The Capture, Storage & Consumption of the MeerKAT Radio Telescope's Sensor Data

Suleiman Hoosen
Software Site Support Engineer - SARAO





Objectives

- MeerKAT Telescope
- The Control and Monitoring (CAM) Sub-System
- CAM Tech Stack for MeerKAT
- KATCP
- Sensors & Samples
- Katstore64
- Future Endeavours





MeerKAT Telescope

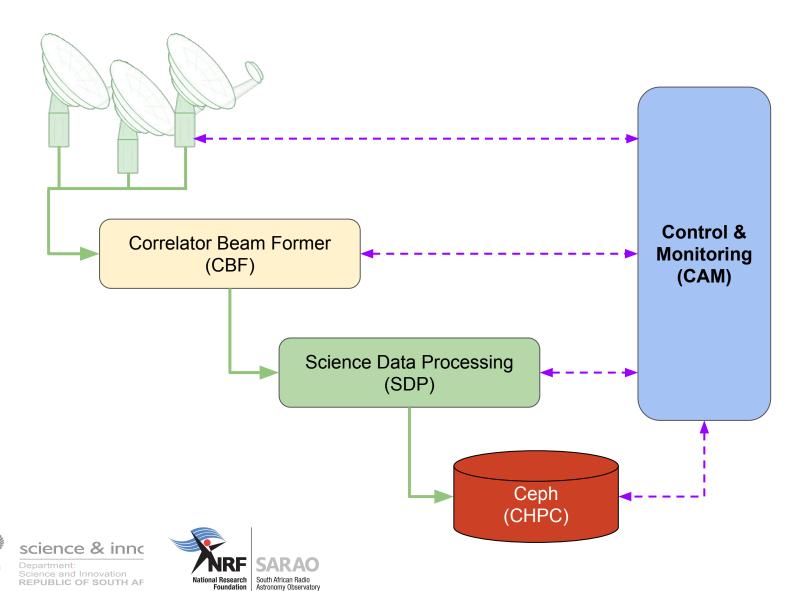
- 64-dish Radio Telescope
- Site: Losberg, Northern Cape
- 13.5m Diameter Dishes with Offset Gregorian Config
- 70% Antennas are in the Core (1km)
- Designed and Built by SKA-SA
- Precursor to SKA1
- Currently the most sensitive Radio
 Telescope in the world







The Control and Monitoring Sub-system



CAM Tech Stack for MeerKAT

















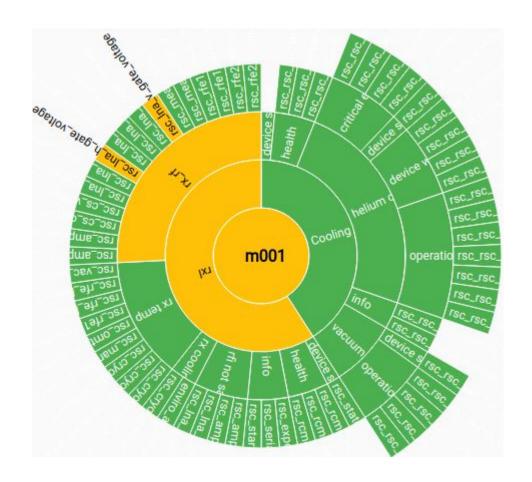








Receptor Health







Weather







KATCP

- Karoo Array Telescope Communication Protocol
- Simple ASCII Protocol
- Layered on top of TCP/IP Application Layer
- Preferred mechanism for interprocess communication
- Supports flexible, run-time configuration





KATCP

- Dynamic Discovery
 - Monitoring point -> KATCP Sensor
 - Control command -> KATCP Request
- CAM implements Device Translators for external components that do not have a KATCP Interface
- https://katcp-python.readthedocs.io/en/latest/





Sensors

Sensor Types

- Integer
- Float
- Boolean
- Timestamp
- Discrete
- Address
- String

Sensor Statuses

- Unknown
- Nominal
- Warn
- Error
- Failure
- Unreachable
- Inactive





Sensor Samples

- Uniquely identified by sensor name and sample time
- Any additional fields are optional
- JSON format
- Sampling Strategy
 - Period (Fixed Time)
 - Event (Value Change)





Sample Example

```
"sensor"
            : "anc_mean_wind_speed"
"sample_time": 1568174400.458403,
"value"
            : 2.1362775,
"status" : "nominal",
"value_time" : 1568174399.487174,
```





Katstore64

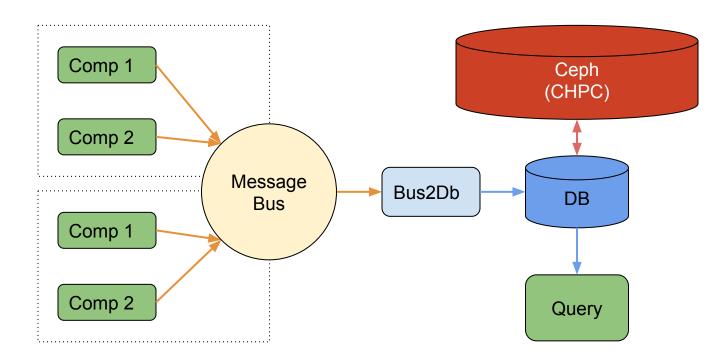
- Development began in 2016
- Deployment to Site in September 2019

- Time Series Data Store
- Fixed Index on Time
- Data is immutable





Katstore64



- NATS Messaging System
- A PostgreSQL extension
- FDW & Bus2DB Python3

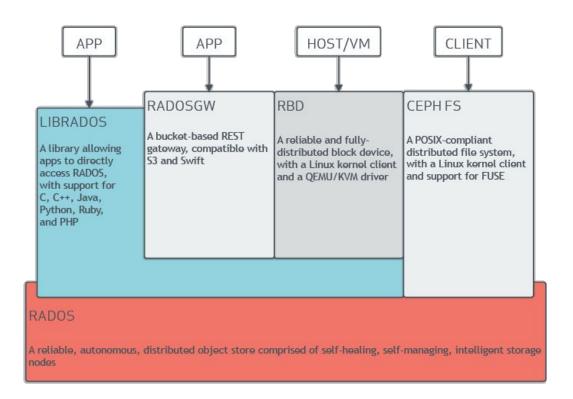
- Heavy lifting in PostgreSQL
- librados library
- Consistently process 40K s/s





Librados

- librados provides low-level access to the RADOS service







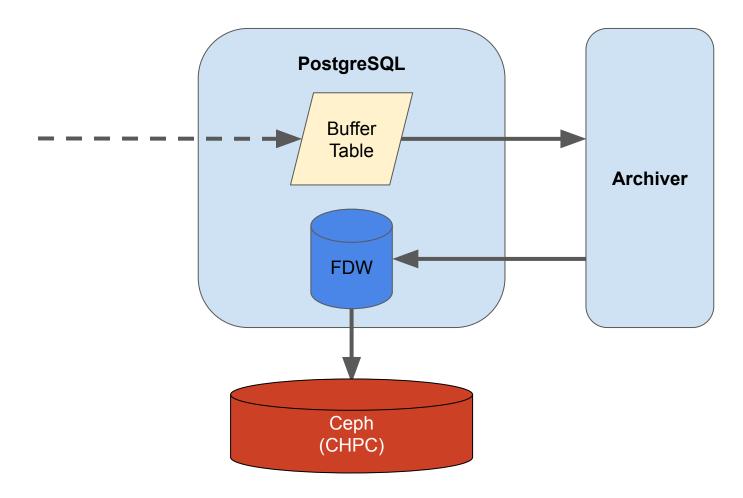
Katstore64 Samples Archive

- PostgreSQL Short Term Storage
- Ceph Long Term Storage
 - Distributed Storage System
 - Fault Tolerant
 - One Large Storage System many hard drives from multiple servers
 - Scalable to several petabytes
 - Leverages Science Archive Tech





Katstore64 Archive Interface







Structure of Unsynced Data

18233:anc_mean_wind_speed



Storage format:

- CBOR Concise Binary Object Representation
- Compression: Blosc using zstd





Repacking the Data

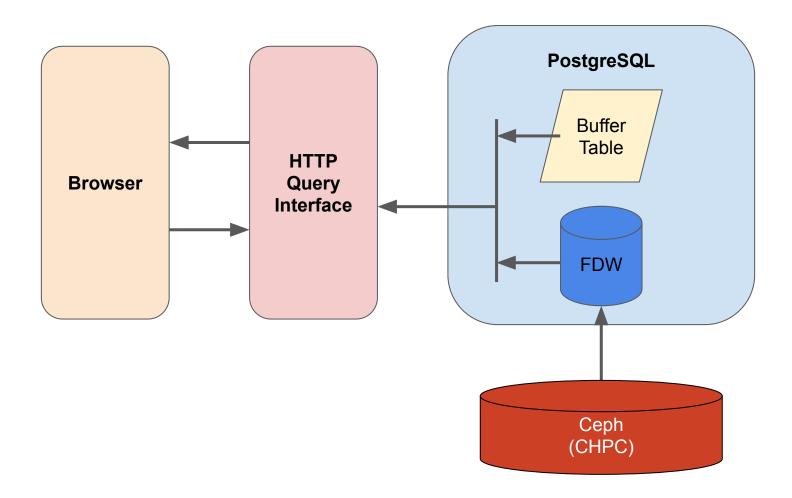
18233:anc_mean_wind_speed

			Data
	_	Header	Data
	Frame		Data
			Data





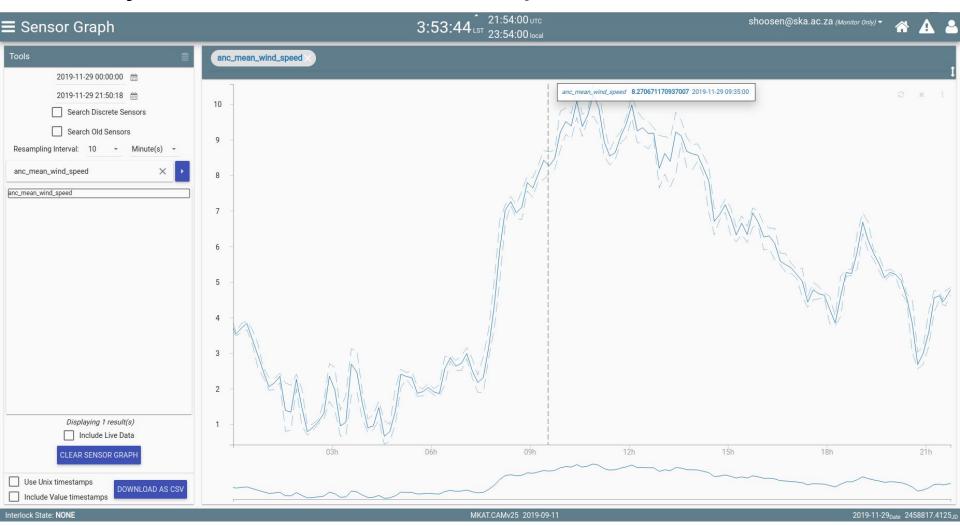
Katstore64 Retrieval Interface







Query Interface - Sensor Graph







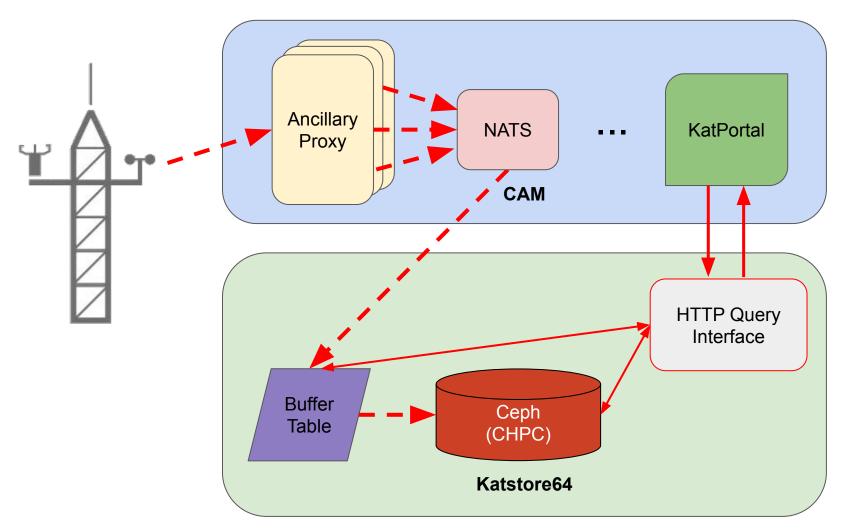
Query Interface - Katstore64 API

```
// 20191130000648
2
       // http://portal.mkat.karoo.kat.ac.za/katstore/api/query/?include_value_time=True&start_time=1574985600&sensor=anc_mean_wind_speed&end_time=1575065128
3
         "title": "Sensors Query",
          "url": "/katstore/api/query/?include_value_time=True&start_time=1574985600&sensor=anc_mean_wind_speed&end_time=1575065128",
         "sensor_name": "anc_mean_wind_speed",
          "data": [
9
10
              "sample_time": 1574985600.639859,
11
              "value": 3.8893429167,
              "status": "nominal",
13
              "value_time": 1574985600.602961,
14
              "sensor": "anc_mean_wind_speed"
15
           },
16
17
              "sample_time": 1574985601.640297,
              "value": 3.8890570833,
18
19
              "status": "nominal",
20
              "value_time": 1574985601.602876,
              "sensor": "anc_mean_wind_speed"
23
24
              "sample_time": 1574985602.639827,
25
              "value": 3.8883483333,
26
              "status": "nominal",
27
              "value_time": 1574985602.602997,
28
              "sensor": "anc_mean_wind_speed"
29
30
31
              "sample_time": 1574985603.647245,
32
              "value": 3.8875546056,
33
              "status": "nominal",
34
              "value_time": 1574985603.603334,
35
              "sensor": "anc_mean_wind_speed"
36
37
              "sample_time": 1574985604.646989,
38
              "value": 3.8859975,
39
              "status": "nominal",
40
41
              "value time": 1574985604.602876,
```





Sensor Example







Katstore64 Performance

• Samples per Second: 6500, Benchmarked at 40,000

- Simultaneous Queries: 64, Benchmarked on Dev environment
- Return Rate per Second: 90,000 samples

Karoo to CHPC Throughput Speed: 10 Gbps

Total Storage Capacity

Karoo: 400 TBCHPC: 10 PB





Future Endeavours

- MeerKAT+: Adding 20 SKA Dish Antennas to MeerKAT
- Machine Learning Project using sensor data





Ceph BoF

Hosted by Martin Slabber and Sean February

Time: 15:30

Location: Heathrow





Info

References:

- https://katcp-python.readthedocs.io/en/latest/
- SCALABLE TIME SERIES DOCUMENTS STORE
 M.J. Slabber, F.J. Joubert, M.T. Ockards SKA SA, Cape Town, South Africa

Thank you

Suleiman Hoosen

Email: shoosen@ska.ac.za



