The UK Square Kilometre Array Regional Centre (UKSRC) Node – How we get there

Contents

An introduction to the UK SRC Node for the UK SKA Community....1
What is the UKSRC and SRCNet?....1
What will the Node offer?....2
How, and when, will the UK develop and deploy the Node's Services?....3
What is does this mean for you ?....3
Timelines....3
The Roadmap of Expected Functionality and Community Inputs/Activities....4

An introduction to the UK SRC Node for the UK SKA Community

The STFC Open Call to create a UK SRC nodes states the following:

"A consortium to be established through an STFC Open Call for a proposal to deliver the UKSRC Node of the SRCNet"

The UKSRC Node will be deployed in a staged way over the next 2-3 years. This is an engineering and technical project that has to interact with, and be driven by, the evolving requirements of the UK SKA Science Community.

The main vehicles for these interactions will be:-

1. The Science Demonstrator Cases
2. Community Governance, including the Community Director and Community Manager
3. You/We

What is the UKSRC and SRCNet?

The SKA Observatory will produce data products that will need substantial further processing to meet the goals of SKA science programmes and users. To enable this processing a global network of SKA Regional Centre (SRCNet) Nodes will provide infrastructure, services and expert support that will enable a global capability to distribute, process and curate the data from the two SKA telescopes.

In order to develop, test and deploy the proposed UK SKA Regional Centre (UKSRC) Node and the global SRCNet itself, the UKSRC and the SRCNet have begun a three-year programme of prototyping, implementation, and early commissioning to deliver an operating SRCNet by 2025, which will achieve 10% capacity and 80% functionality of what is needed for full operations.

The UKSRC Forum, and the future UKSRC Project, has a 2-fold mission:

- To design, deliver and support a UKSRC Node that will be compliant with the global SRCNet architecture and standards, whilst providing local operational infrastructure commensurate with the scientific needs and priorities of the UK Science Community.
- To collaborate with SKAO and Members to ensure components of the SRCNet provide a collection of services and infrastructure that will comprise a global capability to distribute, process and curate the data from the SKA telescopes.

From November 2022 a UKSRC Node will be established that will incrementally deliver and offer SRCNet functionality. This Node will deliver the UK's contribution to the SRCNet and the National UKSRC Project and UK SKA Science Community.

What will the Node offer?

The UKSRC Node will offer access to:

- The hardware, networking and enabling software required for researchers to exploit and process data from SKA and pathfinder/precursors ;
- A UK Operations and Federated Services capability to allow the UK SKA Science Community to access UK-based resources as a precursor to using the wider SRCNet resources;
- Software, Workflows and Data Services that can work in the proposed Federation environment;;
- The UKSRC Science Data Archive Service;
- Training resources;
- A programme of UK-focussed Science User Engagement activities.

These will complement the SRCNet activities, but will be focussed on generating UK specific deliverables and milestones. This focus is needed to ensure the UK SKA science community is prepared for, lead/drive and exploit early the opportunities afforded by the SKA.

This access will be supplied by several technical layers. An SRCNet Node must have these layers.

- The Science Platform and User Engagement Layer provides the portal by which Research Projects and users engage with the SKAO data and generate scientific data and analyses for publication and dissemination.
- The Applications and Workflow Layer contains the software required to perform processing, analysis, modelling, and curational functions on the data.
- The Science Data Archive Layer allows all data products to be stored, catalogued and curated (e.g. addition of Meta Data). This layer must be composed of software common to all SRCs.
- The Data Logistics Layer (data transport and Networking) is the key functionality to ensure the SRCNet and SRC Nodes can function. The software must be common to all SRC Nodes.
- The Federated Computing and Operations Layer allows access to the computational resources and the required data (which may be distributed among several sites), so allowing workflows and their constituent applications to be run.

    - Although not marked as being common to all layers, the cloud technology middleware layer should be common to all SRC Nodes at the very least, otherwise no federated services will operate.

- The Compute, Storage and Middleware enable the data and applications to be stored and computed, therefore generating the advanced data products for scientific analysis.

How, and when, will the UK develop and deploy the Node's Services?

The UKSRC Node will be deployed in a staged way over the next 2-3 years. This is an engineering and technical project that has to interact with, and be driven by, the evolving requirements of the UK SKA Science Community.
The main vehicles for these interactions will be:-

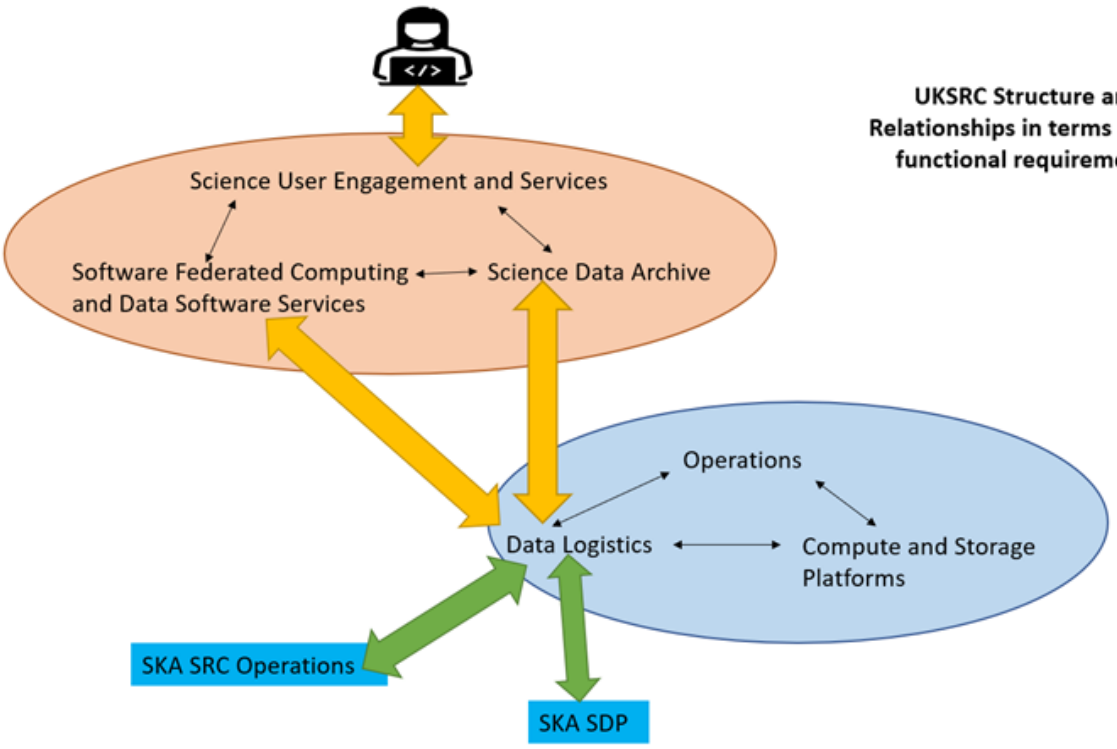
1. The Science Demonstrator Cases
2. Community Governance, including the Community Director and Community Manager
3. You/We

What is does this mean for you ?

- All of the UKSRC Node's services will be available through your laptop/desktop.
- It is envisaged that the resources you need can only be provided by compute clusters.
- This is a huge change for many of you, but familiar to those of the community who started out in the 1980 and 1990s.
- The technology being developed for the STFC IRIS Larger Laptop service will be used to preserve ways of working.
- You will use remote supercomputing services through laptops as though being run on the laptop.
- Notebooks such as Jupyter Notebooks are being configured to work with SRCNet services and resources.
- You will be given training and support for the above.

Timelines

- The Figure below shows the relationships between You, the UKSRC Node and the SRCNet.



- This will have to be built up incrementally over the next three years, and this can only happen with your collective help and participation.
- This is why a set of Demonstrator Cases with particular relevance to the scientific interests and requirements of the UK community, complete with timelines, that define the user requirements (i.e. what the researchers do) for working through these Demonstrator Cases, are requested.
    - There should be 6-8 of these and should encompass the UK Science Community's research interests.
- The Demonstrator Cases will be used to test the proto-SRC's technical functions and efficacy.

Setting up a UKSRC Node and providing access to it is a key Engagement Activity. The UKSRC Node and its services will therefore develop iteratively and incrementally using modern agile methodologies.

The Science Demonstrator Cases supply the workflow requirements we need to design, test and deploy systems and services. The only way to test these are to use unit tests for component functionality and benchmarks (science outcomes) to test the correctness of workflow and applications when they utilize UK systems and services.

The Roadmap of Expected Functionality and Community Inputs/Activities

Table 1: Sample UKSRC Node Evolution

Timeline (month)	**Demonstrator Case Requirement/Activity**	**UKSRC Node Activities and Deliverables**

June 2022	Science Demonstrator Cases: • Workflow functional requirements • Resource Requirements • Preferred way of working • Programme of Work • Creation of Demonstrator Consortia	Design and implementation of Prototype Service: • Workflow Story Boards • Collation of first benchmarks Building The Prototype Service: • Hybrid Cloud between 3 UK sites (RAL, Cambridge,and Manchester) • Larger Laptop Service and Jupyter Notebooks • Authentication and Authorisation • Science Platform with first set of applications and workflows • Federated Data Management • Simple Data repository
Jan 2023	• First Use of Prototype Service by Demonstrator Consortia • Update of Demonstrator Cases and new requirements updates and additions	• Addition and optimization of new functionality to the Prototype Service • Creation of first Science Data Archive • Movement of large datasets into the Archive from UK and International Sources
July 2023	• Update of Demonstrator Cases and new requirements updates and additions	• Increase the Compute and Storage Platform capability • Update the Science Platform(s)
Jan 2024	• Update of Demonstrator Cases • Add new requirements updates and additions in the light of progress for AA0.5 (the first prototype SKA systems)	• Update to enable the reception and processing of technical commissioning data. • Review performance and adjust service accordingly.
July 2024	• Update of Demonstrator • Add new requirements updates and additions in the light of progress for AA1(the second prototype SKA systems)	• Update to enable the reception and processing of technical commissioning data. • Review performance and adjust service accordingly.