



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani

PIANO NAZIONALE  
DI RIPRESA E RESILIENZA

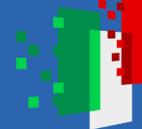


Centro Nazionale di Ricerca in HPC,  
Big Data and Quantum Computing



# Roadmap generale per l'integrazione delle piattaforme cloud Spoke8 e Spoke0 Giacinto DONVITO

Spoke 8 – Bologna, 25/06/2024

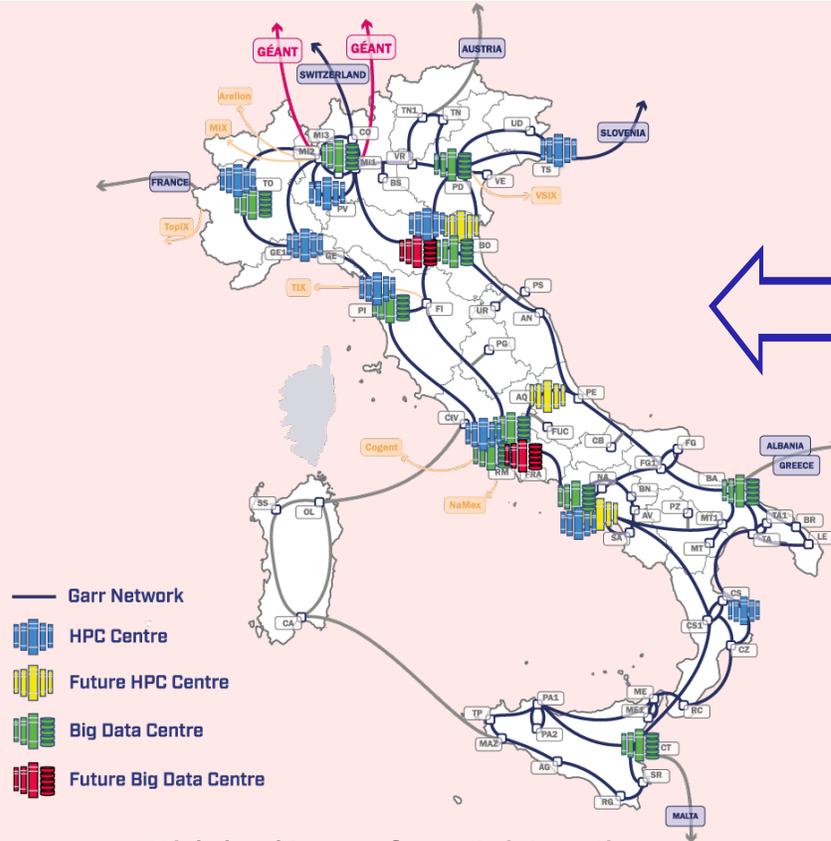


# Il centro nazionale HPC, BD e QC



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

## 0 SUPERCOMPUTING CLOUD INFRASTRUCTURE



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

ISTRUZIONE E FORMAZIONE, IMPRENDITORIALITÀ, TRASFERIMENTO DI CONOSCENZE, 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

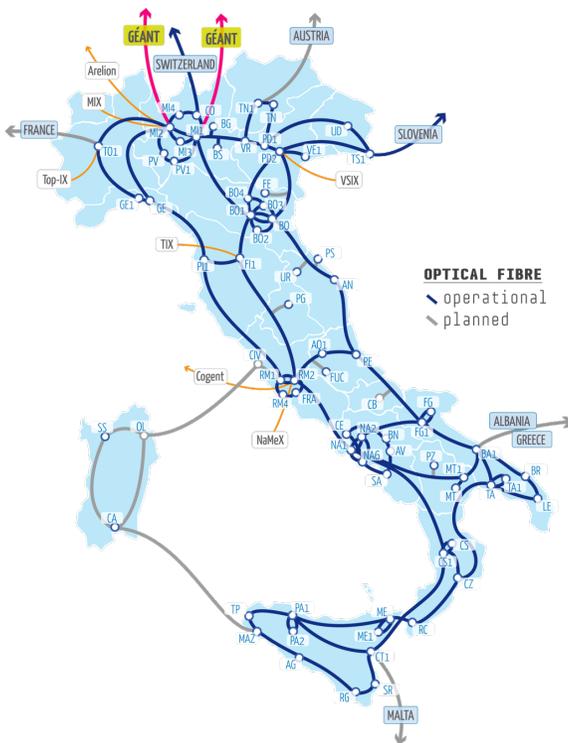
ICSC include  
10 spokes tematici  
1 spoke infrastruttura



# Le infrastrutture di ricerca di TeRABIT



## GARR-T



EuroHPC Joint Undertaking



PARTNERSHIP FOR ADVANCED COMPUTING IN EUROPE

## PRACE-Italy



Galileo100 – cluster HPC Ospitato al CINECA - Bologna



WLCG Worldwide LHC Computing Grid

## HPC-BD-AI - INFN Cloud





# L'infrastruttura di calcolo nazionale

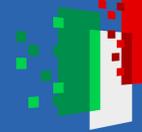
ICSC e TeRABIT prevedono la creazione di una infrastruttura cloud nazionale

Lo scopo è accedere alle risorse in modo trasparente ed efficiente

Gli attori principali sono: **INFN**, **CINECA**, **GARR**

Ma anche: CMCC, ENEA, SISSA, IIT, UniTO, Sapienza, ...



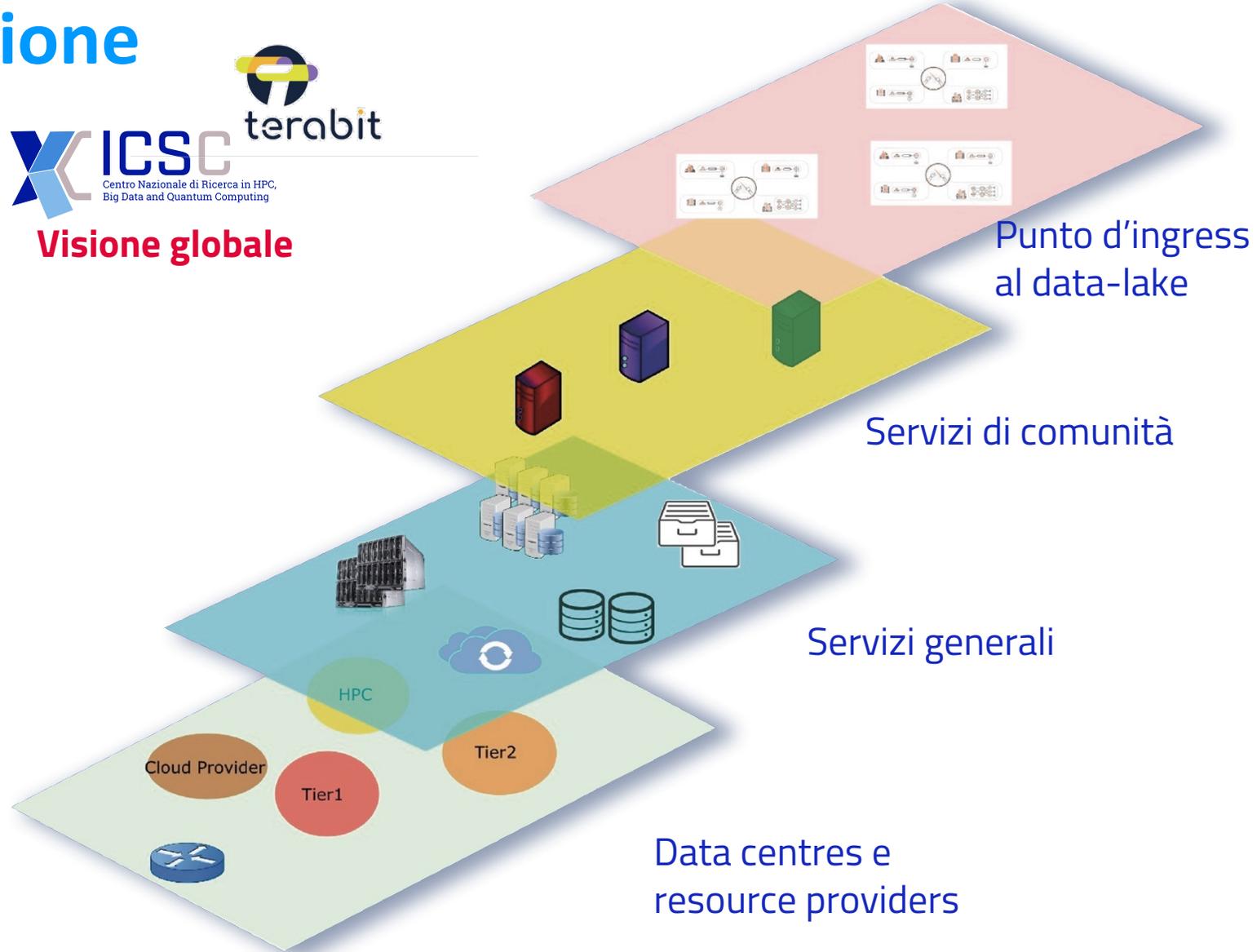


# I principi della federazione

Inclusività, attraverso una federazione "leggera" e l'adozione di standard

Facilità d'uso, attraverso un orchestratore PaaS e una dashboard

Flessibilità, grazie a meccanismi ibridi di allocazione delle risorse

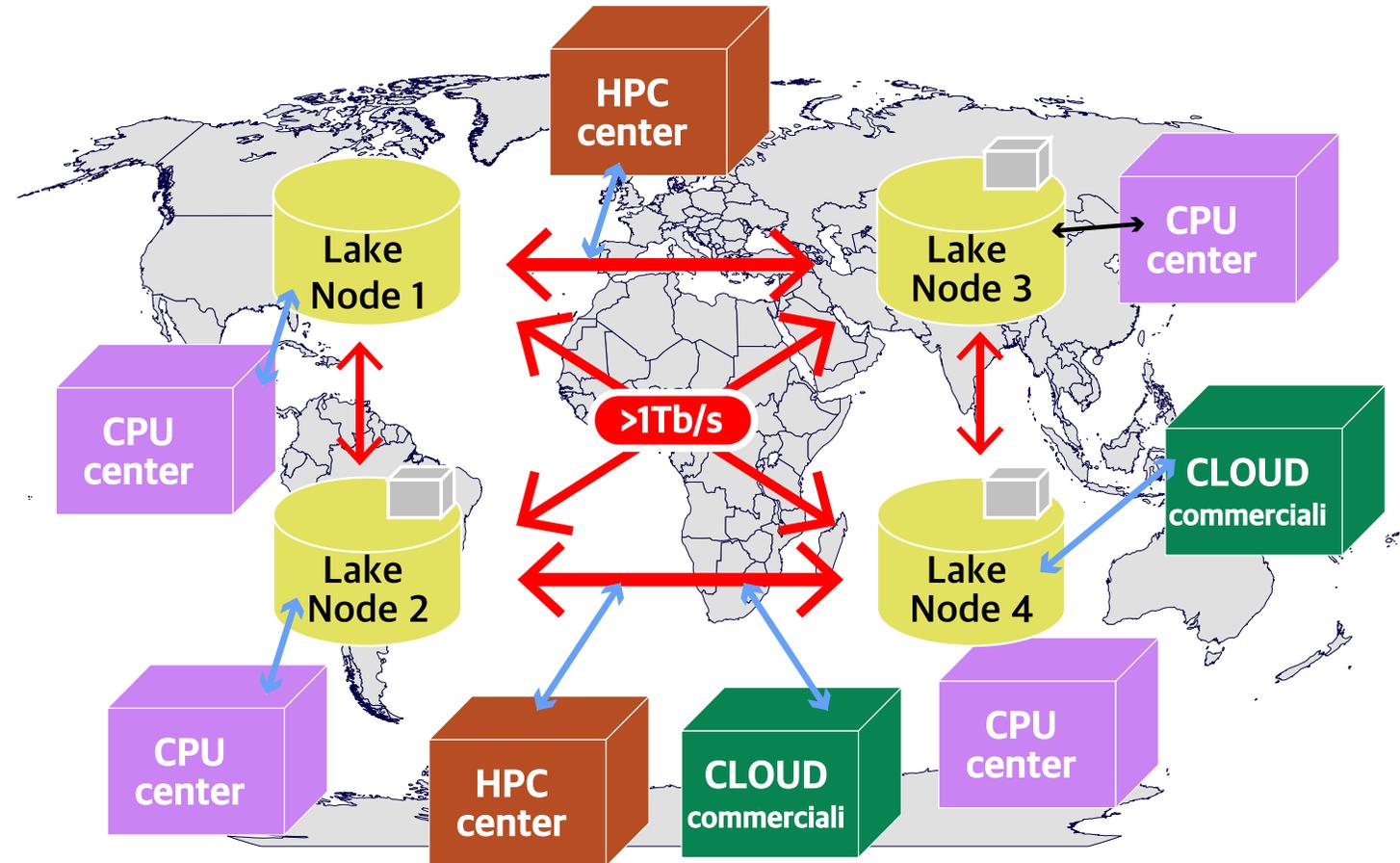


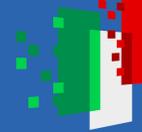
# Modello data-centrico

Disaccoppiamento di storage e CPU

Nodi storage interconnessi tramite una rete a banda larga

Nodi eterogenei possono accedere ai dati in modo indipendente dalla locazione





# Inclusività

La federazione includerà data centres che sono già in produzione, e parte di comunità internazionali

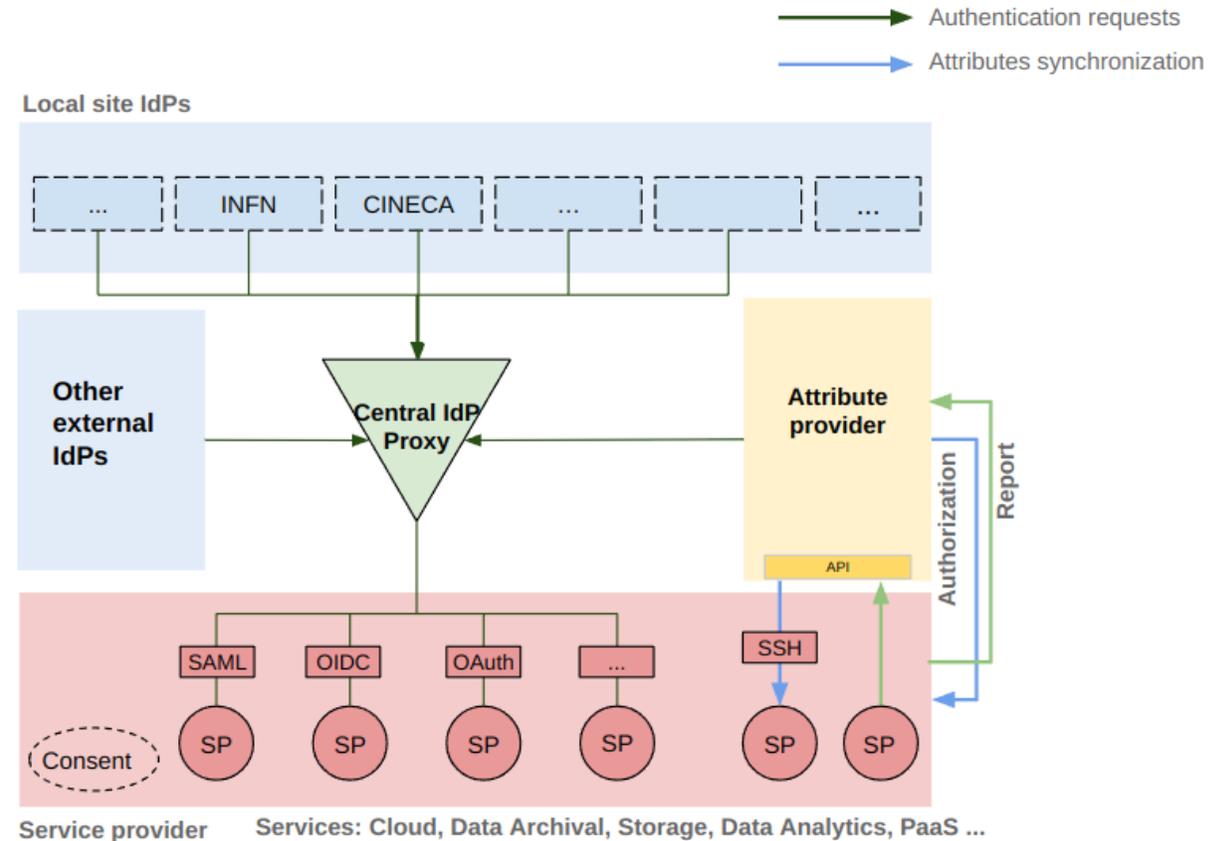
Le procedure per entrare nella federazione devono essere non intrusive

Bisogna usare standard quando possibile, e sviluppati quando non ci sono

La federazione servirà utenti di diverse organizzazioni in diversi campi

Le procedure di ingress devono essere il più semplice possibile

Ad esempio tramite l'uso di federazioni di identità



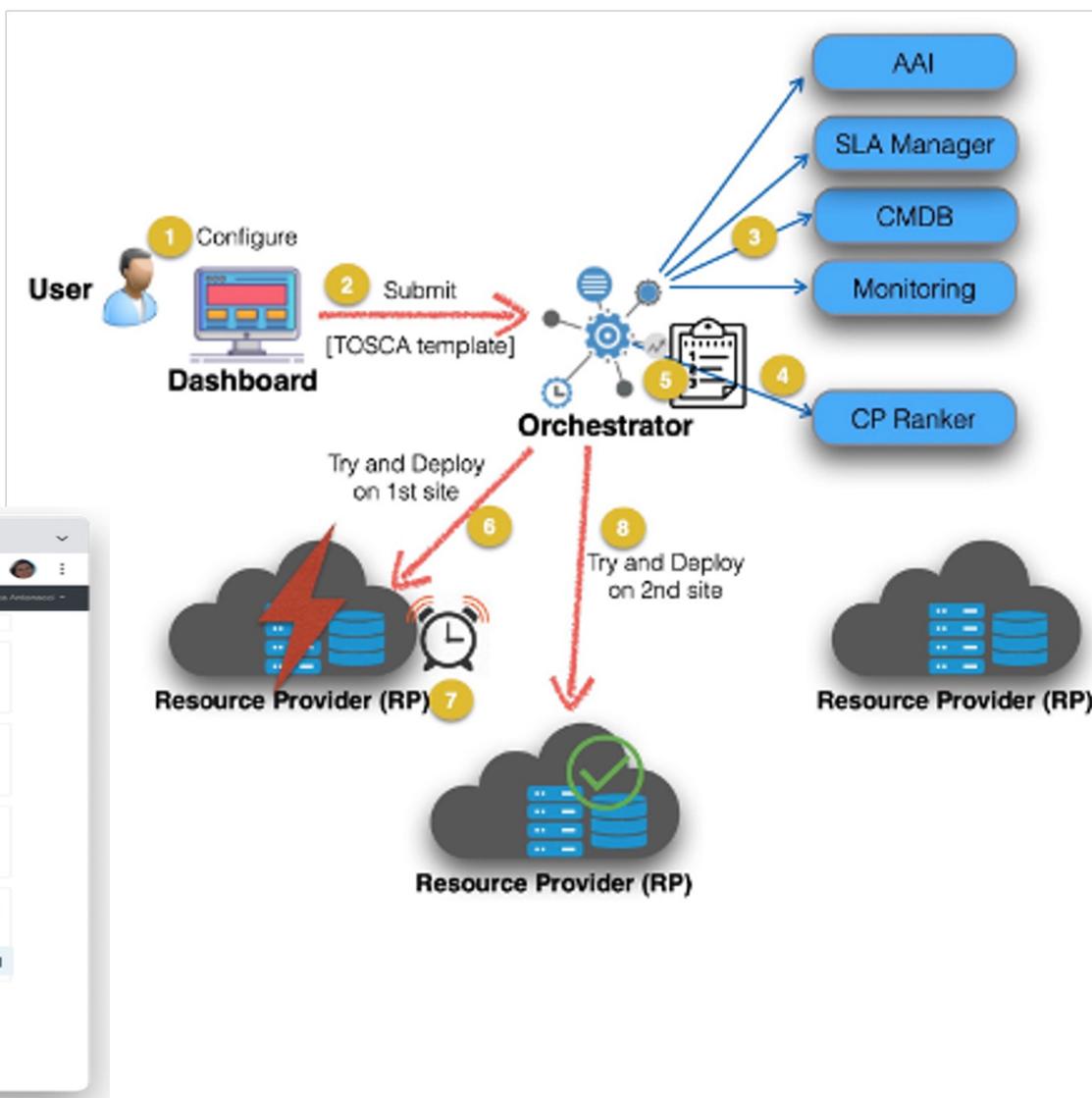
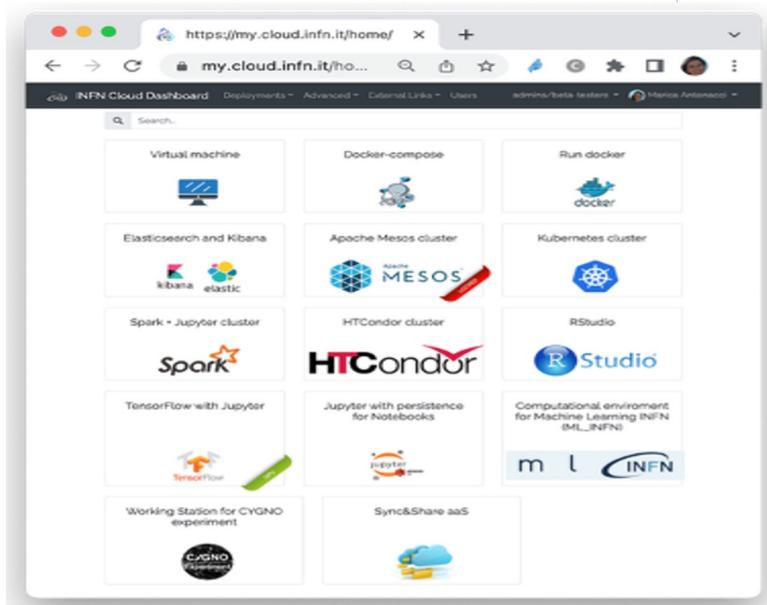


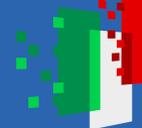
# Facilità d'uso

La federazione servirà utenti con diverse competenze sul calcolo

Nascondere all'utente finale la complessità dell'infrastruttura sottostante

Esperti dei diversi campi sviluppano piattaforme che consentono l'uso efficace delle infrastrutture attraverso la composizione di servizi e risorse





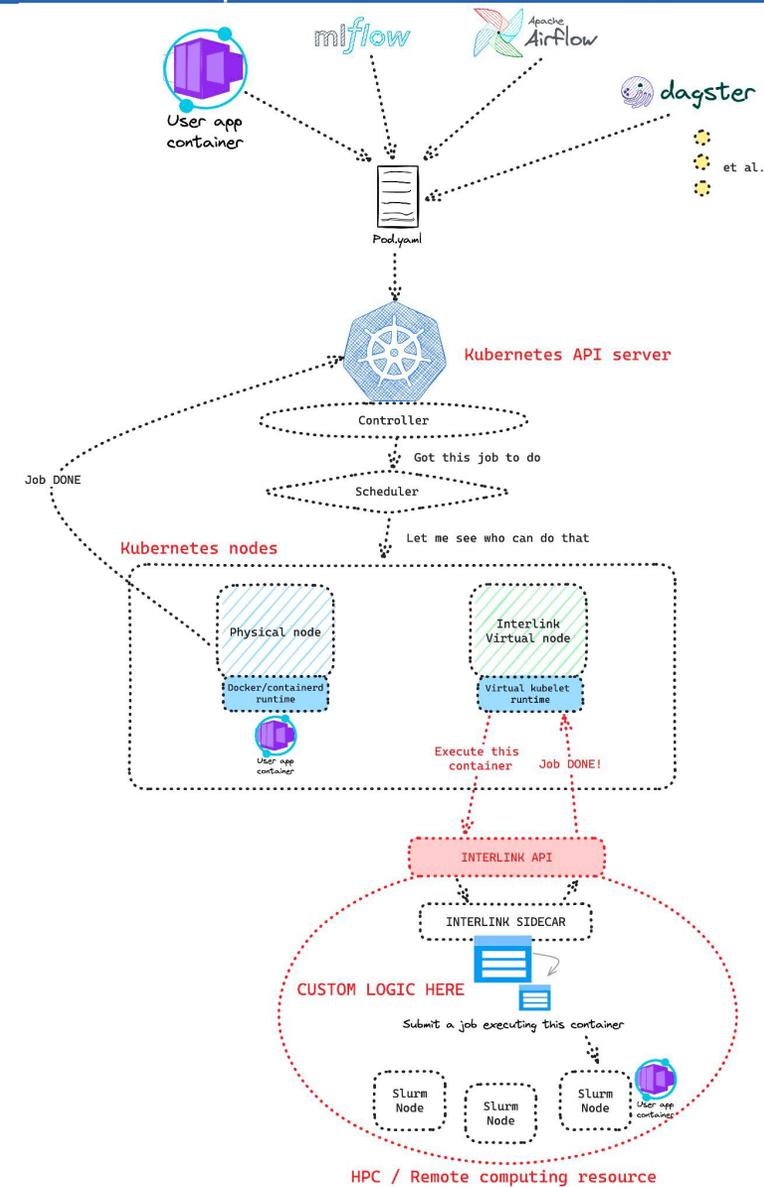
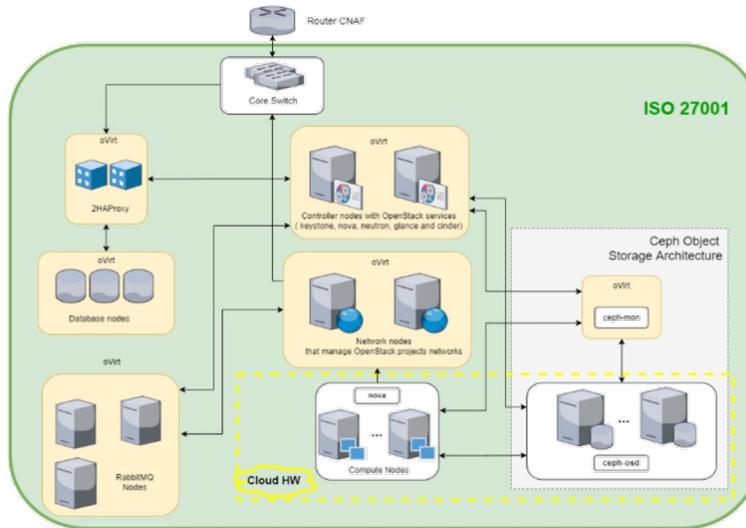
# Flessibilità

Supporto a diversi metodi di accesso alle risorse, con attenzione a:

- a. Trasparenza e facilità d'uso
- b. Efficienza ed efficacia

## Supporto a requisiti specifici delle applicazioni

Ad es. Piattaforme con requisiti particolari sulla privacy

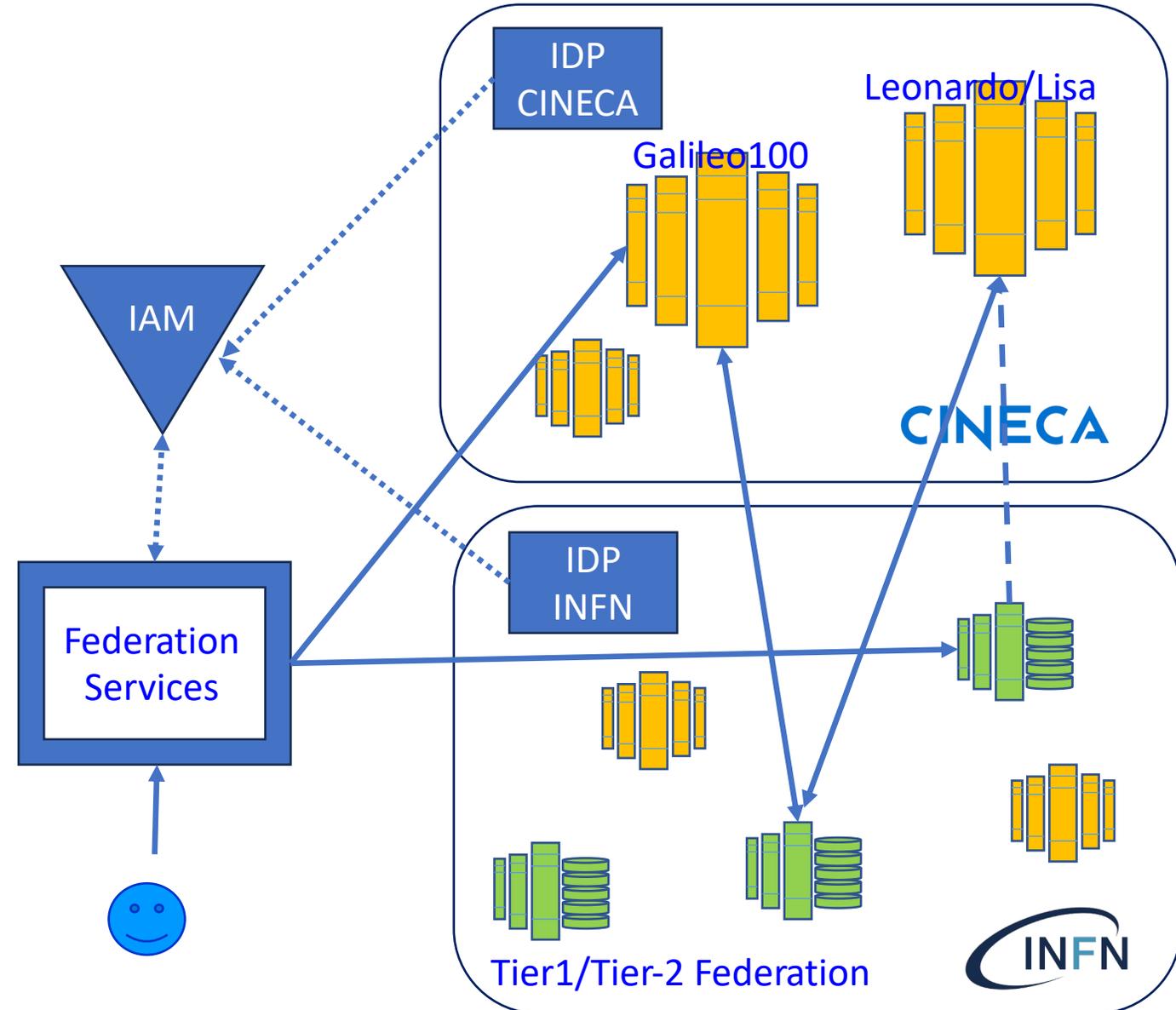


# I piani

Giugno 2024: PoC con funzionalità limitate e utenti «cavia»

Giugno 2025: architettura adottata sull'infrastruttura di produzione con funzionalità via via crescenti

Nel frattempo: accesso alle risorse con meccanismi specifici delle infrastrutture



# Le sfide

## Autenticazione e autorizzazione federata

Costruire il *trust* fra istituzioni e con le federazioni di identità

Definire i meccanismi per comunicare il livello di *assurance* delle operazioni di identificazione

REFEDS Assurance Framework

Adottare sugli strumenti della federazione e nelle istituzioni questi meccanismi

## Accesso alle risorse di calcolo

Utilizzo di soluzioni standard e supporto da parte degli strumenti di federazione

OpenStack, Kubernetes, ...

## Accesso allo storage e gestione dei dati

Supporto a diversi tipi di storage (S3, ...) oggi in produzione

Mantenere efficienza nella comunicazione fra storage e CPU

Tutto questo su infrastrutture in produzione!

# EPIC

## Enhanced Privacy and Compliance Cloud



Enhanced Privacy and Compliance Cloud is an ISO certified cloud platform

A region of INFN Cloud with a certified Information Security Management System



EPIC Cloud offers an IaaS Community Cloud for the communities of

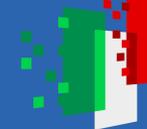
Biomedical and genomic researchers  
Industrial researchers



Site locations: Bologna (active now), Bari and Catania sites will be added in June 2024 enabling for high availability and disaster recovery



Resource available today: about 700 TB of storage, 1440 cores, 10 TB RAM, 6 GPU A100  
On going expansion with 3M euro of NRRP resources and 4M euro of funds from other projects



# Lo Scope del certificate EPIC Cloud Multisito

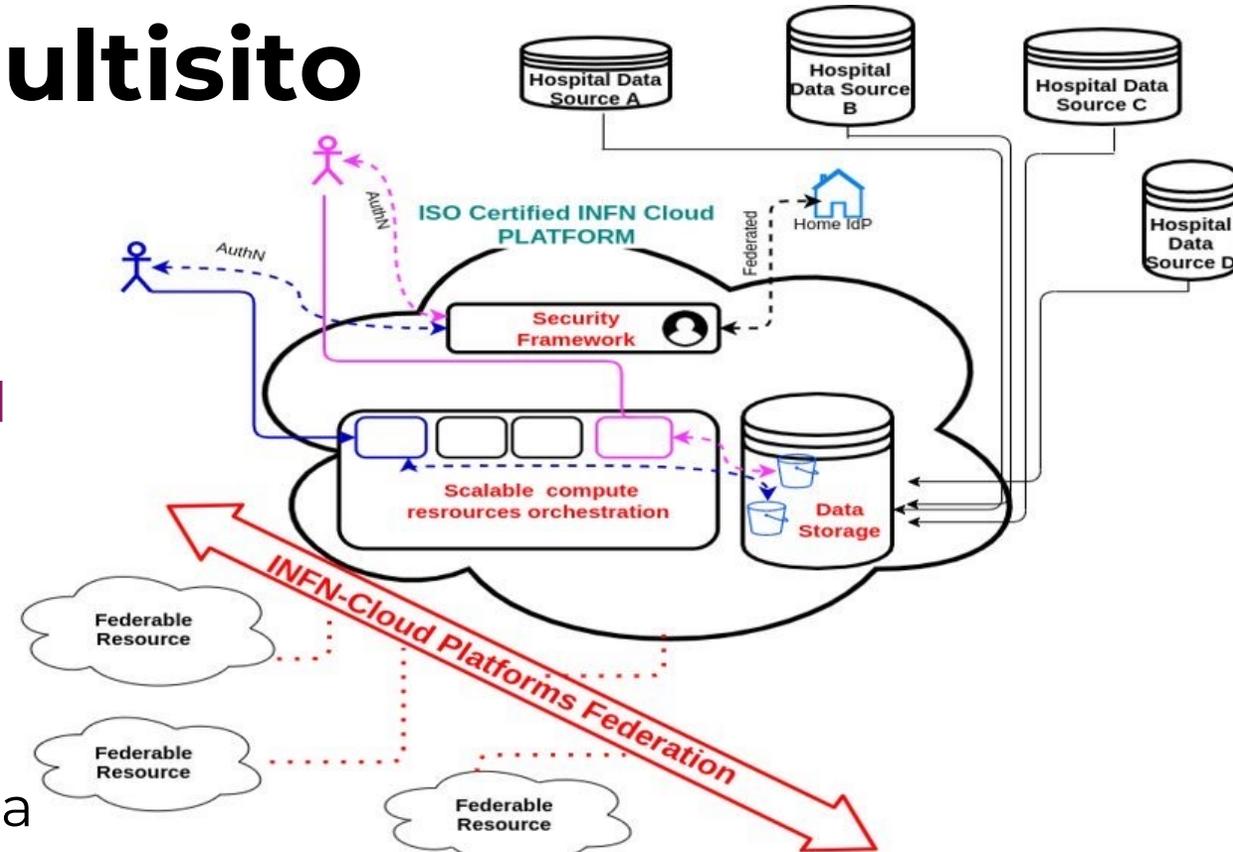
Co-progettazione, sviluppo e manutenzione di soluzioni software DataCloud per il settore della ricerca.

Erogazione di servizi DataCloud IaaS, SaaS e PaaS con modello di deployment Community Cloud.

# Il DataLake ed i servizi di EPIC Cloud Multisito

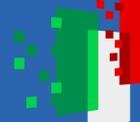
Cosa è disponibile:

- Una piattaforma di **genomica computazionale** basata su BOSCO di IRCCS AOU di Bologna (Sant'Orsola)
- Una piattaforma **Molecular Tumor Board** sviluppata da Health Big Data
- Una piattaforma di **radiomica** sviluppata da INFN Next-AIM
- Prototipi di **HPC-Bubble certificate** per utilizzare l'HPC via cloud
- Prototipi di **piattaforme IoT** per la raccolta di dati da sensori e dispositivi *edge*



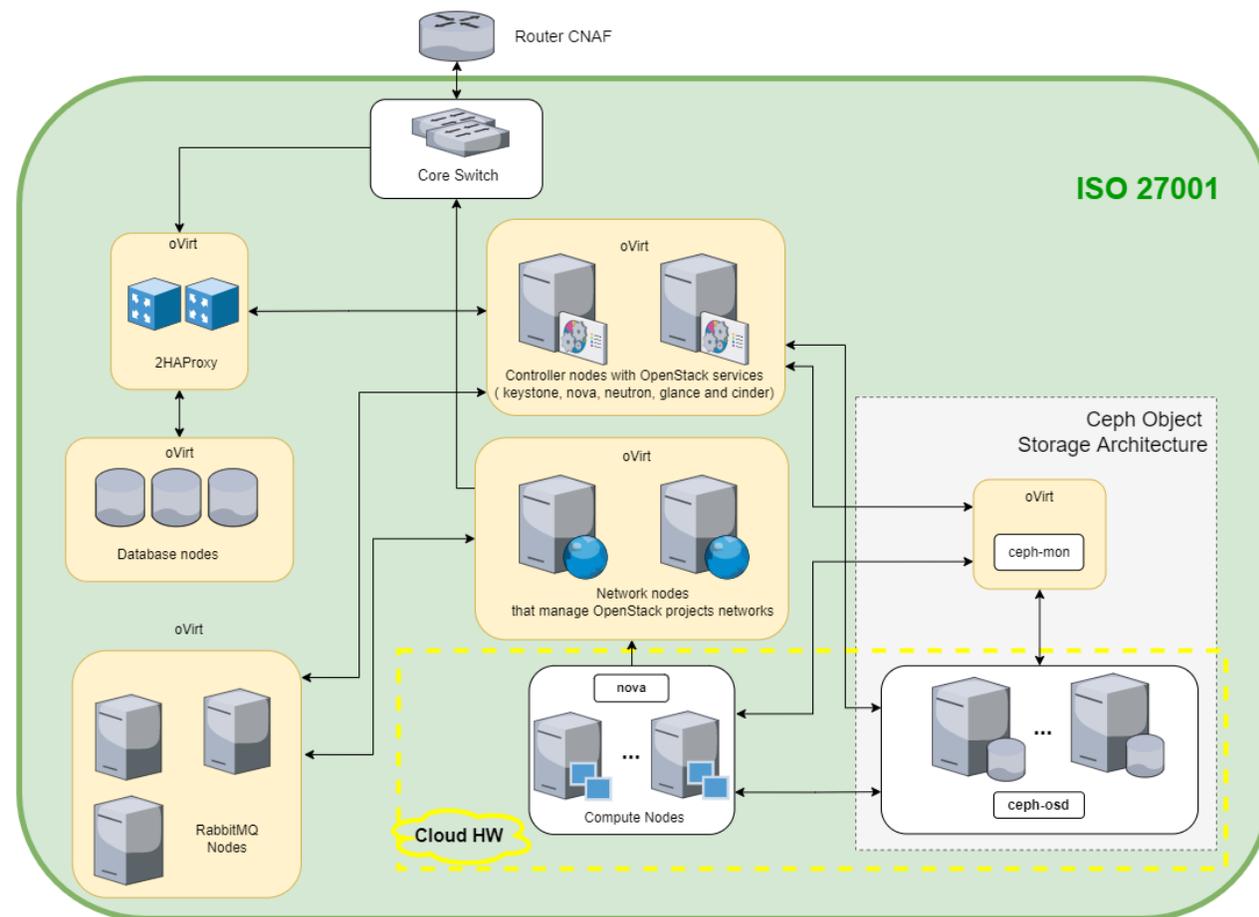
[https://www.physicamedica.com/article/S1120-1797\(21\)00320-3/fulltext](https://www.physicamedica.com/article/S1120-1797(21)00320-3/fulltext)

Le piattaforme saranno integrate nella federazione INFN Cloud e ICSC mutuando le tecnologie di federazione INDIGO IAM, PaaS Orchestrator e RUCIO-FTS



# The EPIC Cloud infrastructure

- All services in HA (Availability is one aspect of security)
- It currently hosts 5 Projects
- ~ 1440 ~10TB RAM, ~700TB Disk
- Tenant/domain segregation
- Physical security (defined perimeters, controlled access to racks, TVCC)
- Network isolation from Tier-1 resources, Next-Generation Firewall (NGFW) in place



OpenStack Infrastructure



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

*Supercomputing  
shaping the future*