INFN-DataCloud a distributed infrastructure supporting multi purpose Scientific data analytics services

Giacinto Donvito

(donvito@infn.it)

EGI 2024

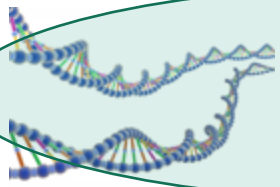
Physics experiments

Il Large Hadron Collider (LHC) al CERN

1 LHC Experiment ~2020: ~200.000 CPU Cores; ~200 PB disk; ~350 PB tape → 1 High-Luminosity LHC Experiment ~2028: ~15M CPU Cores; ~15 EB disk; ~26 EB tape

DUNE: ProtoDUNE in 2019 collected 3 GB/s; real DUNE expected 80x at the end of the 2020s.

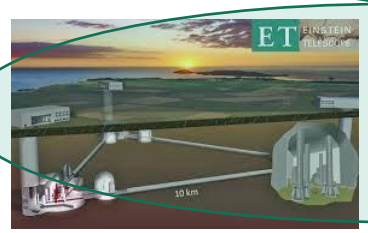
SKA: up to 2 PB/day , to be collected and processed at "complex" locations.



Genomics: a single genome ~100 GB. Any population study (>1M people) over 100 PB

CTA: ~ 10 PB/y in 2025+.

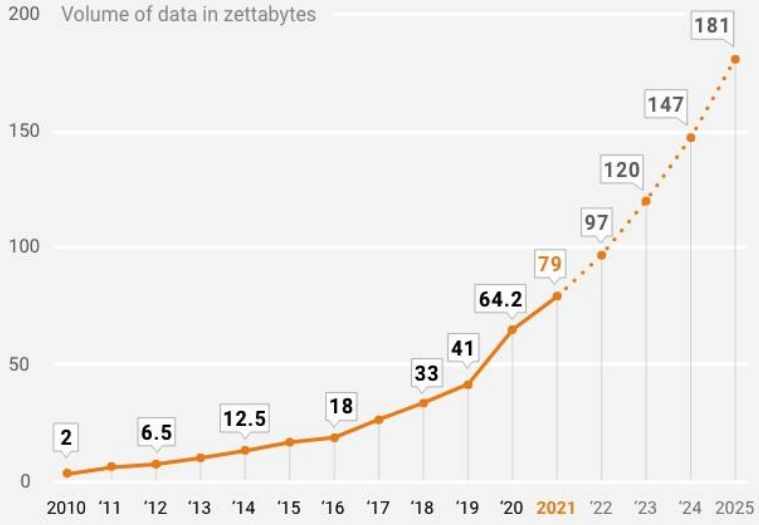
Virgo: ~10% of a LHC experiment.



Einstein Telescope: aiming at being ~10% of a HL-LHC experiment.

Volume of data created, captured, copied, and consumed worldwide

The volume of data generated, consumed, copied, and stored is projected to exceed 180 zettabytes by 2025



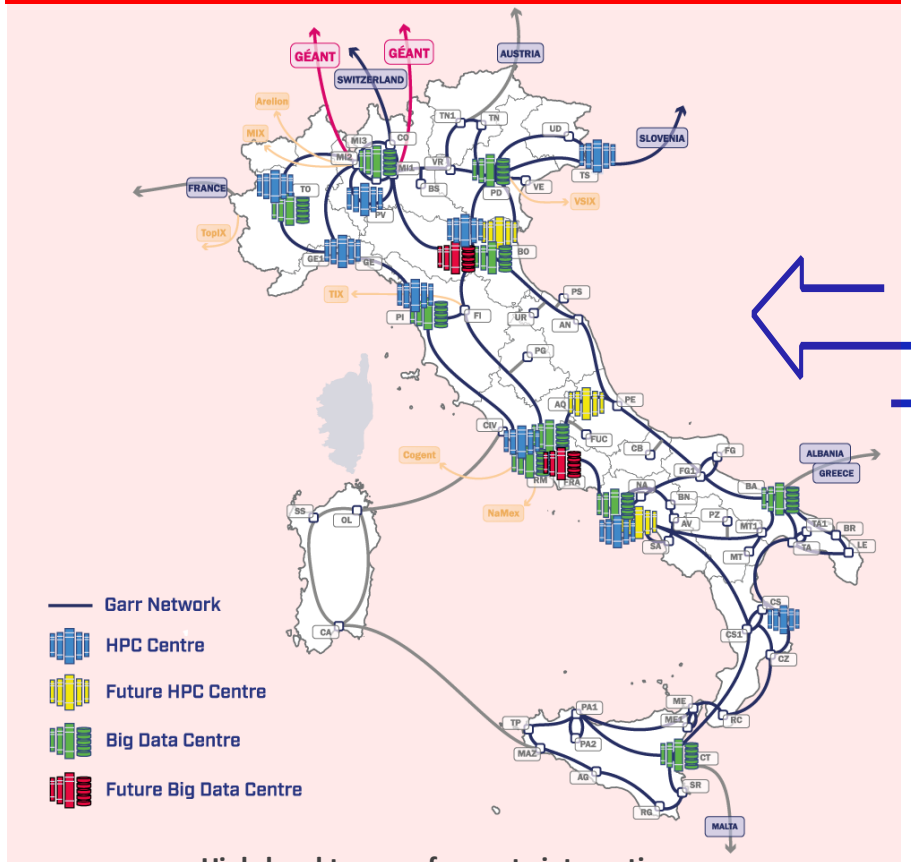
The INFN **Facilities**

ICSC includes

10 thematic spokes
1 infrastructure spoke

25 universities
12 Research institutes
14 Strategic private companies

0 SUPERCOMPUTING CLOUD INFRASTRUCTURE



High-level teams of experts integrating
the Spokes working groups (mixed cross-sectional teams)

EDUCATION & TRAINING, ENTREPRENEURSHIP, KNOWLEDGE TRANSFER, POLICY, OUTREACH

1 FUTURE HPC & BIG DATA

2 FUNDAMENTAL RESEARCH & SPACE ECONOMY

3 ASTROPHYSICS & COSMOS OBSERVATIONS

4 EARTH & CLIMATE

5 ENVIRONMENT & NATURAL DISASTERS

6 MULTISCALE MODELING & ENGINEERING APPLICATIONS

7 MATERIALS & MOLECULAR SCIENCES

8 IN-SILICO MEDICINE & OMICS DATA

9 DIGITAL SOCIETY & SMART CITIES

10 QUANTUM COMPUTING

The TeRABIT project

The Terabit network for Research and Academic Big data in Italy (TeRABIT) envisions the creation of a a distributed, hyper-connected, hybrid HPC-Cloud environment that offers services designed to meet the evolving needs of research and innovation.

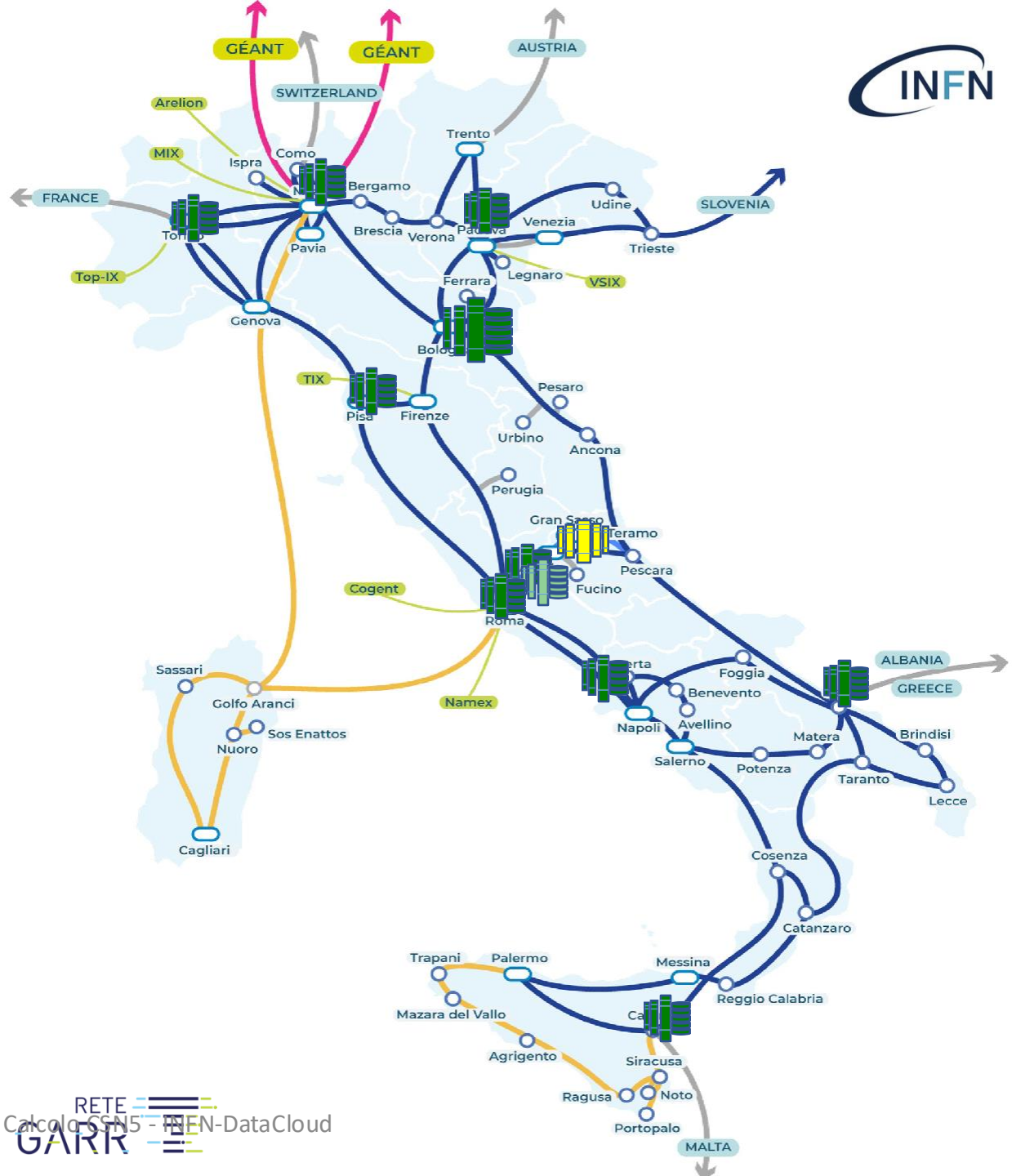
The environment will federate and strengthen the three existing research infrastructures GARR-T, PRACE-Italy and HPC-BD-AI (HPC-Big Data-Artificial Intelligence), leveraging their existing of connections to other national and European research infrastructures and data spaces through the GÉANT backbone.

Main objectives:

- Enable widespread data transfer, up to Terabits per second, and services on a national scale in Italy, connected to Europe;
- Innovate the central HPC node of PRACE-Italy;
- Innovate the HPC services offered to researchers, beyond the centralized calculation model, adding distributed "HPC-Bubbles"

DataCloud is the Infrastructure for INFN Scientific Computing

- Tier-1 (CNAF)
- Tier-2's (BA, CT, LNF, LNL/PD, NA, MI, PI, RM1, TO)
- INFN Cloud
  - Backbone and federated clouds
- HPC4DR (LNGS)
- (Tier-3)

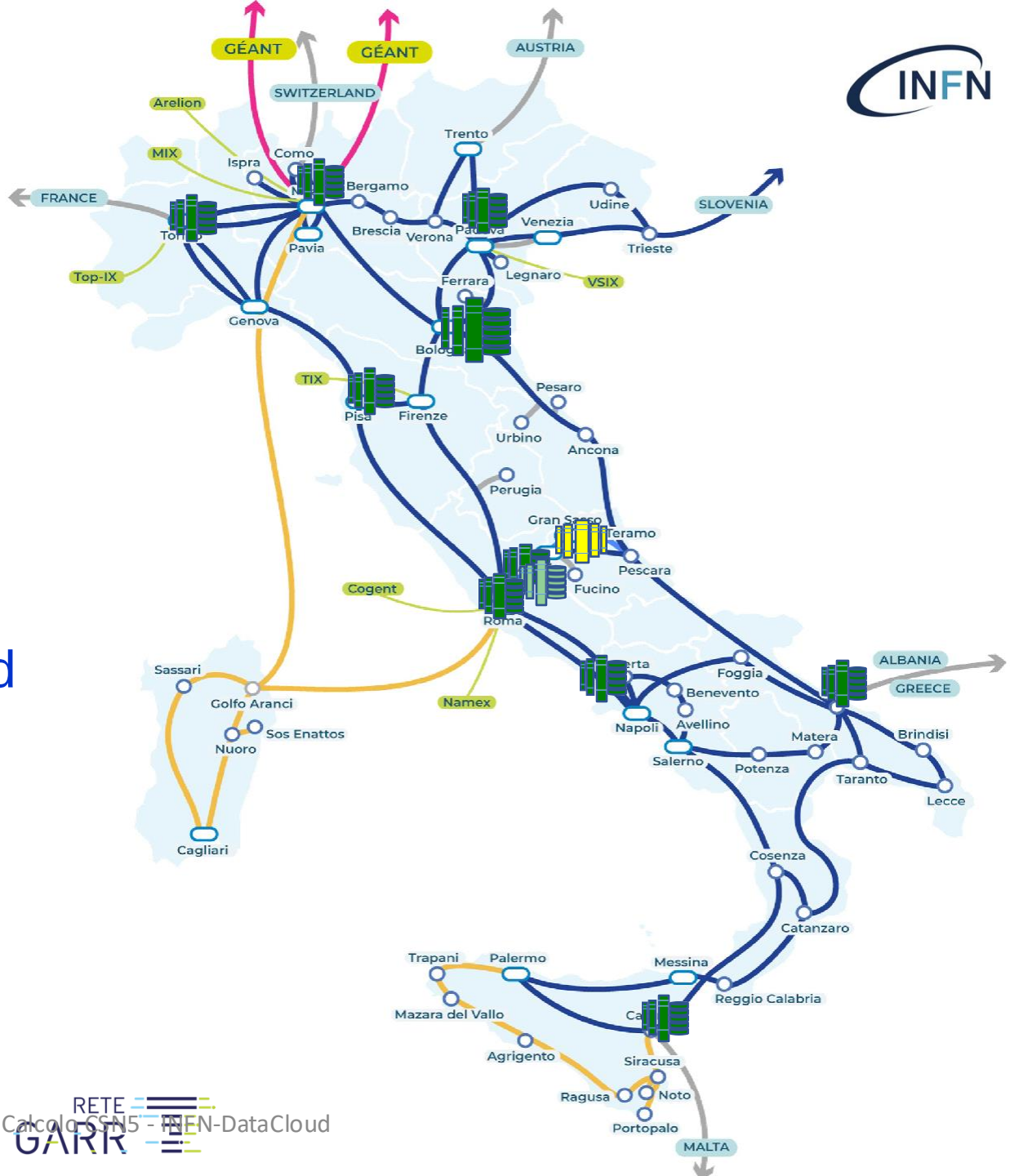


DataCloud addresses the needs of INFN research projects

- Internal projects: from CSN's
- External projects: regional, national and international projects, collaborations

The competences developed in the past years have brought to INFN visibility at national and international level

External projects have become more important



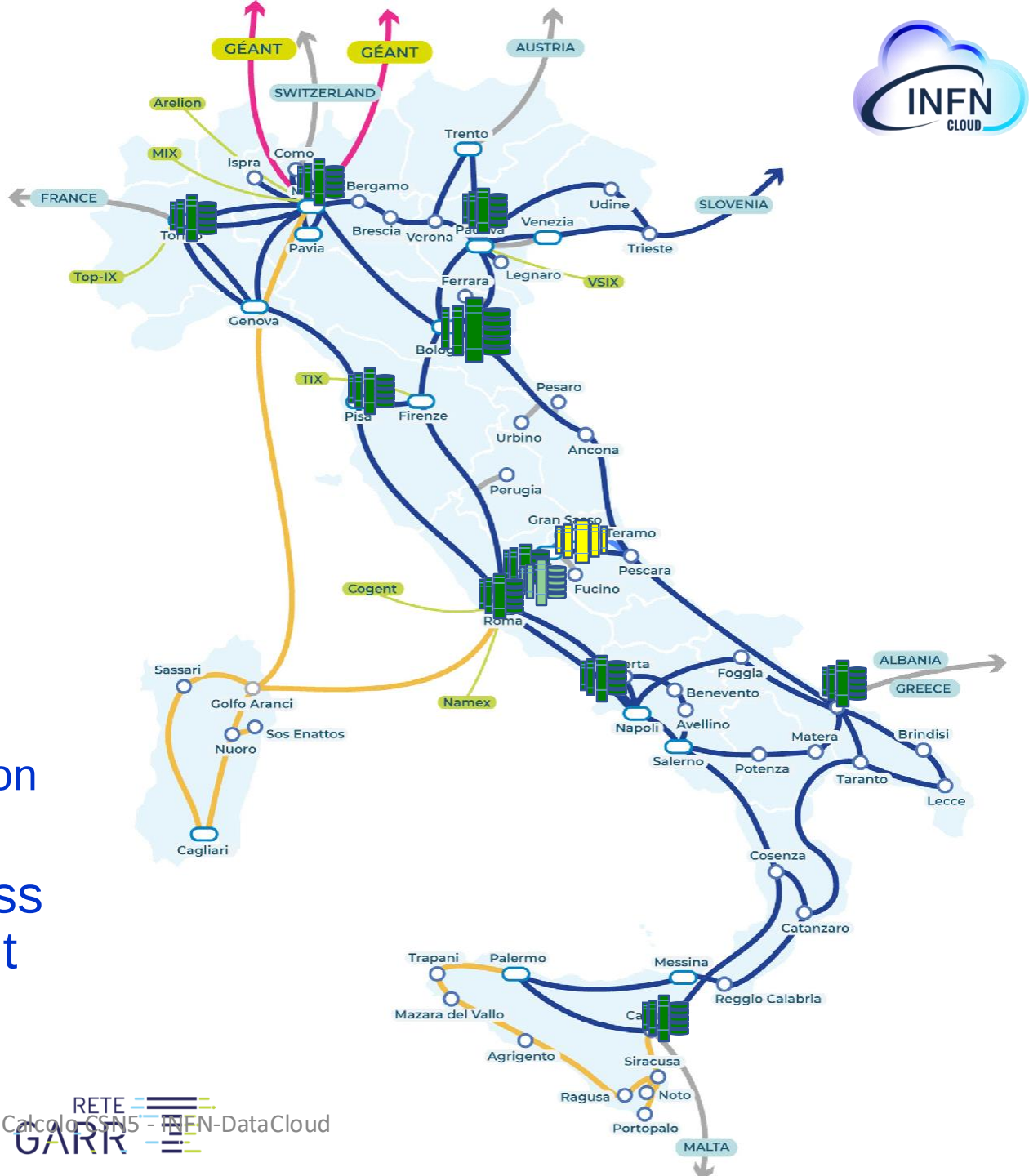
DataCloud is evolving into a Cloud Federation

Following the INFN Cloud model, resources are being made available through Cloud interfaces

- Inclusivity, through a lightweight federation model and the adoption of standards
- Ease of use, through the PaaS orchestrator and dashboard
- Flexibility, thanks to hybrid resource allocation mechanisms

Traditional (Grid and batch system) access remains as needed and when convenient

- E.g. through Virtual Kubelets, …



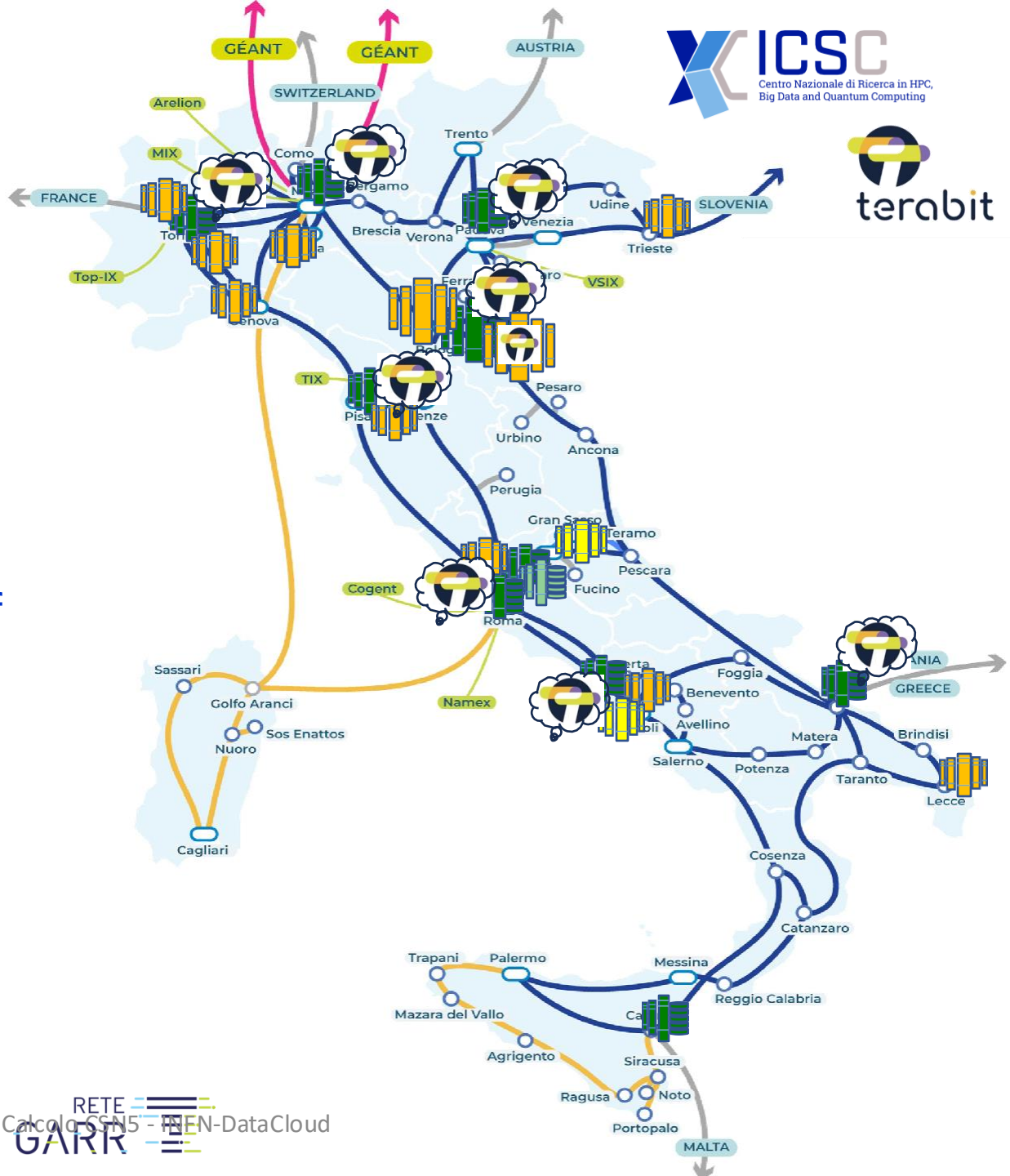
DataCloud is the basis for the Italian Cloud Federation

In the framework of the current NRRP projects, in particular ICSC and TeRABIT, INFN has a leading role in the creation of the Italian Cloud Federation

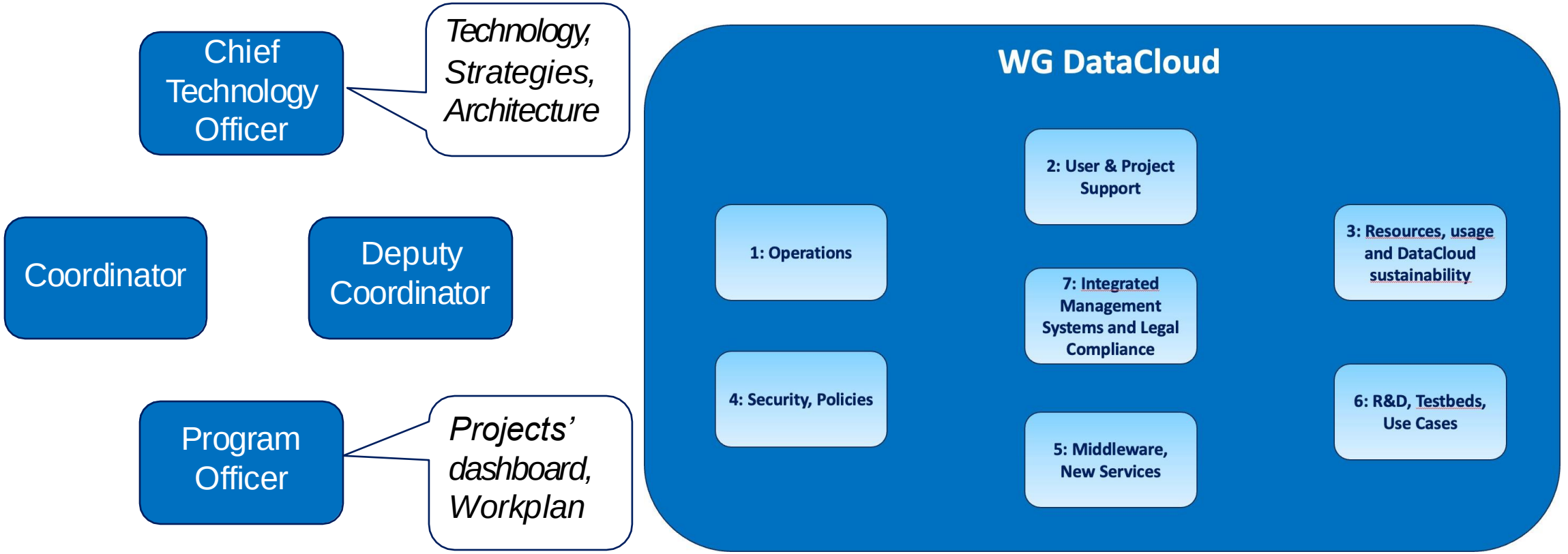
The goal is to access all Italian scientific computing resources through uniform interfaces

Main players: INFN, CINECA, GARR

But also: CMCC, ENEA, SISSA, IIT, UniTO, Sapienza, …



A new organization for DataCloud - structure



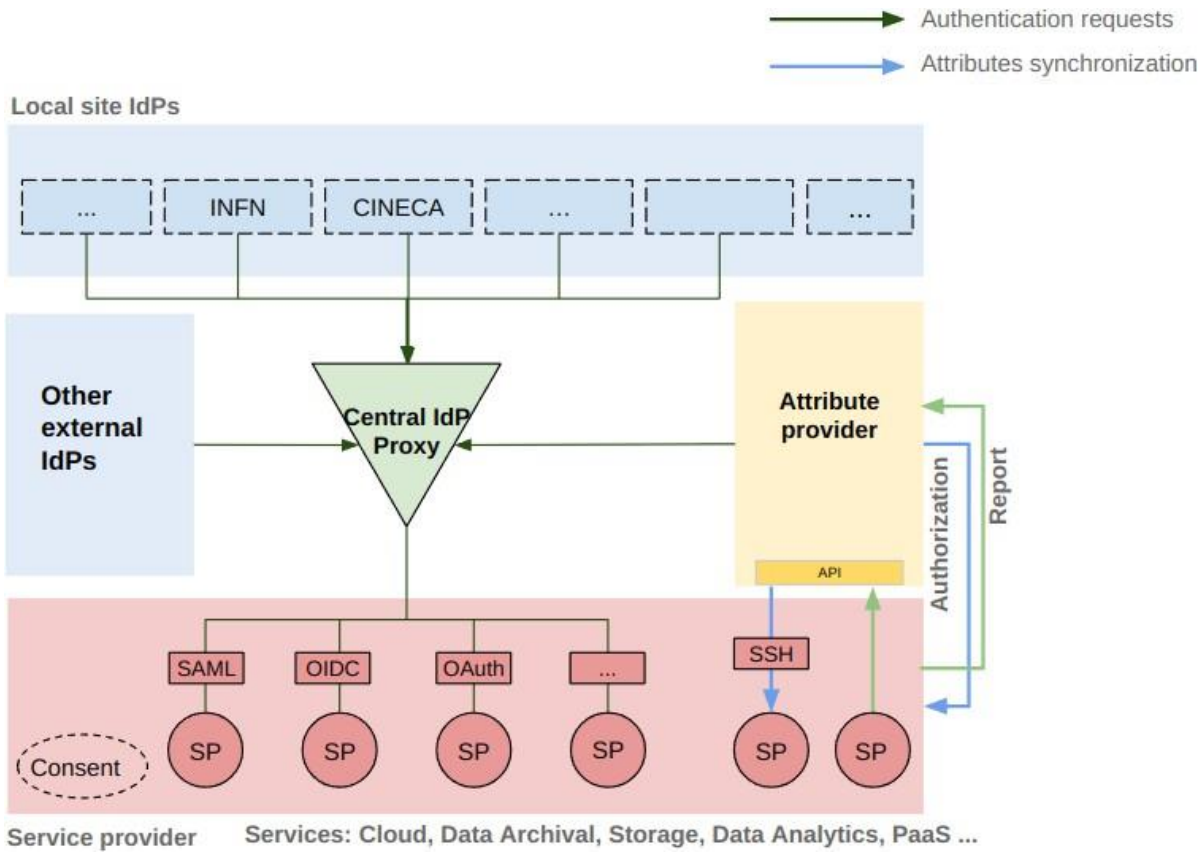
Inclusivity

The federation will include data centres that are already in production, and part of international communities

- The procedures for joining the federation must be non-intrusive
- Standard must be used whenever possible, and developed when missing

The federation will serve users of several fields and organizations

- The procedures for user's onboarding must be as simple as possible
  - E.g.: use of Identity Federations



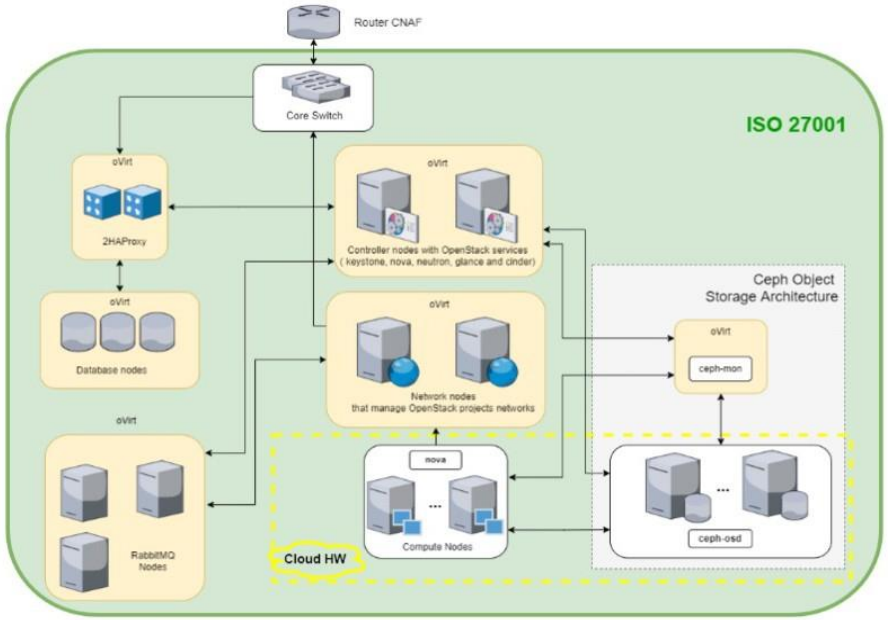
Flexibility

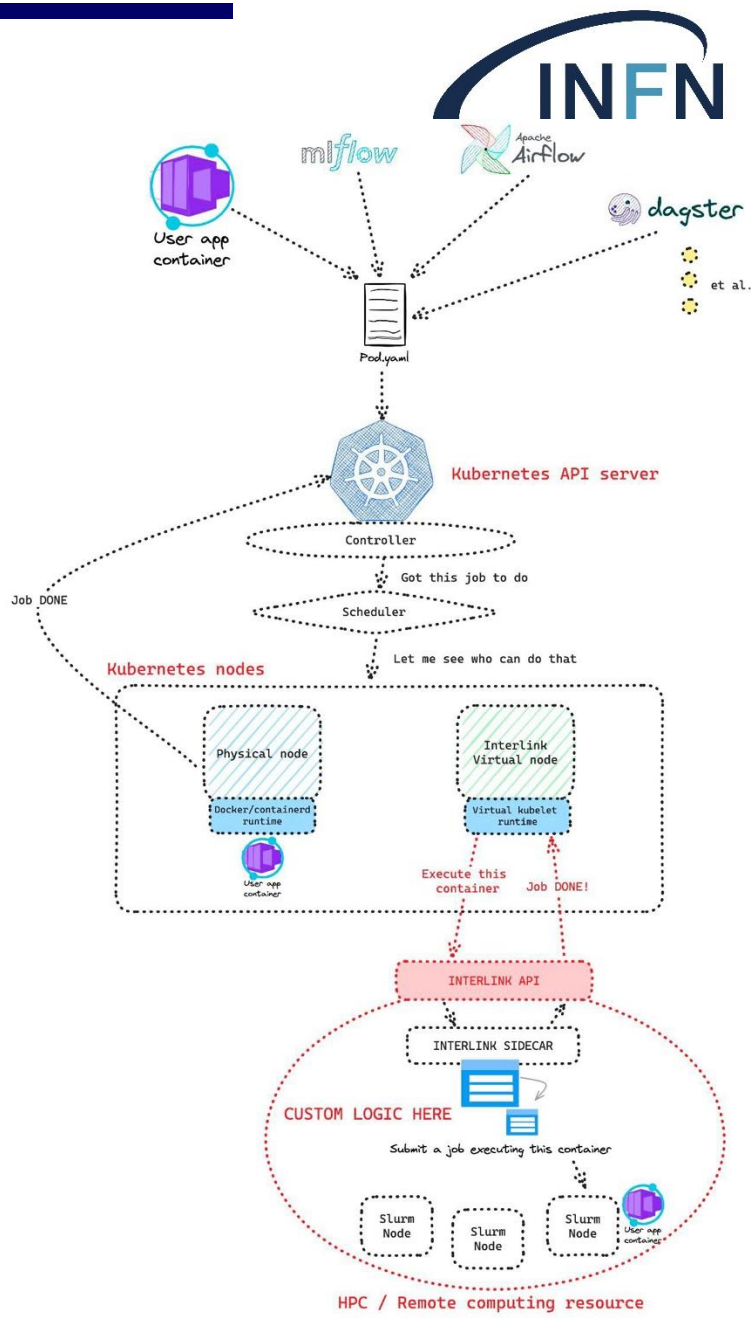
Support multiple access methods to the resources, oriented to:

a. Transparency and ease of use
b. Efficiency and effectiveness

Support application-specific requirements

E.g. enhanced privacy



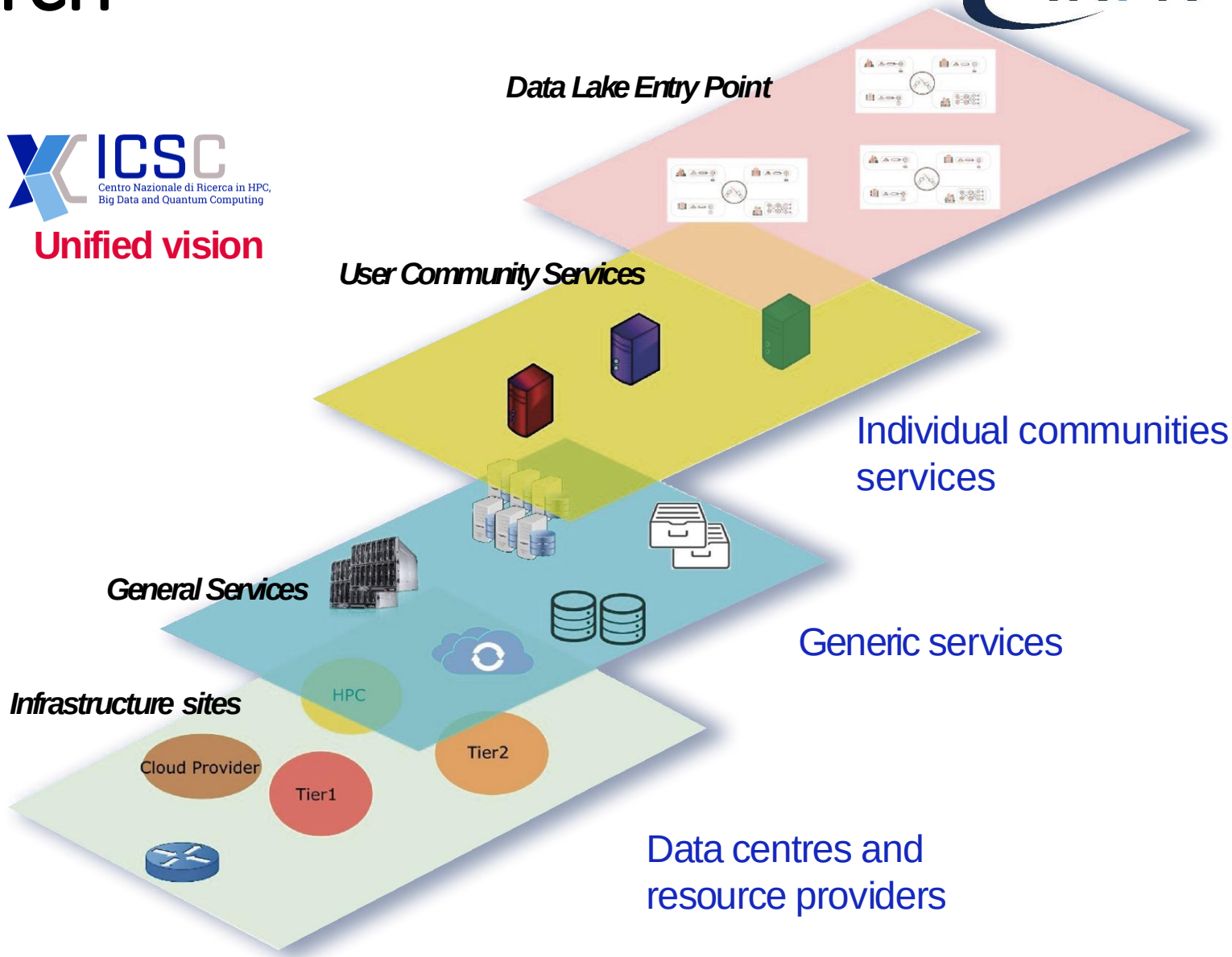


A data lake for research

- Existing infrastructures aggregation, upgraded and made available to scientific domains
- A dynamic model, where infrastructures and domains can also be temporary
- A clear separation between the physical and the logical levels
- A high-speed network interconnection to hide the actual resource locations
- A unified vision (when needed) of an Italian research data-lake

ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Unified vision



Individual communities services

Generic services

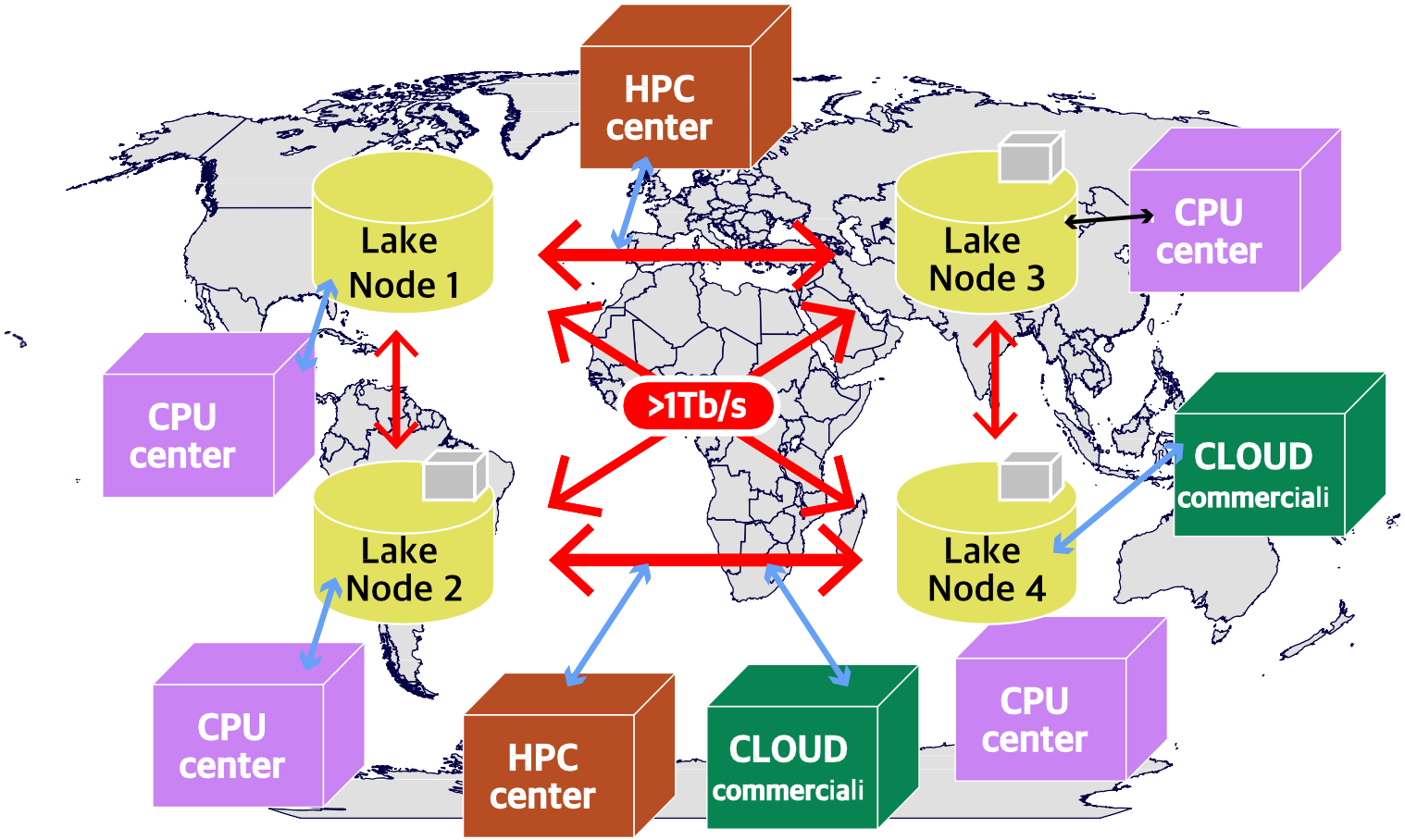
Data centres and resource providers

Data-centric model

Decouple storage and CPU

Storage nodes interconnected with high bandwidth network

Heterogeneous computing nodes can access data wherever they are

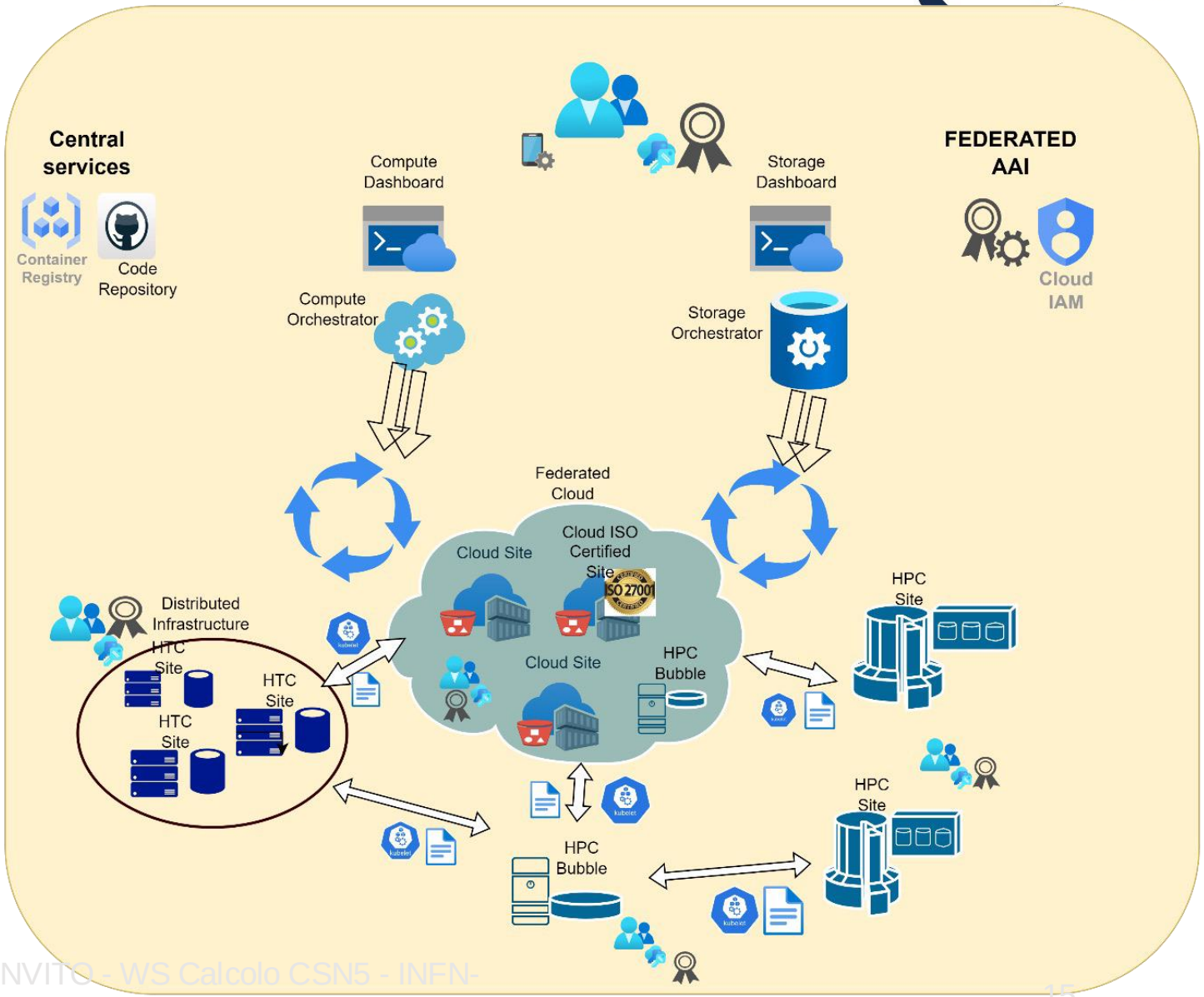


Etherogeneity

Integration of a diverse set of resources, providers, and solutions

We call it:

Computing *continuum* from Edge, to Cloud, to HPC



INFN Cloud

ARCHITECTURAL FOUNDATIONS

NO VENDOR LOCK-IN

Open-source, vendor-neutral architecture

FEDERATION

of existing Cloud infrastructures for both compute and data

DYNAMIC ORCHESTRATION

of resources via the INDIGO PaaS Orchestrator

CONSISTENT AUTHN/AUTHZ

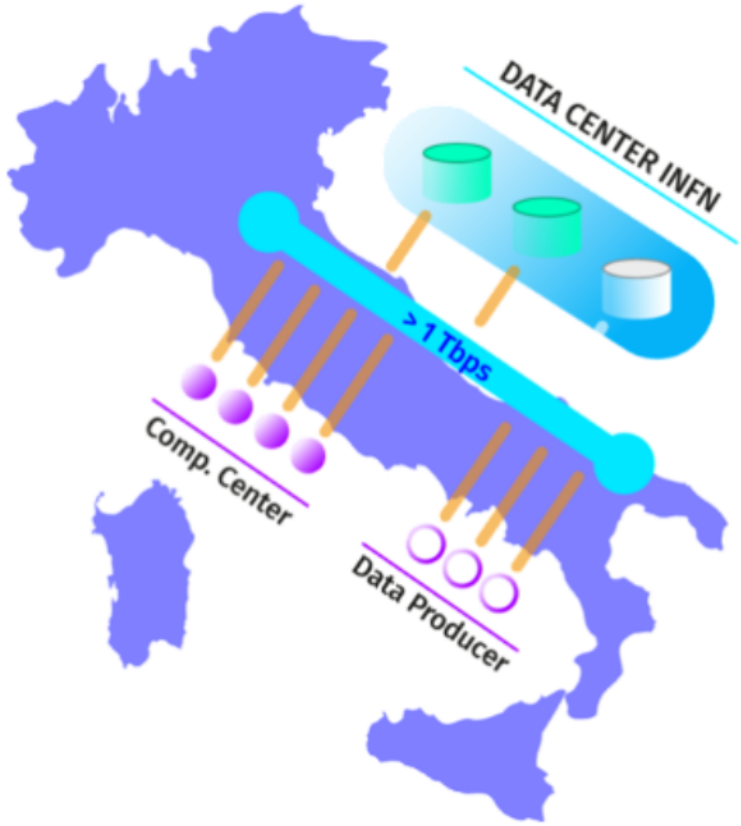
at all cloud levels via OpenID-Connect/OAuth2

Federation

INFN CLOUD IS DESIGNED AS A FEDERATION OF PRE-EXISTING INFRASTRUCTURES

- The **Backbone** of the INFN Cloud is made up of two closely linked federated sites, BARI and CNAF.
- A scalable set of **satellite sites**, geographically distributed across Italy and loosely coupled, expand the resources offered by the backbone.

INFN Cloud core services and some centralised, fully managed, high-level services are hosted on the Backbone. This allows us to leverage **high-availability** and **disaster recovery** capabilities to ensure that these critical services are always available and operating at peak efficiency.

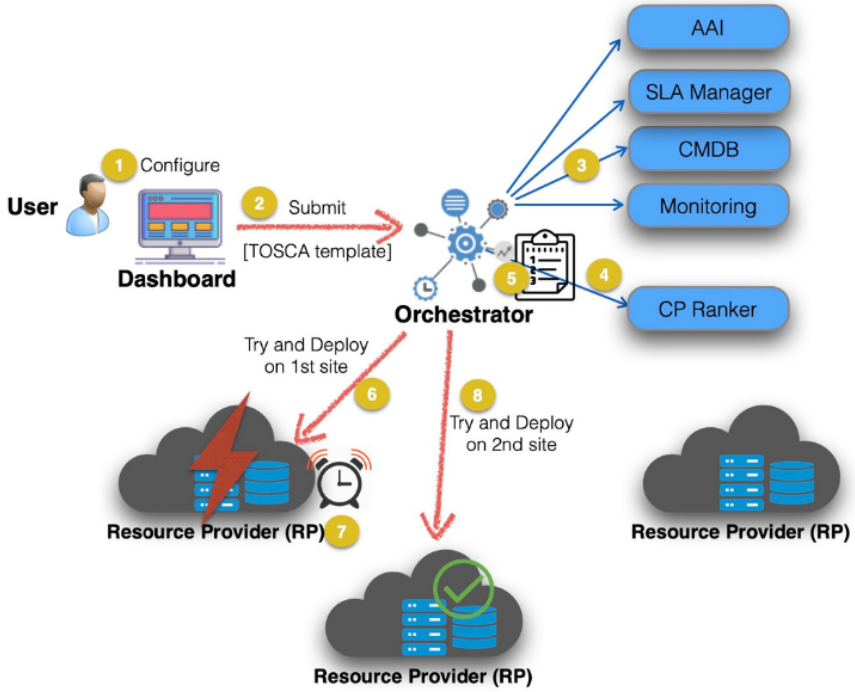


Middleware

THE FEDERATION MIDDLEWARE

The INDIGO PaaS Orchestrator enables the federation of distributed and heterogeneous compute environments: clouds, docker orchestration platforms, HPC systems.

- **Smart scheduling** → Automatic selection of the best provider
  - based on compute/storage requirements vs provider capabilities including the following criteria:
    - Resource quotas (SLA)
    - Monitoring data
    - Support for specialized hardware (GPU, Infiniband)
    - Data location
- Support for **hybrid deployments** and network orchestration
- **Client interfaces** for advanced users (REST APIs, CLI, python bindings) and end-users (web dashboard - no skills required)



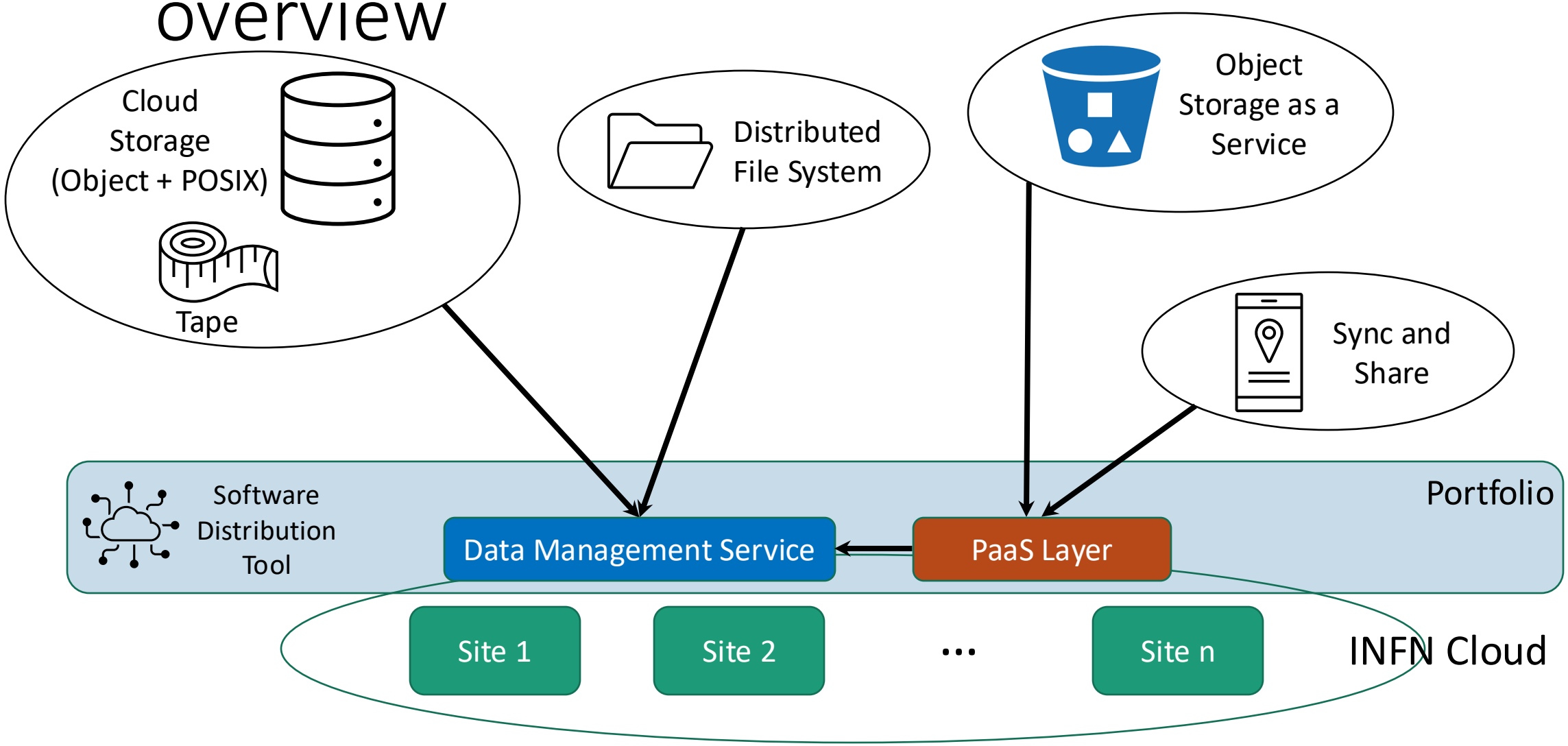
INFN-Cloud services

Centralized services (SaaS):
•INFN Cloud Registry service (Harbor)
•INFN Cloud object storage service (based on rados-Gateway)

PaaS services:
•Virtual machine
•Docker run
•Docker compose
•Kubernetes cluster
•HTCondor mini
•HT Condor cluster
•Jupyter with persistence for Notebooks
•Jupyter + Matlab (with persistence for Notebooks)
•Spark + Jupyter cluster
•Working Station for CYGNO experiment
•Computational enviroment for AI_INFN
•Elasticsearch and Kibana
•INDIGO IAM as a Service
•Sync&Share aaS

IaaS services:
•Start and Stop
•Hostname choice
•Manage VM ports

Large scale datalake: Architectural overview



Implementation: adaptive solutions

THE SERVICE IMPLEMENTATION STRATEGY

Keyword:
Service Composition

The employed strategy is based on the **Infrastructure as Code paradigm**.
Users describe "**What**" is needed rather than "**How**" a specific service or functionality should be implemented.
The adopted technologies enable a Lego-like approach: **services can be composed and modules reused** to create the desired infrastructure.

OASIS TOSCA

TOSCA is used to model the topology of the whole application stack

ANSIBLE

Ansible is used to automate the configuration of the virtual environments

docker

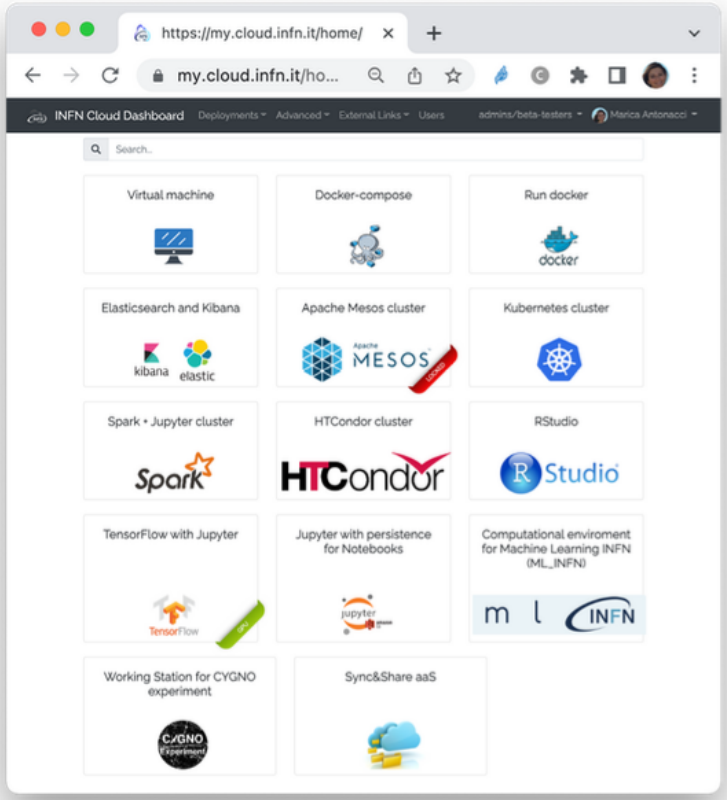
Docker is used to encapsulate the high-level application software and runtime

Front-ends

THE PAAS DASHBOARD

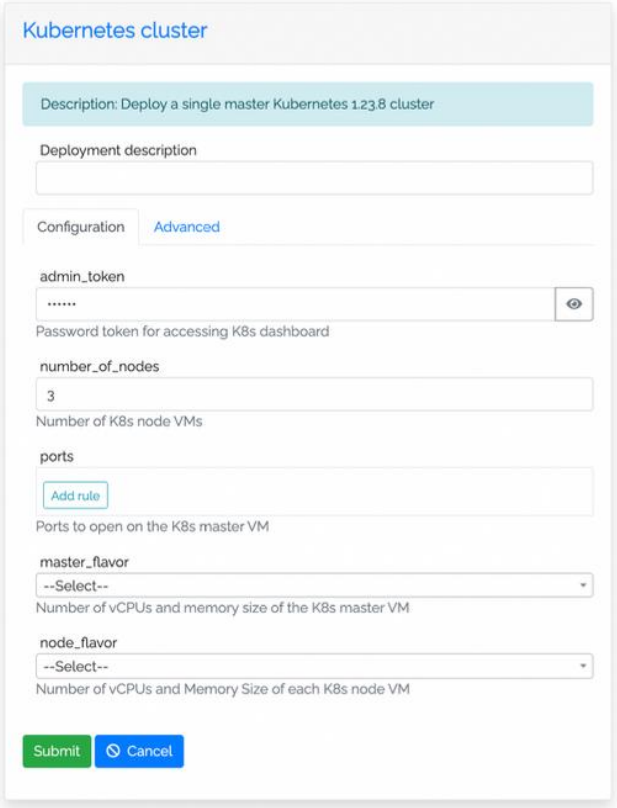
The INDIGO PaaS Dashboard is a web-based user interface that enables users to manage and monitor their deployments without requiring any TOSCA knowledge.
The dashboard hides all technical details and provides an intuitive interface for managing service deployments.

- OpenID-Connect Authentication
- Multi-tenancy
- Secrets management (via Vault integration)
- Dynamic view of service catalog (depending on the user group membership)



Self-provisioning

REQUEST SERVICES WITH JUST A FEW CLICKS

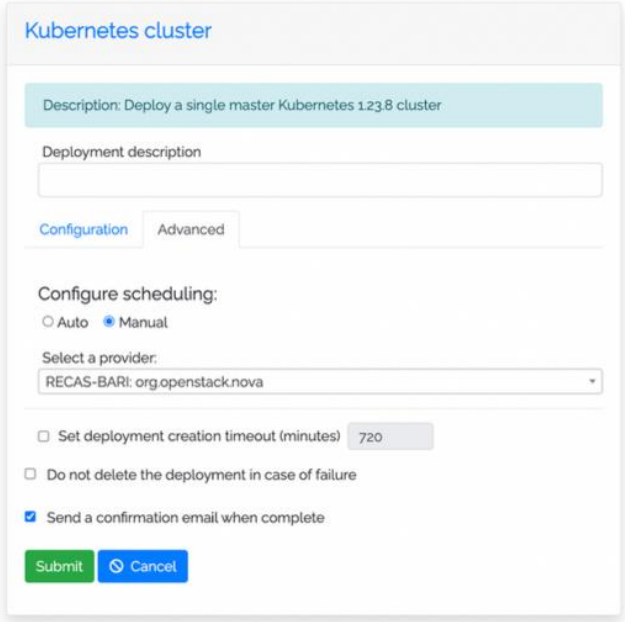


Customize your deployment

through the deployment input parameters

Choose the Scheduling strategy

- automatic: let the Orchestrator select the best provider
- manual: choose the provider from the drop down menu automatically created by the Dashboard with the list of providers returned by the SLA Manager service



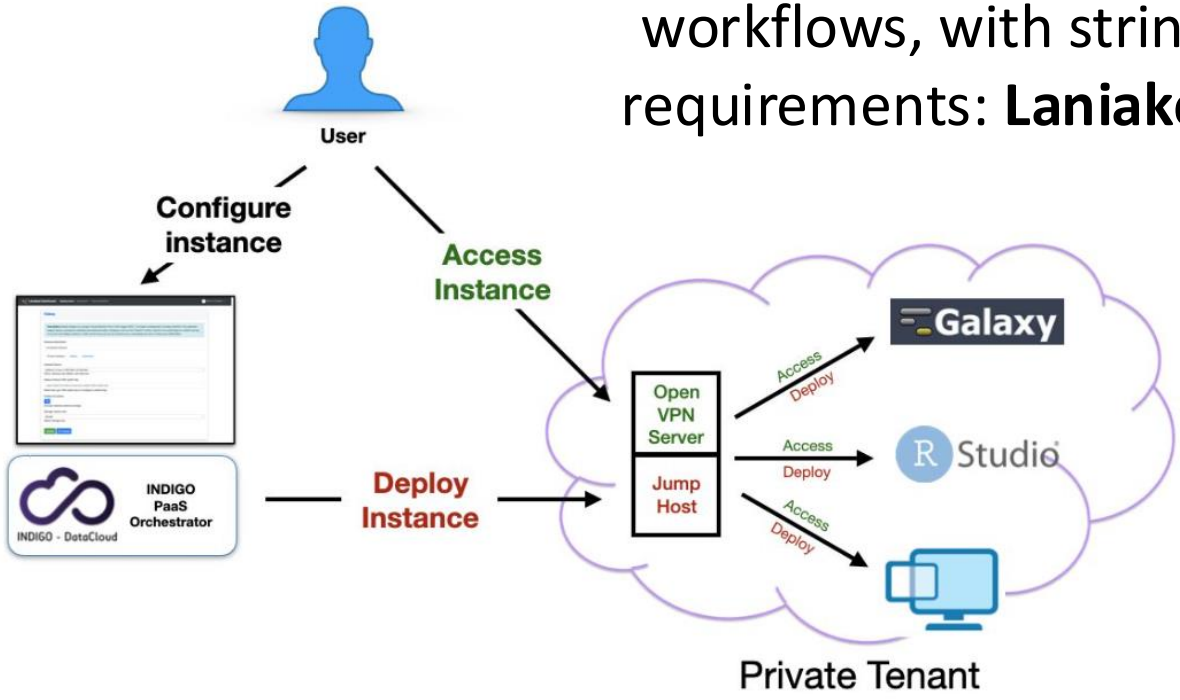
Security: multiple isolation levels

Deployments under VPN

VPN isolated environments - Automatic deployments of virtual environments on private networks.

Isolation is reached using Tenant and security groups properties, granting the access only through VPN authentication.

User authentication to the VPN using the same Laniakea credentials.

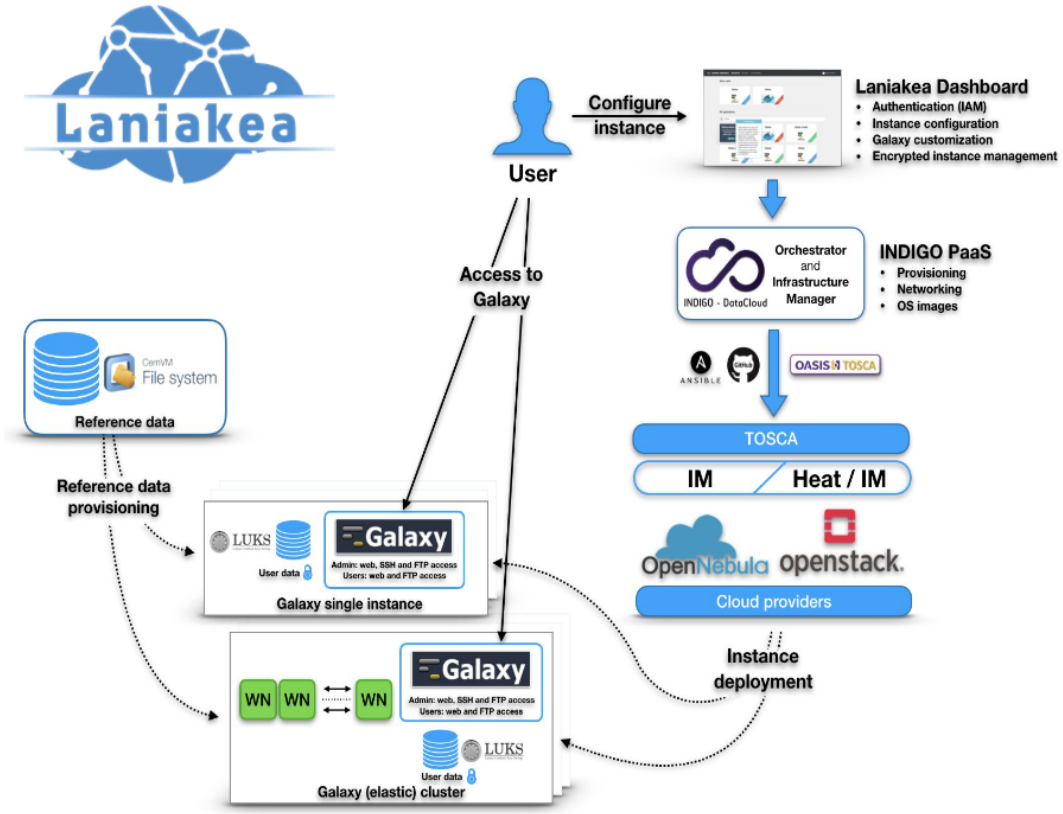


An *example* of the **Service Composition** approach: extending INFN Cloud to support complex workflows, with stringent security requirements: **Laniakea** (Elixir Italy)

41

(Some of) The gory details

Laniakea architecture



- **Dashboard -** User friendly access to configuration and and launch of a Galaxy instance.
- **IAM -** Authentication and Authorization system.
- **INDIGO PaaS -** Galaxy automatic deployment.
- **Cloud Providers -** (INFN) ReCaS-Bari and others.
- **Persistent storage** - With/without encryption.
- **Reference data availability** - With CERN-VM FileSystem.
- **CLUES -** Elasticity manager.

EPIC Cloud

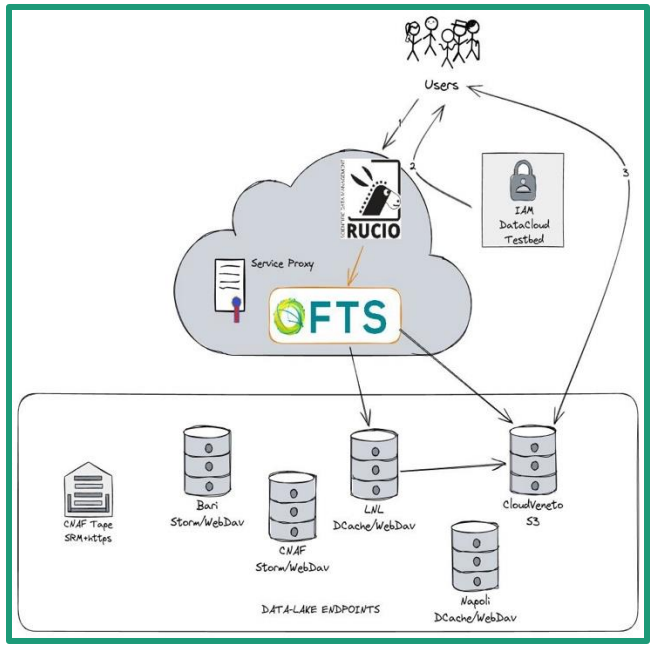
Enhanced Privacy and Compliance Cloud – The INFN Cloud partition for personal and confidential data processing

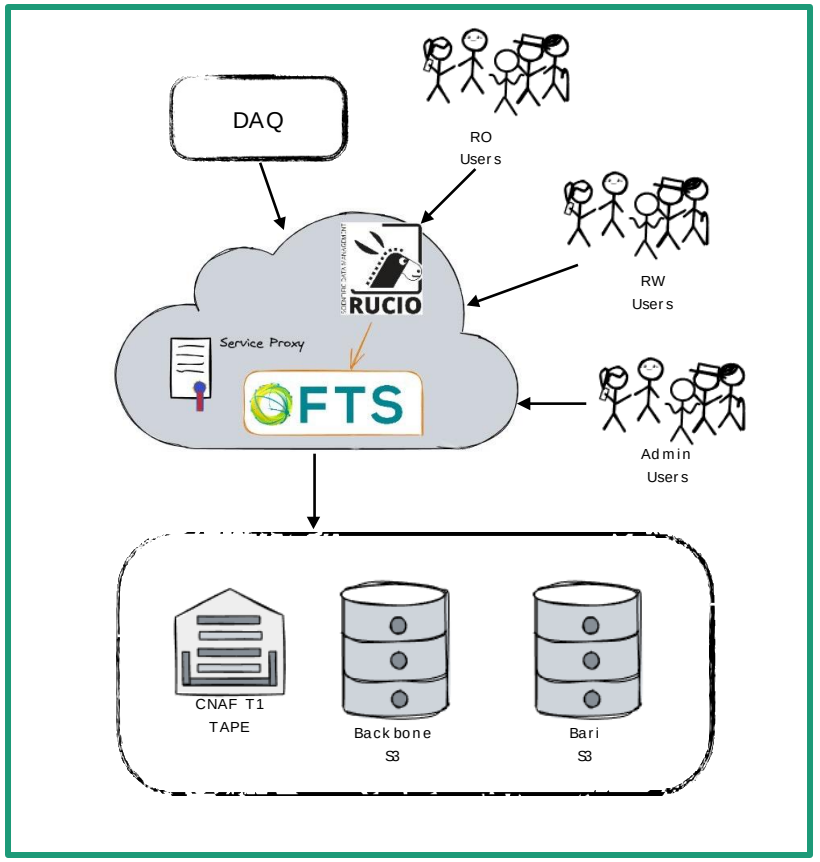
- The GDPR states that Clinical and medical data (for instance, genomic) is personal data; i.e., it fits in the Art.9 special categories of personal data.
  - Genomic data is mostly impossible to be anonymized → GDPR shall always be applied
  - ISO/IEC 27001 is the main certification mechanism compliant with GDPR requirements (Art. 43, 58, 63)
- In order to comply with the requirements of health research projects INFN is involved in, we created **a region of the INFN Cloud infrastructure**, applied specific organizational and technical security measures, and certified it ISO/IEC 27001, 27017, 27018.
  - This is **EPIC Cloud**: a reference Cloud implementation for the treatment of sensitive data at INFN.

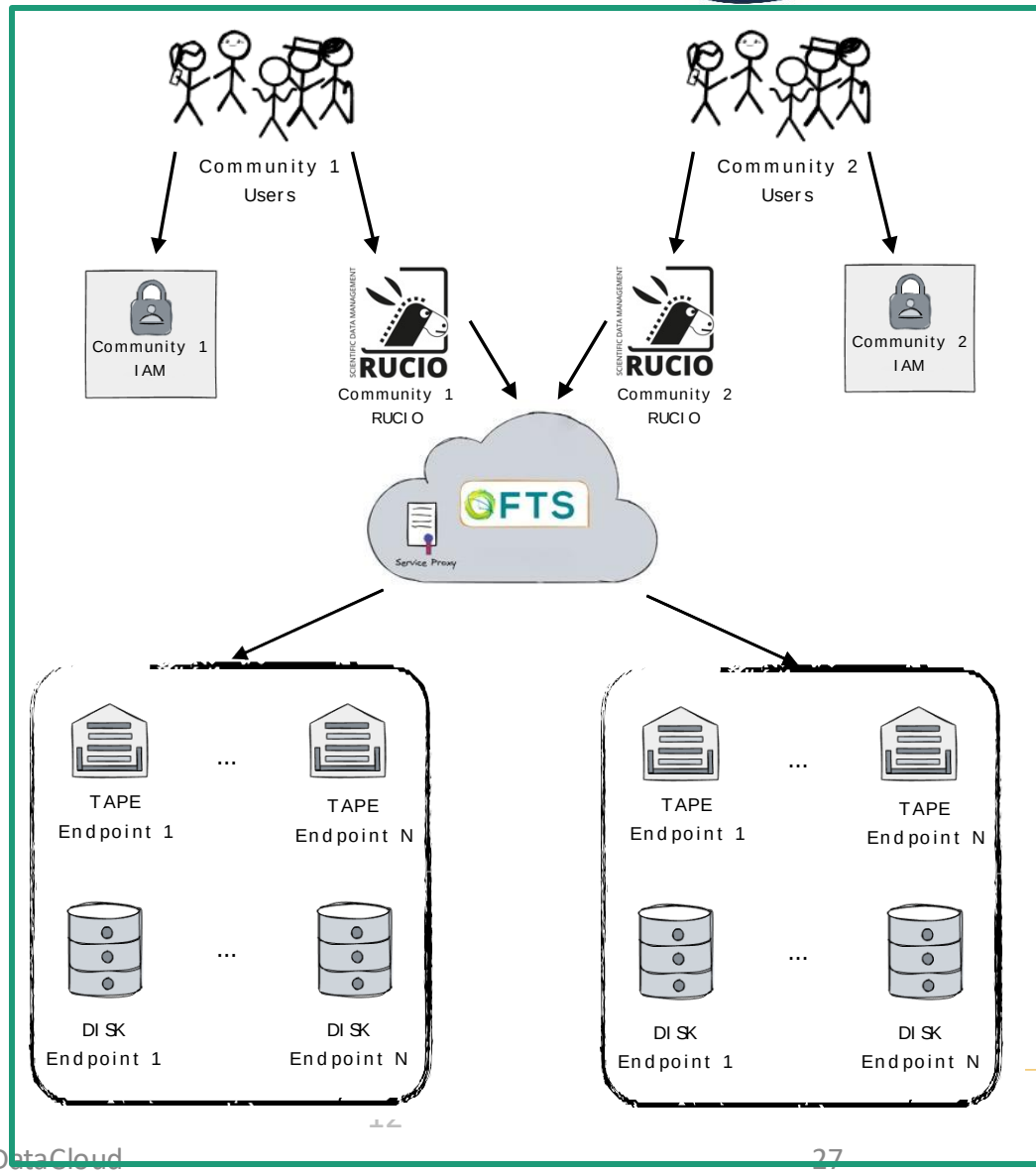
https://indico.egi.eu/event/6441/sessions/5209/#20241001

From the Data Controller side, the fact that EPIC Cloud is ISO-certified is a way to demonstrate that processing is performed in accordance with the GDPR.

Data Lake: overview







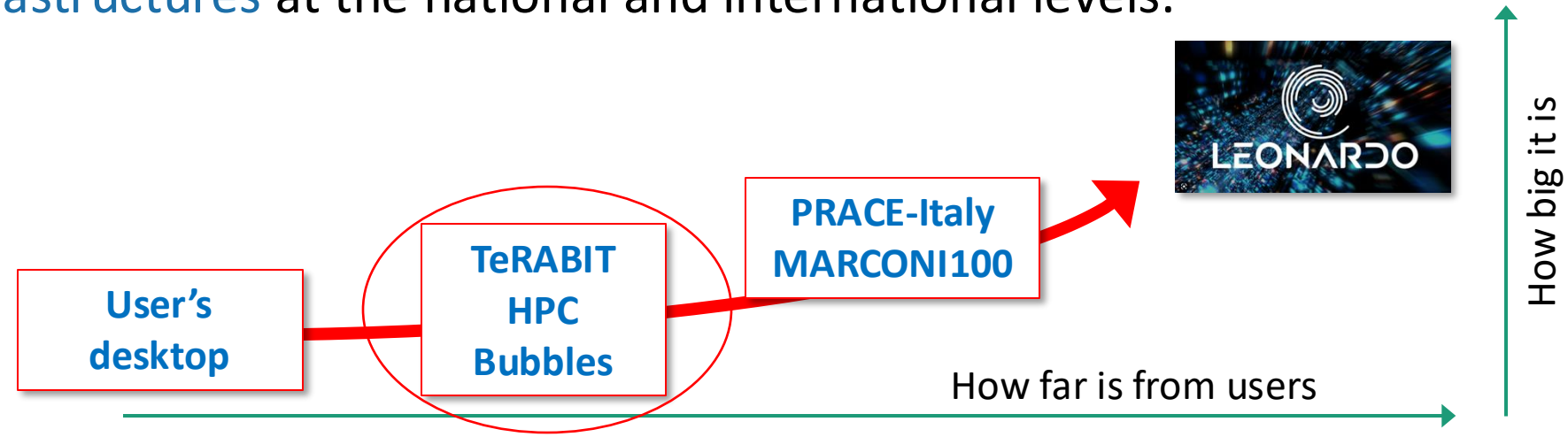
Recap so far

INFN Cloud is the architectural foundation for the **evolution of the distributed infrastructure** managed and operated by INFN (HPC-BD-AI).

1. This is true for **all** our computing-related engagements with PNRR projects, including ICSC, TeRABIT, DARE, and others.
2. This covers **both hardware acquisitions and the Cloud service portfolio**, in accordance with our service composition architecture. Concretely, this means that we are:
   - Expanding hardware resources across the entire INFN DataCloud;
   - Extending the number of ISO-certified DataCloud regions in Italy;
   - Increasing the solutions offered by INFN Cloud.

An example: the “HPC Bubbles”

- As part of the TeRABIT PNRR project, we are implementing several highly innovative concepts:
  - Availability of scalable HPC resources and services on INFN Cloud through Cloud-native interfaces at the IaaS, PaaS and SaaS levels.
  - Strong interaction between network, data and HPC/HTC resources.
  - Communication and federation between the HPC Bubbles and other HPC infrastructures at the national and international levels.



An example: the “HPC Bubbles”

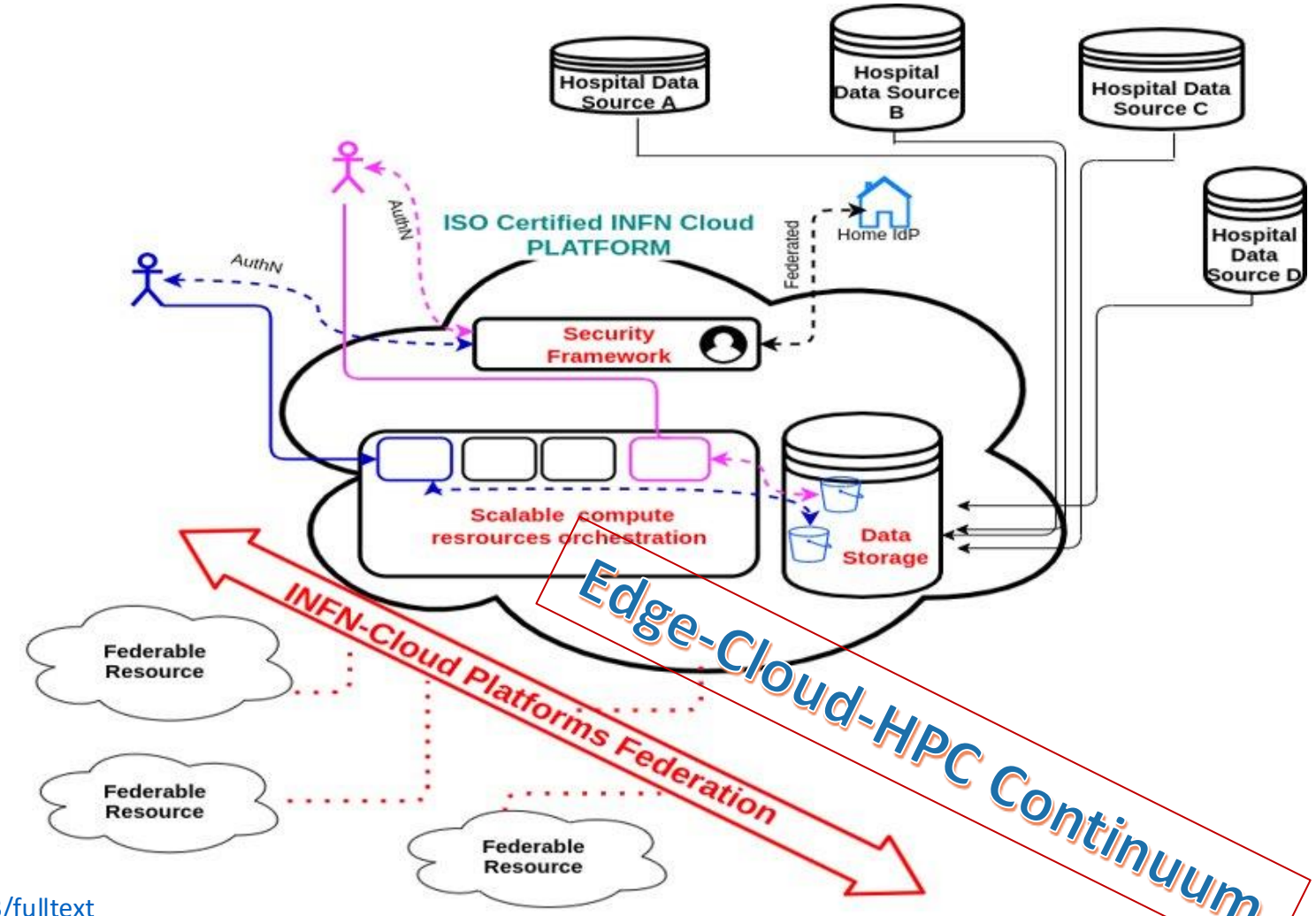
CPU Node	192 real core 1.5TB RAM DDR5 IB NDR 400G 20TBL (SSD)
GPU Node	Same CPU + 4x NVIDIA H100 SXM5 con minimo 80GB e memoria HBM2e
FPGA Node	32core RAM 768GB DDR5 IB NDR 440G 4 x XILINX U55C o 4 x TerasicP0701
Storage node (CEPH Bricks)	64 core fisici 1TB RAM DDR5 384 TBL HDD + 25.6 TBL NVMe

From 8 node per bubble (infiniband connected) => min 3000 HT Cores per single cluster

The goal: a federated datalake

Multiple ways to ingest and process data are possible. For example, to handle sensitive data (e.g., in the nation-wide Health Big Data project), we are working on supporting these options:

1. **Central harvesting** of data generated remotely
2. **Edge-level anonymization**, followed by central ingestion and analysis of data
3. **Edge-level feature extraction**, followed by central ingestion and analysis of features
4. **Federated learning** based on edge-level training, followed by publishing of the trained methods and by inference performed either centrally or at other edge locations.



https://www.physicamedica.com/article/S1120-1797(21)00320-3/fulltext

Recap

1. Due to <u>increased needs</u> and <u>substantial new opportunities</u>, INFN is expanding its computing infrastructure and services. The overall technological approach is to **abstract from where resources are, leveraging *aaS* models** to build a Cloud-native, [trans-]National, federated structure.
   - Some of the **challenges**: distributed data management, cloud abstraction & policies, opportunistic extensions, proper handling of sensitive data, security, HPC in the Cloud.

2. In doing so, INFN has the ambition to **create, operate and develop a vendor-neutral, open, scalable and flexible "data lake"** that serves much more than just INFN users and experiments.
   - This will become a **key asset** for fundamental, applied and industrial research in Italy and beyond.