Scale Challenges of the MeerKAT Radio Telescope

Thomas Bennett

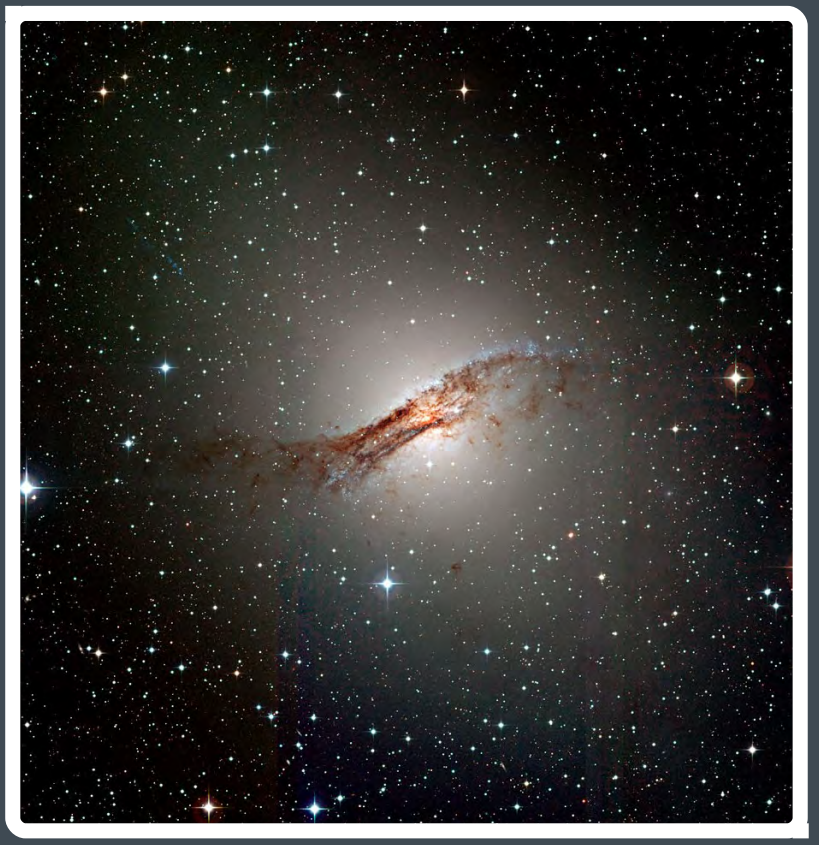
MSST 2018

Mesozoic observational astronomy

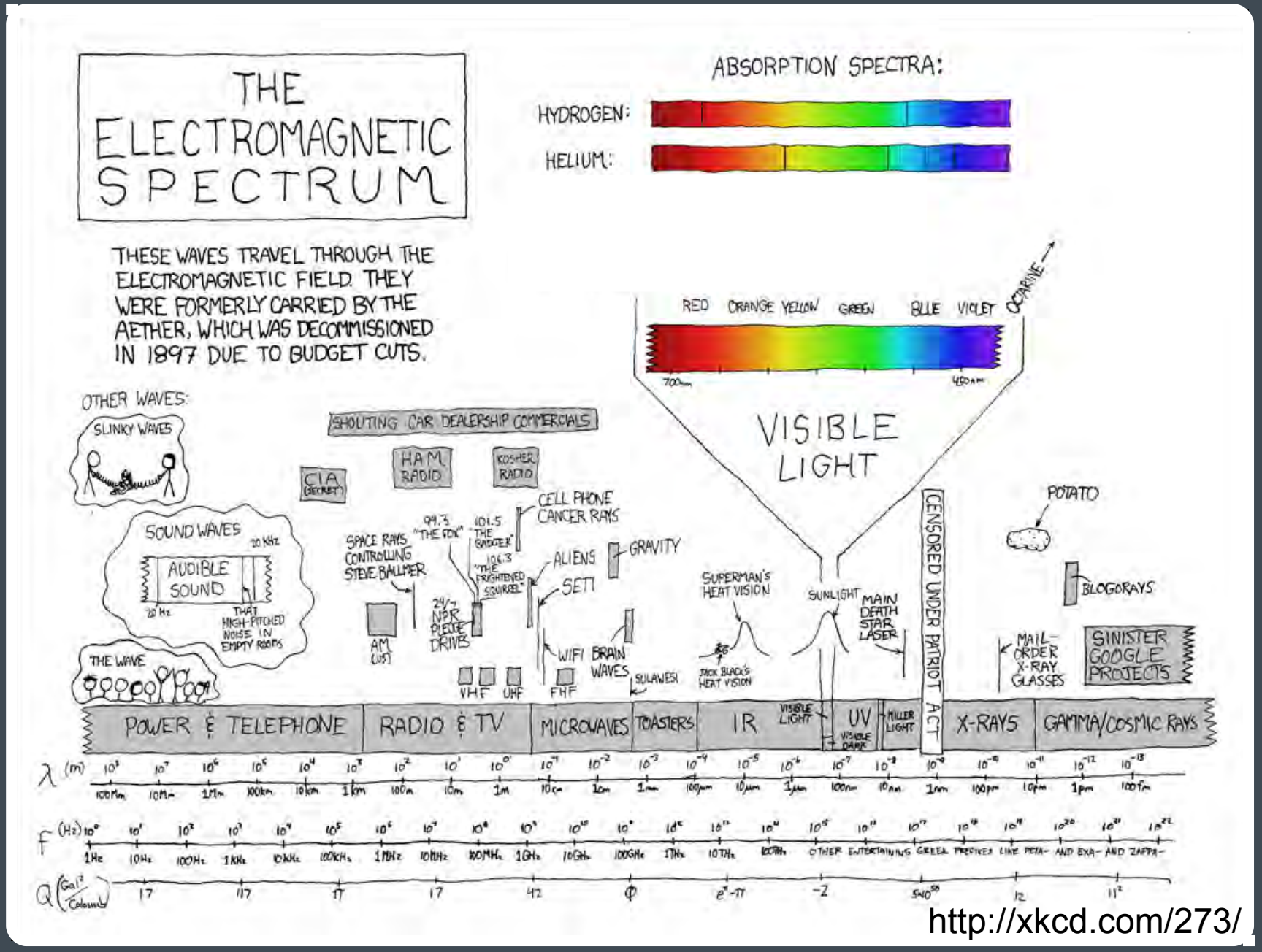
Cenozoic observational astronomy

ESO/MPG 2.2-m telescope, La Silla

=



Everything we know so far



http://xkcd.com/273/

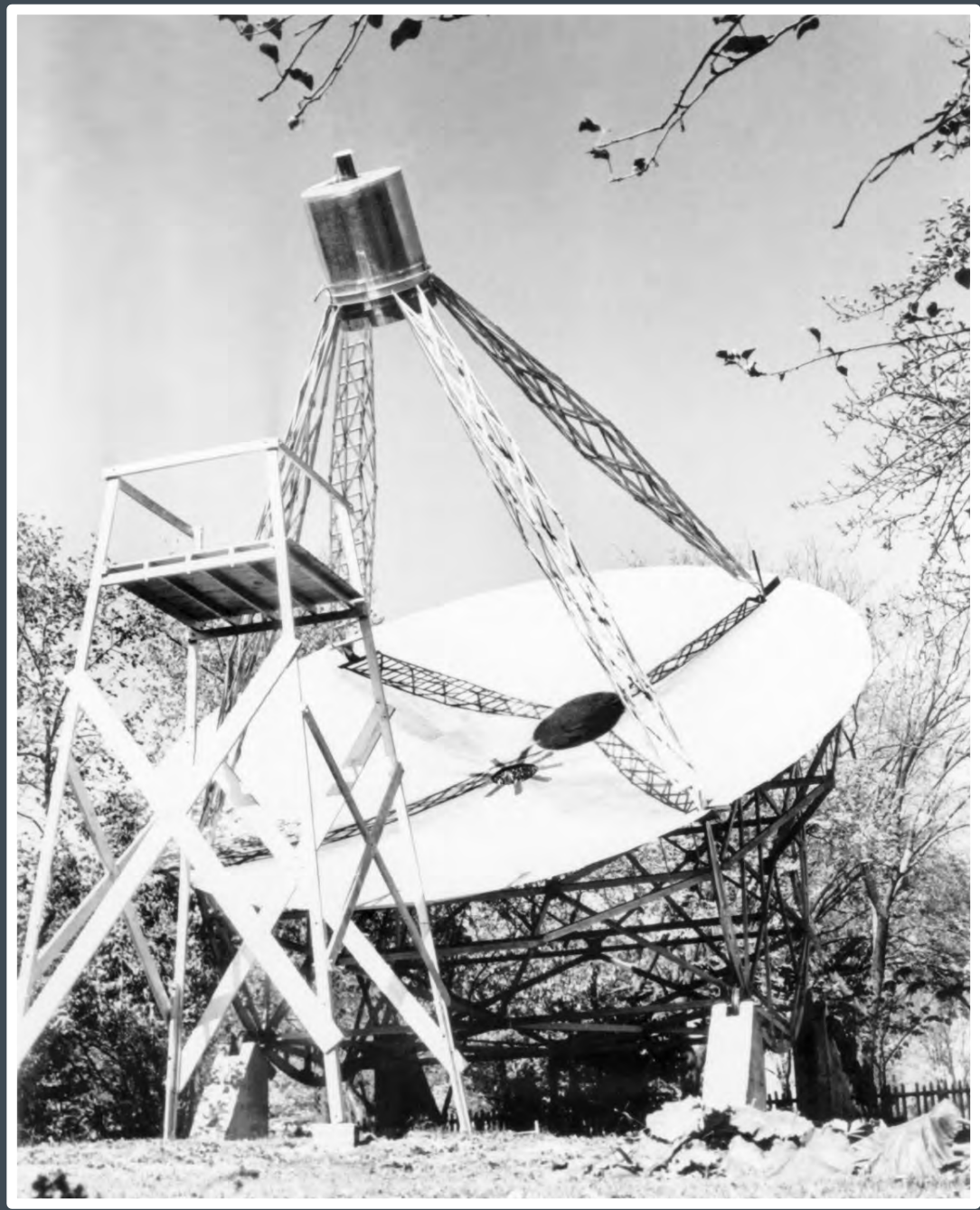
A fuller picture

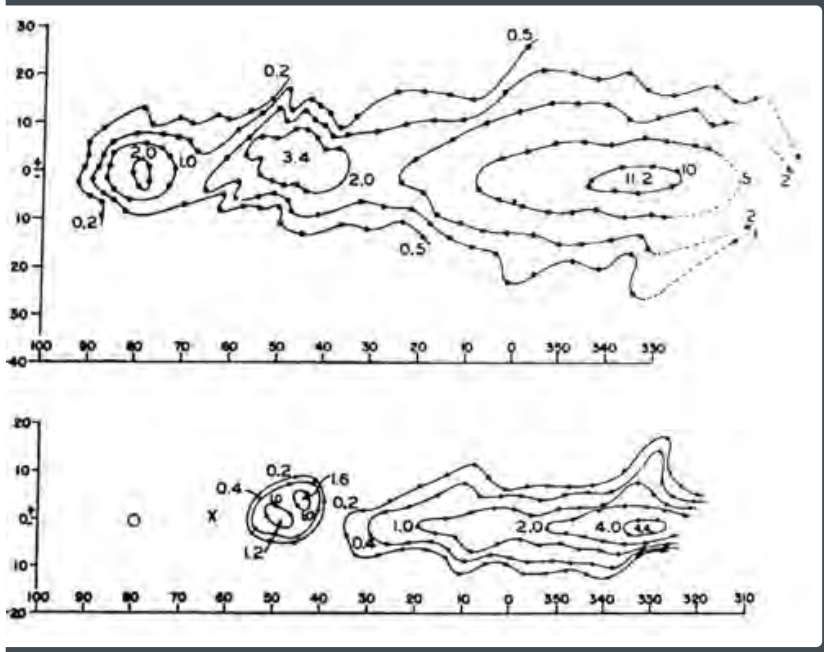
=

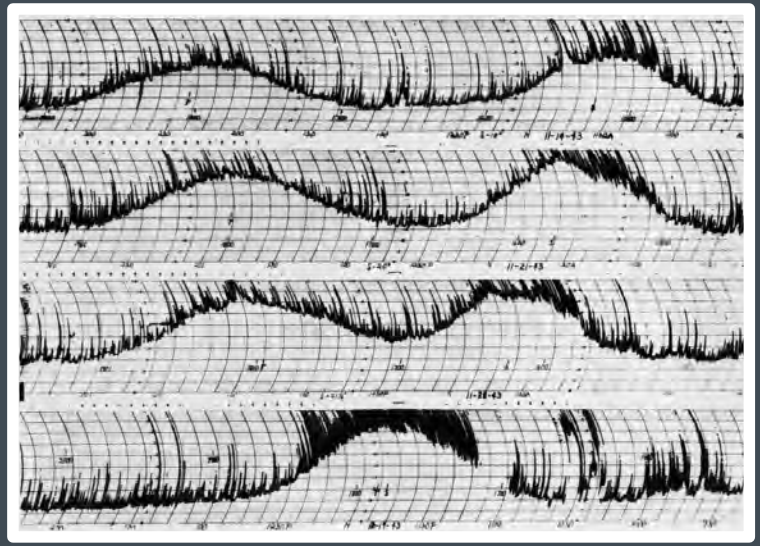
+

Jansky VLA, New Mexico

=



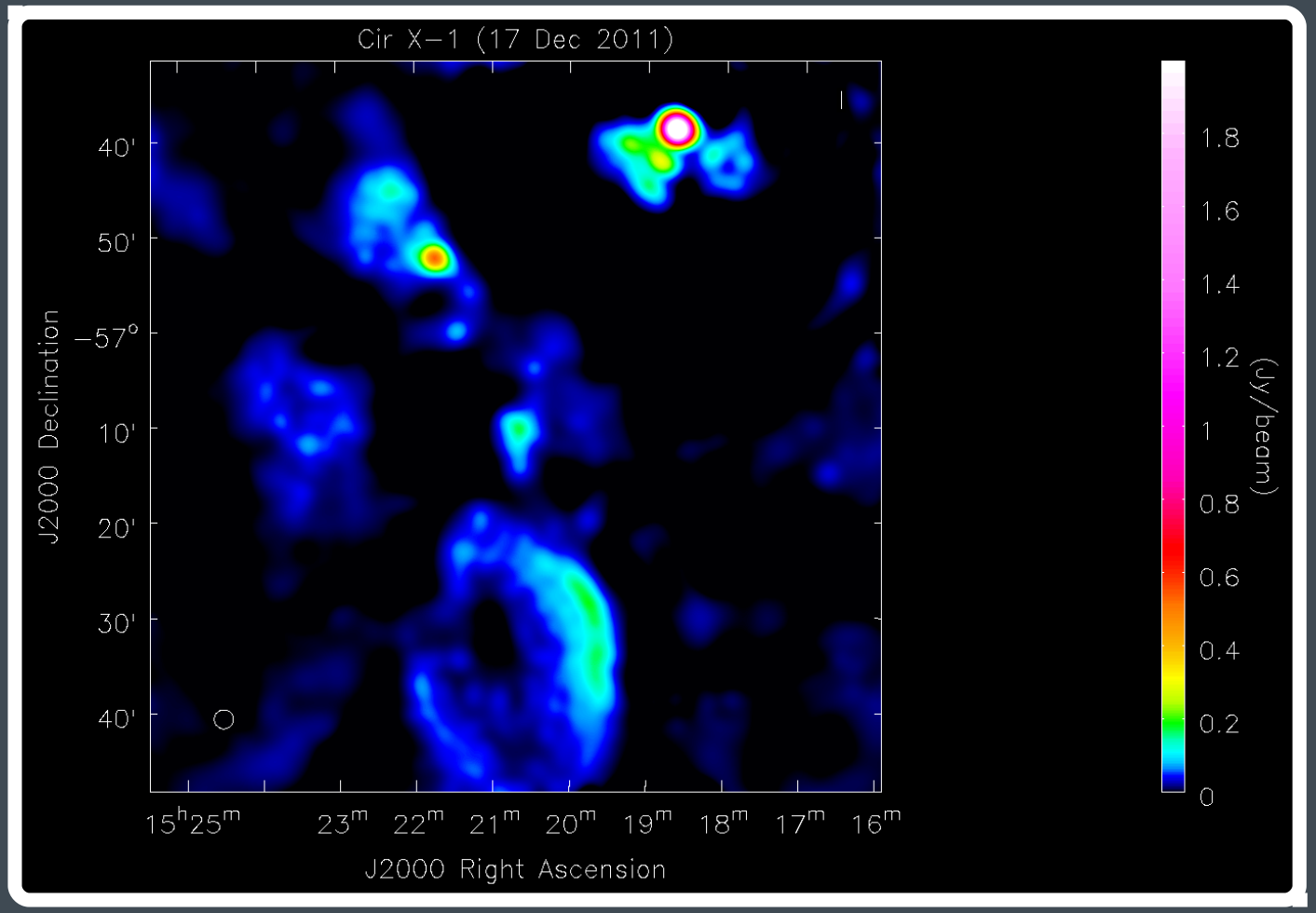




Grote Reber 1937

Watt?

$$1\ \mathrm{Jy} = 10^{-26}\ \mathrm{W m^{-2} Hz^{-1}}$$

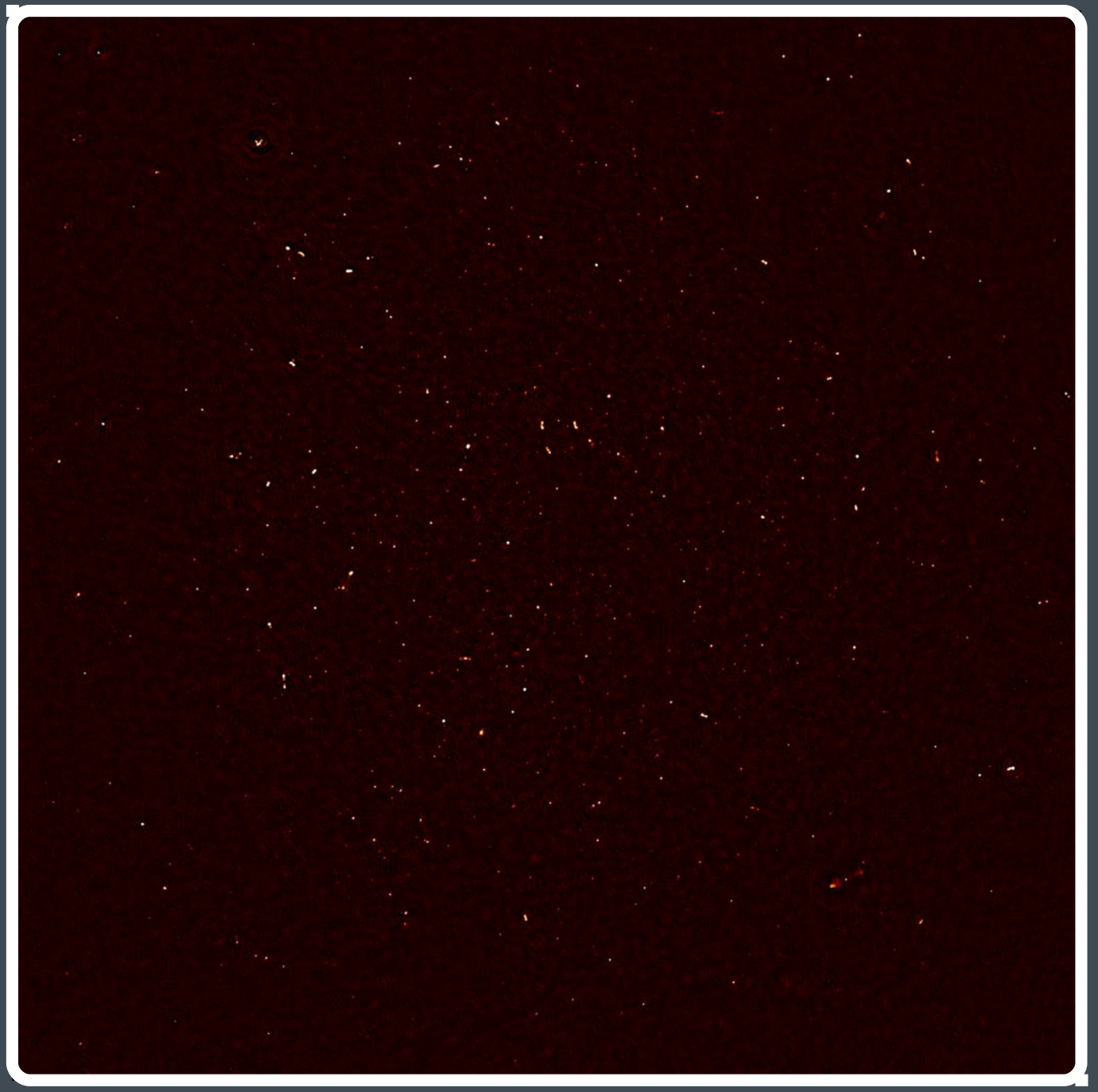


I cannae do it, captain, ye cannae change the laws of physics

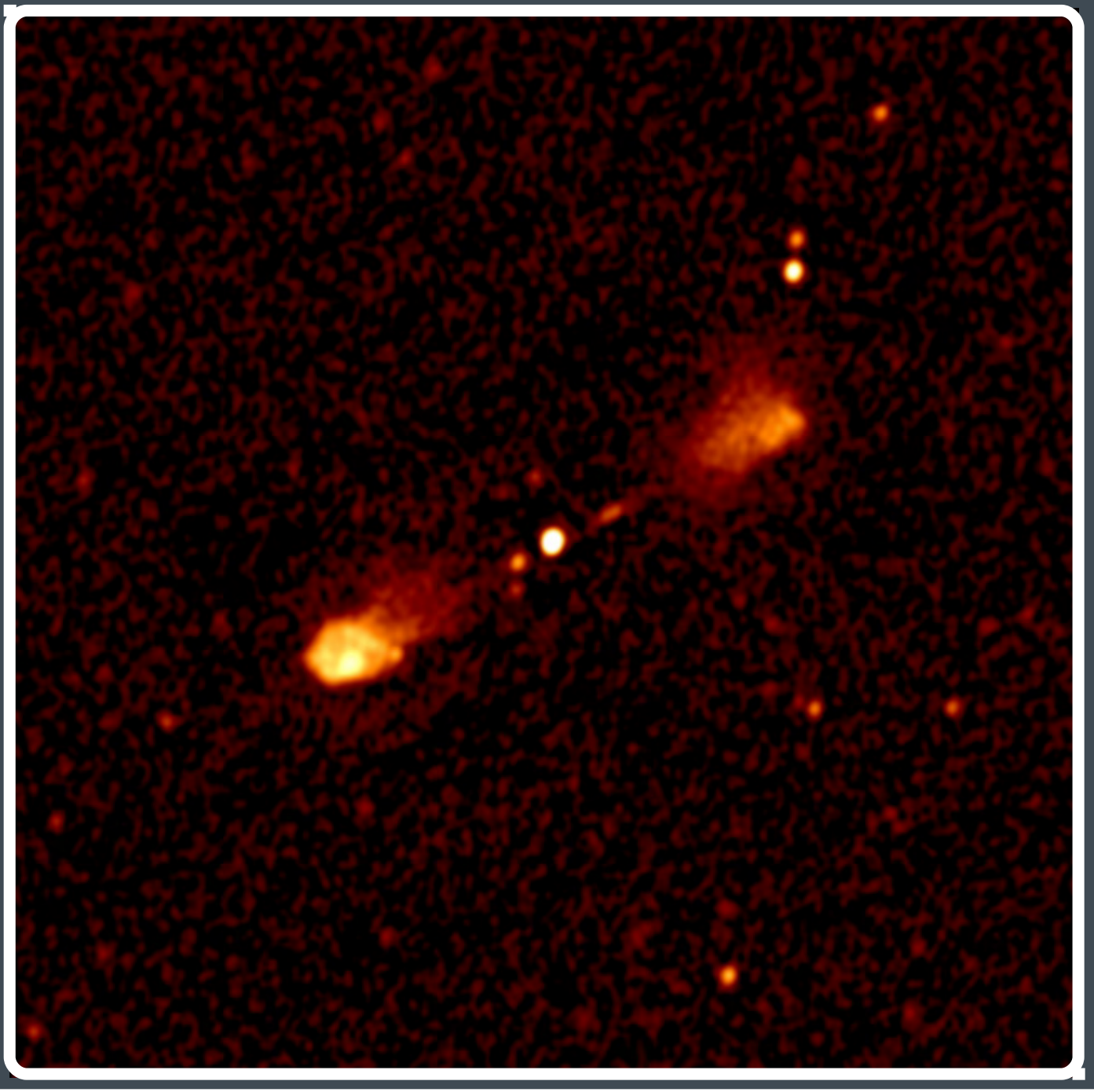
MeerKAT

64 antennas

1300 galaxies and counting



Expect the unexpected



The SKA

SKA Phase 1 (SKA1)

Cost: €650M, construction start ~~2017~~ 2019 2018

SKA1_MID

254 Dishes including:

64 x MeerKAT dishes

190 x SKA dishes

SKA1_LOW

Low Frequency Aperture Array Stations

SKA1_SURVEY

96 Dishes including:

36 x ASKAP

60 x SKA dishes

Exploring the Universe with the world's largest radio telescope

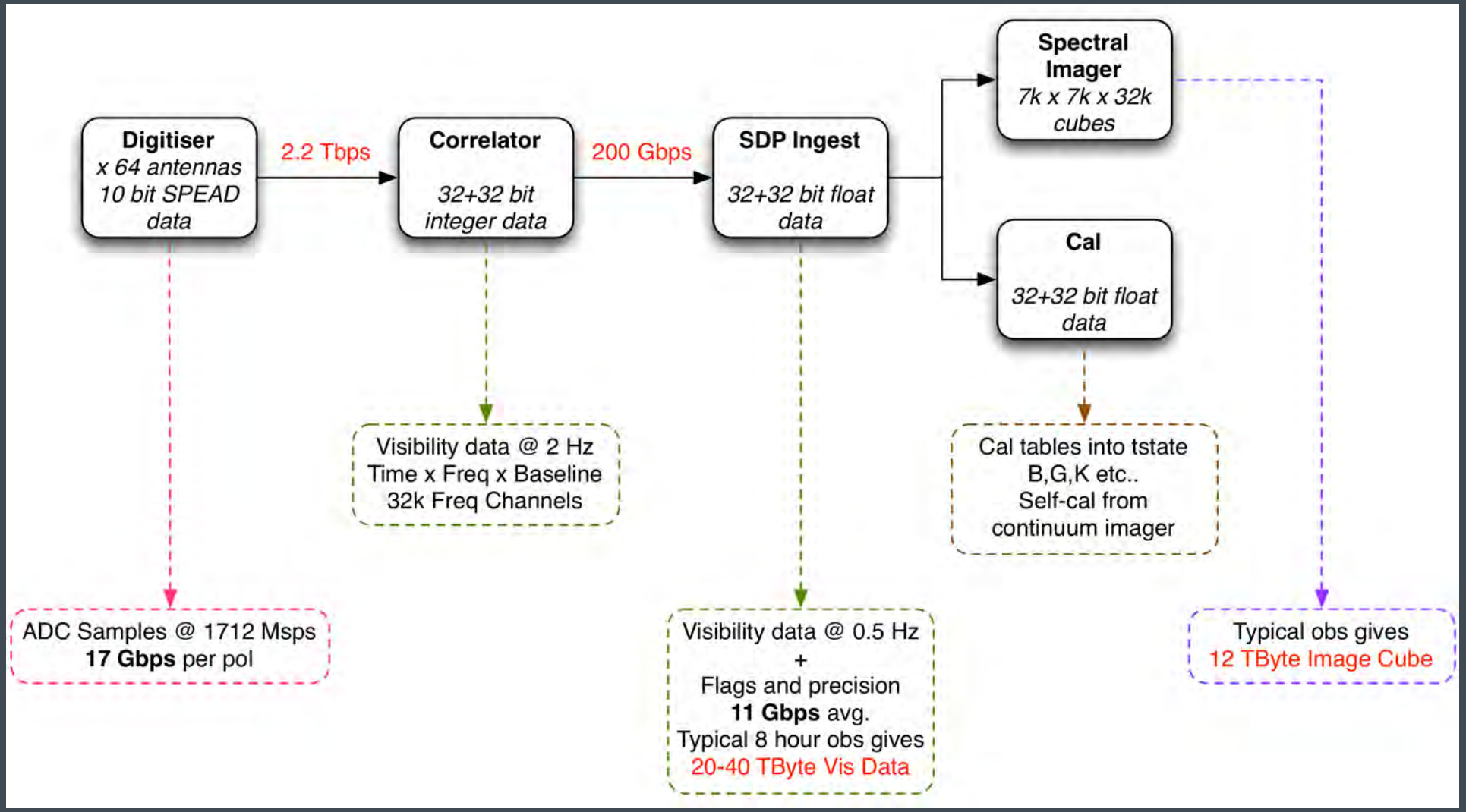
An African proposition

Radio Telescope - Deconstructed

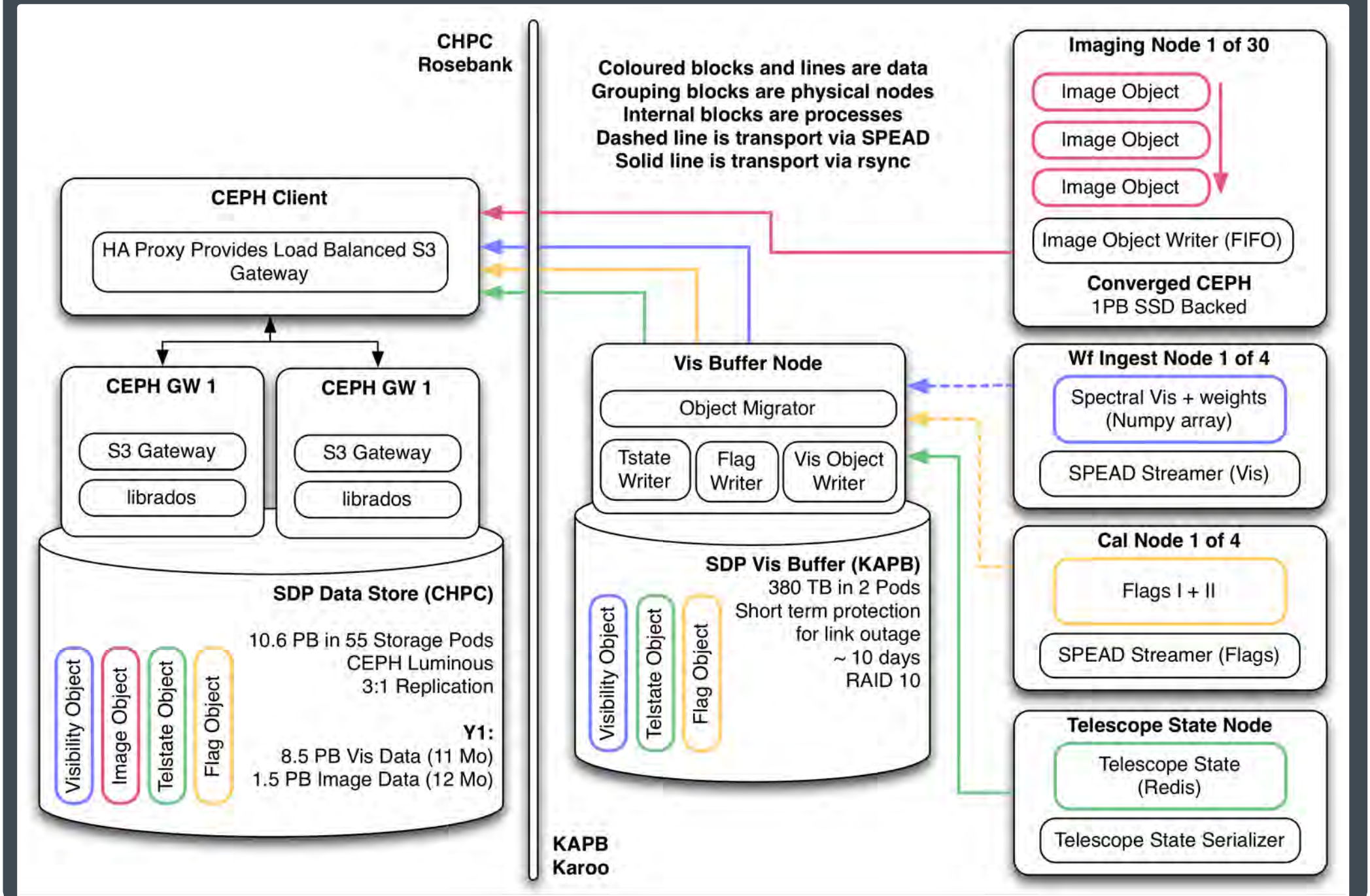
$$\vec{V}_{ij} = M_{ij} B_{ij} G_{ij} D_{ij} E_{ij} P_{ij} T_{ij} \vec{V}_{ij}^{\,IDEAL}$$

PB / PFLOPS / MW / G$

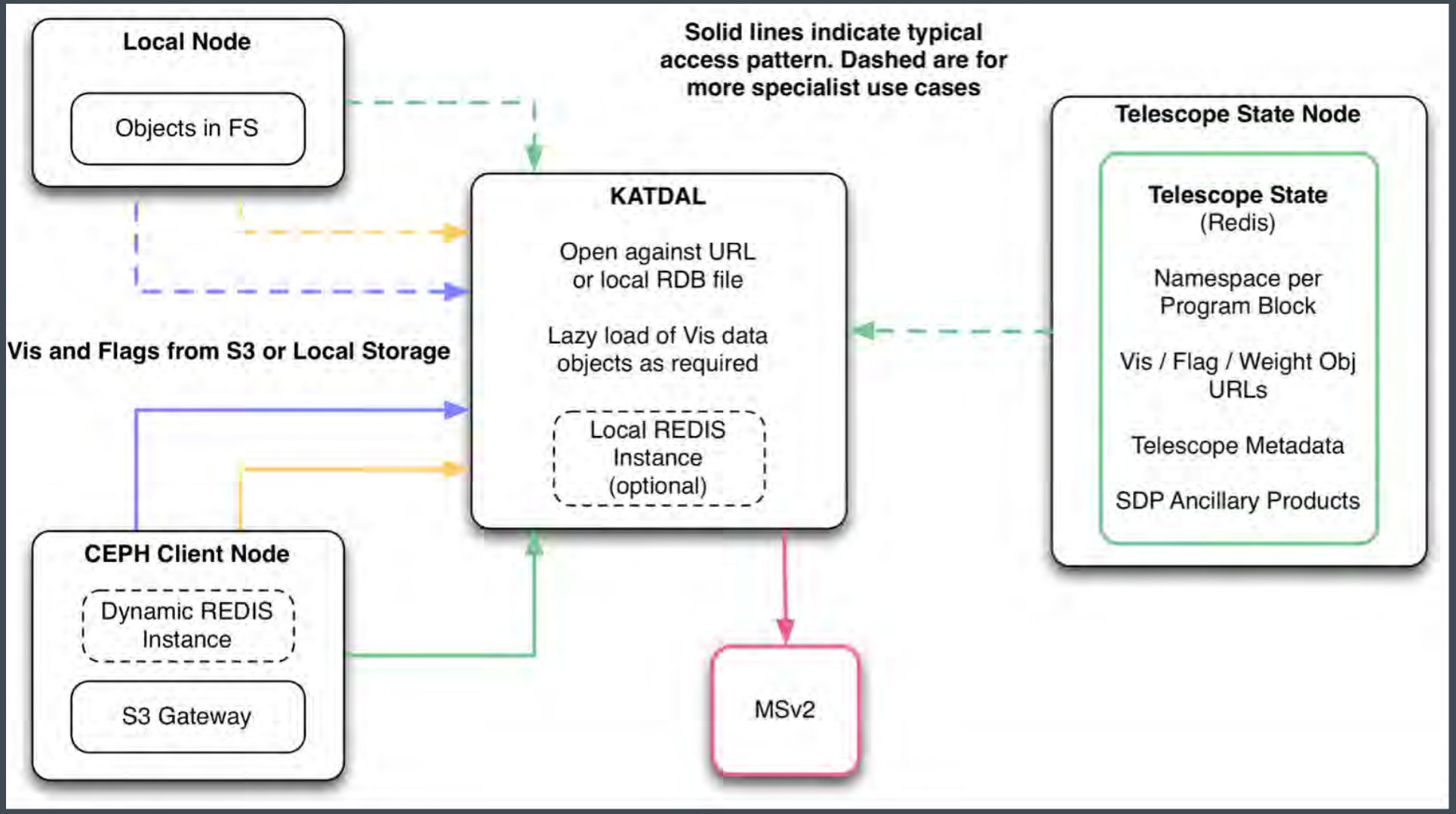
Data Rates


Digitiser
x 64 antennas
10 bit SPEAD data
2.2 Tbps
Correlator
32+32 bit integer data
200 Gbps
SDP Ingest
32+32 bit float data
Spectral Imager
7k x 7k x 32k cubes
Cal
32+32 bit float data
Visibility data @ 2 Hz
Time x Freq x Baseline
32k Freq Channels
Cal tables into tstate
B,G,K etc..
Self-cal from
continuum imager
ADC Samples @ 1712 Msps
17 Gbps per pol
Visibility data @ 0.5 Hz
+
Flags and precision
11 Gbps avg.
Typical 8 hour obs gives
20-40 TByte Vis Data
Typical obs gives
12 TByte Image Cube

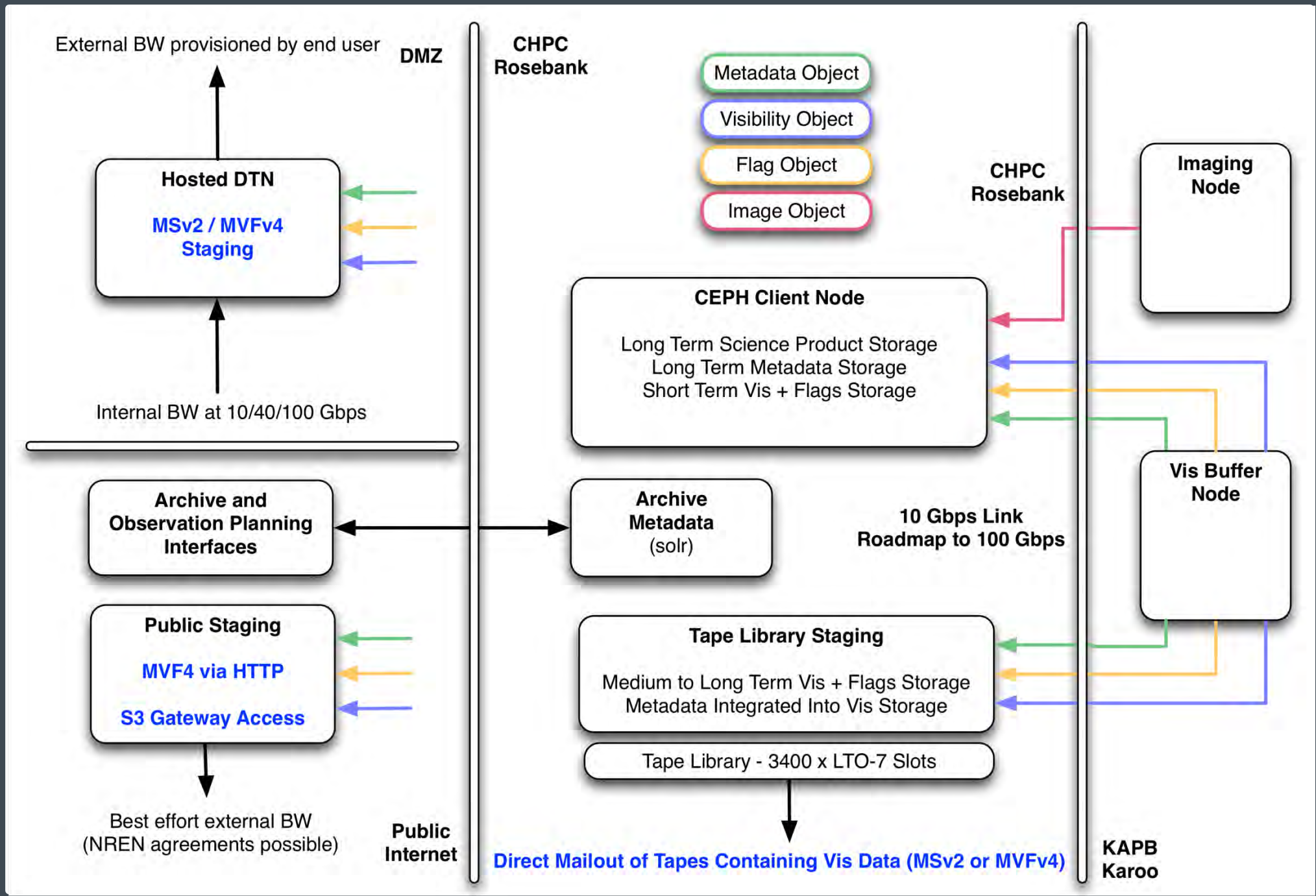
Data Storage and WAN Movement



Data Access Layer (KATDAL)

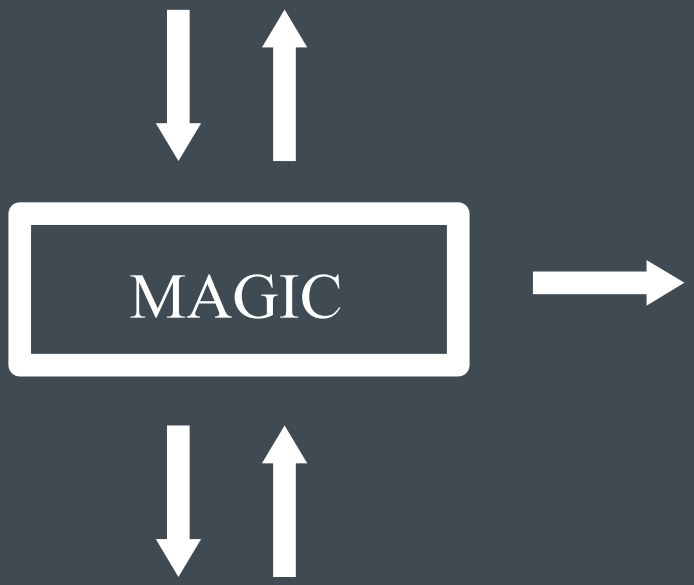


External (to site) data access



The Hardware Challenge

$$\vec{V}_{ij} = M_{ij} B_{ij} G_{ij} D_{ij} E_{ij} P_{ij} T_{ij} \vec{V}_{ij}^{\,IDEAL}$$

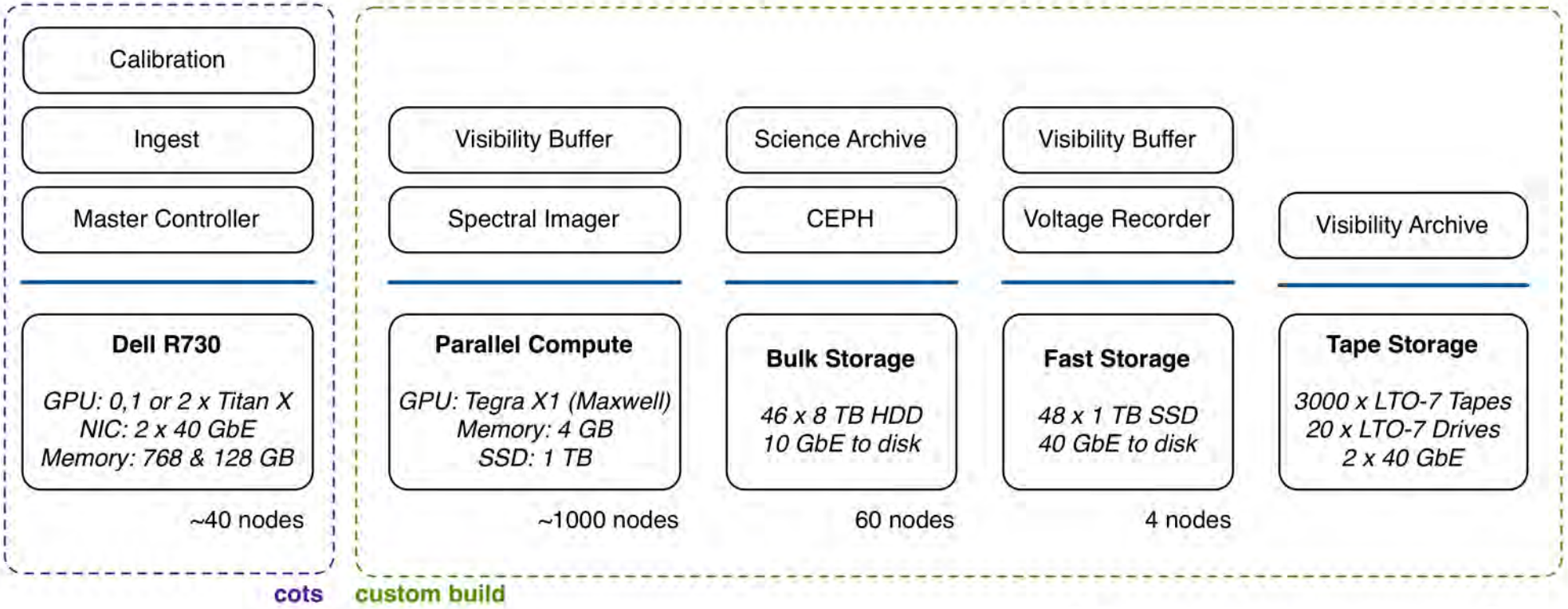


PB / PFLOPS / MW / G$

Republic credits are no good here...

So what to do...

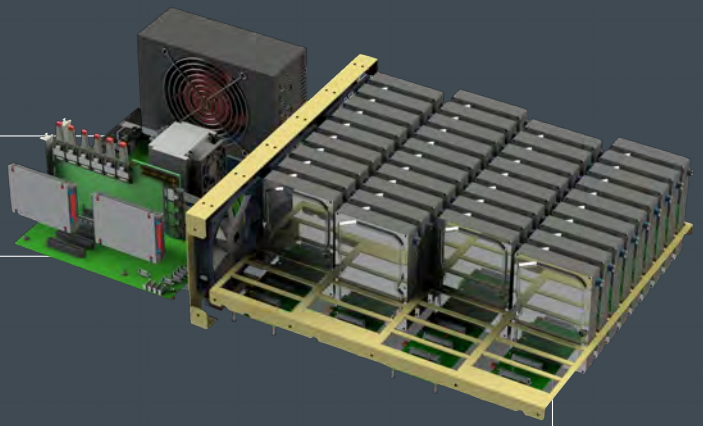
Self Build


Calibration
Ingest
Master Controller
Dell R730
GPU: 0,1 or 2 x Titan X
NIC: 2 x 40 GbE
Memory: 768 & 128 GB
~40 nodes
cots
Visibility Buffer
Spectral Imager
Parallel Compute
GPU: Tegra X1 (Maxwell)
Memory: 4 GB
SSD: 1 TB
~1000 nodes
Science Archive
CEPH
Bulk Storage
46 x 8 TB HDD
10 GbE to disk
60 nodes
Visibility Buffer
Voltage Recorder
Fast Storage
48 x 1 TB SSD
40 GbE to disk
4 nodes
Visibility Archive
Tape Storage
3000 x LTO-7 Tapes
20 x LTO-7 Drives
2 x 40 GbE
custom build

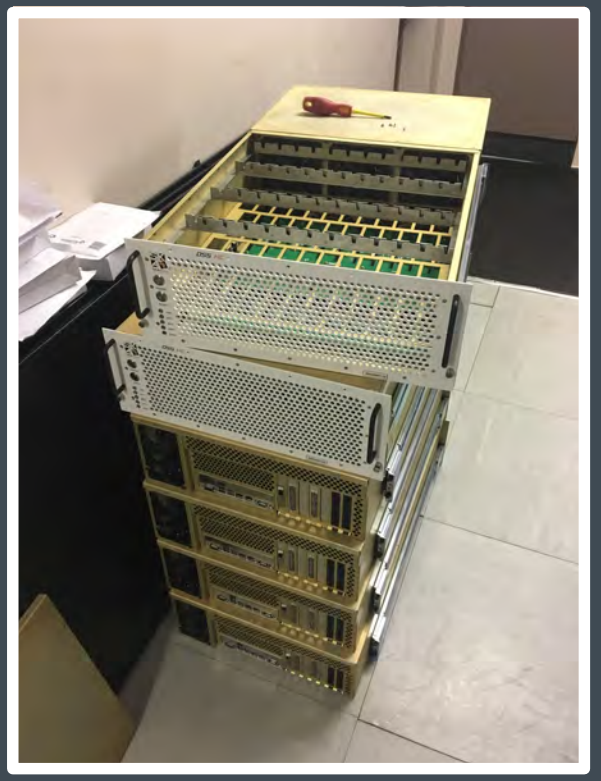
Back Blaze

Peralex Electronic Systems

DSS HC Storage pod	
KEY FEATURES	• 25 GbE (SFP+) • x48 SAS 3 connectivity
PROCESSOR	• Haswell-EP 4C Intel Xeon 4 core 3.7 GHz
MEMORY	• 64 GB per node
NETWORKING	• Mellanox Network Interface Card, 25GbE, Single-Port
DRIVE CONFIG	• 48 x 8TB HDDs • 2 x 512 GB nvme SSDs
FORM FACTOR	• 4U with 1500W PSU




20 PACK




20 PACK

Data Storage in Cape Town

Software Toolset

Ceph Luminous
MAAS - Metal as a service
Ansible (and Ceph-ansible)
Prometheus and Graphana
Apache proxy
SAML

Current Ceph implementation:

20% of final
3 monitors
12 OSD nodes
576 OSDs
25 Gigabit network
~ 4 PB of storage

Planned Ceph implementation:

End of June
3 monitors
55 OSD nodes
2640 OSDs
25 Gigabit network
~ 21 PB of storage

Shameless Advertising

http://www.ska.ac.za