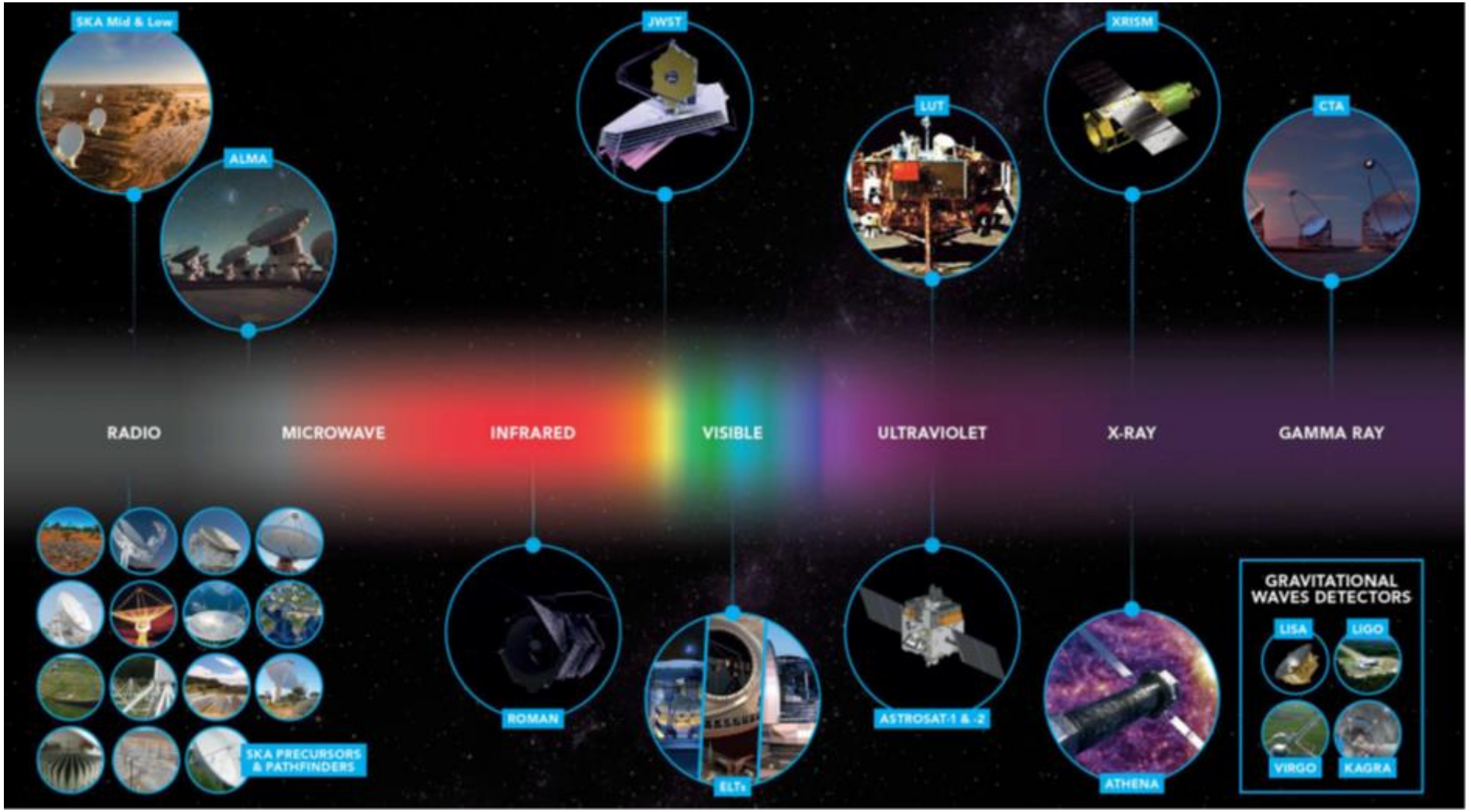
UK SKA Regional Centre

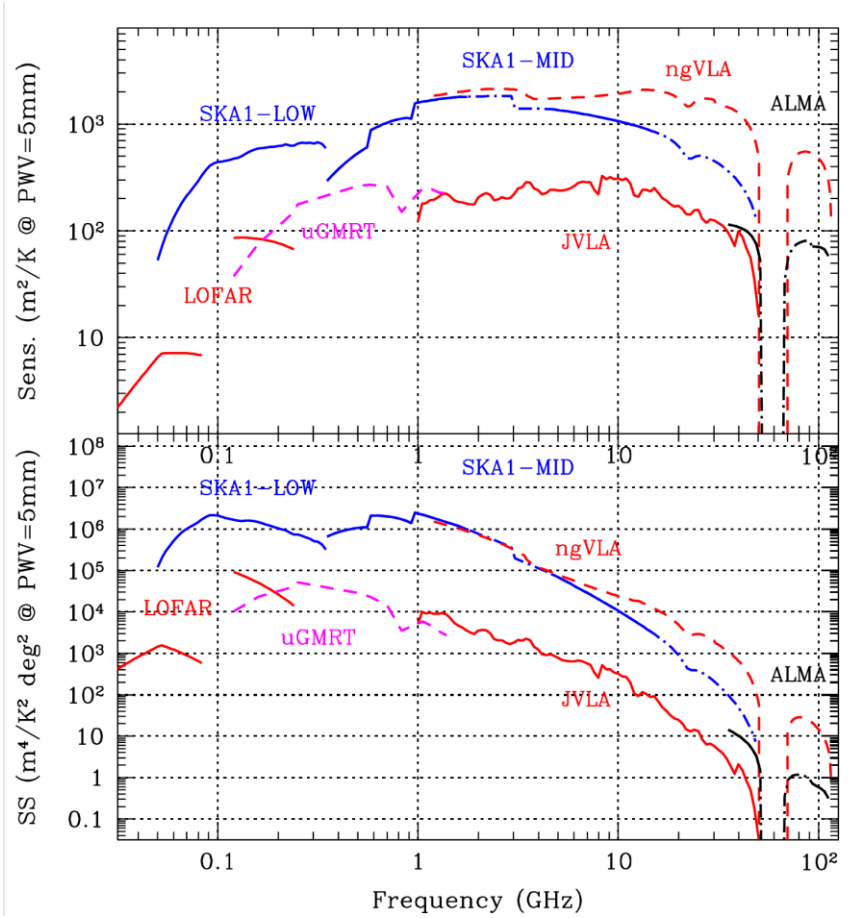
Jeremy Yates, Chair

Where does SKA Fit



Science Drivers

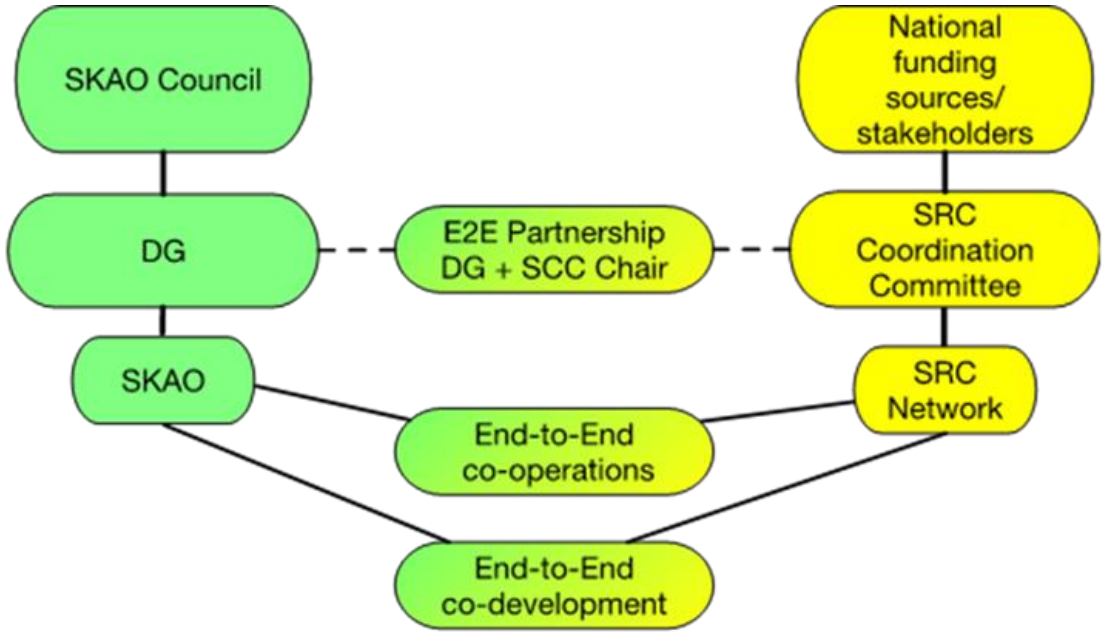
SKA Sensitivity



Time line

	SKA-Low	SKA-Mid
Start of construction (T0)	1ST JULY **2021**	1ST JULY **2021**
Earliest start of major contracts (C0)	AUGUST **2021**	AUGUST **2021**
Array Assembly 0.5 finish (AA0.5) SKA-Low = 6-station array SKA-Mid = 4-dish array	FEBRUARY **2024**	MARCH **2024**
Array Assembly 1 finish (AA1) SKA-Low = 18-station array SKA-Mid = 8-dish array	FEBRUARY **2025**	FEBRUARY **2025**
Array Assembly 2 finish (AA2) SKA-Low = 64-station array SKA-Mid = 64-dish array, baselines mostly <20km	FEBRUARY **2026**	DECEMBER **2025**
Array Assembly 3 finish (AA3) SKA-Low = 256-station array, including long baselines SKA-Mid = 133-dish array, including long baselines	JANUARY **2027**	SEPTEMBER **2026**
Array Assembly 4 finish (AA4) SKA-Low = full Low array SKA-Mid = full Mid array, including MeerKAT dishes	NOVEMBER **2027**	JUNE **2027**
Operations Readiness Review (ORR)	JANUARY **2028**	DECEMBER **2027**
End of construction	JULY **2029**	JULY **2029**

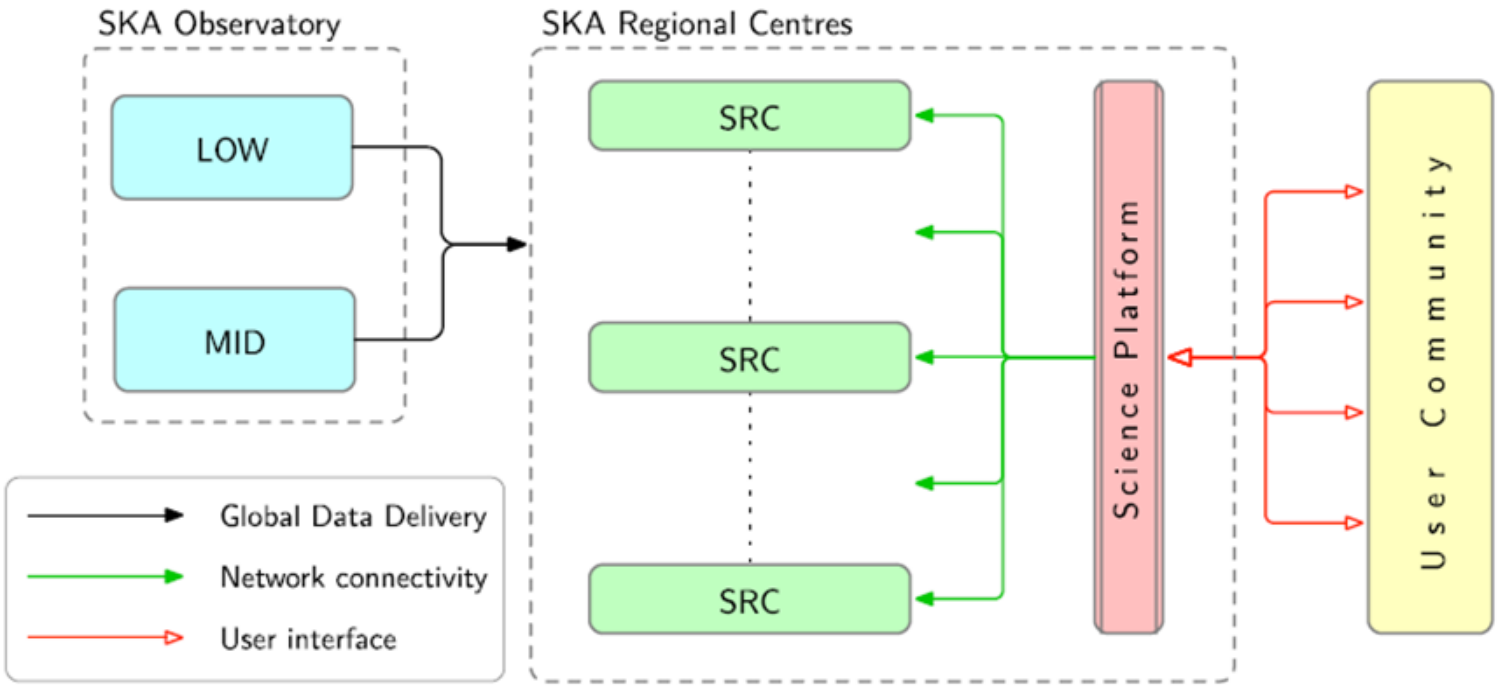
SRC Network



SKAO and SRC Network End-to-End Partnership Governance and Reporting

Figure 1: In the diagram above, and in the rest of the document, green is used to indicate SKAO and yellow to indicate the SRC Network.

A collaborative network of SKA Regional Centres hosts Observatory Data Products produced by the SKA Observatory. Access to these ODPs, as well as the tools and processing power necessary to fully exploit the science potential of those products, is provided by a Science Platform.



SRC Network Activities 2021-2023

- It meets monthly
  - Peter Quinn is the Chair and the Australian member
- Architecture Development
  - Phase 1: April-Dec 2021
    - High level Architecture
    - Requirements gathering
    - November – L1 cross-review
    - Work Breakdown Structure for Prototypes by end of 2021
  - Phase 2: Jan-Oct 2022
    - L2 refinement,
    - L3
    - Protoytping
  - Phase 3: Nov 22-Mar 23
    - Development of Implementation Plan

SRC Working Groups

SRC Working Group	Responsibility	UKSRC Participants
WG0: The Architecture of the SRC Network	Delivering the SRC network Design and Implementation Plan	Jeremy Yates (UCL, Co-ordinator)) Robert Beswick (Manc)
WG1: Data Logistics	Deliver data from the Observatory to SRCs	Keith Grainge (Manc) Robert Hughes-Jones (Geant/Manc)
WG2: Operations	Enabling efficient and collaborative use the SRC network computing, storage and communication resources	Ian Collier (RAL) Jeremy Coles (Camb))
WG3: Software Federated Computing and Data Software Services	programming and federated execution environments, to interact with and analyse the delivered data. This includes data logistics and access across and within the SRC network nodes	Martin Hardcastle (Herts) Anna Scaife (Manc)
WG4: Science Data Archive	Provide a functional and persistent SKA Science Archive incorporating FAIR and VO services	Paul Harrison (Manc) Rachael Ainsworth (Manc)
WG5: Compute and Storage Platforms	Provide the computational and storage resources and services	Jeremy Yates (UCL)
WG6: Science User Engagement	Engage the science community in tuning the requirements for the SRC network in order to maximize the science return. Guide the scientific community at large towards the new end-to-end procedures that the SKA era will require for performing a scientific program.	Robert Beswick (Manc) Leah Morabito (Durham)

Prototyping Projects

UKSRC RO /HEI	Prototype Activity
MANC, CAM, RAL	Prototype 1a and 1b: Data products replication, distribution, and synchronization across multiple locations (Rucio (1a) and CADC (1b) approaches) This would include moving data from SDP to the SRC network and between SRCs. This is the higher priority prototype (to be converted to operational). That could include different aspects like Direct copy and third-party copies (between two SRCs but initiated by a third one) or Advanced data products replication. For these prototypes there is a need of roadmap synchronisation with the SDP roadmap
RAL, UCL	Prototype 2: Federated Authentication and Authorization API Definition and adoption of a federated Authentication framework and Authorization API (needed for a secure data access)
MANC, HERTS, Dur	Prototype 3: Data Processing Notebooks Jupyter Notebook (or equivalent) integration for multi-user access, flexible processing and use of computing resources (including GPUs), data access and A&A federated systems.
MANC, HERTS, DUR	Prototype 4: Visualization of SKA data with high volume of users and high amount of data Instantiation and deployment of visualisation services such as CARTA or other visualisers for a high user demand and data volume environment, using flexible local computational resources. This includes storage sharing, data access, federated authentication, and integration with workload manager.
MANC, UCL	Prototype 5: Distribution of software, tools and services Provide an infrastructure that allows to centralise (or to distribute) all software, tools, services and documentation with version control. This scenario provides a marketplace (catalogue), including software (tools, etc.), containers (and their definitions), packages, documentation, etc. for both the SRCNet system (certified/validated) and individual users, for all components of the system. It will be used to deploy services, perform data processing, feed workflows-pipelines, instantiate functions as service on notebooks, etc. at each network site while maintaining provenance information related to software.

UKSRC FORUM

- Meets every 2 weeks
- Has begun Strategy Case : Timeline Below

Timeline	WG Tasks	UKSRC Task
April-May 2021	Requirements Gathering	State our own particular roles in each WG and what we contribute to them
June-July 2021	Generate draft document for SRC Network Participation, Governance, Development and Operations	Reports from WG activities and discussion of what the UK can contribute
Aug-Oct 2021	Prepare SRC Network Participation, Governance, Development and Operations documents for Council	Generate shape of UKSRC Strategy
Oct-Dec 2021	Present SRC Network Participation, Governance, Development and Operations plans to Council and production of final documents and plans	Generate formal UKSRC Strategy
Jan-Feb 2022	Possible start of SRC Network Participation, Governance, Development and Operations Plan	UKSRC Strategy is reviewed and a final version is produced.
March-May 2022		Open Call to implement to the UKSRC Strategy
June 2022		Winning Consortium begins work

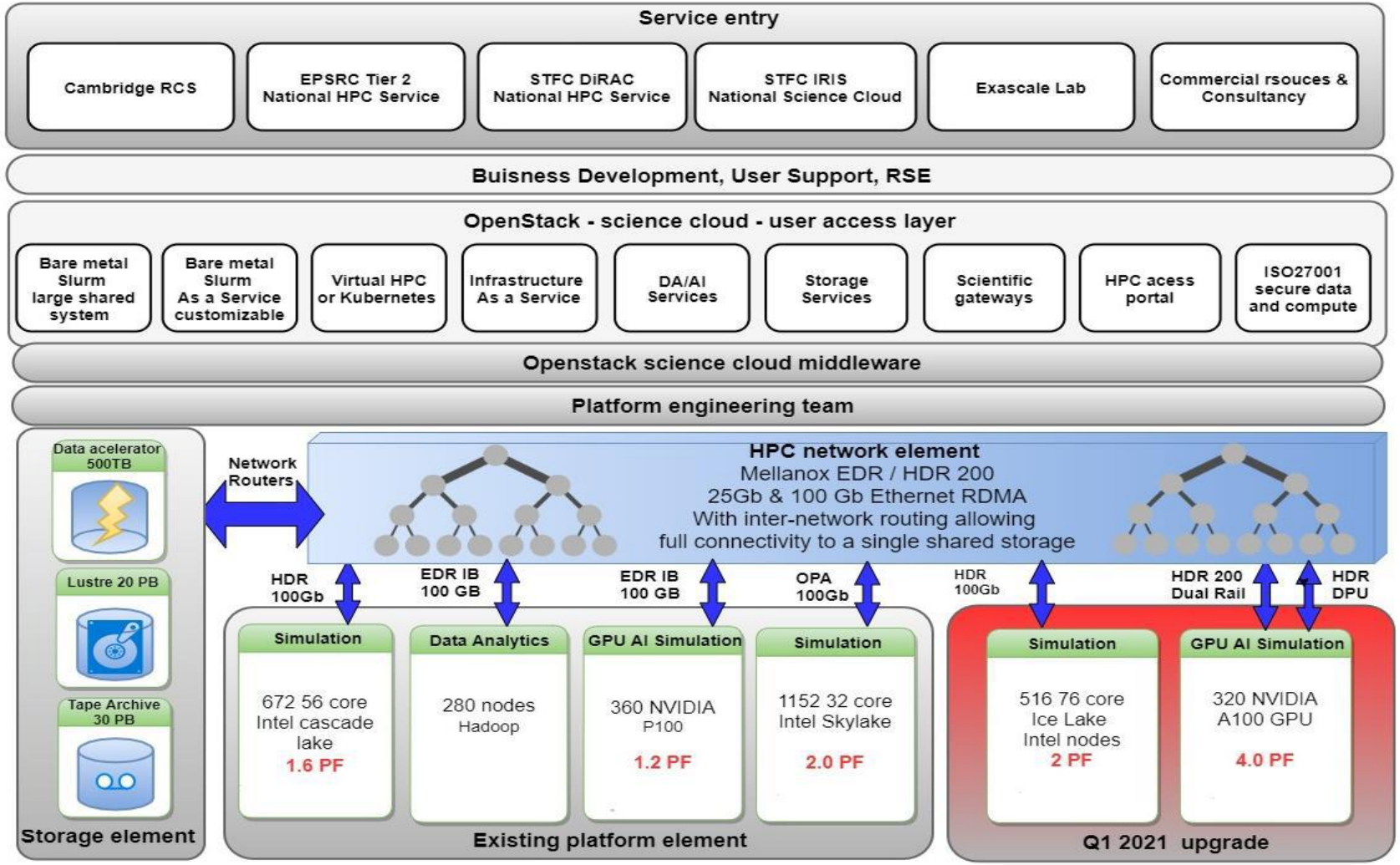
IRIS Digital Research Infrastructure Digital Assets Awards

DRI proposal	Deliverables
UK SRC Proposal to Establish a High-Performance Data Transfer Node Infrastructure at a Supercomputing Site – Hughes-Jones, Grainge, Yates	Build and demonstrate a data transfer infrastructure at the Cambridge IRIS Supercomputer site capable of receiving SKA Observatory Data Products (ODP) from the two telescopes and exchanging Advanced Data Products (ADP) between SKA Regional Centres.
Designing and Creating SRC workflows to run on native cloud and the supporting repository infrastructure – Harrison, Scaife	A repository that would enable curation and archiving of cloud instances.
Enabling IRIS Access for Radio & mm-wavelength Astronomy – Beswick, Scaife, Hardcastle, Ainsworth	**D1**: Domain specific (Radio astronomy/interferometry) IRIS (cloud & data centre) user & access documentation, including guides, online tutorials and community dissemination activities. **D2:** Documented example use-cases, including suggested workflows and steps towards implementation. **D3:** Evaluation report and implementation plan for development of interfacing between Science Platforms
	Total

IRIS Infrastructure Inputs

- 2018 estimate of UKSRC requirement (Scaife, Coles, Ashdown et al)
  - Heterogeneous (GPU+CPU) solution is still correct and the unified GPU+CPU memory models that should be available on a 23-24 timescale will deliver the requirements
    - Native Cloud Heterogeneous Supercomputer at Cambridge now built and operational
  - System interconnect (networking) and Data (storage) Infrastructure requirements require major updating (a lot has changed in 2 years)
    - DiRAC3 and ExCALIBUR (the IO testbed at Cambridge) provide useful starting points and expertise.
  - Basic assumption is that there will be a:-
    - Tier 1 RAL for loosely parallel work and the big tape store
    - Tier 1 Cambridge for parallel work and visualisation
    - Tier 3 Manchester for Development and Operations
    - GridPP and DiRAC Tier 2s taking work as appropriate

We have a modern cloud native science platform in STFC IRIS and DIRAC located at Cambridge



- Heterogeneous Computing
- Cloud native computer
- Enables Development and testing
- Supports Preliminary Science
- Also the ExCALIBUR Hardware and Enabling Software testbeds are available - https://excalibur.ac.uk/testbeds