

Istituto Nazionale di Fisica Nucleare

# Organizzazione del Calcolo INFN

G. Carlino - INFN Napoli

Corso Nuovi Assunti PNRR

G. Carlino - Organizzazione del Calcolo INFN

LNF 09/10/23

09/10/2023

1

L'INFN è stato fondato  
nel 1951  
e ha una lunga e  
prestigiosa  
storia che discende  
da Enrico Fermi  
e i ragazzi di via  
Panisperna

G. Carlino - Organizzazione del Calcolo INFN

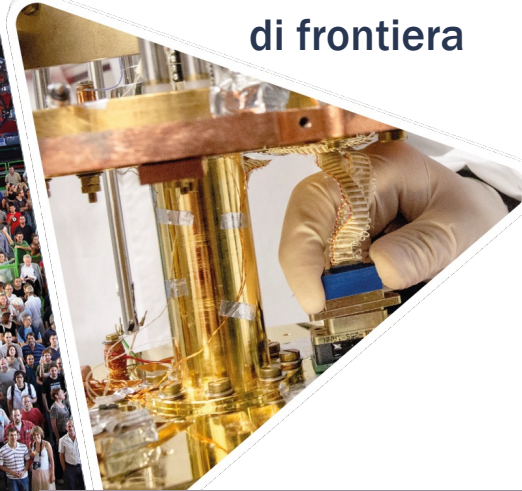
**Spingersi oltre  
le frontiere della  
conoscenza**  
I segreti del Big Bang

La nostra  
**missione**

**Formare**  
gli scienziati e gli  
ingegneri di domani



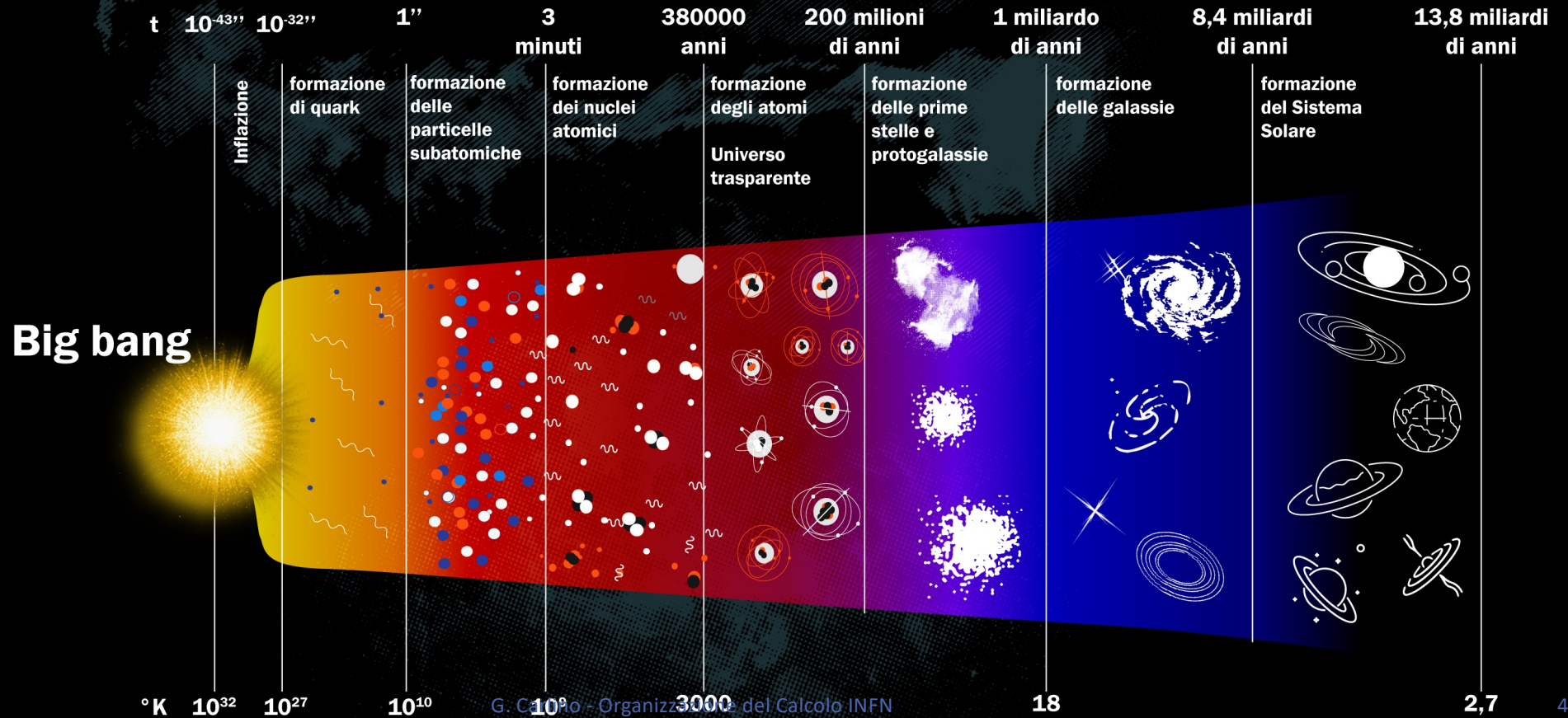
**Sviluppare**  
nuove tecnologie  
di frontiera



**Lavorare insieme**  
con i giovani e i ricercatori  
di tutto il mondo

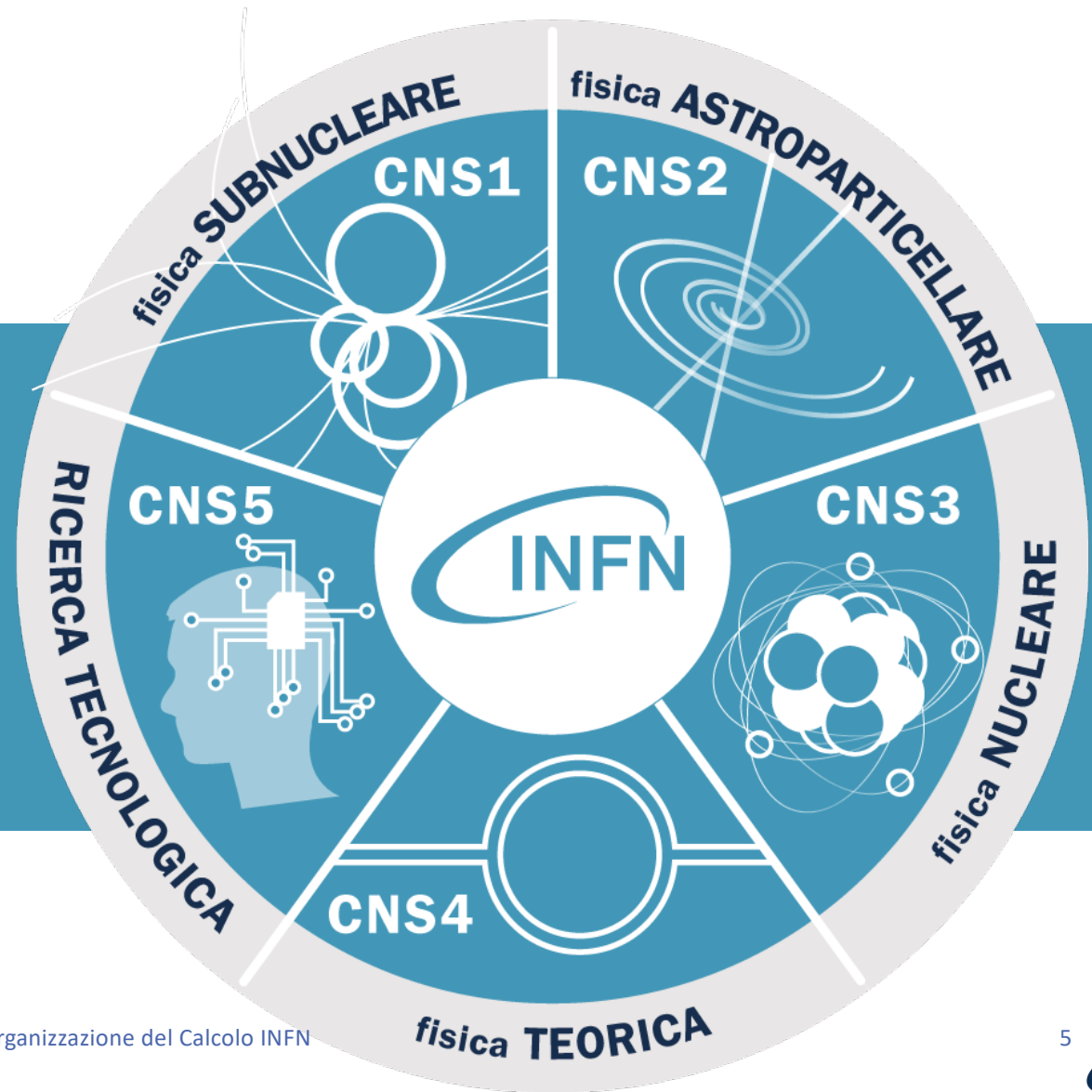
# La sfida

comprendere l'Universo nei primi momenti dopo il Big Bang



G. Caprio - Organizzazione del Calcolo INFN

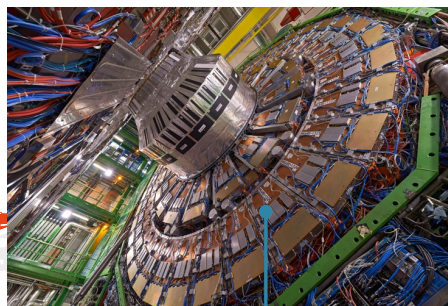
Le 5 linee  
di ricerca  
e le commissioni  
scientifiche  
nazionali



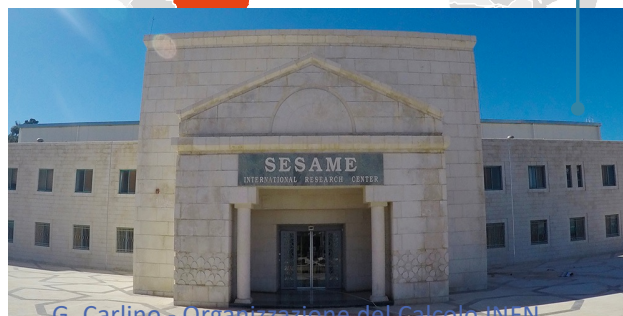
G. Carlino - Organizzazione del Calcolo INFN

- 4 Laboratori Nazionali
- 20 Sezioni
- 6 Gruppi associati
- 3 Centri Nazionali e Scuole
- 1 Consorzio internazionale

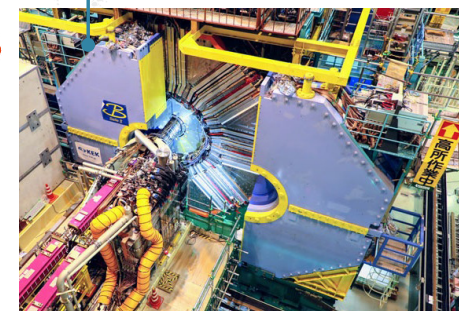


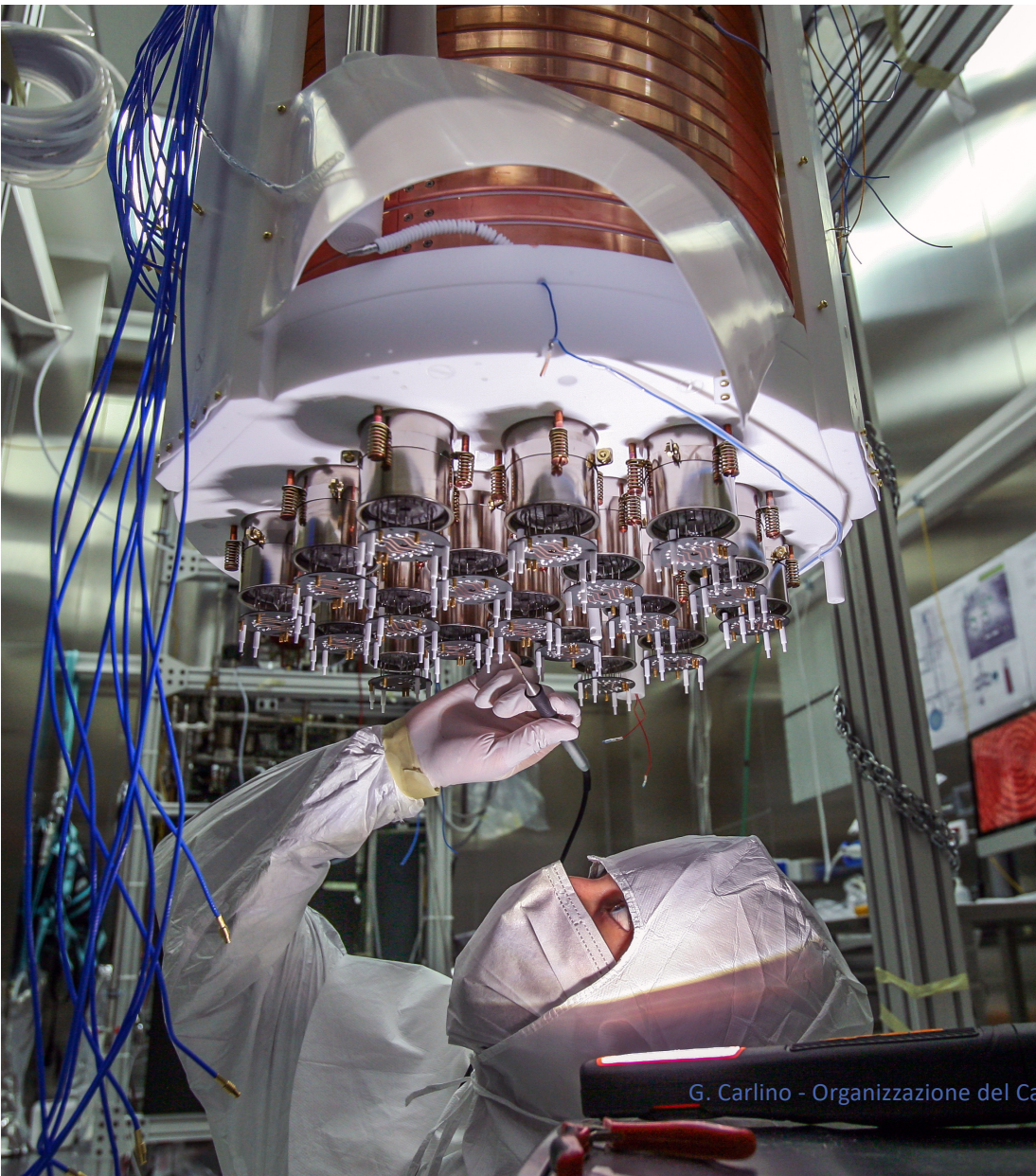


# Un DNA internazionale



G. Carlino - Organizzazione del Calcolo INFN





G. Carlino - Organizzazione del Calcolo INFN



L'INFN è ...

una comunità di  
oltre **6000** persone

di cui ~ **25%** di assegni di ricerca,  
borse di studio e borse di dottorato





# Numero del personale a Tempo Indeterminato / Tempo Determinato



**TEMPO INDETERMINATO 2106**

uomini 1550

donne 556

TI



**TEMPO DETERMINATO 126**

uomini 71

donne 55

TD



# Distribuzione del personale a Tempo Indeterminato



**Ricercatori 699**

**Tecnologi 447**

**Tecnici 606**

**Amministrativi 352**

**Dirigenti Amministrativi 2**

**Totale personale TI**

**2106**

# Distribuzione del personale Associato



G. Carlino - Organizzazione del Calcolo INFN

<b>Personale associato totale</b>	<b>4503</b>
<b>Assegnisti</b>	<b>366</b>
<b>Borsisti</b>	<b>38</b>
<b>Dottorandi</b>	<b>1143</b>
<b>Incarichi di ricerca</b>	<b>816</b>
<b>Altre associazioni</b>	<b>1975</b>
<b>Associazioni tecniche</b>	<b>165</b>

Aggiornato al 31/12/2022

# Budget 2022

**579** mln di euro

Risorse finanziarie  
esclusi Fondi Vincolati

**292** mln di euro



Fondi Vincolati  
a progetti specifici

**121** mln di euro

PNRR

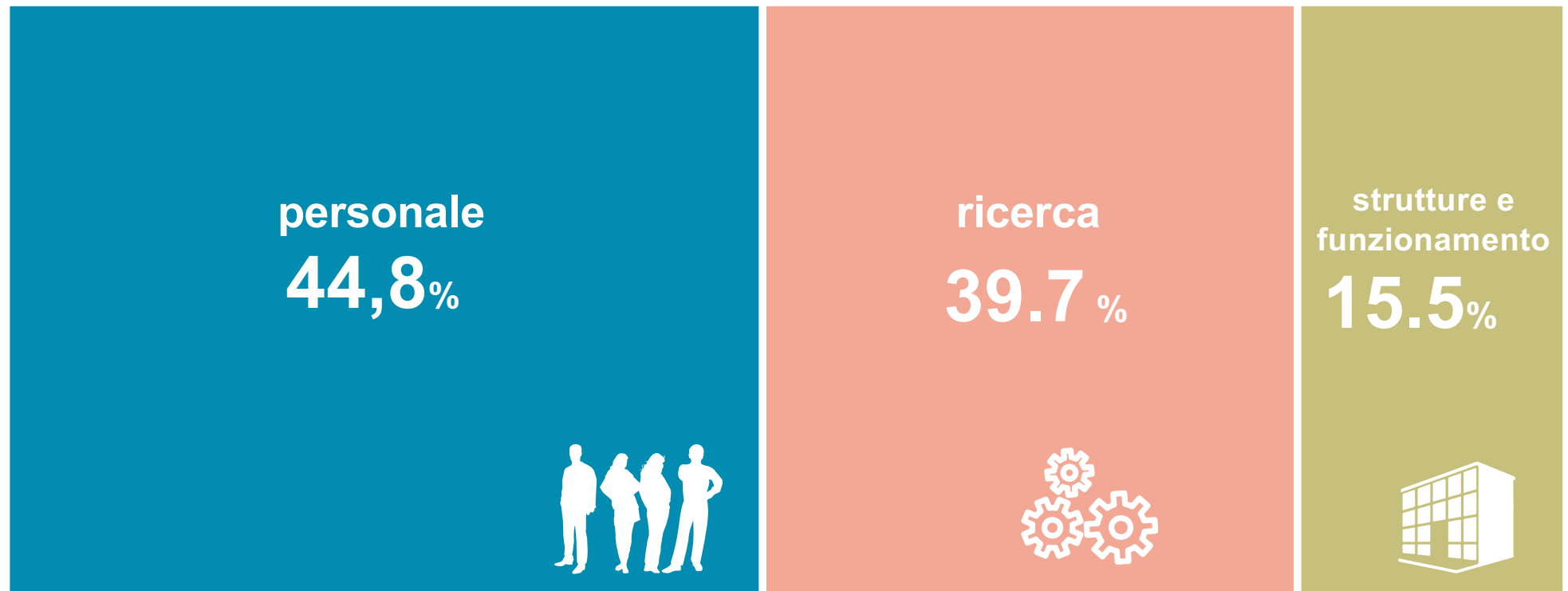
**166** mln di euro

G. Carlino - Organizzazione del Calcolo INFN



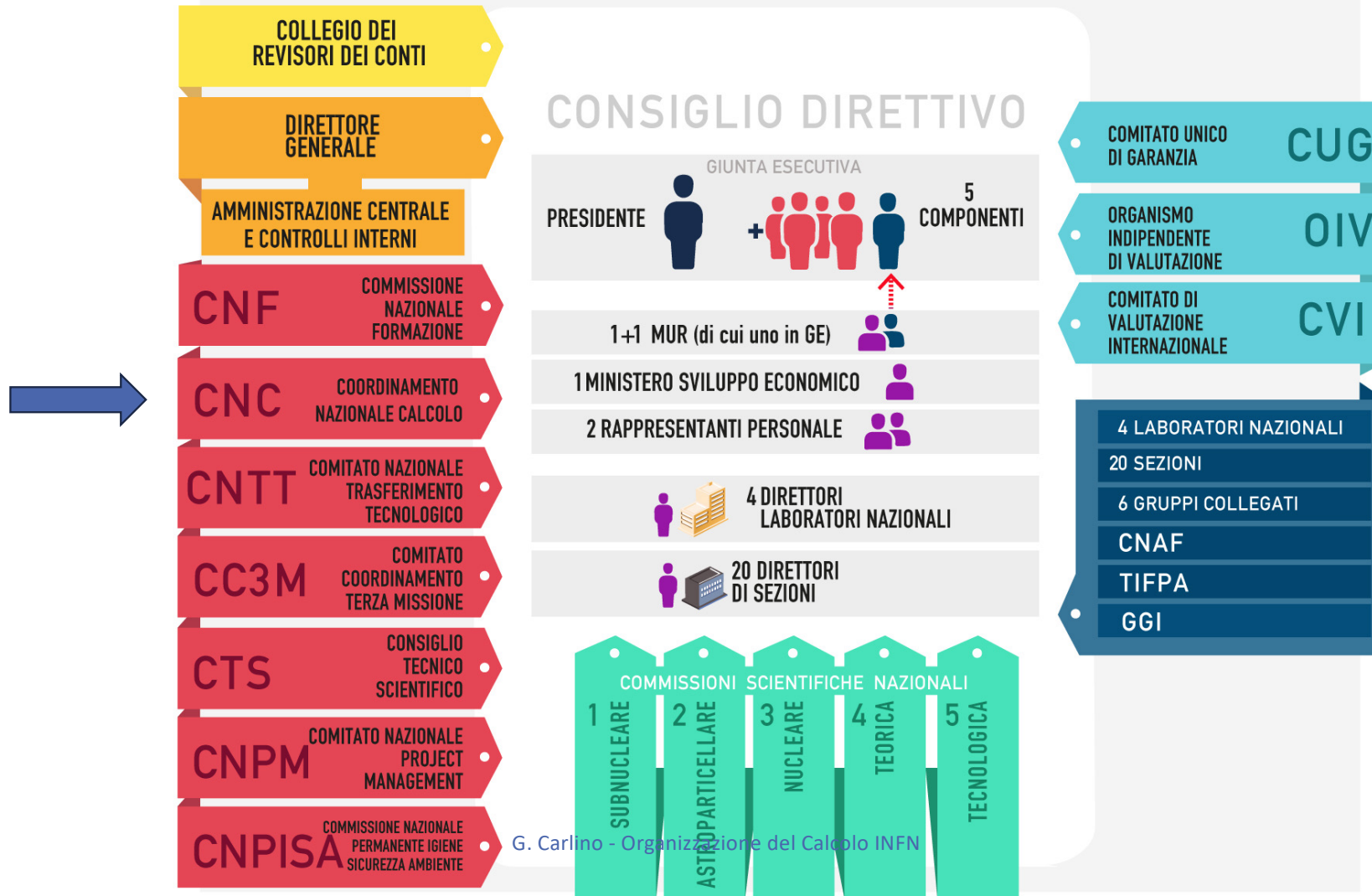
# Bilancio 2022

Profilo di spesa per  
ricerca, personale e funzionamento



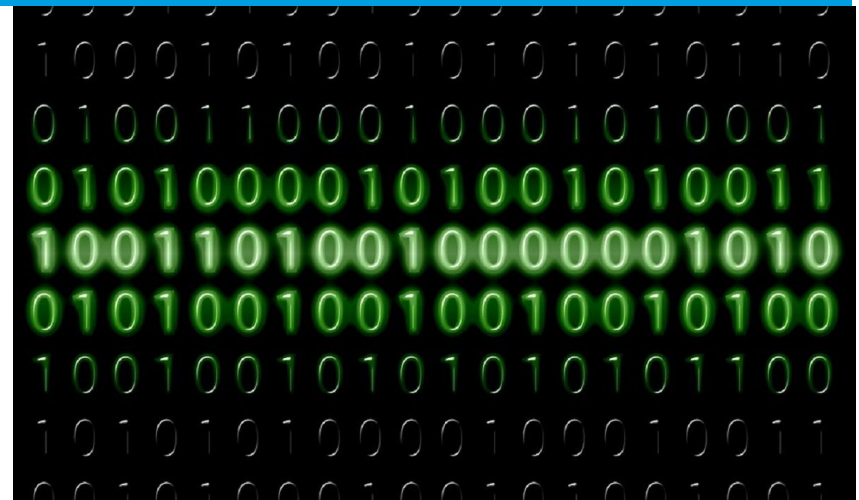
G. Carlino - Organizzazione del Calcolo INFN

# L'organigramma dell'INFN



G. Carlino - Organizzazione del Calcolo INFN

# Il Calcolo nell'INFN



Tutte le attività INFN, sia scientifiche che amministrative, richiedono un'infrastruttura di calcolo e un insieme di servizi altamente efficienti e personale altamente qualificato

# L'Infrastruttura di Calcolo Scientifico INFN



L'Infrastruttura di Calcolo INFN, in attività da più di 20 anni, è costituita da 10 centri:

- 1 Tier1: CNAF @ BO
- 9 Tier2: BA – CT – LNF - LNL/PA - MI – NA – PI - RM1 - TO

connessi con link dedicati a 240 Gbps (Tier1) e a 10 / 100 Gbps (Tier2) al backbone del GARR

Non solo centri INFN

- HPC @CINECA: da anni l'INFN utilizza i sistemi di calcolo ad alte prestazioni sia per l'HTC che per attività che richiedono alta parallelizzazione

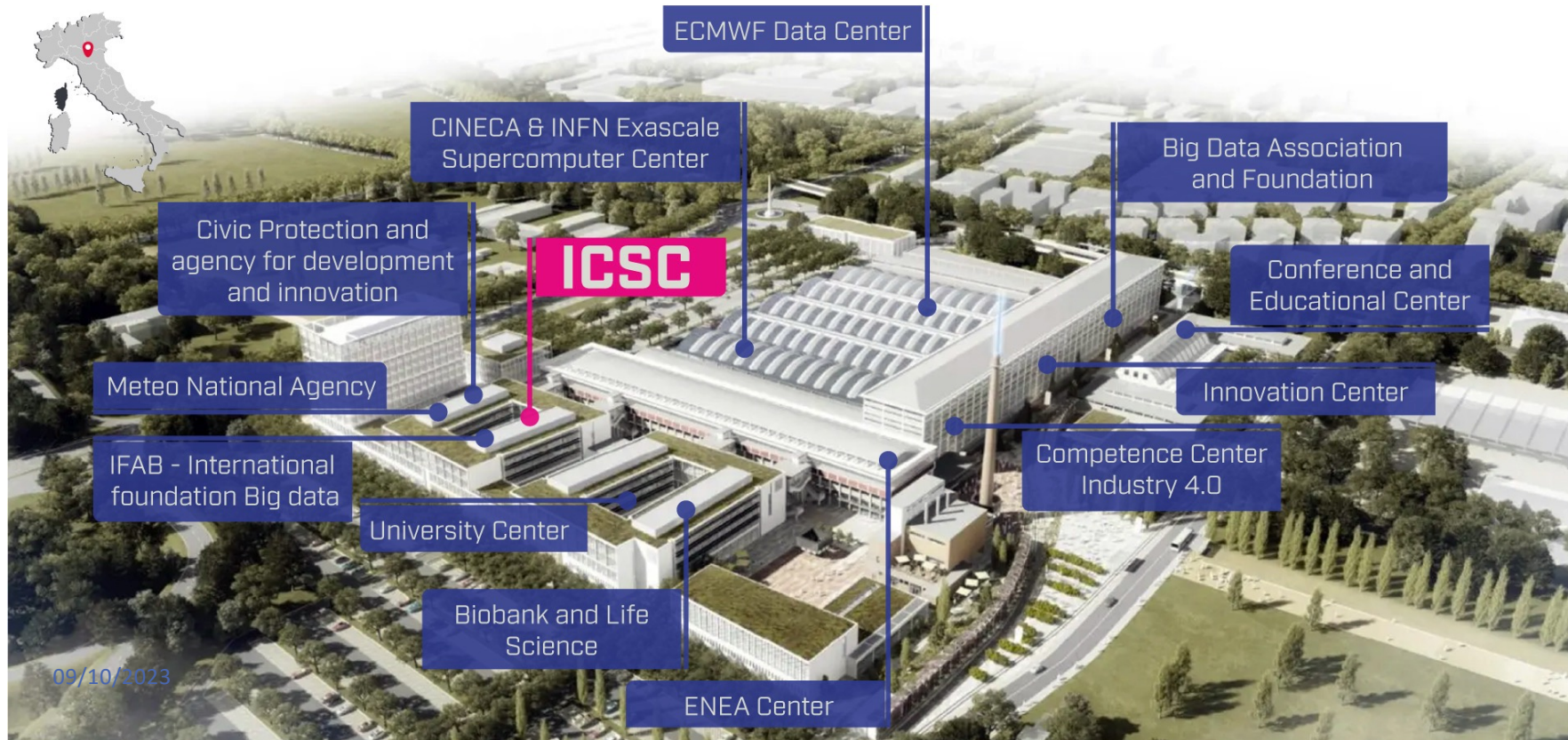
09/10/2023

G. Carlini - Organizzazione del Calcolo





# Calcolo Scientifico @ Tecnopolo



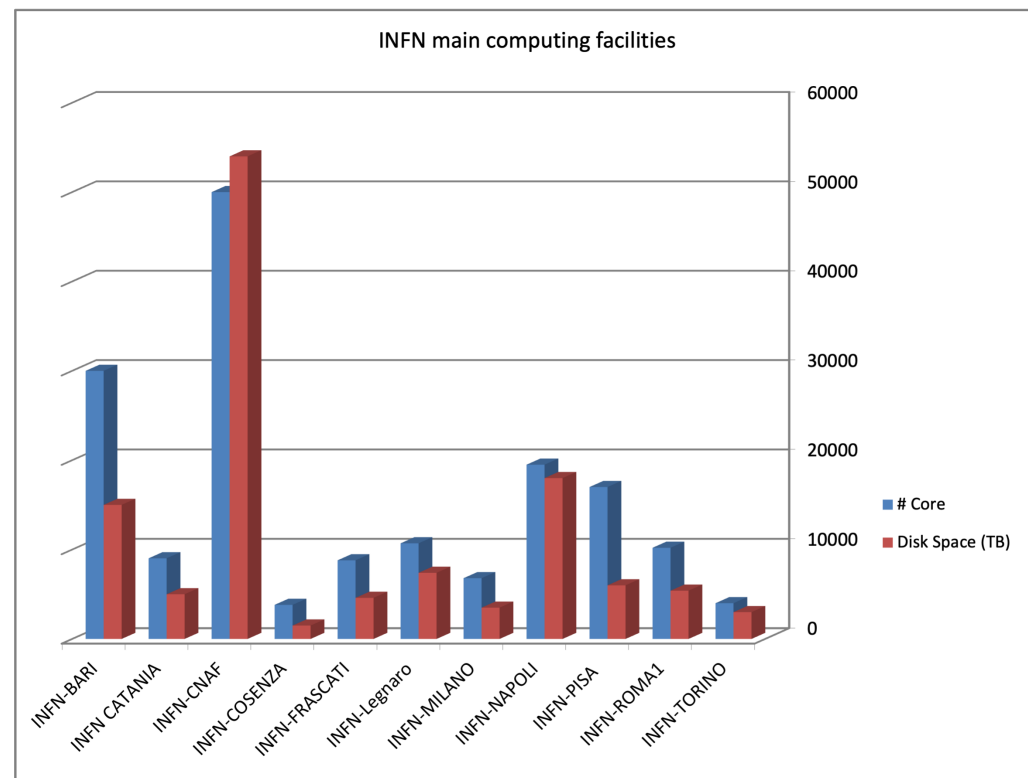
# L'Infrastruttura di Calcolo Scientifico INFN



Risorse di calcolo e storage nei centri:

- Calcolo
  - 170.000 cores
  - 2.200.000 HepSpec06
- Storage
  - 125 PB Disco
  - 220 PB Nastro

L'infrastruttura si sta arricchendo di architetture alternative (GPU, FPGA, ARM) per supportare applicazioni parallele e studiare come aumentare l'efficienza energetica dei centri

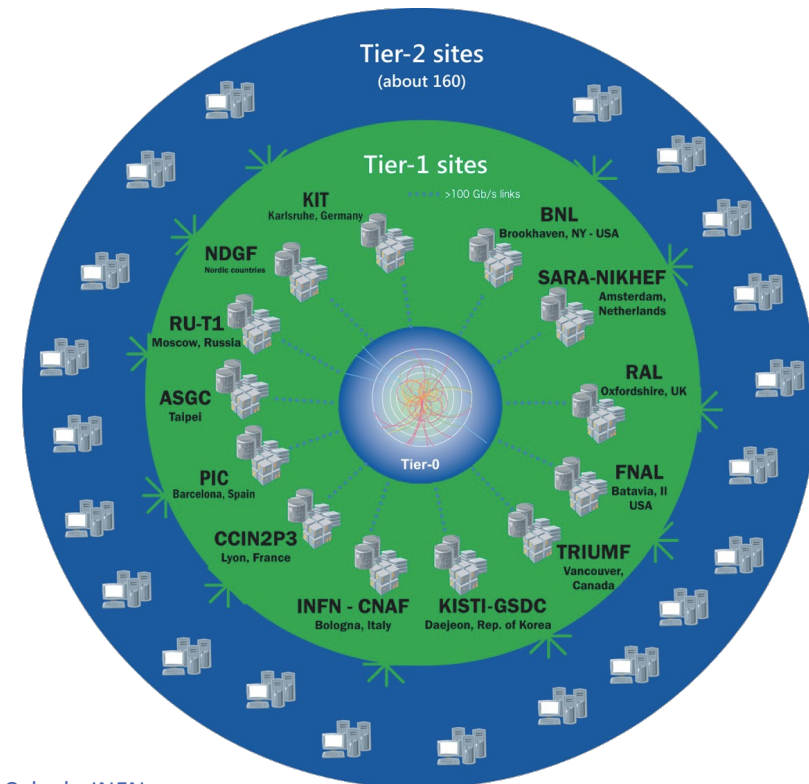


# WLCG – Il calcolo distribuito



L'Infrastruttura di Calcolo Scientifico INFN è parte dell'infrastruttura mondiale di calcolo distribuito gestita da WLCG in attività dall'inizio degli anni 2000

- Il progetto WLCG (Worldwide LHC Computing Grid) è una collaborazione internazionale di circa 170 centri in più di 40 paesi che nasce con la missione di fornire risorse di calcolo e storage in una singola infrastruttura, per distribuire e analizzare i dati raccolti da LHC
- Evoluzione di WLCG per includere esperimenti non LHC o non HEP (astroparticles)



# Le sfide del calcolo scientifico



La prossima generazione di esperimenti HEP richiederanno enormi risorse di calcolo e storage (e rete) per la crescente complessità degli eventi (track multiplicity) e la quantità di dati raccolti (quasi un ordine di grandezza maggiore a LHC)

- Risorse di calcolo necessarie x20 rispetto a oggi ma il budget a disposizione resterà sostanzialmente costante (flat budget) **e non c'è sempre un PNRR**
- Nuovi esperimenti richiederanno risorse paragonabili a HL-LHC

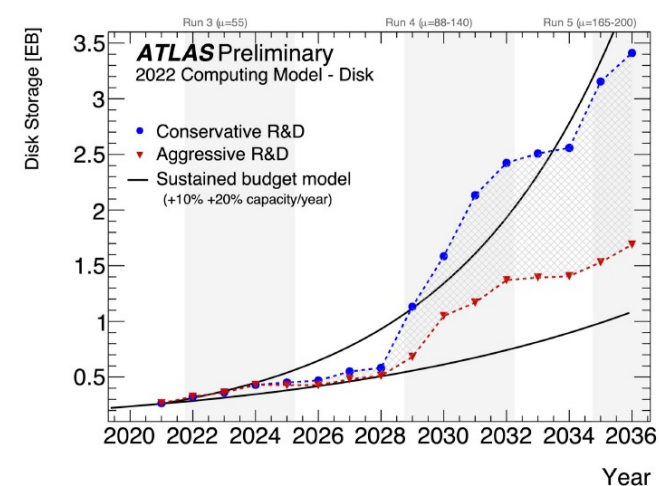
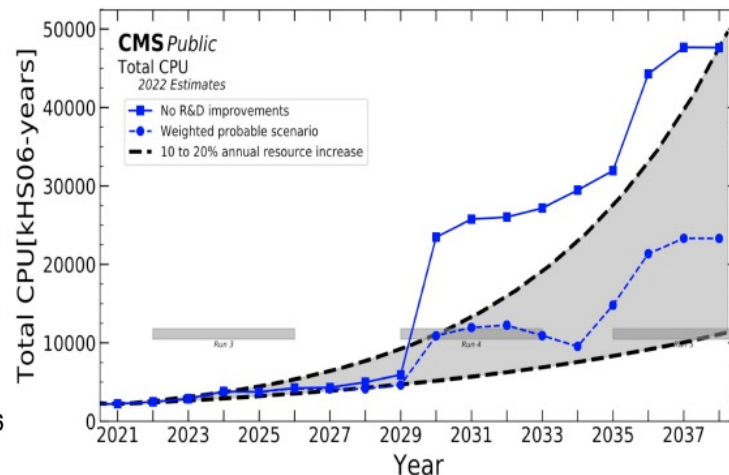
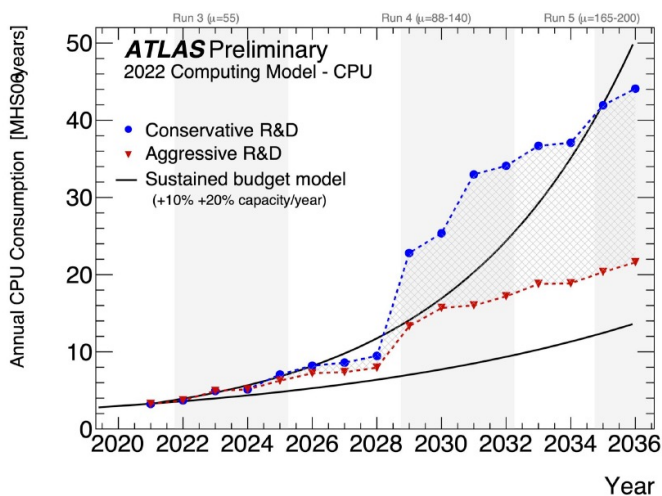
Parameter	unit	2023 Run3	2029 Run 4	2033 Run5
Interaction/crossing	max $\mu$	55	140	200
Integrated luminosity	$fb^{-1}y^{-1}$	100	300	450
LHC ready for physics	$10^6$ s	7	7	7
Rate	kHz	1.4	10	10
Recorded events	$10^9$	10	70	70

L'evoluzione della tecnologia e delle infrastrutture «standard» usate fino al Run2 di LHC non aiuta a ridurre i costi!

È necessario apportare innovazioni in

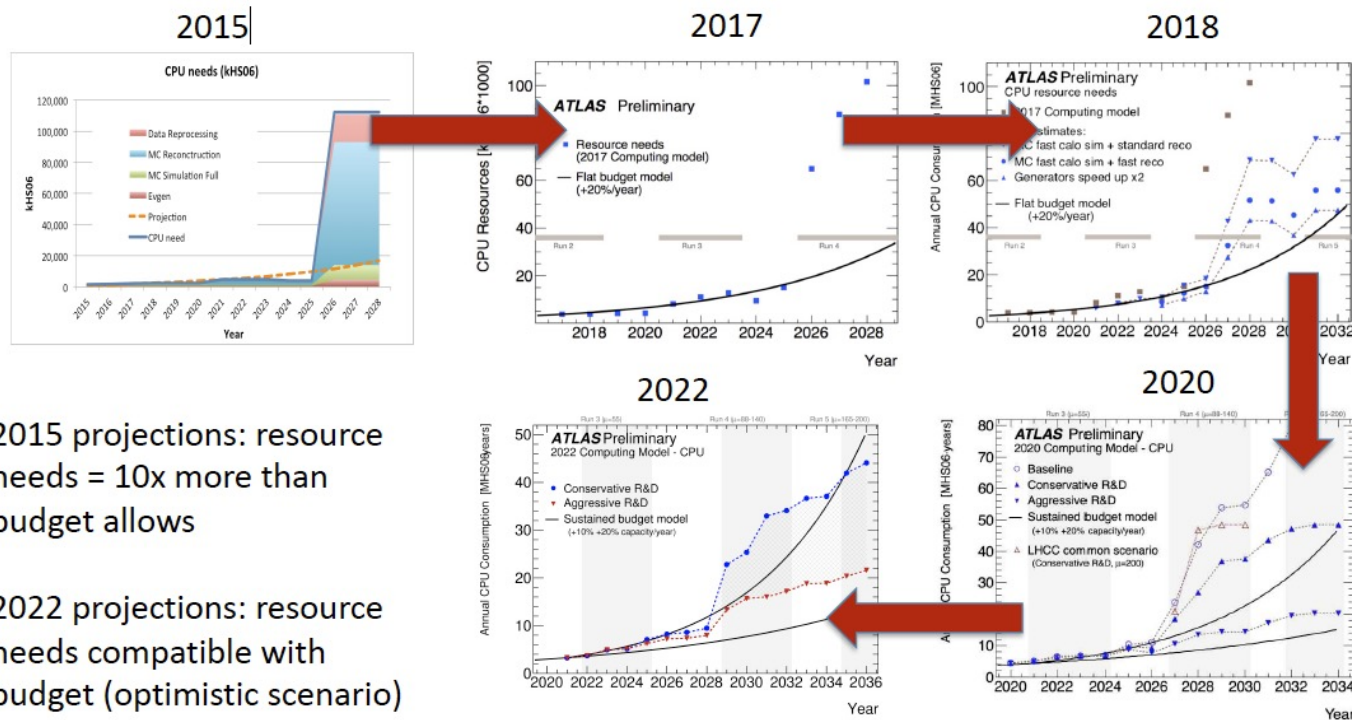
- Tecnologia
  - Parallelizzazione (HP06/\$\$), eterogeneità (HP06/Watt)
- Infrastruttura
  - **Data Lake**: consolidamento dello storage in pochi centri, CPU da centri HPC o cloud commerciali, riduzione costi operations (QoS)
- Computing e Analysis Model degli esperimenti

# Stima necessità risorse (\$\$)



- Le esigenze di risorse di calcolo per ATLAS e CMS sono compatibili con l'ipotesi di flat budget (nelle ipotesi più aggressive)
- La possibile inclusione di risorse con acceleratori per i work-flow centrali permetterà di ridurre le esigenze rendere i modelli più sostenibili riducendo i consumi

# Evoluzione necessità risorse



Simone Campana  
CHEP 23

2015 projections: resource needs = 10x more than budget allows

2022 projections: resource needs compatible with budget (optimistic scenario)

# Una nuova sfida: Il PNRR - ICSC



Il PNRR comporta una sfida enorme per il calcolo INFN

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



L'intera infrastruttura di calcolo INFN è inserita nel Centro Nazionale

- Il Tier1 e tutti i 9 Tier2 vengono potenziati in termini di impianti, rete e risorse

→ I Tier2 dovranno essere tutti attivi per almeno i prossimi dieci anni. Non si prevede la dismissione di nessun centro. Importante l'immissione e la formazione di personale nuovo negli anni

→ sarà comunque possibile differenziare i ruoli

# The ICSC aim and objectives

Create the **national digital infrastructure** for research and innovation, starting from the existing HPC, HTC and Big Data infrastructures ...



... evolving towards a **cloud datalake** model accessible by the scientific and industrial communities through flexible and uniform cloud web interfaces, relying on a high-level support team ...

... form a globally attractive **ecosystem based on strategic public-private partnerships** to fully exploit top level digital infrastructure for scientific and technical computing and promote the development of new computing technologies



# Public Research Institutions Founding Members: a pervasive initiative throughout Italy



### National Institutions

Logos of National Institutions: INFN (Istituto Nazionale di Fisica Nucleare), CINECA, ENEA, Consiglio Nazionale delle Ricerche, GARR (Consortium), INAF (Istituto Nazionale di Astronomia), INGV.

### ICSC founders: a public-private partnership

25 Universities	12 Research institutes	14 Strategic private companies
--------------------	---------------------------	-----------------------------------

09/10/2023 25

# Private 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 at the national level

**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 DEVELOPMENT

*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*

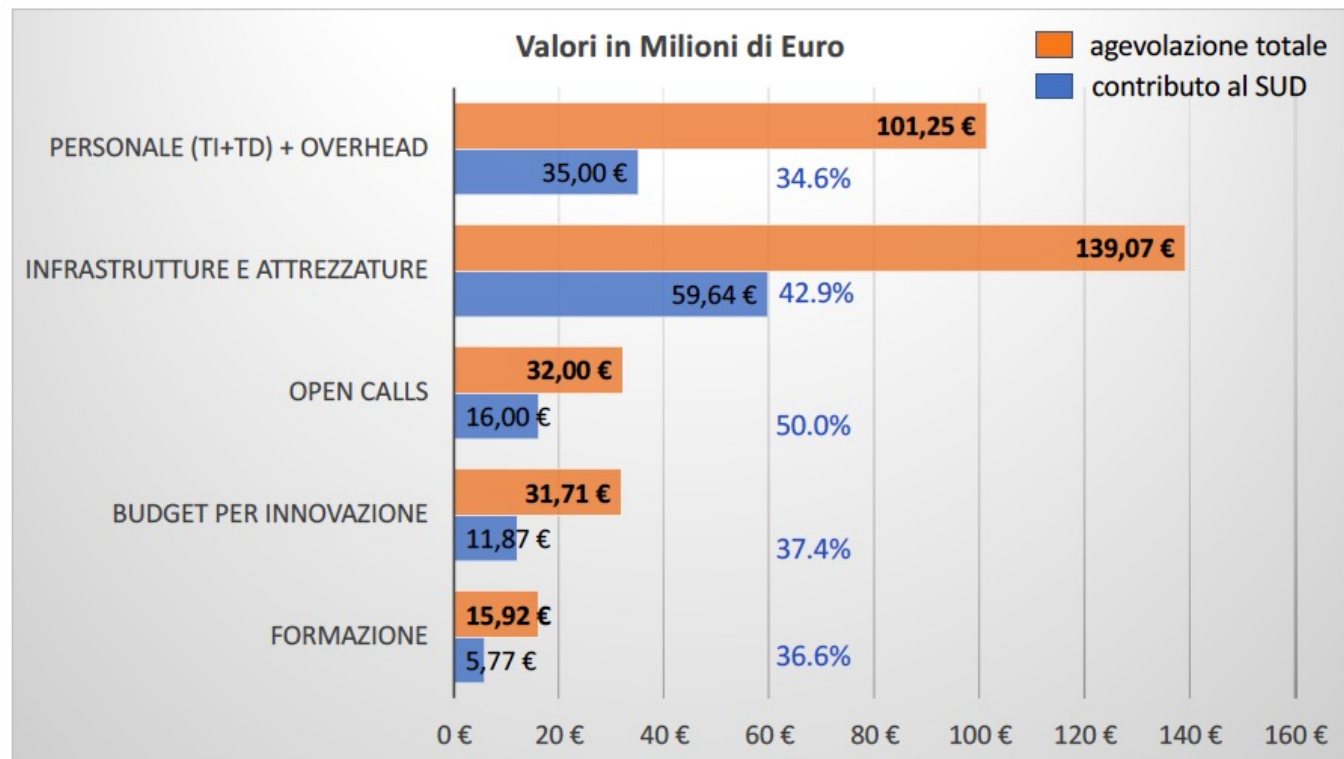
09/10/2023

26

# ICSC - Budget



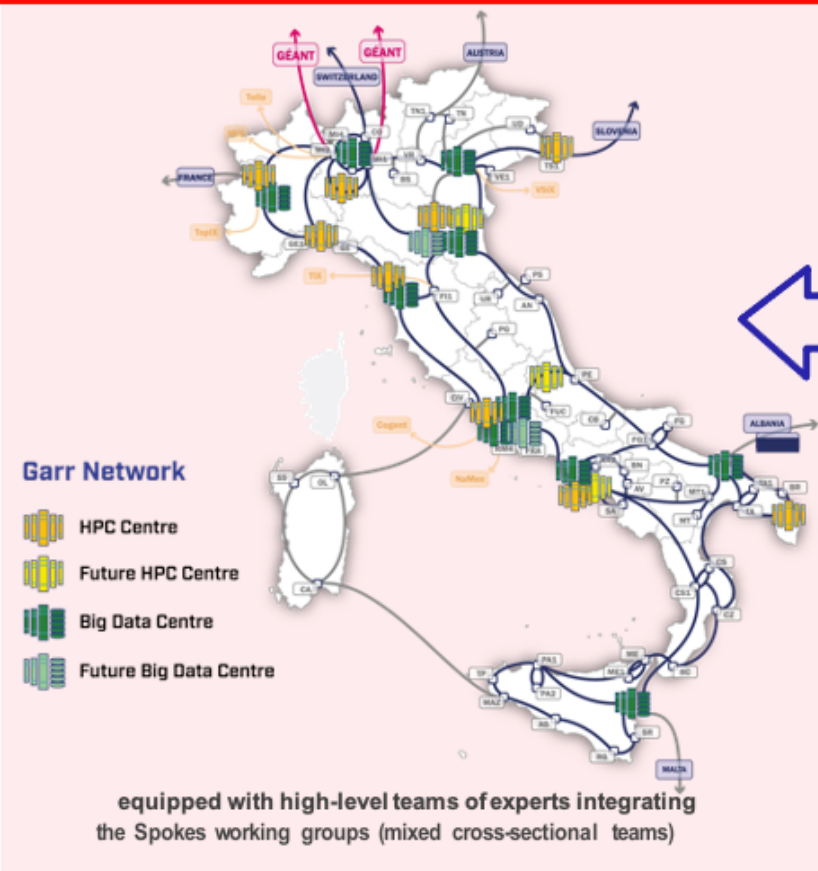
Finanziamento Totale 320 M€



# The ICSC will include 10 thematic Spokes and 1 Infrastructure spoke



## 0 SUPERCOMPUTING CLOUD INFRASTRUCTURE



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

- |   |   |
|---|---|
| <b>1</b><br><b>FUTURE HPC &amp; BIG DATA</b>              | <b>2</b><br><b>FUNDAMENTAL RESEARCH &amp; SPACE ECONOMY</b>           |
| <b>3</b><br><b>ASTROPHYSICS &amp; COSMOS OBSERVATIONS</b> | <b>4</b><br><b>EARTH &amp; CLIMATE</b>                                |
| <b>5</b><br><b>ENVIRONMENT &amp; NATURAL DISASTERS</b>    | <b>6</b><br><b>MULTISCALE MODELING &amp; ENGINEERING APPLICATIONS</b> |
| <b>7</b><br><b>MATERIALS &amp; MOLECULAR SCIENCES</b>     | <b>8</b><br><b>IN-SILICO MEDICINE &amp; OMICS DATA</b>                |
| <b>9</b><br><b>DIGITAL SOCIETY &amp; SMART CITIES</b>     | <b>10</b><br><b>QUANTUM COMPUTING</b>                                 |

# Il Calcolo INFN



## Come si organizza l'INFN?

Nel 2019 si è avviato un processo di analisi e revisione dell'organizzazione del calcolo INFN per razionalizzare il modello di gestione e nel 2022 è stato istituito il Coordinamento Nazionale del Calcolo (CNC)

Per il calcolo scientifico, l'obiettivo è definire le strategie per affrontare le sfide proposte da HL-LHC e dai progetti emergenti:

- Evoluzione dell' infrastruttura INFN
  - Ruolo del Tier1 e dei Tier2
    - Nell'approccio Data Lake, l'infrastruttura attuale costituita da 9 Tier2 ha senso? I Tier2 vanno differenziati assumendo ruoli diversi? Sono tutti necessari?
  - Interazione con il CINECA
    - Il Tecnopolo a Bologna, idea nata nel 2015
    - Leonardo, uno dei due supercomputer pre-exascale Europei
  - Sviluppo del middleware per il Centro Nazionale e coordinamento dei progetti PNRR



# Il C3SN



La gestione strategica e di coordinamento del calcolo INFN viene svolta in un comitato di steering ristretto, il Comitato di Steering del CNC (C3SN)

## Compiti:

- elabora e propone al management INFN la strategia del calcolo e l'evoluzione dell'infrastruttura
- definisce le linee guida e armonizza le attività che vengono svolte nei WG
- mantiene i contatti con le Commissioni Scientifiche, i gruppi di ricerca e i Fondi Esterni
- valuta e armonizza la partecipazione italiana a progetti internazionali e nazionali
- indirizza la rappresentanza dell'INFN nei confronti delle organizzazioni nazionali (CINECA, ICDI, GARR, EPR, ....) e internazionali (WLCG, EOSC .....
- gestisce i finanziamenti del CNC

# La CCR



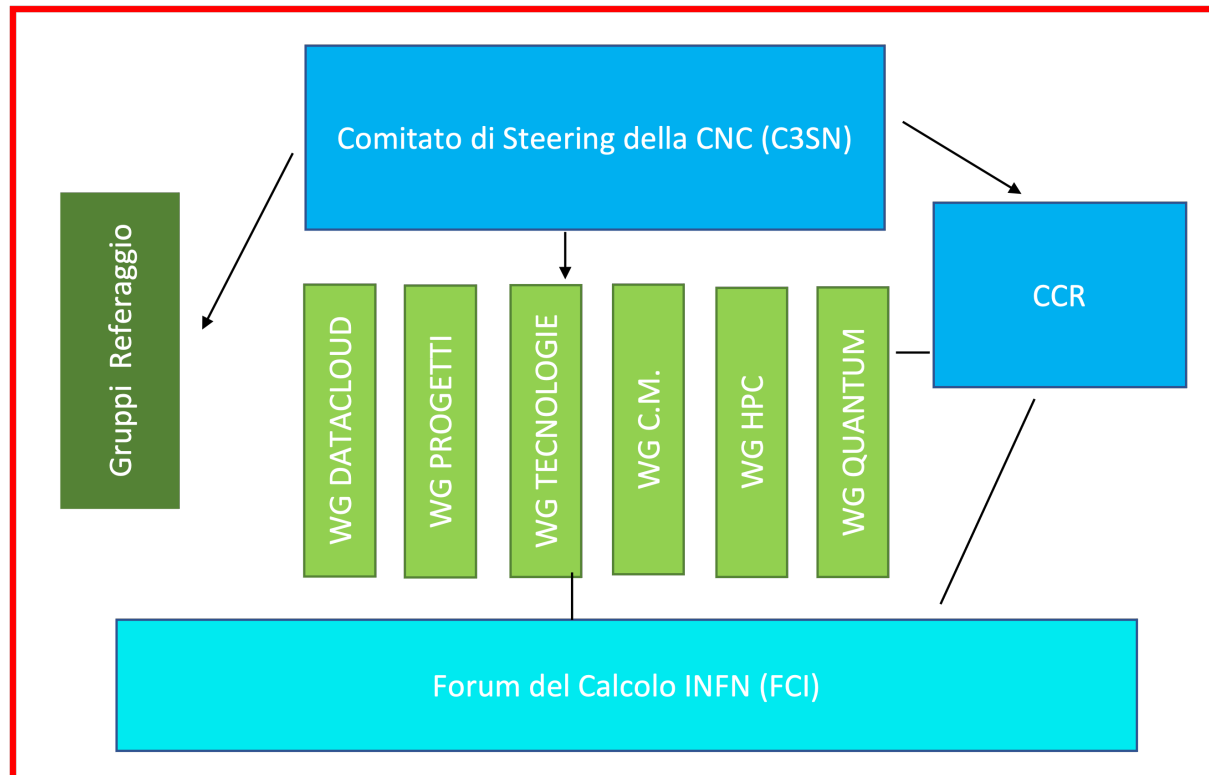
Coordina le attività che riguardano i Servizi di Calcolo e i Servizi Nazionali, l'infrastruttura di rete nazionale e le tematiche di sicurezza informatica

## Compiti:

- definizione e armonizzazione delle attività dei Servizi di Calcolo Locali
- Infrastruttura di Rete nazionale e rapporti tecnologici con il GARR
- Infrastruttura di Business Continuity e Disaster Recovery
- SSNN (Infrastruttura di Autenticazione e Autorizzazione (AAI), mailing, servizi multimediali, servizi web, strumenti collaborativi)
- Sicurezza informatica (di concerto con il RTD)
- Policy di accesso alle risorse informatiche
- Acquisti centralizzati di software e manutenzioni apparati di rete
- Supporto al SI



# I Gruppi di Lavoro



# I Gruppi di Lavoro



---

## **WG DataCloud**

Infrastruttura: sviluppo e gestione del modello architetturale di Data Lake e delle infrastrutture ISO, supporto alla gestione e all'operazione dei siti, agli esperimenti e agli utenti, sviluppo nuovi servizi

## **WG Progetti Esterni**

supporto, direttive e linee guida per organizzare le attività nei progetti esterni

## **WG HPC**

coordinamento delle attività in ambito HPC, porting dei codici e sviluppi sw, interazione con CINECA e utilizzo Leonardo

## **WG Quantum Computing**

coordinamento delle attività di calcolo quantistico

## **WG Computing Model**

studio dell'evoluzione dei CM degli esperimenti in corso o futuri, ottimizzazione dei modelli per l'utilizzo efficace sull'infrastruttura INFN

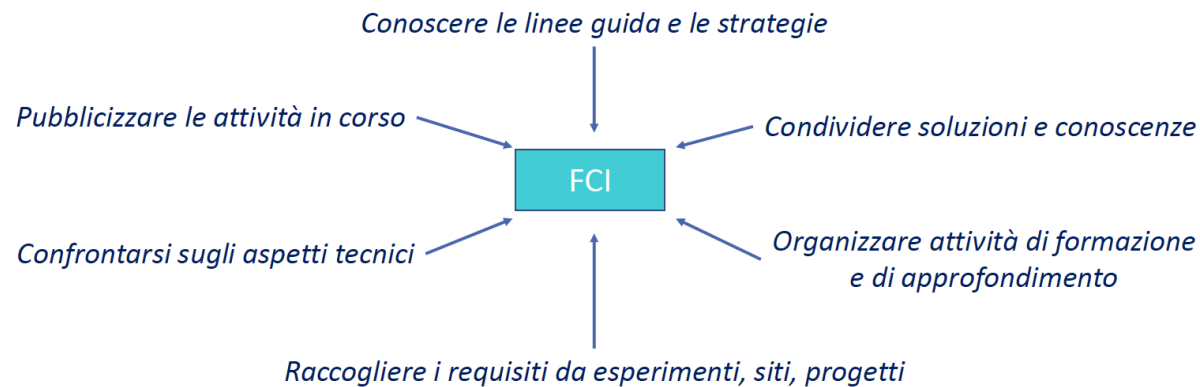
## **WG Tecnologia Informatica**

studio dell'evoluzione della tecnologia informatica collaborando con istituzioni internazionali (Cern, Hepix), supporto tecnico agli esperimenti nella preparazione delle gare e al gruppo di referaggio, R&D

# Il Forum del Calcolo



- Luogo dove tutte le comunità coinvolte nel calcolo (esperimenti, sviluppatori, siti .....) possono confrontarsi per discutere, rendere pubbliche e condividere le attività
- discussione delle linee guida e delle strategie sul calcolo sviluppate nel C3SN e loro implementazione
- discussione delle attività e delle esigenze tecnologiche delle varie parti: esperimenti, siti, progetti, etc
- confronto tecnico sulle attività in corso nei WG affinché siano funzionali alle esigenze delle parti
- condivisione delle soluzioni tecnologiche adottate e delle esperienze specifiche



# Conclusioni



Il personale coinvolto nel calcolo INFN

- ricercatori INFN e universitari, tecnologi e tecnici
  - esperimenti, progetti, siti, sviluppo software e servizi, operations, servizi calcolo .....
  - tempo indeterminato e determinato, assegnisti, borsisti, dottorandi, post-doc .....
- è di alcune centinaia di unità

Tutti sono impegnati per vincere le sfide che ci vengono proposte a livello mondiale dagli esperimenti e dai progetti europei, dal management e dalle strutture locali.

I progetti PNRR sono una sfida enorme e contiamo molto sul contributo del nuovo personale assunto