Extreme Computing for the SKA Telescope

Peter Braam
Oct 2017

This talk

- What is the SKA telescope?
- Science cases for the SKA
- What is the Science Data Processor?
- Models of the computation
- Parallel software & system architecture

My role

- Computation project is led by Cambridge University
- I've been an advisor & consultant to SKA since
- Acknowledgement:
  - A large group of people (~500) are working on this project
  - Background is publicly available, e.g. SDP Preliminary Design Review
  - Much material comes from other SKA presentations & documents

What is the SKA?

The Square Kilometre Array (SKA)

Next Generation radio telescope – compared to best current instruments it offers ...

- ~100 times more sensitivity
- ~ 10^6 times faster imaging the sky
- More than 5 square km of collecting area over distances of >100km

Will address some of the key problems of astrophysics and cosmology (and physics)

- Builds on techniques developed originally in Cambridge
- It is an Aperture Synthesis radio telescope (“interferometer”)

Uses innovative technologies...

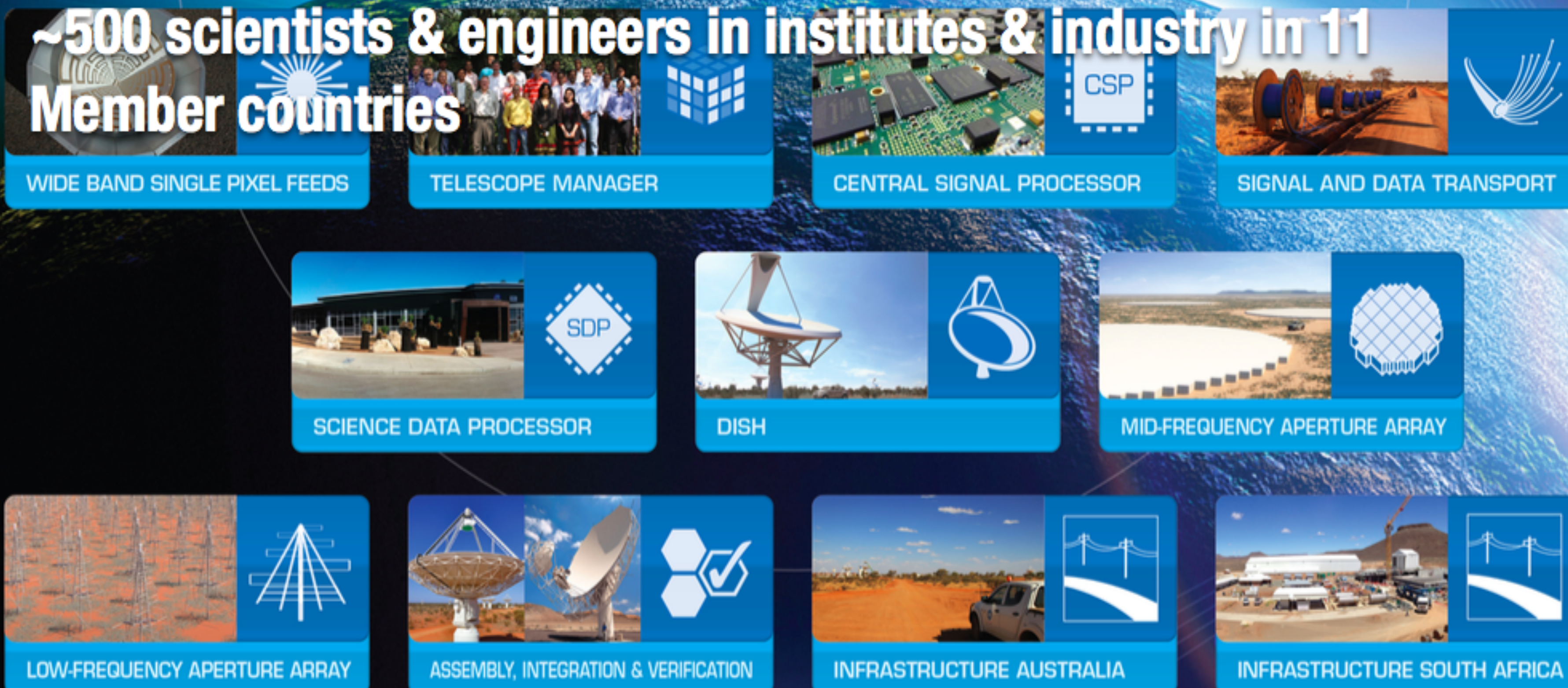
- Major ICT project
- Need performance at low unit cost

SPDO/Swinbur

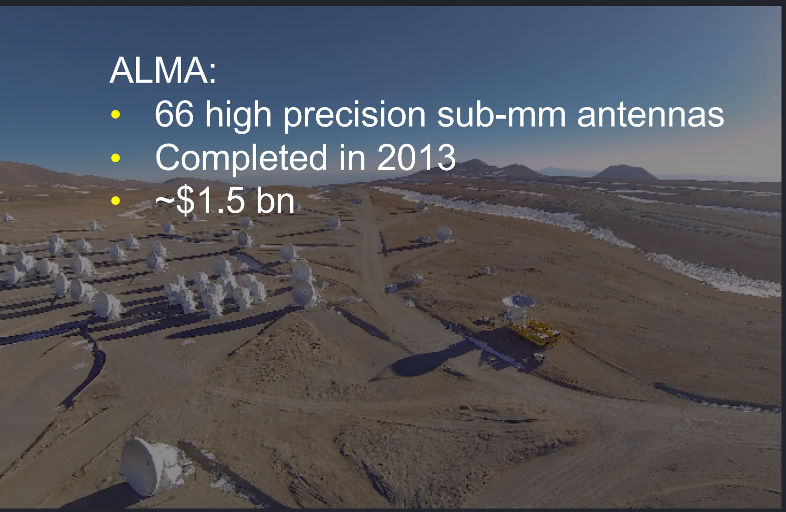
SKA International Design Consortia

Project Management and System Engineering Team based at JBO (UK)

~500 scientists & engineers in institutes & industry in 11 Member countries

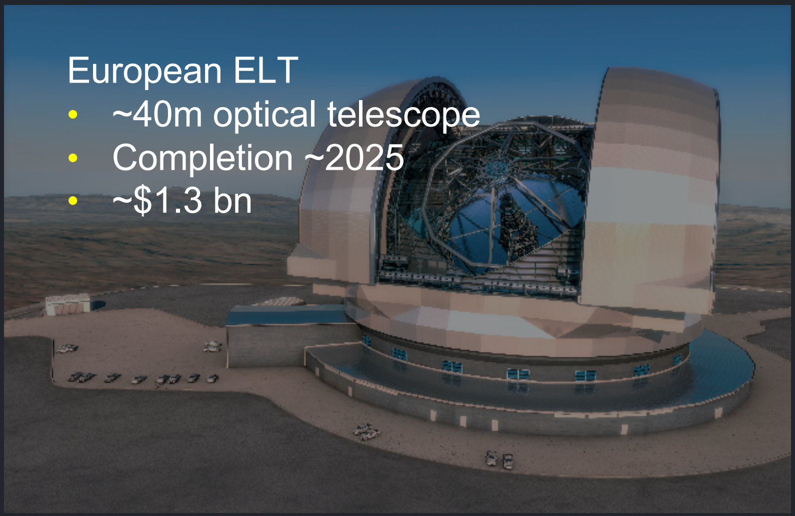


SKA – a partner to ALMA, EELT, JWST



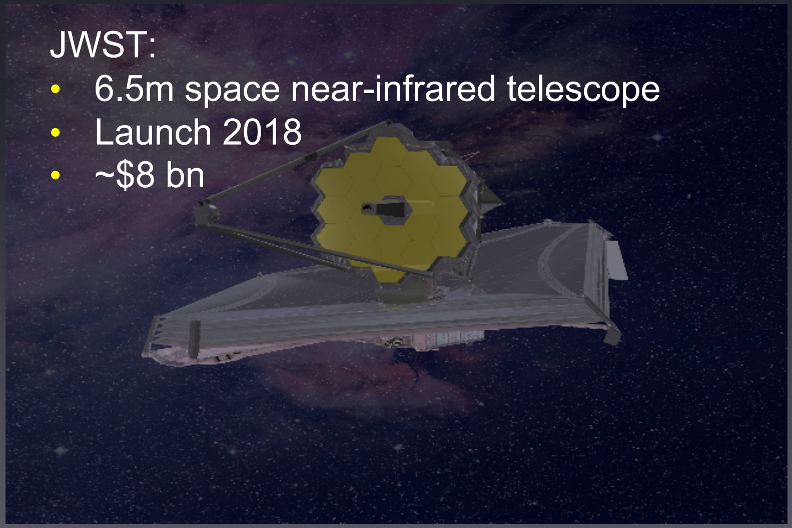
ALMA:
- 66 high precision sub-mm antennas
- Completed in 2013
- ~$1.5 bn

Credit:A. Marinkovic/XCam/ALMA(ESO/NAOJ/NRAO)



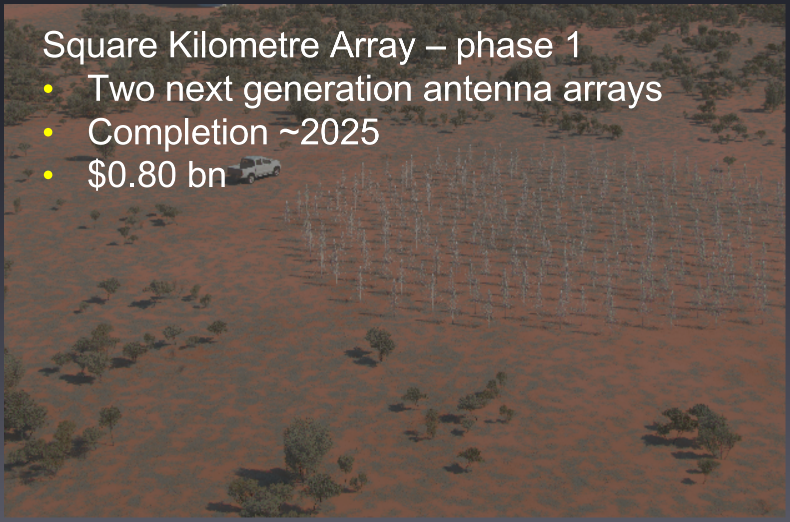
European ELT
- ~40m optical telescope
- Completion ~2025
- ~$1.3 bn

Credit:ESO/L. Calçada (**artists impression**)



JWST:
- 6.5m space near-infrared telescope
- Launch 2018
- ~$8 bn

Credit: Northrop Grumman (**artists impression**)

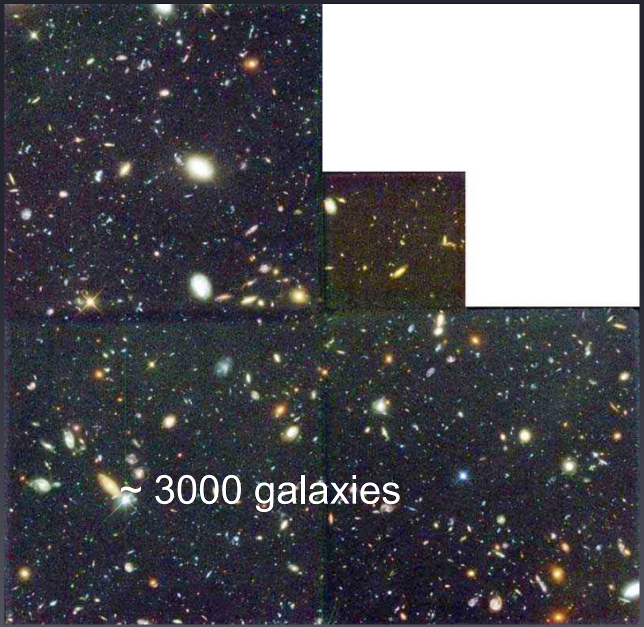


Square Kilometre Array – phase 1
- Two next generation antenna arrays
- Completion ~2025
- $0.80 bn

Credit: SKA Organisation (**artists impression**)

What will we see

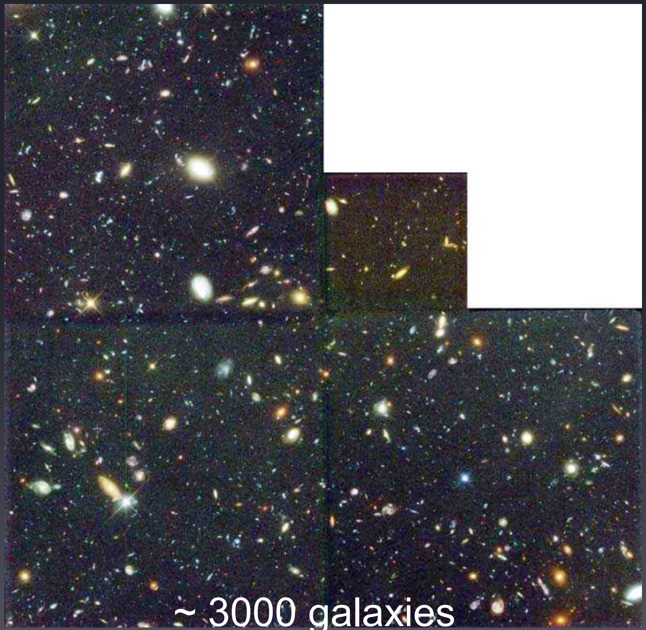
- Hubble Deep Field (HDF)



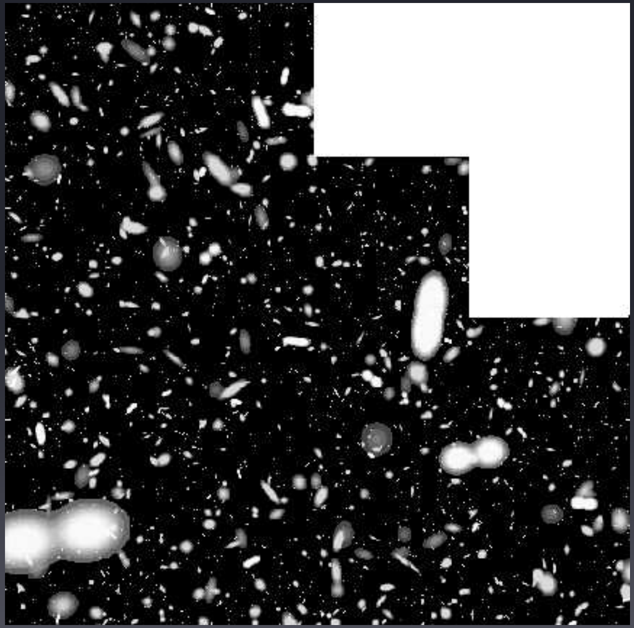
- Very Large Array

Hubble - SKA

Hubble Deep Field (HDF)



Simulation of SKA Observation



In summary …

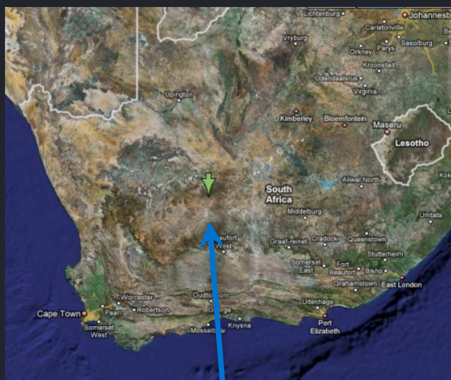
- SKA aims to be an “instrument” like CERN

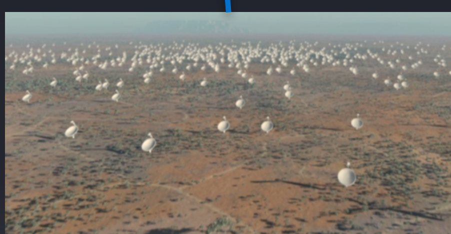
- This discussion focuses on SKA1 - 2025
  - Support from participating countries (budget ~$0.7B)
- SKA2 should have 10x more antennas – 2030?

- Caveat
  - Ongoing changes
  - Some inconsistencies in the numbers

  - Cf. NYT Oct 2016:
    “Maybe there are 100x more galaxies than we previously thought.”

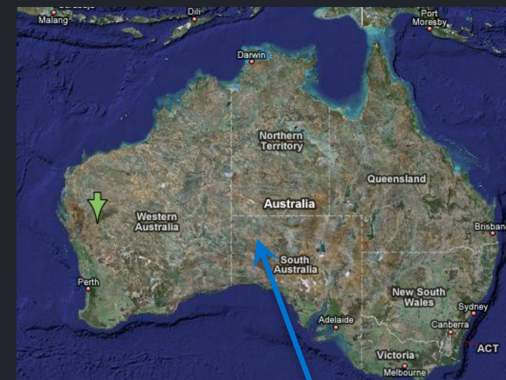
SKA1 Implementation





Mid Frequency Array

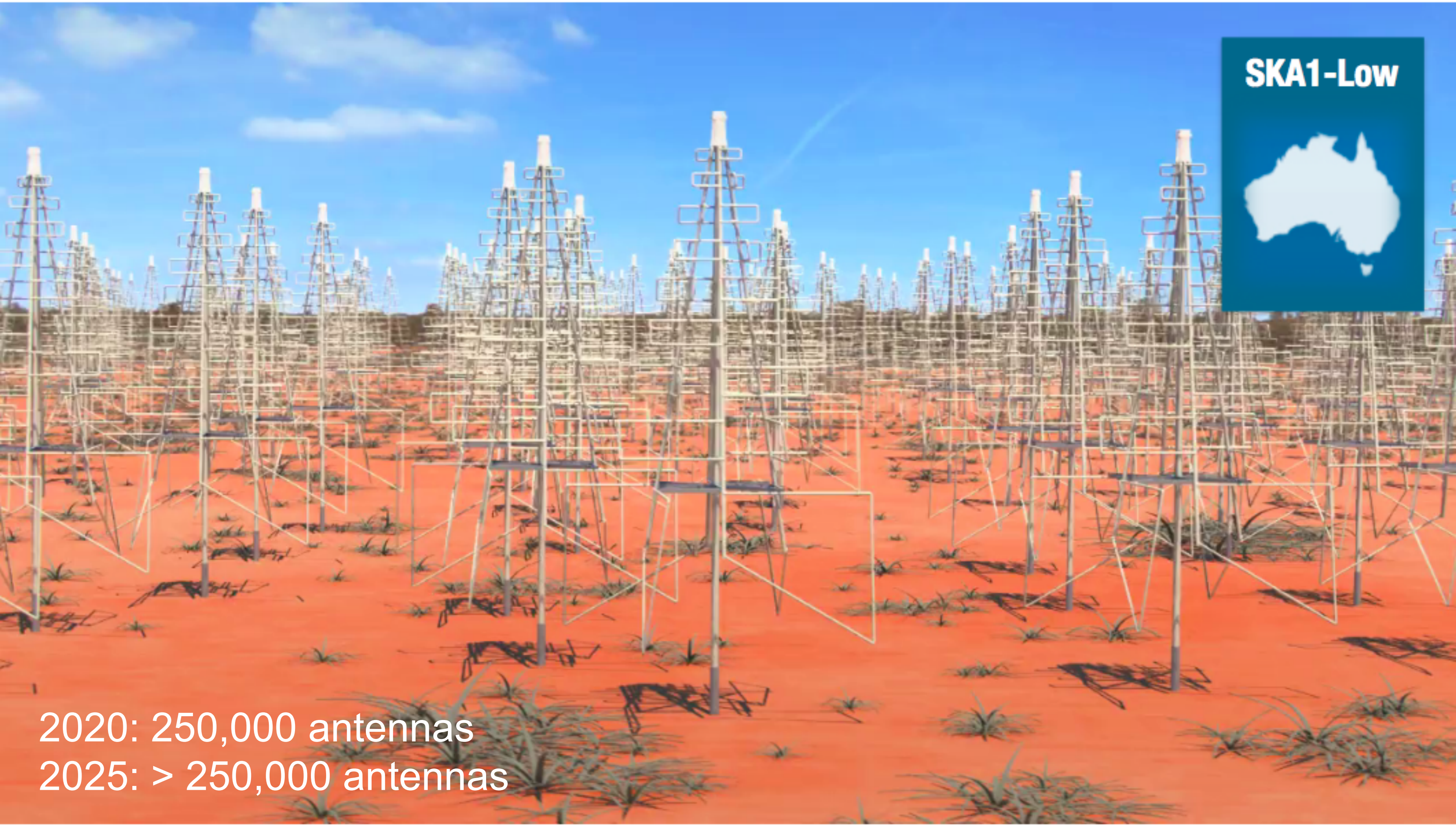
250 dishes with single receiver
Karoo, SA - 3 humans / sqkm
Compute in Cape Town (400 km)



Low Frequency Aperture Array

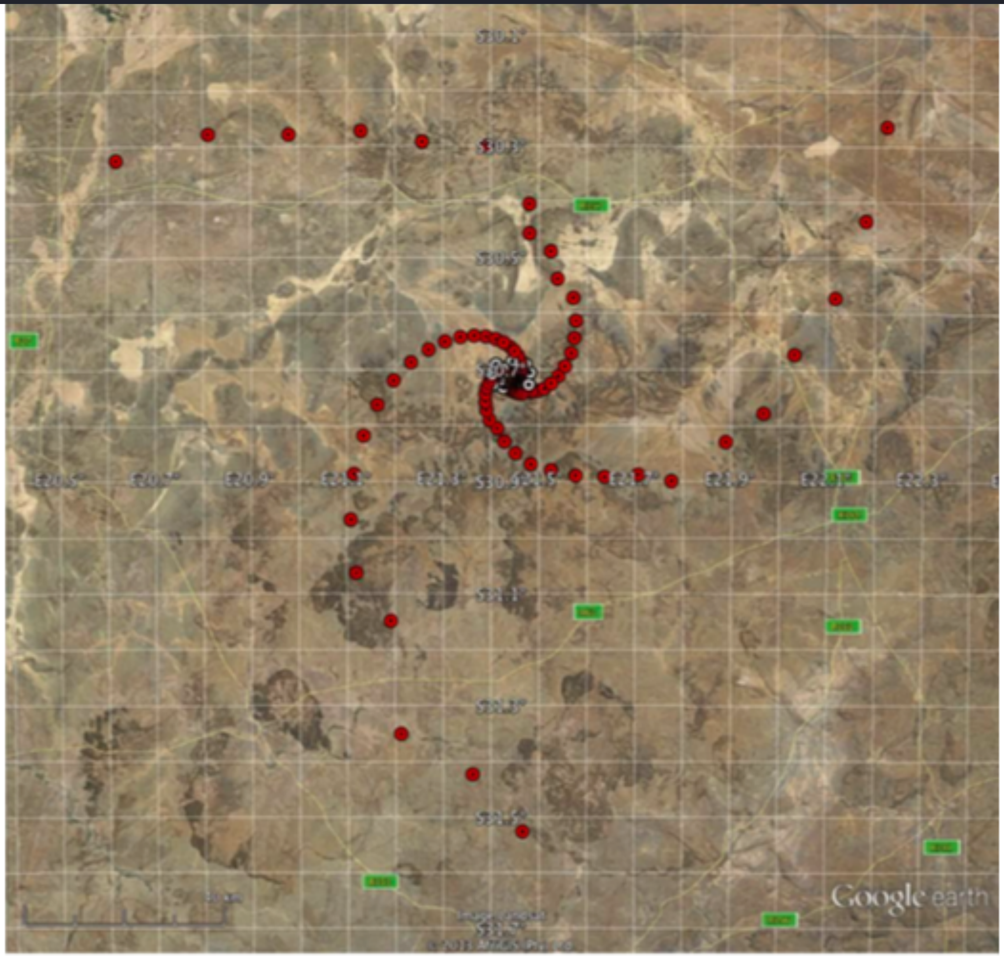
1000 stations 256 antennas each
Murchison, AU - 0.05 humans / sqkm
Compute in Perth

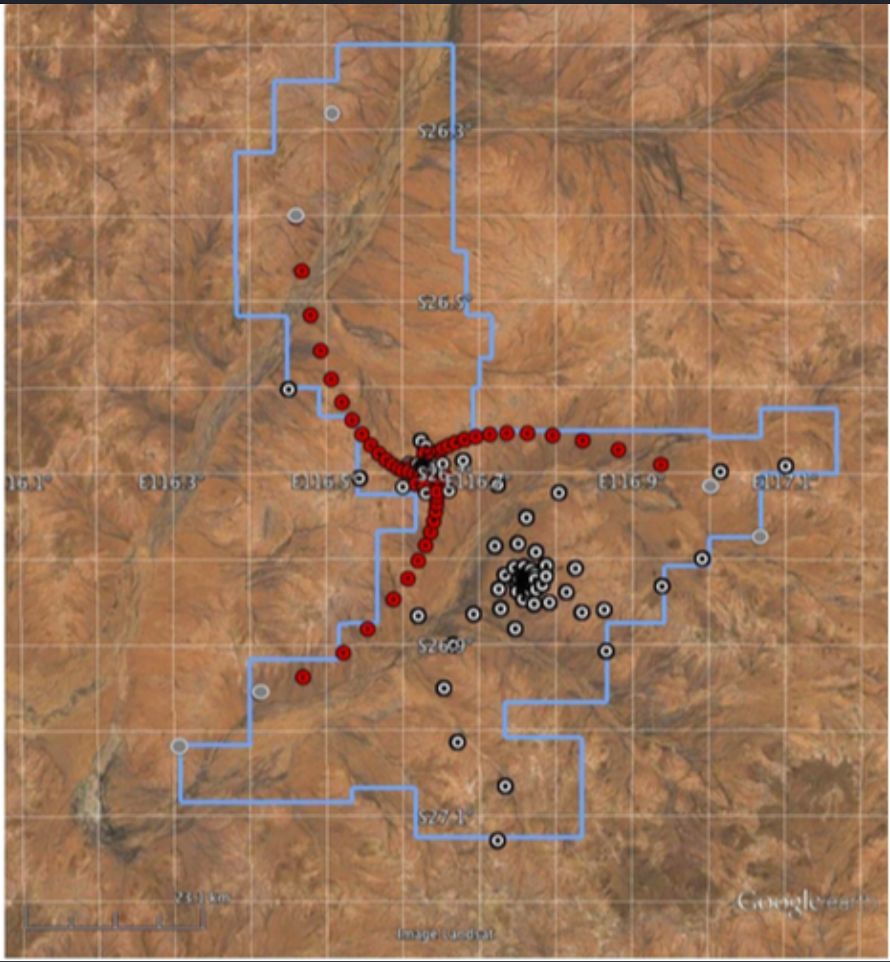
SKA Telescopes



SKA Telescopes

Antenna array layout





SKA1–MID, –LOW: Max Baseline = 156km, 65 km

Science cases

Science Headlines

Fundamental Forces & Particles

Gravity

- Radio Pulsar Tests of General Relativity
- Gravitational Waves
- Dark Energy

Magnetism

- Origin and Evolution of Cosmic Magnetism

Origins

Galaxy & Universe

- Cosmic dawn
- First Galaxies
- Galaxy Assembly & Evolution

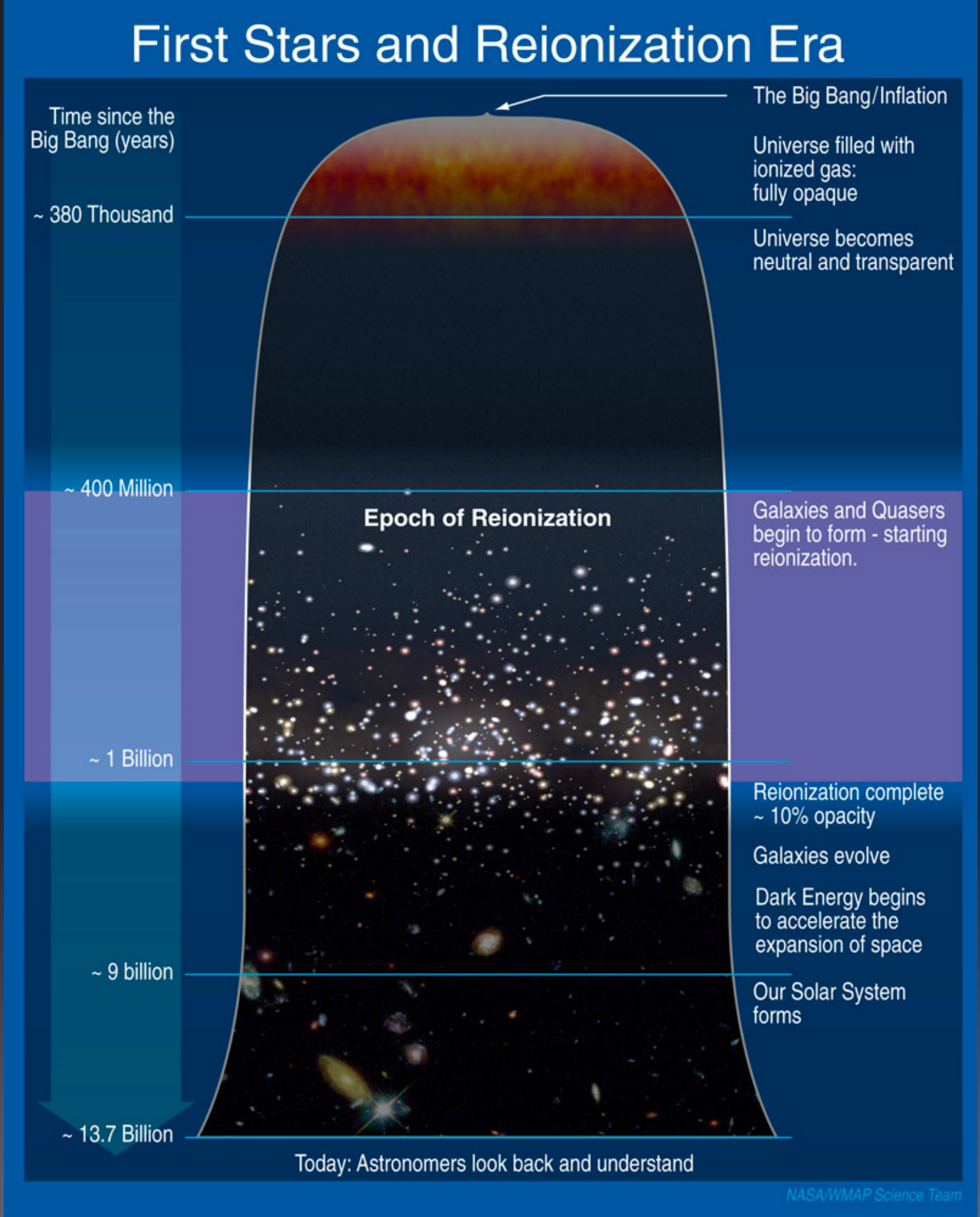
Stars Planets & Life

- Protoplanetary disks
- Biomolecules
- SETI

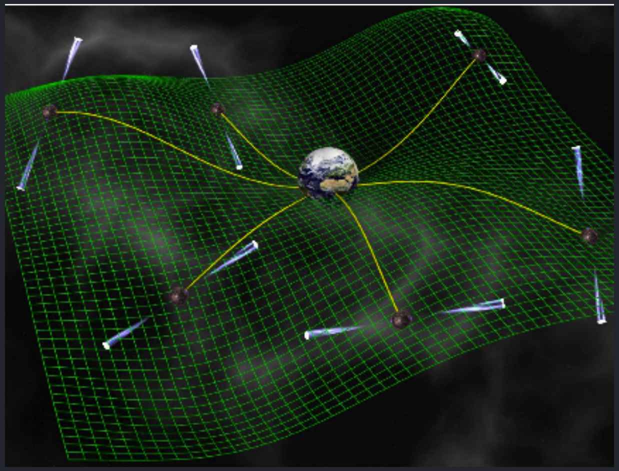
skatelescope.org –
two very large books (free!) with science research articles surrounding SKA

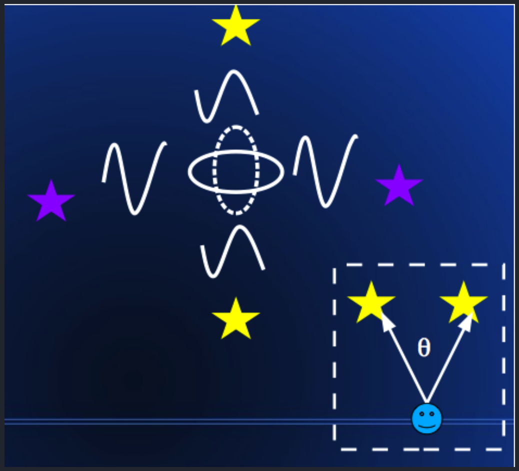
Epoch of Re-Ionisation

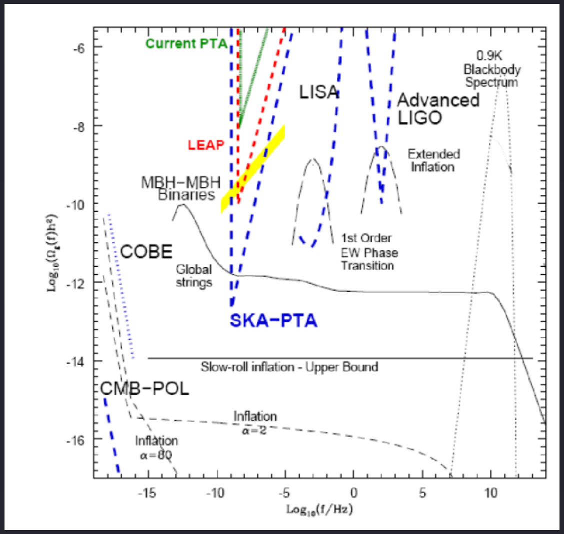
- 21 cm Hydrogen spectral line (HI)
- Difficult to detect,
- Tells us much about the dark ages: 400,000 – 400,000,000 years



Pulsar Timing Array



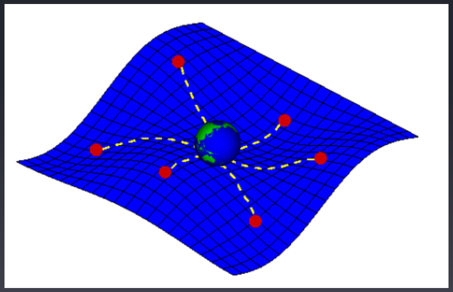




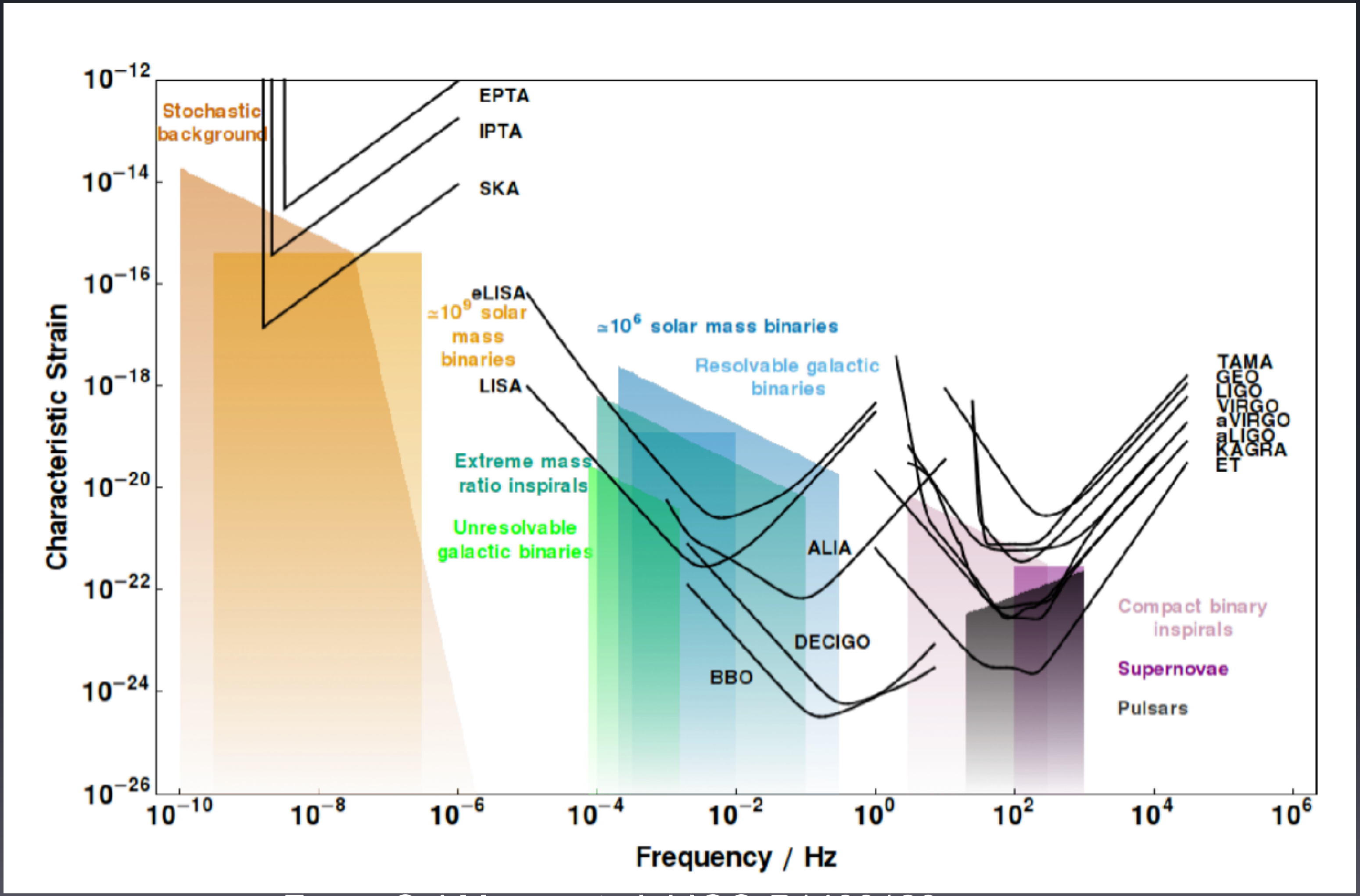
What can be found:

- gravitational waves, cosmic censorhisp, “no-hair” hypothesis

- Nano-hertz range for frequencies, ms pulsars, fluctuations of 1 in 10^20



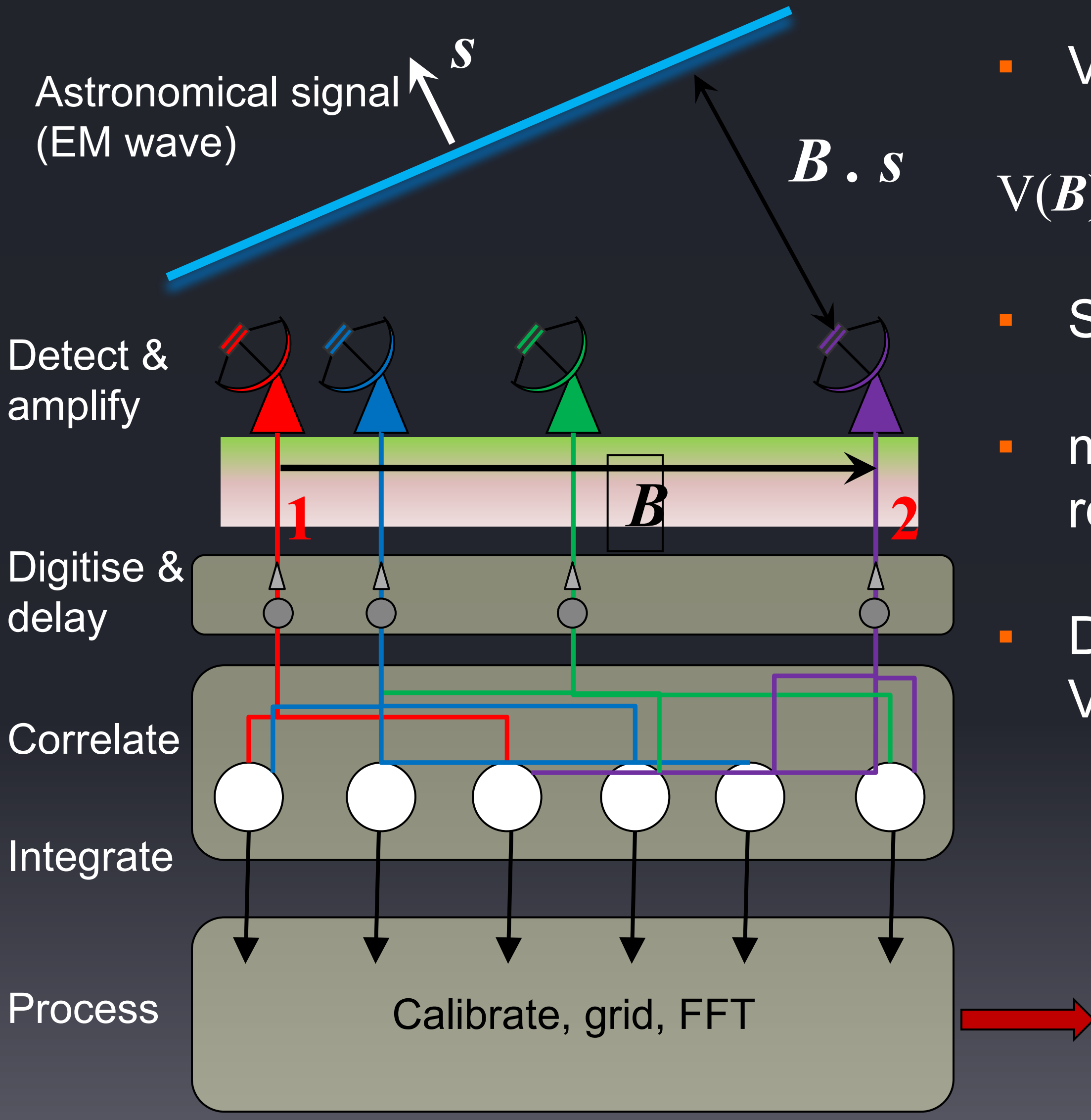
Gravitational Waves & SKA



From: C J Moore et al, LIGO-P1400129.

Computing in the SKA?

Standard interferometer



- Visibility – what is measured

$$V(\boldsymbol{B}) = E_1 E_2^* = I(\boldsymbol{s}) \exp(i\, \omega\, \boldsymbol{B}.\boldsymbol{s}/c)$$

- Solve for I(s)
- maximum baseline gives resolution: $\theta_{max} \sim \lambda / B_{max}$
- Dish size determines Field of View (FoV): $\theta_{dish} \sim \lambda / D$

SKY Image

Challenge

Turn telescope data into science products

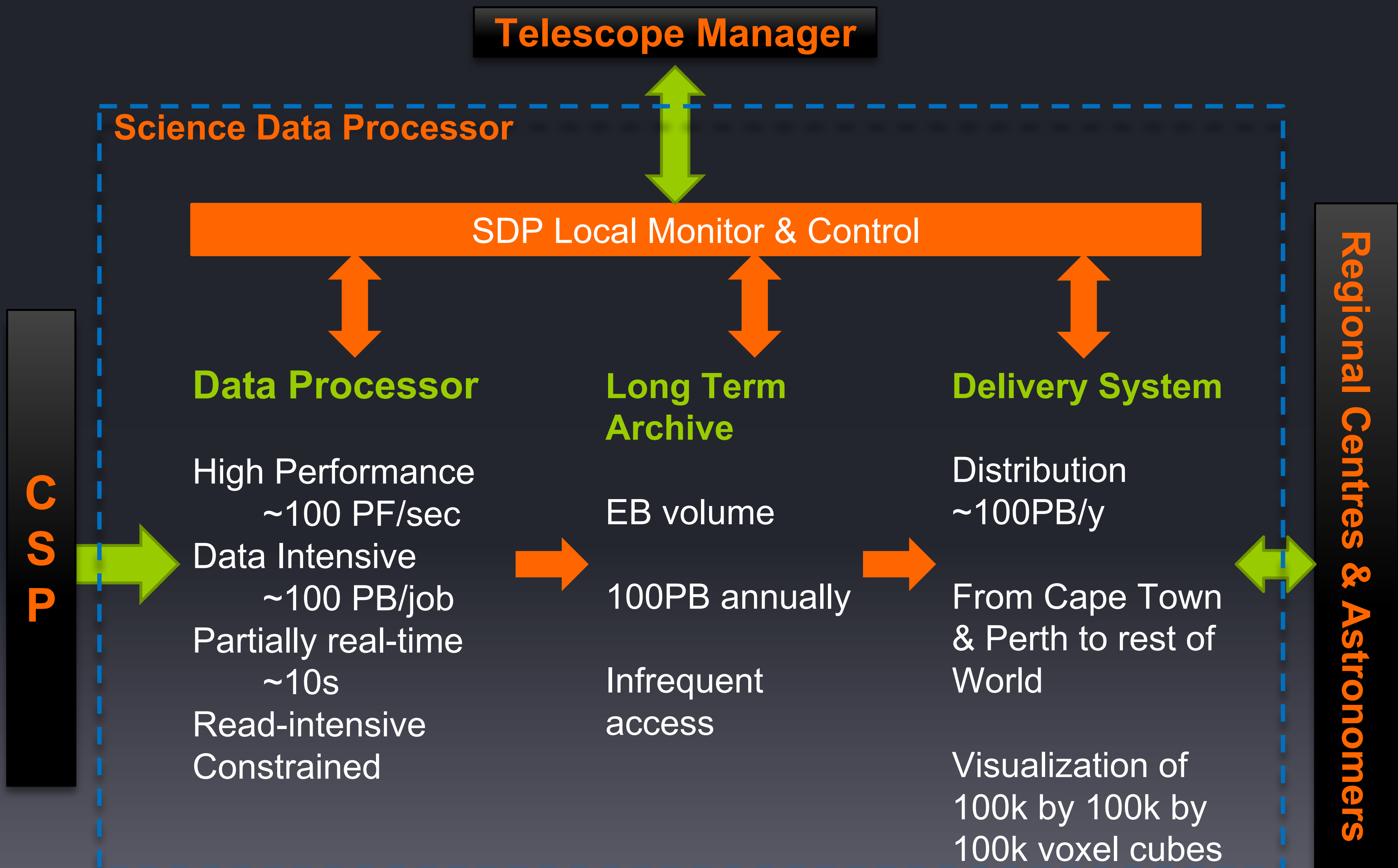
Scientists will consume this worldwide

The SKA telescope will probably live ~50 years

SDP computing hardware will refresh ~every 5 years

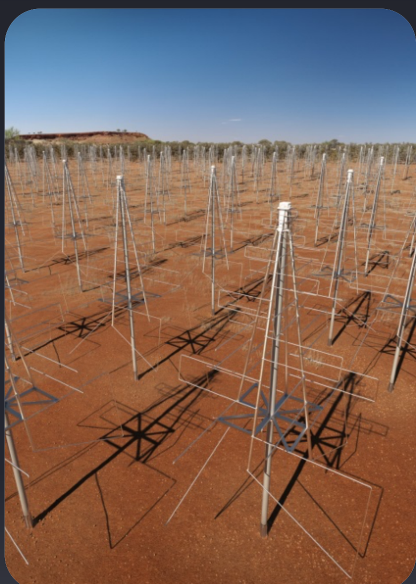
The initial 2023 computing system should exploit the SKA instrument sufficiently to deliver a competitive instrument

SDP top-level compute challenge

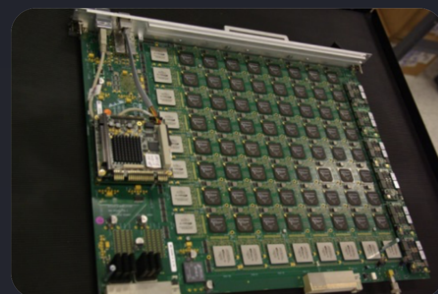


SKA – data schematic

Antennas



Central Signal Processing (CSP)



Transfer antennas to CSP
2024: 20,000 PBytes/day
2030: 200,000 PBytes/day

Over 10’s to 1000’s kms

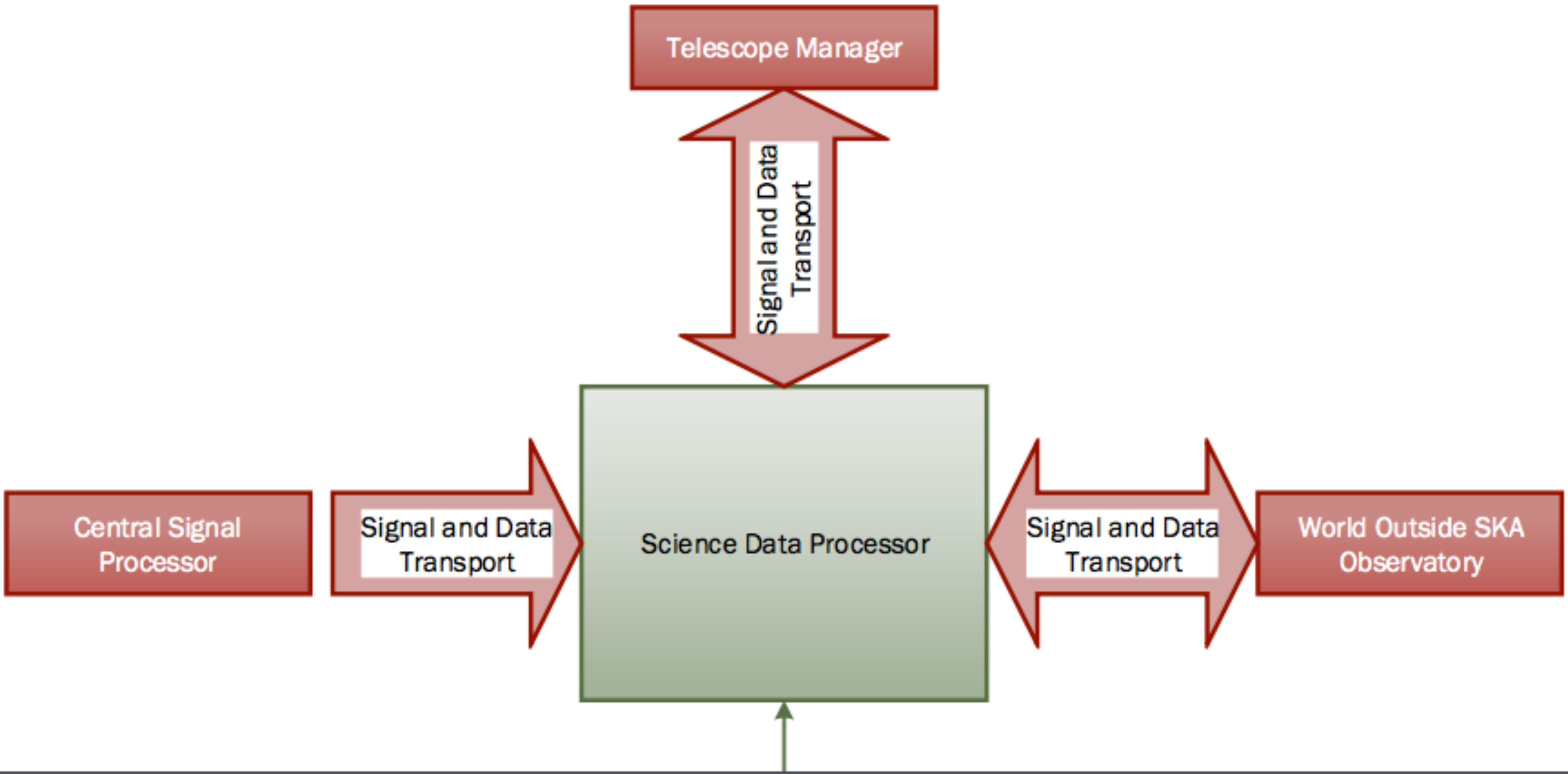
Imaging (SDP) – HPC problem
2024: 100 PBytes/day
2030: 10,000 PBytes/day

Over 10’s to 1000’s kms

High Performance Computing Facility (HPC)

HPC Processing
2024: 300 PFlop
2030: 30 EFlop

Orchestration – interfaces



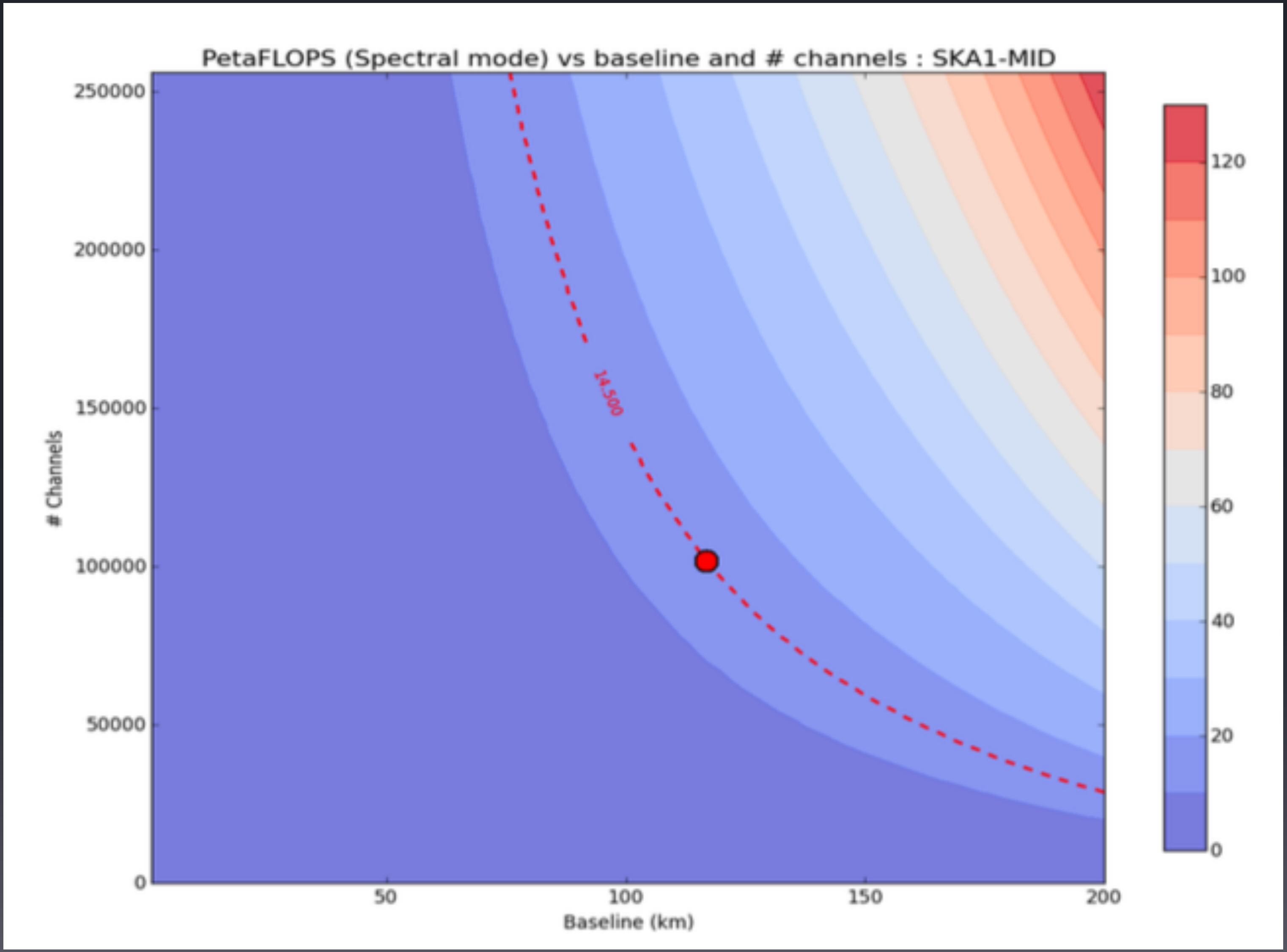
Radio astronomy 0.101

- **@Antennas**: wave guides, clocks, beam-forming, digitizers

- **@Correlator (CSP):** DSP for antenna data
  - Delivers data ***for every pair of antenna's (a "baseline")***
  - Dramatically new scale for radio astronomy ~500K baselines
  - Correlator averages and reduces data, delivers sample every 0.3 sec
  - Data is delivered in frequency bands: ~100K bands
  - 3 complex numbers delivered / band / 0.3 sec / baseline
  - Do math: ~ 1 TB/sec of so called ***visibility data***

- **@Science Data Processor (SDP)** – process correlator data
  - Create images & find transients – these are "science products"
  - Adjust for atmospheric and instrument effects ***calibration***
  - A total measurement lasts up to 6 hours, transients detected in ~10s

Flops vs. #channels & baseline



Baseline distribution

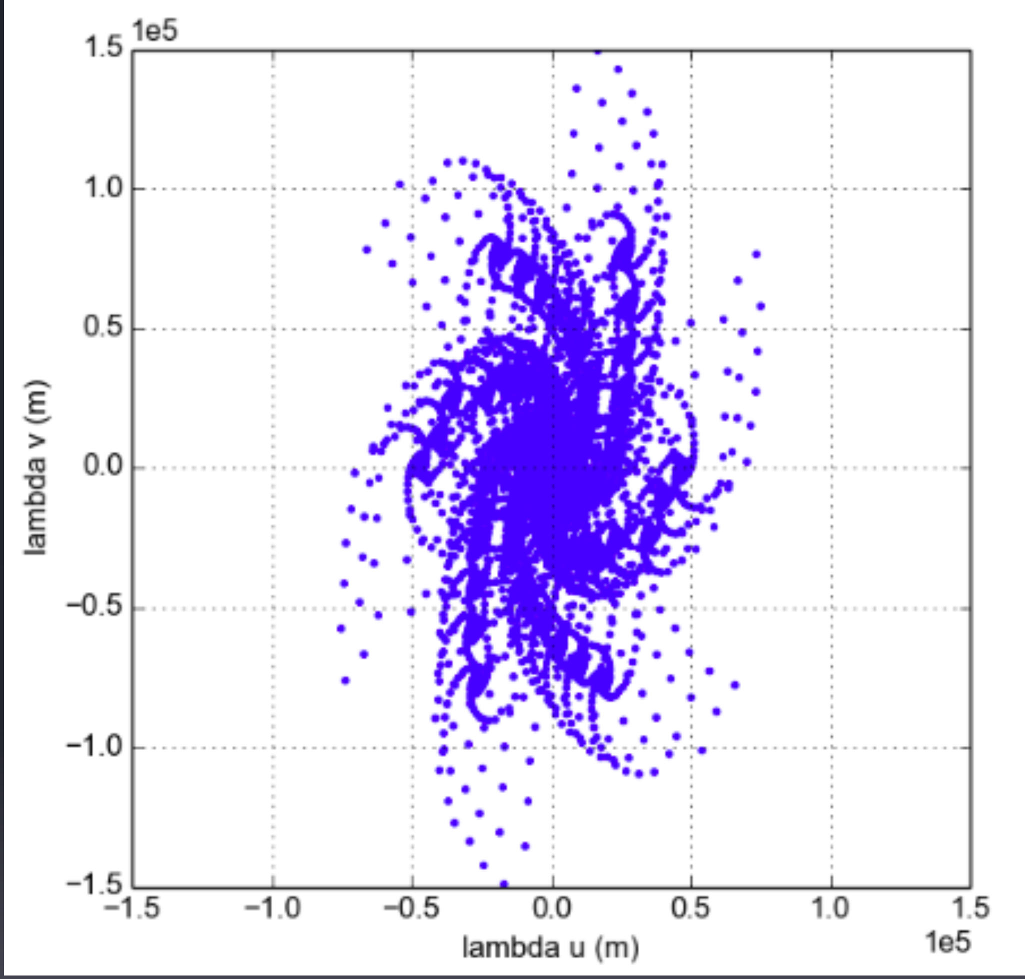
Each pair of telescopes has a baseline

Baselines rotate as time progresses

Each baseline has associated visibility data ("sample")

Baselines are sparse & not regular, but totally predictable

The physical data structure strongly enables and constrains concurrency & parallelism



Simulated data from 250 SKA1-MID dishes

Understanding data & computation: Parametric model

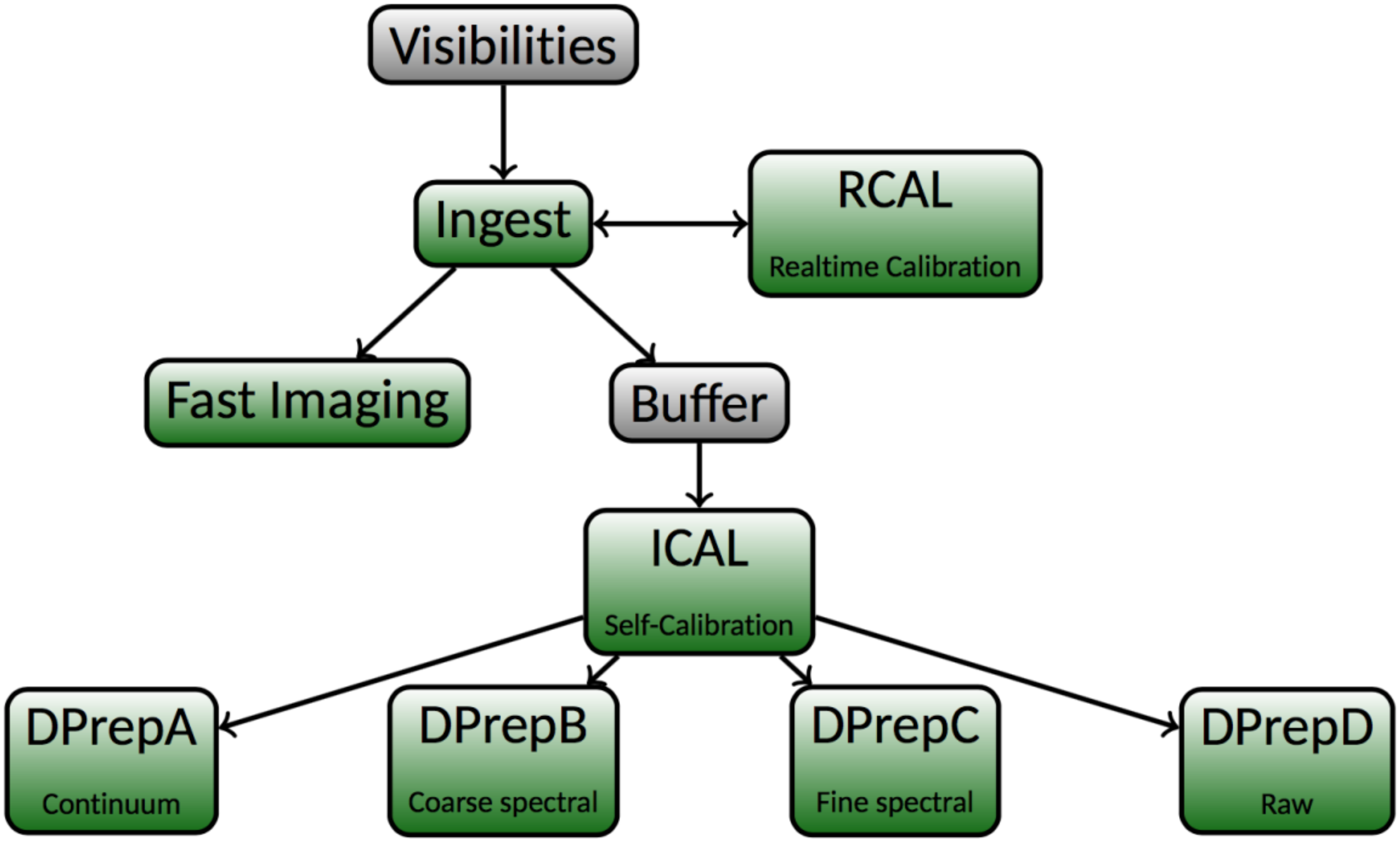
SDP pipelines computing

- **Solve for:**
  - Imaging of the sky – every ~6 hour period
  - Transients – to be found within ~5 sec
  - Effects of the atmosphere, imperfect sampling and imperfect telescope mechanics/electronics ("calibration")

- **In soft-real time**

- **In order to:**
  - Find/measure very faint signals
  - Correct for some of the atmospheric/mechanical problems in real time
  - Produce "science-ready" data products

Pipelines Structure



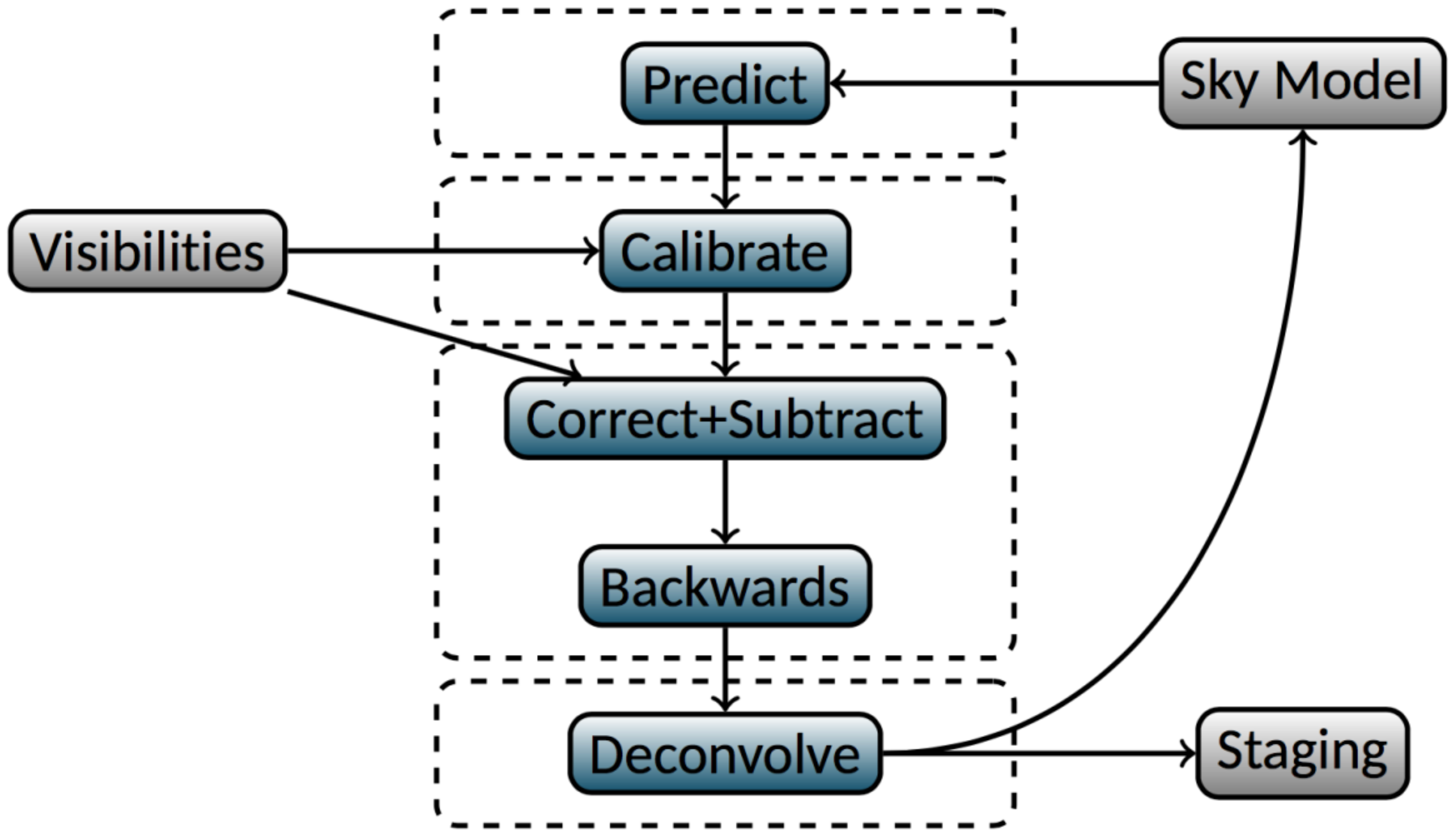
Follows architecture, allows running multiple data preparations.

Pipelines

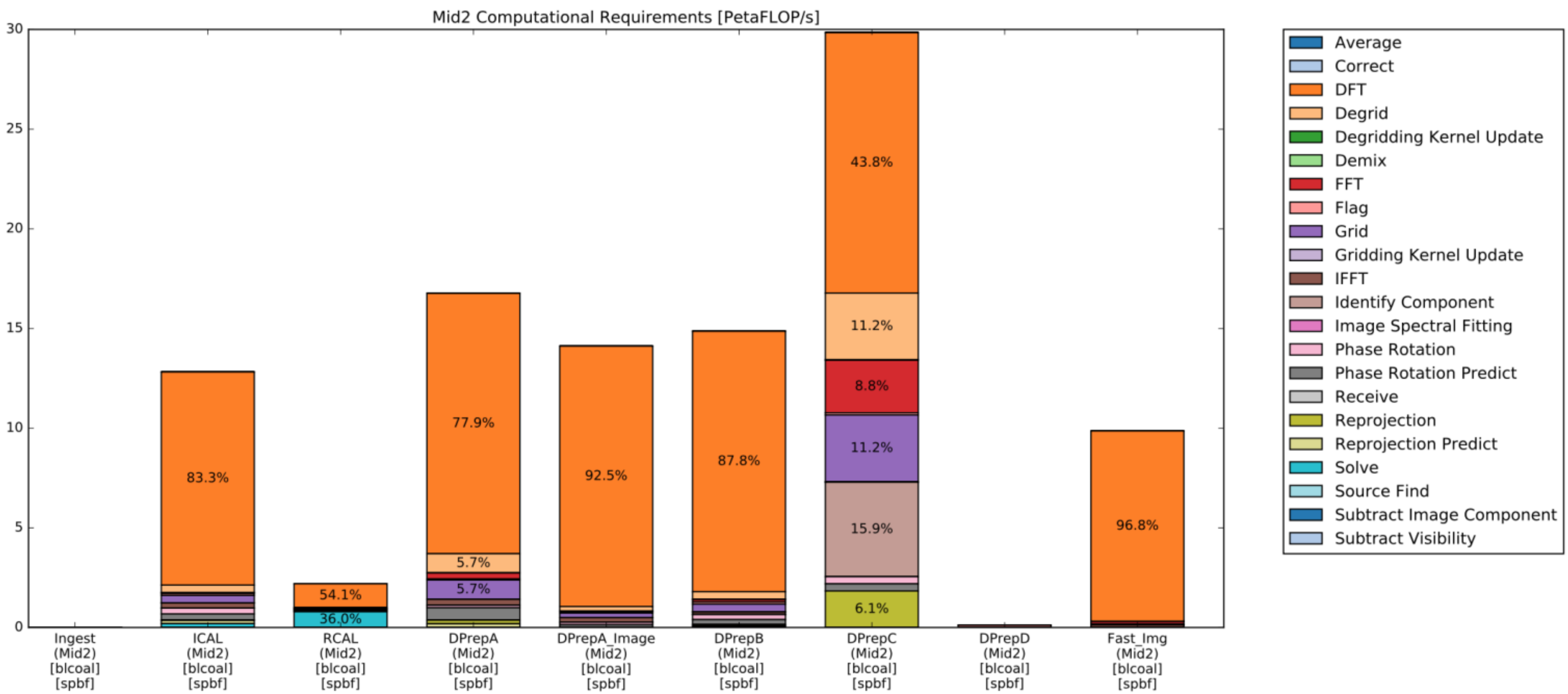
- Many similarities with other image processing
- Each step is
  - Convolution with some kind of a “filter”
  - Fourier transform
  - E.g. “gridding” – approximating irregularly sampled data with a regular sample
- Why new & different software?
  - The input data is sampled not on a grid, but on baselines
  - The scale of the problem is much larger than RAM

Pipeline Internals

Rough structure and distribution pattern of most pipelines:



Relative kernel cost



Imaging analysis summary

- About 10,000 operations per byte of I/O input
- Iterate through the dataset about 10 times
  - read = 10x write – need 10TB/sec read IO
- Need about 100 PetaFLOP/second
- 0.5 Flops / byte read from RAM
  - 200 PB/sec memory BW

Efficiency

Computation demands 89 PF/sec sustained

Efficiency is estimated at 9 – 30%.

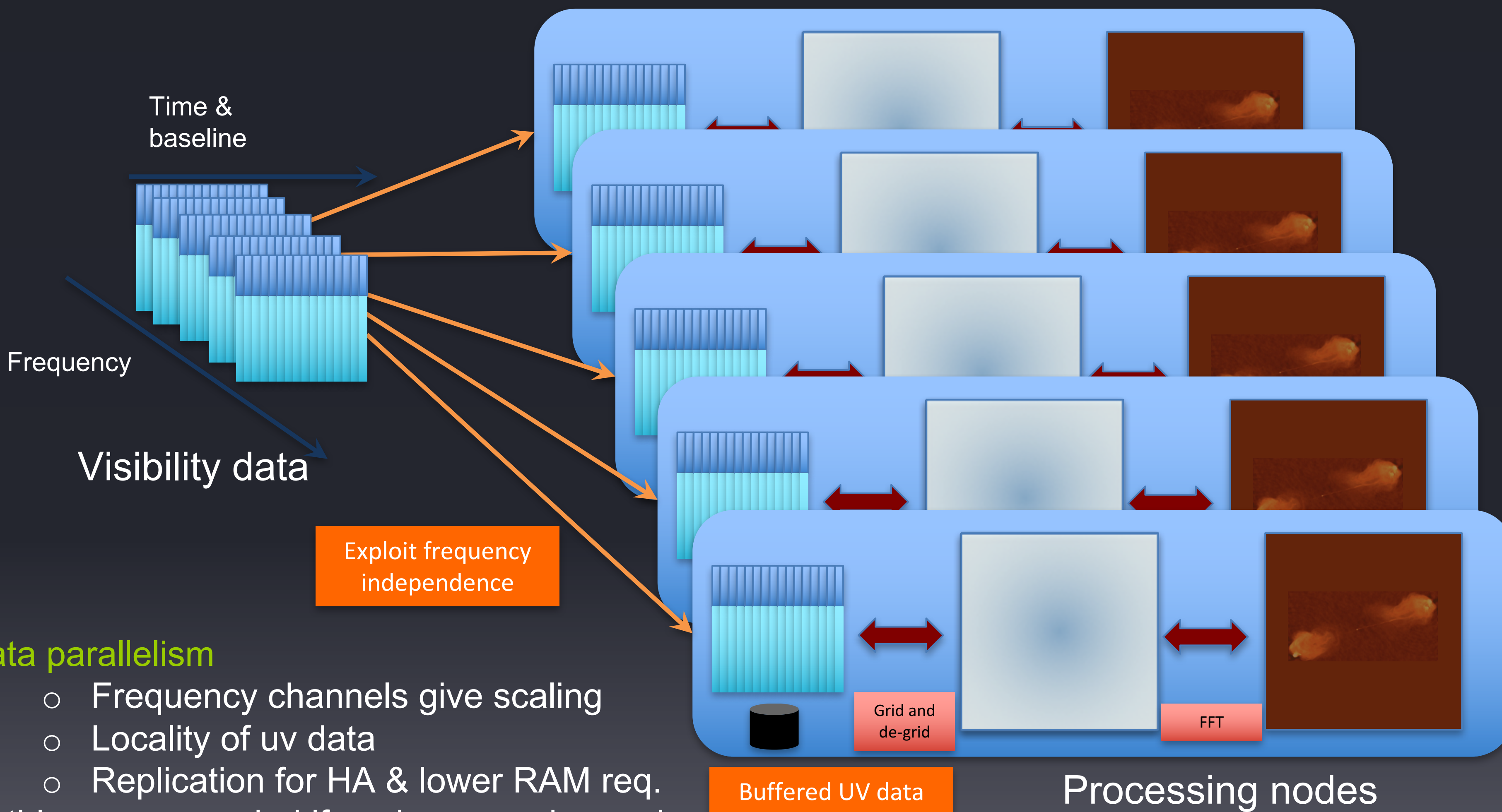
Tied to **energy related hardware parameters** &
Physical data layout & cache re-use

Supercomputer parameters

2023	LFAA (AU)	Mid (SA)
FLOPS	100 PF	360 PF
Buffer Ingest	7.3 TB/s	3.3 TB/s
Budget	45 M€	50 M€
Power	3.5 MW	2 MW
Buffer storage	240 PB	30 PB
Storage / node	85 TB	5 TB
Archive storage	0.5 EB	1.1 EB

Parallel implementation & system architecture

Data Flow on System Architecture



Data parallelism

- Frequency channels give scaling
- Locality of uv data
- Replication for HA & lower RAM req.

Nothing more needed if each processing node can process a frequency channel completely

Principles in parallelism

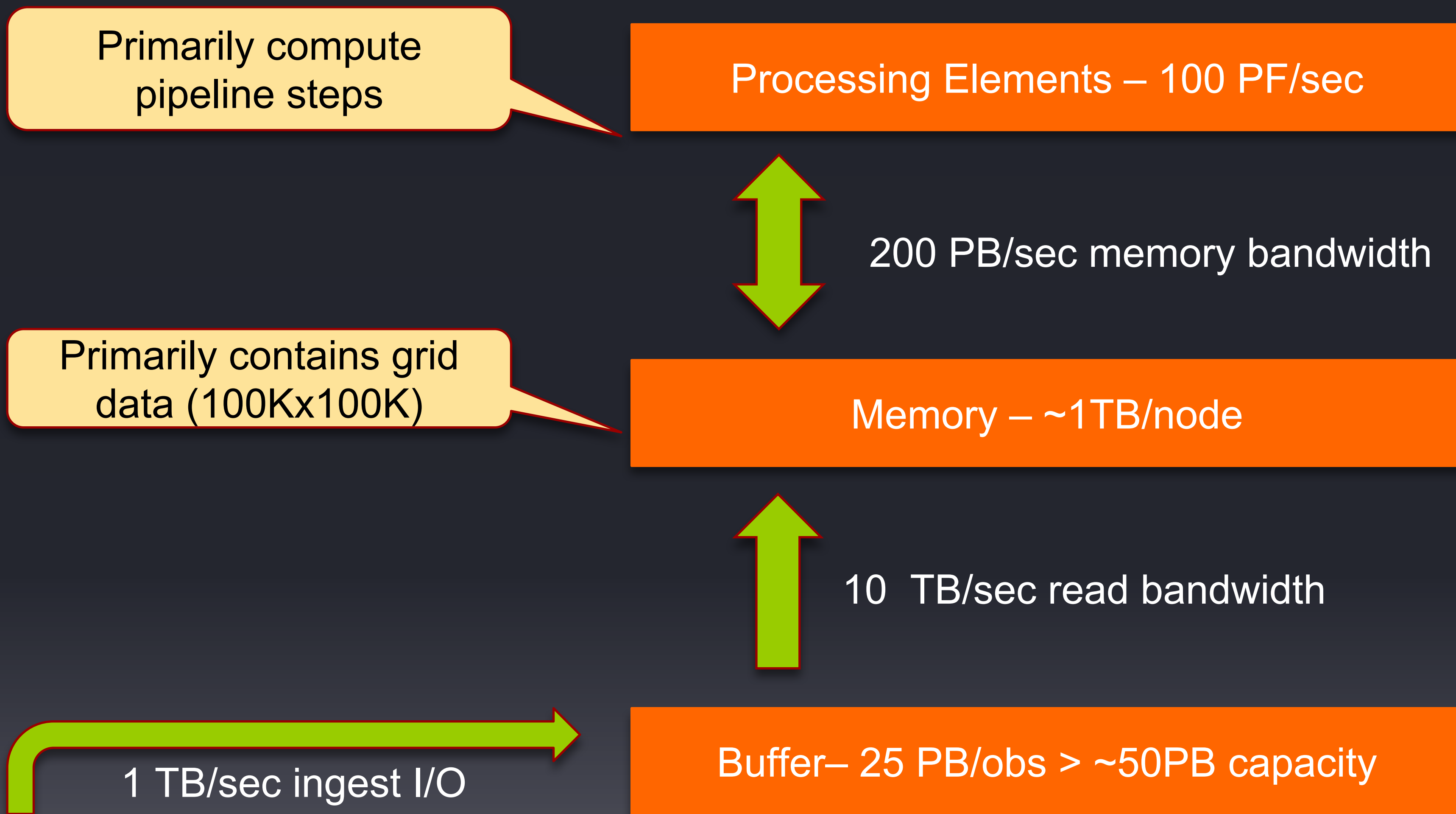
- FFTW - ~1999
  - Massive search to find the most efficient sets of instructions
  - Very successful

- Halide ~2012
  - Explore many strategies for so called “tiling problems”
  - Minute 15:01-17:28 of Halide Movie

Data in the computation

- Two principal data types
  - input is visibility – irregular, sparse uv grid of baselines
  - output regular grid is sky image
  - Messages can be “GB size”
- Different kinds of locality
  - Splitting the stream by frequency
  - Grouping visibilities by region
  - Duplicating input data for fail over
  - Duplicating input data for faceting – less memory use / more I/O

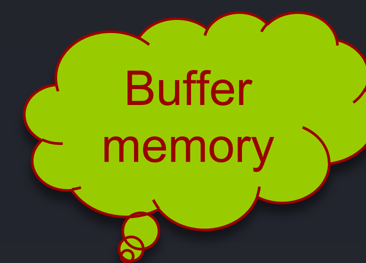
Problem size – locality view



Long vs short buffer question

Processing requires up to 6 hours of ingest – buffer that.

21,600 TB – “unit of data ingest” to compute on



Overlapping ingest and compute: double buffer ?

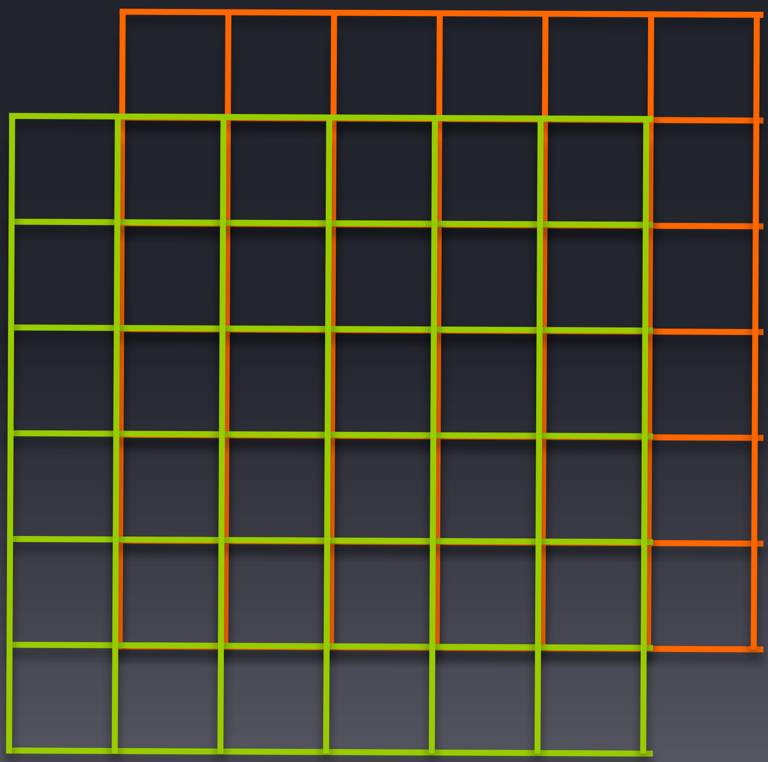
Double Buffer: ~100PB, write 1TB/sec, read 10TB/sec
But processing time is uneven – double buffer:
minimizes storage cost, increases compute cost

Memory Technology

The memory bandwidth of 200PB/sec remains the most problematic.

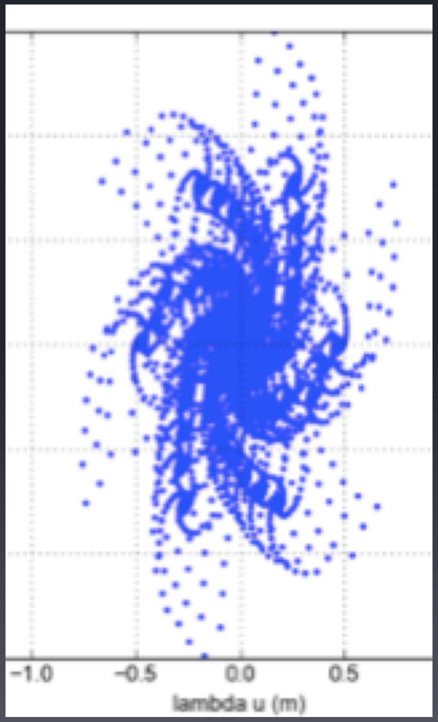
- it probably requires on-package memory (HMC, HBM)
- High Bandwidth Memory / Hybrid Memory Cube
- this offers ~10x BW of RAM
- it consumes too much energy (~50 pJ/byte)
- today we would look at 200K modules
    – somewhat too many

Visibility gridding & cache re-use



Time rotation of UV grid.

Only fetch edges
Re-use core



Stream fusion

- Some kernels exchange too much data
- Solution: deviate from pipeline actors by doing more operations and less data movement.
- Few compilers / frameworks support this with automatic computation
- Doing it manually is awkward for portability

Software approaches

- Creating software is a very high risk part of the project
- Ideal perspective:
  - Execution framework from 3rd party
  - Domain specific application language for pipelines
  - Automatic optimization
- Many approaches:
  - Adapting existing packages – MPI C++ applications
  - Use a big-data framework like Spark
  - Use HPC frameworks like Swift/T, Legion
- Remains undecided.

Storage hardware

NVM solutions - xPoint and other

Could deliver ~50 GB/sec / node

200 nodes could get 10TB/sec I/O BW

But distributed storage remains hard

Summary

- SKA next generation telescope
  - Order of magnitude improvements
  - Key science drivers: gravitational waves, ionization of primordial gas by first galaxies
- Timeline
  - Design finishes in 2018
  - Construction starts 2019
  - Commissioning 2021
  - Full operations 2024
- Science Data Processor
  - Aligned with computational model & industry data

Thank you! Questions?

Sensor

HPC Data / Big Data

Sensor big data

- Radio astronomy
- Remote sensing
- Earth observation
- Geophysics
- Medical (imaging or other)
- Internet of Things?
  - 10 Hz sampling air temperature inside your fridge?

Tentative big data classification

Application	Input data volume	Input data density	Computational density	Message rate
HPC simulation	low	high	high	medium
Map reduce analytics	high	high	medium	low
Graph analytics	medium	high	low	high
Sensor data	high	low	medium	low

Characteristics of sensor big data

Noisy	• Information content << sampling rate • Combination/correlation of data necessary to find signals of interest -> volume!
Multiple Streams of Input Data	• Volume/shape/ characteristics of data known in advance • High degree of inherent parallelism at the front-end of processing
Calibration effects	• Large sensors networks can not be made perfect • Allow biases in measurement and find and correct for these in post-processing
Incomplete/ imperfect sampling	• Poor control over “experiment” • Expensive to precisely specify the sampling points • Non-parametric statistics from incomplete/imperfect data