SKA Regional Centres SRCSC Update

Nov 2019

Peter Quinn, SRCSC Chair

SRCSC

Quinn, Peter (Chair, Australia), An, Tao (China), Barbosa, Domingos (Portugal), Bolton, Rosie (SKA), Chrysostomou, Antonio (SKA), Conway, John (Sweden), Gaudet, Séverin (Canada), van Haarlem, Michiel (Deputy Chair, Netherlands), Klockner, Hans-Rainer (Germany), Andrea Possenti (Italy), Simon Ratcliffe (South Africa), Scaife, Anna (UK), Lourdes Verdes-Montenegro (Spain), Vilotte, Jean-Pierre (France), Wadadekar, Yogesh (India)

Meetings: 8 May 2019 (F2F), 5 August 2019, 2 October 2019, Shanghai Nov 2019
Regular Update/Reports: SEAC and SKAO Board meetings

Timeline: SRC White Paper for SEAC (Feb 2020) and SKAO Board (March 2020)

DFAP => SRCCG => SRCSC

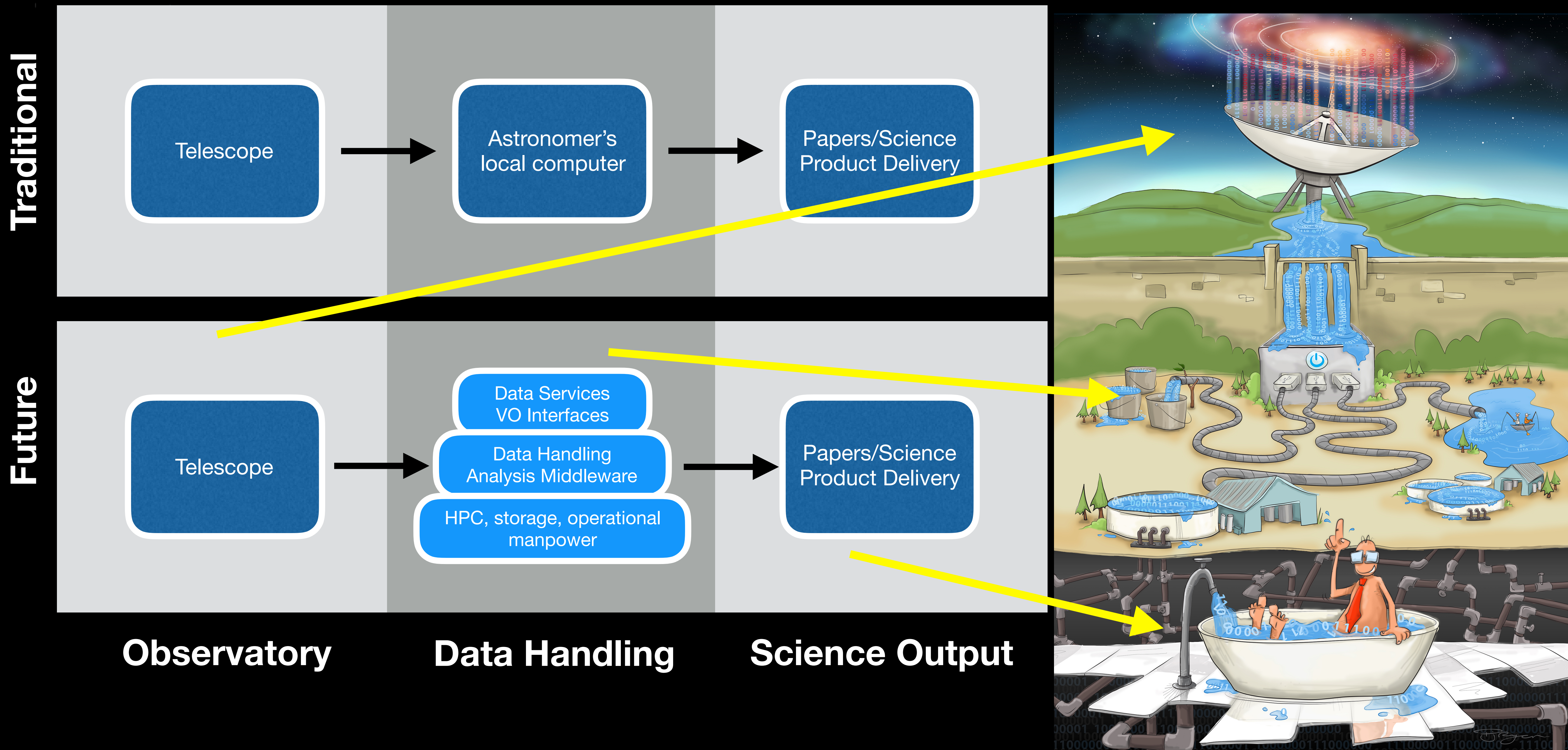
SRCSC Mission:

The mission of the SRCSC is to define and create a long-term operational partnership between the SKA Observatory and an ensemble of independently-resourced SKA Regional Centres.

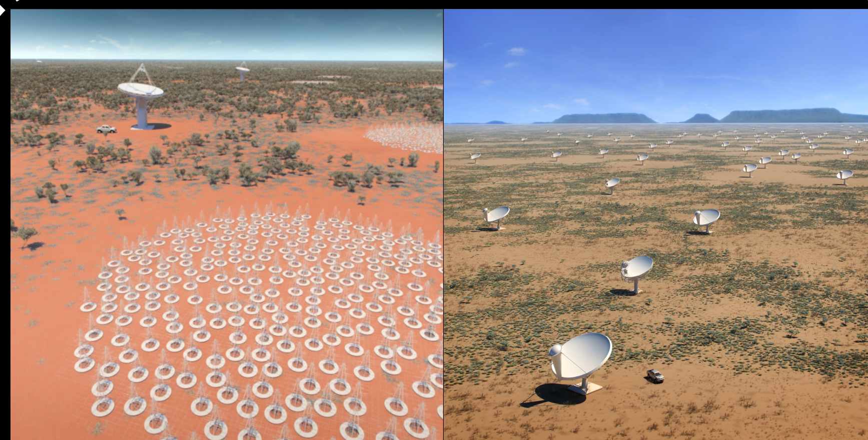
- The SRCSC will be superseded in due course by the operational partnership that is formed as a result of its work.
- The SRCSC will report to the DG and provide a report to the SKAO Board/Council - input and review will be sought from the SEAC
- The SRCSC will function as a round-table partnership consisting of the SKAO and individuals appointed by the SKA partners
- The SRCSC Chair will be elected from the members and serve for 2 years.

The challenge

The challenge



The challenge



SKA 1 Observatory

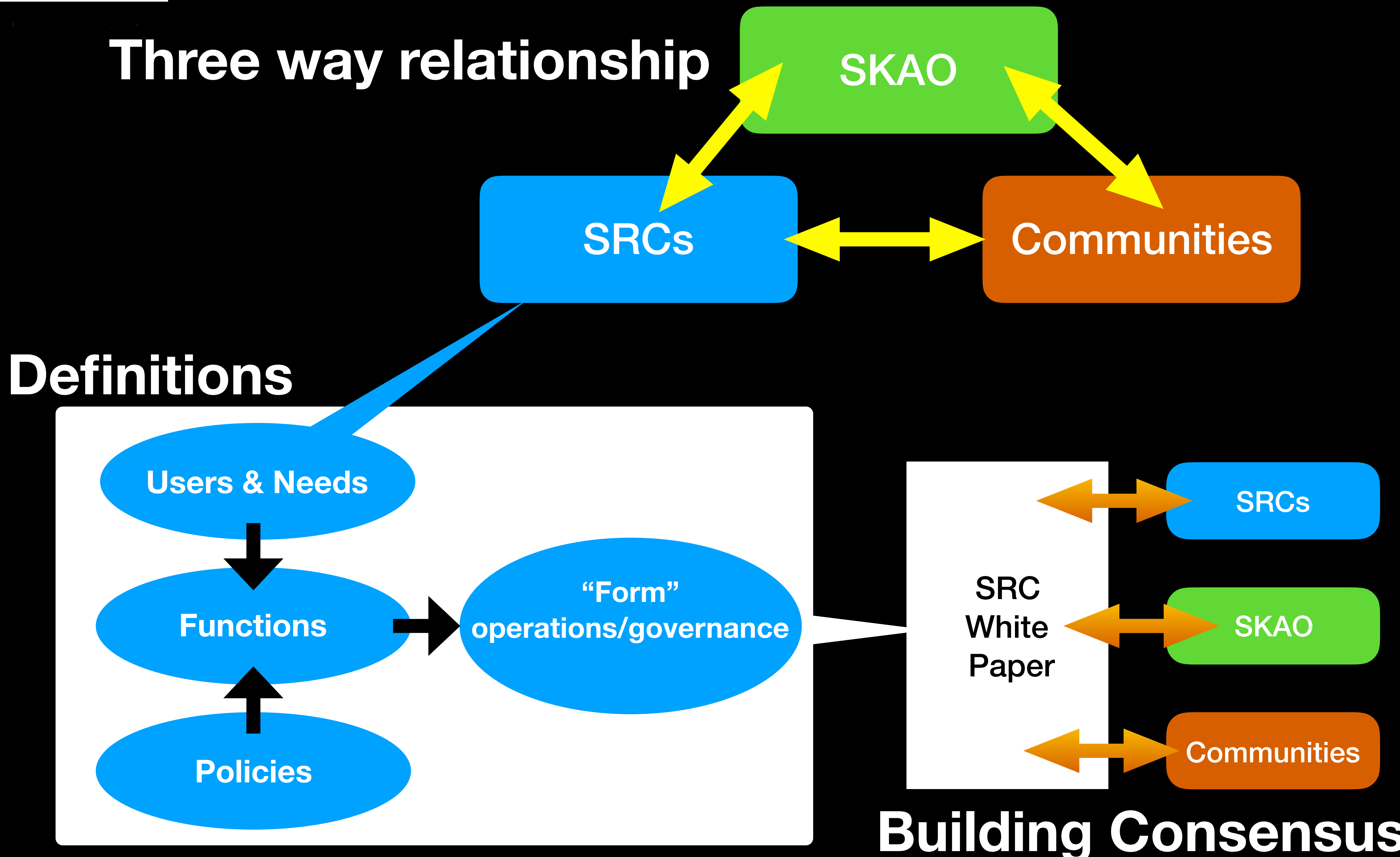
- Online processing and storage through the CSP-SDP chain

Construction and Operations $$

Community $$

- SKA Regional Centres (SRCs) will host the SKA science archive
- Provide access and distribute data products to users
- Provide access to compute and storage resources
- Provide analysis capabilities
- Provide user support
- Multiple regional SRCs, locally resourced and staffed

Finding a common language



White Paper Structure

A. Purpose, audience, assumptions, SRC scope

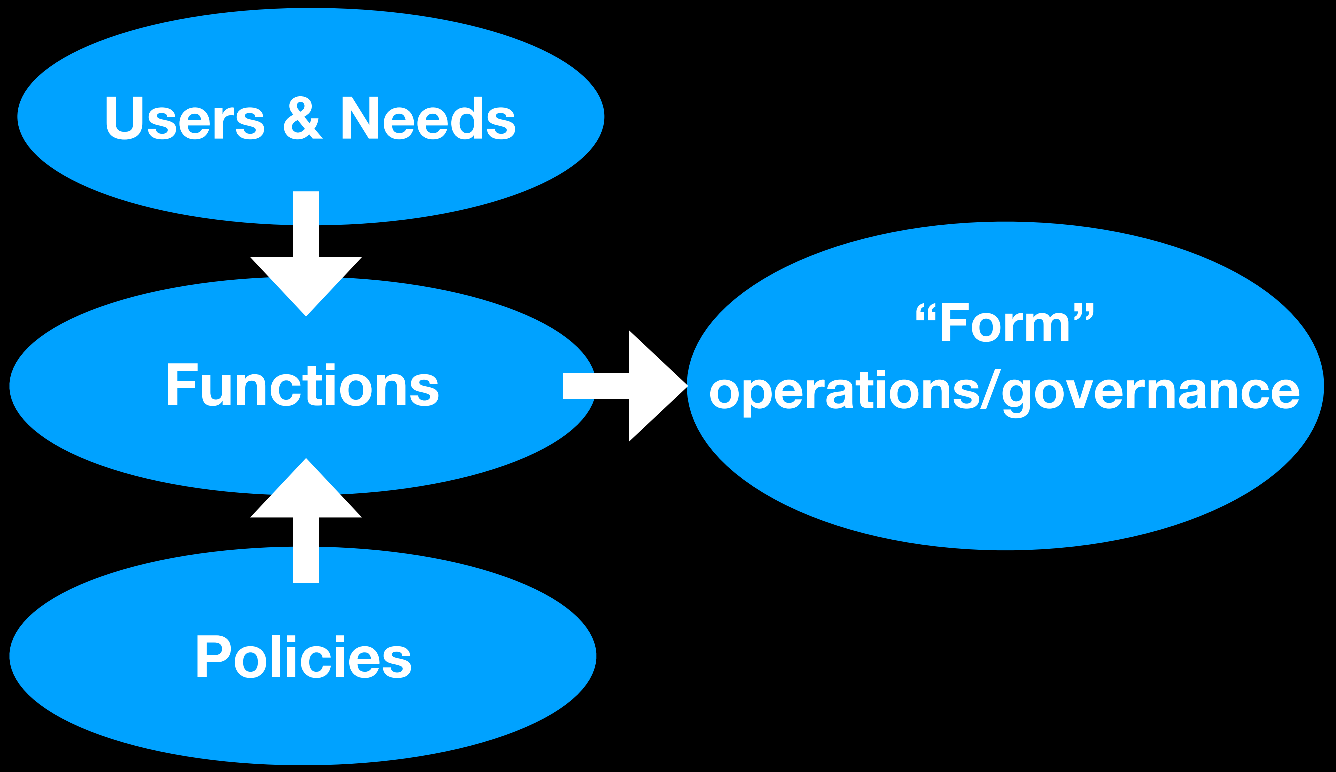
B. Users and their needs

C. Required functionality

D. Operations, interfaces and governance

E. Policies/Principles/Open issues

F. SRCSC Work Plan, timeline, working groups, community engagement

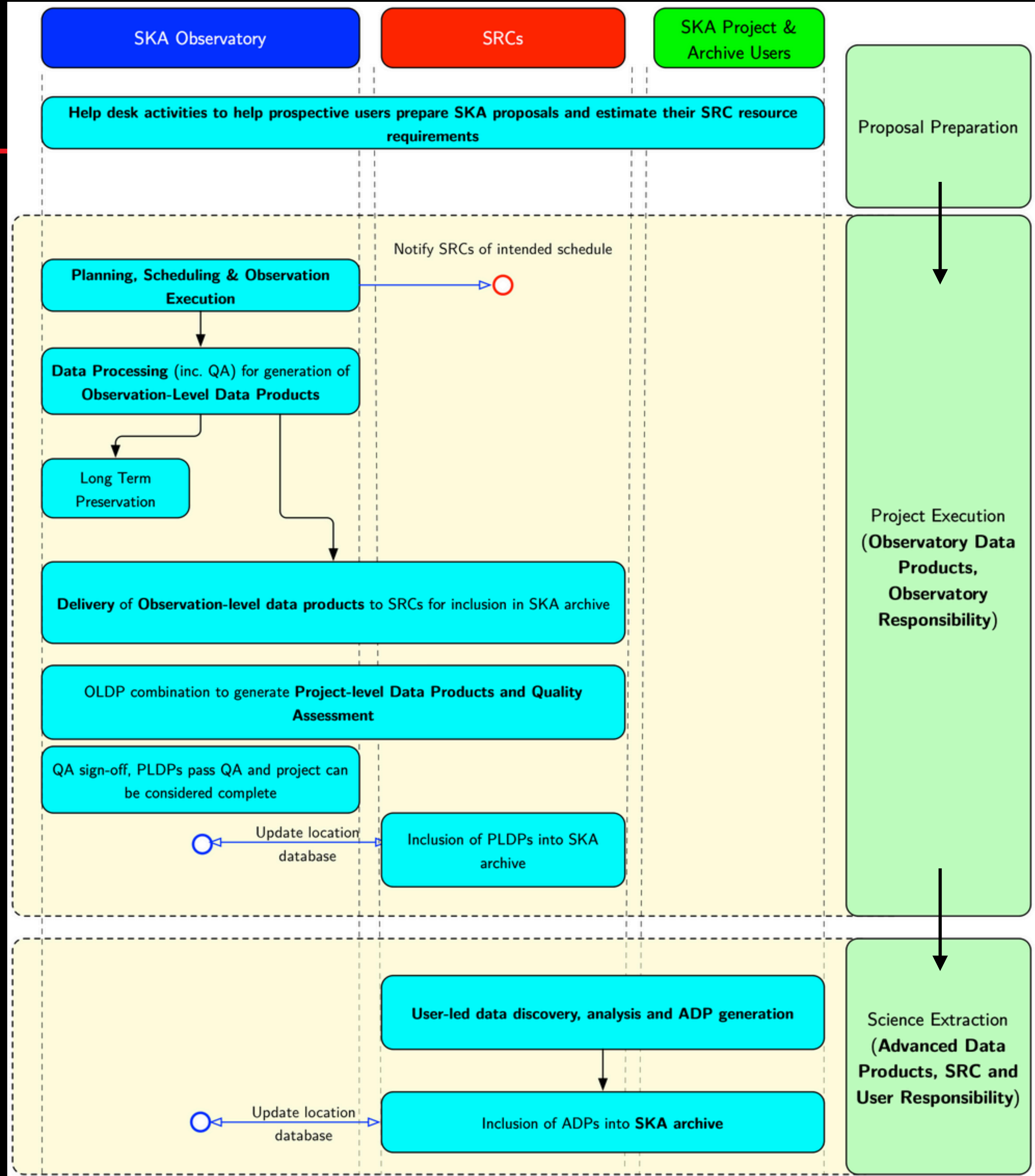


Users

User Group	Description	Set needs?
SKA Programme users	These can be users at the stage of Proposal preparation, or (astronomy) science users either in the Project Execution stage or in the Science extraction stage.	Yes
Non-astronomy users	These are users of SRC resources who are undertaking work which is neither astronomy related nor typically using data products in the SKA science archive.	No
Tools users	These are SRC users who wish to access software tools made developed for or used by the SRC, but for use on external systems	No
Software developers	These are software developers who are interested in making use of the distributed computing tools at SRC or in ensuring that their software can be run effectively in the SRC environment.	Yes
SKA Archive users	These are (astronomy) science users typically accessing public data in the SKA archive	Yes
SKA Observatory	The SKA Observatory (via its Operations staff) will likely be a user of SRC facilities itself	Yes
Commercial users	Commercial users are those who have a commercial relationship with an SRC	No

Users, functions and processes

Programme User Process flow, interfaces and functions



One of these diagrams for each user type

SRC High Level Functions

- **Data Flow** - delivering data from the observatory to those that have been given time to acquire it
- **Data Processing** - the resources necessary to work on the data after it's delivered
- **Data Curation** - providing a performant and persistent science archive that allows discovery and new science
- **User Support** - supporting all users with all of the things above
- **Commonality** - support a common and minimum set of tools/interfaces/systems to enable users to work at SRCs
- **Resource Management** - enable and support an interface to observatory TAC and operations processes to ensure maximal use of distributed SRC resources

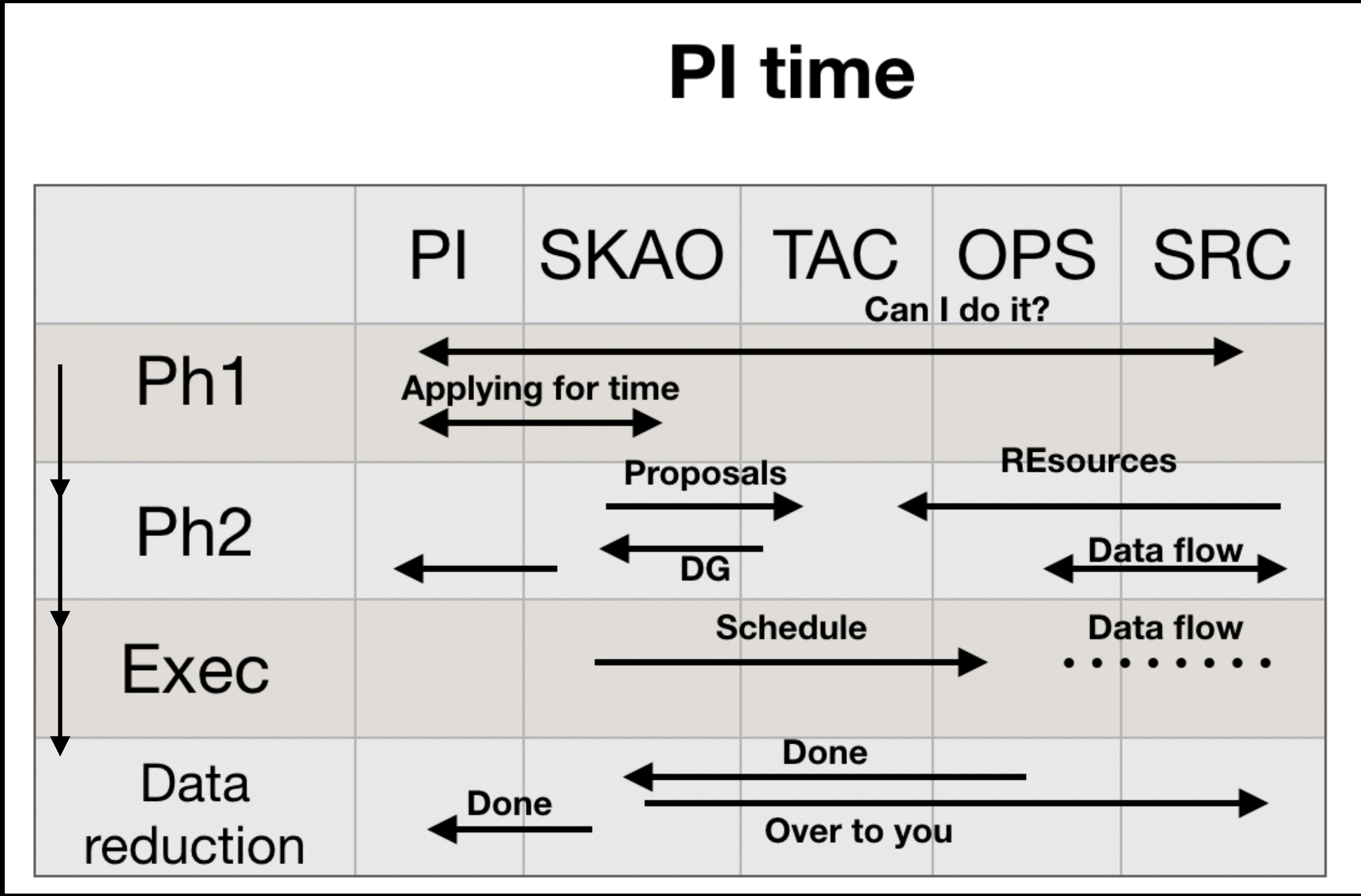
Functional Processes

We need to map the important operations processes and interfaces : e.g. PI time

a) How can SRCs fit into SKA time allocation process?

b) Do KSPs need separate resource model in SRCs c.f. PI / general time?

c) How are SRCs structured to make decisions about allocation of their resources?



One process flow, interface, function diagram for each critical process

Principles and Policies

- Data placement in SRCs will be driven by optimising science, not by politics
- There must be an integrated resource allocation process considering SKAO resources alongside SRC resources
- FAIR principles for data and methods
- SKAO/SRC/NREN/ Cyber infrastructure network must be via co-design
- SRCs will evolve with engagement with precursors/pathfinders
- Contributions to the SRC collective pool will be codified in terms of capability and not only in terms of cost
- There does need to be an SRC body/entity
- Data management will be undertaken by the SRC body
- There will be a minimum set of SW enabling a common SW platform
- IVOA: access through IVOA services - data in SRCs will be IVOA compliant
- ……

“Form” : Ops/Gov

There needs to be an SRC “entity” that enables the necessary decision making and coordination to support an operational SRC network that interfaces efficiently and effectively with the Observatory and community.

- Individual SRC resources and services must be provided in a globally uniform and consistent manner to the SKA community
- Individual SRC must appear to be globally uniform and consist to SKA-1 operations

Form of an SRC Entity

An SRC Entity

- Provides the coordination between national SRCs and the interaction between the SRCs collectively and the SKAO.
- Needs to be persistent
- Cannot own or directly control the funding and resources of any individual SRC that are owned and controlled at the national level by diverse sets of funders and stakeholders
- Needs to ensure consistency and uniformity of service to SKAO and the community. The Entity must be responsible to the SKAO and the community for this operational consistency and uniformity

Options

There are a number of options and models to examine for the Form of an SRC Entity

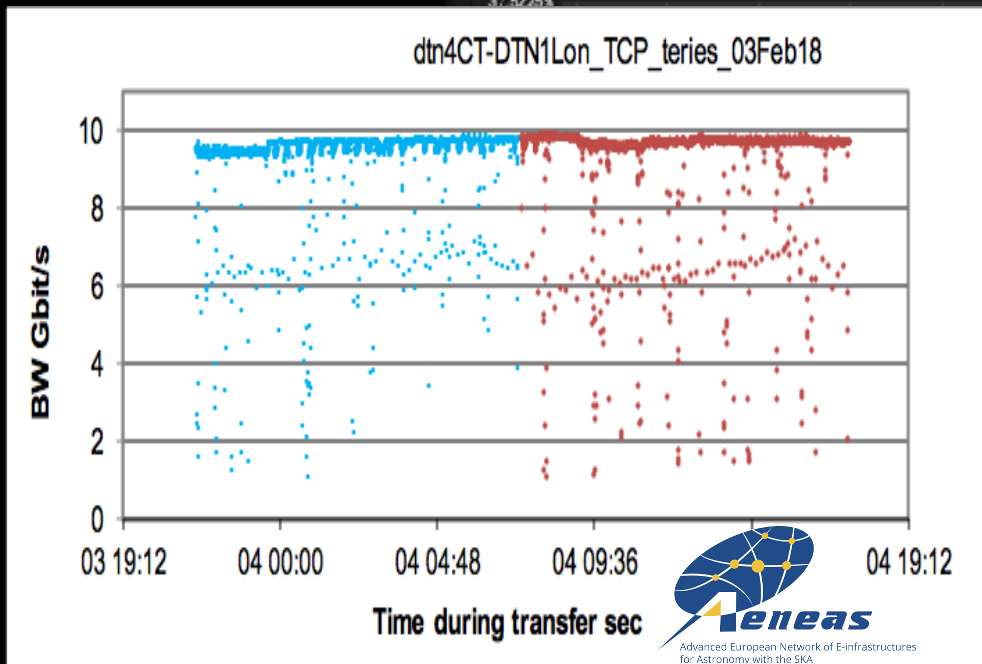
- The SKAO itself
- A new structure connected to the SKAO Council
- A new stand-alone structure (Alliance, Joint Venture, other…)
- An existing structure

The White Paper will examine the +/- for each of the options - persistence and ownership of resources will be major factors.

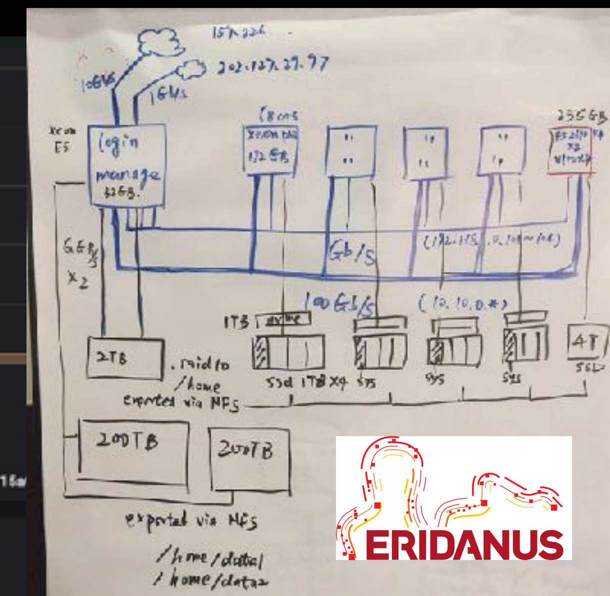
SRCSC Working Groups

- Users (Long term group)
- Operations (SRC + SKAO)
- Software, including services
- Global networking and Data logistics
- Archive, including IVOA
- Prototyping and SRC data challenges
- FAIR

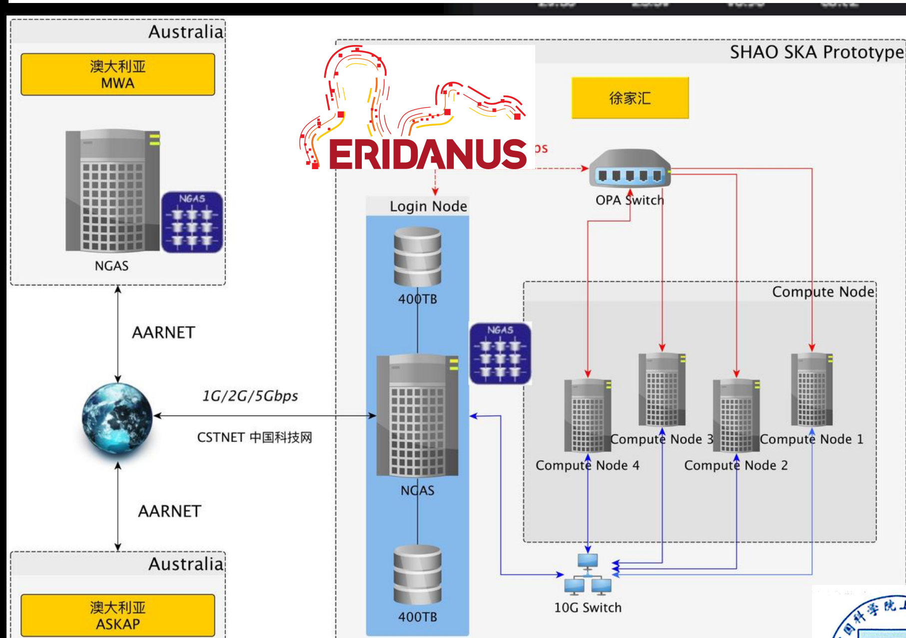
Results from SRC early work



SANReN Cape Town to GÉANT London

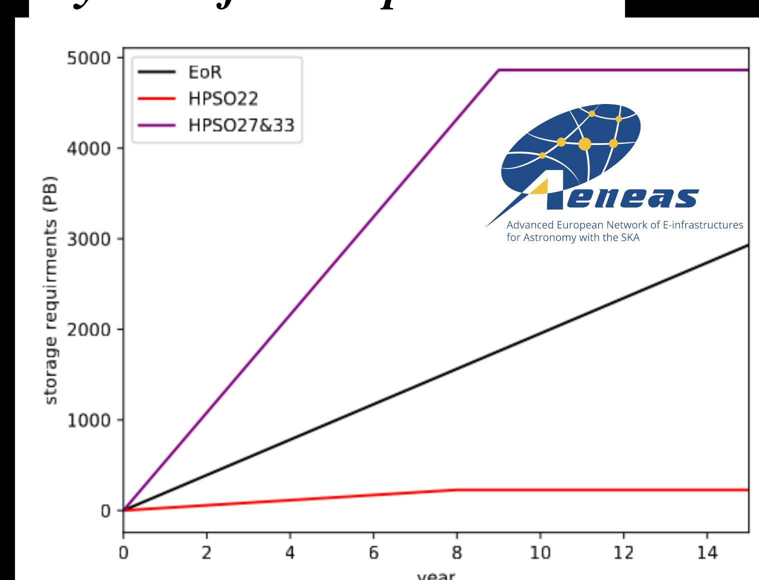


- 3 Gbps data transfer Perth (Pawsey) to Shanghai (SHAO)
- Automatic MWA and ASKAP data transfer (using NGAS)
- Automatic pipeline deployment & execution on data arrival



NGAS archives in Shanghai and Perth

~9 ExaBytes over first 15 years of SKA operations



Processing in IDIA cloud

Global SRC Effort

SRC	FTE	M Euro	Timeline
Aus	7	2.5	2019-22
Canada	9	7.5	2019-24
China	10	6	2019-22
France	awaiting CP		
Germany	5		2020
India		5.5	2019-23
Italy	10	4	2020-22
Netherlands	10		2020
Portugal	3		2020
South Africa	2		2019-23
Spain	5		2019-22
Sweden	3	4	2020-26
UK	3		2019-21
SKAO	2		2019-22
	53 (69)	25.5 (29.5)	

SUMMIT supercomputer

Summit Overview

OpenPOWER

Components

IBM POWER9
- 22 Cores
- 4 Threads/core
- NVLink

NVIDIA GV100
- 7 TF
- 16 GB @ 0.9 TB/s
- NVLink

Compute Node

2 x POWER9
6 x NVIDIA GV100
NVMe-compatible PCIe 1600 GB SSD

25 GB/s EDR IB- (2 ports)
512 GB DRAM- (DDR4)
96 GB HBM- (3D Stacked)
Coherent Shared Memory

Compute Rack

18 Compute Servers
Warm water (70°F direct-cooled components)
RDHX for air-cooled components

39.7 TB Memory/rack
55 KW max power/rack

Compute System

10.2 PB Total Memory
256 compute racks
4,608 compute nodes
Mellanox EDR IB fabric
200 PFLOPS
~13 MW

GPFS File System

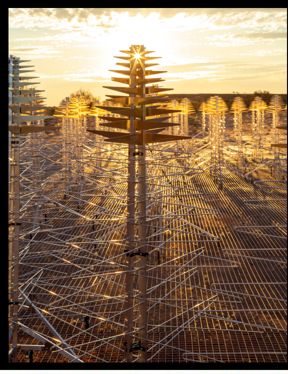
250 PB storage
2.5 TB/s read, 2.5 TB/s write

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

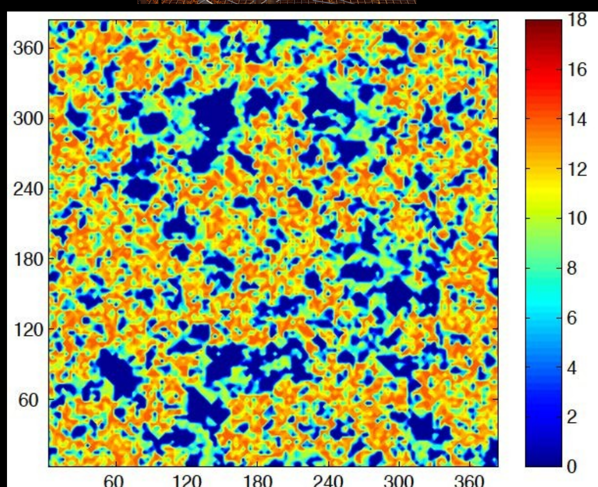
~220 million Euro
~24 million Euro/yr power

~200,000 cores
~200 PFLOPs

SKA1-low simulation



Telescope Model



EoR Sky Model

OSKAR2
Correlator
simulation

SPEAD2

7.3 billion vis/sec
400 GB/s

Ingest Pipeline

Averaging

3 hr run time = 6 hrs SKA1-low

Measurement
Sets

106 m vis/sec

disk
150 TB
10 GB/s

Imaging Pipeline

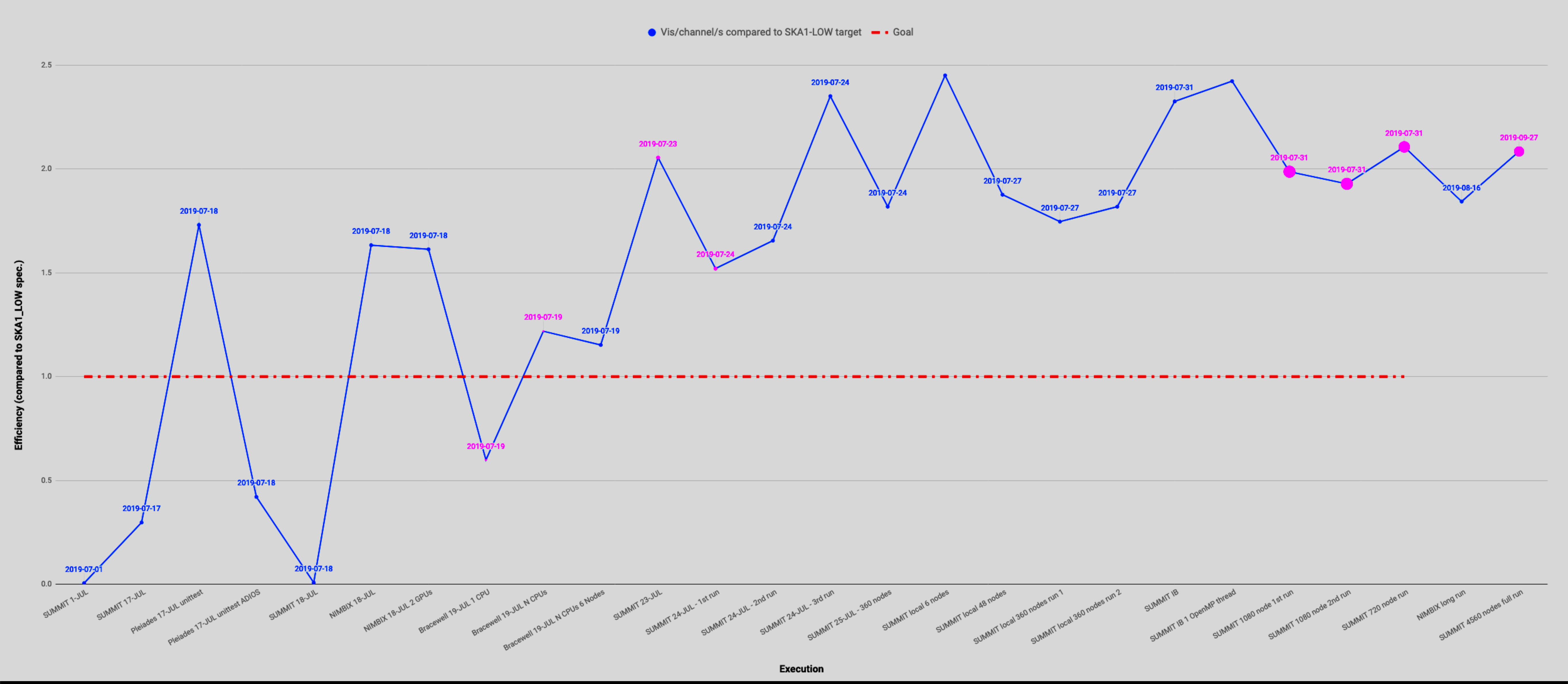
ASKAPSOFT
DALiuGE

Image Cube

SKA1-scale Processing Workflow on World's fastest Supercomputer SUMMIT

Credit: ICRAR ,ORNL ,SHAO

SUMMIT performance



Next Steps

- **2019 Q4 Shanghai Meeting**
- **2019 Q4 First draft of White Paper**
- **2020 Q1 Present White Paper to SEAC/Board**
- **2020 Q1 Initiate working groups**
- **2020 Q1-4**
  - Support national business cases
  - National SRC meetings
  - Progress SRC network establishment (Entity ops/gov)
  - Coordinate national developments
- **2021 Q1-4**
  - Prototype SRC network (precursor focus, SRC data challenges)
  - Workshops/conferences
- **2022 - 2025**
  - Proto-SRC precursor science
  - Alignment with SKA commissioning needs

SRCSC → **SRC Entity**