Finanziato dall'Unione europea
NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA

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

The Scenario Why a National research Center on HPC, BD & QC?

With the current Data explosion...

- An unprecedented amount of data is going to be produced
- The real competitiveness challenge is extracting value from data
- Supercomputing, simulation, AI, high-performance data analytics and Big Data are essential for innovation and growth in a data-driven society

... need for an ambitious Italian strategy ...
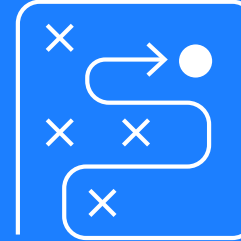
- Europe has a clear strategy (e.g. EuroHPC, EOSC, EPI, Chip Act, Quantum Flagship) - European Data Strategy
- People, businesses and organisations should be empowered to make better decisions based on insights from data

... to "close the gap" with best in class
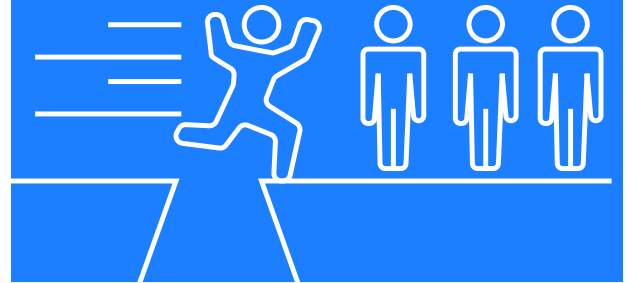
- First actions from 2015: Bologna's Technopole, ECMWF Data Centre, Leonardo pre-exascale supercomputer
- A step forward based on 5 pillars

- Build a **world-class supercomputing** cloud infrastructure to store, manage and process all the produced data

- Set **up centers of excellence** with teams of high-level experts to develop domain applications
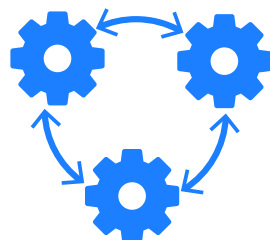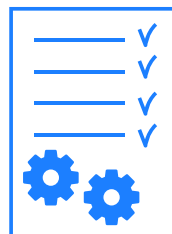
- Set up strong **links** between **Academia, Industry** and **Public Administration**

1001100010010
1010100100001
1010100100101

- **Train** the next generation of data scientists and managers to become **experts** in the digital transition

5 pillars of the action plan

- Implement **structural measures for innovation** and for **dissemination**

https://www.supercomputing-icsc.it/

Next Generation EU funds

191.5 B€ in Italy

- 30.88 B€ for research and education

  - 11.44 B€ "From research to business"

    - 1.6 B€ for R&D *Champions* in Key Technologies

      - **320 M€ for ICSC**

  - 1.58 B€ for Research Infrastructures

    - **41 M€ for TeRABIT**

  - …

ICSC Working model
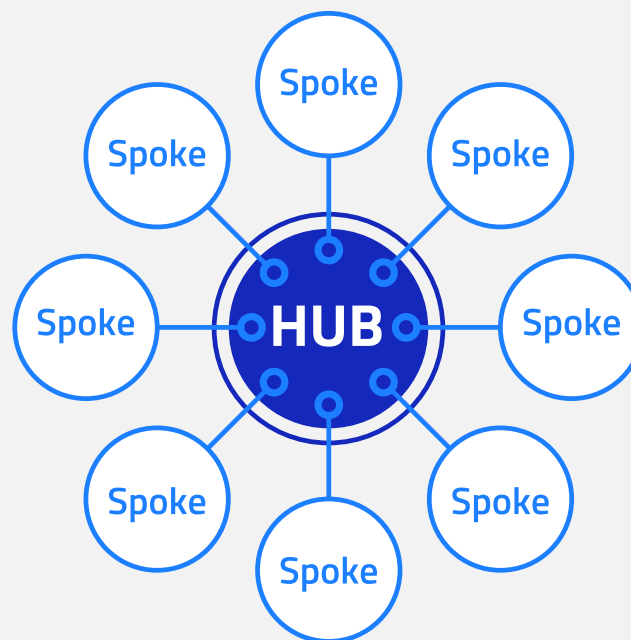
Networks of universities, research institutions, public and private entities aggregated in consortia in «HUB&SPOKE» mode

Started: September 2022

Hub & Spoke model

- Governance structure: Hub and Spokes

- Hub purpose: management and coordination

- Spokes purpose: CN activities execution (research, development, infrastructures and research material hosting, etc.).

- Spoke Leader/Co-Leader: lead the scientific activities coordination. The initial set of Spoke Leader e Co-leader will remain in charge for 4 years and each person could be nominated again only once

L'ICSC includes
10 thematic spokes
1 infrastructure spoke

0 SUPERCOMPUTING CLOUD INFRASTRUCTURE

Garr Network
HPC Centre
Future HPC Centre
Big Data Centre
Future Big Data Centre
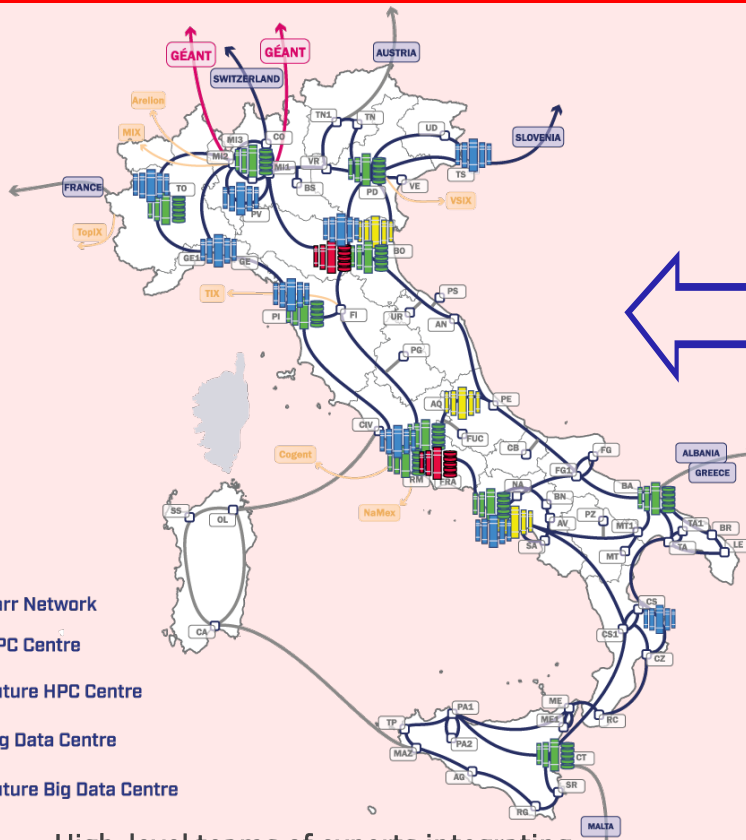
High-level teams of experts integrating the Spokes working groups (mixed cross-sectional teams)

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

Finanziato dall'Unione europea
NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA

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

ICSC founders:
a public-private partnership

25
Universities

12
Research institutes

14
Strategic private companies

Finanziato dall'Unione europea
NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA

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

Private companies Founding members:
strategic players for digital transformation

Highly-qualified group of large leading companies covering most of the strategic industrial sectors involved by digital transformation in Italy

fondazione innovazione urbana

Strategic partner to implement and develop the digital twin pilot case of an urban complex system

iFAB INTERNATIONAL FOUNDATION BIG DATA & ARTIFICIAL INTELLIGENCE FOR HUMAN DEVELOPMENTE

Industry-driven not-for-profit international organization aimed at: (1) aggregating   companies, including SMEs, to engage with ICSC through a structured partnership, (2) funding research and innovation projects, (3) promoting the Big Data Technopole
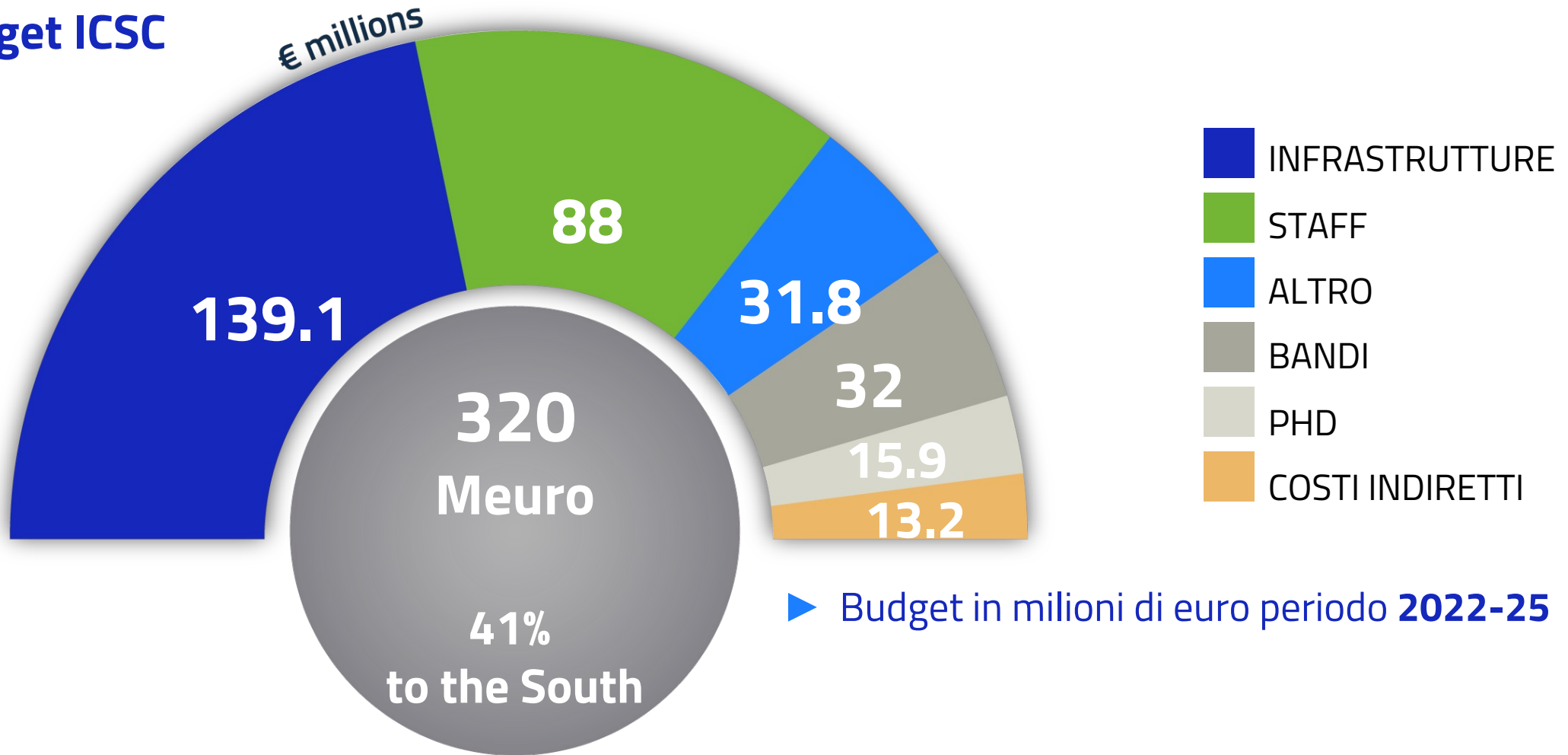
ICSC: resources to bring **Research results to Business**

1.500
Personnel shared by partners

250+
New researchers

320 M€
Total funds

250+
New PhDs

32 M€
Open call

32 M€
Innovation grant

ICSC Italian Research Center on High-Performance Computing, Big Data and Quantum Computing   Missione 4 ▪ **Istruzione e Ricerca**

The Bologna Big Data Technopole

Bologna

ECMWF Data Center

CINECA & INFN Exascale Supercomputer Center

Big Data Association and Foundation

Civic Protection and agency for development and innovation

ICSC

Conference and Educational Center

Meteo National Agency

Innovation Center

IFAB - International foundation Big data

Competence Center Industry 4.0

University Center

Biobank and Life Science

ENEA Center

Finanziato
dall'Unione europea
NextGenerationEU

Ministero
dell'Università
e della Ricerca

Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA

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

Role of INFN

INFN has been chosen by the Italian Ministry for University and Research (MUR) for driving the preparation and execution of the ICSC project

Acknowledgement of the experience in computing technologies and Big Data in particular of **INFN** and **HEP** in general

Strategic partners for the Supercomputing Infrastructure Cloud:
CINECA for HPC
GARR for networks

INFN leading role also in spoke 2 and 3:

Fundamental Research & Space Economy

Astrophysics & Cosmos Observations
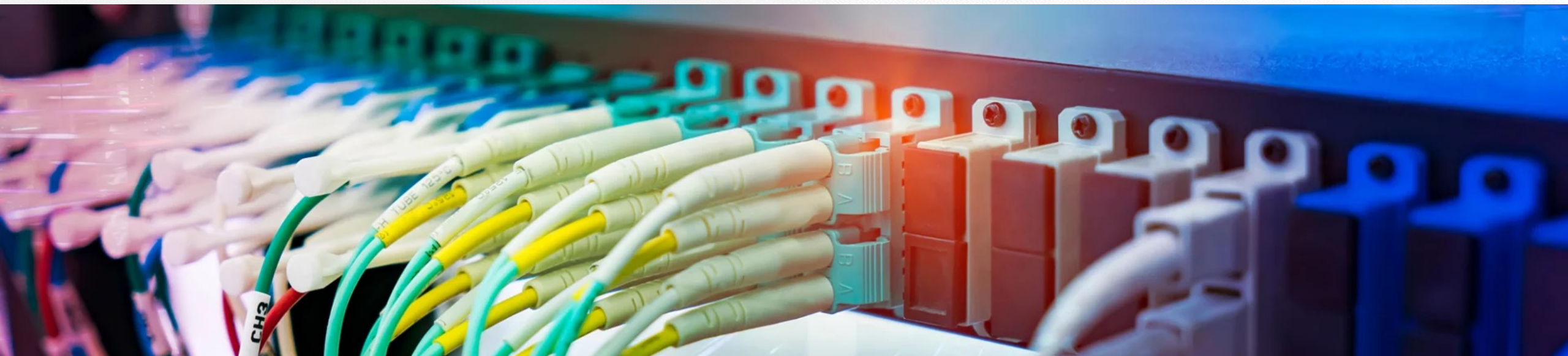
https://www.terabit-project.it/

TeRABIT: Terabit Network for Research and Academic Big Data in ITaly

TeRABIT is a Research Infrastructure project synergic with ICSC, started in January 2023

Partners are the same of the ICSC Spoke-0 (Supercomputing Cloud Infrastructure):

INFN, CINECA and GARR

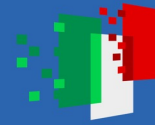Covers areas complementary to those of the ICSC infrastructure

TeRABIT research infrastructures:

GARR-T
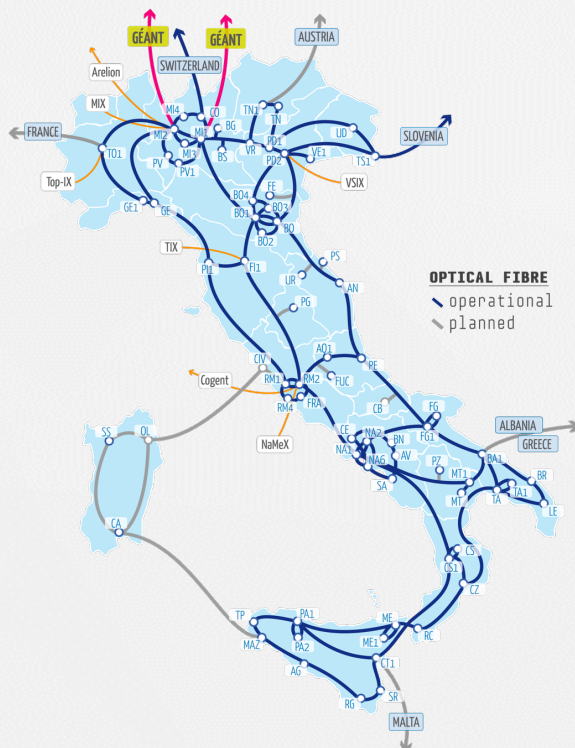
PRACE-Italy
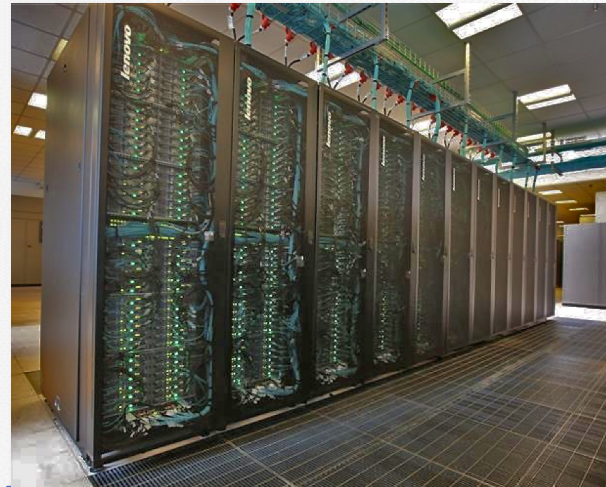
Galileo100 – cluster HPC
Ospitato al CINECA - Bologna

HPC-BD-AI - INFN Cloud
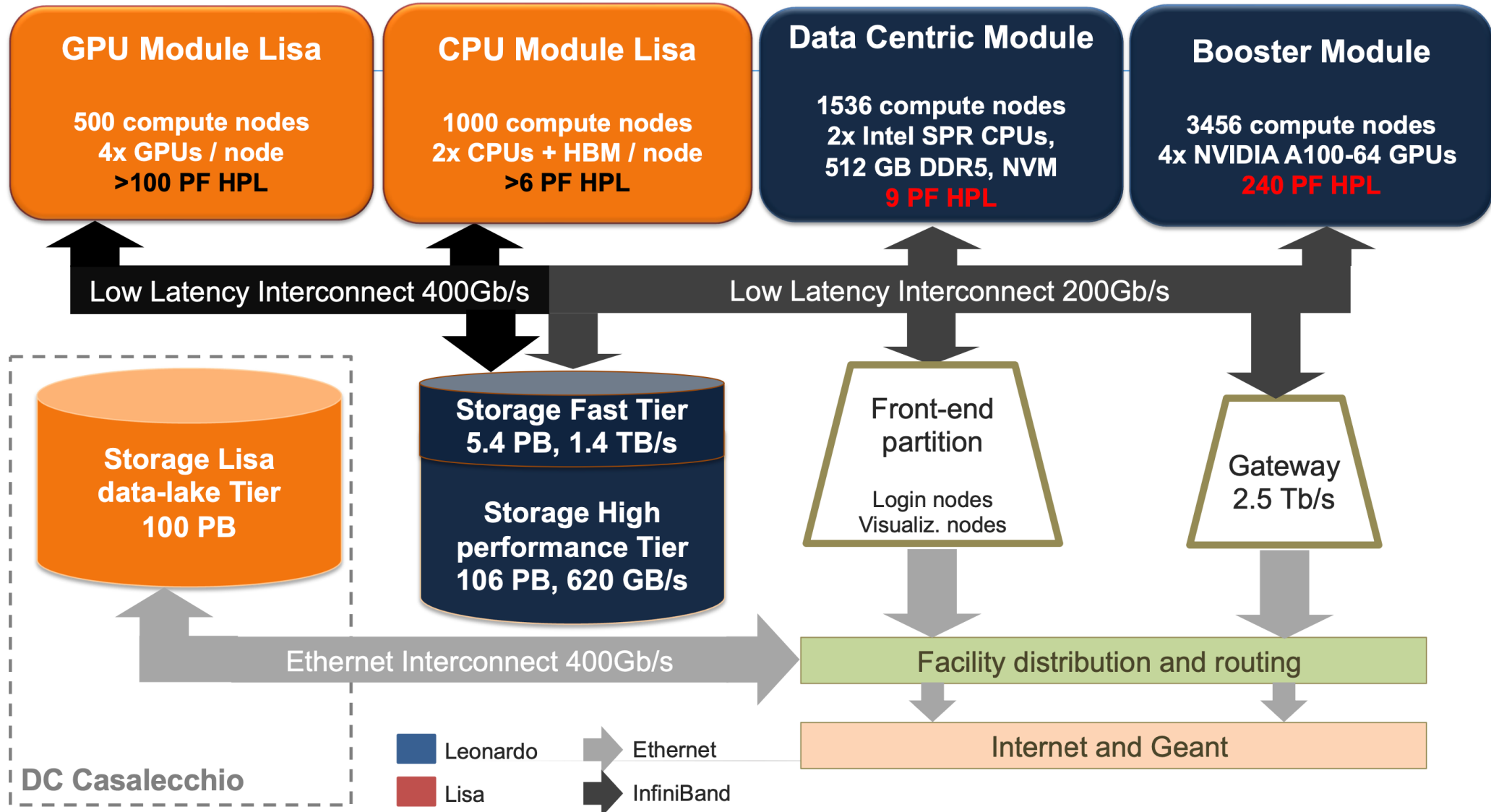
Network

GARR-T
Upgrade of the optical network centre-north OLS+DCI (100G+, 400G+)

ICSC
Upgrade of the GARR-X Progress network (OLS) upgrade (100G+, 400G+)

TeRABIT
Acquisition of optical fibre in Sardinia and interventions in southern regions
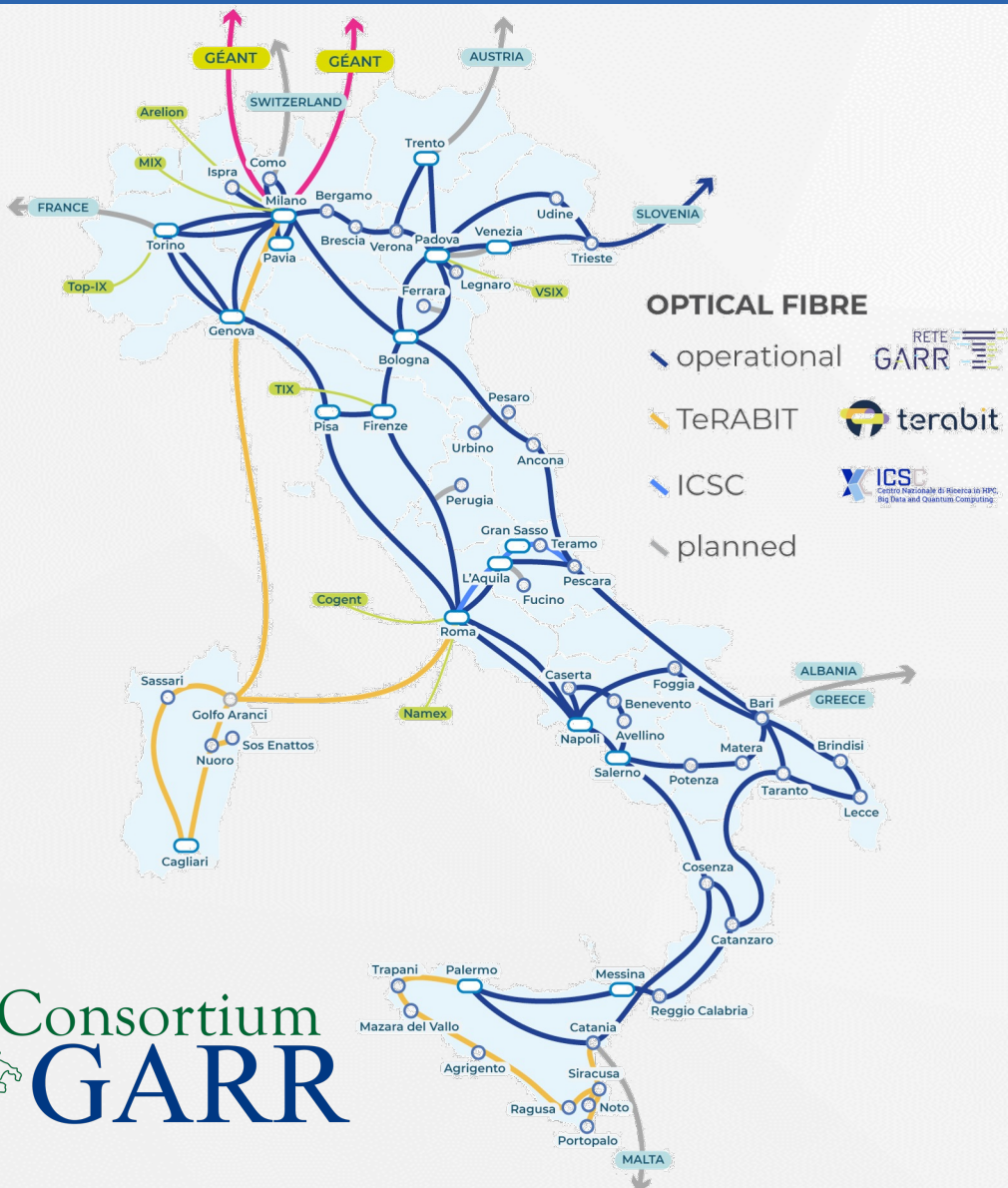
Big Data and Federated Cloud

INFN WLCG Tier-1 & Tier-2 infrastructure

Currently about 100,000 CPU cores, 100 PB disk (net), 150 PB tape

About 100,000 more CPU cores, 80 PB disk (net), >30 PB tape + a new library at CNAF

30 M€ investment in ICSC

HPC bubbles: HPC systems in a selected number of sites, equipped with CPUs, GPUs (Nvidia H100), FPGA, fast storage, Infiniband

~10 M€ investment in TeRABIT

New data centres for Disaster Recovery (Gran Sasso) and Space Economy (Frascati)

9 M€ investment in ICSC

A *data lake for research* – High-level view

The proposed model is based on:

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

Unified vision

Data Lake Entry Point

User Community Services

Individual communities services

General Services

Generic services

Infrastructure sites

HPC

Cloud Provider

Tier1

Tier2

Data centres and resource providers
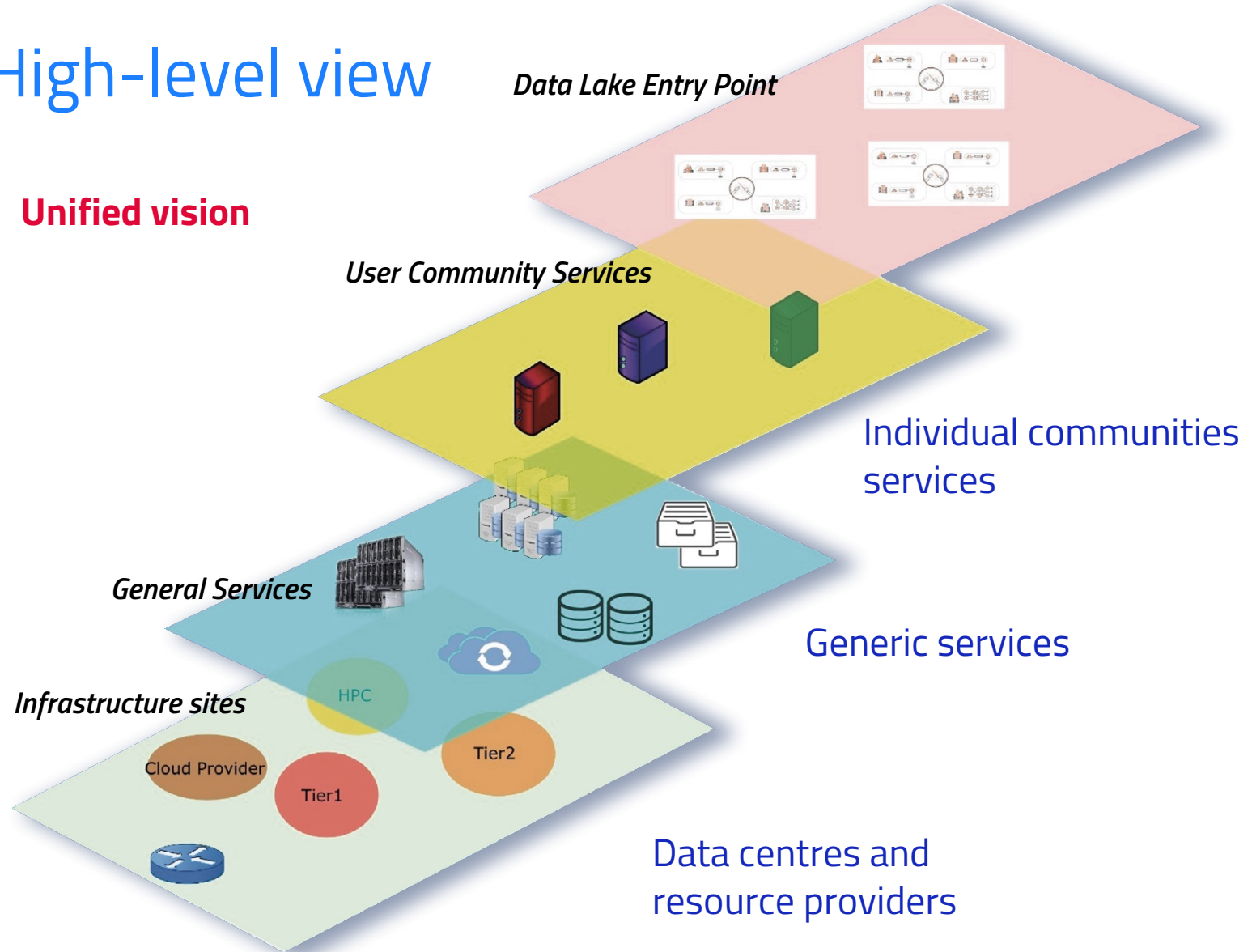
Physical and logical levels

1. The data-lake of a specific scientific domain is connected to a unique entry point (the Italian research data-lake entry point); for example for multi-domain activities or international links
2. The data-lake of a specific scientific domain is defined as the sum of the services that provides (portals, SW services, CPU/Disk/Tape resources, ...)
3. On the other side of the wall, every physical resource provider may decide to support a specific scientific domain via the publication of «capabilities»
4. Each scientific domain service are deployed on suitable resources via a match-making process

External physical connections

External logical connections

Centro 2

Sc. Domain B | Sc. Domain C

Functionally flat network connection

Centro 1

Data LAKE ENTRY POINT

Sc. Domain A

Sc. Domain C

Italy

Centro 3

Sc. Domain A | Sc. Domain C

Centro 4

Sc. Domain A | Sc. Domain B

WALL: separates the logical (services) view from the physical (resources) view

**Resources view
(Physical, as seen by the resource provider)**

**Services view
(Logical, as seen by the researcher)**

Finanziato
dall'Unione europea
NextGenerationEU

Ministero
dell'Università
e della Ricerca

Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA

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

Enhanced Privacy and Compliance Cloud

Medical applications (and not only) require treatment of data protected by GDPR

The EPIC Cloud at CNAF addresses the requirements of projects and experiments dealing with clinical, biomedical and genomic data

The cloud is certified according to information security standards like ISO/IEC 27001:2013, ISO/IEC 27017:2015 and ISO/IEC 27018:2017

A multi-site certification is being extended to the Catania and Bari sites

European Commission

GDPR

INFN applications (thematic spokes)

Science driven use cases

    from Theoretical, Collider and Astroparticle Physics

    from Medicine

Technology driven use cases

    Single node optimization: GPU, FPGA, alternative architectures

    Machine Learning, AI

    Distributed computing: high-rate analyses, data management for smaller experiments, …

    Cross boundary initiatives, with a strong link to Space Economy satellite data

    Sensitive data management

Technological research

    Quantum Computing (systems, algorithms)