SKAO REGIONAL CENTRE NETWORK

SKAO and the SKA Regional Centres Network

Jesús Salgado - SRC Architect
Rosie Bolton - Head of Data Operations
SRC Product Manager

IVOA Interop Oct 2022 - Radio IG session

SKA1-low – the SKA's low-frequency instrument

The Square Kilometre Array (SKA) is a next-generation radio astronomy facility that will revolutionise our understanding of the Universe. It will have a uniquely distributed character: **one** observatory operating **two** telescopes on **three** continents. Construction of the SKA will be phased and work is currently focused on the first phase named SKA1, corresponding to a fraction of the full SKA. SKA1 will include two instruments – SKA1-mid and SKA1-low – observing the Universe at different frequencies.

SKA SQUARE KILOMETRE ARRAY

Location: Australia

Frequency range: **50 MHz** to **350 MHz**

~131,000 antennas spread between **500 stations**

Total collecting area: **0.4km²**

Maximum distance between stations: **>65km**

SKA1-low

Data transfer rate: **7.2 Terabits** per second

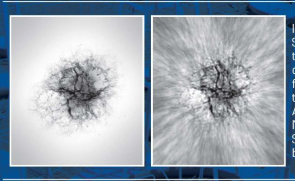


Image quality of SKA1-low (left) versus the best current facility operating in the same frequency range, the LOw Frequency ARray (LOFAR), in the Netherlands (right). SKA-low's resolution will be similar to LOFAR.

Compared to LOFAR Netherlands, the current best similar instrument in the world

25% better resolution

8x more sensitive

135x the survey speed

www.skatelescope.org @SKA_telescope SKAtelescope ska_telescope Square Kilometre Array ska-organisation

SKA1-mid – the SKA's mid-frequency instrument

The Square Kilometre Array (SKA) is a next-generation radio astronomy facility that will revolutionise our understanding of the Universe. It will have a uniquely distributed character: **one** observatory operating **two** telescopes on **three** continents. Construction of the SKA will be phased and work is currently focused on the first phase named SKA1, corresponding to a fraction of the full SKA. SKA1 will include two instruments – SKA1-mid and SKA1-low – observing the Universe at different frequencies.

SKA SQUARE KILOMETRE ARRAY

Location: South Africa

Frequency range: **350 MHz** to **15.3 GHz** with a goal of 24 GHz

197 dishes (including 64 MeerKAT dishes)

Total collecting area: **33,000m²**

or **126 tennis courts**

Maximum distance between dishes: **150km**

SKA1-mid

Data transfer rate: **8.8 Terabits** per second

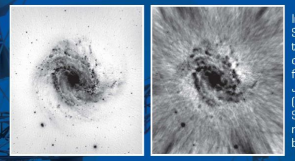


Image quality of SKA1-mid (left) versus the best current facility operating in the same frequency range, the Jansky Very Large Array (JVLA) in the United States (right). SKA-mid's resolution will be 4x better than JVLA.

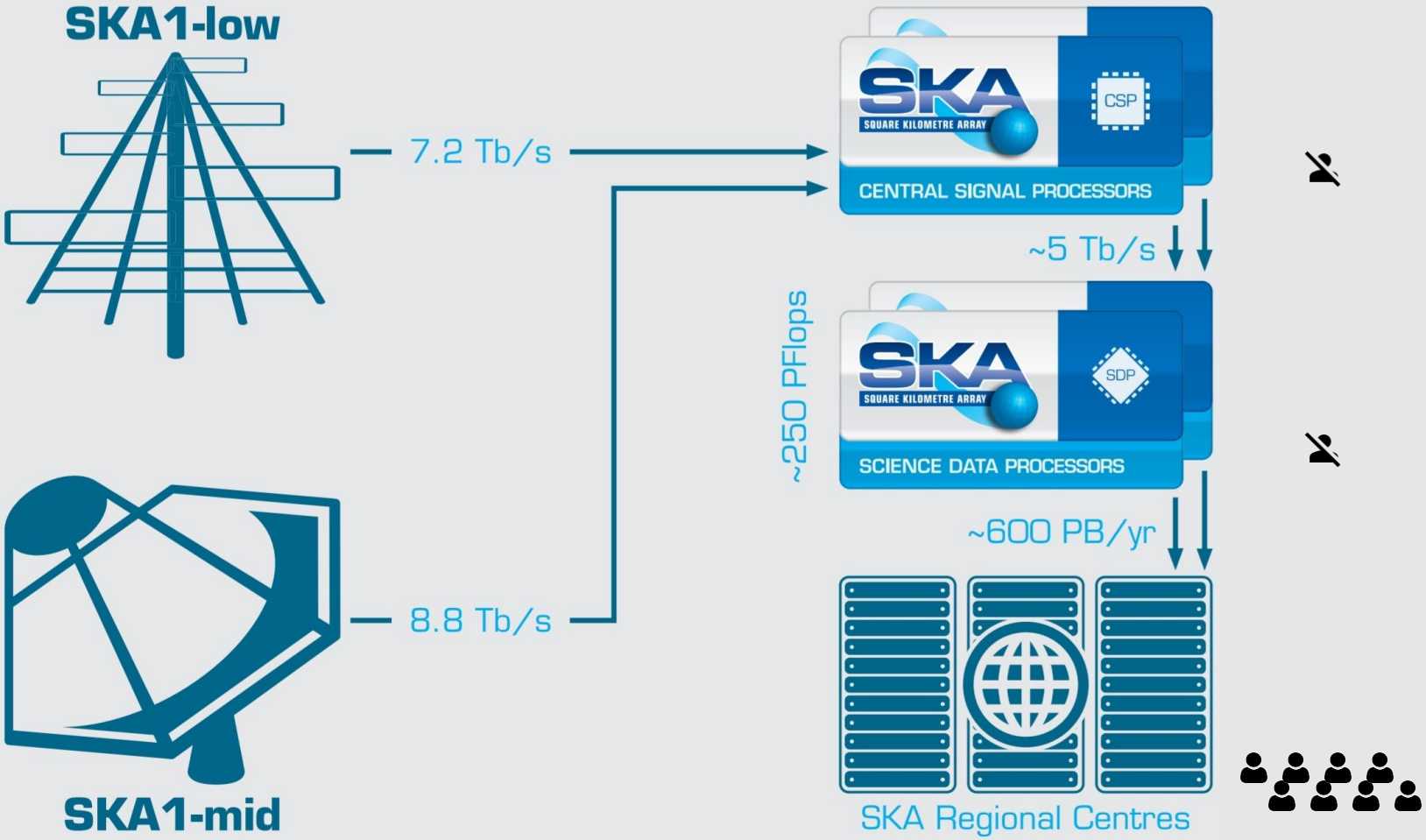
Compared to the JVLA, the current best similar instrument in the world:

4x the resolution

5x more sensitive

60x the survey speed

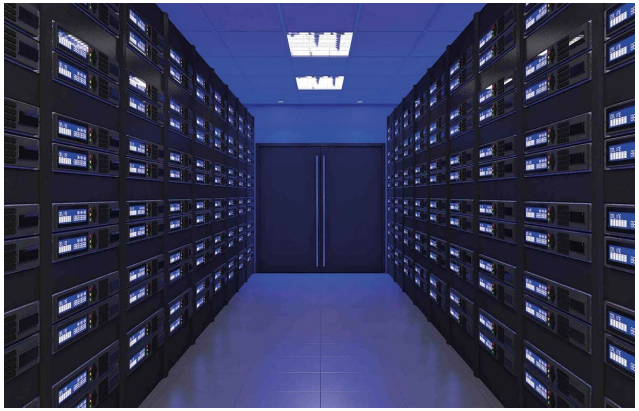
www.skatelescope.org @SKA_telescope SKAtelescope ska_telescope Square Kilometre Array ska-organisation

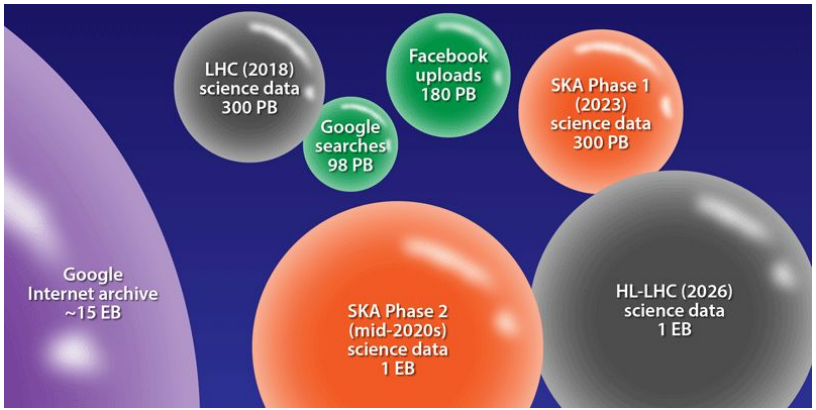

SKA1-low
7.2 Tb/s
SKA1-mid
8.8 Tb/s
SKA
SQUARE KILOMETRE ARRAY
CSP
CENTRAL SIGNAL PROCESSORS
~5 Tb/s
~250 PFlops
SKA
SQUARE KILOMETRE ARRAY
SDP
SCIENCE DATA PROCESSORS
~600 PB/yr
SKA Regional Centres

SKA Regional Centres **(SRCs)**

SKA Regional Centres (SRC) Net in numbers

- Around 600 PB/year of Scientific Data
- Up to 15 countries involved
- 40 FTEs during prototyping phase
- Up to 100 FTEs during development phase
- More than 6 main data locations
- HPC availability for users
- Collaboration agreements with CERN, CNRS, Vera Rubin and others

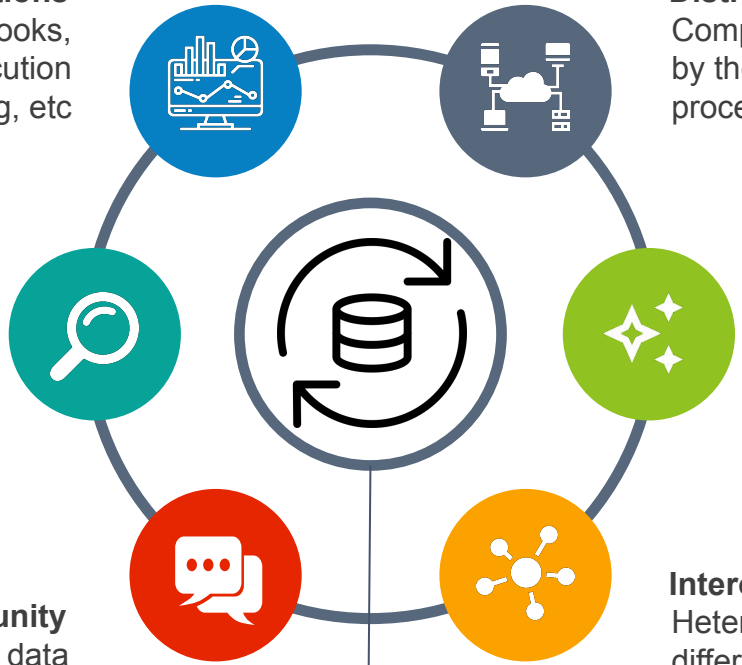




SKA Regional Centre Capabilities Blueprint

Science Enabling Applications
Analysis Tools, Notebooks, Workflows execution Machine Learning, etc

Distributed Data Processing
Computing capabilities provided by the SRCNet to allow data processing



Data Discovery
Discovery of SKA data from the SRCNet, local or remote, transparently to the user

Visualization
Advanced visualizers for SKA data and data from other observatories

Support to Science Community
Support community on SKA data use, SRC services use, Training, Project Impact Dissemination

Interoperability
Heterogeneous SKA data from different SRCs and other observatories

Data Management
Dissemination of Data to SRCs and Distributed Data Storage

See Yan Grange presentation in GWS session

SRC Network global capabilities

Every node is an instance of the blueprint

Interconnections are done using agreed APIs, using FAIR and VO protocols where available

Collectively meet the needs of the global community of SKA users

Anticipate heterogeneous SRCs, with different strengths

Data Intensity vs. User Flexibility

"Size" or Data intensity

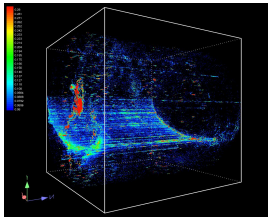


Image cubes

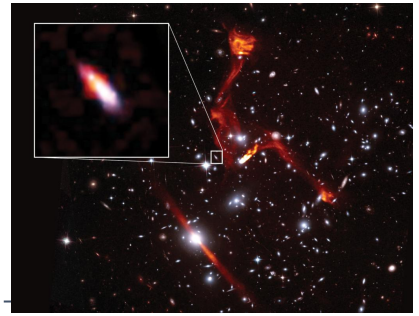
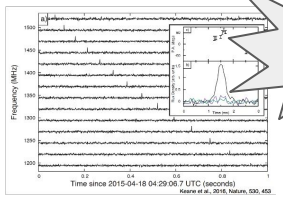
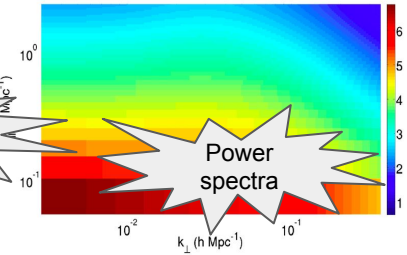


Image cut-outs



Time series data

Catalogues / Source List

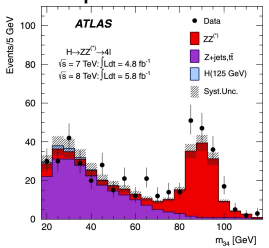


Power spectra

Credit: Heywood et al.; Sophia Dagnello, NRAO/AUI/NSF; STScI.

Paul, Sourabh et al. (2016). ApJ 833. 10.3847/1538-4357/833/2/213.

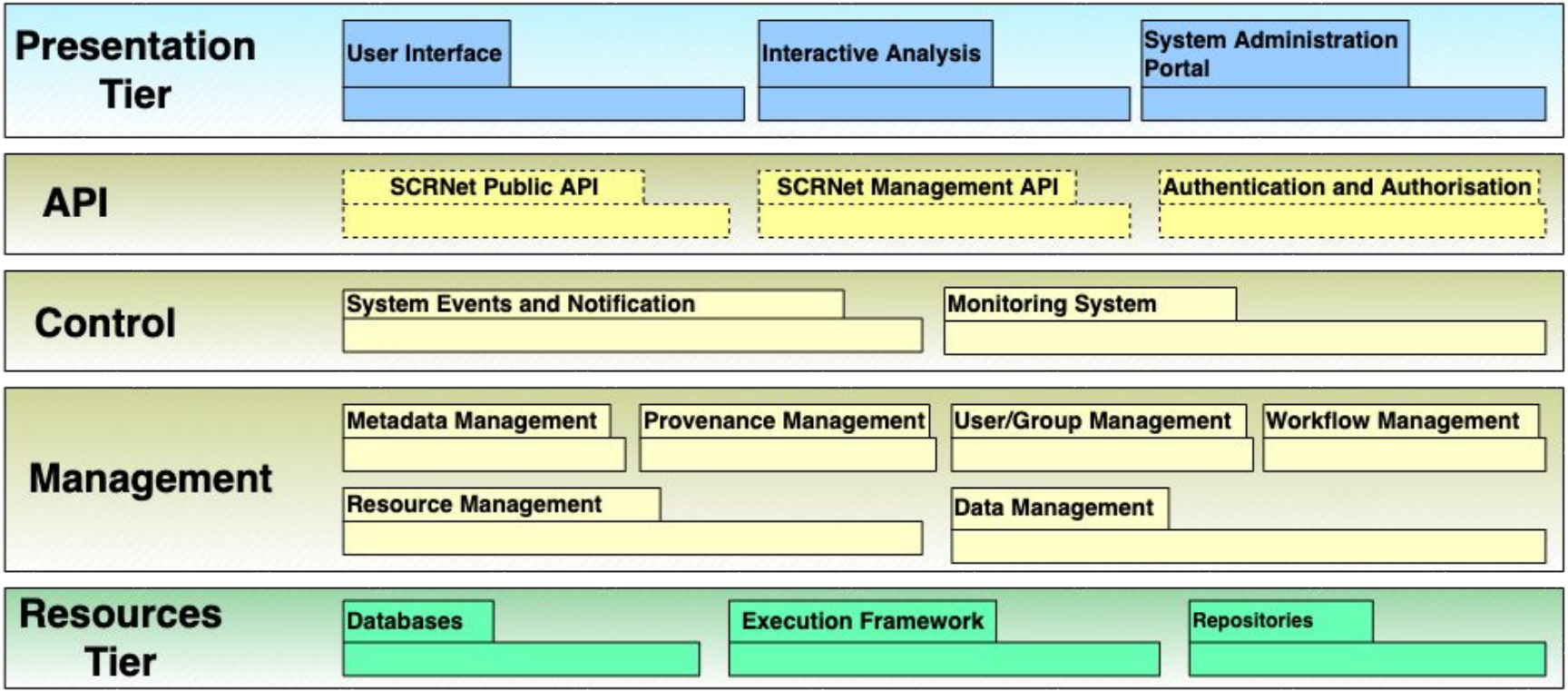
Plots and scripts for publication



"Pre-defined" (Observatory Data Products)

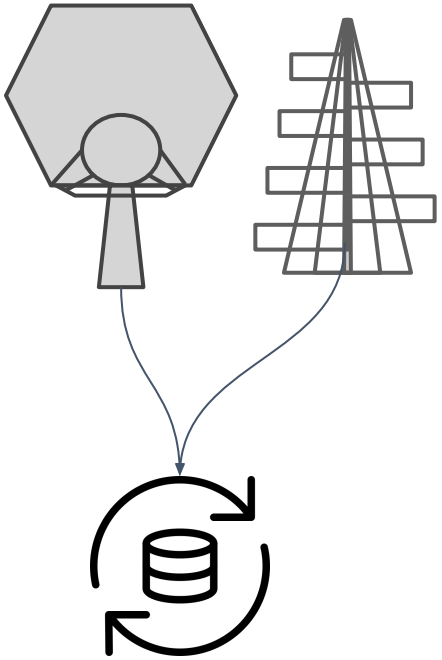
"User-defined" (Advanced Data Products)

SRC Node Architecture Simplified View



SRCNet principles: Data Management

- Storing SKAO data growing at up to 600 PBytes each year will be a challenge
  - (plus user-generated data toox)
- Roughly 5-10 million dollars per year in new data, for one copy
- Global data management within SRCNet should enable best possible use to be made of available storage resources
- Avoid unnecessary duplication and transfers
- Support mirroring of popular data products to enhance user experience



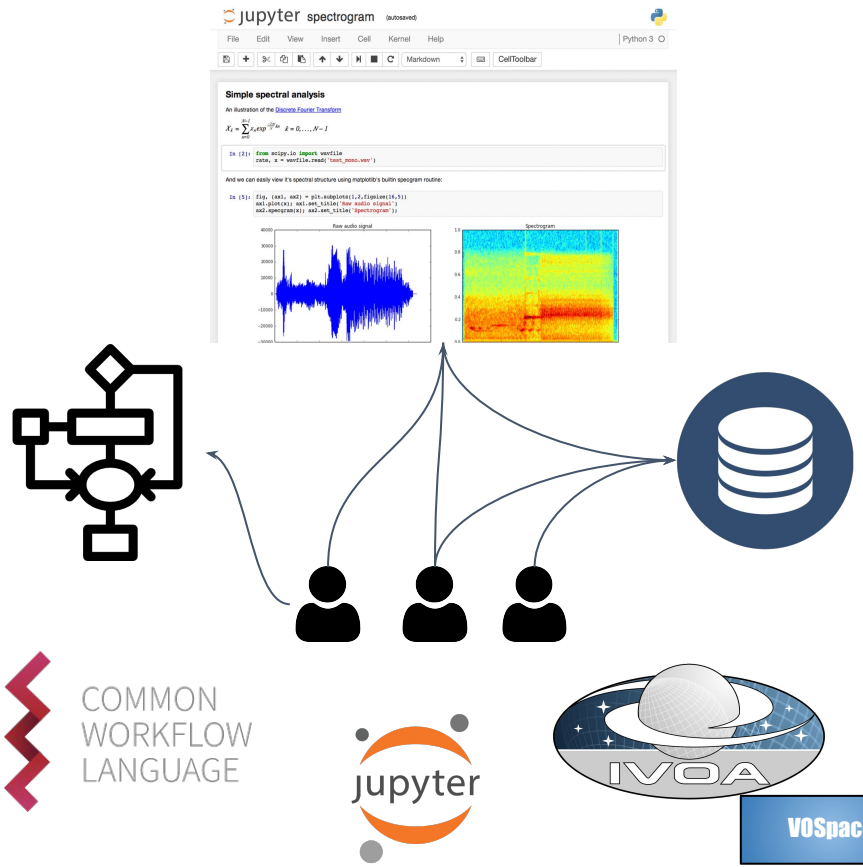
- Exploration of data managements systems able to handle Exabytes

SCIENTIFIC DATA MANAGEMENT
RUCIO

SRCNet principles: Collaboration and Reproducibility

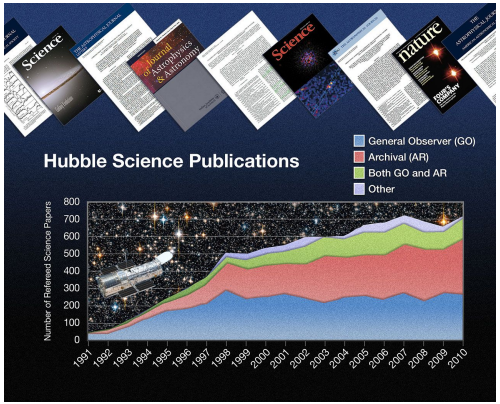
- Most SKA projects will be collaborative
- SRCs will provide collaborative tools
  - Sharing components
  - Single Sign-on
  - Authorisation System

- Support to workflows
- Provenance metadata
- Science Reproducibility at the level of workflows is essential as data should not be downloaded



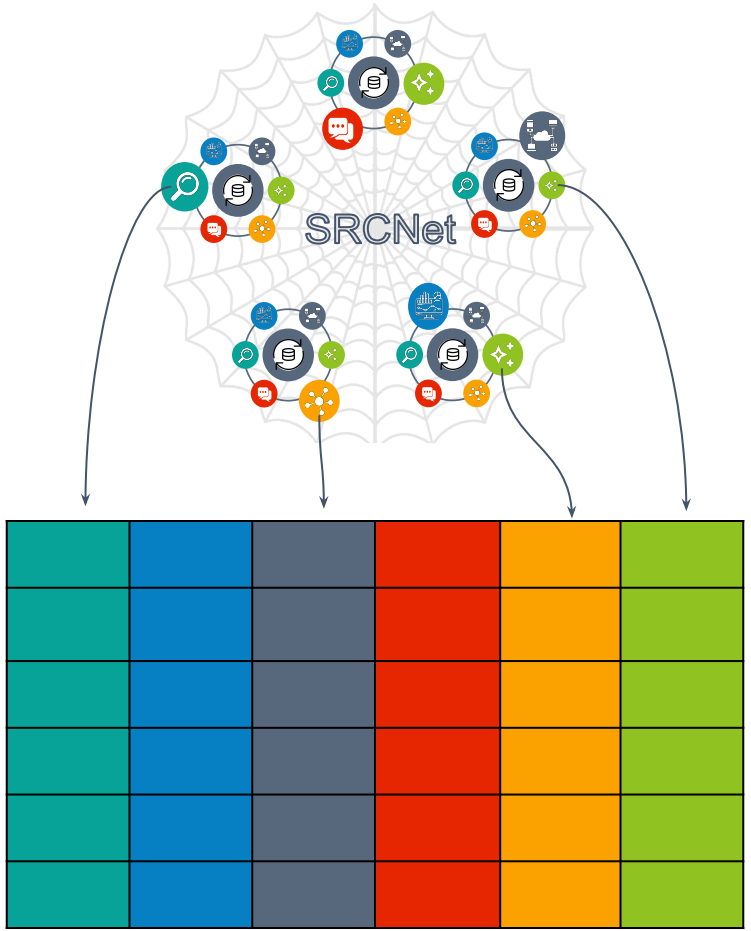
SRCNet principles: Use of Standards

- Build SKA science archive around FAIR and IVOA standards
- Ensure interoperability with other archives and other experiments
- Strong adherence to the FAIR principles
- Give credit appropriately to all contributors to a team

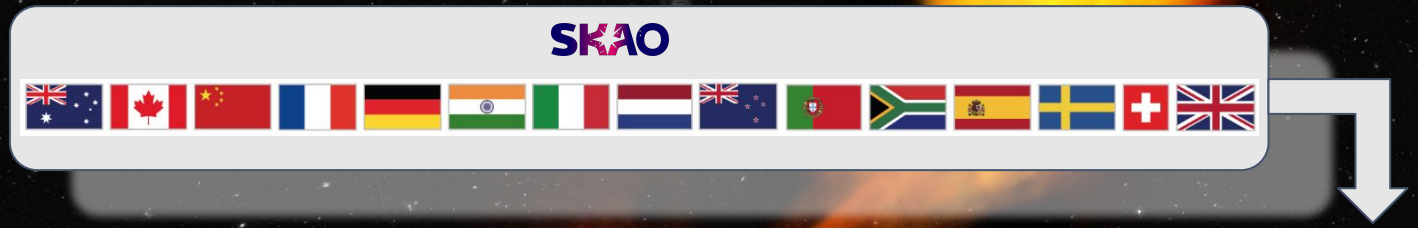


SRCNet principles: Pledging

- Each SRC to pledge resources into global pool to support SRCNet activities
- Each SRC should be able to contribute a total effort that is proportional to their SKA fraction
- Additional resources at an SRC could be given to the pool or prioritised to support national interests



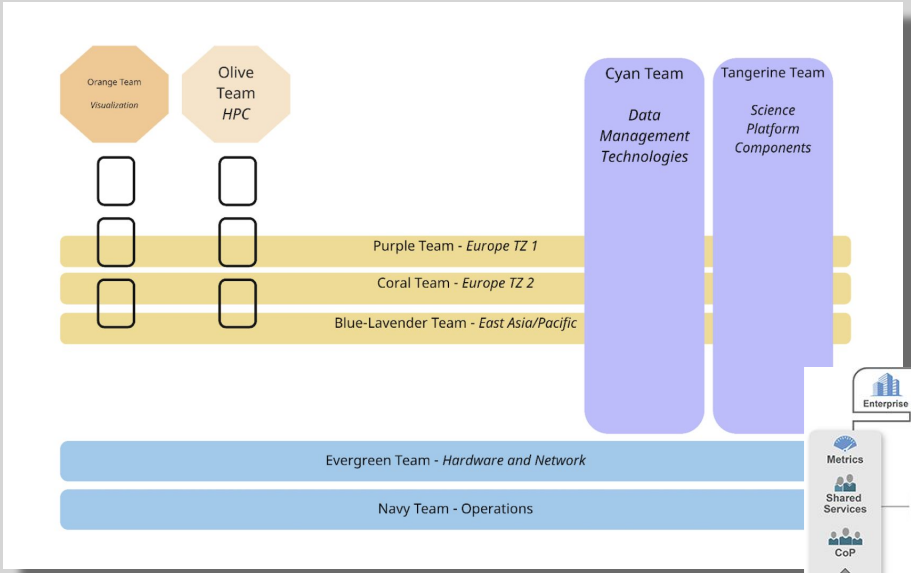
Multinational Development Teams

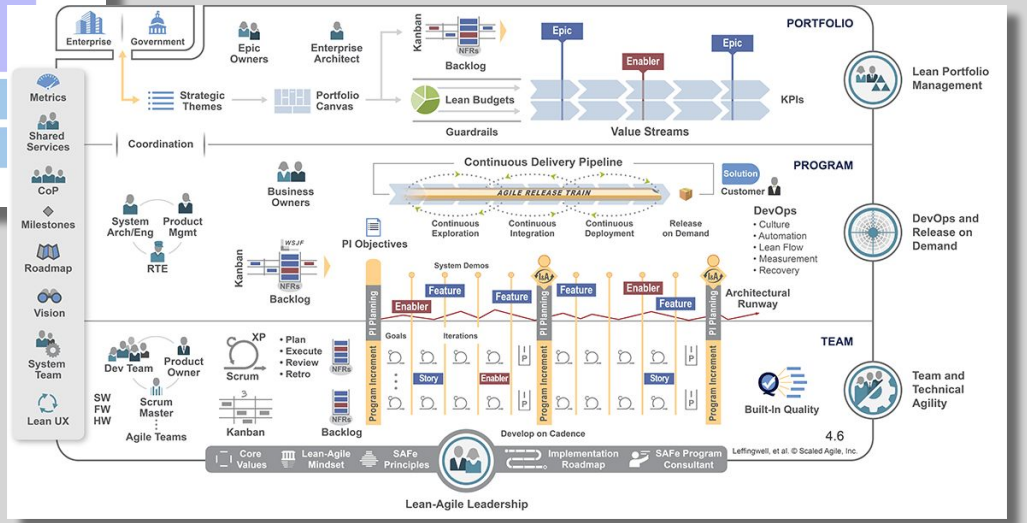


15 Countries involved in the first development phase of the SRC Net

visualization high
cloud services
performance computing
science platform components
components authentication authorization
deploy entity
platform components authentication
new src node
protosrc
management technology science
hardware network operations
data management technology
technology science platform
authentication authorization

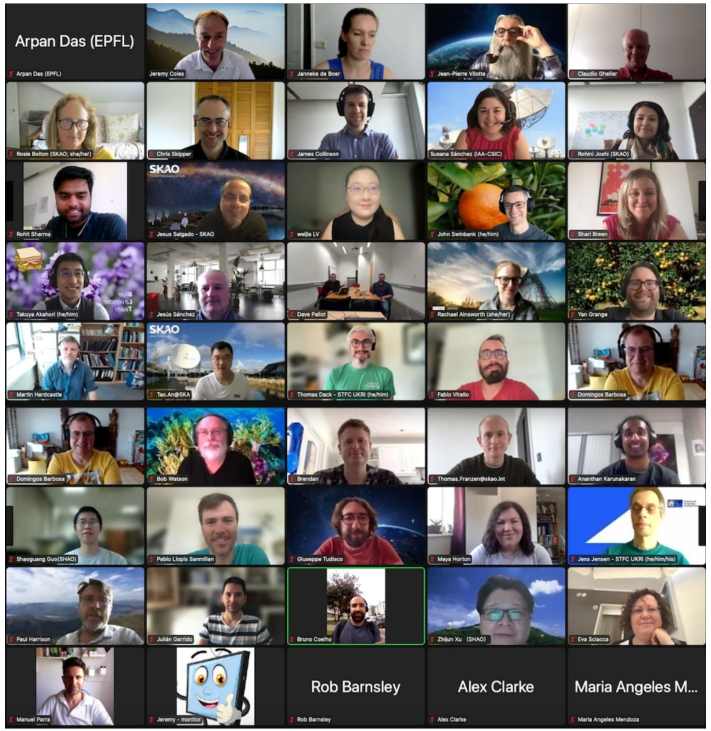
Prototyping Teams





SRC Prototyping

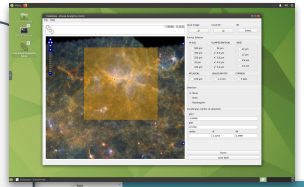
- Start of the Prototyping phase with a team of teams (50 developers, 20 observers) on the SRC ART
- Program Team: Rosie Bolton, Jeremy Coles, Jesus Salgado
- Science user engagement WG driving an improved understanding of the use cases

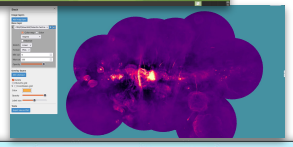


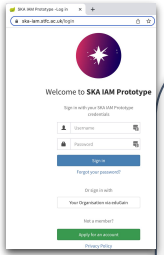
Prototype efforts on different modules

Exploration of Visualisation Tools for SKA data Demos and integration studies

Orange

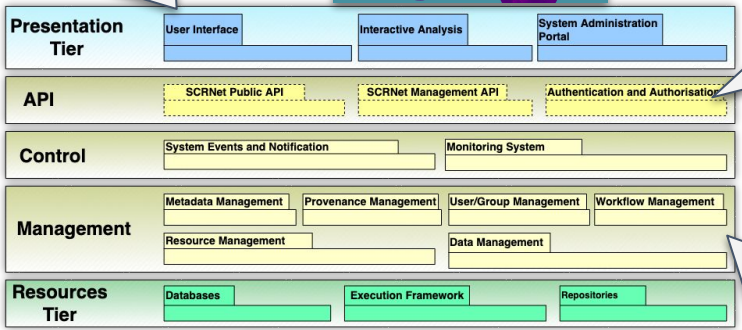






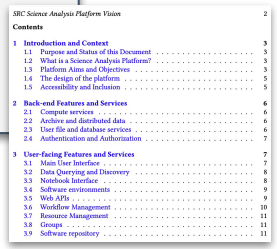
"IAM"* service deployed in UK (Rutherford Appleton Lab) https://ska-iam.stfc.ac.uk/login

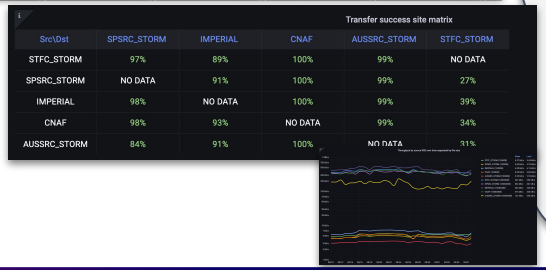
Purple



Comparison of Science Platforms Vision Document

Tangerine





Rucio authentication tokens and Improved monitoring (SKAO)
New storage in Spanish SRC (Granada) and Australian SRC (Perth)

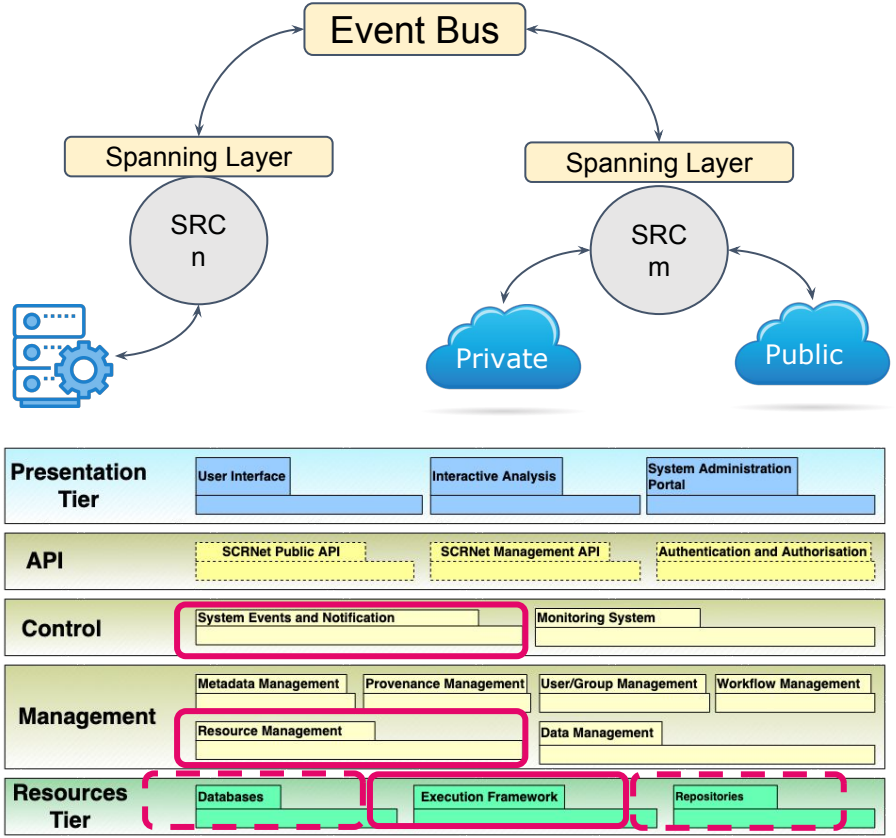
Cyan
Coral
Blue-Lavender

Federated Infrastructure

- Discussions on a federated software platform
- Solve wide-area multi-stage workflows (High Performance Data Analytics, AI, Visualisation)
- Heterogeneous platforms (Cloud, Grid, Hardware, etc)
- Workflow as a Service
- Spanning Layer
  - Execution Framework abstraction

Architecture targets during PI16:

- Common understanding
- Existing solutions

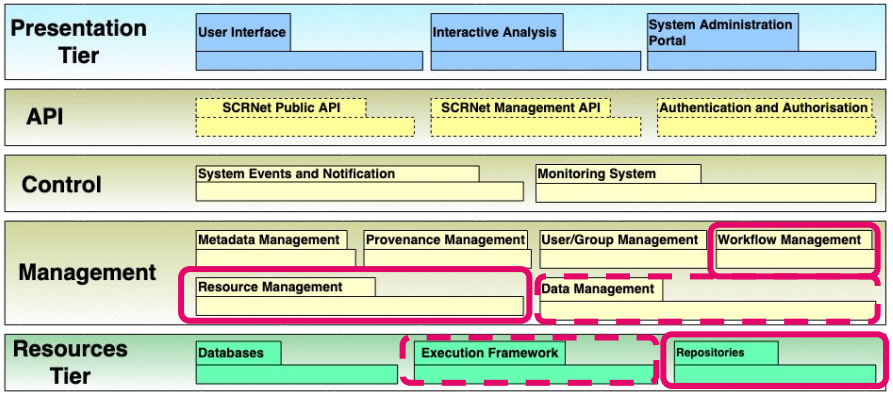


Container Management System

- Software Containers are a possible solution for many questions within the SRCNet
  - Coordinated Deployment
  - CI/CD
  - Containerised Science Workflows
- Common system is a must
- Federated system is a nice to have
  - High availability
  - Better performance
  - SRC delegation

Architecture targets during PI16:

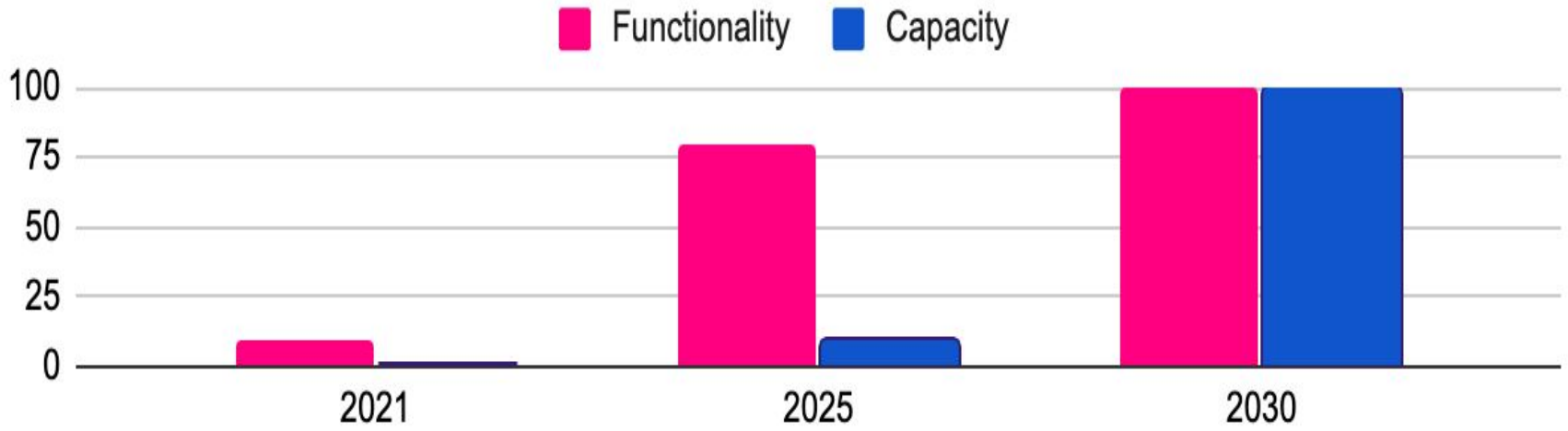
- Common understanding
- Create a container repository
  - GitLab, Docker hub, etc
- Study federated solutions



IVOA standards impacting SKAO

- Radio Data Interest group standards
- Provenance Data Model
  - ODPs and ADPs workflow characterization
- Image/Cube data model
- Proposal DM and ObsLocTAP protocol
- Single Sign-On Profile, Credential Delegation, Group Membership Service
- VOEvent
- SODA: Server-side Operations for Data Access
- TAP/ADQL
- HiPS, hierarchical multi-resolution astronomical maps
- VO and the Cloud

Timeline



Thanks for your attention