



Assessment of Rucio to manage the SKA data lake

Rucio Community Workshop, San Diego

James Collinson (SKAO)

Rosie Bolton (SKAO), Rob Barnsley (SKAO), James Walder (STFC)



Outline

- The SKA Project
 - **SKA Regional Centre Network (SRCNet)** - enabling science with an **exabyte-scale** data archive
 - Timeline to **science verification**
- SRCNet data management principles
- Options for managing the SRCNet data lake
 - Overview of the candidate solutions
 - Why **Rucio** and **FTS** were selected for **SRCNet v0.1**



SKAO Mission

“The SKAO’s mission is to build and operate cutting-edge radio telescopes to transform our understanding of the Universe, and deliver benefits to society through global collaboration and innovation.”

We acknowledge and recognise the Indigenous people and cultures who have traditionally lived on the lands on which our facilities are located.

SKAO



The SKA project in numbers

**€1.3
BILLION**

CONSTRUCTION
COST (2021 €)

**131,072
ANTENNAS**

IN WESTERN AUSTRALIA

**710 PETABYTES
PER YEAR**

OF SCIENCE DATA DELIVERED
TO SCIENCE USERS

**€0.7
BILLION**

FIRST 10 YEARS OF
OPERATIONS COST (2021 €)

**197
DISHES**

IN SOUTH AFRICA
(INCLUDING 64 MEERKAT DISHES)

**1 GLOBAL
NETWORK**

OF DATA CENTRES TO DELIVER SCIENCE-
READY DATA PRODUCTS TO END-USERS

8 YEARS

TO CONSTRUCT

16 COUNTRIES

PARTICIPATING IN 2023

50+ YEARS


OF TRANSFORMATIONAL SCIENCE




SKAO Science Drivers



Cosmology



Cradle of Life



Epoch of Reionization



Extragalactic Continuum



Extragalactic Spectral Line



Gravitational Waves



High Energy Cosmic Particles



HI Galaxy Science



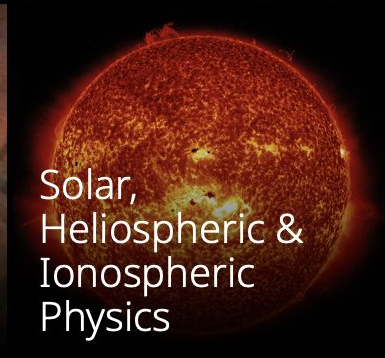
Magnetism



Our Galaxy



Pulsars



Solar, Heliospheric & Ionospheric Physics

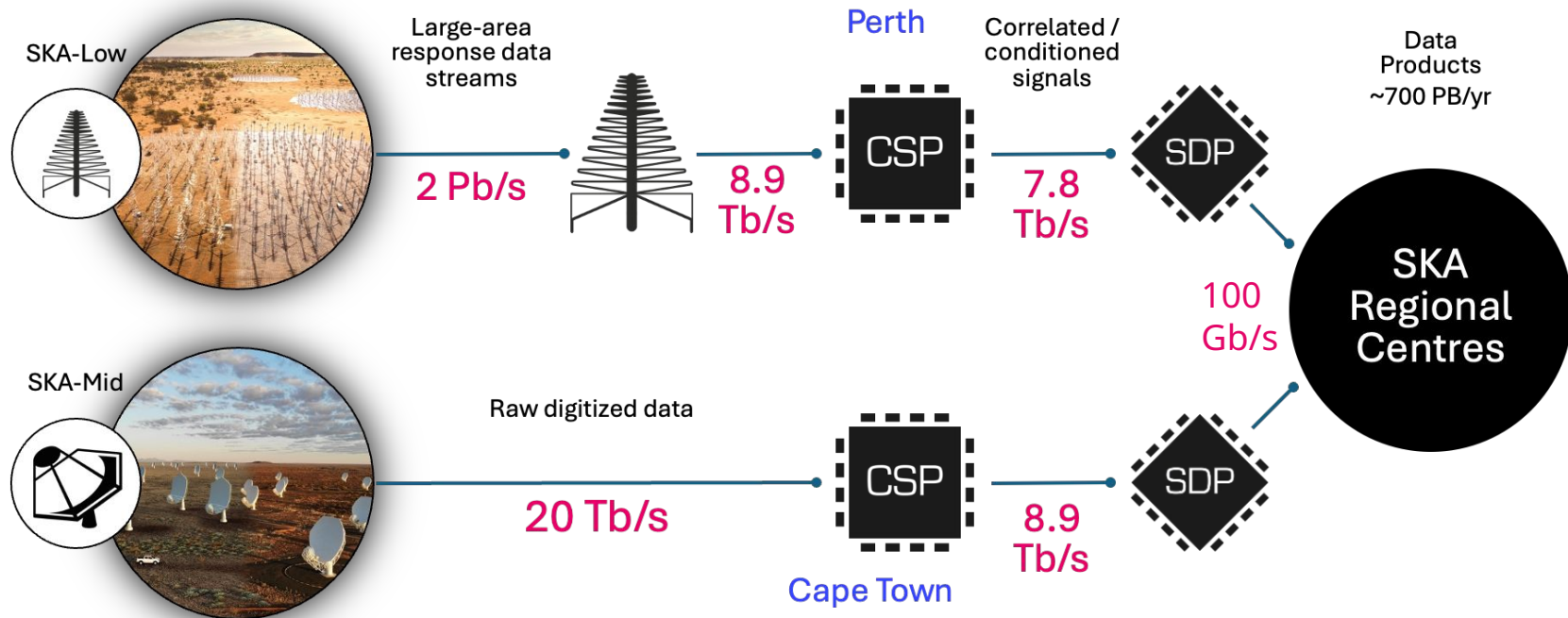


Transients



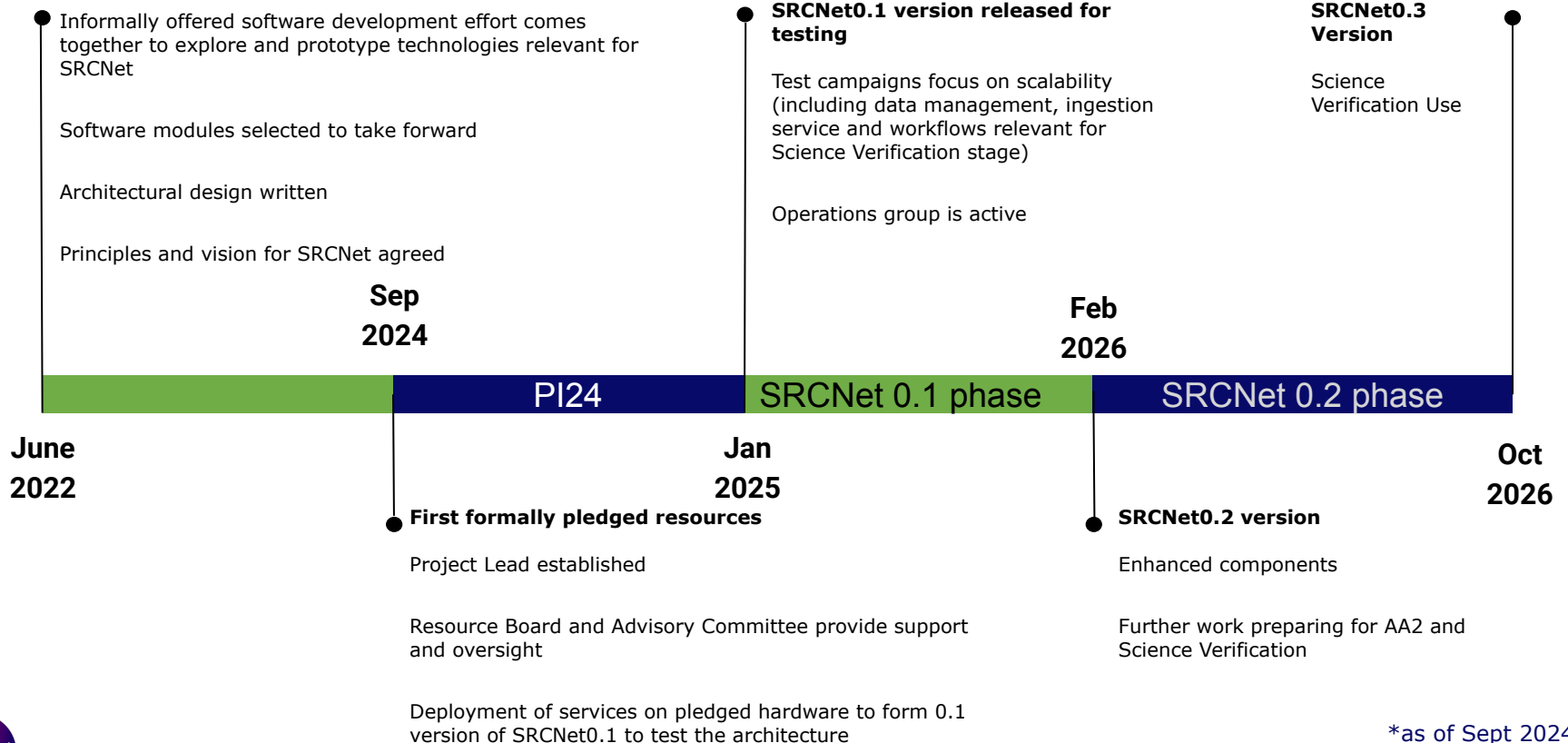
VLBI

SKA Data Rate to the SKA Regional Centres



SRCNet timeline*

SRCNet Software development collaboration begins



*as of Sept 2024

SRCNet Data Management

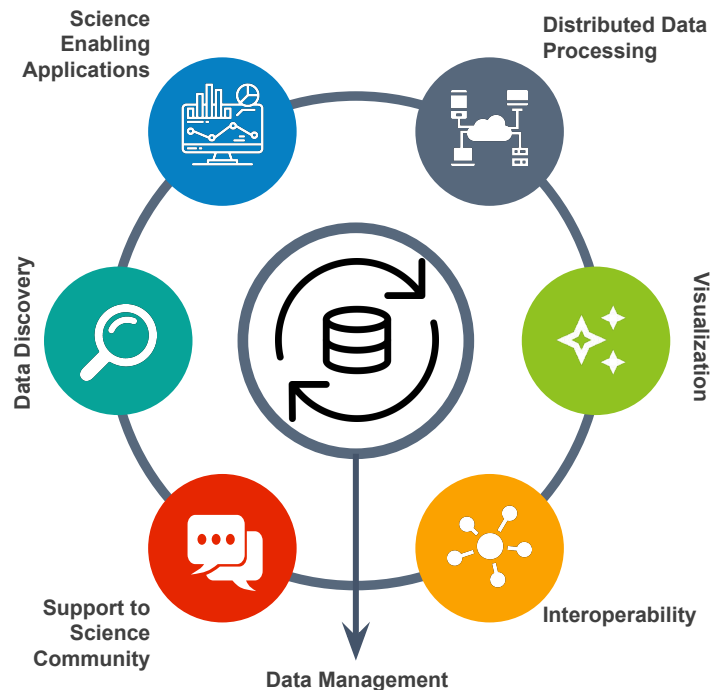
A core component of the SRCNet software stack, managing the **SKA Science Archive** globally, ensuring **storage efficiency**, and **data availability and integrity**

Fundamental principles:

- SRCNet will be **only point of access** to SKA data for the user community
- SRCNet **Operations** will be supported in a distributed fashion and **coordinated** by the **SRC Operations Group**
- Resources pledged into the SRCNet will become part of a **global federated pool**
- Network resources will be allocated to **projects or teams** rather than individuals or groups
- The physical location of SKA data products will be determined to **optimise access** and **minimise data redistribution**
- The lifecycle of science data products will be **managed by the SRC Operations Group**

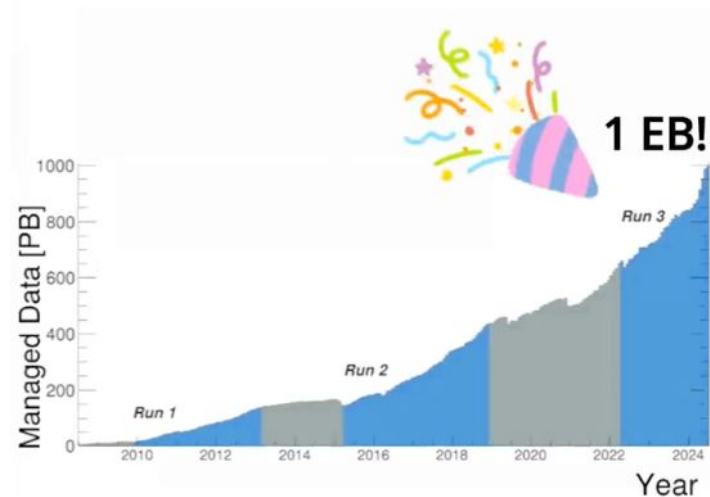
Data Management challenges:

- **Long haul network links** between SRCNet sites
- Number of files and **data volume**
- Data lifecycle and **quality of service**



Candidate 1: Rucio

- Rucio is a **large scale** scientific **distributed data management** tool
 - Built by/for the **ATLAS** experiment but since has been adopted by **CMS**, **Belle II**, and astronomical observatories such as **Rubin** and **CTAO**
 - ATLAS alone recently hit **1 Exabyte** under management
- Rucio architecture: server, daemons, middleware, applications
 - **Server** exposes a REST API, handles authentication, maintains DB
 - **Daemons** handle asynchronous work to keep the physical system in the requested state, e.g. transfers, deletions to ensure data locations match database
- Data lake state is parameterised by declarative statements ("**rules**") describing data policies
- **Storage agnostic**: Rucio is not responsible for managing storage - as long as there's a communication protocol to talk to it, it can be connected
- Uses separate **File Transfer Service** (FTS) to broker file replications via third-party copy

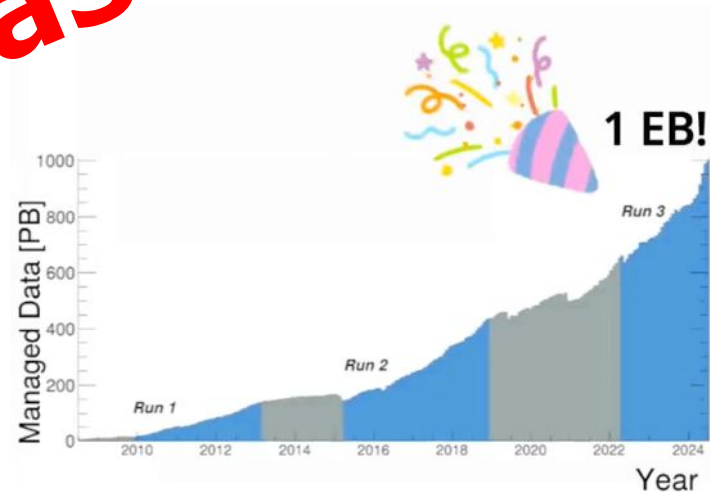


https://www.linkedin.com/posts/atlas-collaboration_new-milestone-for-the-atlas-experiment-weve-activity-7217502790237974530-Mpt2



Candidate 1: Rucio

- Rucio is a **large scale** scientific **distributed data management** tool
 - Built by/for the **ATLAS** experiment but since has been adopted by **CMS**, **Belle II**, and astronomical observatories such as **Rubin** and **CTAO**
 - ATLAS alone recently hit **1 Exabyte** under management
- Rucio architecture: server, daemons, middleware, applications
 - **Server** exposes a REST API, handles authentication, maintains DB
 - **Daemons** handle asynchronous work to keep the physical system in the requested state, e.g. transfers, deletions to ensure data locations match database
- Data lake state is parameterized by declarative statements ("**rules**" describing data policies)
- **Storage agnostic**: Rucio is not responsible for managing storage - as long as there's a communication protocol to talk to it, it can be connected
- Uses separate **File Transfer Service (FTS)** to broker file replications via third-party copy

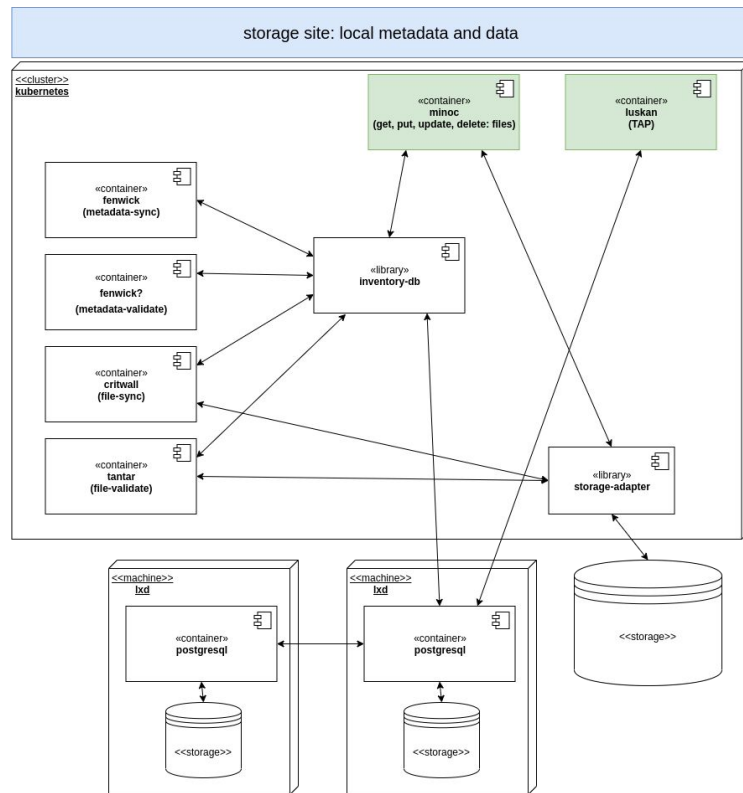


https://www.linkedin.com/posts/atlas-collaboration_new-milestone-for-the-atlas-experiment-weve-activity-7217502790237974530-Mpt2



Candidate 2: CADC Storage Inventory

- **Storage Inventory** has been developed to manage a federated data lake (5 PB) across **3 sites in Canada**
- Comprises a collection of services which **each site operates**, to manage a **subset of the total data archive**
 - Services for data discovery, metadata synchronisation, enacting transfers
- Central '**Global**' **database** tracks **data locations** across the **individual storage sites**
- Closely integrated with the recommended standards set out by the International Virtual Observatory Alliance (**IVOA**)
- Decentralised approach
 - + Fault tolerance in event of site outage
 - ! Higher operational burden at each site



Relating to the SRCNet principles

- Noteworthy principles:
 - *Resources pledged into the SRCNet will become part of a **global federated pool***
 - *The physical location of SKA data products will be determined to **optimise access** and **minimise data redistribution***
 - *The lifecycle of science data products will be **managed by the SRC Operations Group***
- These are **more closely aligned to** the **Rucio** principle of centrally managed replication rules and data policies
 - CADC SI would require further development to meet these, since **global requirements are not guaranteed by design**
- Therefore Rucio was determined to be the more appropriate design for adoption by SRCNet v0.1



Rucio: Additional strengths

- Ability to adapt for **astronomical metadata**
 - Developed metadata plugin which enables data discovery using existing astronomical client libraries (**see Rob Barnsley's talk**)
- **Loose coupling** between Rucio and FTS allows component switch out for adaptability and maintainability
- **Proven to scale** to Exabyte scale in ATLAS
- Adopted by an ever-increasing range of projects, including observatories
- Ease of onboarding sites (range of technologies/protocols supported)
- **Strong development community**
 - Relatively low barrier for new developers to start contributing
 - 'Good first issues'
 - Code guidelines, dev environment, open dev meetings...



Rucio: Worry list

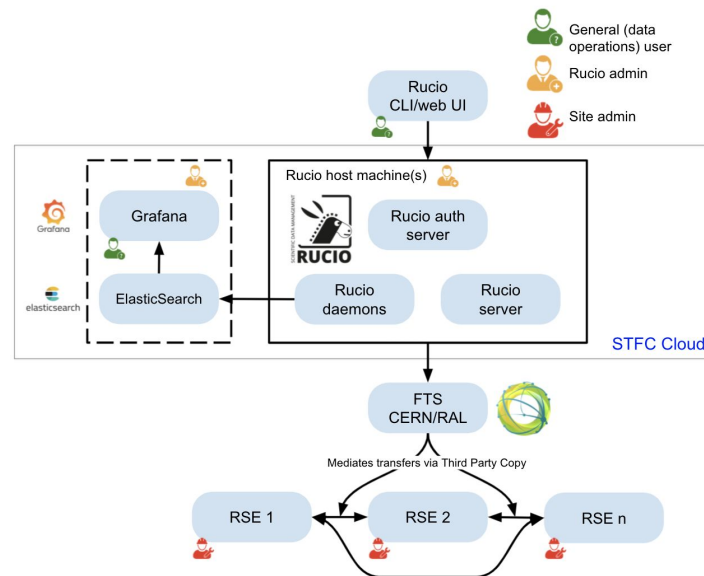
- Differing **development timescales** across different projects
 - Has proven an issue when certain **PRs are blocking further dev***
- **Certificate bundles** - is this going to bite us in the future!?
 - Should we plan to maintain our own certificate bundles in the future?
- **Hierarchical data** remains a challenge
 - Possible solutions discussed but disadvantages of each

*Example: Issue <https://github.com/rucio/rucio/issues/6372> with fix <https://github.com/rucio/rucio/pull/6373> is currently blocked by <https://github.com/rucio/rucio/issues/6666>



SRCNet Rucio instance

- We're running a **K8s**-based deployment at STFC Cloud (RAL, Oxfordshire, UK)
- Recently switched to in-house FTS and **Indigo IAM** instance for authentication via OIDC
- National SRCs with RSEs configured in data lake (in addition to SKAO):
 - Canada, China, Italy, Japan, Netherlands, Spain, Sweden, Switzerland, UK
- **Testbed** for testing and development



Towards an SRCNet 'Data Challenge'

- SRCNet v0.1 will be used to run a series of data challenges
- Ingest representative SKA data in to the data lake and replicate across the network
 - Performance and functionality tested

Transfer failure site matrix ⓘ

Src\Dst	STFC_STORM_ND	STFC_STORM	SPSRC_STORM	NLSRC_PROD_DC	NLSRC_DCACHE	KRSRC_STORM	JPSRC_STORM	IMPERIAL
._XRD_DEVCEPHFS	0%	0%	0%	0%	0%	100%	0%	0%
STFC_STORM	100%	-	0%	0%	0%	100%	0%	0%
SPSRC_STORM	100%	100%	-	0%	0%	100%	0%	0%
RC_PROD_DCACHE	0%	0%	0%	-	0%	100%	0%	0%
NLSRC_DCACHE	0%	0%	0%	0%	-	100%	0%	0%
JPSRC_STORM	100%	100%	0%	0%	0%	100%	-	0%
IMPERIAL	0%	0%	0%	0%	0%	100%	0%	-
CNAF	100%	100%	0%	0%	0%	100%	0%	0%
CASRC_XRD	0%	0%	0%	0%	0%	100%	0%	0%



Summary

- **After comparing two candidate distributed data management systems for the SRC Network use case, we have adopted Rucio for version 0.1**
 - Rucio shows a close match to the architectural design principles of the SRCNet
 - Proven to scale to SKA project data scales
 - Is adaptable to astronomical use cases
 - Relatively low operational overhead is advantageous
 - Healthy development community!



Thanks!

