The SKA telescope, memory bandwidth and POSITs

Prof. Peter J. Braam
Physics Department
Oxford University
peter@braam.io

Mar 2019

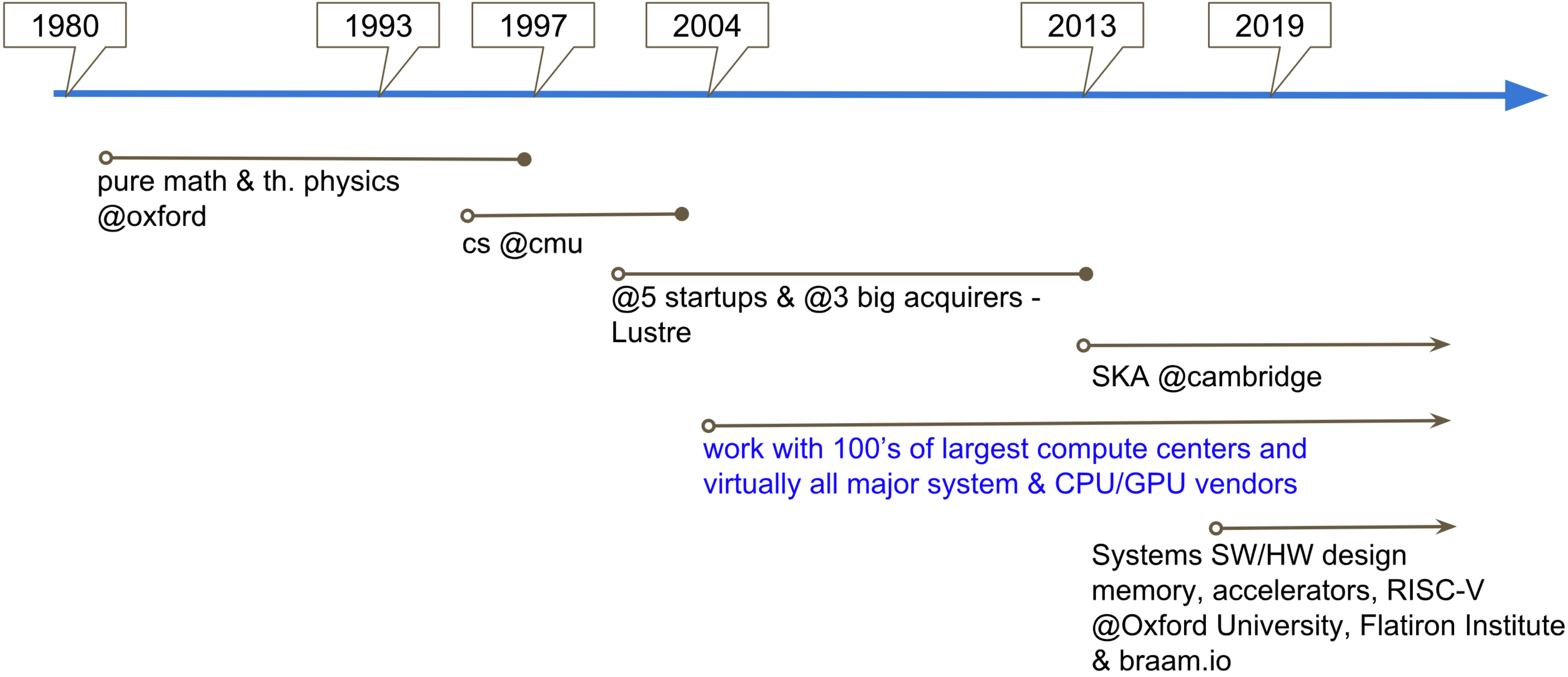
Acknowledgement

A large group of people (~500) are working on SKA design 2013-2019
Most information is publicly available, but very technical
This presentation re-uses much from other SKA efforts

Particularly I'm using a few slides from Peter Wortmann
Background: skatelescope.org

My role:
- Long term consultant & visiting academic in the Cambridge group: 2013-2018

me


1980
1993
1997
2004
2013
2019
pure math & th. physics
@oxford
cs @cmu
@5 startups & @3 big acquirers -
Lustre
SKA @cambridge
work with 100's of largest compute centers and
virtually all major system & CPU/GPU vendors
Systems SW/HW design
memory, accelerators, RISC-V
@Oxford University, Flatiron Institute
& braam.io

Message from this talk

1. SKA telescope is a grand challenge scale project

2. Synergy between scientific computing and industry for performance

    Hardware – particularly memory, energy

    Software – agility, parallelism, energy

3. Pre Hardware use of POSITs can dramatically reduce memory bandwidth

4. Deep questions and alternative approaches may have long term value

What is the SKA?

The Square Kilometre Array (SKA)

Next Generation radio telescope – compared to best current instruments it will offer …

- ~ 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

SKA International Design Consortia

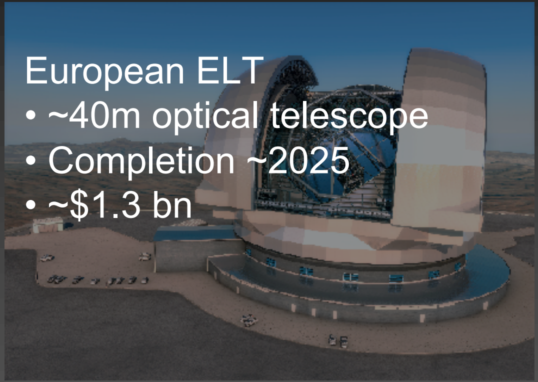
Project Management and System Engineering Team based at JBO (UK)

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

SKA – a partner to ALMA, EELT, JWST

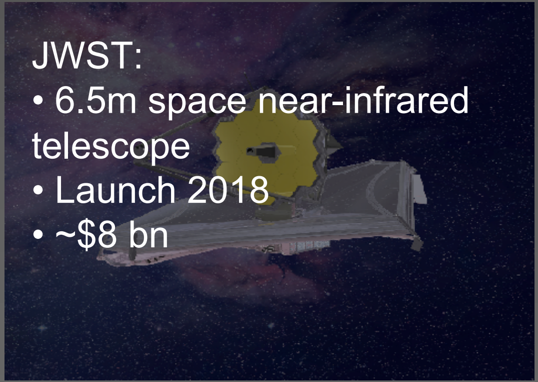
ALMA:
- 66 high precision sub-mm antennas
- Completed in 2013
- ~$1.5 bn

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)

In summary …

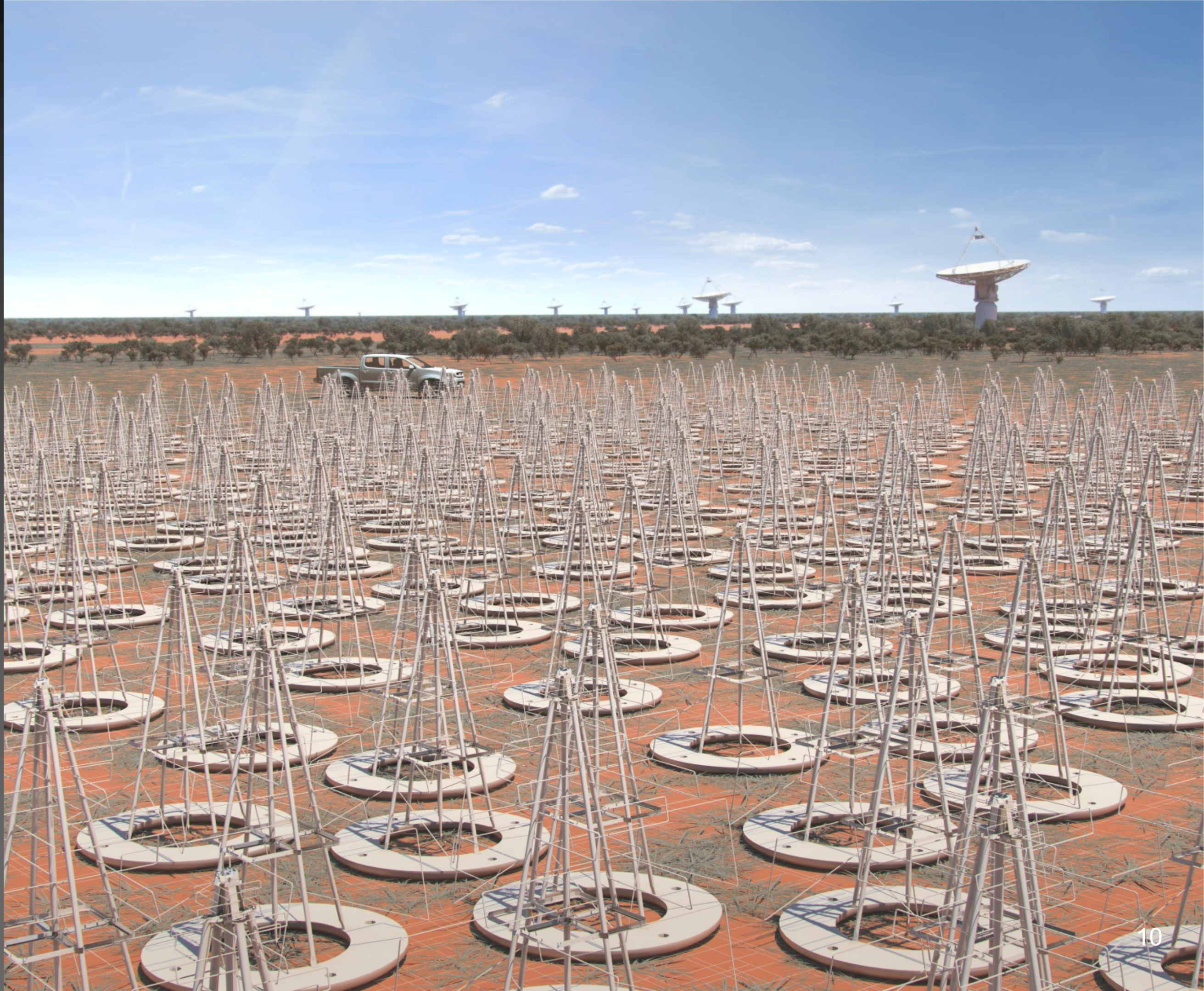
- SKA aims to be a world class “instrument” like CERN
- SKA Phase 1 – in production 2025. Design wrapping up in 2019.
- SKA Phase 2 – likely 10x more antennas – 2030’s?
- This presentation focuses on SKA1
- Caveat
  - Ongoing changes
  - Some inconsistencies in the numbers

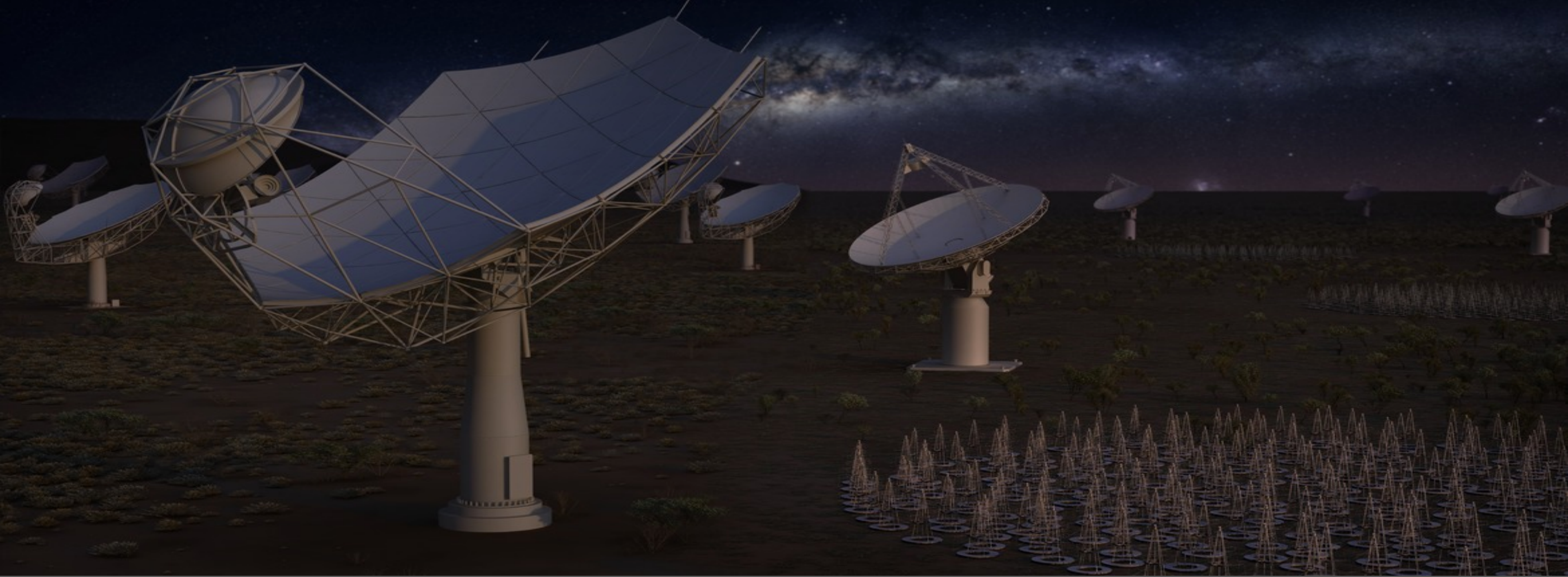
Low Frequency Aperture Array
0.05 – 0.5 GHz

Australia

~1000 stations
256 antennas each
phased array with
beamformers

Murchison Desert
0.05 humans/km^2
Compute in Perth



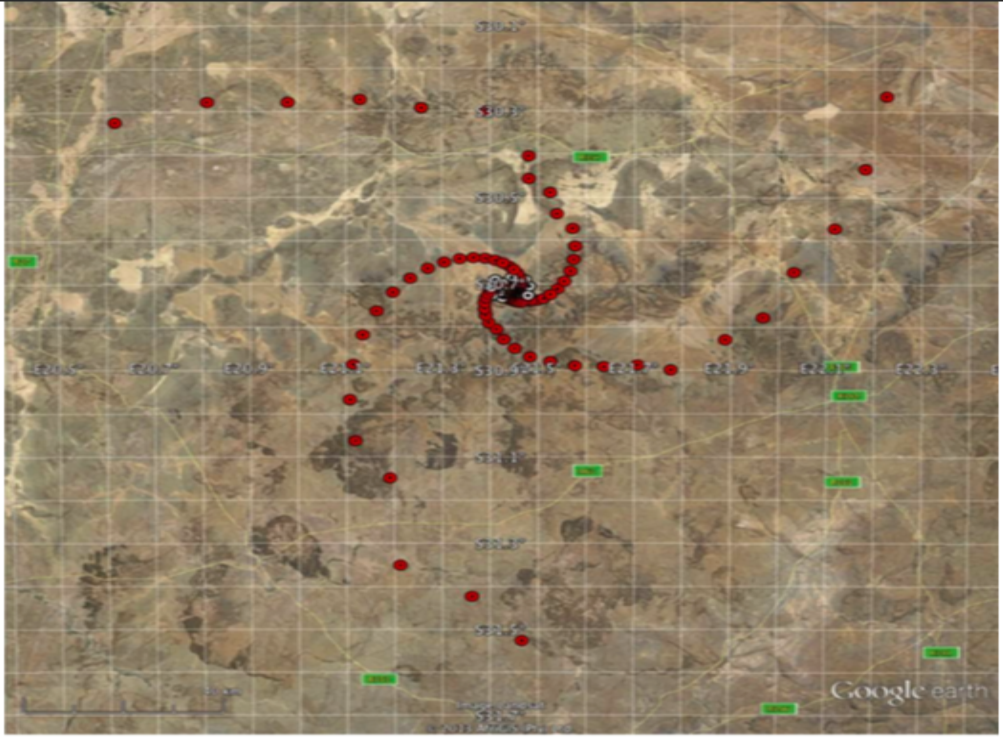


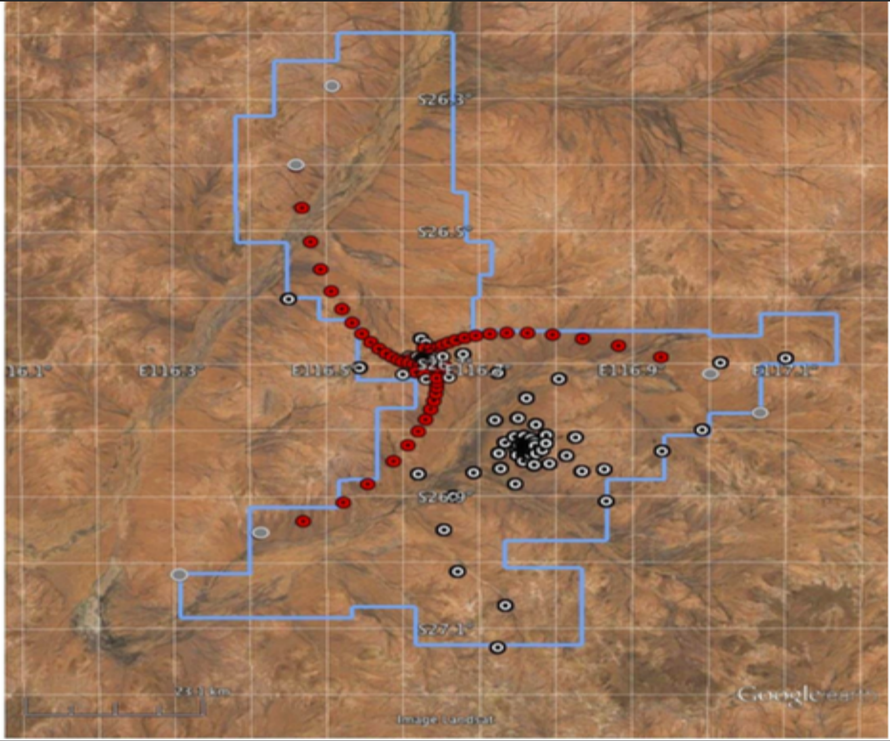
Mid Frequency Telescope

South Africa

250 dishes with single receiver
Karoo Desert, SA - 3 humans / km^2
Compute in Cape Town (400 km)

Antenna array layout





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

Science

Science Headlines

Fundamental Forces & Particles

Gravity

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

Magnetism

- 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 research papers surrounding SKA

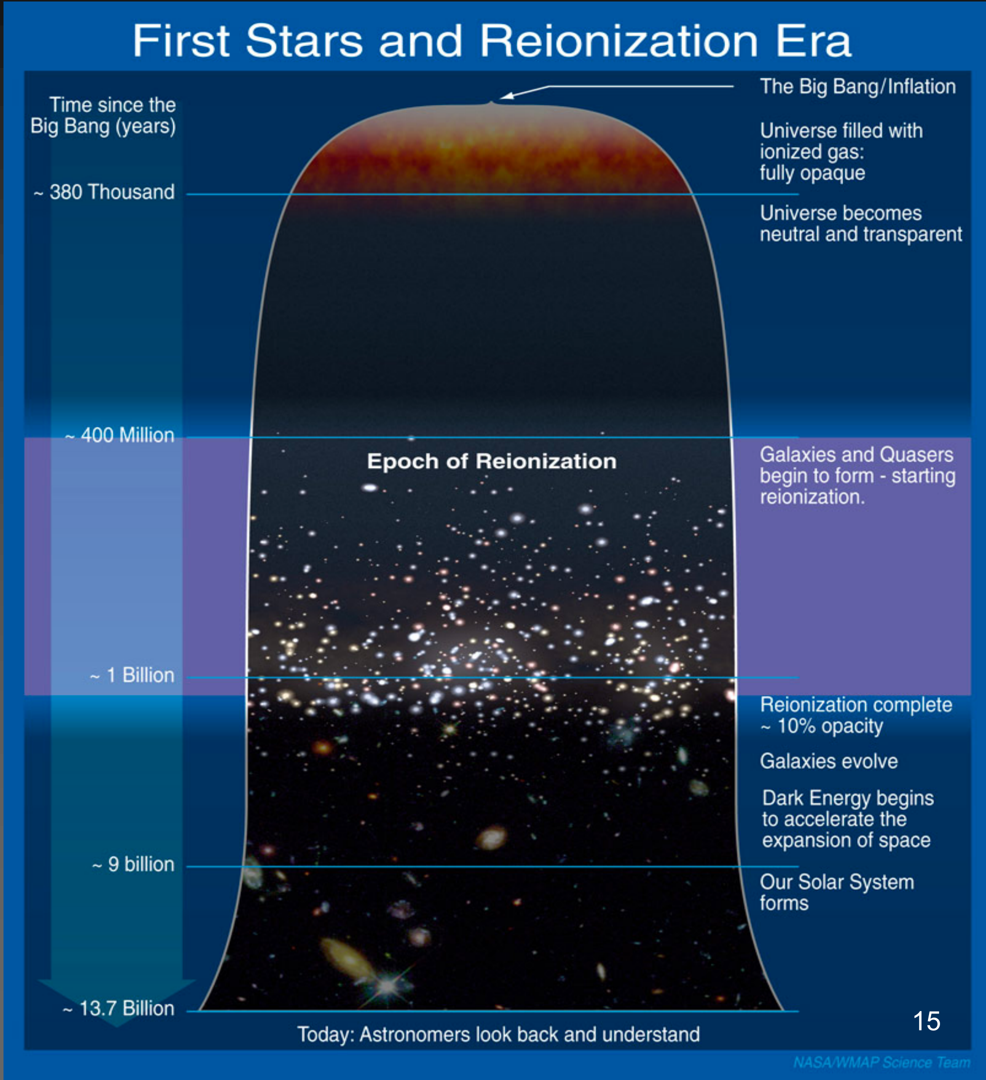
Epoch of Re-Ionisation

21 cm Hydrogen
spectral line (HI)

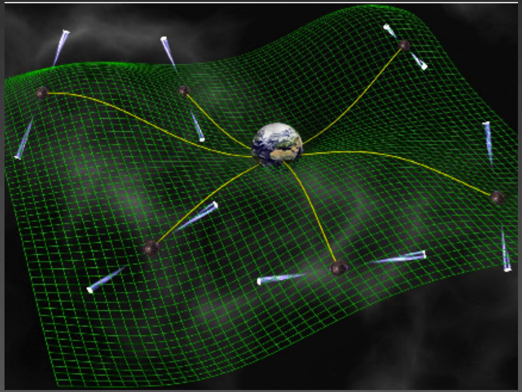
Difficult to detect

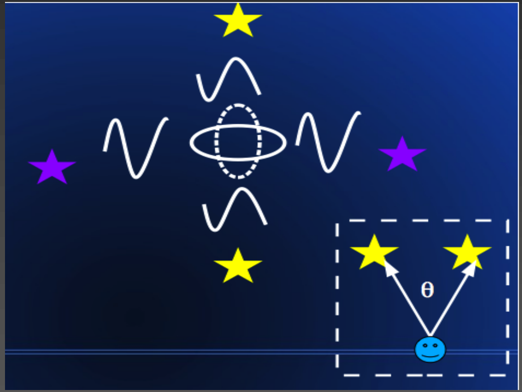
Tells us about the dark age:

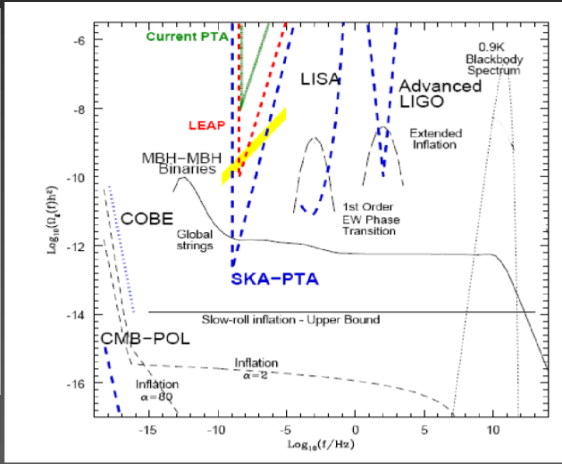
400K – 400M years
(current age =13.5G years)

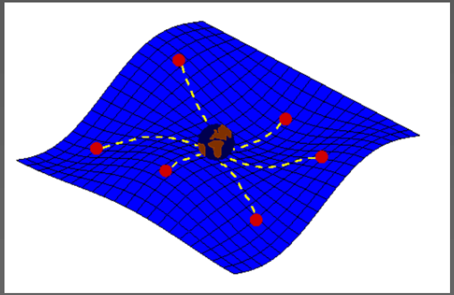


Pulsar Timing Array









What can be found:

- gravitational waves
- Validate cosmic censorship
- Validate “no-hair” hypothesis

- **Nano-hertz frequency range**
- ms pulsars, rotation fluctuations of 1 in 10^20

- SKA1 should see all pulsars (estimated ~30K) in our galaxy

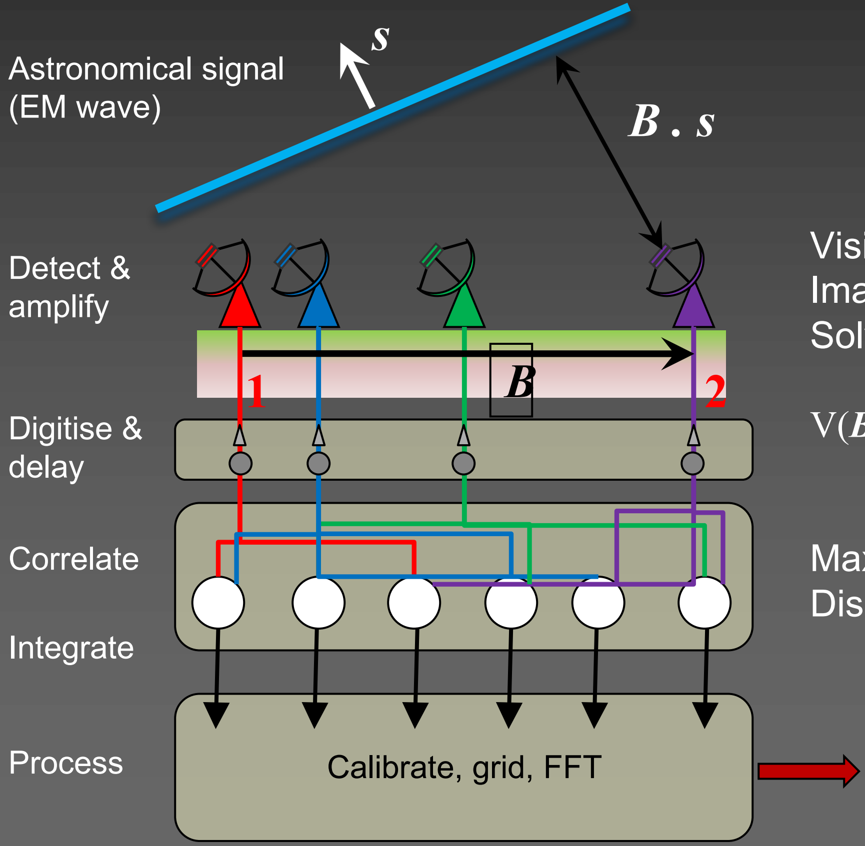
Physics & Astrophysics

Many key questions in theoretical physics relate to astrophysics

Rate of discoveries in the last 30 years is staggering

Imaging Problem

Standard interferometer



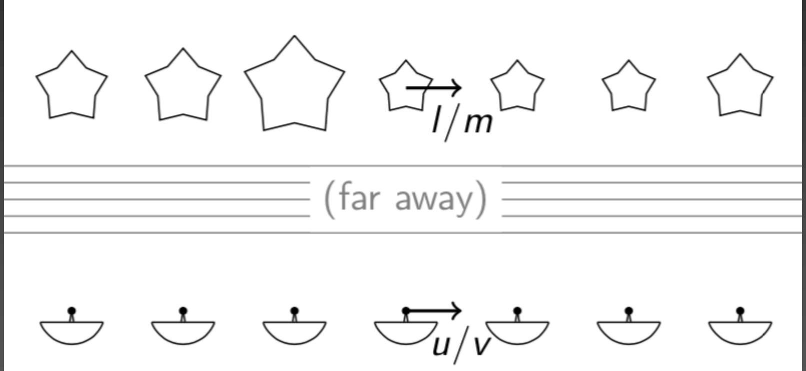
Visibility V(B): what is measured on baselines
Image I(s): image
Solve for I(s)

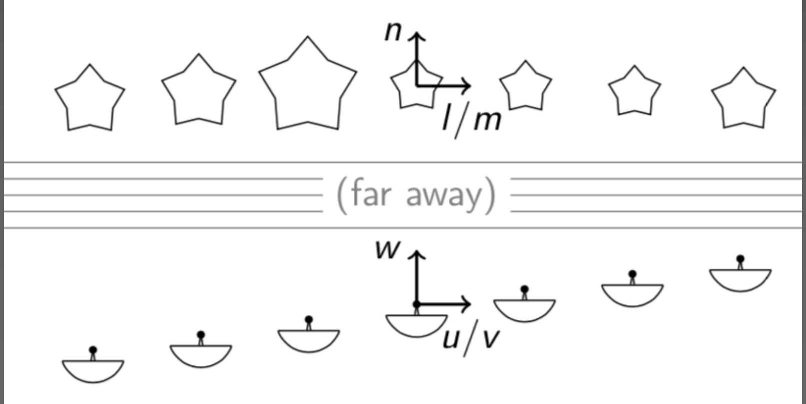
$V(\boldsymbol{B}) = \; E_1 E_2^* = I(\boldsymbol{s}) \exp(i\, \omega\, \boldsymbol{B.s}/c)$ – image equation

Maximum baseline gives resolution: $\theta_{max} \sim \lambda / B_{max}$
Dish size determines Field of View (FoV): $\theta_{dish} \sim \lambda / D$

SKY Image

Interferometry radio telescope





Simplified

Sky is flat
Earth is flat

Telescope to image is Fourier transform

Actually

Sky is sphere, earth rotates, atmosphere distorts

Now it is a fairly difficult problem:
1. Non-linear phase
2. Direction, frequency, baseline dependent gain factor

Computing in radio astronomy - 101

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

@Correlator (CSP central signal processing): == DSP for antenna data

Delivers data *for every pair of antenna's (a "baseline")*

Dramatically new scale for radio astronomy ~100K baselines

Correlator averages and reduces data, delivers sample every 0.3 sec

Data is delivered in frequency bands: ~64K bands

3 complex numbers delivered / band / 0.3 sec / baseline

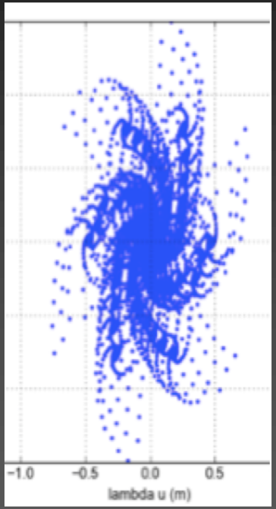
Do math: ~ 1 TB/sec input of so called *visibility data*

@Science Data Processor (SDP) – process correlator data

Create images (6 hrs) & find transients (5 secs) – "science products"

Adjust for atmospheric and instrument effects *calibration*

Data in the computation



Two principal data types

inputs: **visibility** – irregular, sparse uvw - grid of baselines

output: **image grid** - regular grid in sky image

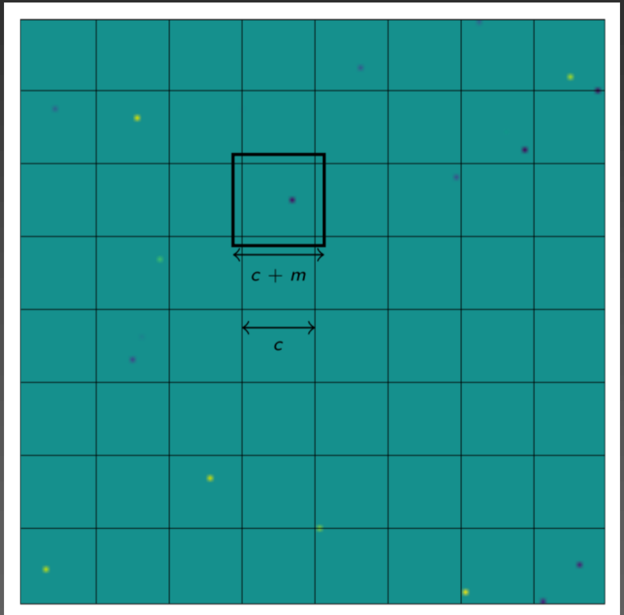
Different kinds of locality

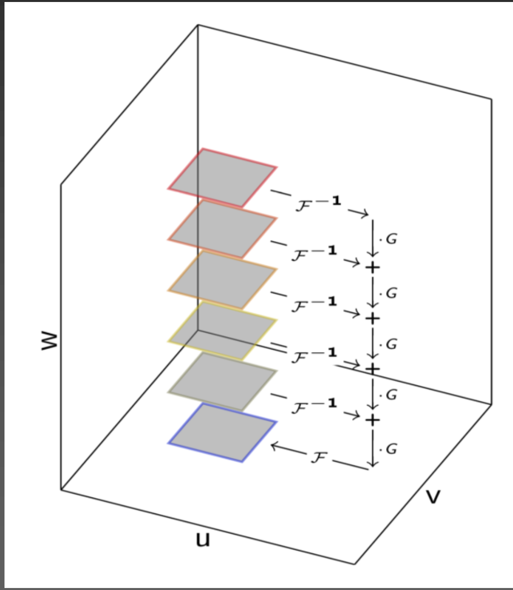
Splitting the stream by frequency

Visibility “tile” data is highly irregular, but computable

Analyze visibility structure – 0, sparse, dense: separate tile strategies

Reducing to 2D





Try to go back from 2D to 3D problem by relating (~100) different w values.
Domain-specific optimization.

Grid size is 64K x 64K for 64K frequencies – problem is large
Full FFT is O(k log k), sparse FFT: O(#nonzero log #nonzero). This approach is close to this.

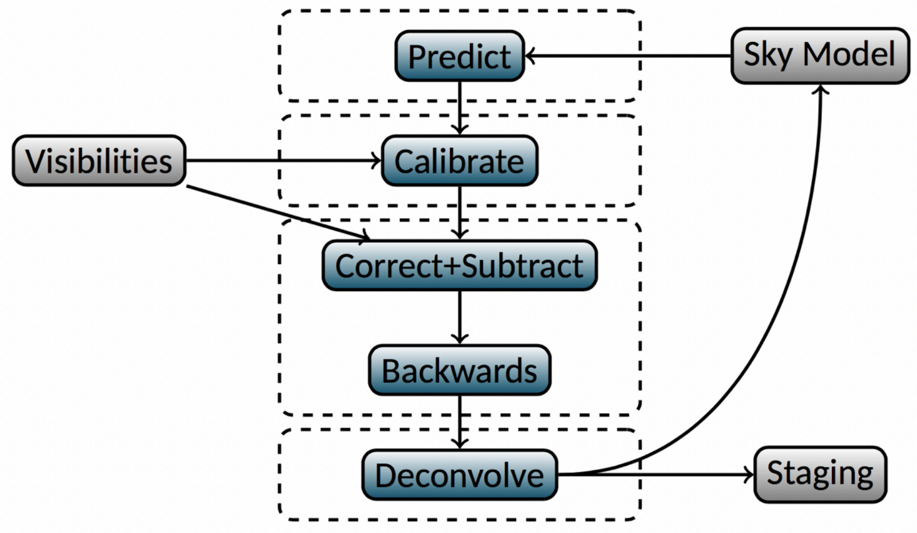
Outline of algorithm

About 5 different analysis algorithms:
e.g. spectral vs continuum imaging

Imaging pipelines:

- Iterate until convergence – approximately 10 times
- Compares with an already known model of the sky
- Incorporates and recalculates calibration data

Rough structure and distribution pattern of most pipelines:



SDP specific Pipelines

Algorithmic similarities with other image processing

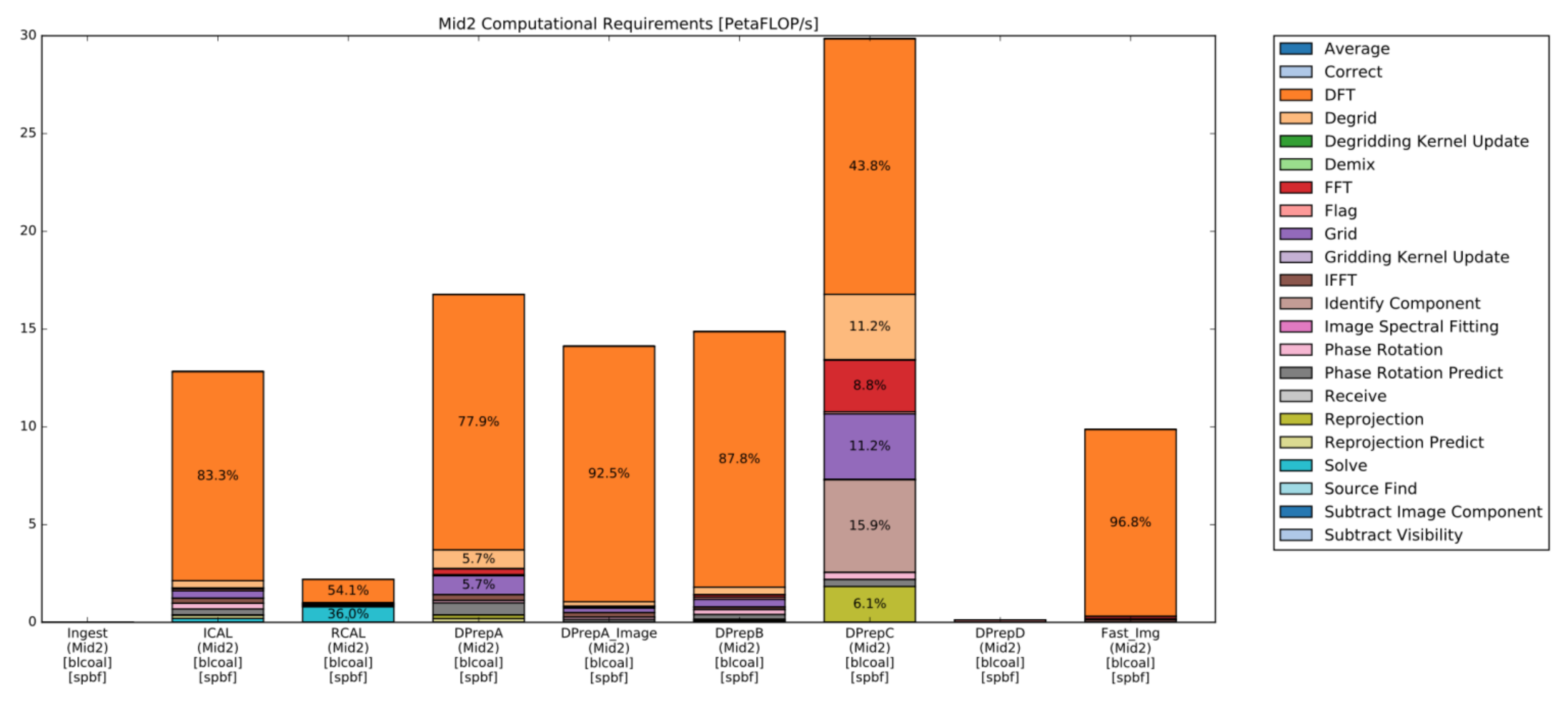
Each step is

- Convolution with some kind of a “filter” – e.g. “gridding”
- Fourier transform
- All-to-all for calibration

Why new & different software?

- Data is very distinct from other image processing
- Problem is very large – much bigger than RAM
- Reconstruction dependencies: sky model & calibration

Relative kernel cost



SDP “performance engineering” approach

Conservative - this is not computing research

Known-good algorithms, hardware

Perhaps deep math question remains: is problem really O(#antennas^2)?

Parametric model of the computation

Detailed FLOPs, memory use, data movement, energy

Key outcome: 100 PF/sec & move 200 PB/sec from HBM to CPU

@50 PJ / byte this is ~10MW power

10-30*% CPU efficiency*

Software

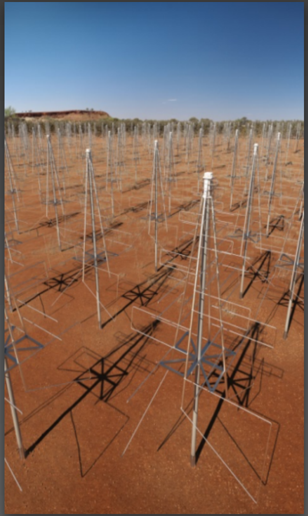
Agility: telescope life = 50 years, many HW refreshes

Agility: every new telescope has led to serious adaptations of algorithms

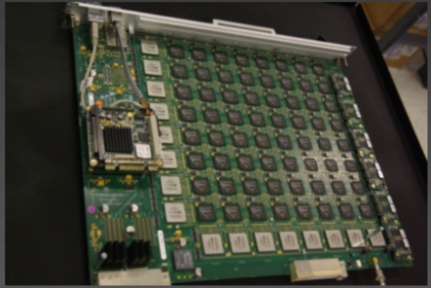
Samples of Data Processing Considerations

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 ~500 km WAN

High Performance Computing Facility (HPC)

HPC Processing
2024: 300 PFlop
2030: 30 EFlop

in: 20 EB in -> out: 100 TB (10^5 reduction)

SDP top-level compute challenge

Telescope Manager

Science Data Processor

SDP Local Monitor & Control

CSP

Data Processor

High Performance
~100 PF/sec
~200 PB/sec
Data Intensive
~100 PB/job
Partially real-time
~10s
Read-intensive
Constrained

Long Term Archive

EB volume

100PB annually

Infrequent access

Delivery System

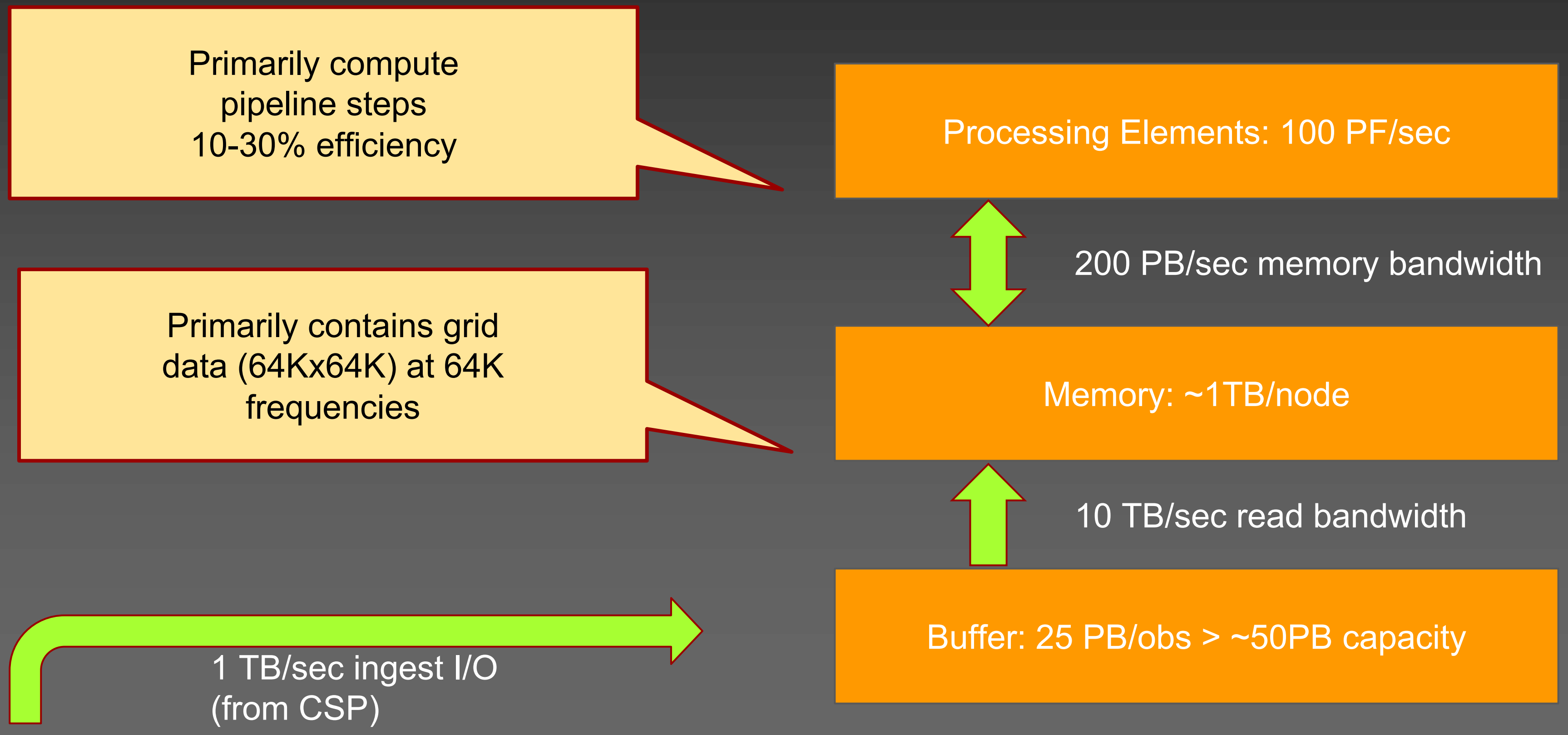
Distribution
~100PB/y

From Cape Town & Perth to rest of World

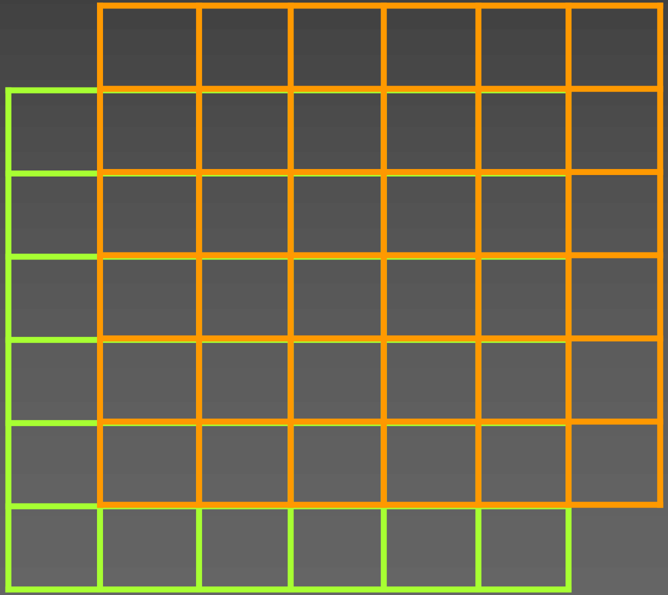
Visualization of 100k x 100k x 100k voxel cubes

Regional Centres & Astronomers

Data Movement in Science Data Processor

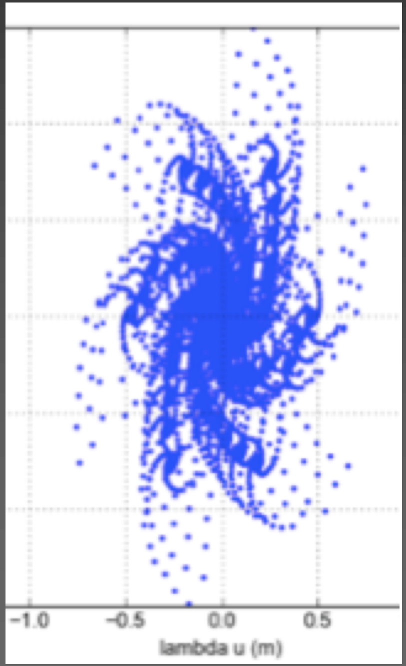


Visibility gridding & cache re-use



Time (earth) rotation of UV grid.

Only fetch edges
Re-use core



Early supercomputer parameters

2025	LFAA (AU)	Mid (SA)
FLOPS	100 PF	360 PF
Memory bandwidth	200 Pb/sec	200 Pb/sec
Buffer Ingest	7.3 TB/s	3.3 TB/s
Budget	45 M€	45 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

Memory Bandwidth
- Cost
- Energy
- 10x first ExaFlop targets

Memory … SKA’s biggest challenge

High Bandwidth Memory (HBM) is becoming dominant for HPC

In 2013 the problem looked perhaps out of reach

HBM is 2.5/3D, on package, memory technology, 10x BW of RAM

Delay in SKA the deliverables has been very helpful

A **domain specific accelerator** similar to Google’s TPU v3 could deliver SKA bandwidth with 10,000 nodes. (TPU’s are 10x ahead of GPUs for this problem)

UNUM's to the rescue?

- Precision and error analysis went out of fashion (very active in 1960's).
    - World converged on IEEE 64 bits
- Compression of non-meaningful error data is difficult
- These issues are now beginning to get attention they need

- For SKA:
- Antenna signals from telescopes have roughly 1 bit of information
- Correlator output – 32 bit is planned, but much less would suffice
- Science Data Processor – planned 64 bit, but requires much less

Opportunity – data movement!

- Build system to **move** data of appropriate precision
  - 50%-75% reduction in memory bandwidth (200PB/sec -> 100PB/sec)
  - 50% reduction in IO bandwidth (10TB/sec -> 5TB/sec)
  - 50% reduction in wide area networking bandwidth (1TB/sec -> 0.5 TB/sec)
- We are talking about significant percentages of the computing, storage and WAN cost.
- Both capacity and bandwidth are affected

Opportunity – compute accurately

- Long term: mixed precision UNUM systems, well understood accuracy
  - POSIT roadmap

- Short / intermediate term
  - Use spare compute cycles (70%) to convert from wire format to high IEEE precision, after memory bus transfers.
  - Make sure these big numbers don’t go back to RAM, only use them during compute (requires cache controls – sorely missing high level API !!!).
  - Compute accurately, write back results in appropriate but much smaller number formats

Evaluate

- I formed a group of experts working on this
  - John Gustafson, telescope experts, pipeline experts, compiler experts
  - Selected evaluation case: Fourier transforms of large 2D images (wide applicability !!!)
  - Target: demonstrate no flaws in thoughts, suggest strategy of using number formats
- And for the POSIT community
  - This is a grand challenge problem – there will be many similar ones
  - Note that bit level (not 2^n bit level) precision & range control have serious HW cost impact

Conclusions

Computing is extremely central in SKA, well beyond the instrument

e.g. applying AI / ML to analyzing the science data

Astrophysics has everyone's attention – this project must succeed

SKA will succeed based on astrophysics

but its computing lies on the frontier of big data handling

POSIT arithmetic could likely play a key role: cost reduction and agile architecture

Thank you.

questions?

skatelescope.org

peter@braam.io