

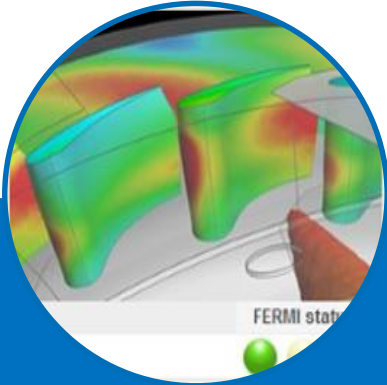
Cineca HPC services for OGS users with Galileo100

Mirko Cestari - CINECA

2021-09-16



Principal activities at CINECA



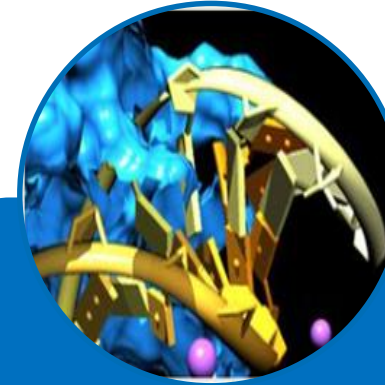
SUPPORT FOR RESEARCH & INNOVATION

High Performance Computing
and Data Analytics,
Scientific Visualization,
Virtual Interactive
Environments,
Research and development,
Education and Training



SERVICES FOR UNIVERSITY & MIUR

Workflow Innovation,
Dematerialisation,
Security,
Services and Solutions for
University Administration and
Ministry of Education,
University and Research



TECHNOLOGY TRANSFER

Numerical Experiments,
Prototyping,
Problem Solving,
Machine Intelligence,
Health and Research Clinics,
Numerical Forecasting

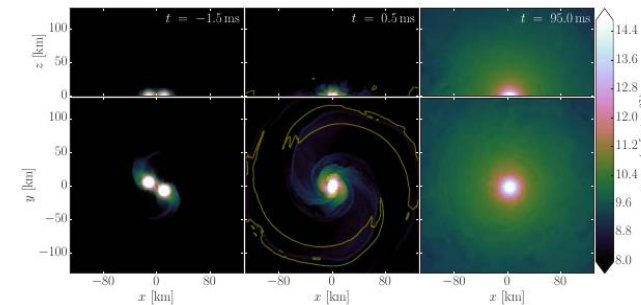
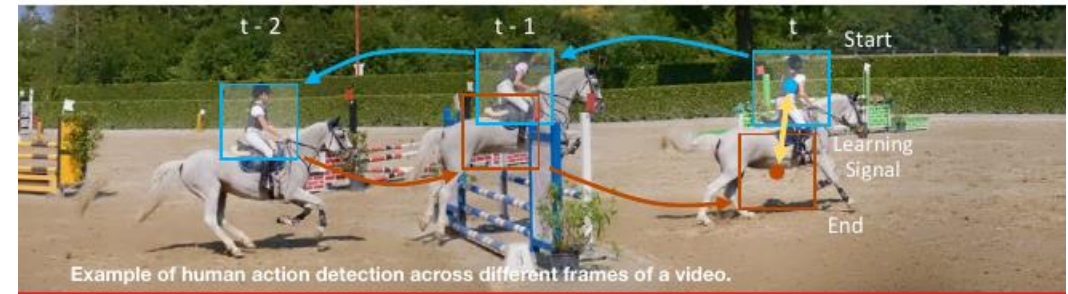
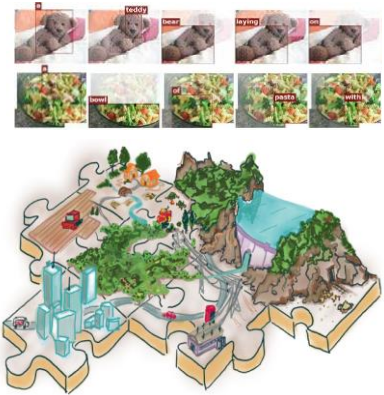
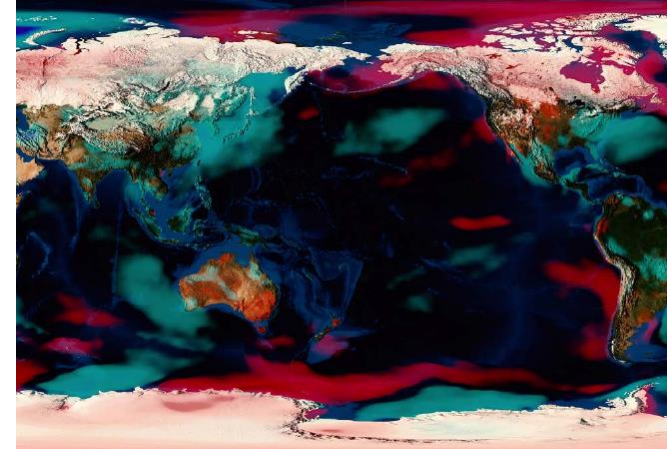
CINECA - HPC

10 billions of computing hours over last 3 years

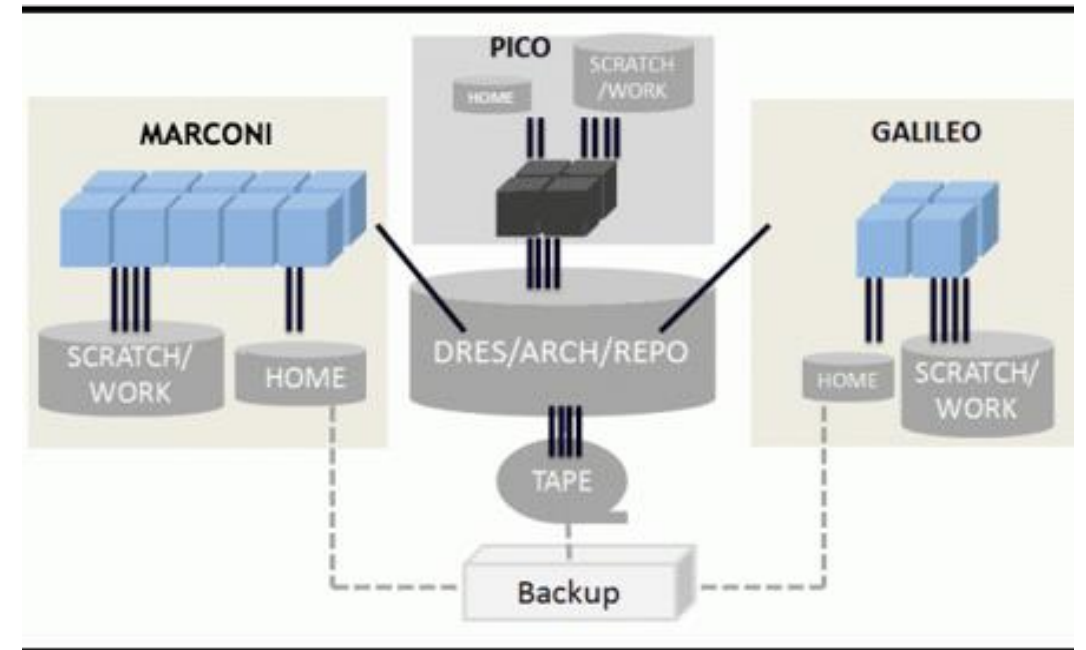
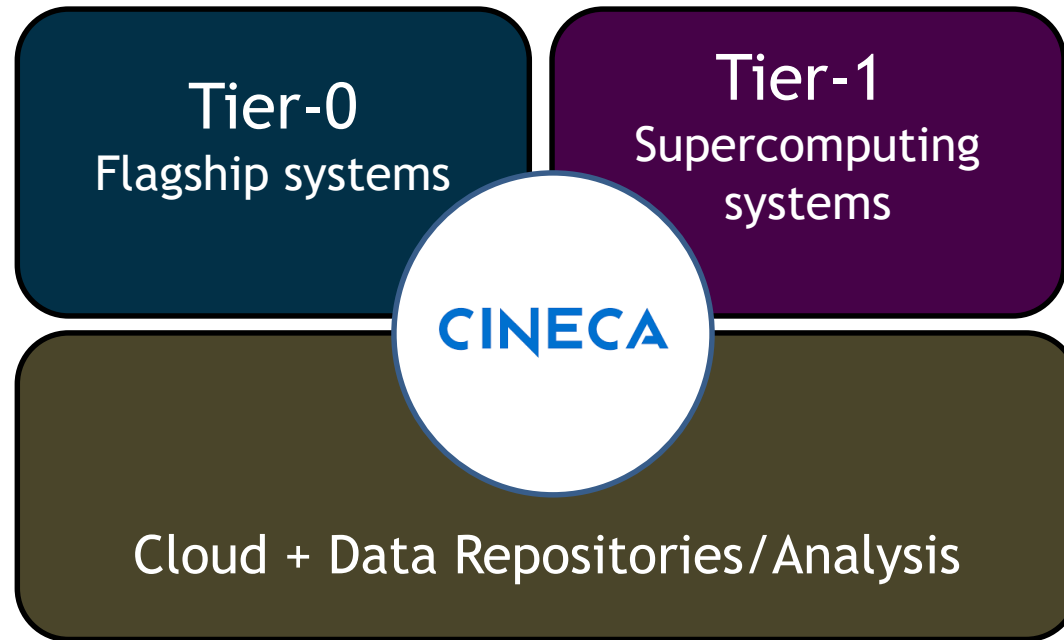
500 students on our Courses and schools

4000 active users

2000 active projects



CINECA HPC Infrastructure



PICO: Data Analysis system and Cloud

Compute

- **Processor:** Xeon E5 2670 v2 (2.5 Ghz)
- **Computing nodes:** 66
- **Visualization nodes:** 2, **GPU:** 2 Nvidia K40
- **Big insight nodes:** 4, **32 TB local disk**
- **Big mem nodes:** 2, **GPU:** Nvidia K20, **RAM:** 512 GB

Storage

- 90 TB SSD, 6PB GSS, 12 PB TAPE (LTFS)

Pico



Intel
Ivy Bridge

2015

2021

202?

Meucci: Weather forecast, K8s, and Cloud

Compute

- **Processor:** Xeon E5-2697 v4 (2.3 Ghz)
- **RAM:** 256 GB
- **Cloud nodes:** up to 98 (77)
- **Meteo:** ~120 nodes
- **K8s:** 4, to be upgraded to 8

Storage

- GSS through NFS, CEPH storage

ICEI-Fenix



Convergent

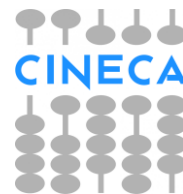


G100 genesis

G100 Introduction



- Common technical specifications
 - Derived from scientific use case requirements
 - Basis for the (coordinated) HW tender specifications
- Procurement of HW with EC funds
- Procurement started at the end of 2019
- Procurement ended November 2020
- Contract awarding January 2021
- System installation and deployment January-August 2021



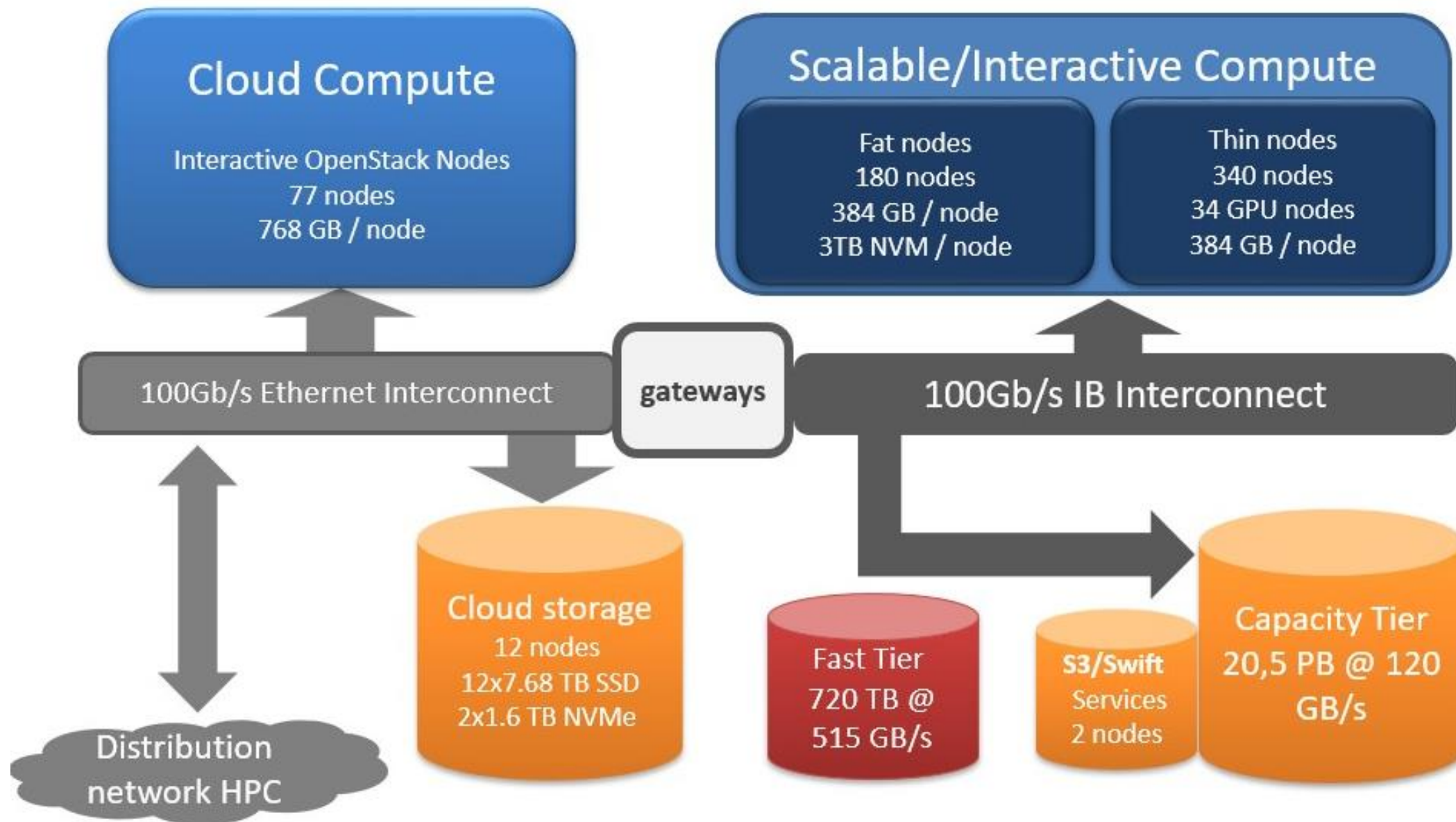
www.cineca.it



G100 Introduction

- Cineca elected to provision an infrastructure particularly devoted to data analysis and storage
 - Large memory servers
 - GPU servers
- Infrastructure able to offering multiple computing services
 - Mid-size parallel jobs
 - Interacting computing services
 - Flexible computing environments
- Very balanced in terms of computing and storage offerings
 - Intensive I/O workloads
- Tightly coupled services

Galileo-100: ICEI Fenix Infrastructure



Data Analysis and sharing

Scalable compute

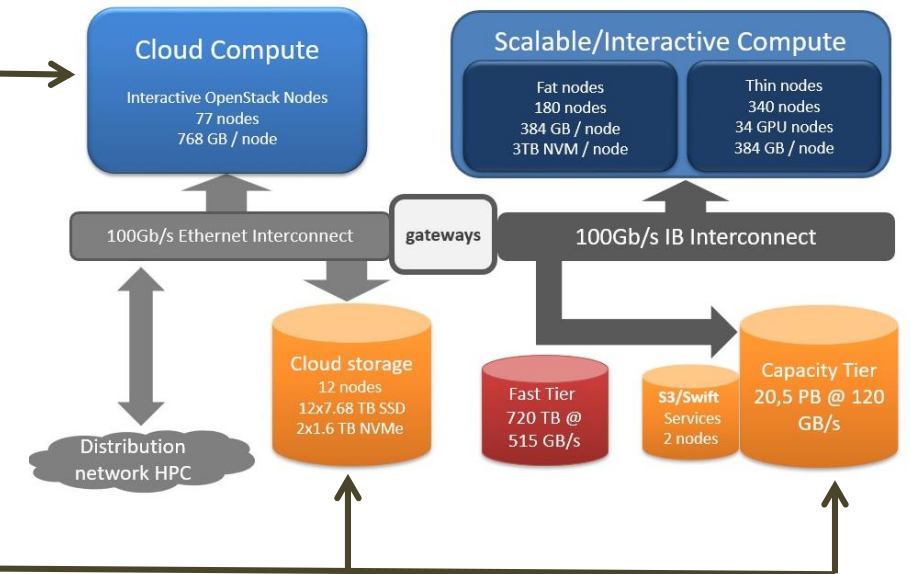
- 2x CSL Intel CPUs (48 cores)
- IB 100 Gbs (2:1)

Data Analysis/Interactive Computing

- NVIDIA GPUs
- Intel Optane: 3 TB Dimm memories
- Large RAM: 384 GB

Cloud provisioning

- VMs
- Storage
 - Hot-storage: 700 TB of high IOPS (full flash)
 - Archive storage through Lustre/NFS/Swift (maybe cinder)



G100 SCC and IAC services

Scalable computing and Interactive Computing can leverage:

- 340 thin nodes:
 - 2x CPU 8260 Intel Cascade Lake, 24 cores, 2.4 GHz base frequency (3,90 GHz Turbo)
 - 384 GB RAM DDR4 2933MT/s
 - 480 GB SSD local disk
- 180 Data processing Fat nodes
 - 2x CPU 8260 Intel Cascade Lake, 24 cores, 2.4 GHz base frequency (3,90 GHz Turbo)
 - 384 GB RAM DDR4 2933MT/s
 - 2 TB SSD local disk
 - 3 TB Intel Optane (12x256GB Intel Optane DCPMM)
- 34 GPU nodes
 - 2x NVIDIA GPU V100
 - 2x CPU 8260 Intel Cascade Lake, 24 cores, 2.4 GHz base frequency (3,90 GHz Turbo)
 - 384 GB RAM DDR4 2933MT/s
 - 2 TB SSD local disk



G100 cloud services

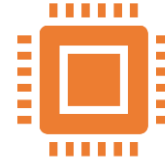
G100 Cloud services

Cloud services complement the HPC (SCC and IAC) infrastructure

On-Demand Self-Service	Easy Access	Resource Pooling	Rapid Elasticity	Measured Service
Project (tenant) Auto provisioning	Only network required Simple clients (ssh, web)	Flexible use of resources	Capacity to grow to satisfy user demand	Metrics to check usage

G100 Cloud services (cont.)

- Openstack <https://adacloud.hpc.cineca.it>
- User Guide: <https://wiki.u-gov.it/confluence/display/SCAIUS/ADA+CLOUD+CINECA+HPC>
- Available features
 - Encryption of cinder volumes
 - Manila (FS shared among VMs)
 - Magnum (Engine for container orchestration)
 - Zabbix for VM health monitoring
- Multiple flavors provided
- Num. of public IP available for all the infrastructure: 1024 addresses
- Ceph storage: 750 TB "Raid 6" full flash (SSD/NVMe)



7500 vCPUs



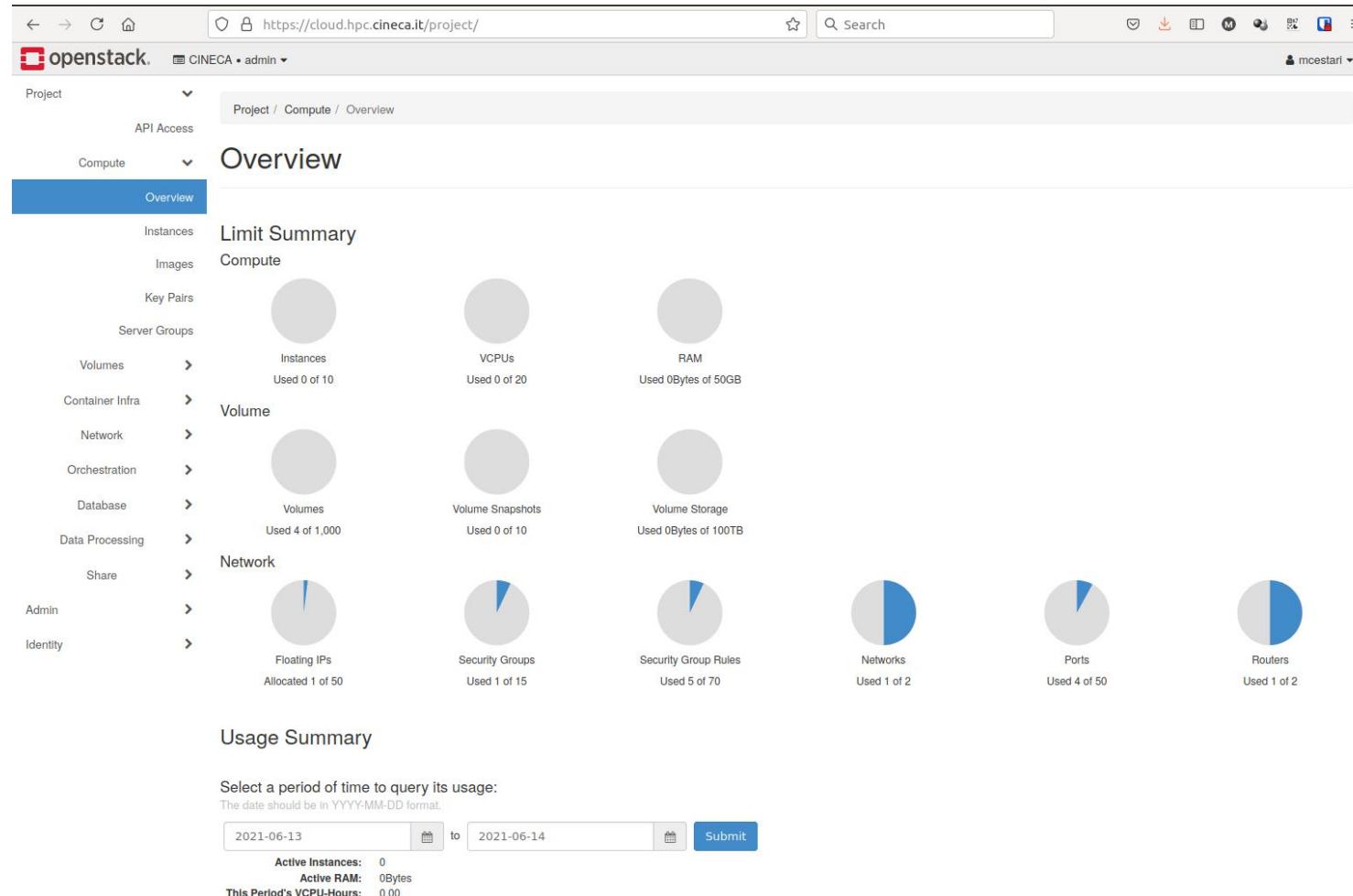
**77 Cascade Lake servers
upgradable to 150**



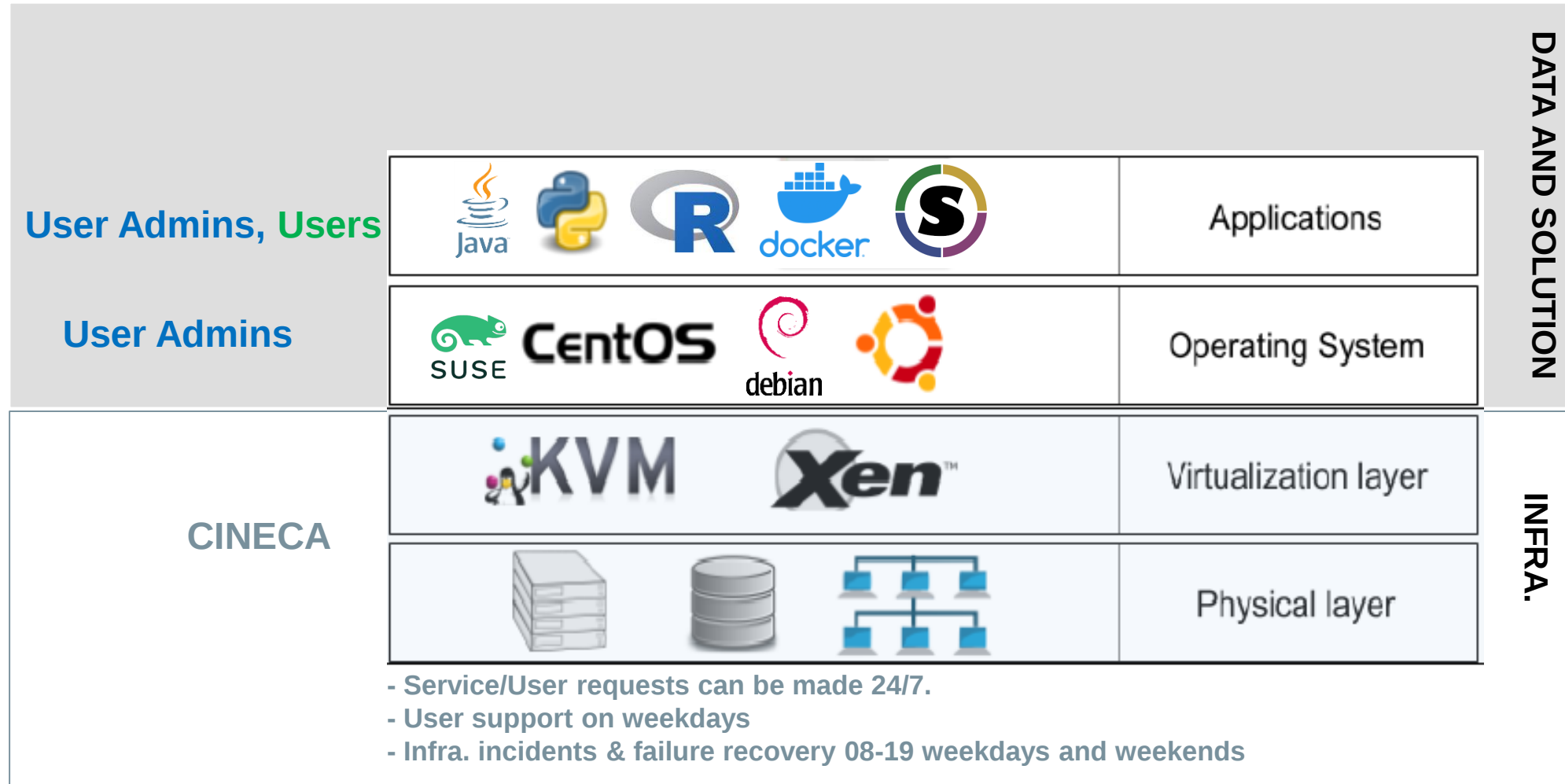
**750 TB Raid 6 CEPH
20 PB via client/NFS (based on
DDN Exascaler)**

G100 Cloud services (cont.)

accessible via: <https://adacloud.hpc.cineca.it>



Cloud User Support and Resource Management



CLOUD Layout

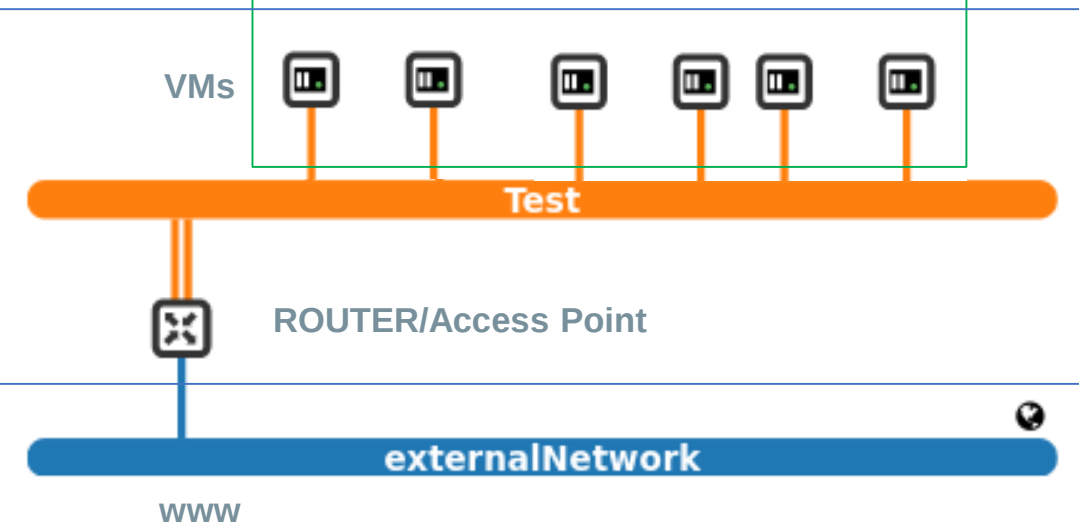
- **Admin creates VMs**
 - Admin A1 can access all VMs created by Admin A1
 - Admin can install a given OS dist.* on the VMs
 - Admin can give users access to VM
- **Each user of VM**
 - can be given root access to their own VM
 - can install their own dev. stack on their VM

Admin. can

- create VMs
- give users access to VM
- decide network topology
- decide security protocols
- create data volumes

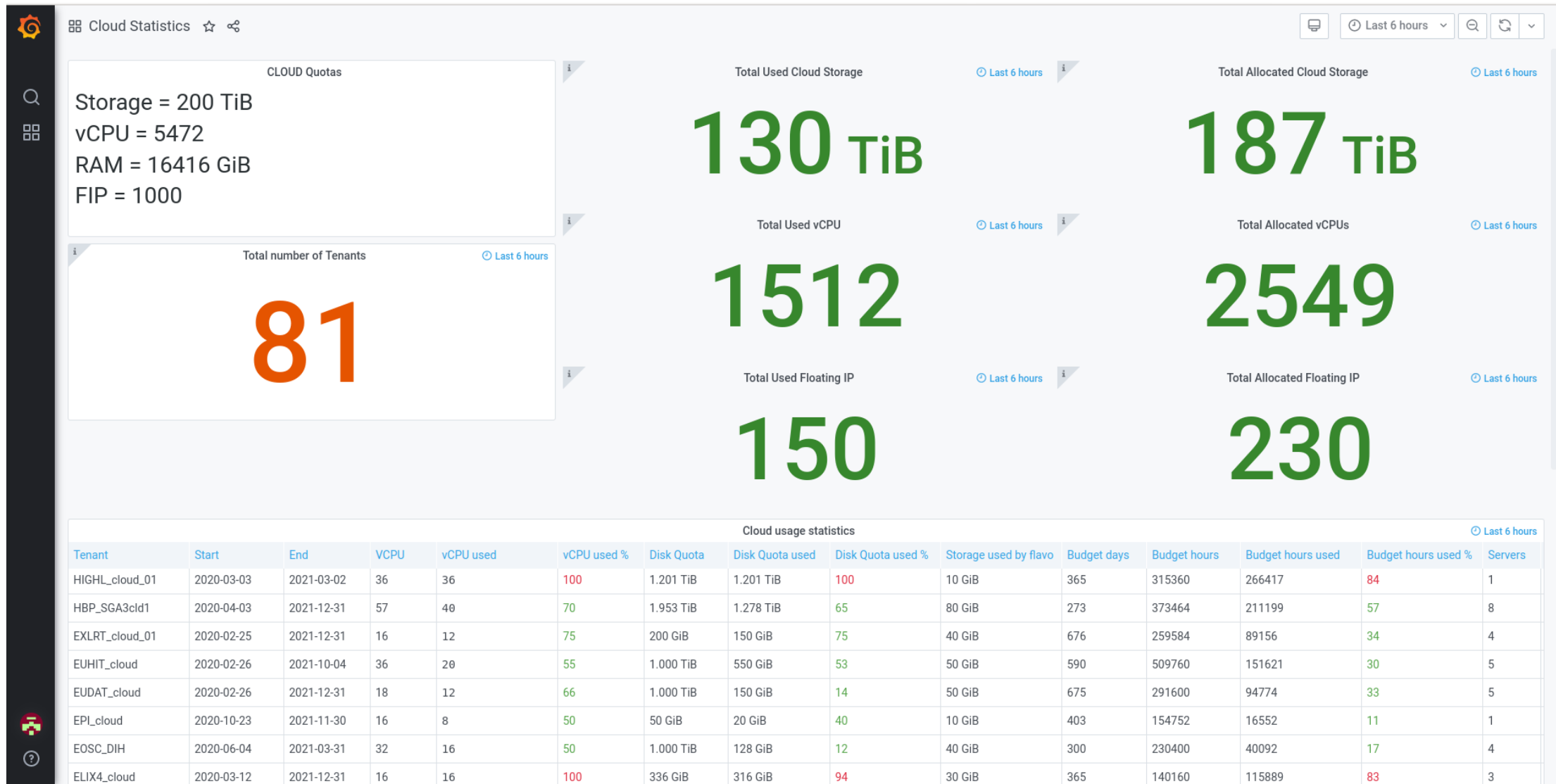
Users can

- install dev. stack
- login into their VM
- work on their solution

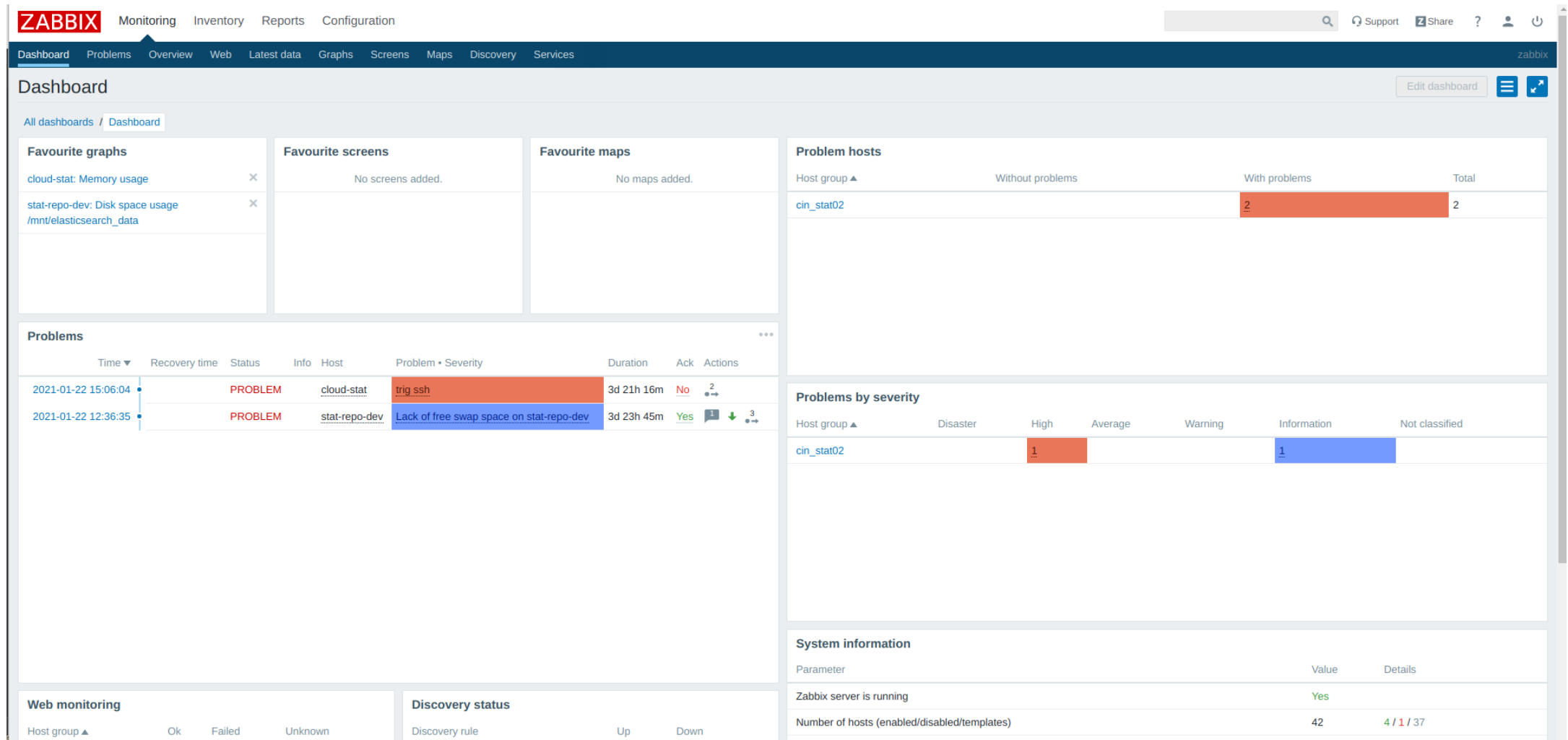


* standard linux: Ubuntu, Suse, Debian, CentOS

Current CLOUD-HPC CINECA



Current CLOUD-HPC CINECA





Click to add text

G100 user env

G100 Productivity Environment

Login interactive sessions

- data movement, archiving, code development, compilations, basic debugger
- are limited to 10 minutes of cputime

Batch sessions

- Workload manager is Slurm (<https://slurm.schedmd.com/>)
- SCC and IAC will exploit Slurm **partitions** and **QoS** to access diverse resources in a timely manner
- Partitions represent pool of nodes
- QoS define classes of batch jobs on (possibly) the same partition

G100 Productivity Environment

[Home](#) > [Resources](#) > [Software](#)

Application Software for Science

Content:

- [Chemistry](#)
- [Deep Learning](#)
- [Physics](#)
- [Life Science](#)
- [Bioinformatics](#)
- [Biodata](#)
- [Engineering](#)
- [Astronomy](#)
- [Visualisation](#)
- [Maths Libraries](#)
- [Data Libraries](#)
- [All Software](#)

Application Software for Science & Technology

Cineca offers a variety of third-party applications and community codes that are installed on its HPC systems.

Most of the third-party software are installed using the software modules mechanism.

The packages available and detailed descriptions of them can be **viewed for discipline** by selecting the menu on the left.

If you want to see the full catalog, please get [all Software in alphabetical order](#).

The information in this list may not reflect all software products available at Cineca. If you do not see an application you are interested in, or if you have questions about software that is currently available, please contact the [Help Desk](#)

Software catalogue

- <https://www.hpc.cineca.it/content/application-software-science>

G100 Productivity Environment

- module command

Command	Action
module avail	show the available modules on the machine
module load <appl>	load the module <appl> in the current shell session, preparing the environment for the application.
module load autoload <appl> ...	load the module <appl> and all dependencies in the current shell session
module help <appl>	show specific information and basic help on the application
module list	show the modules currently loaded on the shell session
module purge	unload all the loaded modules
module unload <app>.....	unload a specific module

All **profiles, categories, and modules available** on G100 system via “modmap”

> modmap -all

G100 Productivity Environment



- Software auto-provisioning
 - SPACK <https://github.com/spack/spack>
- `$ module load spack/<vers>`
 - By loading this spack module, `setup-env.sh` file is sourced. Then `$SPACK_ROOT` is initialized to `/cineca/prod/opt/tools/spack/<vers>/none`, `spack` command is added to your `PATH`, and some nice command line integration tools too.
 - A folder is created into your default `$WORK` space (`$USER/spack-<vers>`) in order to contain some subfolders created and used by spack during the phase of a package installation:
 - sources cache: `$WORK/$USER/spack-<vers>/cache`
 - software installation root: `$WORK/$USER/spack-<vers>/install`
 - module files location: `$WORK/$USER/spack-<vers>/modulefiles`

G100 Productivity Environment

- Spack currently has 5646 mainline packages
https://spack.readthedocs.io/en/latest/package_list.html
- Strong community effort
 - Users can rely on work done by others
 - Easier deploy of project software stack





Use Cases

Weather forecast

COSMO simulations (Fortran+MPI)

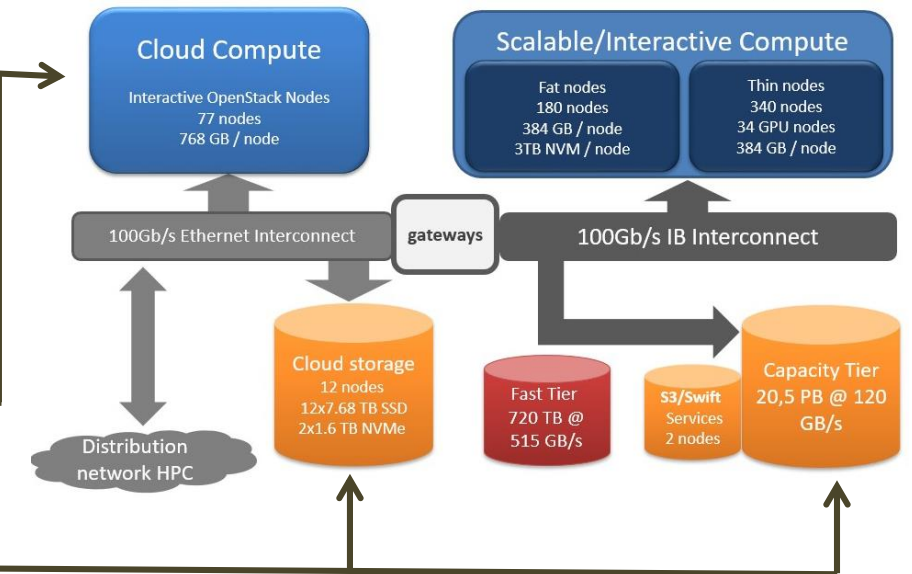
- Grid: 5km Mediterranean, 2,2km Italy
- simulation time 18-72 hours
- 2 runs every day

Hybrid simulations

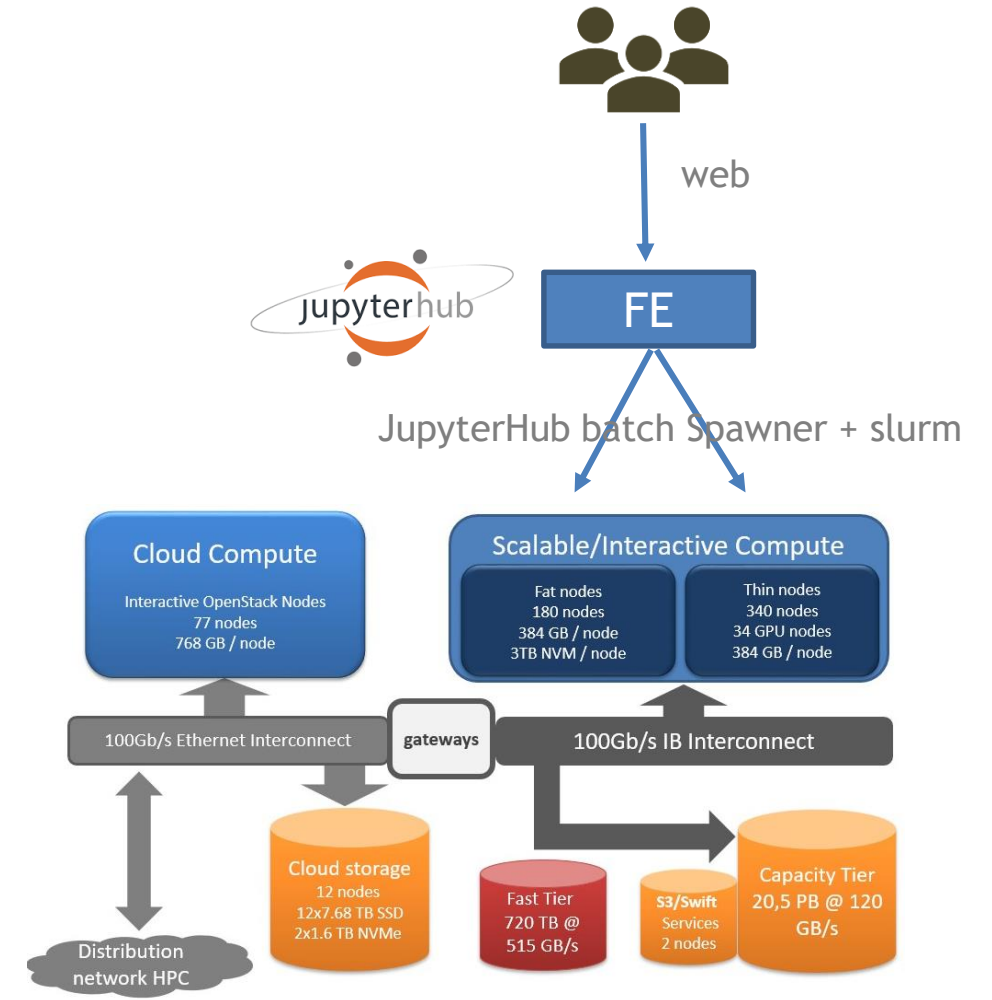
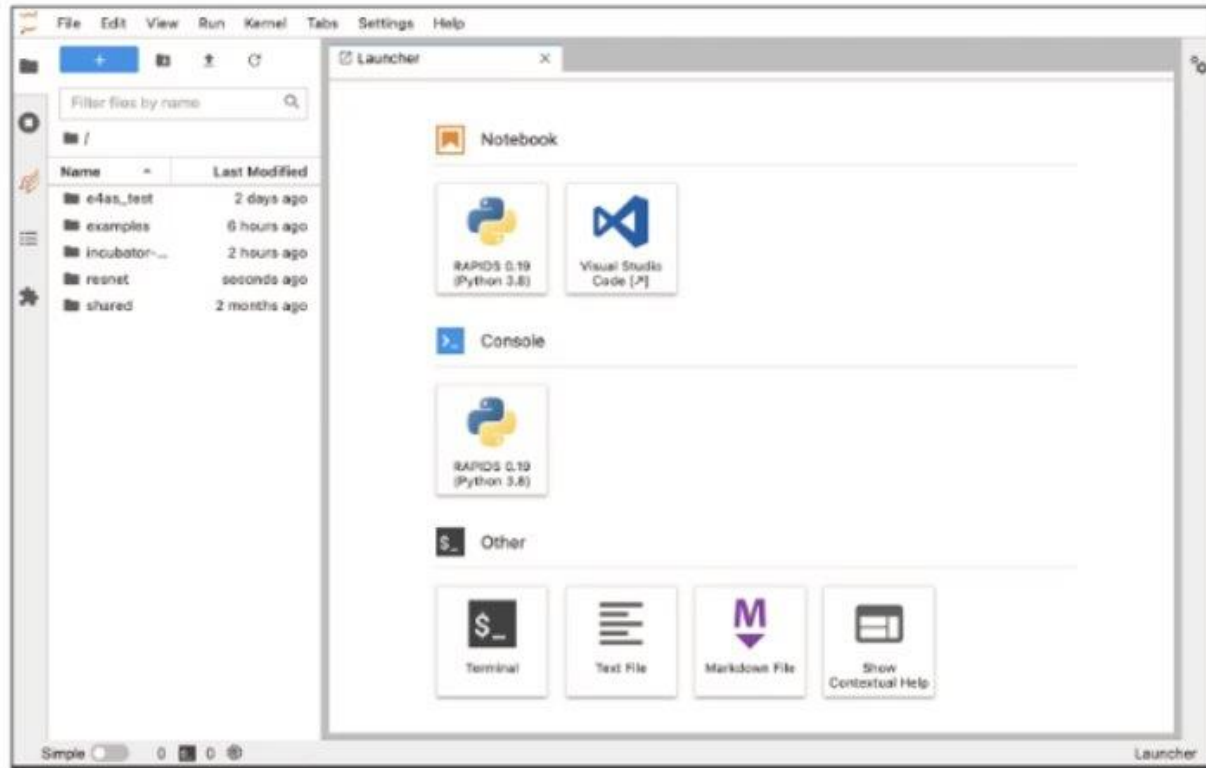
- Application ported on NVIDIA GPUs
- Evaluation of the stability of application/model still required
- Results need to be periodically checked with CPU simulations

Virtual hosting

- dissemination of data/maps
 - VMs+IO
- 250 TB of archival space, tightly integrated
- <https://meteohub.hpc.cineca.it/>



G100 Interactive Computing



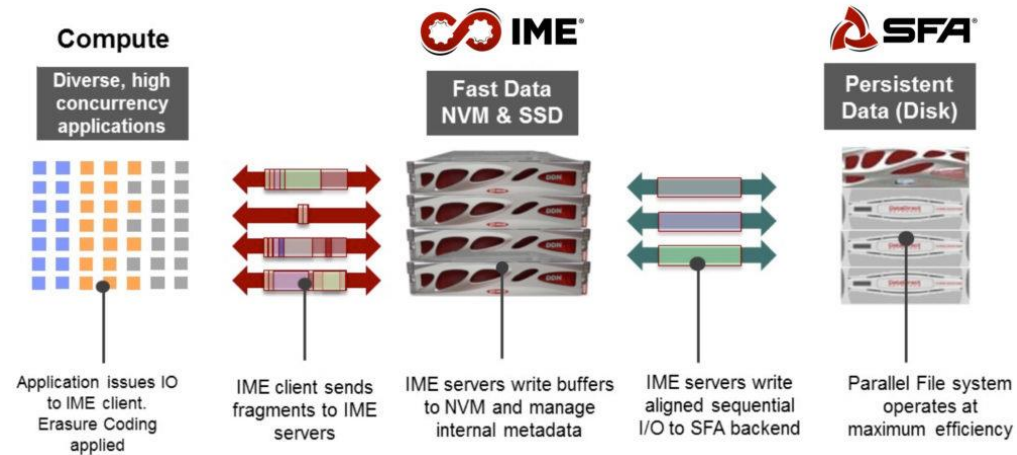


G100 storage

Active Data Repositories (ACD)

- Based on DDN solutions

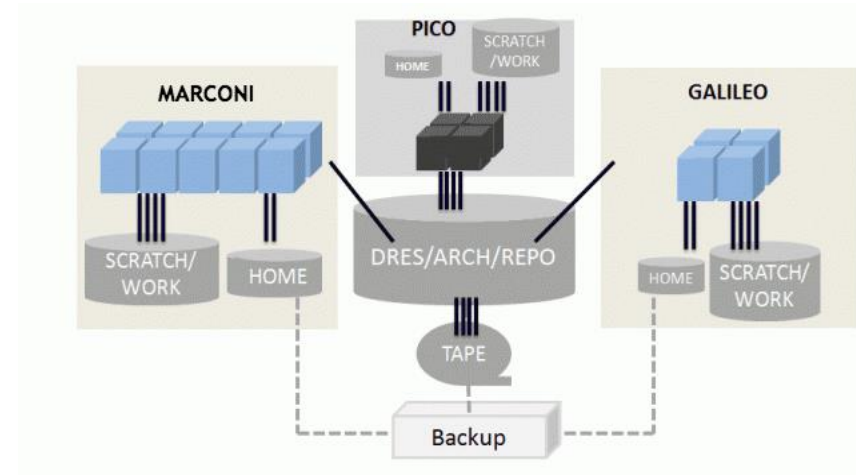
- IME appliance
- Large capacity storage
Lustre Filesystem



- Enhanced IO performance
- Enhanced security and data segregation (via nodemap)
- (Eventually) data *at rest* encryption

Active Data Repositories (cont.)

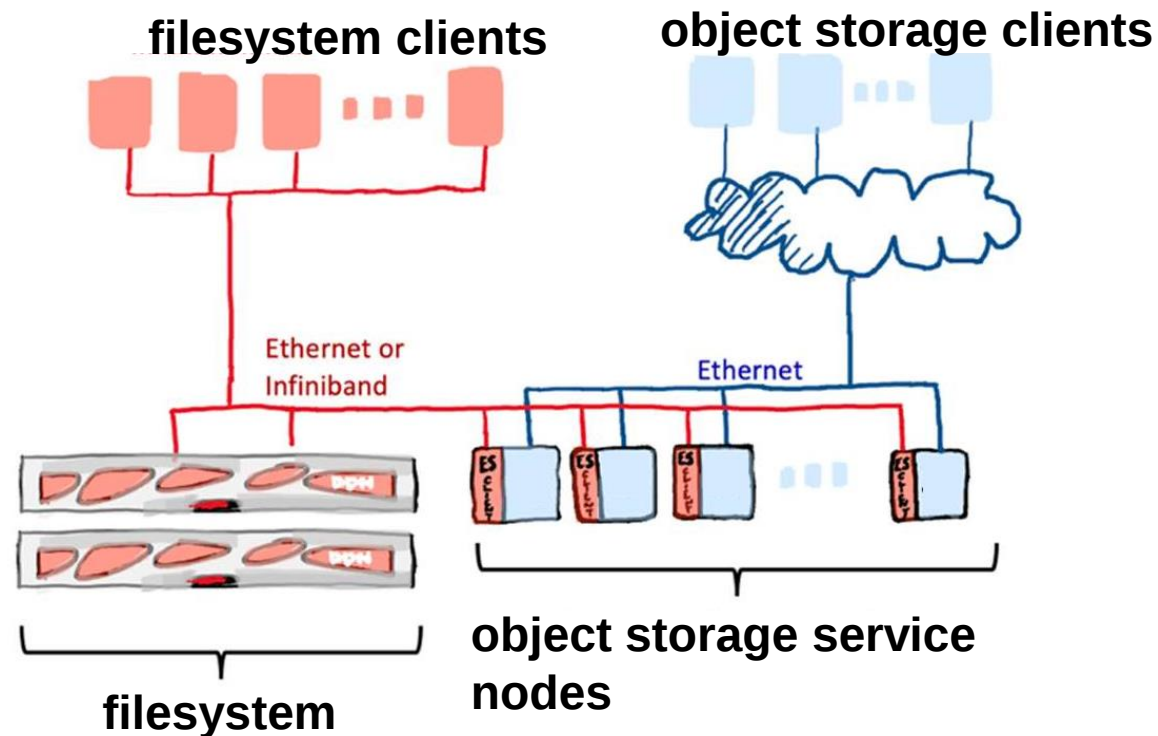
- temporary/permanent
- user/project specific
- local/shared
- Data Storage architecture
 - \$HOME: permanent/backed up, user specific, local
 - \$WORK: permanent, project specific, local
 - \$CINECA_SCRATCH: temporary , user specific, local
 - \$TMPDIR: temporary, user specific, local
 - \$DRES: permanent, shared (among platforms and projects)
 - \$TAPE: permanent, user specific, shared



Archive Data Repositories (ARD)

■ Object storage service

- Scalable service to store massive amount of data (structured and unstructured)
- Data repository accessed via HTTP calls
- Ideal for single write multiple reads (storing sensors data)



Thank you

Dr. Mirko Cestari

HPC and Cloud Technology team coordinator

CINECA, www.cineca.it

Mail: m.cestari@cineca.it

