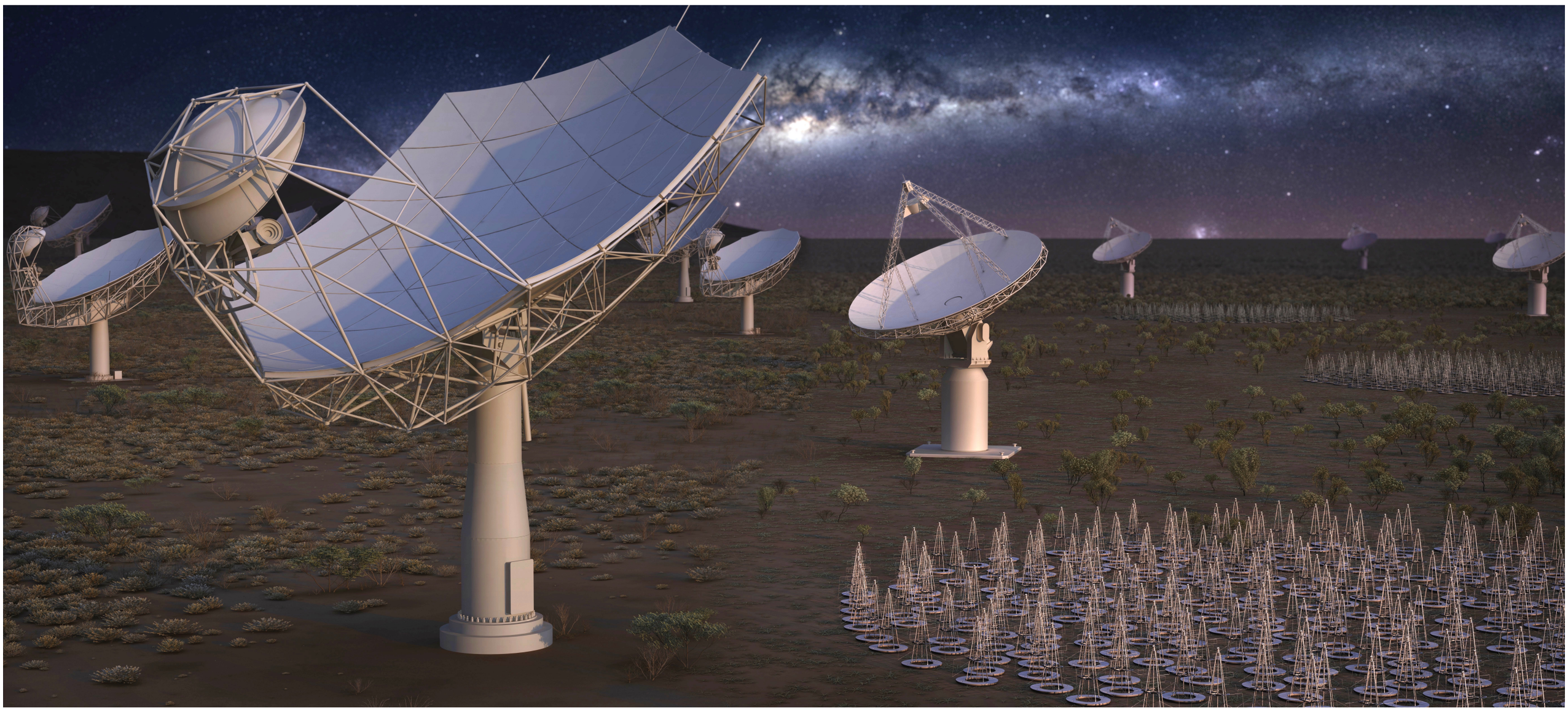
Project Update

SKA, SRCCG and SDP



SQUARE KILOMETRE ARRAY

Exploring the Universe with the world's largest radio telescope

Antonio Chrysostomou & Rosie Bolton

Square Kilometre Array Organisation

www.skatelescope.org

a.chrysostomou@skatelescope.org

@astroant

Outline

Progress update from SKA office

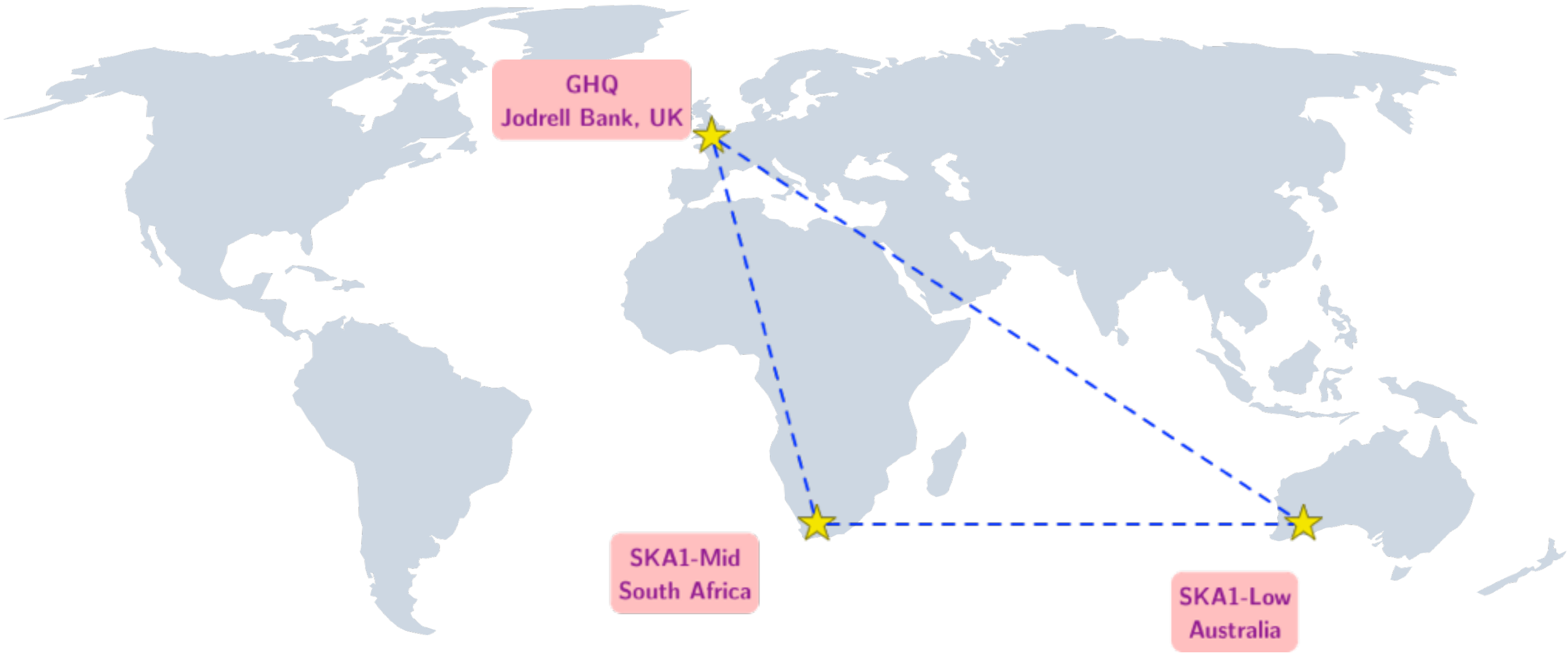
Latest news and documentation from SRCCG

News and progress from SDP

10 member countries: AUS, CA, CH, IN, NZ, RSA, SWE, NL, UK

Currently in discussion with others: FRA, GER, JP, KOR, POR, SPA, SWI

Operational Model



1 Observatory

The SKA

2 Telescopes

SKA1-LOW
SKA1-MID

3 Sites

Australia (LOW)
South Africa (MID)
United Kingdom (GHQ)

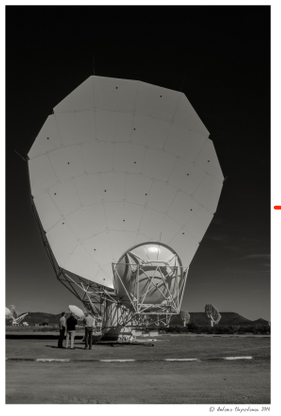
Principles

one observatory
optimal operation
minimise duplication
autonomy & authority

Distributed Operations

SKA1-LOW

SKA1-MID



~ 2 Pb/s

LOW-FREQUENCY APERTURE ARRAY

~ 7 Tb/s

CENTRAL SIGNAL PROCESSOR

300 PB/yr

~ 130 PFlops

SCIENCE DATA PROCESSOR

Memory Bandwidth:

~ 200PB/s

SKA

Regional Centres

~ 9 Tb/s

CENTRAL SIGNAL PROCESSOR

~ 5 Tb/s

SCIENCE DATA PROCESSOR

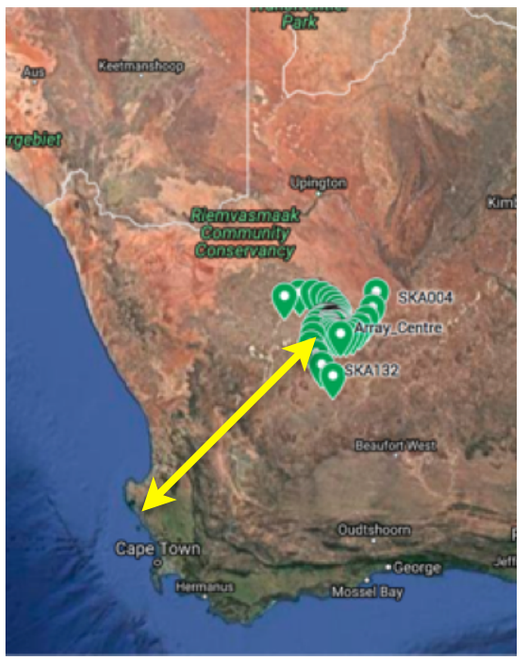
~ 130 PFlops

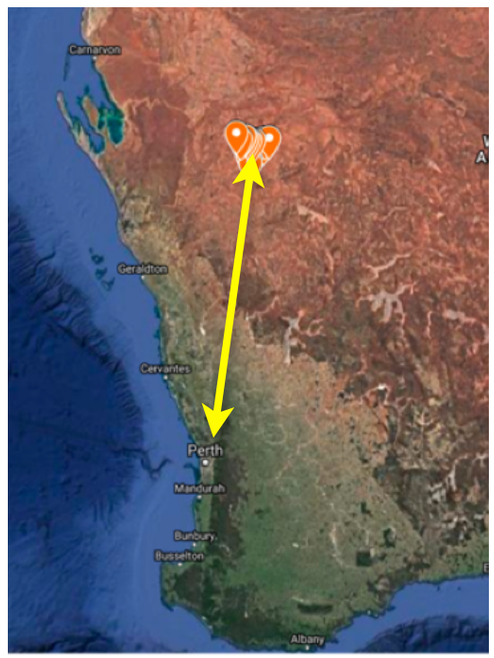
300 PB/yr

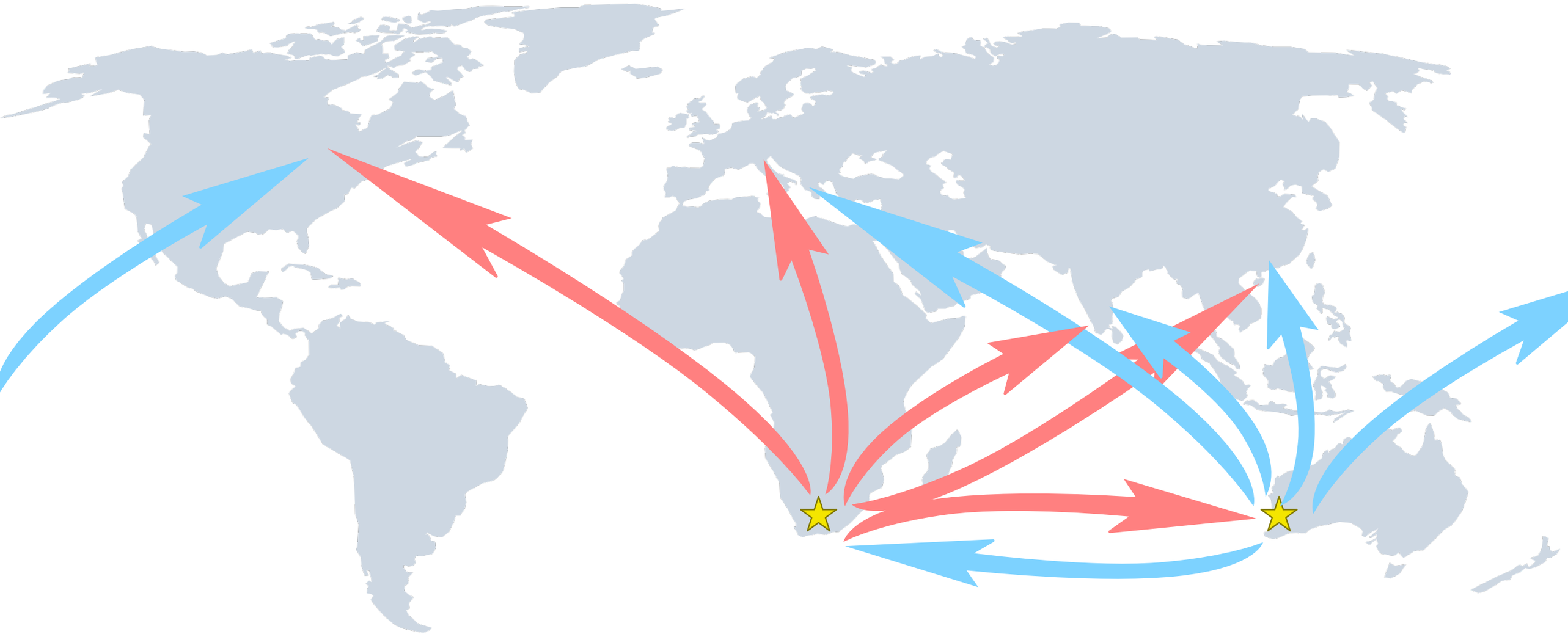
Telescope Operator located ~600km away from the telescope.

Correlator located ~600km away from SDP.

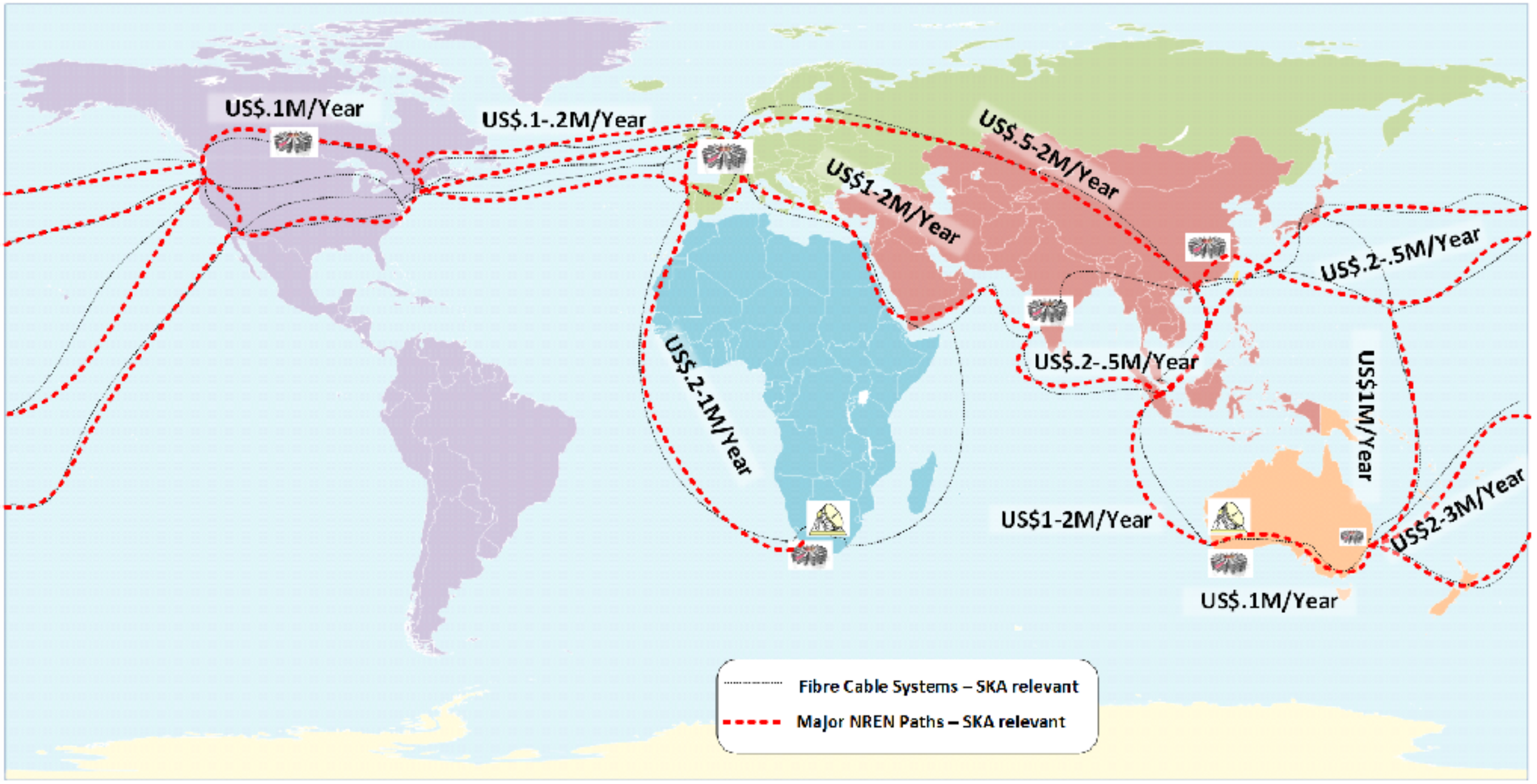
Real-time calibration.





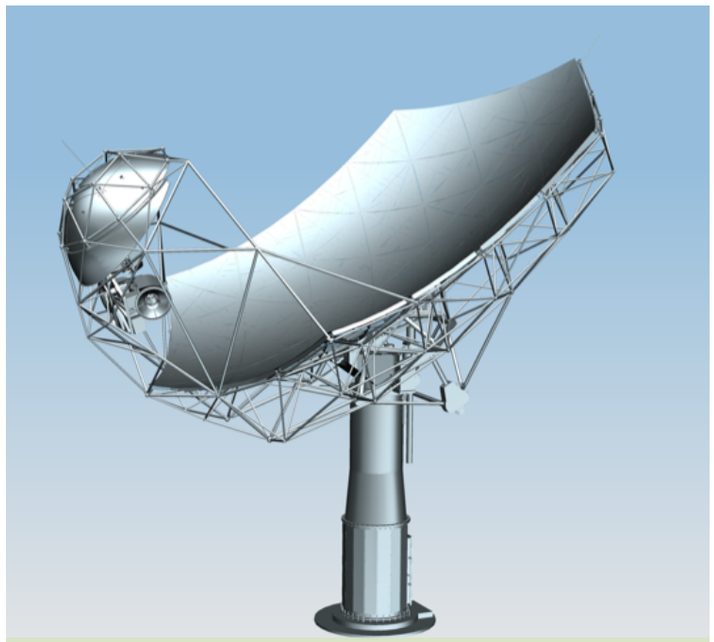


Observatory Data Products flow from the Science Data Processors in Perth and Cape Town to SRCs around the globe

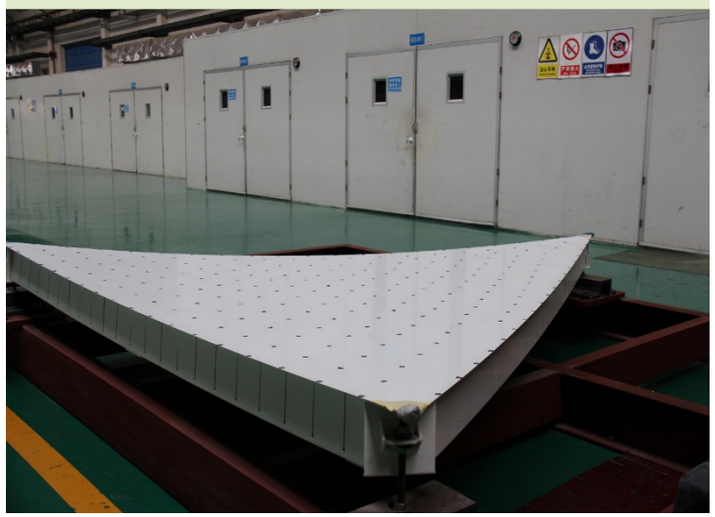


Observatory Data Products flow from the Science Data Processors in Perth and Cape Town to SRCs around the globe

Technical Progress



Chinese (CETC54)/German (MTM) Dish Design



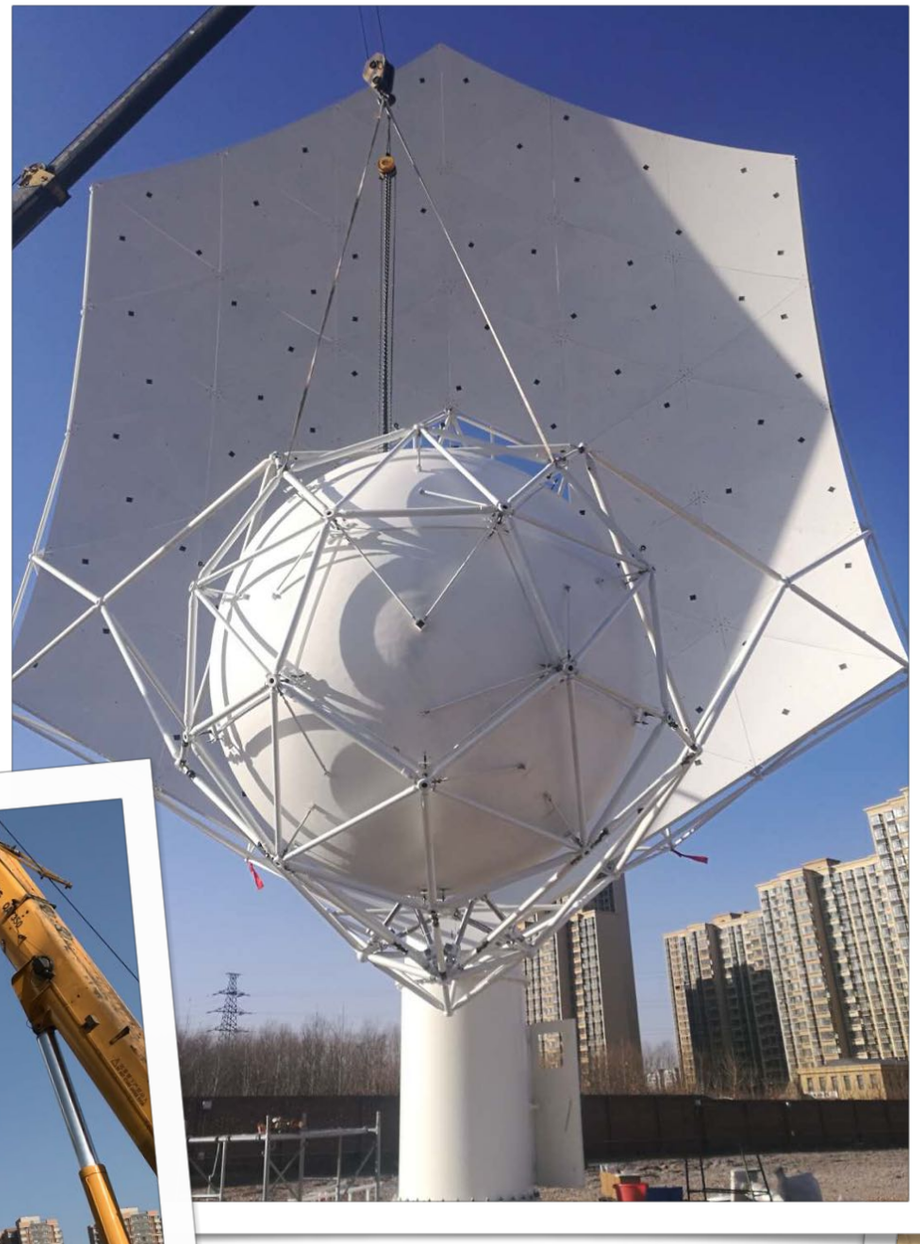
China : Subreflector

Italy: Feed indexer

Technical Progre

First prototype dish constructed

China/Germany/
Italy/South Africa



Technical Progress

SKA-LOW prototype antenna station deployed

Solar power station: 2.6 MW-hr Lithium-ion battery

Design/Deployment Baselines

	Design Baseline	Deployment Baseline	Re-instatement '+' means add to system
SKA1-Mid			
No. dishes	133 + 64*	130 + 64*	+3 dishes at 150 km
Max. Baseline	150 km	120 km	+ infra to 150 km
Band 1 Feeds	133	130	+3 Band 1 Feeds for 3 dishes
Band 2 Feeds	133	130	+3 Band 2 Feeds for 3 dishes
Band 5 Feeds	133	67	+66 Band 5 feeds
Pulsar Search (PSS)	500 nodes	375 nodes	+125 nodes
SKA1-Low			
No. stations	512	476	+36 stations (18 stns at 49 & 65 km)
Max. Baseline	65 km	40 km	+infra to 65km
Pulsar Search	167 nodes	125 nodes	+42 nodes
Common			
Compute Power	260 PFLOPs	50 PFLOPs	+210 PFLOPs

* 64 MeerKAT dishes

	Cost €M
Design Baseline	798
Deployment Baseline	675
Operations (/yr)	89
Ops Deployment (/yr)	77

Result of Cost Control Project in 2017 & approved by Board, July 2017

Design baseline for which CDRs will be carried out is unchanged

- the Deployment Baseline is scoped to match the construction budget
- commitment to re-instate all items in the Design Baseline as funding permits

SRCCG Progress

In Jun17 we released Rev02 of the Background and Framework document

- provides high-level view for interested parties

In Sep17 the **first** version of SRC requirements were released

A register of risks has been developed and will be regularly reviewed

Timeline of milestones established

SKA REGIONAL CENTRES: BACKGROUND AND FRAMEWORK

Document number SKA-TEL-SKO-0000706
Document Type ..... RSP
Revision 02
Author ..... A. Chrysostomou, SRCCG
Date 2017-06-06
Document Classification ..... UNRESTRICTED
Status ..... Released

Name	Designation	Affiliation	Signature
A. Chrysostomou and the SRCCG	SKA Re Cent Coordinati		
A. Chrysostomou	Head of Operations		
G.R. Davis	Direct Operations		
P.J. Diamond	Director		

SKA REGIONAL CENTRE REQUIREMENTS

Document Number SKA-TEL-SKO-0000735
Document Type ..... RSP
Revision 01
Author ..... R. C. Bolton and the SRCCG
Date 2017-09-27
Document Classification ..... UNRESTRICTED
Status ..... Released

Name	Designation	Affiliation	Signature
Authored by:			
Rosie Bolton et al.	SRC Project Scientist	SKAO	R Bolton
Owned by:			
Antonio Chrysostomou	Head Science Operations Planning	SKAO	Antonio Chrysostomou
Approved by:			
Gary Davis	Director of Operations Planning	SKAO	GRD
Released by:			
Phil Diamond	Director General	SKAO	

Selection of SRC Requirements

Membership:

- Membership of the SRC alliance will be awarded if individual prospective SRCs meet and maintain all the criteria set out in appropriate MoUs. The ability of each SRC to meet its criteria, and the criteria themselves will be reviewed annually (TBC).

Interfaces:

- Interfaces between each SRC and the SKAO will be compliant with policies set out by the SKAO

Data Policies:

- Each SRC will preserve and make available to users, the SKA science data products, in adherence to SKAO data access policies and data security standards

Reproducibility, provenance & workflow:

- Each SRC must be capable of saving the complete workflow and provenance associated with any ADP, in such a way that they can be queried, viewed and the associated workflows can be re-used to create new ADPs

Selection of SRC Requirements

Open access:

- The SRC Alliance will enable users to provide public links to SKA science data products in their research publications. Published and non-proprietary data must be publicly available

Storage capacity:

- The SRC Alliance must provide a bare minimum of 600 PetaBytes (TBC) of storage at the start of SKA1 operations

These requirements we expect to evolve over time as we learn more from SRC design work from various partners around the globe

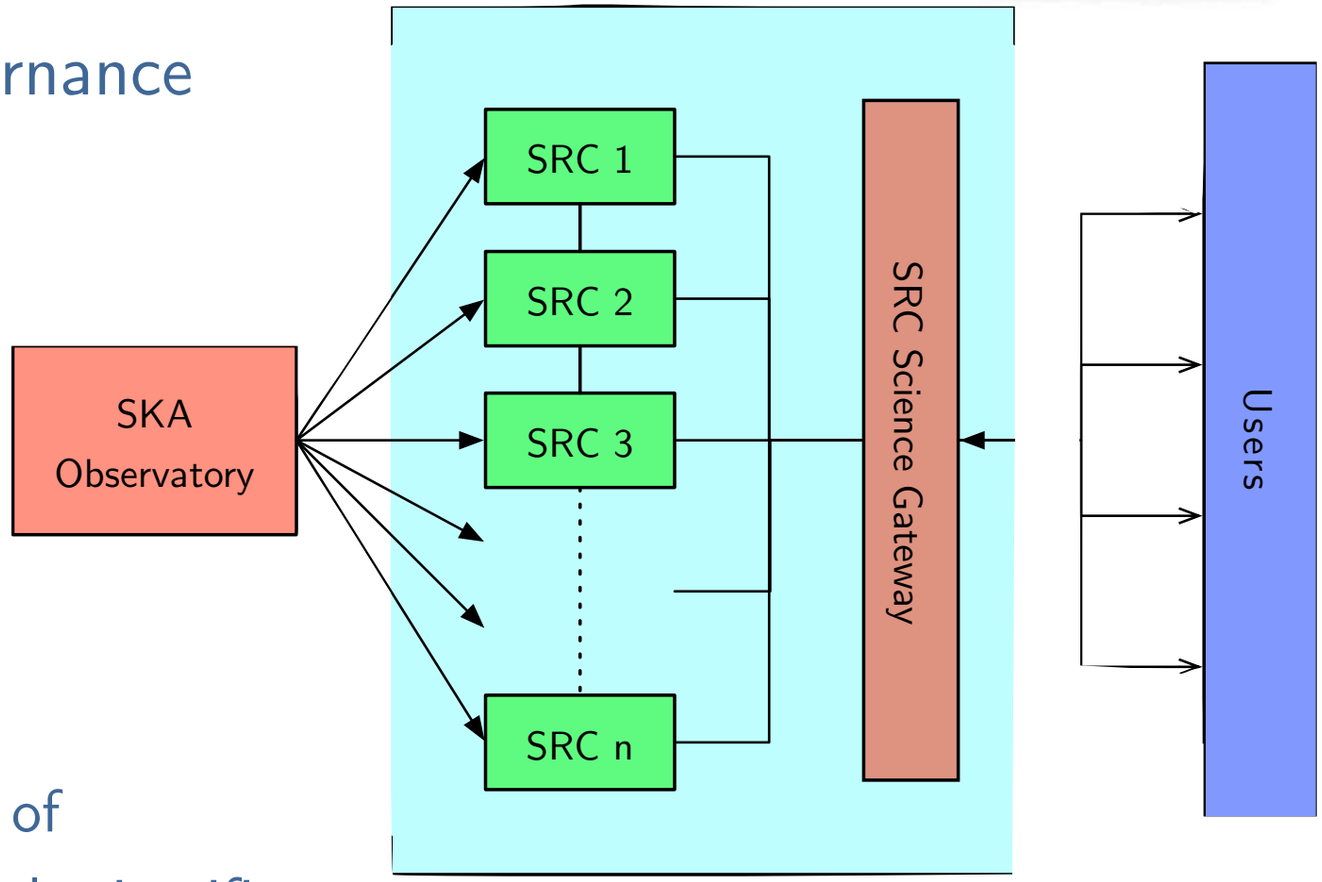
`https://astronomers.skatelescope.org/documents/`

SRC Governance

We are now working to define the governance model for the network of SRCs

Roles and responsibilities of the SRCs to each other, the SKA Observatory and vice-versa

- not forgetting the most important aspect of their collective responsibilities to the global scientific community



Discussions on this topic are planned at this meeting

All feedback is welcome!

SRCCG Members

SKAO Members

- Antonio Chrysostomou - Chair
- Rosie Bolton - SRC Project Scientist
- Miles Deegan
- Nick Rees

Member Representatives

- Séverin Gaudet - Canada
- Jeremy Main - South Africa
- Peter Quinn - Australia
- Yogesh Wadadekar - India
- Michael Wise - Europe
- Shenghua Yu - China

External Advisory Members

- Ian Bird - CERN
- Andy Connolly - LSST
- Lourdes Verdes-Montenegro - IAA/Spain

SDP and the SRCs

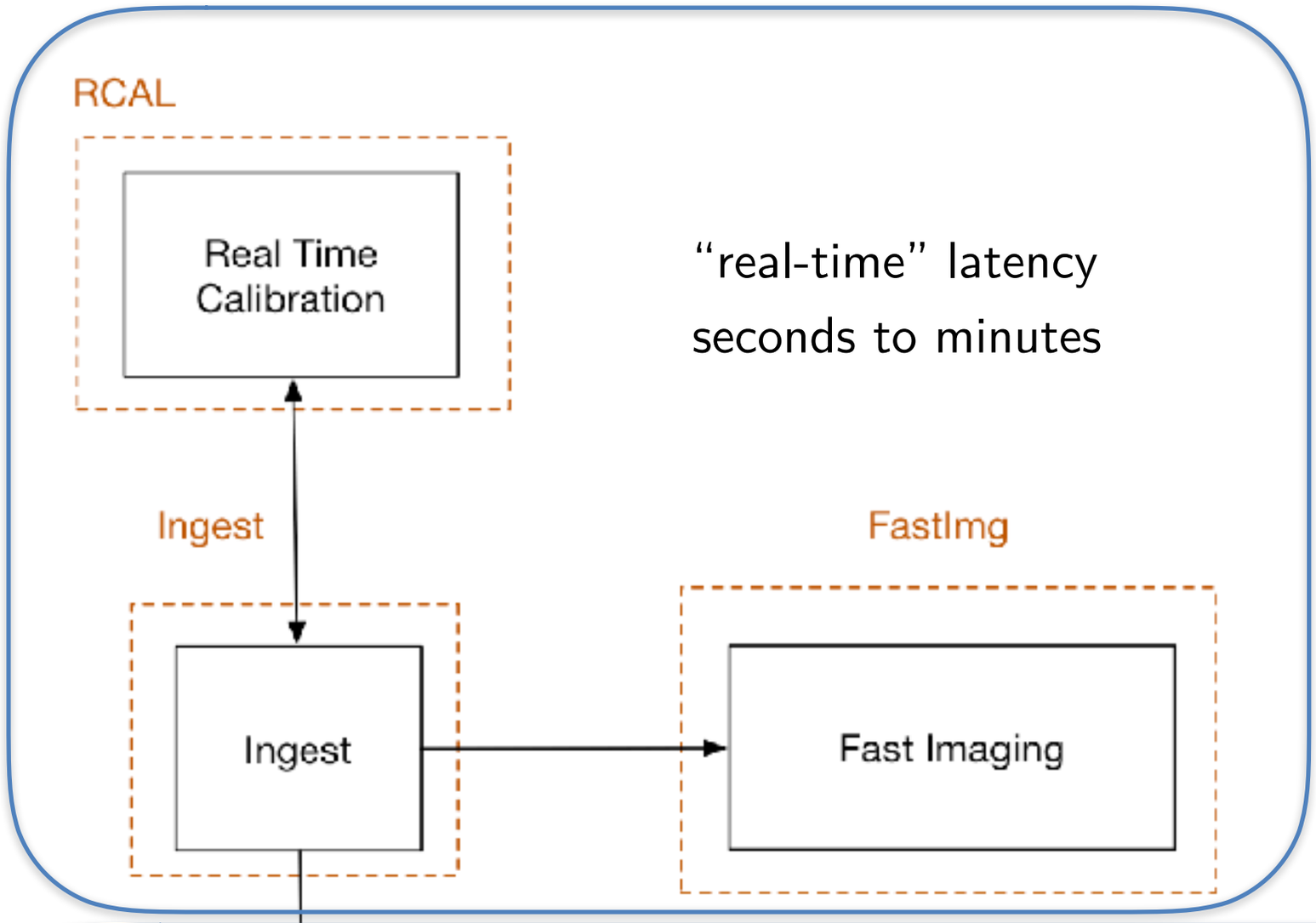
SDP workflow recap

SDP data products

Regional centre data products and pipelines

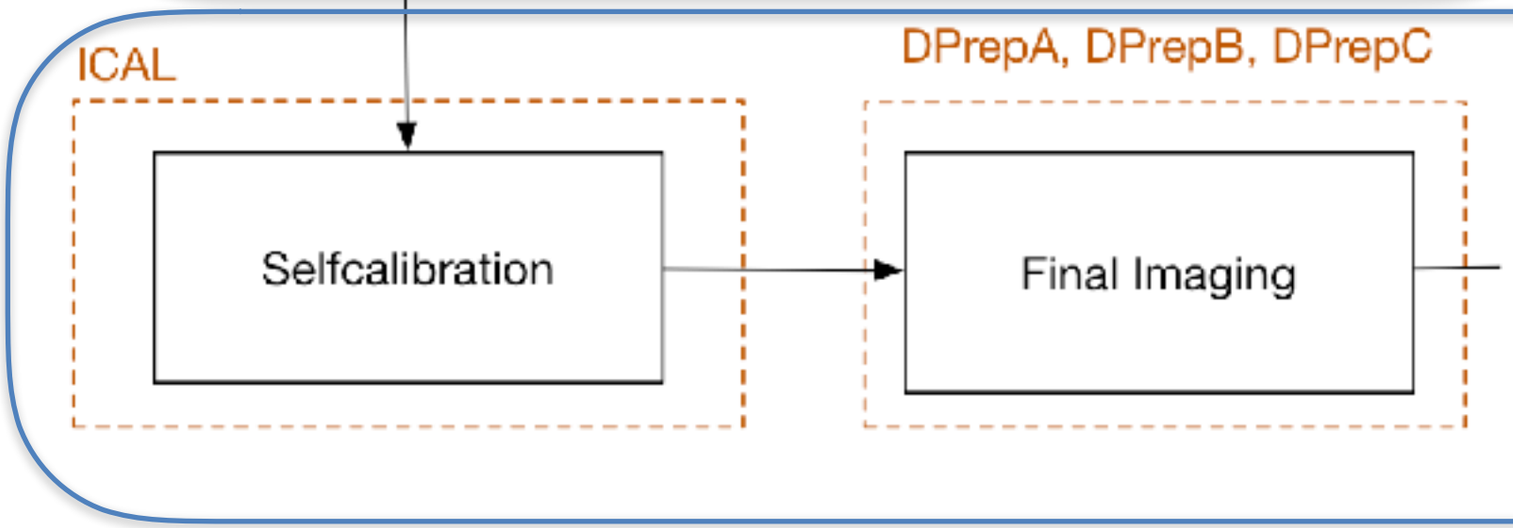
SDP system pressure

SDP Workflow - visibility domain



"real-time" latency

seconds to minutes

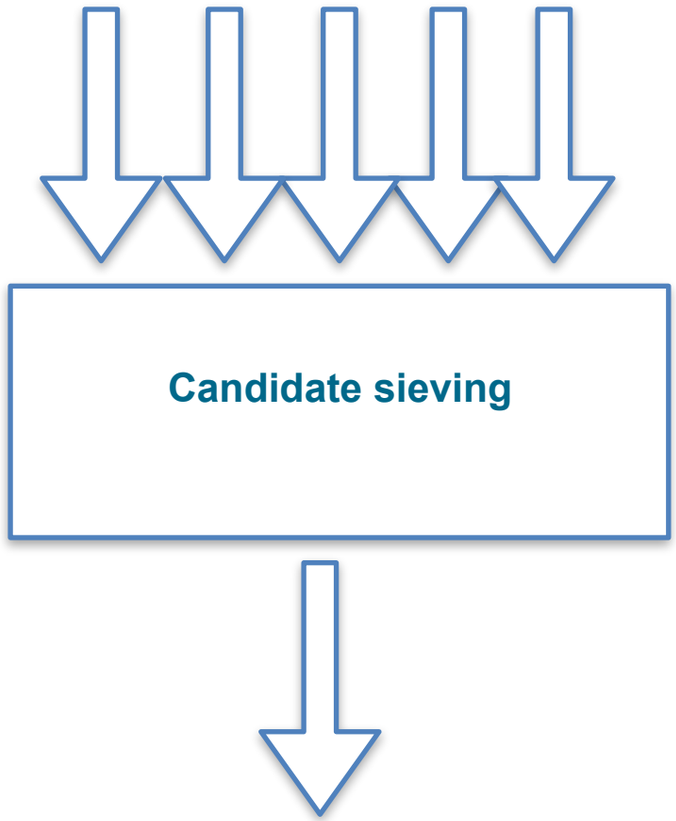


Off line batch system;

latency up to several days

(or longer)

SDP workflow - time domain



Candidate sieving performed off-line

Real time calibration working on visibility data - calibrates array delays to keep pulsar beams coherent

SDP dataflow

1: Raw data from Central signal processor

2: Conditioned data (e.g. some RFI flagging, re-ordering) and possibly calibration solutions

3: Observatory data products

4: Advanced data products

SKA Science Data Products

Two types of science data products:

Observatory data products (3 on last slide) are generated by the observatory, in the SDPs. We have a list of these...

Calibrated Visibilities, **LSM Catalolgue**, Transient Source Catalogue, **Pulsar Timing Solutions**, Transient Buffer Data, Sieved **Pulsar and Transient Candidates**, Science Product Catalogue,

Taylor term continuum image

Residual image in continuum

Clean component image in Taylor terms

Spectral line cube after continuum subtracted

Residual spectral line image (i.e. residuals after clean applied)

Representative Point Spread Function for observations

Gridded visibilities: Calibrated visibilities, gridded onto grids at spatial and frequency resolution required by experiment.

Gridded visibility weights: Accumulated Weights at each uv cell in each grid

Advanced data products

Advanced data products (4) will be generated by the users in the SRCs. When "finished" these will be included in the discoverable SKA archive which will be housed in across the SRCs. We need to understand all of these so we can work out what the SRC workflows look like.

Input from the community is needed to determine what these products should be.

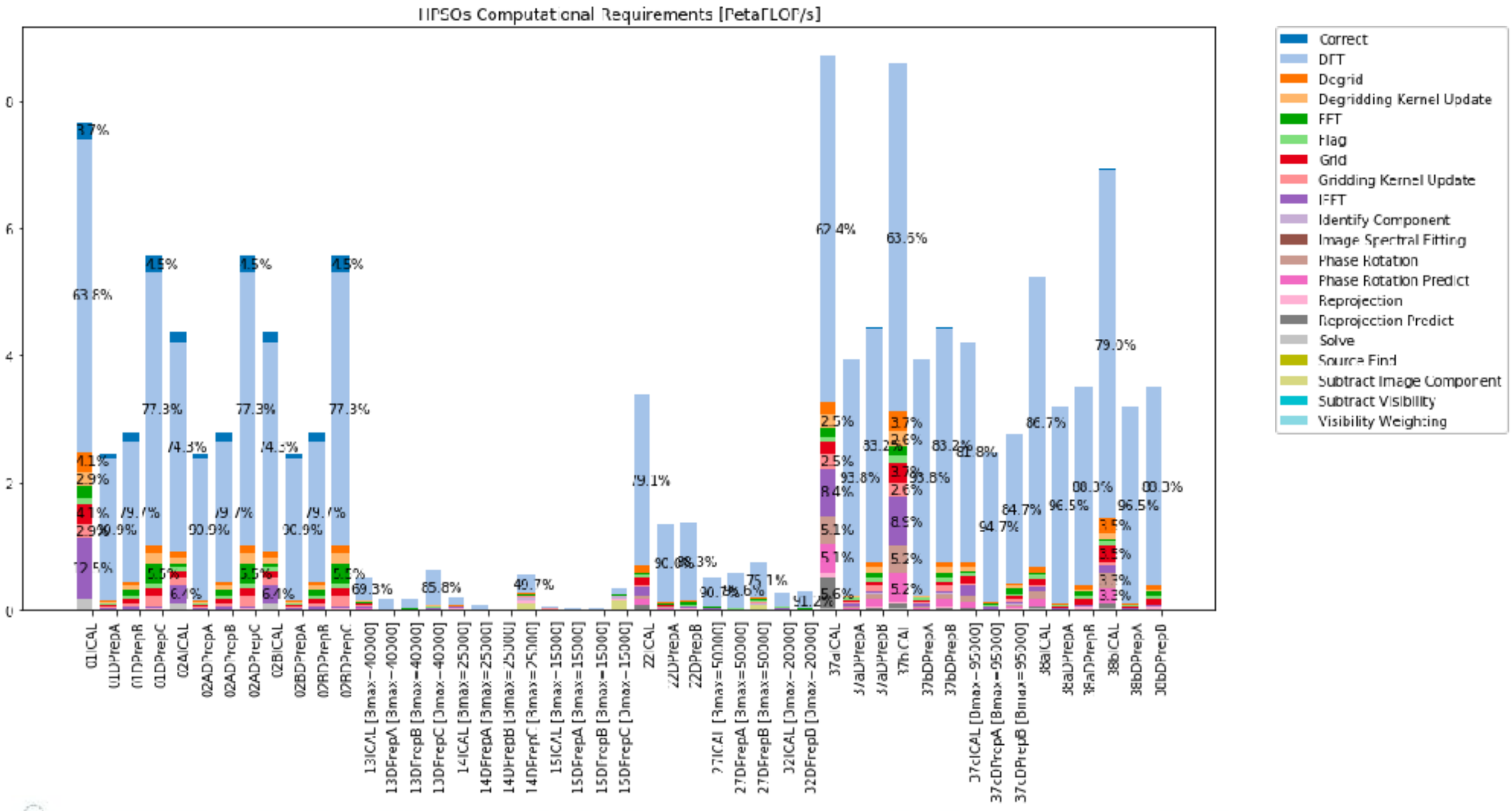
List does not need to be complete - must be able to generate these products from the ODPs...

Co-added images (in all the same formats as the ODPs in principle)
Rotation measure maps
Moment maps
Power spectra
Cutouts
Pulsar timing residuals
Variance maps
Spatially integrated line spectra
Cross-matched catalogues
Shear maps

The pipelines for generating these products from the ODPs will be developed by the user community, presumably with guidance and input from user support at SRCs

SDP - a constrained system

System sizing for SDP is based on parametric model and has many assumptions about the pipelines employed by different experiments.



SDP - a constrained system

System sizing for SDP is based on parametric model and has many assumptions about the pipelines employed by different experiments.

Overall system size estimate: 259 PFLOP/s

- Building this system would require experiments scheduled to have average compute cost less than 259 PFLOP/s - some can be higher but only if others are lower
- SDP is a schedulable resource that must be considered by the TAC

Deployment baseline is currently: 50 PFLOP/s

- Building a system 5x smaller does not exclude any specific pipeline within the SDP *BUT* the system constraints will be very strong
- Demanding experiments will face tough scrutiny and need to promise extremely good science results in order to get scheduled.
  - "Demanding" here can be in terms of antenna time, SDP resource needs, data product size or any other aspect that has implications on finite resources (time, money, power)
  - Globally too, experiments will need to consider the SRC processing and storage requirements as part of the overall cost of science delivery