FENIX

RESEARCH INFRASTRUCTURE

Introduction to Galileo100 system (G100) at CINECA

Mirko Cestari – CINECA

June 2021

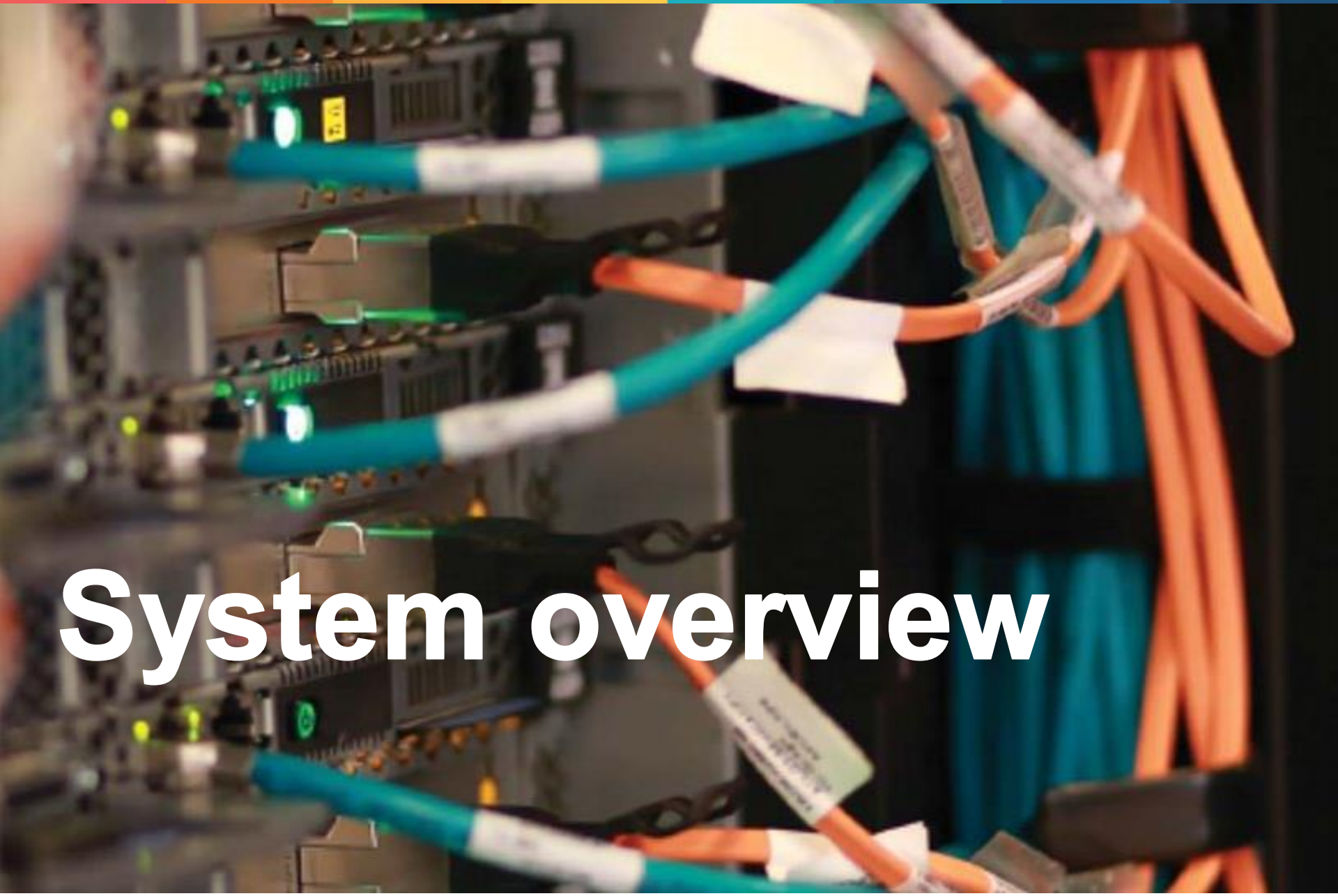
The ICEI project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 800858.

Introduction

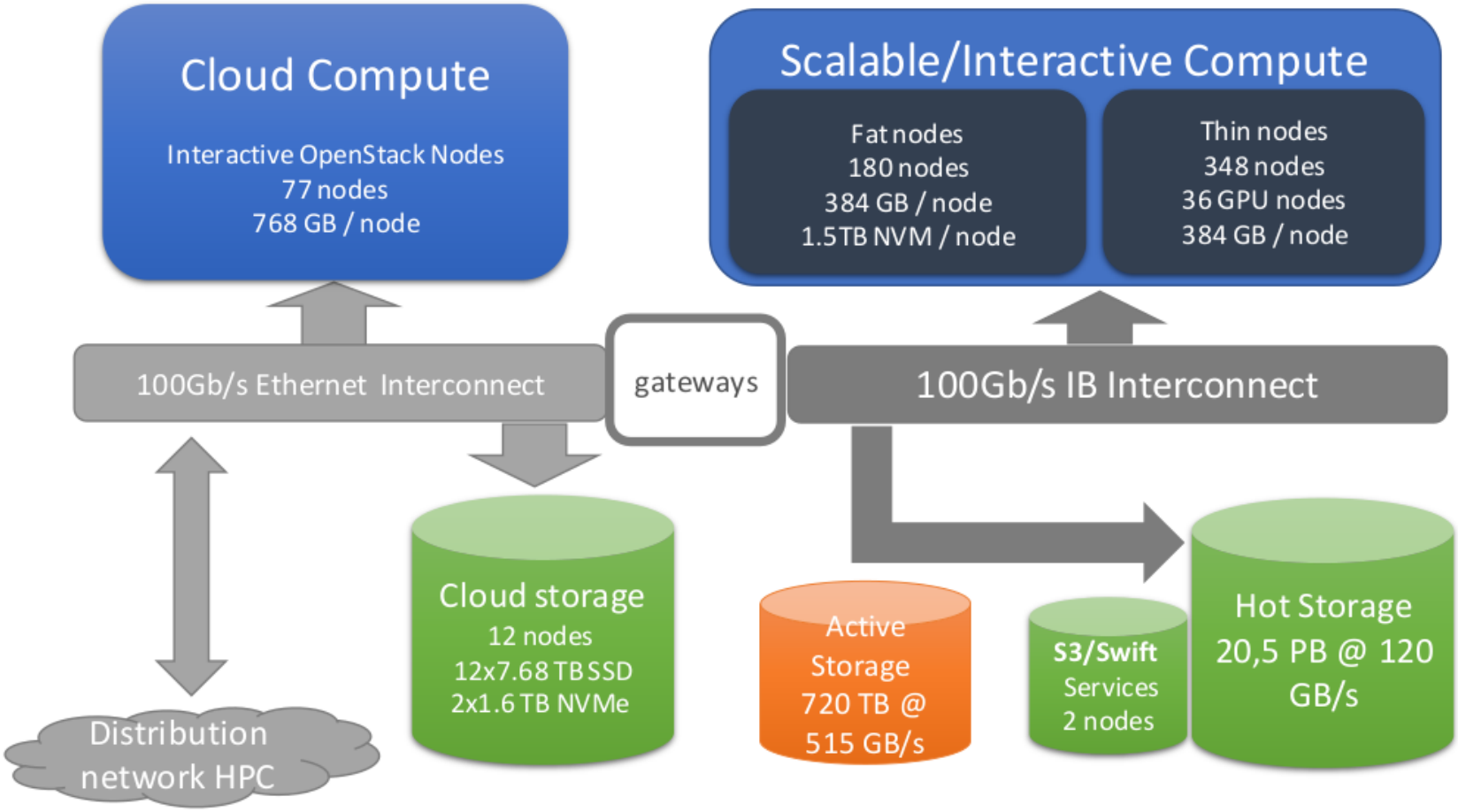
- Common technical specifications
  - Derived from scientific use case requirements
  - Basis for the (coordinated) HW tender specifications
- Procurement of HW with EC funds
- Procurement started at the end of 2019
- Procurement ended November 2020
- Contract awarding January 2021
- System installation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

Introduction (cont.)

- 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


System overview

G100 Infrastructure overview



G100 infrastructure overview (cont.)

Scalable compute

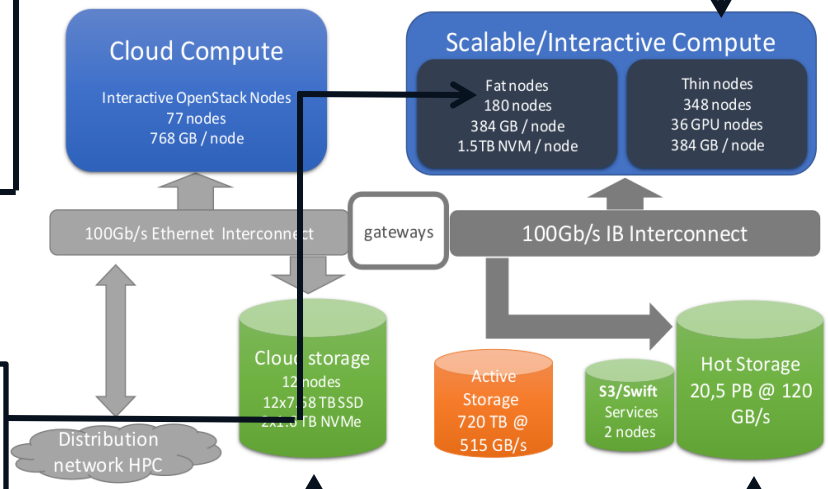
- 2x CSL Intel CPUs (48 cores)
- IB 100 Gbs (2:1)

Data Analysis/Interactive Computing

- NVIDIA GPUs
- Intel Optane: 1,5 TB Dimm memories
- High RAM: 384 GB

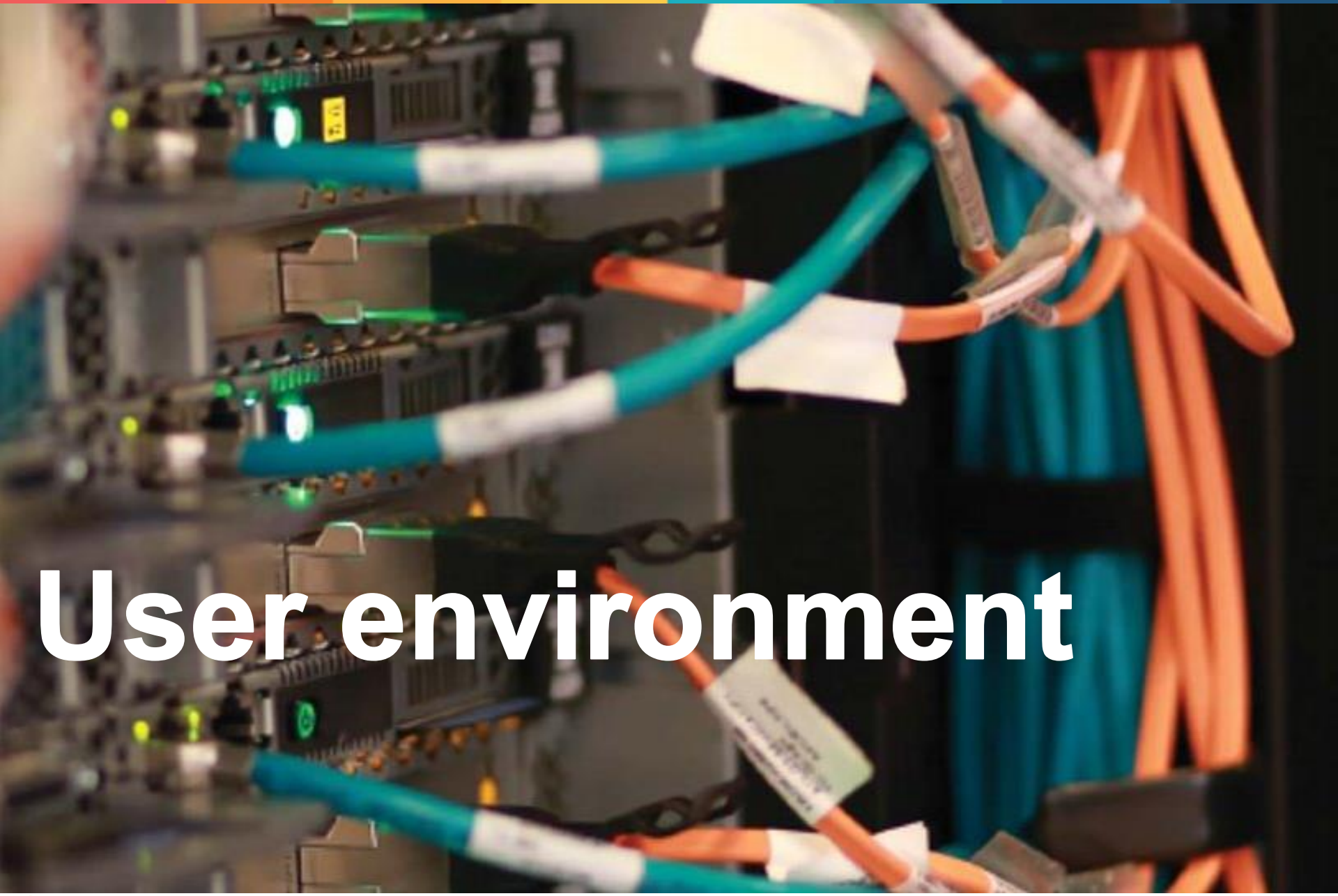
Cloud provisioning

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



G100 SCC and IAC

- Scalable computing and Interactive Computing can leverage:
  - 348 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
    - 1,5 TB Intel Optane (12x128GB Intel Optane DCPMM)
  - 36 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

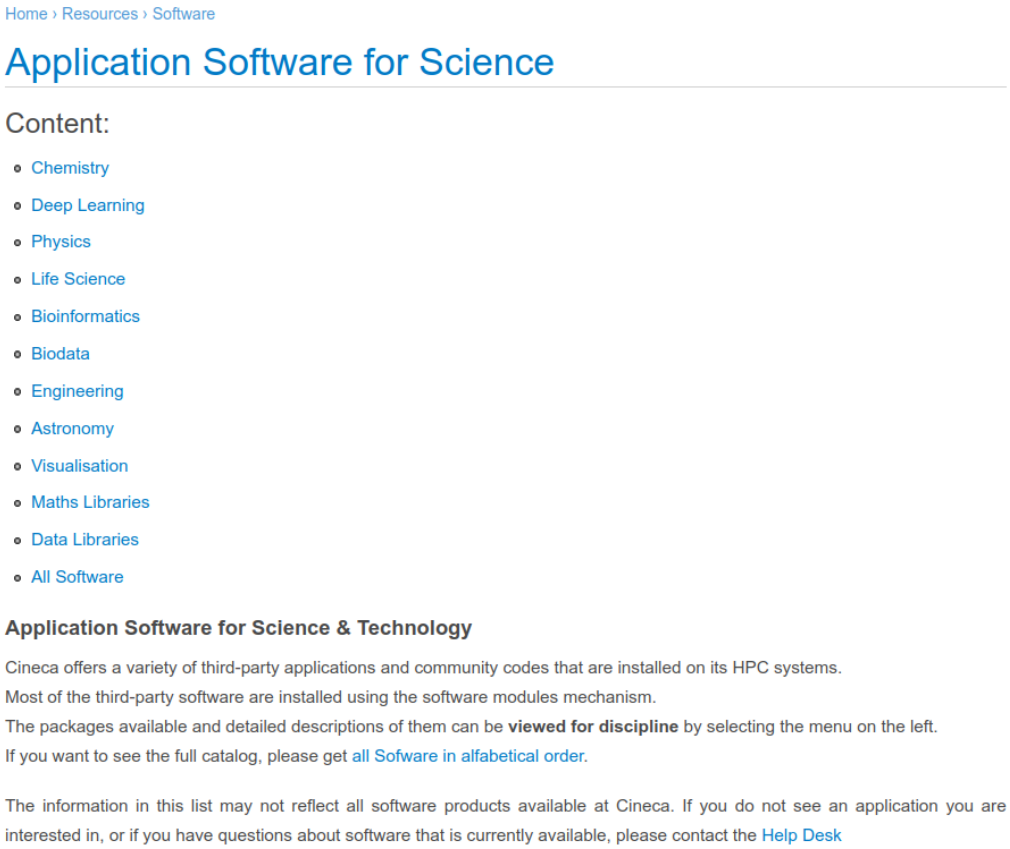

User environment

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 (cont.)

- Software catalogue
  - https://www.hpc.cineca.it/content/application-software-science



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 Sofware in alfabetical 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

G100 Productivity environment (cont.)

- 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

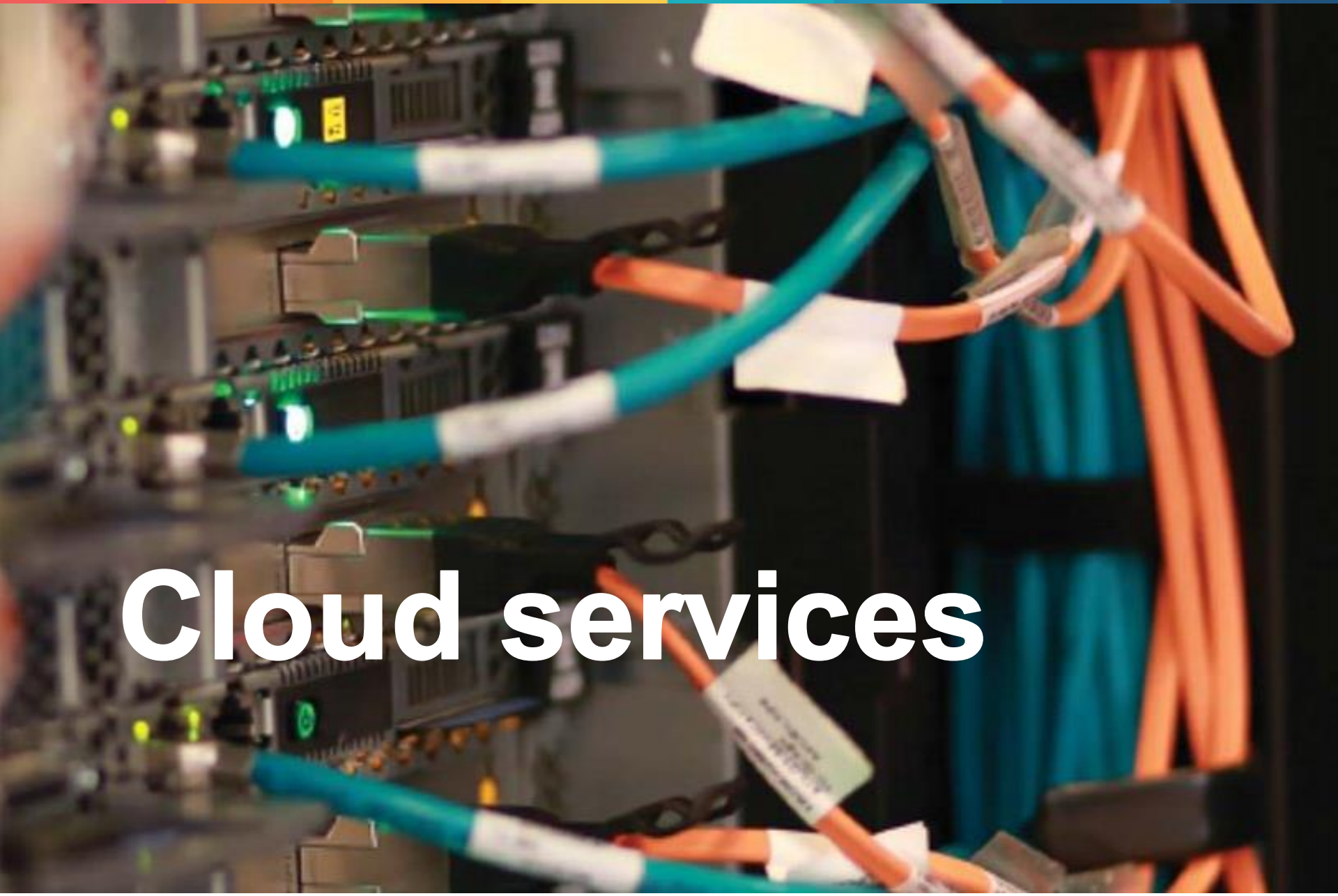
G100 Productivity environment (cont.)

- 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 (cont.)

- 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


Cloud services

Cloud Services

- Cloud services complement the HPC (SCC and IAC) infrastructure

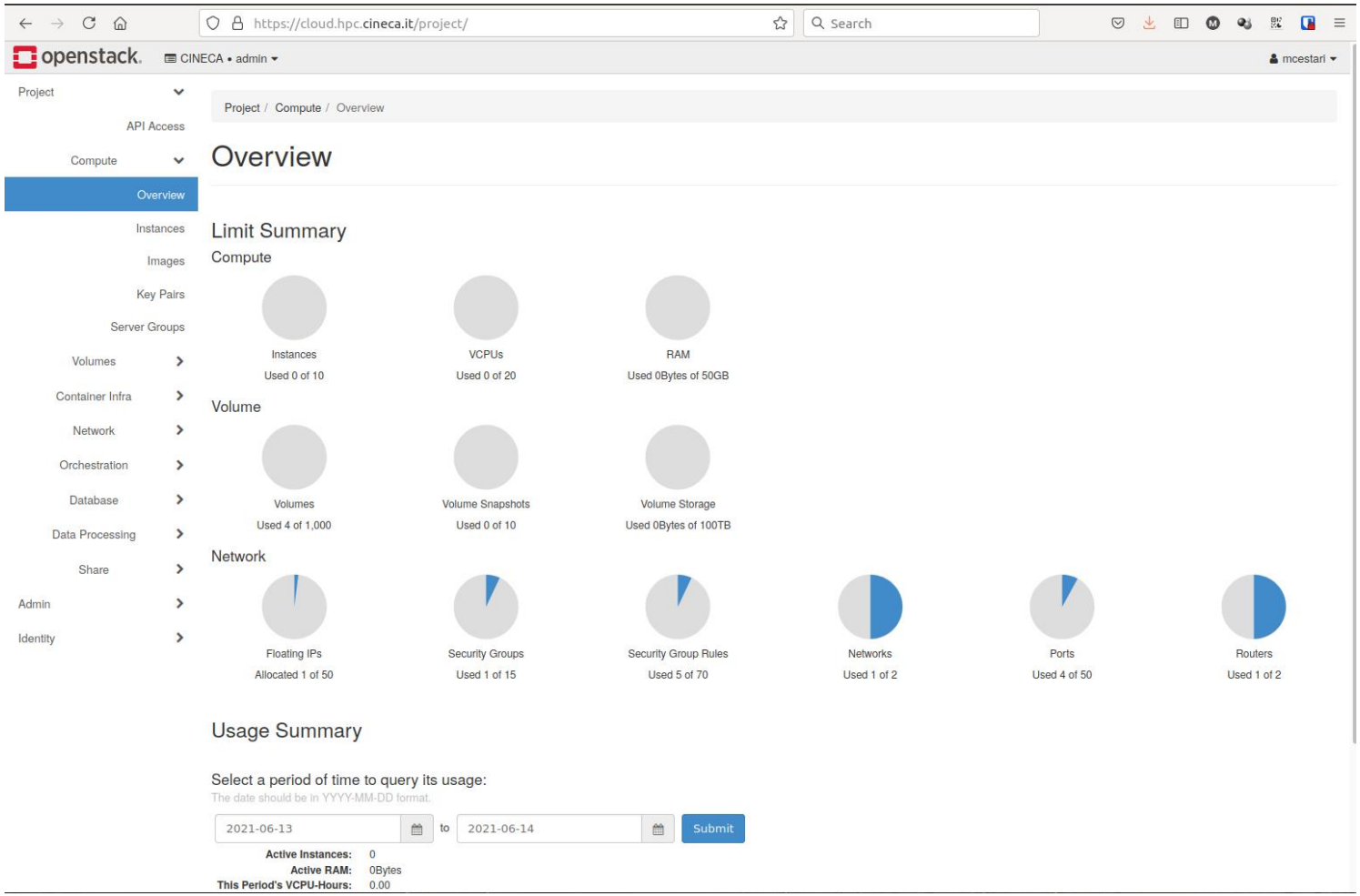
On-Demand Self-Service	Easy Access	Resource Pooling	Rapid Elasticity	Measured Service
Project (tenant) Auto provisioning	Only network required Simple clients (ssh, web)	Flexible use of resources	Capacity to grow to satisfy user demand	Metrics to check usage

Cloud Services (cont.)

- 77 Openstack compute servers, featuring 2x CPU 8260 Intel CascadeLake, 24 cores each at 2.4 GHz, for a total of 48 cores per server. This amounts to a total of 7400 vCPUs for the cloud infrastructure. Each computing server features also 768 GB of DDR4 RAM.
- 12 OpenStack storage nodes providing a total of 720 TB of dedicated CEPH storage full flash (NVMe/SSD) for high IOPS. This storage will host VM root disks, disk volumes and VM snapshots.
- 6 nodes for storage management metadata (CEPH).
- 1024 floating IPs for external (public) connection.
- 3 service nodes to host OpenStack services.

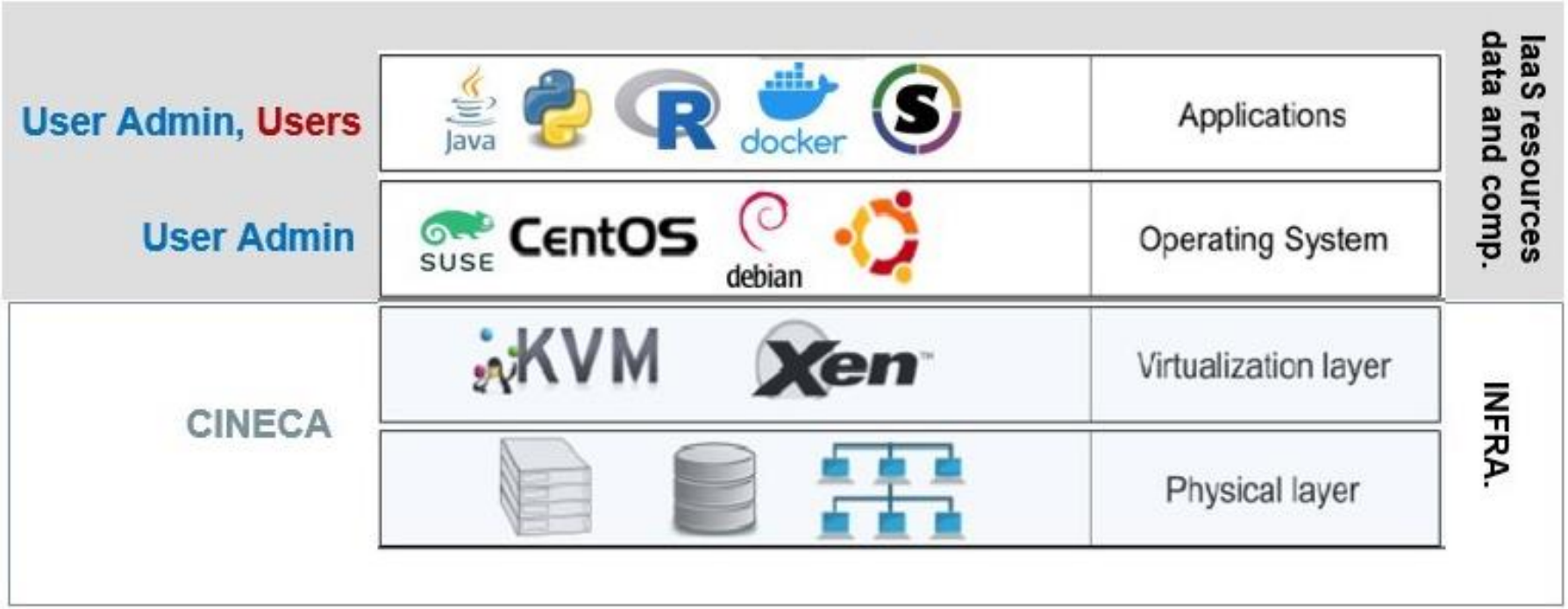
Cloud Services (cont.)

- accessible via: https://cloud.hpc.cineca.it



Cloud Services (cont.)

- UserGuide: https://wiki.u-gov.it/confluence/display/SCAIUS/CINECA+public+cloud



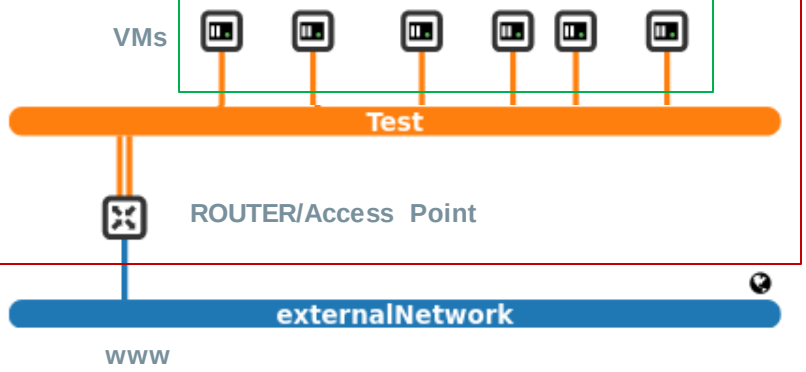
Cloud Services (cont.)

- **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 VMs**
  - **can be given root access to their own VM**
  - **can install their own dev. stack on their VM**

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

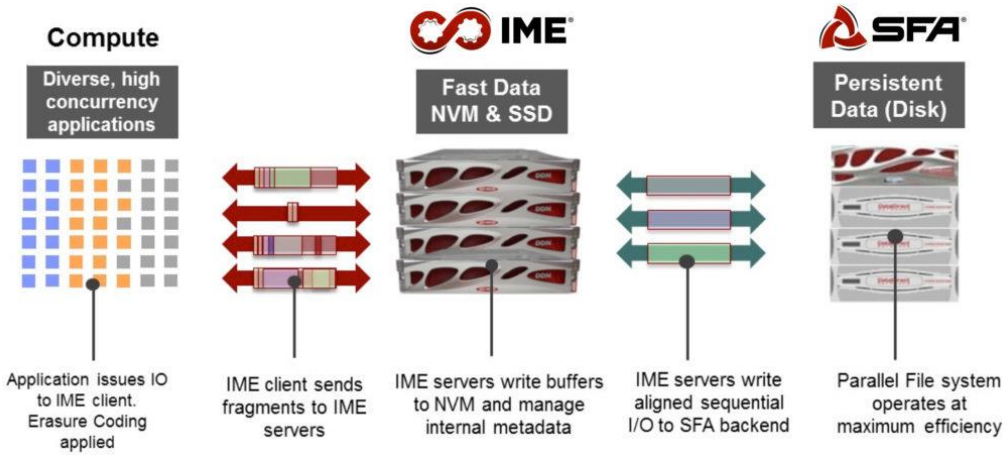
Admin. can
- create VMs
- give users access to VM
- decide network topology
- decide security protocols
- create data volumes




Storage

Active Data Repositories (ACD)

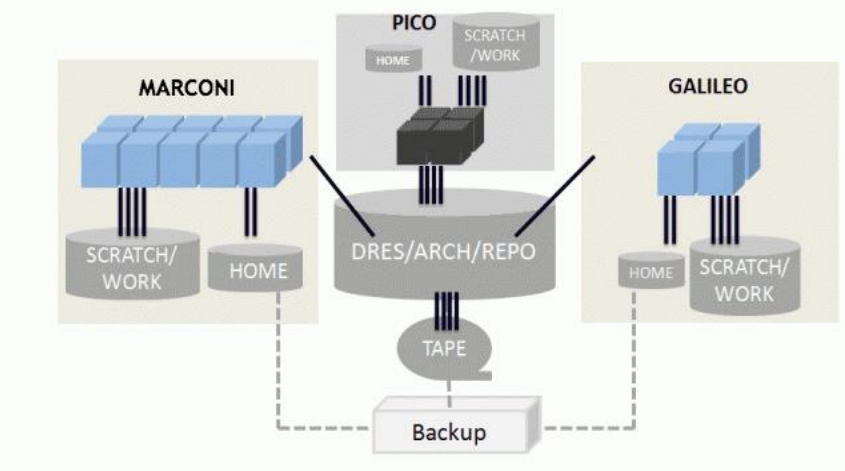
- Based on DDN solutions
  - IME appliance
  - Large capacity storage Lustre Filesystem



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

Active Data Repositories (cont.)

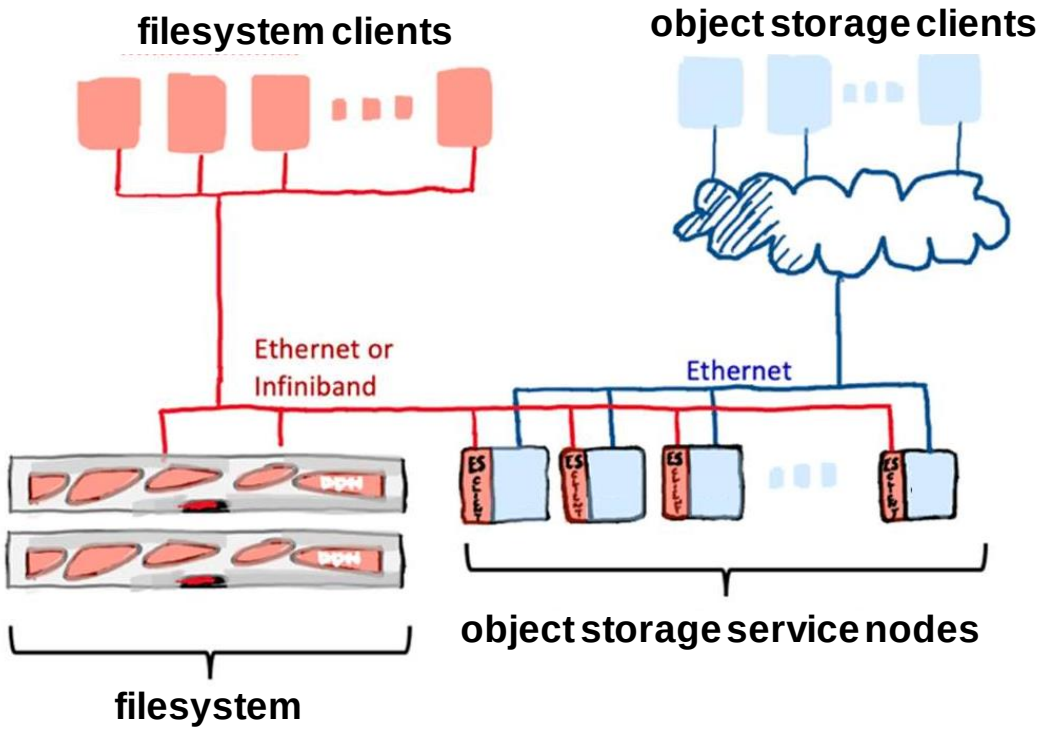
- temporary/permanent
- user/project specific
- local/shared



- 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 for your attention!