The extreme computing and archiving needs of modern radio astronomy: an innovative ground for worldwide collaborations

Andrea Possenti

Osservatorio Astronomico di Cagliari

remote – 11 July 2025

The role of «*Computing*» in astrophysics

The need to process a large amount of data and to have advanced software is becoming **everyday life** even **for astrophysicists**

Simulations and data analysis for the exploitation of modern instruments in **space and on the ground** require increasing computing and data storage capacities (e.g. Gaia, Euclid, CTA, Vera Rubin, etc.)

Controlling and monitoring instruments and experiments also requires advanced computation and software developments

As an extreme case, there is a sector of astrophysics - **radio astronomy** - which requires **calculation capabilities** and **management of large amounts of data** that have always been equal to or **greater** than those of **almost all scientific disciplines**

Ingredients for successful Radio Astronomy

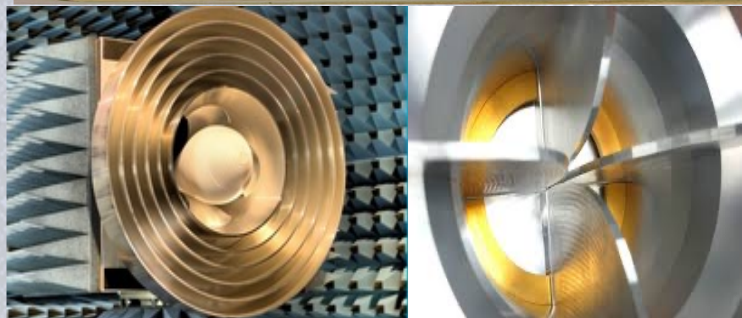
Collecting Area of the telescope

@ FAST

Field of View of the telescope

@ CHIME

Receivers' Large Bandwidth



@ Atnf,MPIfR

DRAGNET @ LO;FAR

Computational power & Data Storage Capabilities

Computational power needed for analysis of very large TeraByte size datasets …

© courtesy Botteon

Filaments, halos, structures in the Galaxy Cluster Abell 2255

… or for Cosmological simulations…

Cosmological simulations are the best tool in astrophysics to ***connect real telescope observations*** *with the* ***past evolution of the Universe****, using a consistent physical picture:*

which were the physical status of the matter billions of years ago and the mechanisms that led to shape the galaxies, including our own Milky Way ?

© courtesy Vazza

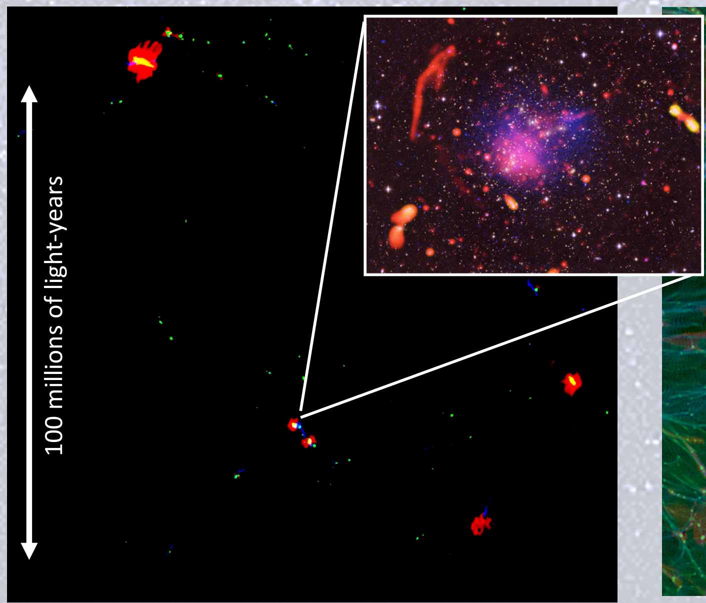
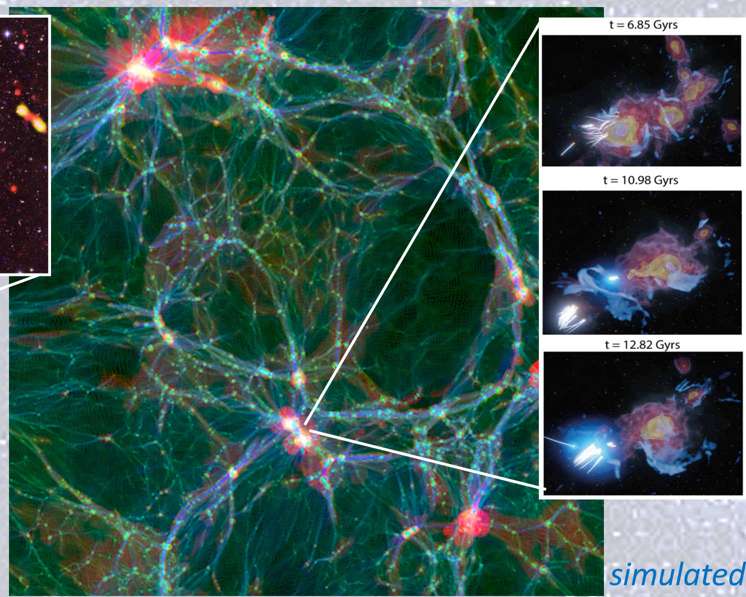


image of a galaxy collision from observed radio data

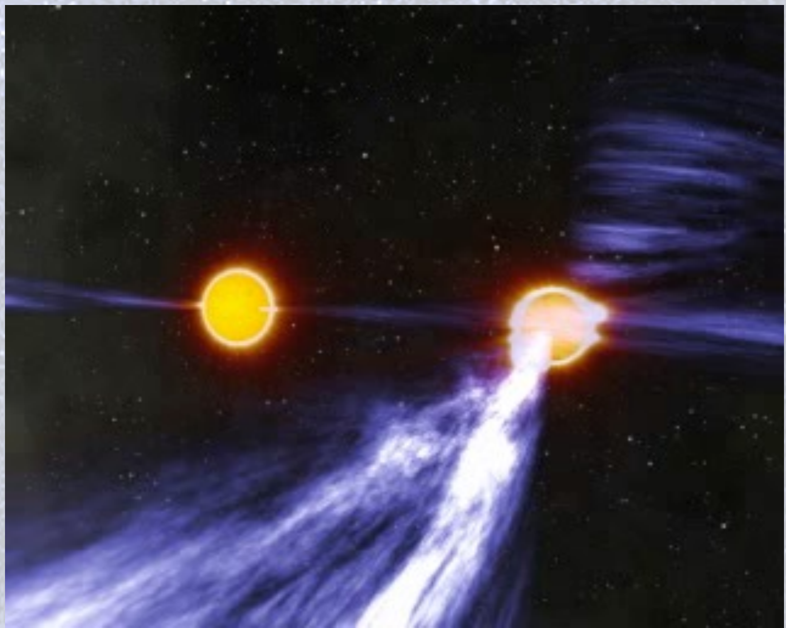


simulated collision

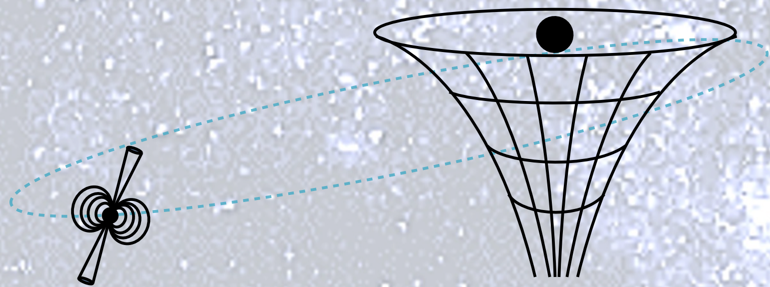
Simulations are memory bound and employ millions of core hours

… or for searching for a pulsar orbiting a black-hole

Binary pulsars in close orbital systems are magnificent tools to perform unique studies of Gravity Theories, mass of a Neutron Star and a Black Hole, nuclear Physics, Plasma Physics.
Finding a Pulsar orbiting a Black-Hole would allow to address with unprecedented detail the question:
is the Einstein's general Relativity theory able to also describe the Black Holes?



Cartoon of the Double Pulsar system J0737-3037. It comprises two neutron stars and it is at the moment the best testbed for the Einstein's radiative theories



[courtesy: A. Ridolfi]

Direct application of FFTs is not possible due to the changing apparent spin period caused by the Doppler effect

… in fact… the exciting perspectives of a PSR+BH …

FINDING AND TIMING A PSR-BH BINARY (AND MAYBE A PSR-MSP BINARY IN A GLOBULAR CLUSTER [Clausen et al. 2014], [NGC1851E, Barr et al. 2024])

From the ordinary PK parameters — *BH mass M with precision < 0.1%*

From precessional effects on semi-major axis and longitude of periastron — *BH spin S with precision < 1%*

From M & S — $(c/G)\ (S / M^2) \leq 1$ — *Test of the* Cosmic Censorship Conjecture [Penrose 1969]

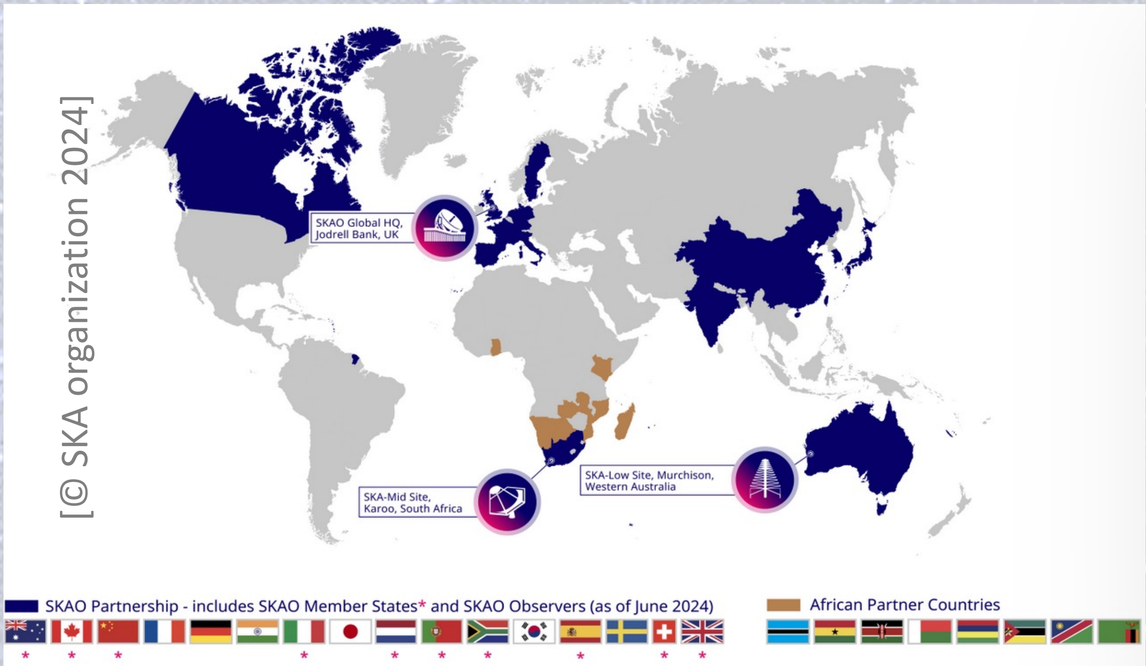
FINDING AND TIMING A PSR CLOSELY ORBITING SGR A*

From only 1 PK parameter — *BH mass M with precision < 0.001%*

From BH oblateness — *BH quadrupole moment Q with precision ~ 1%*

From M & Q — $Q = -\ (1 / c^2)\ (S^2 / M)$ — *Test of the* No Hair theorem

Behind the corner a new huge step in computing an data handling …. the SKA project

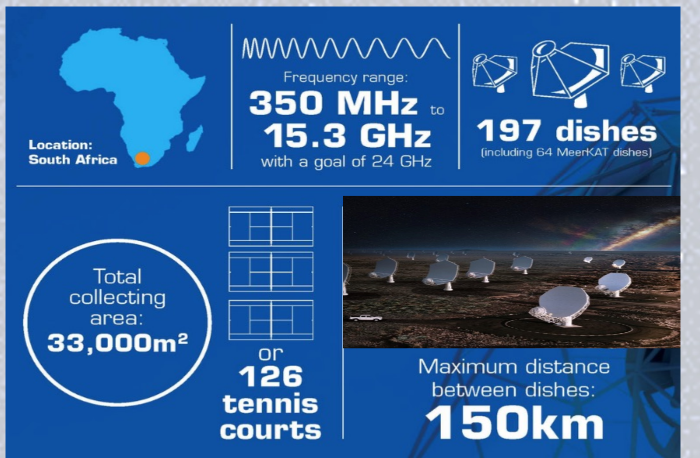


Signature for the IGO (inter-governmental organization) responsible for delivering the construction and operation of the SKA occurred on 12th March 2019 here in Rome

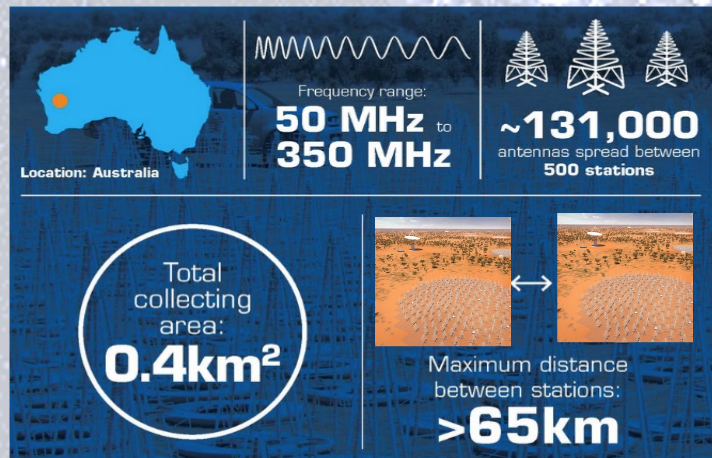
5 countries, among which **Italy and China**

SKAO in summary

Staged construction: SKA-LOW AA* by 2029, SKA-MID AA* by 2031, full deployment 2035?



SKAO



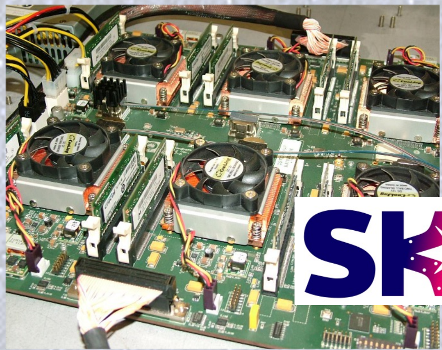
The computing challenges of SKAO

Element	SKA ~ 2030
Dishes, feeds, receivers	~200
Aperture arrays	~130,000
Signal transport	~1 Pb/s
Signal processing	~exa-MACs
High performance computing	~100s **peta-flops**
Data storage	Exa-byte capacity
Power requirements	~10MW

The SKA Observatory will become more and more powerful with the improvement over the years/decades of the computing capabilities

… not only on site: from the antennae to the Regional Centers

CSP: Central Signal Processor



e.g. FPGAs in the ASKAP correlator

5 + 9 Tb/s data buffer of 2 minutes

Adapted from Philippa Hartley (SKAO)

5 Tb/s data buffer of 2 weeks

SDP: Science Data Processor

e.g. SDP prototype, Cambridge

SRC net

SKAO regional centre network

SRC

SRC

SRC

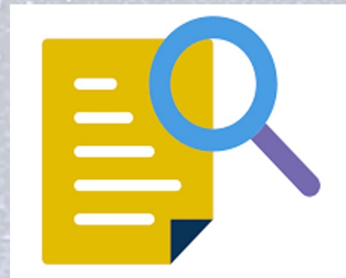
SRC

SRC

SKA portal

Distributed facilities

USERS



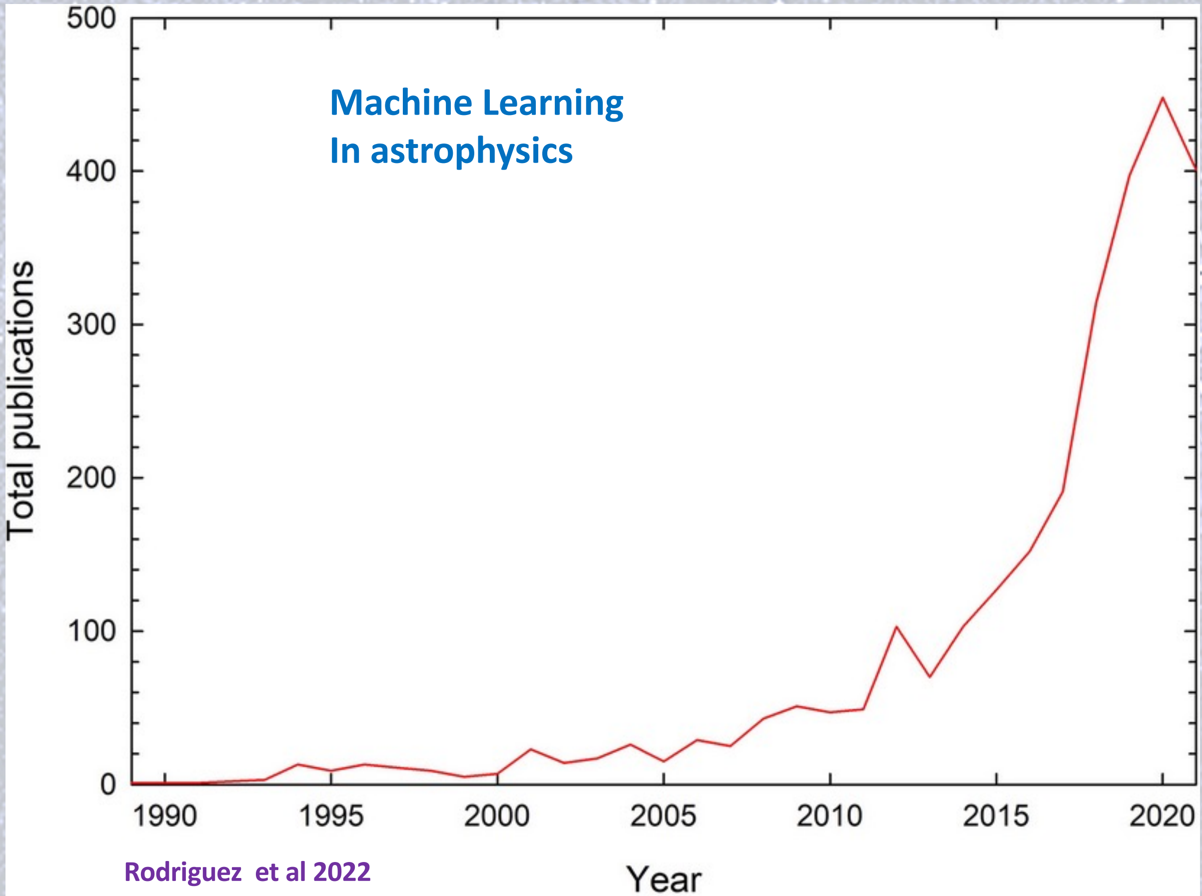
SKAO will generate an **outbound data flow of approximately 200 Gb/s**, with **10+/100+ TB datasets** each to calibrate and analyze, and within next 7 years a total of approximately **320 PB of data per year**

All of that… possible thanks to an unprecedented role of ITC, ML & AI

Visualization, classification, identification of anomalies, inference are key ingredients for extracting the best of the science from antennae, network and receivers

Machine Learning & AI

Machine Learning techniques take advantage of increasingly powerful GPUs and with ever more memory

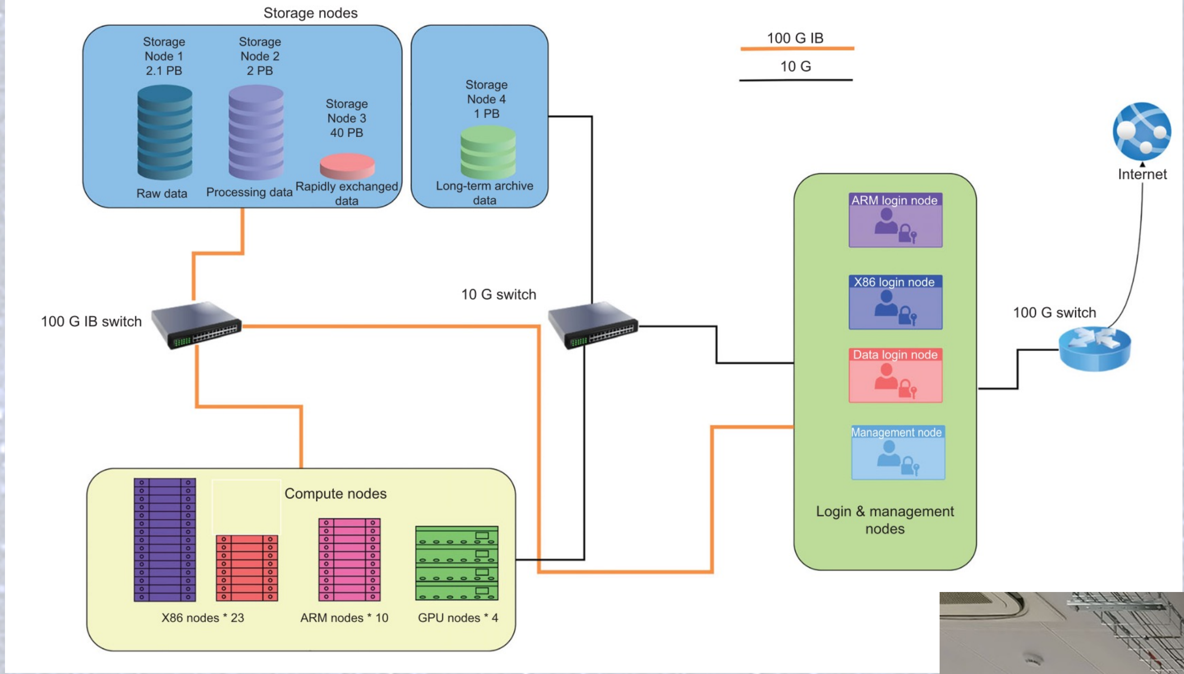


The incoming SKA Regional Center (SRC) of INAF (2026)

Acquisition of ≈**1.5 PetaFlops** (with a combination CPU and GPU) **and ≈11 PBy** (combined between fast disks for computing and tapes for long-term preservation) computing system, to be installed inside of one of the CINECA areas at **the Bologna Technopole**

The Technopole is already hosting the European weather centre ECMWF

The already set up Chinese SRC node



Italian & China involvement

① **Italy and China are two of the 5 initial partners**

② **Italy and China are each hosting one of the first nine nodes of the SKA Regional Data Center network**

Thanks