SKA Regional Centre development in Spain

Fortalecimiento y actualización de la infraestructura informática para el nodo español del SKA Regional Centre

S. Sánchez-Expósito, L. Verdes-Montenegro, A. Alberdi, J. Garrido, L. Darriba, M. Parra-Royón, J. Moldón, J. Sánchez, M.A. Mendoza

Instituto de Astrofísica de Andalucía - CSIC

AstroHEP-PPCC24

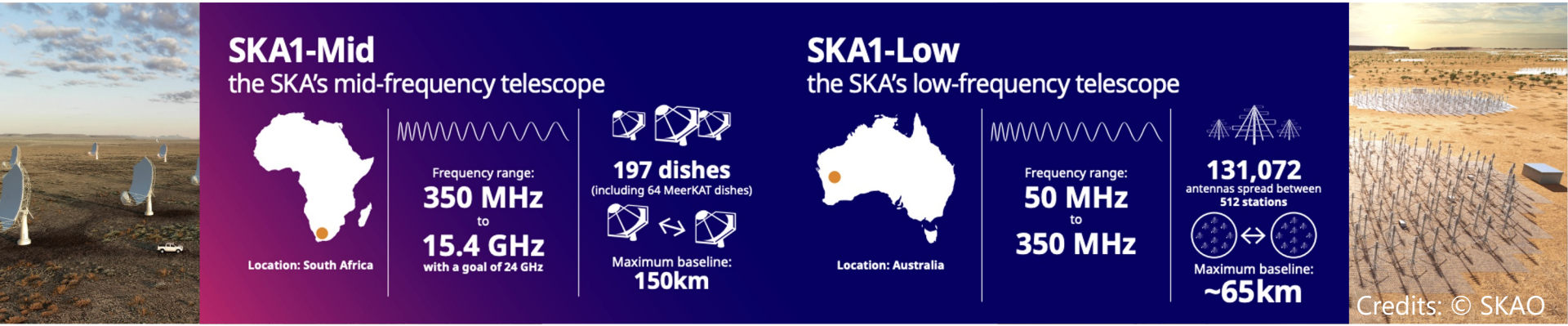
5-7 Junio 2024, Benasque

Table of content

The SKA Observatory

Open key questions in Astrophysics, Astrobiology and Fundamental Physics

- Formation of the 1st galaxies in a dark Universe dominated by atomic gas
- Evolution of the atomic gas and star formation till the current epoch
- Strong field tests of gravity using pulsars black holes
- Active Galactic Nuclei and the Galactic Centre
- Extrasolar planets (proto-planetary disks, biomarkers)



Spanish participation in the SKA

- Spain's involvement in SKA dates back to 1990
- **2011**: Ministry applies for Spain to become a SKA Observer.
- **2018**: Spain became a Member of the SKA *Organisation.*
- **2023**: Spain became a Member of the SKA *Observatory*.

IAA INSTITUTO DE ASTROFÍSICA DE ANDALUCÍA

Coordinator of the Spanish scientific and technological participation in SKA **since 2011**

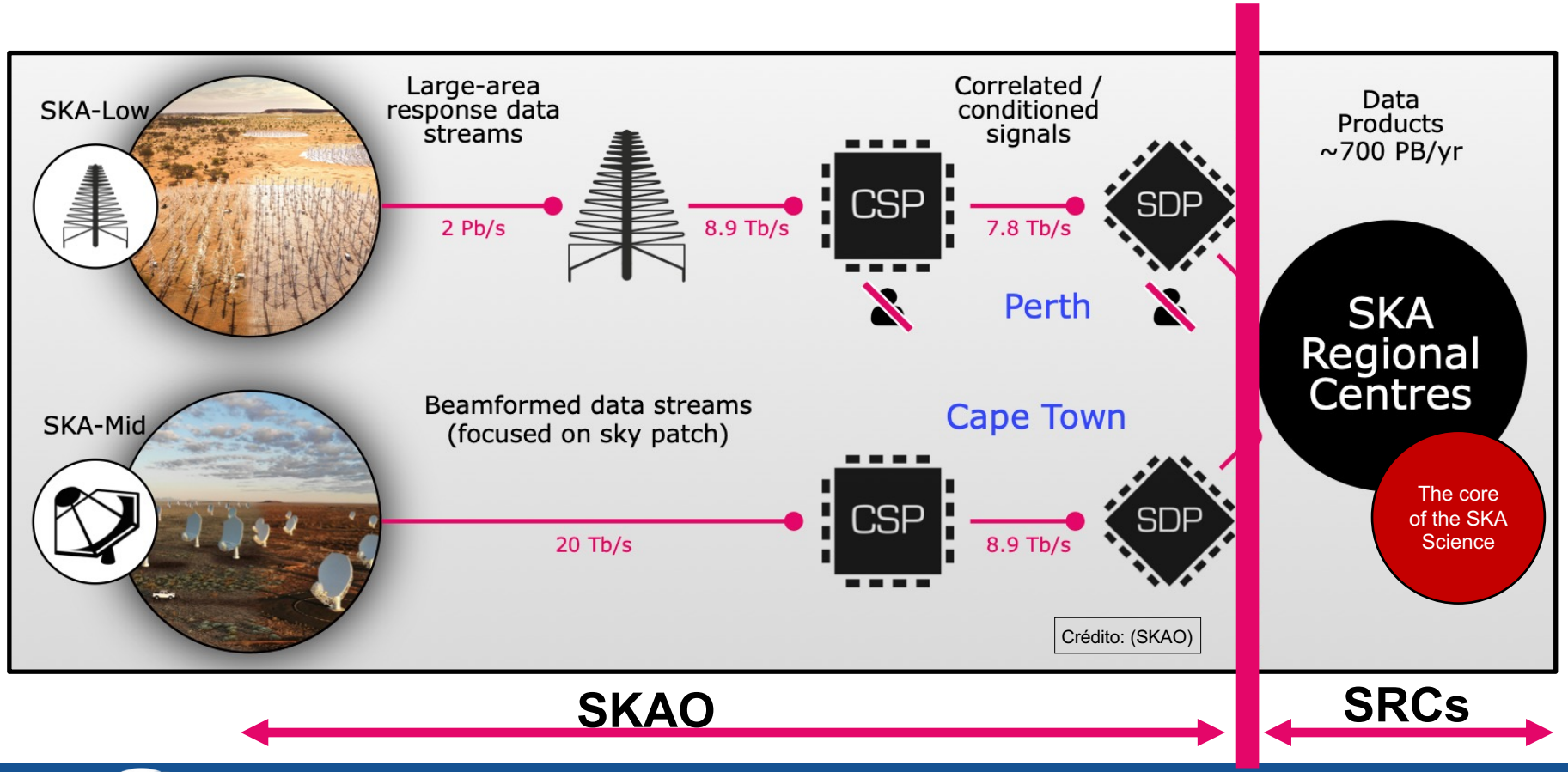
65 Spanish members (from 18 institutions) in the SKAO Science Working Groups

- **2014 - SKA Science Book** (10% of chapters with Spanish contribution)
- **2015 - Spanish SKA White Book** (120 researchers from 40 centres)

12 centres and 12 companies in the Design Consortia:

- Disk Antennas
- Central Signal Processor
- Signal and Data Transport
- Telescope Manager
- Science Data Processor
- Infrastructure Australia
- Infrastructure South
- Africa and Phased Array Feeds

The SKA Regional Centre Network (SRCNet)



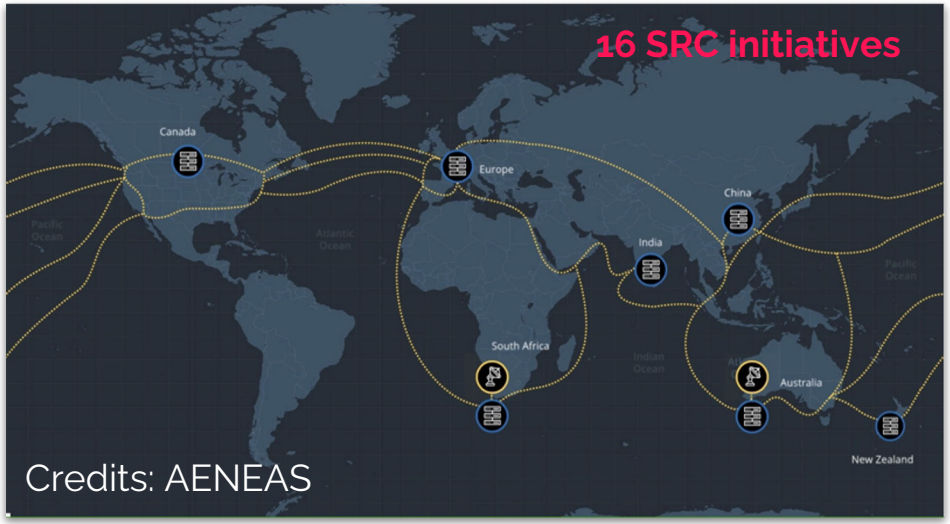
The SKA Regional Centre Network (SRCNet)

Main functions:

- Host the SKAO Science Archive
- Provide access to SKAO data for authorized users
- Provide computing and storage resources required for data processing
- User support and community training

SRCNet Technological challenges

Big Data, e-Science, Data management, Remote visualisation, Science Analysis Algorithms, AI, End-to-End reproducibility, Green computing

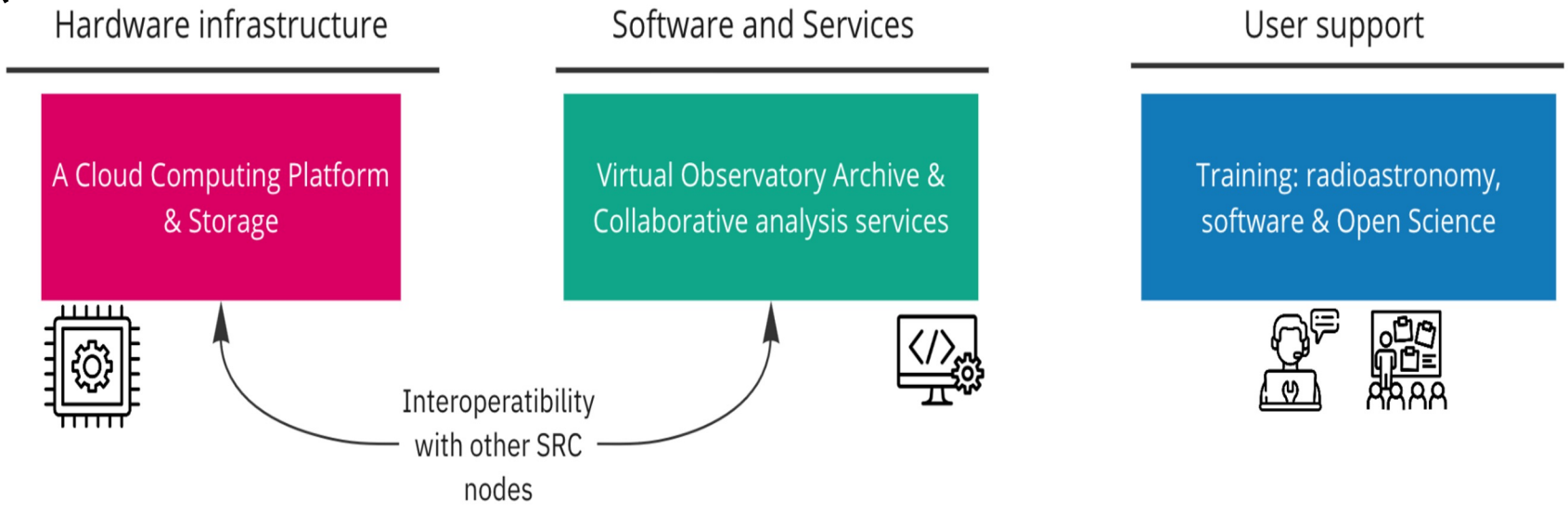


The Spanish SRC prototype: **esp**SRC

EXCELENCIA SEVERO OCHOA

Beyond a computing cluster

Started in 2018



Started in 2018 with the funding support from the **Severo Ochoa Grant**. The espSRC is also funded by other grants from **MICIU, the Junta de Andalucía, the European Regional Development Funds (ERDF) and by the European Union NextGenerationEU/PRTR.**

The Spanish SRC prototype: **esp**SRC

Beyond a computing cluster

- Enable the community to extract the utmost scientific value from the SKAO
- Maximize participation in SKAO Key Science Projects
- Facilitates preparatory SKA science with SKAO precursors and pathfinders telescopes
- Support the community to acquire the scientific and technical skills for the new SKA paradigm and for the Open Science
- Reinforce Multifrequency / Multimessenger synergies

Started in 2018 with the funding support from the **Severo Ochoa Grant**. The espSRC is also funded by other grants from **MICIU, the Junta de Andalucía, the European Regional Development Funds (ERDF) and by the European Union NextGenerationEU/PRTR.**

The **esp**SRC computing infrastructure

Credits: © AMIGA Group

Computing nodes: **Dell PE R640**

240 CPU Cores

6 nodes with 2 Intel Xeon 6230

* 5 with > 9GB RAM/core
* 1 with 1 TBytes of RAM

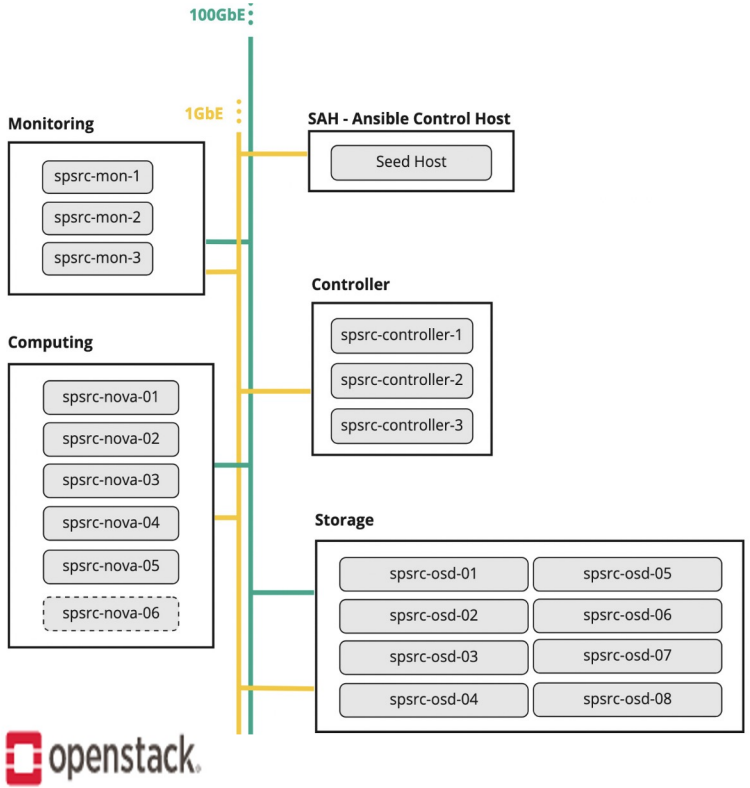
Storage nodes: **Dell PE R740 XD**

Credits: © Dell

8 nodes, ~ **1.5 PBs of raw storage**

* 4 with 32 x 7.68TB SSD 12Gbps
* 4 with 32 x 3.68 TB SSD 12 Gbps

The **esp**SRC computing infrastructure



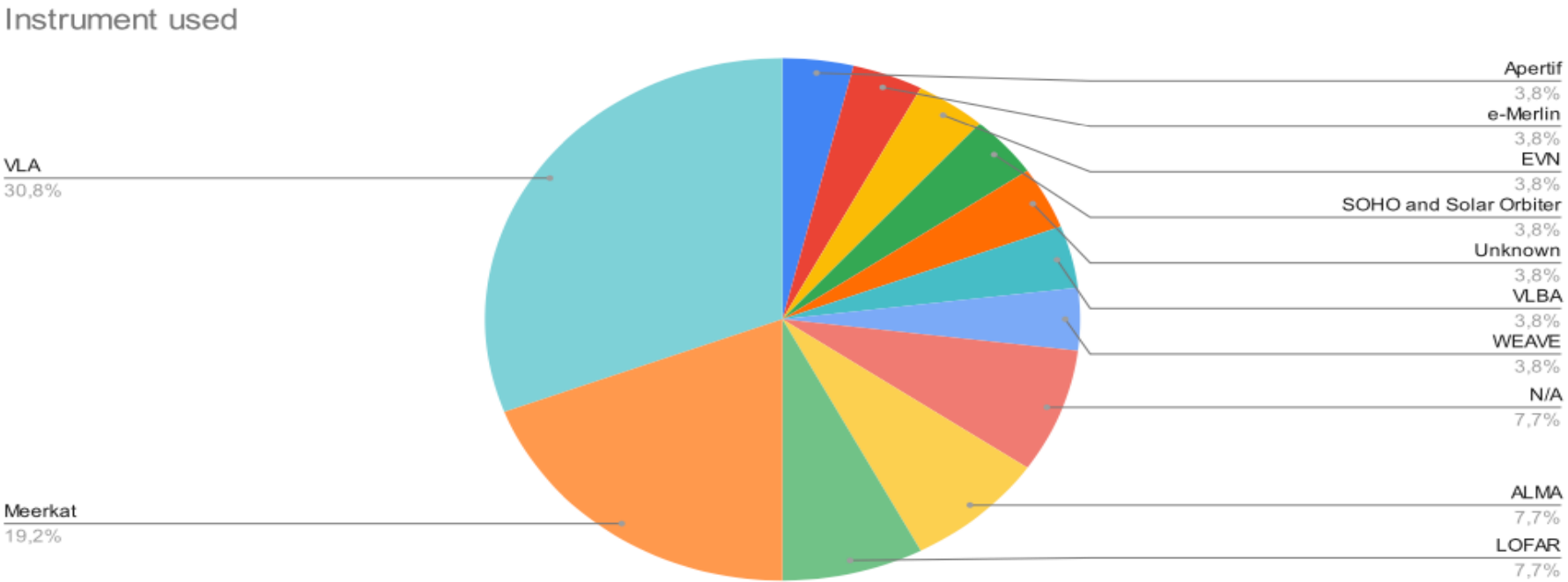
- **On demand Virtual Machines**
- **Kubernetes Clusters and Services Hub**
  - DaskHub
  - JupyterHub
- **Batch processing with Slurm**
- **Container catalogue for software deployment**

ANTELOPE

slurm workload manager

Science at the **esp**SRC with SKA Precursors/pathfinders

60 projects: 23 Research, 25 Development, 12 Training



Science at the **esp**SRC with SKA Precursors/pathfinders

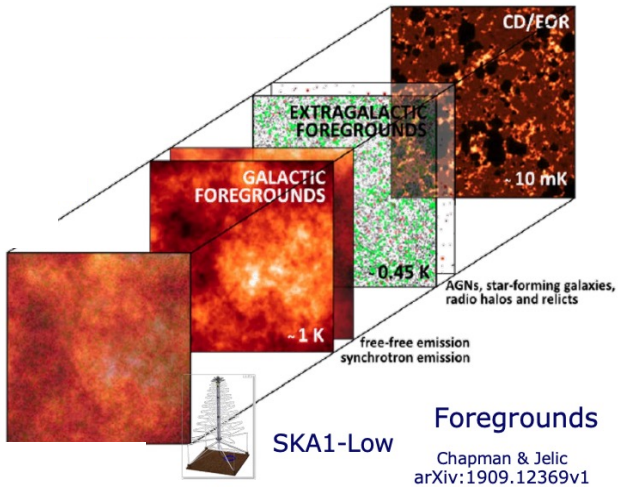
SKA Data Challenge 2:

Spanish (IAA) team: 5th place + Gold Medal on Reproducibility

SKA Data Challenge 3a: EoR

- 33 teams. Spanish one led by IFCA
- espSRC & CESGA supported teams from Spain, South Korea and India

SKA Data Challenge 3b: EoR



Contribution to the SRCNet Development

SRCCG, SRC Coordination Group (2016-2018)

→ L.Verdes-Montenegro (IAA-CSIC) invited as external advisor

SRCSC, SRC Steering Committee (2019-2023)

→ L.Verdes-Montenegro (IAA-CSIC) Spanish representative, designated by the Ministry

SRCSC Working Groups (2020-2022) – Design phase

- WG0 - Architecture
- WG1 - Data Logistics
- WG2 - Operations
- WG3 - Software Processing/Work Flow
- WG4 - Science Archive/VO/FAIR
- WG5 - Compute
- WG6 - Science user engagement

Main outputs
SRC Network requirement gathering (206)
A high-level architecture

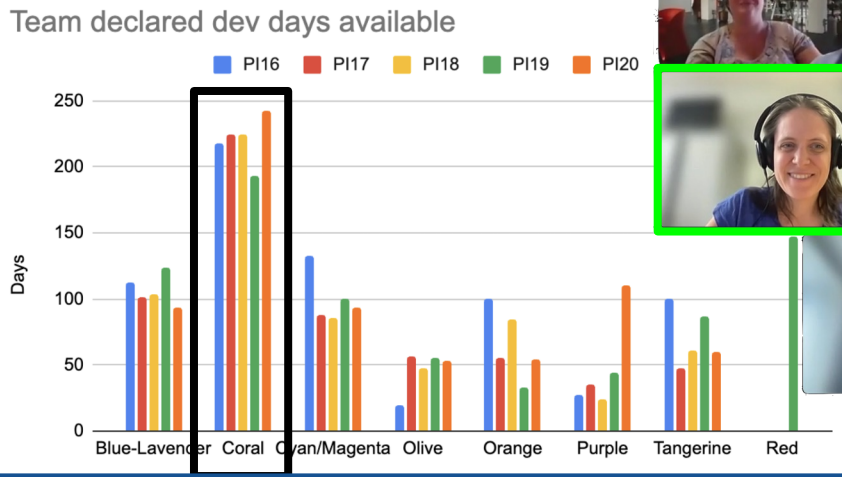
→IAA in the 7 WGs; BSC in WG3; IFCA, ICE, OAN in WG6

SKAO REGIONAL CENTRE NETWORK

Contribution to the SRCNet Development

CORAL Team @ SRCNet - Builder of **testbeds** for technologies

SRCNet prototyping phase
(Since April 2022)



The Mini-SRCNet demonstrator

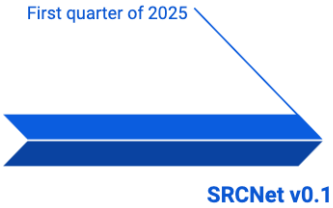
Mayor components

- Integrated A&A using prototype **IAM** service
- **Data management** and **dissemination,** and Science archive (OpenCADC Storage Inventory)
- **Science platform:** user container images, notebooks, visualization, desktop sessions (CANFAR)
- User **collaboration spaces** (mountable storage, User Interface, VOSpace implementation)

Canadian Astronomy Data Centre

Towards implementation phase: SRCNet v0.1

Roadmap Timeline



Milestone	Description	SRC Net Functionality	Scope (users)
SRCNet v0.1 First quarter of 2025	Opportunity to engage SRCNet with AA0.5 data transfer and access.	• Test data (and some precursors data) disseminated into a prototype SRC Net • Data can be discovered through queries to the SRC Net • Data dissemination to SRC nodes • Data can be accessed through a prototype data lake • Data replication. Data can be moved to a local SRC area where non-connected local interactive analysis portals (notebooks) could allow basic analysis • Unified Authentication System for all the SRCs • Visualisation of imaging data	SRC ART members Members of SKA Commissioning team

Source: J. Salgado. SRCNet v0.1 Development Workshop (Shanghai, 22nd March 2023)

From prototyping to implementation: SRCNet v0.1

espSRC among the **9** EoI sent to SKAO to participate into the SRCNet v0.1

EOI ANSWER

Requirement	Spanish Response
Software Stack	Compliance demonstrated during prototyping activities
Personnel contribution	Able to provide 1 FTE and respond in < 3 working days
Network	Shared 10Gbps link to RedIris-NOVA IPv6 is still not implemented 100Gbps internal network
Storage and computing resources	Procurement process to extend our platform and provide resources according to the Roadmap (0.42 PB storage capacity, 0.01 PFLOPS) • Funding is secured • End of procurement process (expected) : October 2024
SRCNet Node Validation Tests	Ability to run validation tests

Willingness to collaborate with other countries to define a common SRCNet deployment node.

IAA INSTITUTO DE ASTROFÍSICA DE ANDALUCÍA · EXCELENCIA SEVERO OCHOA · CSIC

Source: S. Sanchez. SRCNet v0.1 Development Workshop (Shanghai, 22nd March 2023)

Conclusions

SKAO will consist of two radio interferometers with capacity of making revolutionary contributions to Astrophysics, Astrobiology, and Fundamental Physics

The SRCNet is the place where SKAO data will be analysed by the community

The espSRC

- It is contributing to the SRCNet
- It is supporting the community with the preparatory SKA science
  - With a especial effort on Open Science
- PPCC will contribute, together with other funds, to make the espSRC one of the nodes of the SRCNet v0.1

Thanks

With the financial support from a) , b), c):

a)

b) The grant **CEX2021-001131-S** funded by MCIN/AEI/ 10.13039/501100011033

c) from the grant PID2021-123930OB-C21 funded by MICIU/AEI/ 10.13039/501100011033 and by ERDF/EU

References

[1] SKA Regional Centre Prototype at IAA-CSIC: building an Open Science platform based on cloud services. Sánchez-Expósito, S.; Luna, Sebastián; Garrido, Julián, et al. ADASS 2020. Nov 2020.

[2] Toward a Spanish SKA Regional Centre fully engaged with open science. Julián Garrido, Laura Darriba, Susana Sánchez-Expósito et al. Journal of Astronomical Telescopes, Instruments, and Systems, Vol. 8, Issue 1, 011004 (November 2021). https://doi.org/10.1117/1.JATIS.8.1.011004

[2] Asymmetric distribution of data products from WALLABY, an SKA precursor neutral hydrogen survey. Manuel Parra-Royón, Austin Shen, Tristan Reynolds, et al. ADASSXXI. Nov 2021. Preprint: https://arxiv.org/abs/2303.11670

[3] Cloud services at the Spanish SKA Regional Centre prototype (SPSRC) and what could be used by the SRCNet. Susana Sánchez Expósito, Lourdes Verdes-Montenegro, Julian Garrido, et al. SRC WG5 Workshop: Native Cloud Computing and Software Defined Clusters. Mar 2023.

[4] A fully-reproducible workflow for the SKA Data Challenge 2 HI-FRIENDS solution. Javier Moldón; Laura Darriba; Lourdes Verdes-Montenegro, et al. Reproducibility and Open Science in Astronomy, May 2022.

[5] A Spanish Prototype of a SKA Regional Centre. Javier Moldón. SEA Conference 2022, Sep 2022.

[6] Integration of storage endpoints into a Rucio data lake, as an activity to prototype a SKA Regional Centres Network. M Parra-Royón, J Sánchez-Castañeda, J Garrido, S Sánchez-Expósito et al.. ADASS XXII. Nov, 2022. Preprint: https://arxiv.org/abs/2303.07524

[7] Open Science and green computing to support sustainability at the SKA Observatory and its Regional Centres. Julián Garrido, Lourdes Verdes-Montenegro, Susana Sánchez. Astronomers for planet earth Symposium 2022. Nov, 2022.

[8] An approach to provide serverless scientific pipelines within the context of SKA. Manuel Parra-Royón (presenter), Carlos Ríos-Monje, Javier Moldón, et al. IWSG 2023, Tubingen, Germany (Jun 2023).

[9] Empowering SKA Data Challenges: A homogeneous platform for enhanced collaboration and scalability fully aligned with Open Science. Manuel Parra-Royón, James Collinson, Jesús Sánchez-Castañeda, Pablo Llopis-Sanmillan, Carolina Lindqvist, Susana Sánchez-Expósito, Laura Darriba and Lourdes Verdes-Montenegro. ADASSXXIII. Nov 2023.