cube is 24 TB

Once stable for QA, effort switches to improving science quality under guidance from LSPs.

Deployed Infrastructure

Realtime Cluster

Ingest + Calibration

100 TFLOPs, 6TB RAM

Batch Cluster

Spectral + Continuum Imaging

1.1 PFLOPs, 4TB RAM, 0.8 PiB buffer

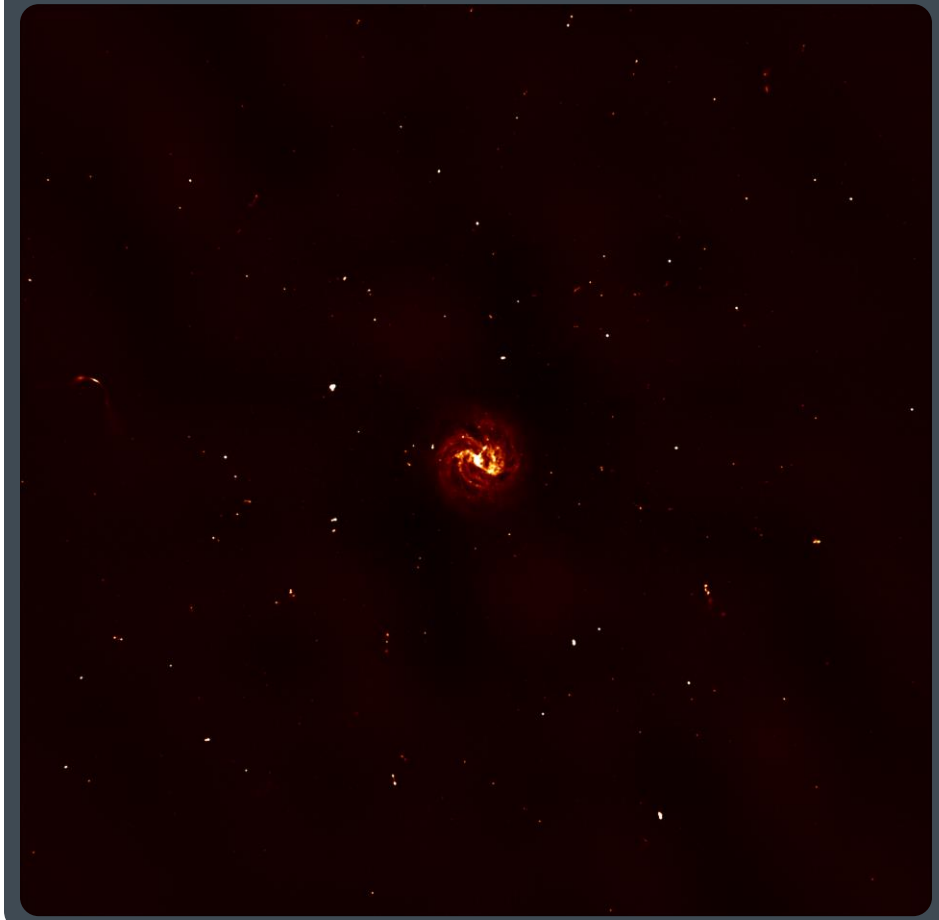
Storage

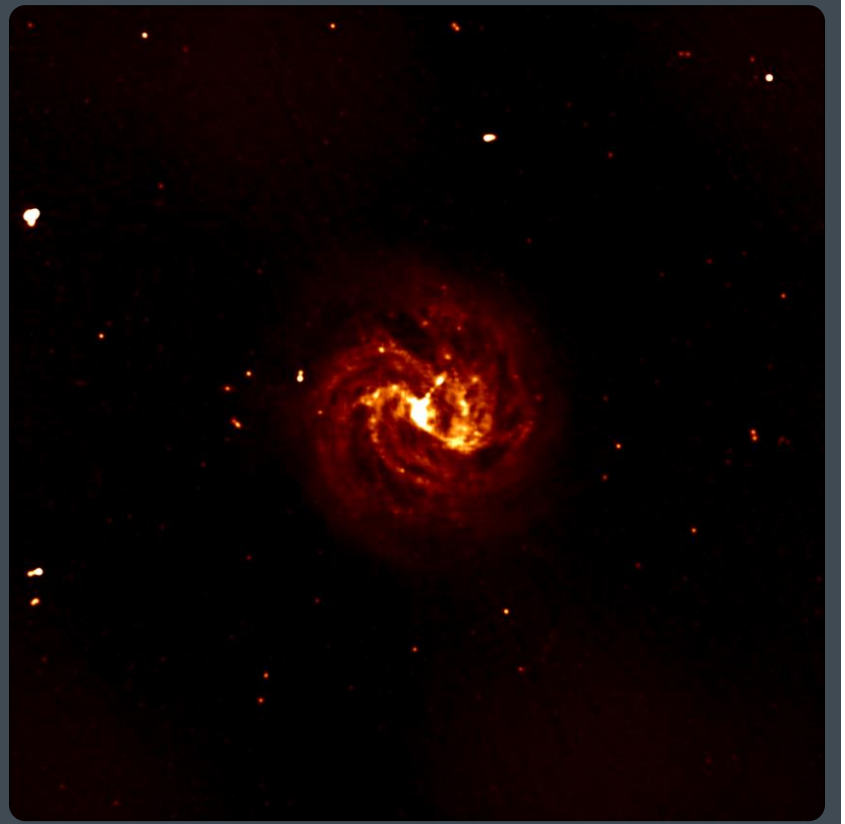
Vis Data + Science Products

object store: 20.5 PiB

tape: 19.2 PiB

Our first hands-free image (M83)





Relentless Imaging

(the first 7 days)

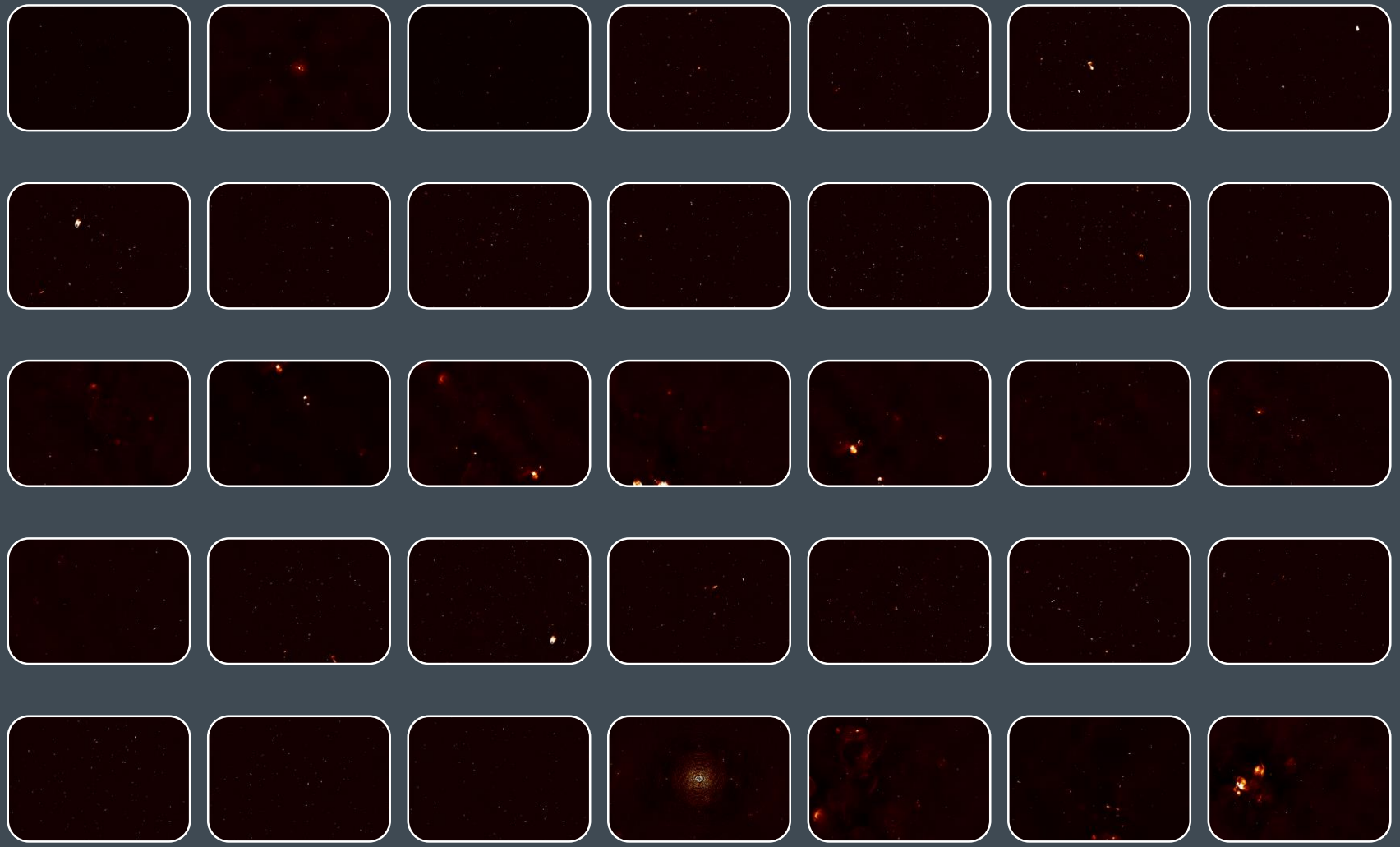
Imaging Pipeline

Deployed into production on 11 Nov. First week saw 76 fields imaged and in the archive.

Average time to image in archive: 1.4x target duration

No flux calibration (yet – check again next week)

Only a handful of spectral cubes so far – mostly limited by target time (< 3 hours)



So what next…

Commissioning starts at first light and pretty much never ends.

Tight integration with engineering activities very helpful, but can be frustrating for commissioning group.

Don’t underestimate the tooling needed, both from a scale and performance aspect.

Backpressure on feature development is important – the “greater good” needs a high level view.