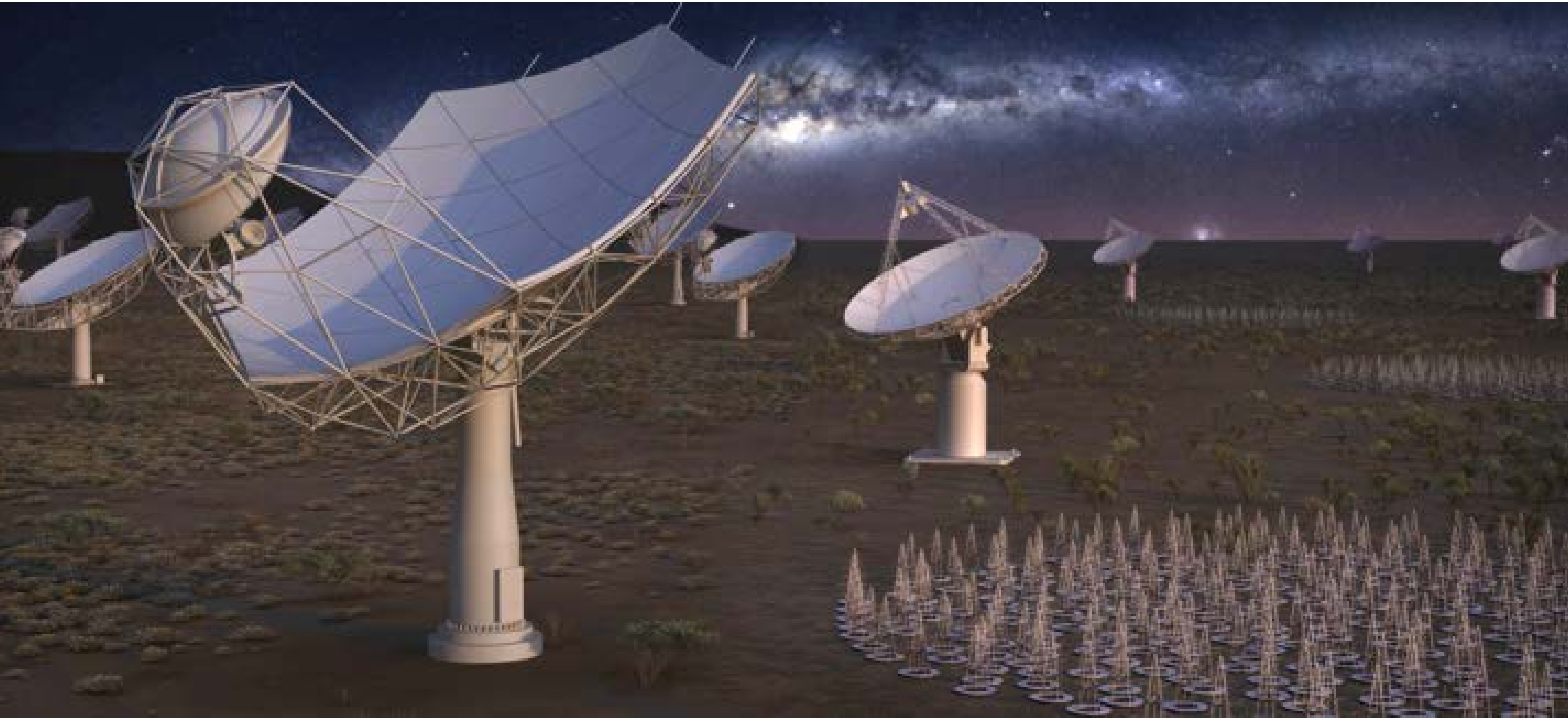
SKA Regional Centres

Background and Framework



SQUARE KILOMETRE ARRAY

Exploring the Universe with the world's largest radio telescope

Dr Antonio Chrysostomou

Head of Science Operations Planning

(a.chrysostomou@skatelescope.org | @astroant)

Outline

Introduction to the Square Kilometre Array

The data flow that drives us to a model for SKA Regional Centres

Model for collaborative network of SRCs

The SKA Regional Centres Coordination Group

One Observatory

The Square Kilometre Array

Two Telescopes

SKA-LOW

SKA-MID

Three Sites

Australia (LOW)

South Africa (MID)

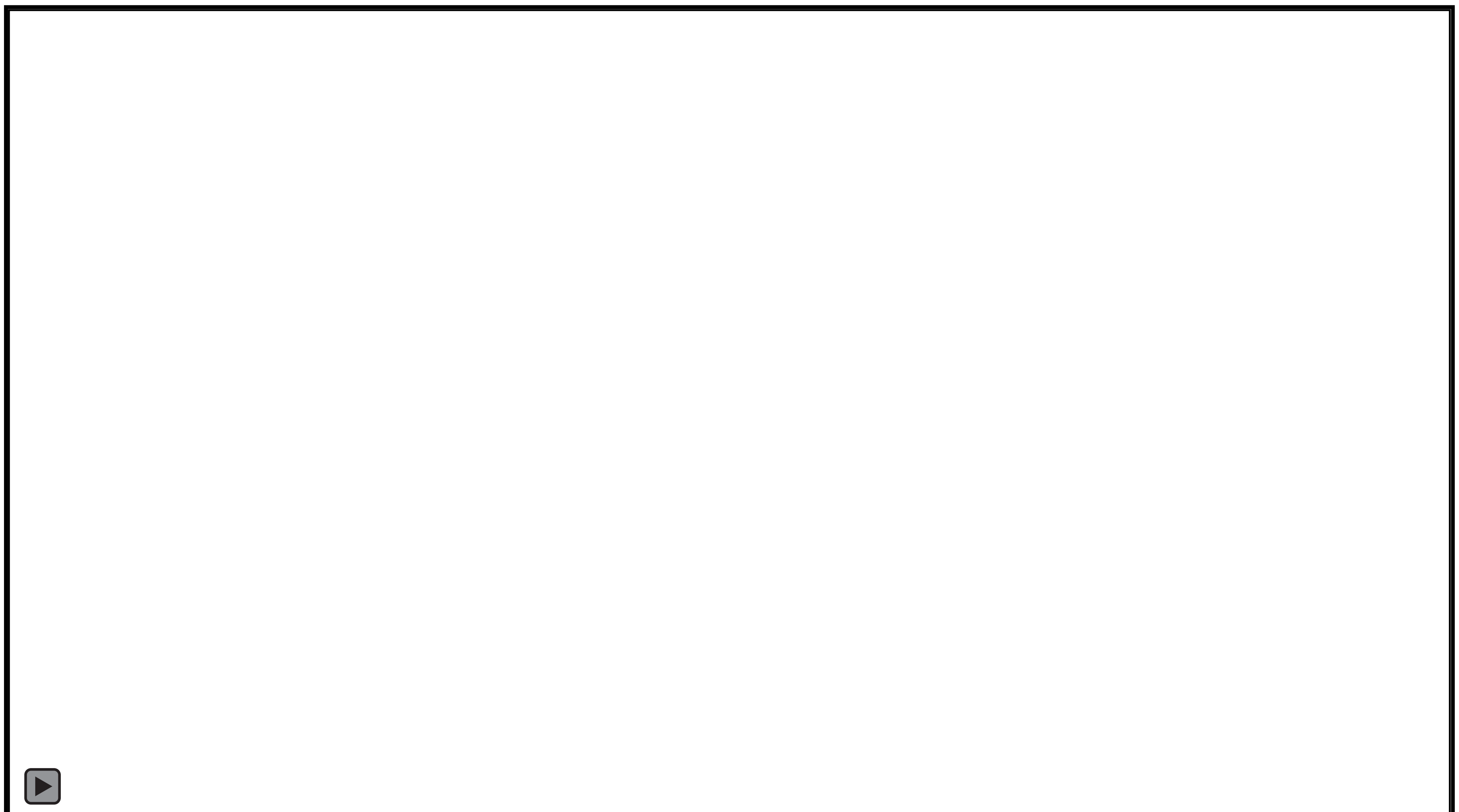
UK (GHQ)

SKA1 LOW - Western Australia

131,072 antennas : 512 stations of 256 antennas, core + 3 spiral arms, 65km baselines

50 → 350 MHz full instantaneous bandwidth

Raw Data output approx. 2 Pbit/s → 7 Tbit/s into the correlator

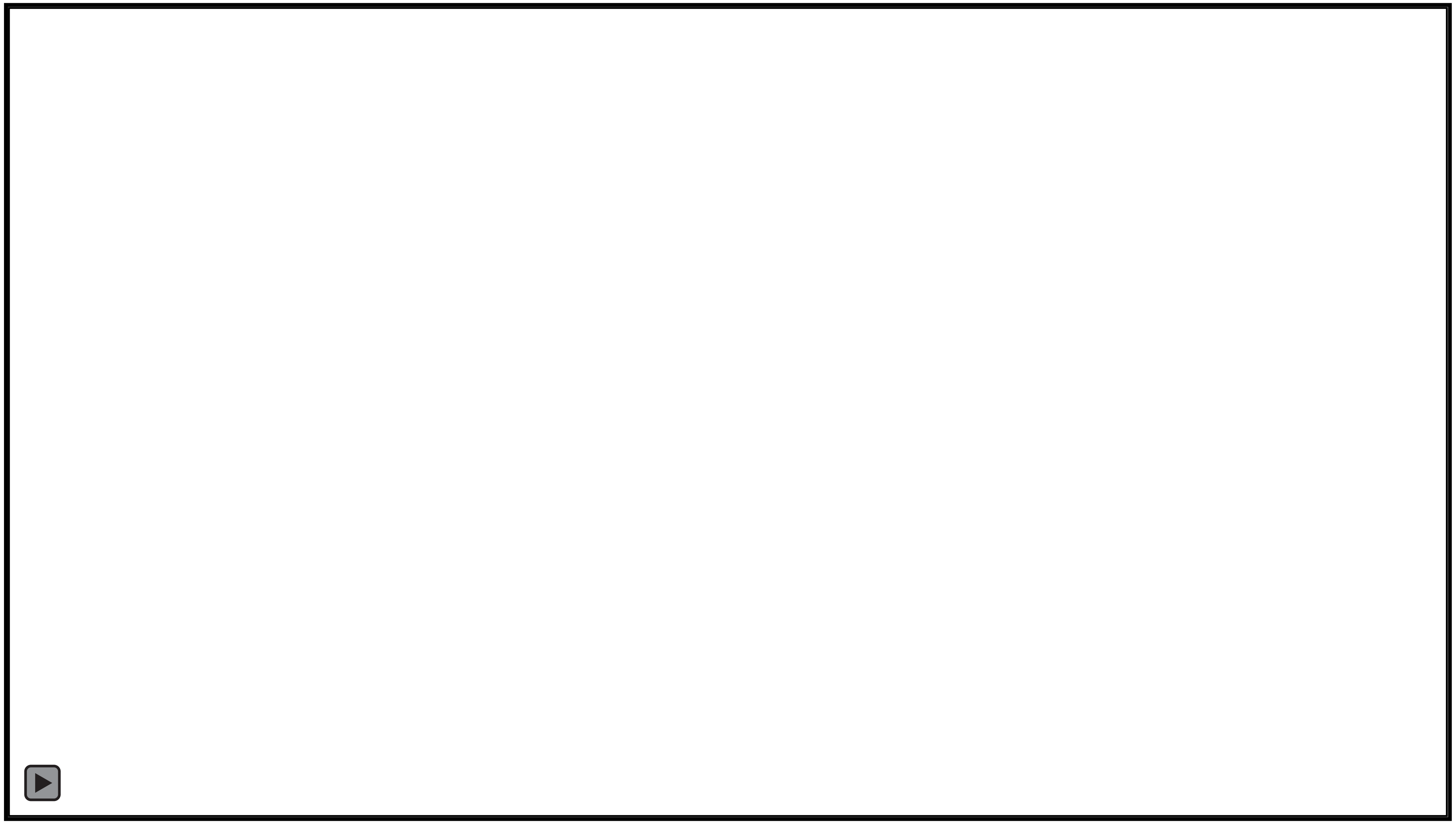


SKA1 MID - Karoo, South Africa

133 SKA1 dishes (15m), 64 MeerKAT (13.5m), core + 3 spiral arms, 150km baseline

0.35 → 15GHz covered in 5 bands

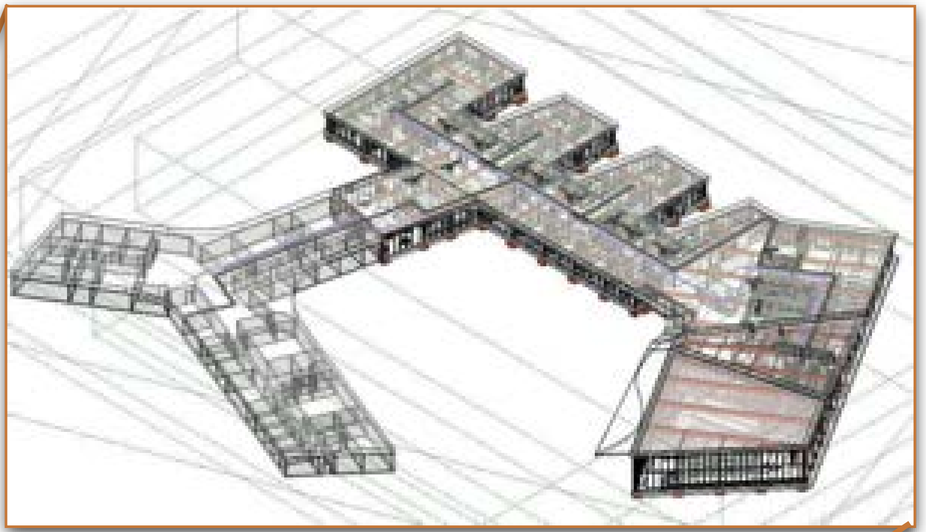
Raw Data output approx. 9 Tbit/s into the correlator



SKA GHQ

New building, starting this year to be completed by May 2018

State of the art facilities



Data flow

SKA1-LOW

~2 Pb/s

LOW-FREQUENCY APERTURE ARRAY

7.2 Tb/s

SKA1-MID

8.8 Tb/s

SIGNAL AND DATA TRANSPORT

CENTRAL SIGNAL PROCESSOR (CSP)

~50 PFlops

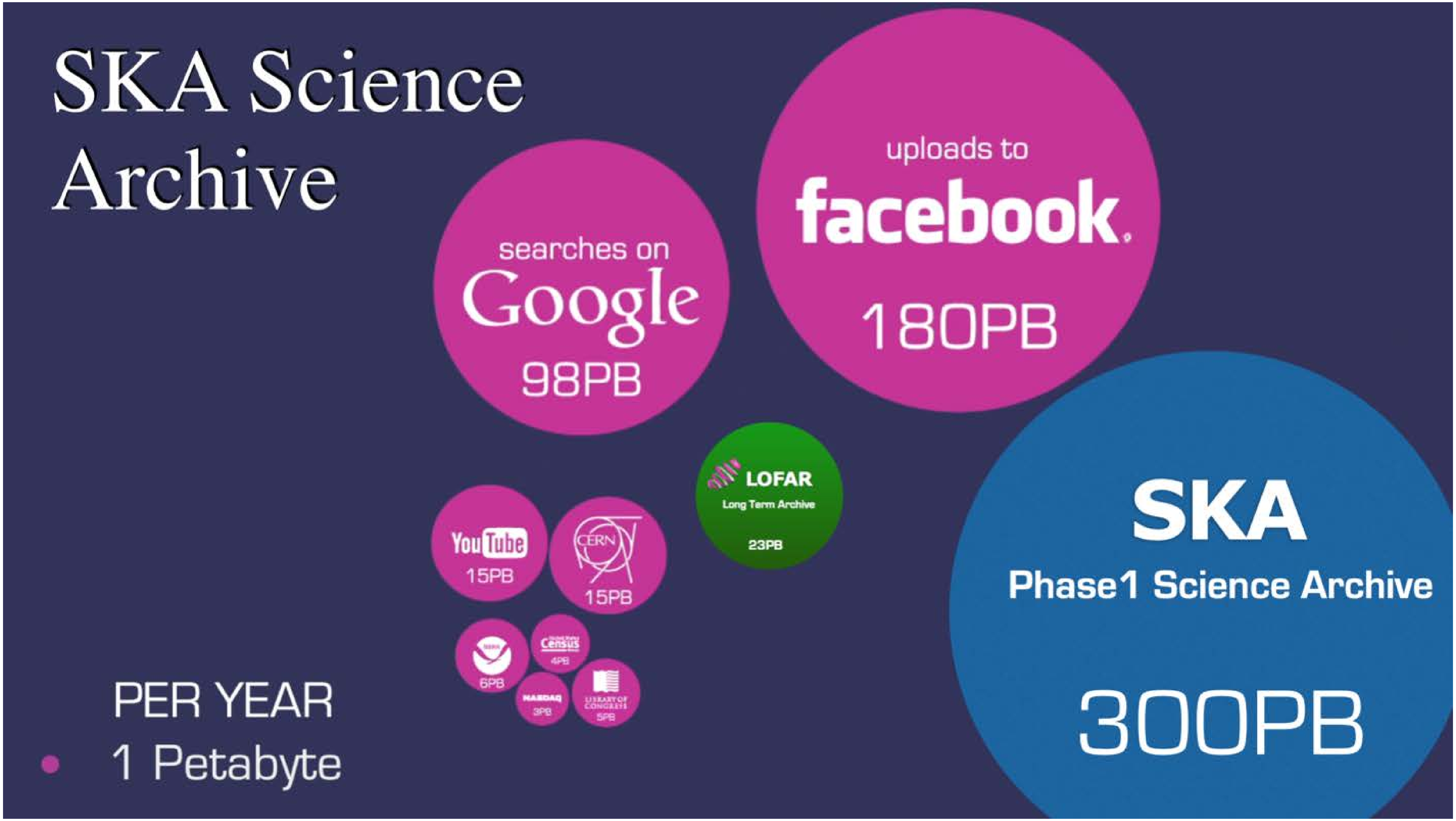
~5 Tb/s

SCIENCE DATA PROCESSOR (SDP)

~250 PFlops

~300 PB/yr

Some perspective



A collaborative model of SRCs

Three main factors that lead to a model of a collaborative network of SRCs

(1) The science data products that emerge from the SKA observatory are not in the final state required for science analysis

(2) The data volumes are so large that direct delivery to end users is unfeasible

(3) The community of scientists working on SKA science data products will be geographically distributed

A collaborative model of SRCs

Three main factors that lead to a model of a collaborative network of SRCs

(1) The science data products that emerge from the SKA observatory are not in the final state required for science analysis

generation of advanced data products not in scope of project

SDP must maintain throughput matched to input data rate

combination & further analysis of data products outside of observatory boundaries

(2) The data volumes are so large that direct delivery to end users is unfeasible

(3) The community of scientists working on SKA science data products will be geographically distributed

A collaborative model of SRCs

Three main factors that lead to a model of a collaborative network of SRCs

(1) The science data products that emerge from the SKA observatory are not in the final state required for science analysis

(2) The data volumes are so large that direct delivery to end users is unfeasible

does not account for possible future "discovery" archive

final data volume for each project will exceed that delivered by the observatory

downloading data to local machines/cluster expensive and unfeasible in long term

"take processing to the data"

(3) The community of scientists working on SKA science data products will be geographically distributed

A collaborative model of SRCs

Three main factors that lead to a model of a collaborative network of SRCs

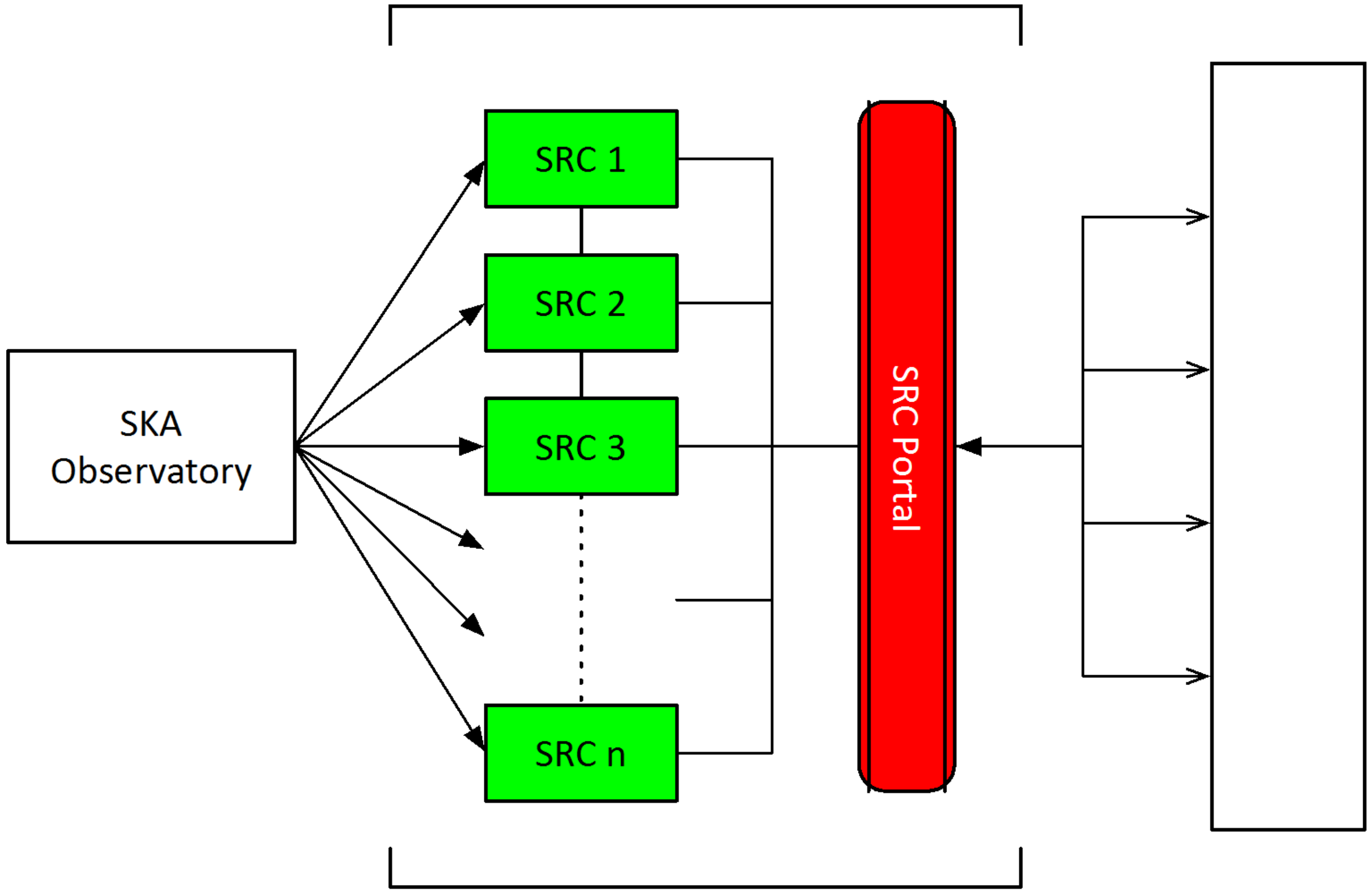
(1) The science data products that emerge from the SKA observatory are not in the final state required for science analysis

(2) The data volumes are so large that direct delivery to end users is unfeasible

(3) The community of scientists working on SKA science data products will be geographically distributed

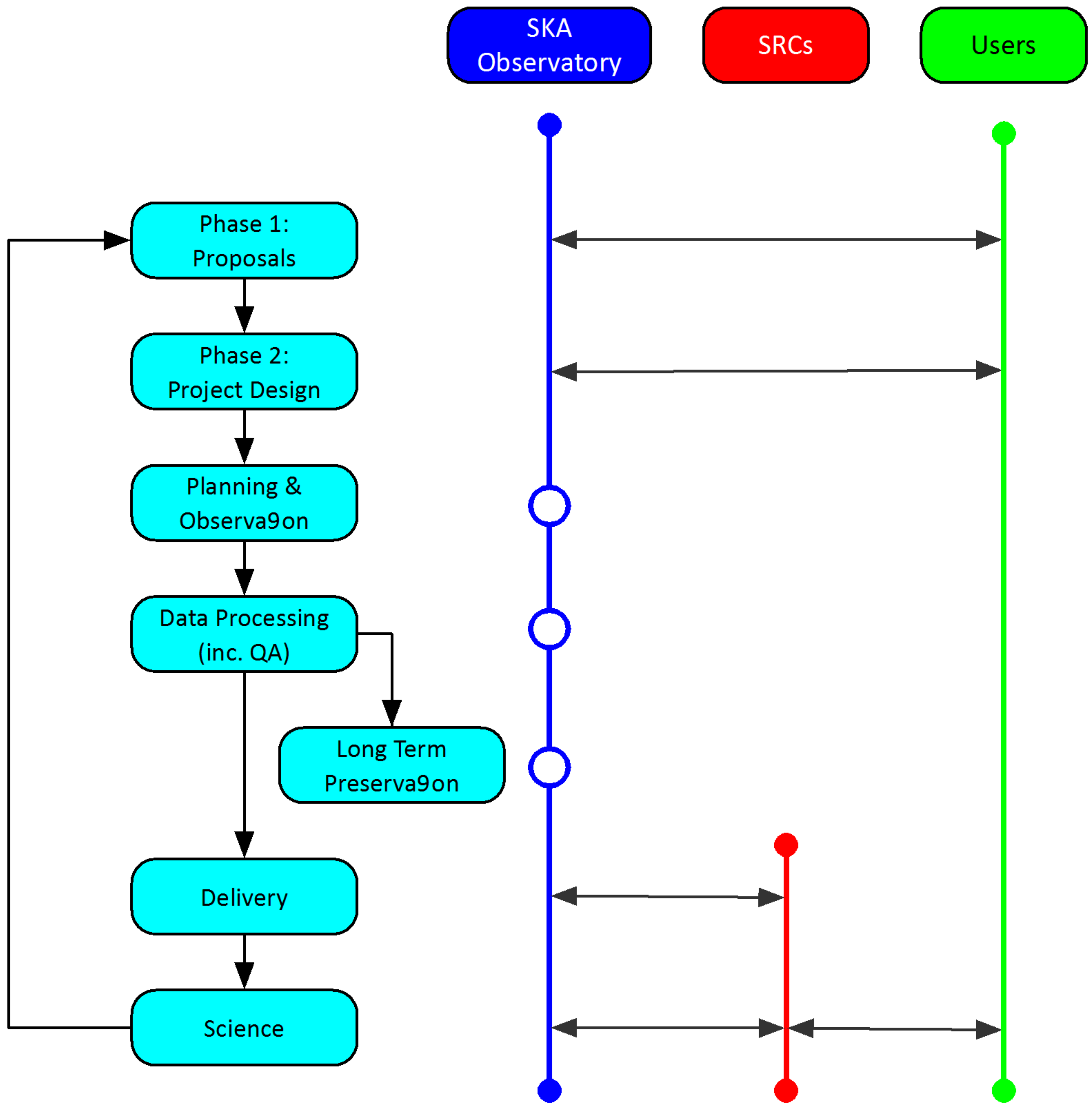
KSPs with 1000s of hrs of observing time will dominate the science programme
large international teams drawn from across the membership
need new methods, algorithms and techniques
driven by the community so they need a platform on which to do this

Model for collaborative network of SRCs



Simplified description but highlights important factors

- a collaborative network
- transparent and location agnostic interface to SRCs for users
  - no SKA user should care where their data products are
  - all SKA users should be able to access their data products, irrespective of whether their country or region hosts a regional centre



The observatory executes the observing projects, the calibration, QA and generates the data products

Users interact with the observatory to write proposals and prepare projects for observing

SRCs receive data products from the observatory and provide resources to users to carry out higher level analysis

SKA Regional Centres Coordination Group

Essential Functions of SRCs:

- provide transparent access to SKA science data products & user support
- provide computational resources for post processing (analysis & visualisation)
- provide platform for development of software tools
- provide long-term science archive

SRCCG (abridged) instructions:

- define minimum set of requirements for SRCs
  - individual and whole network, including links with SKA telescopes
- draft MoUs between SKAO and the SRCs, and an accreditation process
- ingestion and curation of science archive for user-generated data products
- data challenges

Also need to recognise that requirements of KSP and PI projects will differ

SKA Regional Centres Coordination Group

SKAO:

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

Members:

- Séverin Gaudet (NRC, Canada)
- Jasper Horrell (SKA-SA)
- Peter Quinn (ICRAR, AUS)
- Yogesh Wadadekar (NCRA, India)
- Michael Wise (ASTRON, NL)
- Shenghua Yu (BAO, China)

Externals:

- Ian Bird (CERN)
- Andy Connolly (LSST, UWash)
- Lourdes Verdes-Montenegro (IAA, Spain)

Note that this is a Coordination Group and not a Working Group

- subgroups will be formed to study and report on specific issues

SQUARE KILOMETRE ARRAY

Exploring the Universe with the world's largest radio telescope