

22 MARCH 2024 – REMOTE

# Italy possible contributions to SRCnet v0.1

ANDREA POSSENTI



# Italian expected outcome of the SRC network foundation



- ✓ 1. The identification of a kernel of “**modi operandi**” in the interactions among the various actors to secure an **efficient, persistent, and always developable** science-needs driven system
- ✓ 2. The establishment of a **SRC network with a significant node** located in Italy
- ✓ 3. The recognition of the local investments **in both hardware and human expertise, and its conversion into incentives as soon as possible**

# The Italian current contribution to the global effort

expertise

Fabio	Vitello	ITASRC	0.30
Giuseppe	Tudisco	ITASRC	0.50
Eva	Sciacca	ITASRC	0.15
Andrea	Lorenzani	ITASRC	0.20
Alessandra	Zanichelli	ITASRC	0.15
Vincenzo	Galluzzi	ITASRC	0.15
Gianluca	Marotta	ITASRC	0.20
Robert	Butora	ITASRC	0.40
Marco	Molinaro	ITASRC	0.10
Sara	Bertocco	ITASRC	0.10
Gianmarco	Maggio	ITASRC	0.10
Claudio	Gheller		0.10

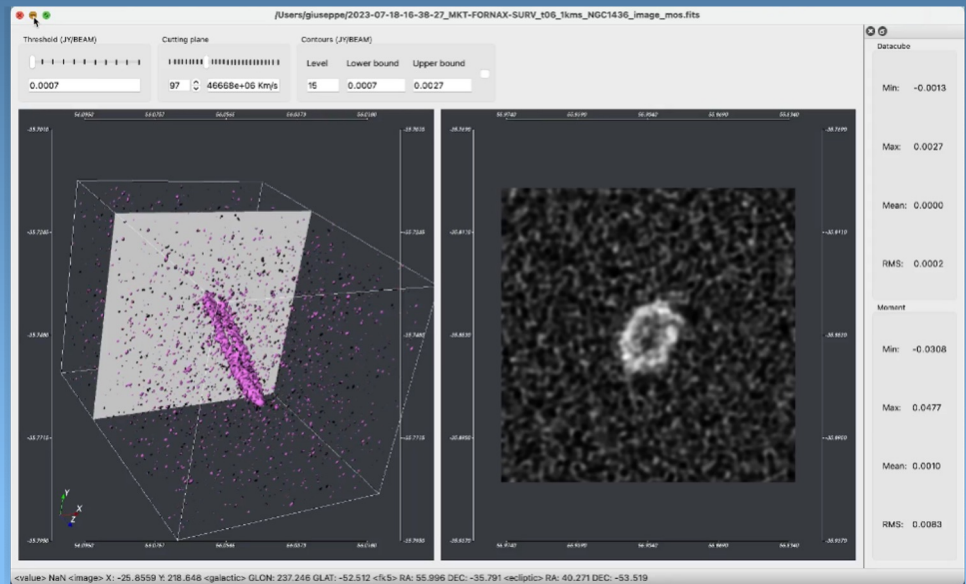
Anticipated Fractional FTE for Pl22 (13 March 24-14 June 24): total = 2.45 FTE

Most of this in the Orange Agile Team

# The work of the **Orange** Team

## Started with working on Prototype 4: Visualization in PI15 (June 2022)

- Contributing to the **definition** of visualization **use cases** for SRCNet
- **Visualization Tools review** (dependencies, interfaces, workshop)
- **Collection of data products** and data formats from **precursors and pathfinders**
- **Adapting Visualization Tools** to address use cases and work with SRC architecture and its data lake
- Development, testing and deployment of **SODA** (Server-side Operations for Data Access) into SRCNet, integrated with **Rucio Data Lake** and **Discovery services**
- Review of Solutions and Technologies for the **Computing Services API**
- Testing and **deployment** of visualization tools and data access services **into SRC nodes**



# The work of the Orange Team

## the next steps

- Further remote operations on data cubes by extending SODA service
- Make latest version of visualisation tools available via the SRCNet CANFAR science platforms
- Collection and classification of heterogeneous HiPS datasets for testing services and tools
- Enable visualisation based data exploration in the science gateway
- Adapt SODA so that it can process requests through the Computing API (Application Programming Interface)
- Adapt VisIVO to invoke cutout and visualize data through the Computing API (Application Programming Interface)

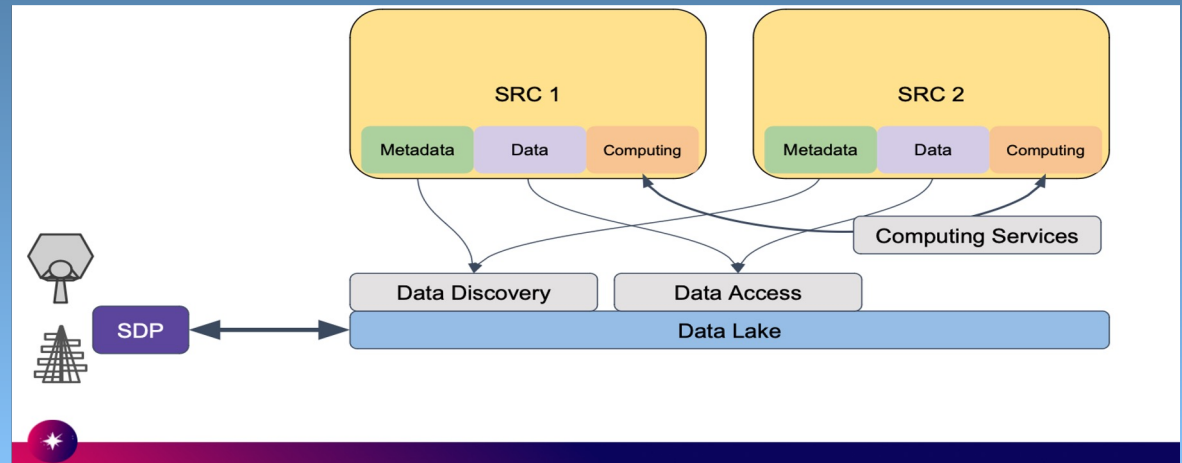


image credits: Jesús Salgado - SKA Regional Centre Architect

# The Italian contribution to the global effort

hardware: already funded assets which could be made available to SKARC 0.1

Acquisition of  $\approx 0.1$  PetaFlop/s (with combining CPU and GPU) and  $\approx 1.5$  PBy (combined between fast disks for computing and tapes for long-term preservation) Tier-3 computing system, to be installed inside of one of the CINECA areas at the Bologna Technopole



*The Technopole already hosts the European weather centre ECMWF, the Leonardo super-computer and will host the United Nations University on Climate Change*

**Towards** the deployment of **a dedicated Tier-2 sizing system** integrated into a Tier-1 sizing system and becoming the **kernel of the Italian node of the SKA Regional Center**. In addition, investment of the CN-PNRR for the needs of INAF and CNR owned by CINECA, with guaranteed (**non-exclusive**) use for INAF.

# Summary of SKARC.IT resources which could be made available for versions 0.1 and 0.2

Item	End 2024	End 2025	
<i>CPU power</i>	0.1 PF/s (Tier 3 – dedicated, CPU only)	1.5 PF/s (Tier 3 – dedicated, CPU+GPU)	15 PF/s (Tier 1 – shared, CPU+GPU)
<i>Storage</i>	0.3 PB on-line disk (S3), 1.2 PB Tape	2 PB on-line disk (S3), 5 PB Tape	10 PB Flash (LUSTRE, shared)
<i>Network</i>	10 Gb/s (LAN + WAN)	100 Gb/s (LAN) 100 Gb/s (WAN)	400 Gb/s (LAN) 100 Gb/s (WAN)

# The development of the INAF node

planned hardware on the longer term

(beyond 2027)



**2<sup>nd</sup> Step in the Global effort:** by mid 2028, a **Tier-2/Tier-1.5 dedicated** system with capability of  **$\approx 2$  Pflops** and  **$\approx 60$  PBy** of storage (20 PBy on-line and 40 PBy near on-line), connected at **100 GB/s** with the other nodes (v 1.0 of the SRCnet node)

**3<sup>rd</sup> Step in the Global effort:** by 2030, a **Tier-1 dedicated** size infrastructure with capability of  **$\approx 3+$  Pflops** and  **$\approx 80$  PBy/yr** of storage (30 PBy on-line and 50 PBy/yr near on-line), connected at **100 GB/s** with the other nodes (v 2.0 of the SRCnet node)



Technopolo - Bologna



# The INAF node: national effort

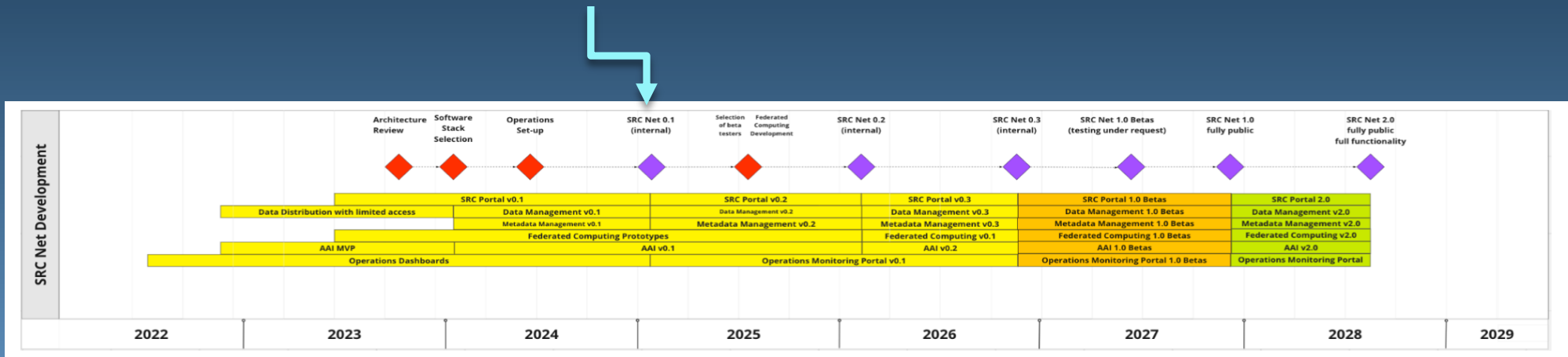
Fabio	Vitello	ITASRC	0.1
Giuseppe	Tudisco	ITASRC	0.1
Eva	Sciacca	ITASRC	0.1
Andrea	Lorenzani	ITASRC	0.1
Alessandra	Zanichelli	ITASRC	0.1
Vincenzo	Galluzzi	ITASRC	0.1
Gianluca	Marotta	ITASRC	0.1
Robert	Butora	ITASRC	
Marco	Molinaro	ITASRC	0.1
Sara	Bertocco	ITASRC	0.1
Gianmarco	Maggio	ITASRC	0.1
Claudio	Gheller		0.1
Cristina	Knapic		0.1
Giuliano	Taffoni		0.1
Matteo	Stagni		0.1

Anticipated Fractional FTE for the National Effort: Total 1.5 FTE

# Roadmap of the INAF node: the national effort

## SRC.IT prototypization

Objective: create a first prototype of Italian SKA Regional Center integrated in SRCnet v 0.1



### On-going tasks:

- Set-up of a data-lake based on object storage solutions (S3+CEPH)
  1. Deploy RUCIO on the top of the distributed repository
  2. Connect to the RUCIO based SRC data infrastructure
  3. Experiment VLKB solution for metadata management

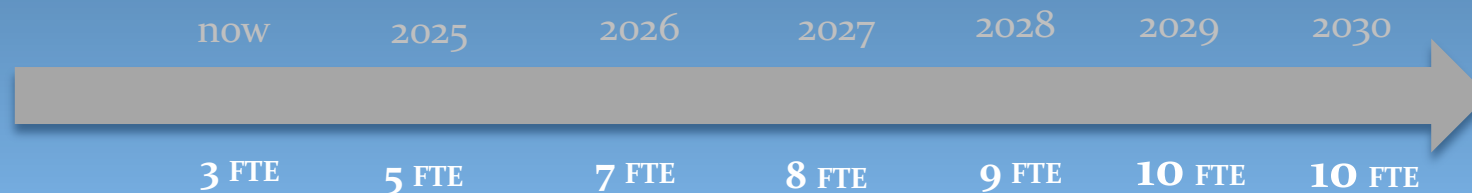
- Deployment of computing services for on site data-local analysis and visualization
  1. Access through the Computing API
  2. Analysis of SKA supported AAI approaches
  3. Remote visualization

# Global and national effort for the INAF node

## the need of personnel



- The aim is to devote at least 6 FTE to the global SRC network at regime
- To run the Italian node of the SRC network, 3-4 additional FTE will be needed at regime



Thank you!



Thank you!