AENEAS All-hands meeting

INAF Infrastructures towards the SKA RC

HTC, HPC, Cloud and new INAF perspectives. Towards the new INAF DATA-STAR Project

U. Becciani, INAF – OACT, C. Knapic INAF- OATS

11-14 November 2019

Utrecht, the Netherlands

INAF Computing Infrastructure

OUTLINES

INAF Tier2-Tier3 Infrastructure
- The CHIPP Project
- The main projects

INAF @ HPC - Cineca Partnership
- MoU
- Tier0 and Tier1 systems
- Dedicated support
- HPC - Cloud Computing

TODAY

INAF Cloud: initial test and and future plans
- Preliminary test on commercial cloud (POCs)
- EGI cloud
- EOSC projects

INAF and other dedicate infrastructures
- LOFAR Computing Infrastructure
- Gaia data archiving and computing
- SRT (future)
- … others

INAF - DATA-STAR PROJECT

The new INAF Infrastructure for Big Data

FUTURE
Big Data era

EUROHPC: *Leonardo*

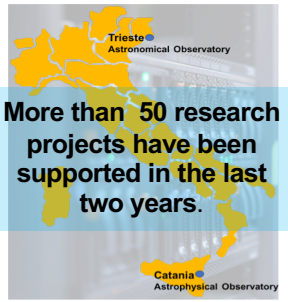
The future PreExascale System

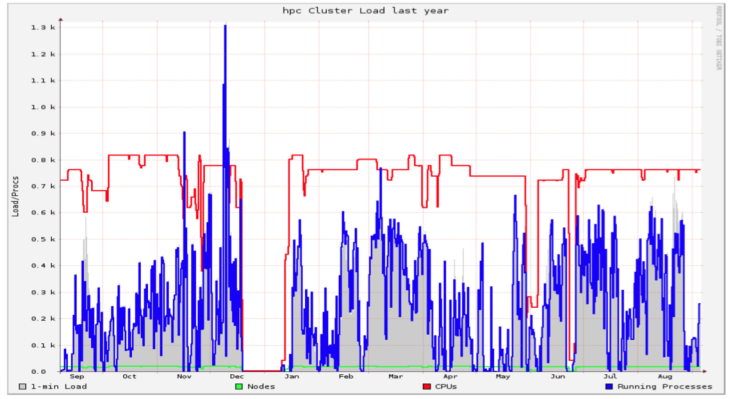
INAF - The CHIPP Project Tier2 – Tier3 Infrastructure

The main purpose of INAF CHIPP project is to **provide HTC and HPC resources (for small/medium size programs) to the INAF community** using the already existing infrastructures. **Period 2017-2020**.

CHIPP Main system INAF Trieste – HOTCAT

Computing nodes: 40 Core INTEL Haswell E5-4627v3 @ 2.60GHz (4 SOCKET); 6GB RAM/Core (256GB RAM)
Total node number: 20 (800 computing cores available) Global RAM 5.1 TB.
Storage: 250TB , 3 I/O nodes:paralklel filesystem based on BeeGFS.
Network: Infiniband ConnectX®-3 Pro Dual QSFP+ 54Gbs
Usability : 40% CHIPP dedicated





CHIPP Main system INAF Catania – MUP

Computing nodes: 12 Core (24 *Hyper-Threading*) Intel® Xeon® E5-2620; 5.2GB RAM/Core (64GB RAM).
Total node number: 16 (192 computing cores available) Global RAM 1 TB
Storage: 60 TB parallel filesystem based on BeeGFS to be implemented.
Network: 10 Gbit network
Usability : CHIPP dedicated

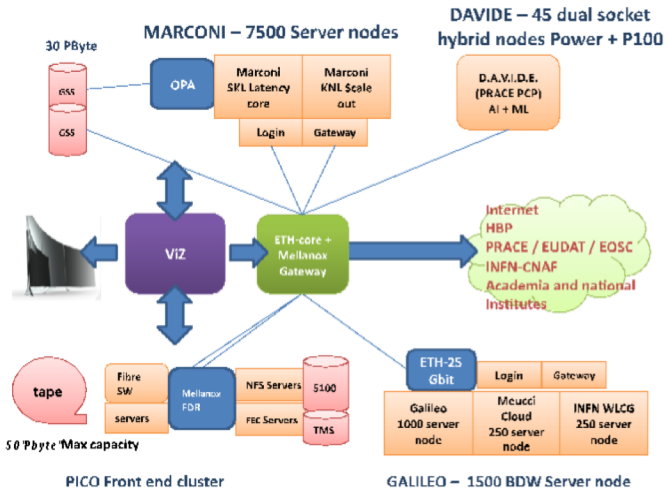
INAF @ HPC: MoU with CINECA for HPC resources

MoU/Framework INAF - Cineca.

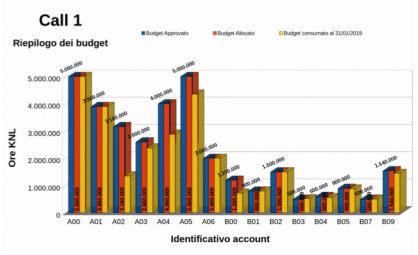
From 2017 to 2020 : reserved to INAF up to **50 Millions cpu/core hours each year computed on Marconi KNL** but usable on all systsems of HPC open to the researchers. **150 TB work space**

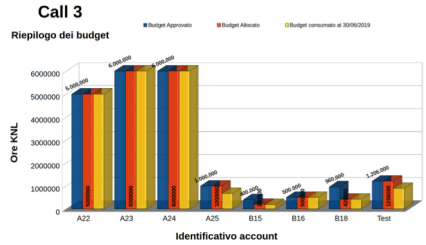
Dedicated support for all INAF groups that needs to optimize codes.

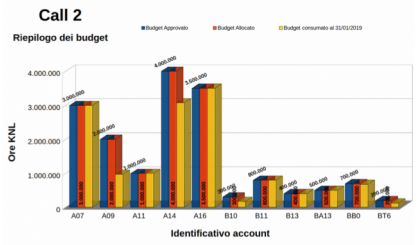
5 CALL FOR PROPOSAL SINCE 2017 - Aprox. ***50 approved projects (4 calls)***, more than **150 researcher involved** in the projects

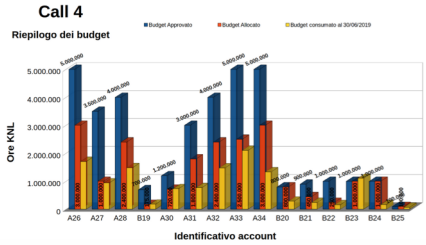


- Global peak performance in excess of 25 Pflops
- Global production capacity > 9000 servers nodes
- Modularity computing combining different microprocessors technology based on the same ISA and different SoT ware stack enabling effec, ve produc, on workflows
- Architecture development driven by the scien, fic challenges and applica, ons domains









INAF Pilot Projects for Commercial Cloud applications

Google Cloud: Proof of Concepts (PoCs)

Evaluate how a commercial cloud solution can be an effective solution for different classes of computational tasks

⇒ 6 use cases have been identified with the aim of measuring different metrics on Google Cloud Platform.

➔ HTC execution of embarrassingly parallel code **DIAMONDS** (Corsaro). The code can be used for any application involving Bayesian parameter estimation and/or model selection problems. Platform performed correctly by executing thousand of instances of the code parallelly. Instance duration 10 minutes

➔ **HPC** (Taffoni): Numerical simulations of gravitationally interacting particles of both dark matter and baryonic matter. We deployed a cluster with different machine types (see figure) to run the OpenMPI based code (GADGET). **Poor results in terms of scalability**.

➔ ***SKA Test*** (Sciacca.) Three different Use Cases tried successfully on the platform : LOFAR prefactor calibration pipeline has deployed using real LOFAR data for 40 frequencies. We used instances with 40 vCPUs and about 256 GB of RAM. **Scalability** of pipeline has been found **very good in function of the number of cores available** on each instance. Software has been ported to the platform using ***Singularity Containers***

INAF @ Cloud. SA-EU and EOSC new challenges

SA-EU Federated Cloud Pilot Project Meeting
EU - IDIA South Africa
CATANIA 10-11 September 2019

The purpose of this meeting was to launch the pilot demonstration project of a South African - European federated cloud to support collaboration on SKA Pathfinder data intensive large programs between South African and European research teams.

- Cloud architecture ➔ IAAS - OpenSatck (IDIA + EU)
- INAF dedicated infrastructure
- First Prototype Dec 2019
- Scientific WG has been setup for System Requirements
- Execution of Scientific usecase starting from 2020

LOFAR-IT Consortium. The distributed infrastructure

UNITO
3 FAT node on OCCAM
4 x Intel Xeon (12 core)
RAM 768 GB DDR4, 1 SSD 800GB,
1 HDD 2TB, 2x10Gb

OATs
4 nodes: Intel Xeon, 256 GB RAM DDR3 6 x 6TB SATA, Infiniband ConnectX®-3 - Hotcat system.
40 TB parallel FS
[Future] more nodes more storage more RAM

IRA BO
- [Actual] 3 nodes: Intel Xeon 384/512 GB RAM, 60 TB
- [future] 1 nodes: 2 socket (40 core) 512 GB RAM, 10gbit ethernet

OACT
- [Actual] 3 nodes: 256GB RAM, 40 core, 10Gbit, 20 TB Storage
- [future] 1 nodes: 512 GB RAM, 40 core, 10gbit 60 TB storage.

LOFAR - LTA

16 TB

500GB

Pre-factor
(calibration)

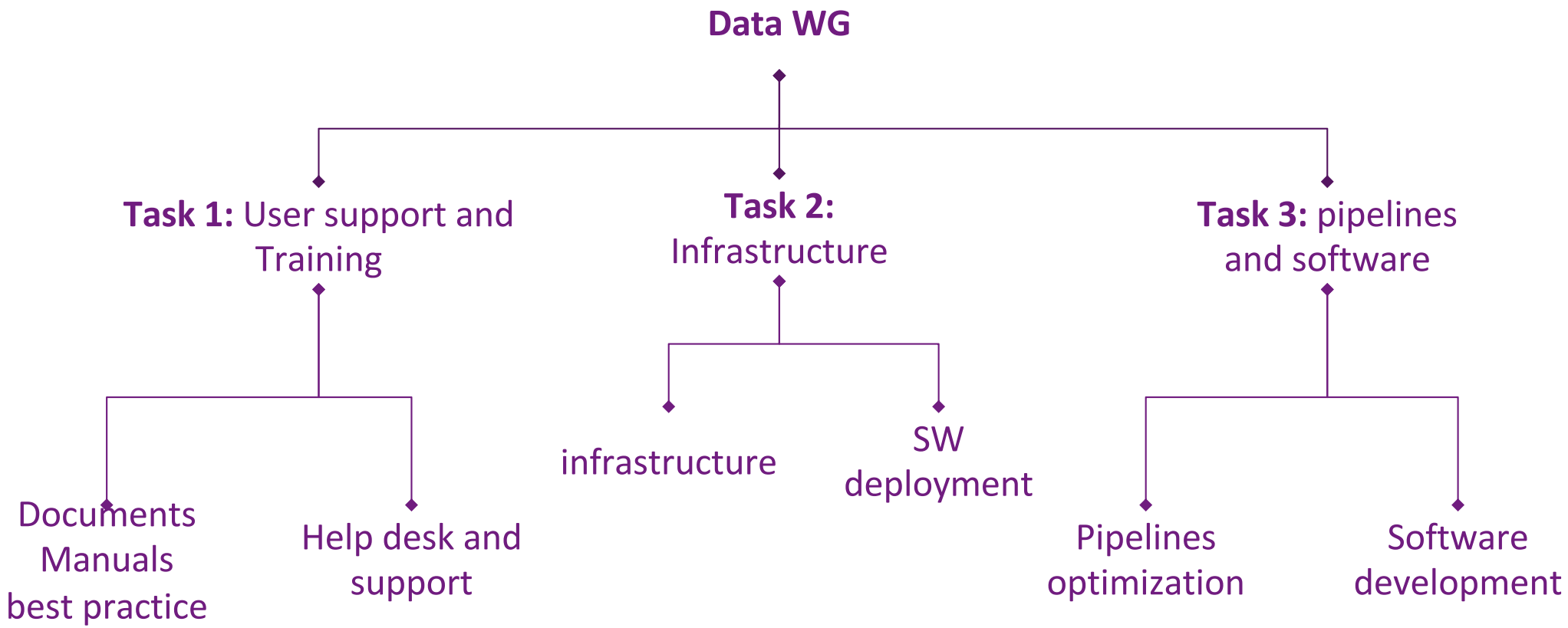
580GB

Initial sub
(sky-model subtraction from vis: FoV im. of the station)

Factor
(low-noise, hr wide-field images)

3 TB op. and **300 GB final**

LOFAR.IT: Working Group
New organization



LOFAR.IT: Data Working Group

- **provide the design** of the hardware and software infrastructure for calibration and **data reduction in Italian LOFAR nodes** and coordination of the infrastructure itself;

- coordinate the **installation, configuration and management** of specific software and pipelines for the reduction of LOFAR data;

- provide **technical support** to users belonging to LOFAR IT through testing, verification, **optimization and development of pipelines** for LOFAR data reduction;

- collaborate with LOFAR developers for further code testing and optimization/parallelization of codes and data reduction pipelines (e.g DDFacet pipeline);

INAF Exascale projects

INAF is one of the leading institutions participating to the design and prototyping of new **Exascale supercomputers** in Europe.

ExaNeSt European funded project (8Meuro) is developing, evaluating, and prototyping the physical platform and architectural solution for a unified Communication and Storage Interconnect, plus the physical rack and environmental structures required to deliver European Exascale Systems.
http://www.exanest.eu/

EuroExa funded project (20Meuro) is brings a holistic foundation from multiple European HPC projects and partners together with the industrial SME to co-design a ground-breaking platform capable of scaling peak performance to 400 PFLOP in a peak system power envelope of 30MW; over four times the performance at four times the energy efficiency of today's HPC platforms. Further, it targets a PUE parity rating of 1.0 through use of renewables and immersion-based cooling.

THE FUTURE INAF INFRASTRUCTURE: DATA STAR

INAF National Center for Computing and Big-Data in Astrophysics and Space Sciences

More than 2 Million Euros of initial funds are foreseen for INAF Technopole Infrastructure… et al

INAF will create a new HPC and Archive facility center with the core placed at the Technopole and few satellite infrastructures.

The ***Bologna Technopole*** represents one of the most important convergence of European, Italian and Regional investments in favour of supercomputing and its application. In fact, thanks to funds from Regione Emilia-Romagna, Italian government and intergovernmental Centre ECMWF, it is hosting one of the most significant community in Europe for supercomputing.

DATA-STAR will be the INAF National Center for Big Data in Astrophysics and Space Sciences. The main site will be hosted in Bologna Technopole in a reserved area for INAF. The infrastructure from 2021 will start to host Archiving and HPC facilities for the main challenges projects in INAF.

DATA STAR ***will host the Italian SKA Regional Center*** and ***<u>will include all the already existing facilities</u>*** (computing and archiving) and expertize: IRA – Bologna, OA Trieste, OA Catania, OA Cagliari

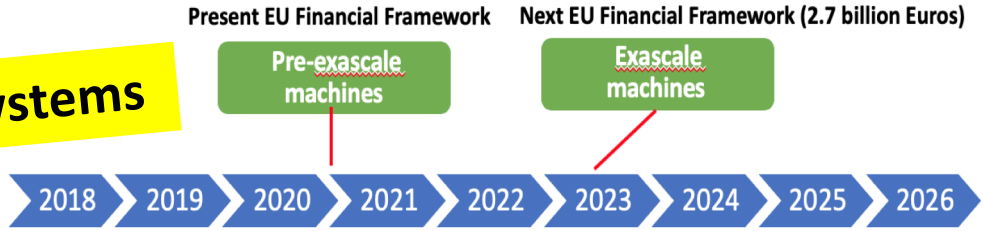
Integration of INAF existing facilities: IA2 and CHIPP

Ugo Becciani

Leonardo supercomputer: EuroHPC - JU

The EuroHPC Joint Undertaking, was approved by the EC in 2017. The main goal is to allow Europe to lead a pre-exascale HPC facility and the exascale phase. **The JU has a budget of about EUR 1 billion up to 2022 .**

Recently assigned the first 3 Pre-Exascale Systems



In Italy the pre-exascale machine named LEONARDO will be hosted by Cineca at the Bologna Technopole for an overall cost of 240 Million Euros.

In Spain the JU supports the **MareNostrum 5**, the future BSC's supercomputer.

The Finland supercomputer will be hosted by CSC in Kajaani, Finland and will be managed by LUMI (Large Unified Modern Infrastructure) consortium

Leonardo supercomputer: a 270 PFlops system

Notation	Description
Booster	Module of the system dedicated to capacity and capability workloads
Data-centric	Module of the system dedicated to high-memory workloads, data visualization and data management
General purpose	Module of the system dedicated to general workload, yet to be adapted to the booster module

INAF WILL PLAY A PRIMARY ROLE IN THE SYSTEM IMPLEMENTATION

System name	Leonardo
Modules	3 (booster, general purpose, data centric)
Number of computing nodes (booster)	3500 (4 accelerators per node)
Number of computing nodes (general purpose)	1000 (> 64 physical cores per node)
Number of computing nodes (data centric)	500 (512 GB DDR and >4TB NVM per node)
Storage (scratch and work space)	Capacity: 150 PB, bandwidth: 1 TB/s
Storage (high IOPS tier and home space)	Capacity: 5 PB, bandwidth: 1 TB/s
HPL Targeted Performance (peak)	150-180 PFlops (210-250 PFlops); Top 3
HPCG Targeted Performance	2.8-3.3 PFlops; Top 3
Interconnect Bandwidth	≥ 200 Gb/s per node
Interconnect Topology	Dragonfly+ or any topology with better full bisection bandwidth
Estimated Power consumption (after PUE)	8-9 MW (8.8-9.9 MW)

INAF has already planned the main challenges for the new system:

SKA Precursors (ASKAP, LOFAR, MEERKAT)

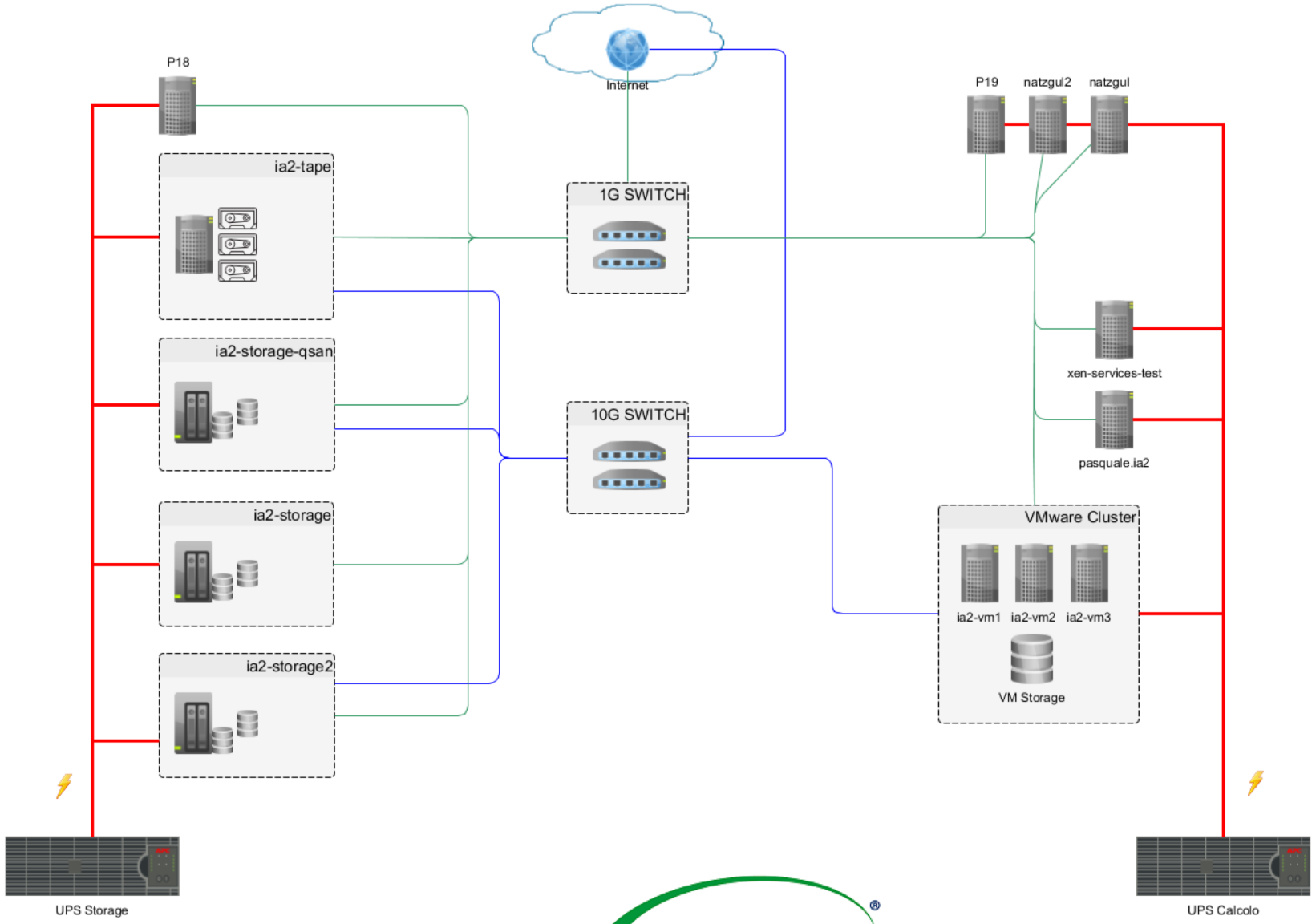
Ground Based and Spaces Missions Observatories (SpaceWeather, Euclid, E-ELT)

Numerical simulations (Black Holes and Primordial Universe, Primordial Galaxies and Gravitational Waves, Large Scale Structure of the Universe)

INAF ARCHIVES: the Italian Astronomical Archives (IA2) facility

Trieste Computational center

Facility Hardware Schema



Servers @OATs:

- 2 VMWare dedicated
  2x10 core - 512 GB

- 1 Xen Servers Citrix (testing)
  2x6 core - 128 GB

Server @ IRA:

- 1 Xen Servers Citrix (testing)
  2x6 core - 128 GB - 80TB

Server @ SRT:

- 1 Xen Servers Citrix (prod.)
  2x6 core - 128 GB - 80TB

IA2 storage capacity

- **Hardware IA2 @ TS**:
  - 800 TB (500 used + 300 free TB)
  - backup : 100 TB for VMs
  - T950 HPE LTO-8 of 1.25 PB expandible to 12.5 PB
  - 10Gb/s Fortinet firewall (in collaboration with OATs SSI - coming)

- **Hardware @ other sites**:
  - IRA : 60 TB on new machine (raid - testing)
  - IRA: 0.5 PB LTO-7 HP Tape Library (coming soon)
  - OACagliari SRT : 60 TB for Radio Distributed Archive (raid)

- **Hardware owned by others**:
  - IRA : 40 TB Radio Distributed Archive
  - SRT : 1 TB (pulsar testing machine)
  - Serra La Nave : 500 GB on site
  - LBT : 12 TB upgraded 1TB /y Full LBT Archive
  - Asiago : 500 GB on site

- **Bandwidth**: 10Gb/s GARR

IA2 services

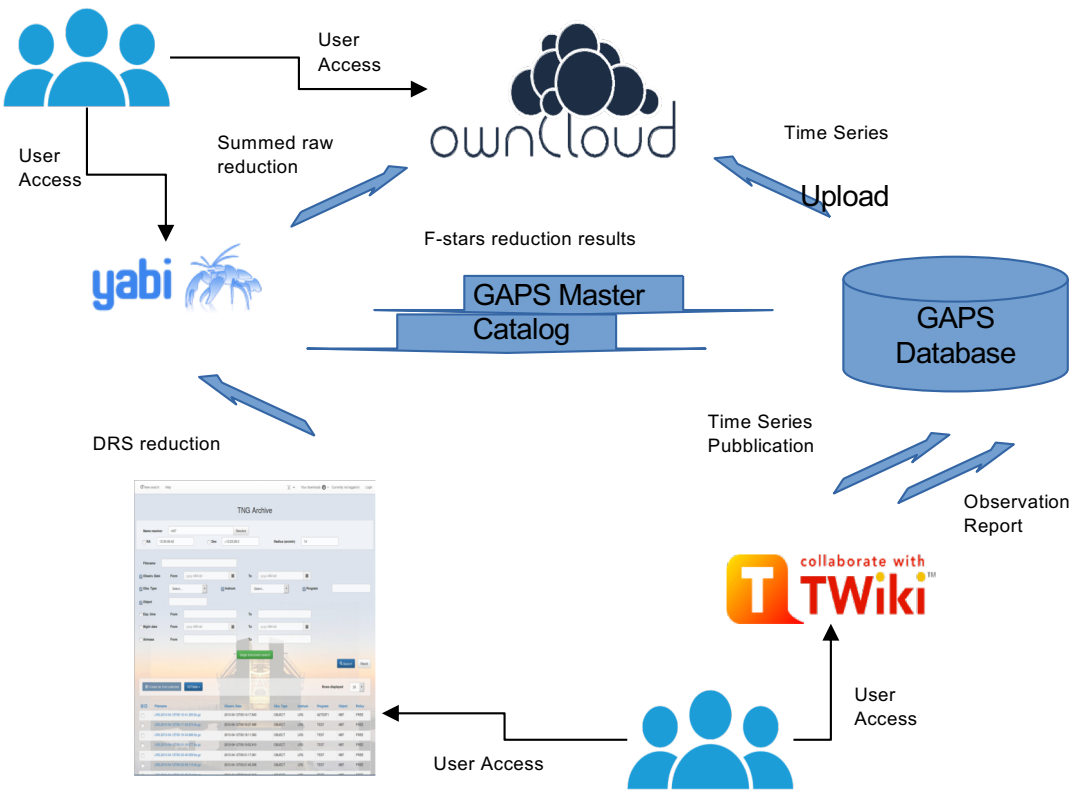
IA2 :

- ***manage INAF national and international telescope archives, projects and surveys;***
- ***hosts quite all the ICT software services (INDICO, OwnCloud, RedMine, ..);***
- ***provide support for data providers;***
- develops software for distributing data;
- hosts services for others;
- provide (limited) computing power for data reduction using workflow management system;
- ***publish data in VO compliant way (TAP services);***
- will soon publish VO services for images and spectra;
- ***is involved in several Italian, EU-H2020 and international Projects;***
- ***host the first release of the SKA SCIENCE DATA CHALLENGES;***
- in collaboration with IRA staff, CADC and OATs computing staff provide the VO compliant authentication and authorization system;

IA2 support to Science

- **Store and preserve** astronomical data (observed or simulated);
- **Support data providers** in correctly set up Archives:
  - from raw data to calibrated one from Telescopes;
  - simulated data (exoclimates, intrigoss, cosmological)
- **Publish public data** through the VO services;
- **Support Astronomers in data retrieval** via web interfaces:
  - search on public data without login;
  - after login, a list of proprietary datasets are presented and filters can be applied;
  - using filters to find data;
  - single file direct download;
  - user space for bunch of files;
  - possibility to download VOTables of queried data;
  - possibility to download a CSV file of queried data;
  - possibility to download a URL list of files to download them using external tools like WGET;
  - name resolver to find coordinates of objects;
  - connection via SAMP HUB to link VO clients like Aladin or Topcat;

Computing platforms



IA2 services

- IA2 manages data in a distributed manner on 3 continents!!
- GAPS experience bring important know-how. Pipelines and workflow management systems will be the must of 2020 Era Telescopes;
- IA2 allow state-of-the-art authentication and authorization mechanisms. Same results will be applied in pilot project for SKA and AENEAS.

Under development: VOSpace for storage and connection with computing facilities

VOSpace implementation compatible with CADC implementation, possible share of authorizations (thanks to S. Bertocco, G.Taffoni, S. Gaudet, P. Douler, B. Major experiments within EgiEngage)

Two levels of computation:

1. user approach to interactive pipeline with no HPC/HTC and small data volumes
   a. RAP + Yabi;
   b. Containers on IA2 infrastructure;
2. user approach using processing with HPC/HTC needs
   a. Containers or VMs on Chipp;
   b. GCloud.

Processing close to Data
+ POCs

Conclusions

- Target of IA2 Data Center:
  - Support science offering Archival services, computing power and maintenance of both

- Incoming services:
  - User space for private data access and computing;
  - collaborative tools: share data, papers, ideas, everything!!
  - Ticketing systems: for reporting problems, ask for help;
  - connection with computational facilities.

Thanks for your attention!