



REGIONAL  
CENTRE  
NETWORK

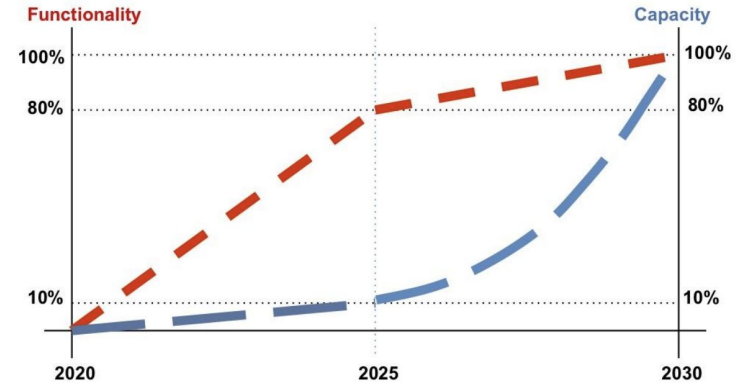
## SRC Net v0.1 Scenarios

Jesús Salgado  
SRCNet Architect



# White Paper SRCNet definition

- Resources to be set-up into SRCNet present into SRC white paper
- Based in continuous deployment with ramp-up in resources
- Functionality percentage difficult to evaluate but resources figures can be calculated
- Storage: 530x2 PB, first year
- Compute: 35 FLOPS



Ramp-up of the SRC network



# SRCNet Roadmap Document

		SRCNet v0.1	SRCNet v0.2	SRCNet v0.3	SRCNet v1.0b	SRCNet v1.0
		Jan 2025	January 2026	Sep 2026	Nov 2027	Jun 2028
Deployment (%)		2.00	10.00	15.00	50.00	100.00
Country	Share (%)	Storage (PB)	Storage (PB)	Storage (PB)	Storage (PB)	Storage (PB)
UK	19	4.03	20.14	30.21	100.70	201.40
South Africa	18	3.82	19.08	28.62	95.40	190.80
Australia	18	3.82	19.08	28.62	95.40	190.80
China	10	2.12	10.60	15.90	53.00	106.00
Canada	7	1.48	7.42	11.13	37.10	74.20
Italy	6	1.27	6.36	9.54	31.80	63.60
India	5	1.06	5.30	7.95	26.50	53.00
France	3	0.64	3.18	4.77	15.90	31.80
Netherlands	2	0.42	2.12	3.18	10.60	21.20
Japan	2	0.42	2.12	3.18	10.60	21.20
Spain	2	0.42	2.12	3.18	10.60	21.20
Portugal	2	0.42	2.12	3.18	10.60	21.20
Switzerland	2	0.42	2.12	3.18	10.60	21.20
Sweden	2	0.42	2.12	3.18	10.60	21.20
South Korea	1	0.21	1.06	1.59	5.30	10.60
Germany	1	0.21	1.06	1.59	5.30	10.60
Total	100	21.20	106.00	159.00	530.00	1060.00

		SRCNet v0.1	SRCNet v0.2	SRCNet v0.3	SRCNet v1.0b	SRCNet v1.0
		Jan 2025	January 2026	Sep 2026	Nov 2027	Jun 2028
Deployment (%)		2.00	10.00	15.00	50.00	100.00
Country	Share (%)	Computing (PFLOPS)	Computing (PFLOPS)	Computing (PFLOPS)	Computing (PFLOPS)	Computing (PFLOPS)
UK	19	0.13	0.67	1.00	3.33	6.65
South Africa	18	0.13	0.63	0.95	3.15	6.30
Australia	18	0.13	0.63	0.95	3.15	6.30
China	10	0.07	0.35	0.53	1.75	3.50
Canada	7	0.05	0.25	0.37	1.23	2.45
Italy	6	0.04	0.21	0.32	1.05	2.10
India	5	0.04	0.18	0.26	0.88	1.75
France	3	0.02	0.11	0.16	0.53	1.05
Netherlands	2	0.01	0.07	0.11	0.35	0.70
Japan	2	0.01	0.07	0.11	0.35	0.70
Spain	2	0.01	0.07	0.11	0.35	0.70
Portugal	2	0.01	0.07	0.11	0.35	0.70
Switzerland	2	0.01	0.07	0.11	0.35	0.70
Sweden	2	0.01	0.07	0.11	0.35	0.70
South Korea	1	0.01	0.04	0.05	0.18	0.35
Germany	1	0.01	0.04	0.05	0.18	0.35
Total	100	0.70	3.50	5.25	17.50	35.00

- Modified from 10% (white paper) to 2% for 2025
- The important numbers are the Totals and the 2028 Capacity. (We accept that SRCs will ramp up at different rates)



# Roadmap Timeline

First quarter of 2025



**SRCNet v0.1**

Milestone	Description	SRC Net Functionality	Scope (users)
<b>SRCNet v0.1</b> First quarter of 2025	Opportunity to engage SRCNet with AA0.5 data transfer and access.	<ul style="list-style-type: none"><li>• Test data (and some precursors data) disseminated into a prototype SRC Net</li><li>• Data can be discovered through queries to the SRC Net</li><li>• Data dissemination to SRC nodes</li><li>• Data can be accessed through a prototype data lake</li><li>• Data replication. Data can be moved to a local SRC area where non-connected local interactive analysis portals (notebooks) could allow basic analysis</li><li>• Unified Authentication System for all the SRCs</li><li>• Visualisation of imaging data</li></ul>	SRC ART members  Members of SKA Commissioning team



# Use cases for v0.1 (after review)

- Main use cases requested by SEAC
- Ingestion and dissemination of relevant data
  - Possible deployment of ingestion nodes at pipeline execution centres (Australia and, if possible, South Africa)
- Science use cases using the distributed data lake to evaluate possible bottlenecks
  - Precursors data and SKA test data
- Execution of “known” science use cases on SKA test data
  - Evaluation of possible problems due to size or format of the pipelines generated data
- Science Data Challenges data and Software associated to analysis
  - Comparison between execution to analysis on data within local node and data lake



# Test data for SRCNet v0.1

- **Precursors data:** Ingesting precursor data enables the execution of complete science use cases
- **SKA-like “test” data:** Synthetic datasets generated by SDP (Science Data Processor) pipelines, mirroring the scale and format expected from real SKA observations
- **Science Data challenges data or similar data:** Data already analysed in a locally controlled scenario using scientific workflows will provide information on the overhead provided by the SRCNet distributed data and resources
- <https://jira.skatelescope.org/browse/SP-4132>



# Expressions of Interest

- Resource level for 2025 requested for all SRCs.
- Three ratings:
  - Possibility for oversubscription
  - In line with Roadmap
  - Less resources but available
- Expressions of interest received by 9 countries to participate into the SRCNet v0.1, the first operational version with stable resources
- The rest of the countries will still be part of the SRCNet ART in development of the software
- <https://confluence.skatelescope.org/display/SRCSC/Summary+of+Expressions+of+Interest+for+SRCNet+v0.1>



# Storage and Compute

	Spain SRC	Netherlands SRC	Sweden SRC	UK SRC	Switzerland SRC	China SRC	Canada SRC	Japan SRC	Italy SRC
Storage (PB)	0.500	0.100	0.290	4.000	0.400	1.000	1.200	0.144	0.300
Compute (PFLOPS)	0.010	0.010	0.011	0.175	0.014	0.175	0.040	0.005	0.100
Percentage Storage (%)	6.302	1.260	3.655	50.416	5.042	12.604	15.125	1.815	4.036
Percentage Compute (%)	1.854	1.854	2.389	32.437	2.595	32.437	7.414	0.927	18.886
Harmonisation Rate *	0.020	0.100	0.036	0.044	0.035	0.175	0.033	0.035	0.333

- Oversubscription compared to roadmap doc
- According with roadmap expectations for this SRC
- Below Roadmap expectations but resources available

\*Computing/Storage (PFLOPS/PB)





# Possible Scenarios

- **Scenario 1: All nodes in the first version**
  - All nodes that has sent a EoI are part of the SRCNet v0.1 as operational deployment nodes
- **Scenario 2: 2-3 Big Nodes + Support on Development**
  - Only the nodes that have more resources are part of the SRCNet v0.1 as operational deployment nodes and the others support operational and monitoring tasks
- **Scenario 3: DevOps Approach, Integration+Operations**
  - Two environment networks are created (operational and integration). Limited number of nodes in the operational network
  - Possibility to include “organically” nodes into operational network from integration nodes between versions



# Scenario 1: All nodes in the first version

- *Benefits:*
  - *High Participation:* All countries have an early participation in the network
- *Challenges:*
  - *Operational Overhead:* Managing and maintaining numerous geographically dispersed nodes can be complex for a first version
  - *Resource Imbalance:* Countries with lower resources might have less capacity to respond to operational challenges of an immature version and limiting contribution to development
  - *Software Management:* Maintaining consistent software versions and ensuring compatibility across all SRCs can be challenging



# Overall: Scenario 1

- Scenario 1 prioritizes inclusivity while acknowledging the challenges of managing a large, geographically diverse network. Also, this ensures maximum representation from all nodes
- However, this approach could be challenging to be implemented due to operational constraints and due to the level of maturity of the SRCNet software
- Also, most of the efforts at this stage are still on development so to solve these challenges could limit resources on this area



## Scenario 2: 2-3 Big Nodes + Support on Operations

- *Candidates for Big Nodes:*
  - UK, China, Canada,...
- *Pros:*
  - Lower operational overhead (fewer nodes to manage)
  - Efficient utilization of resources (higher resource countries contribute compute/storage) with less data transfer latency challenges at this stage
- *Cons:*
  - Limited participation for some countries (might feel excluded)
  - Not getting experience on connecting development and operations
  - Support to operations is unclear



## Scenario 2: Overall

- This scenario prioritizes efficiency but could lead to a less inclusive environment
- It could produce “political” implications for some partners as exposes too much secondary roles for SRCs willing to collaborate on SRCNet v0.1
- Allowing operations tasks on SRCs operational nodes from other SRCs experts could present problems if nodes restrict access
  - Ability to provide support to v0.1 for “external” experts, limited



## Scenario 3: DevOps Approach, Integration+Operations

- *Operational Environment*
  - Limited number of nodes (e.g. UK, China, Canada)
  - This environment provides the core functionalities for scientific analysis. The Ops Environment prioritizes stability and runs well-tested software modules and services
- *Integration Environment*
  - Rest of nodes. Not needed full resources
  - This environment serves as a testing ground for new software modules and services, including tests difficult to be done in single nodes (like data transfers during analysis)
  - Incremental registration of new nodes into the Operational Environment, as per SRCNet successful validation



# Development and Operations for the SRCNet

- *Development and Testing*
  - Developers create (including inputs from scientists) and test new software modules and services within the Dev Environment
- *Integration and Deployment*
  - Once a module or service is stable and reliable, it undergoes an integration process. This involves deployment into Integration Nodes (possibly including new data)
- *Deployment to Ops Environment*
  - Successfully tested modules and services are then integrated into the *Ops Environment*, enhancing its overall capabilities
  - Integration nodes that consistently meet operational standards might be integrated into the *Ops Environment*



## Scenario 3: Overall

- This scenario enables inclusivity of nodes with a proper balance of mitigation of technical challenges
- Development and operations are closely related
- Organic incremental registration ensures confidence on all the phases of the deployment
- Reduce inclusion gap between versions





# Challenges of Scenario 3

- A preliminary decision of the first deployment nodes needs to be done (although this decision is less relevant as new nodes could be registered organically between versions). E.g., the evaluation of the nodes and the decision of adding new ones to the operational environment could be done at the level of PI planning
- As data from the integration and operational environment could be different
  - Either data redistribution or flag at storage level
  - This is not particularly important for SRCNet v0.1



# Summary of Scenarios

- After analysis of the scenarios, I (as architect) consider that scenario 3 is the one more powerful, enables testing of new software (no stopping development) and guarantees a stable/maintained/operational system for testing
- Although the scale of the resources is quite demanding at this time, it could be rescaled by the analysis of the data available and use cases to be tested



# Deployment/Implementation plan relevant dates

- Agree a scenario (today)
- Confirm nodes that are definitely for the first phase (today)
- Discussion at Architecture Forum (present plan and agreements to partners not present in Shanghai) and selection of collaborators to write detailed implementation plan (26 March)
- Beginning of April, distribute deployment/implementation plan



# QUESTIONS DURING NEXT SESSION

- Scenario preferred
  - Any variation?
- Motivation to be part of v0.1
- Impact on current contributions
- Constraints
  - Sites and Implementation Complexity
  - Internal policies
  - Software Stack
  - Procurements
  - ....



**Thanks for your attention**

