



# The Square Kilometre Array

Data Transport, Processing, Archiving and Access

Dr Shaun W Amy  
17 September 2019

CSIRO ASTRONOMY AND SPACE SCIENCE  
[www.csiro.au](http://www.csiro.au)



# Section 1: The SKA

What, Where and When

# SKA Organisation

Australia  
Canada  
China  
Germany (MPIfR)  
India  
Italy  
Netherlands  
New Zealand  
South Africa  
Spain  
Sweden  
UK



- Interested Countries:**
- Switzerland
  - Japan
  - Korea
  - ...



**Members**  
Host Countries: Australia, South Africa, United Kingdom



**African partner countries**

# Timeline (SKA Phase 1)

March 2019	Signing of convention that allows the IGO to be established
December 2019	System CDR: Conclusion of design phase
Q2 2020	SKA Observatory
June 2020	Submission of Construction Proposal and Operations Plan
December 2020	Approval and Commencement of Construction (T0)
December 2027/8	Completion of construction

**1 Observatory, 2 Telescopes, 3 Sites**

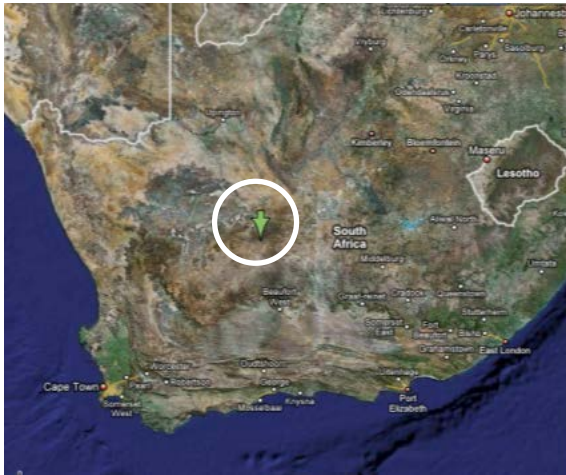
# SKA1-LOW: Western Australia



- SKA1-LOW: 50 – 350MHz
- 131 000 dipoles:
  - 512 stations each with 256 antennas
- 65km baselines.
- Located at Boolardy Station in the Murchison Shire



# SKA1-MID: Northern Cape, South Africa



- SKA1-MID: 350MHz – 14GHz
- 133 SKA dishes + 64 MeerKAT dishes
- 120km baselines.
- Located in the Karoo, Northern Cape

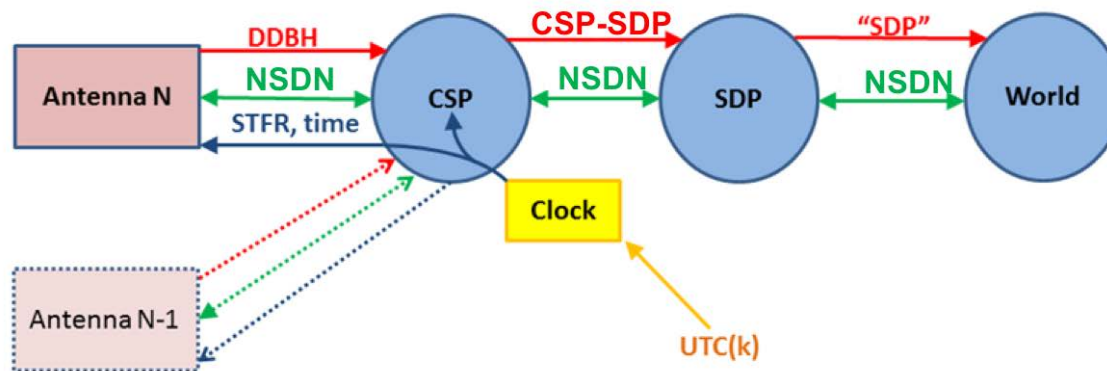


# Section 2: Data

Transport, Processing, Archiving and Access

# SKA Signal and Data Transport (SaDT)

- Three networks:
  - Synchronisation & Timing: UTC, 1pps and frequency reference.
  - Non-Science Data: M&C, general network access, UC etc.
  - Data: Digital Data Backhaul, CSP-SDP and SDP to SKA Regional Centres.



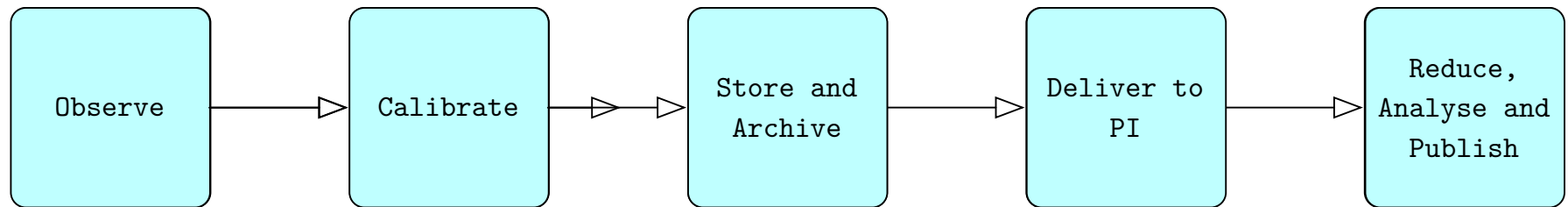


# SKA Phase 1 Data Flows



# Observation to Publication

Historically the Science Data Processor (SDP) would not be considered part of the telescope. The data flow through the SKA is such that this model breaks early in this data delivery chain.

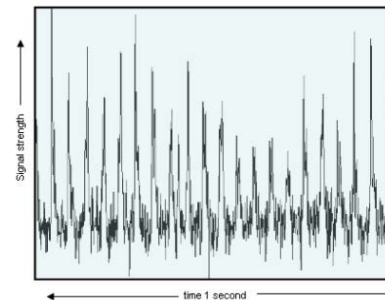
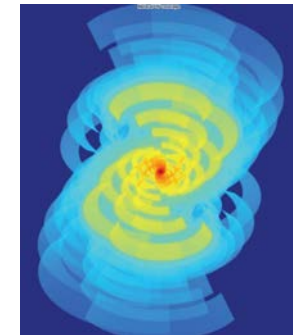
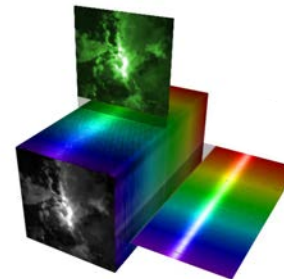


In planning the observing programme of the Observatory the SDP becomes part of the telescope that needs to be scheduled.

**Reduce Data Volume at the SDP**  ***Observatory Data Products***

# Observatory Data Products

- Catalogues:
  - Transient Source Catalogue
  - Science Data Product Catalogue
- Imaging:
  - Image Cubes
  - Gridded Visibilities
- Pulsars:
  - Pulsar and Transient Candidates
  - Pulsar Timing Solutions
  - Dynamic Spectrum
- Transient Buffer Data
- Calibrated Visibilities

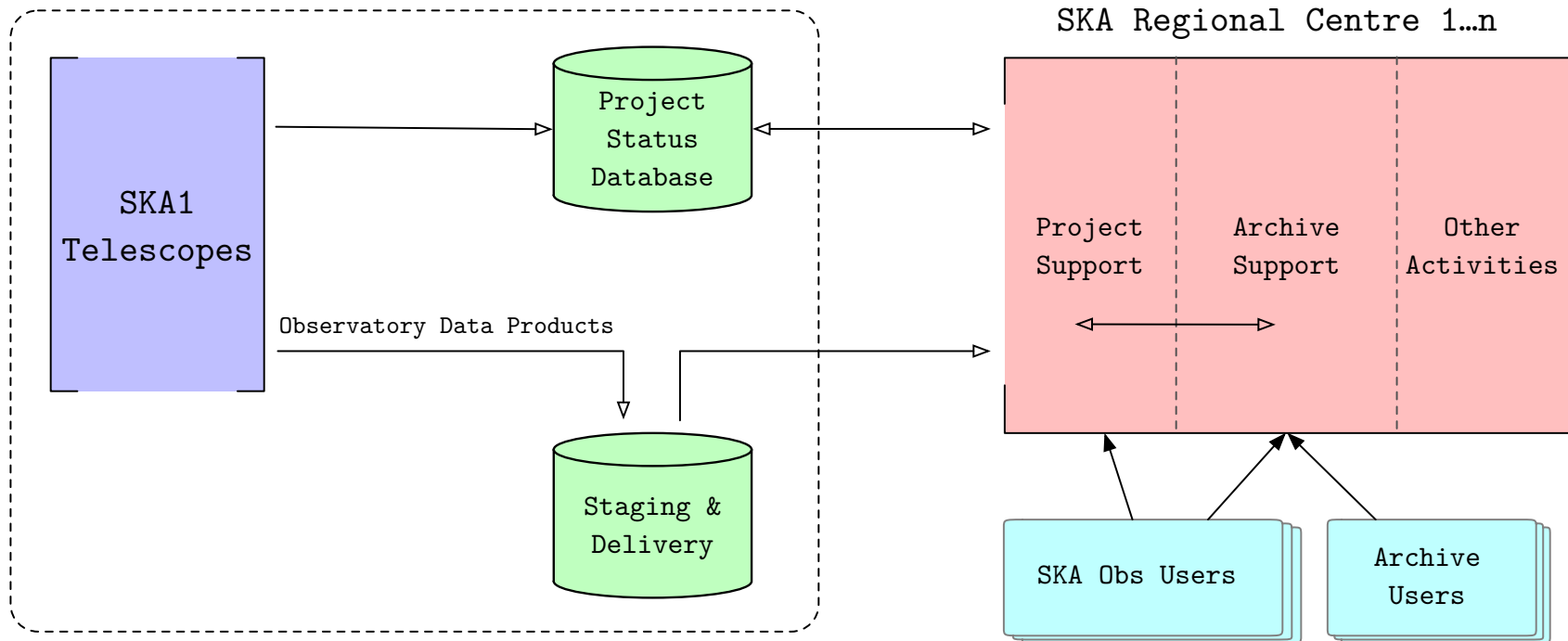


# SKA Regional Centres - 1

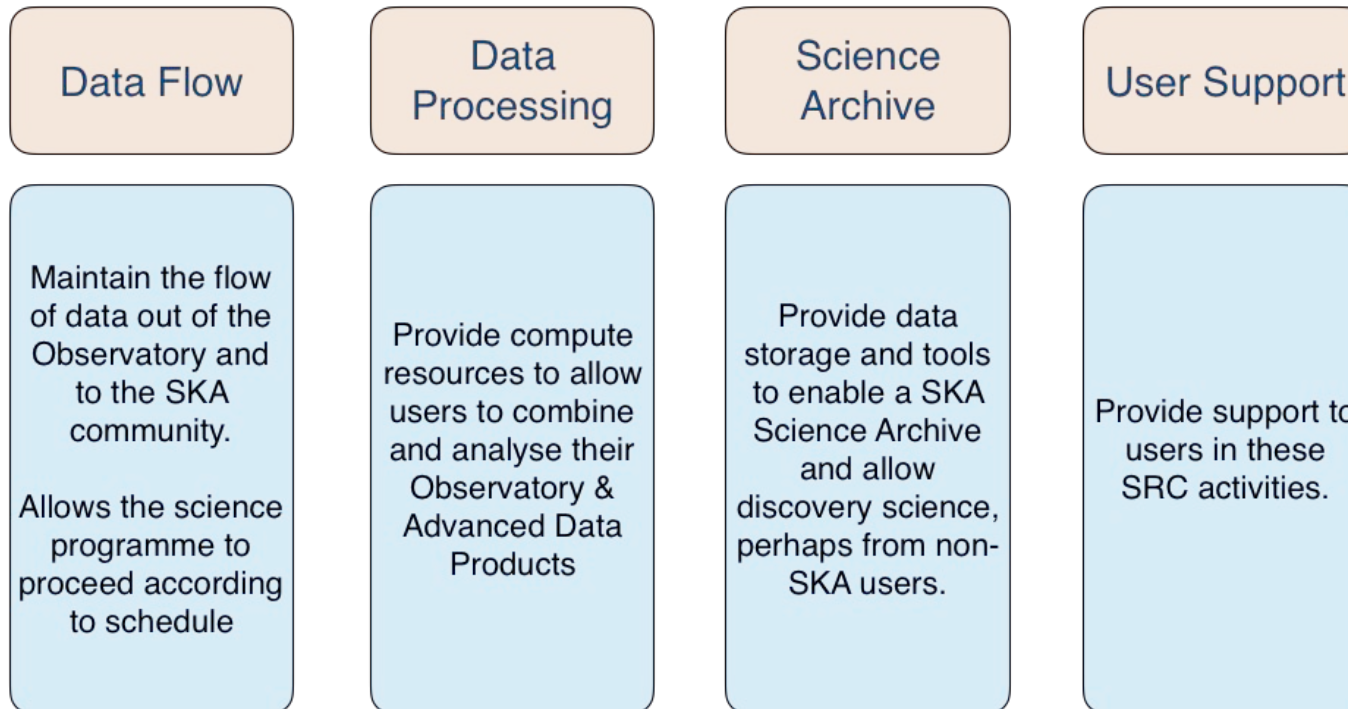
- There are three main factors that lead to a global collaborative model for SKA Regional Centres (SRCs):
  - the Observatory Data Products that emerge from the SKA Observatory are not in the final state required for science analysis and publication,
  - the data volumes are so large that direct delivery to end users is unfeasible,
  - the community of scientists working on SKA science data will be geographically distributed.
- *NOT* funded by the SKA construction budget!

# SKA Regional Centres - 2

SKA Observatory

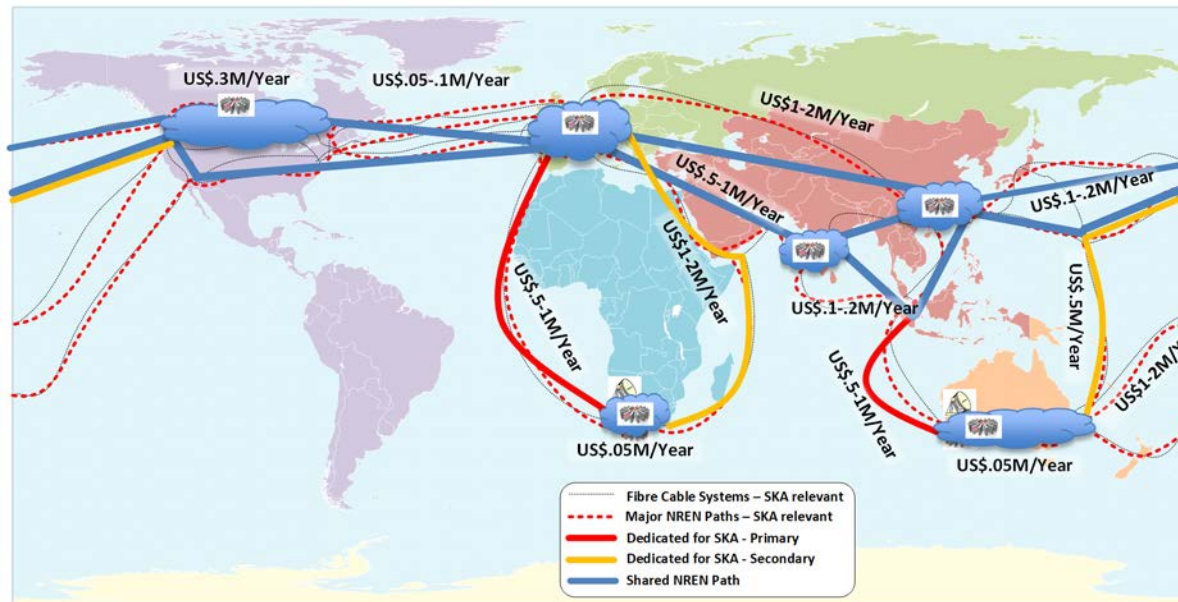


# SKA Regional Centres - 3



# International Network: Dedicated Links and Overlay

- OpEx link and transmission on the academic networks with 10 to 15 year IRUs priced in 2024 USD.
- Assume a dedicated **Primary** 100 Gbit/s link, a dedicated **Backup** 100 Gbit/s link and some use of shared NREN paths.
- I can't emphasize strongly enough the importance of strong partnerships with NRENs (e.g. AARNet and CSIRO).



# Further Thoughts

- Evolution of ethernet and transceivers (> 100Gbit/s)
- Location of equipment:
  - Trade-off between location (power, cooling, space) and long-haul transport.
- (Distributed) Regional Centres:
  - **Deliberately testing single-stream TCP transport**
  - Long-haul data transport (large RTT from Australia to Europe):
    - TCP window size of 1Gbyte
  - Data transfer nodes (e.g. ScienceDMZ is arguably best-in-class) / Security
  - Send multiple copies: “reliable” multicast?
  - Server and NIC performance (100Gbit/s not straightforward)
- FPGA versus GPU for specialized processing.
- Traditional HPC versus Cloud-based services.
- Access to data from elsewhere (surveys, observations at other frequencies)



*We acknowledge the Wajarri Yamatji people as the traditional owners of the Murchison Radio-astronomy Observatory site*

# Thank you

**CSIRO Astronomy & Space Science**

Dr Shaun W Amy (@ausnetman)

t +61 2 9372 4452  
e Shaun.Amy@CSIRO.AU  
w www.atnf.CSIRO.AU

**CSIRO ASTRONOMY AND SPACE SCIENCE**  
[www.csiro.au](http://www.csiro.au)

