Il calcolo nell'era post HL-LHC

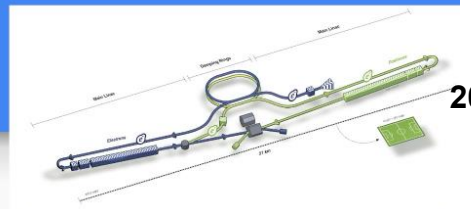
Luca dell'Agnello INFN-CNAF

November 7, 2024

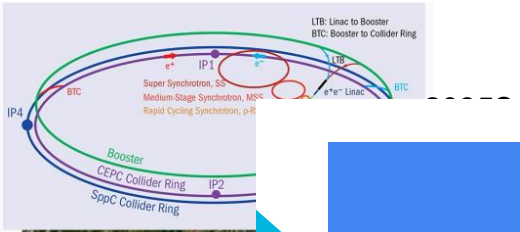
HL-LHC is not the end of the story!

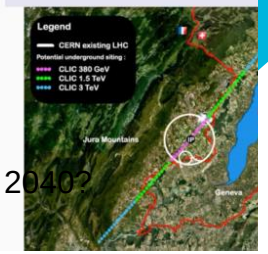
Beyond #1?

- ee machines (CLIC, ILC, FCC-ee,CepC ….)
  - **No major computing problem expected**
  - FCC-ee initial event size estimates are 0.01 - 0.1 the current LHC-pp, and 20 years later
  - Even a huge increase in DAQ channels / interaction rate can hardly be a problem



2040?





2040?

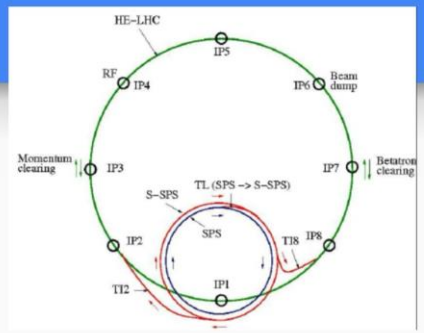
Courtesy of T. Boccali

See: https://doi.org/10.1016/j.revip.2019.100034

Beyond #2?

- hh machines (FCC-hh, HE-LHC, …)
  - …go as high as you want: FCC-hh has (wrt to current LHC)
    - <PU> ~30x (and 5x HL-LHC)
    - Similar collision rate
    - Event sizes not yet known atm
  - But: there is at least a +20y between them, which reduces the problem
  - **HE-LHC** parameters are intermediate between HL-LHC and FCC-hh, but time scale is still at least 2035

- **My thoughts: the step LHC→ HL-LHC in 2026 is the biggest; if we can make HL-LHC computing work, we have a clear path**



2040?

2050?

How are computing resources linked to machine / experiments parameters?

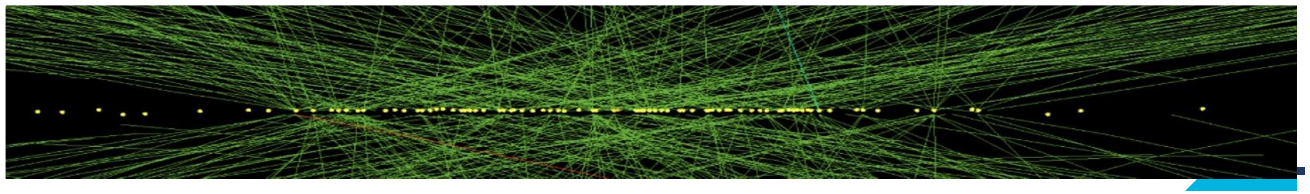
Courtesy of T. Boccali

- # events collected = Experiment live time * Experiment rate to offline
  - LHC RunII: 7Msec * 1000 Hz = ~ 10 B events
- Bandwidth, total storage = # events collected * (1+ f_{MC}) * $F_{STORAGE}$(<PU>)
  - $F_{STORAGE}$(<PU>) ~ linear in <PU>
- Computing power = # events collected * (1 + f_{MC}) * F_{CPU}(<PU>)
  - F_{CPU}(<PU>) superlinear in <PU>
- Storage is also ~ integral with time
- $Storage_{YearN+1}$ = $Storage_{YearN}$ + $Delta_{NEW\ EVENTS}$

In the end, main parameters are

- Trigger rate
- Live time of the Accelerator
- <PU>
- f_{MC} (MC production needs)

PU: the # of pp interactions per single bunch cossing



A brief recap: the data centers infrastructure

- Present INFN computing (mainly) based on the WLCG infrastructure and services
  - Tier1, 9 Tier2s data centers
  - At worldwide level 10 Tier1s and ~200 Tier2s
- Tier1 is the main pillar of INFN infrastructure
  - The new Tecnopolo data center, capable of meeting the requirements of the HL-LHC era and beyond
    - DLC area able to host ~4 MHS06 of computing power
  - ~50% of the INFN computing resources
  - More than 60 supported collaborations
- Tier2s
  - Renewed with NRRP projects (but no DLC!)
- Datacloud
  - INFN initiative to integrate (in the medium term) all the Tier2s and the Tier1 in a common data lake for all INFN experiments At present:
    - Federation of local cloud (e.g. Cloud@CNAF, CloudVeneto,....) + backbone
  - Also, effort to build a distributed management of the cloud

Tecnopolo: a game changer

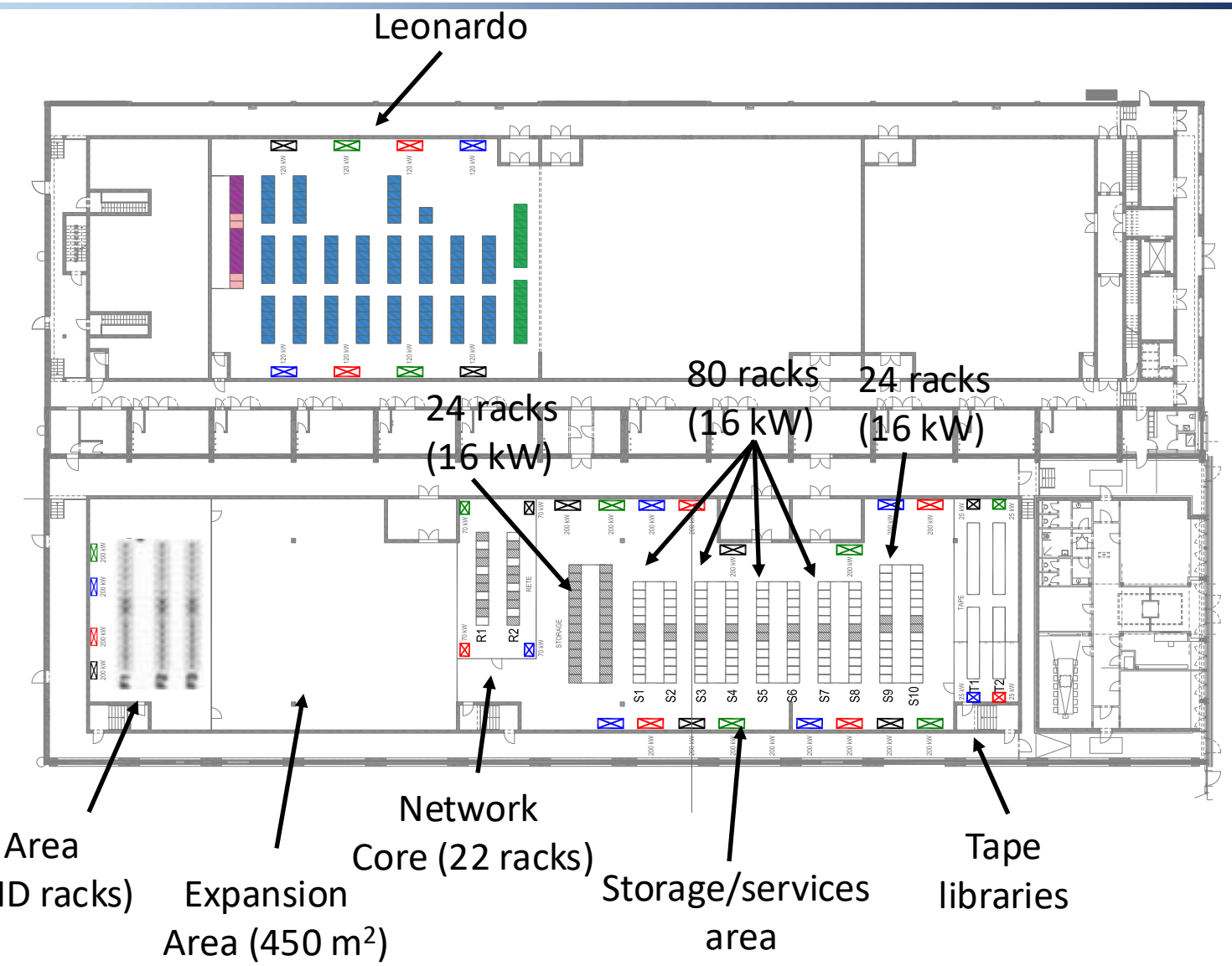
- Joint initiative by INFN and CINECA to build a new data center at the Tecnopolo for both the INFN Tier1, Leonardo and other future systems
- Several opportunities for INFN offered by the new location:
  - Higher electrical power availability (up to 10 MW, currently ~1.6MW) and larger room space (~3000 m^2, currently ~800 m^2) allow for a robust increase in installed resources
    - ICSC resources are also hosted
  - Proximity to Leonardo facilitates its exploitation
    - Soon also Lisa and (probably) the AI Factory (2026)
- **With the new data center, we have an infrastructure capable not only of hosting INFN Tier1 in the HL-LHC era and of meeting the requirements of future experiments, but also of accommodating resources from scientific collaborations and projects in other scientific domains.**

The Tecnopolo

The layout of Tecnopolo data halls

CINECA data hall

B5 (INFN data hall)



Storage area with 132x16 kW racks
- 44 racks for storage (400 PB)
- Cloud infrastructure
- ISO-certified Cloud
- Other ICT services

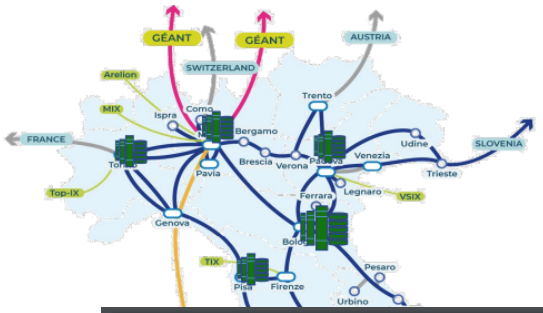
2 options for racks in CPU area:
- 40 kW racks (with cooling plenum)
- Racks with DLC (80-90 kW) in Phase 2 (up to 3.4 MW, 4 MHS06)

Up to 4 tape libraries O(EB) with the current technology

Network links:
- GARR GP and LHC OPN/ONE uplinks
- 1.2 Tb/s LAN extension with CNAF
- 1.6 Tb/s LAN extension with Leonardo
- 4x100 Gbps DCI with CERN

Finanziato dall'Unione europea NextGenerationEU | Ministero dell'Università e della Ricerca

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

RETE GARR

INFN

INFN Istituto Nazionale di Fisica Nucleare

From a talk by C. Grandi at CCR 2024

Finanziato dall'Unione europea NextGenerationEU | Ministero dell'Università e della Ricerca

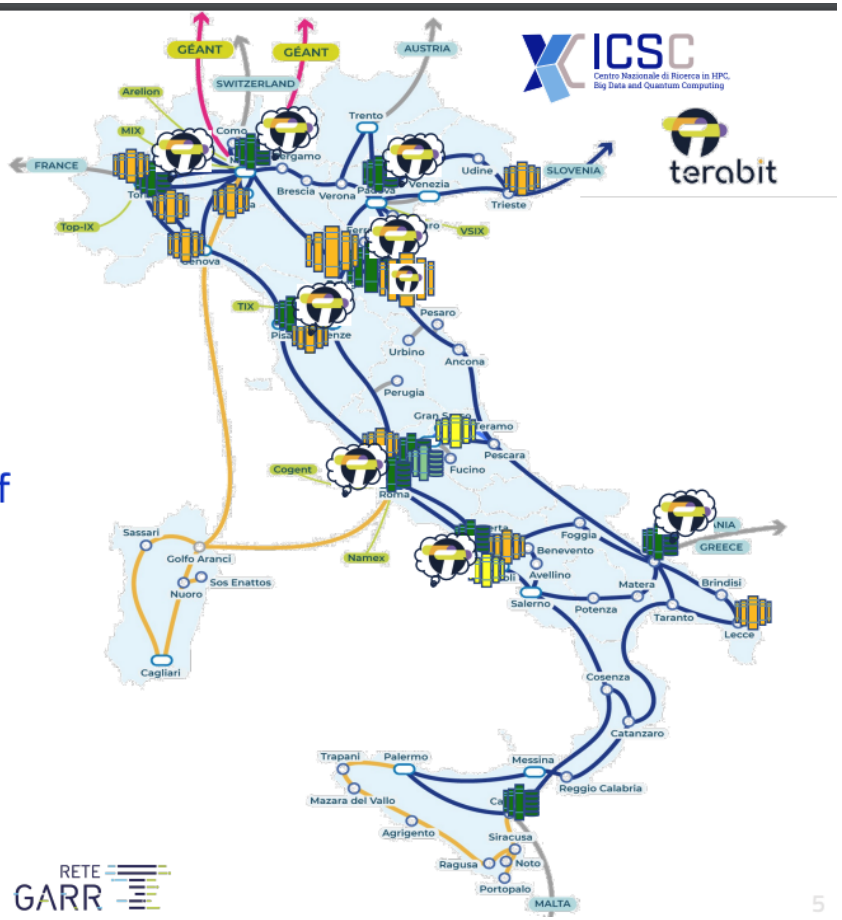
DataCloud is the basis for the Italian Cloud Federation

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, …



RETE GARR

See: https://agenda.infn.it/event/40160/contributions/230276/attachments/119744/173731/Grandi-WS-INFN-Palau.pdf

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

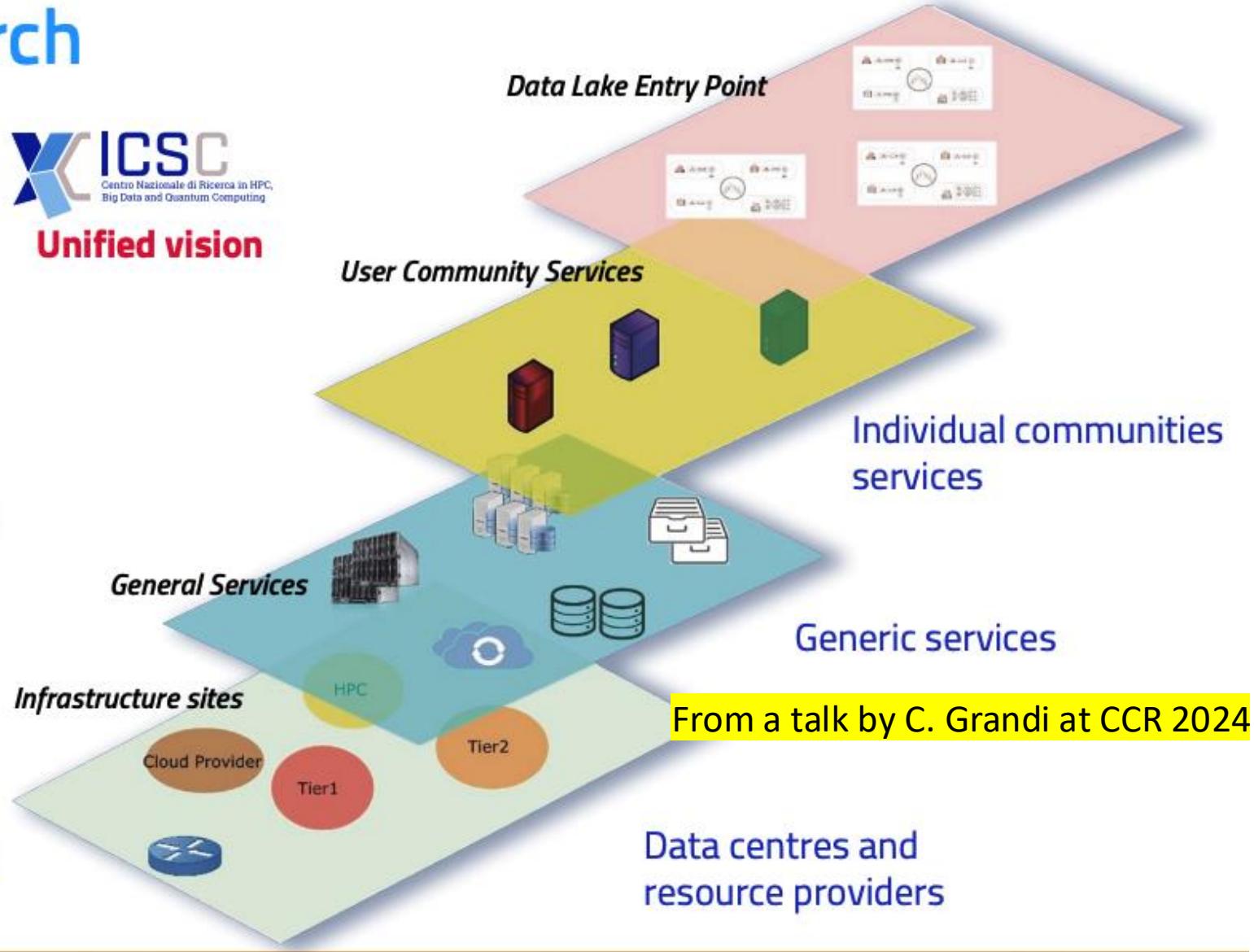
A clear separation between the physical and the logical levels

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



From a talk by C. Grandi at CCR 2024

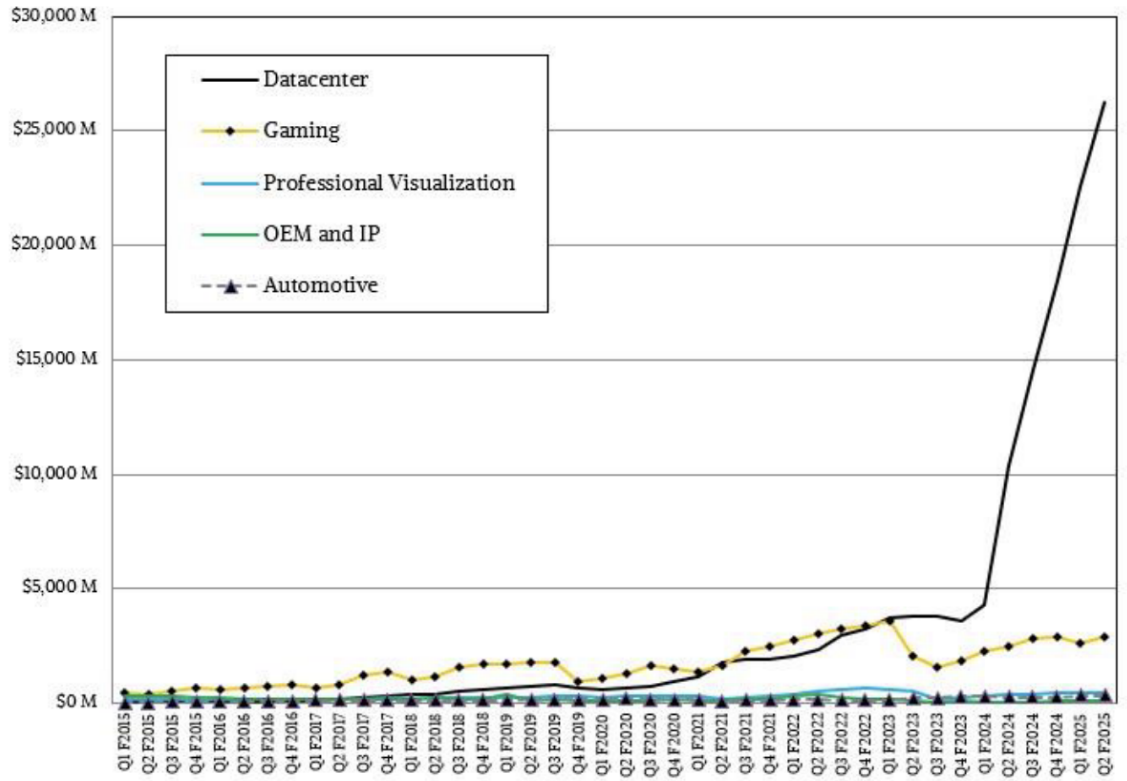
Some (not so small!) points to address

- Several structural issues faced also in the latest strategy
  - The role of HPC
  - The transition from Grid to Cloud
- Costs
  - Resources
  - Operations
  - Middleware
- Ability to keep up with and/or anticipate technological evolution
  - Quantum
  - AI
- The (new?) requirement to take environmental impact into account
  - ARM, RISC-V
  - DLC
- The opportunities given by Datacloud and ICSC
  - Standardization of tools 😀
  - Increase in communities supported ☹️

Can we make predictions?

From A. Sciabà's talk at CHEP 2024

- **Extremely difficult even beyond just 1-2 years**
  - The demand can change unexpectedly. See the case of GPUs
  - E.g., will the AI bubble burst, suddenly? What if profits do not materialize early enough?
- **What are the "hottest" trends? Some examples:**
  - Sustainability and CO_2 emissions
  - Increasing memory bandwidth and latency requirements
  - Increasing power density and liquid cooling
  - Competition between spinning disks and flash memory



Nvidia revenues (Source: The Next Platform)

see: https://indico.cern.ch/event/1338689/contributions/6080127/attachments/2953746/5192935/Technology%20Tracking.pdf

The hardware cost

- “Pledged” resources
  - CPU - 100 €/core (including network)
    - Running costs comparable to hw cost
  - ”Standard” disk: ~100 €/TB-N (i.e., including parity, servers, network)
    - +~15% for software, running costs ~20% over 5 years
  - Tape: 10 €/TB (cartridge only!)
- Core equipment
  - Network backbone (core switches, router)
  - Tape libraries, tape drives
- Ancillary systems (e.g., monitoring etc..)
- **Costs have decreased over the last 20 years, allowing for the adoption of the “flat budget” mantra (with the same amount of money, you can buy 10-20% more resources each year)**
  - **It is not clear if this ”rule” will hold also in the next few years (apparently not!)**
  - **If this is the case, extra effort needed from the experiment to do more with less**

Disk and CPU lifetime: 5-7 years

Unitary costs similar for other Tier1s

The operational model

- Tier1
  - 24x7 service provided (as stated in the WLCG MoU)
    - A best effort really but with full redundancy on infrastructure and all services
      - This is not for free!
    - Same SLA for all collaborations
  - ~21 FTE for the management of INFN Tier1
    - Figures nearly consistent (slightly less) with other equivalent centers (KIT, IN2P3)
    - A significant portion of the staff (33%) is needed for data management
    - Key role played by User Support helping the exploitation of the data center
- Tier2s
  - 24x7 not a requirement
  - Generally, 1-3 FTE/site
- The cloud(s) operated with the contribution of additional people (at CNAF +3 FTE)
- **To be investigated if these figures will at least remain constant in the scenario of INFN Datacloud with common services but with also additional communities (ICSC)**
  - **At least the user support model should be re-discussed/reviewed**

The middleware

- The middleware is the “glue” of the Grid (and now of the Cloud too)
  - Continuous developments since early 2000s
  - Some key tools (e.g., AAI) developed by INFN and widely adopted
  - Specific WLCG tools used also by other communities
- CNAF develops and/or maintains some middleware products (~5 FTE)
  - AAI infrastructure
    - VOMS – since 2003, now phasing out (maintenance still needed)
    - IAM – in production and active development
  - Data Management
    - StoRM Webdav

Used by WLCG and other communities

Used by INFN Tier1 and other few WLCG data centers

- Other products developed in the framework of Datacloud project
  - E.g. the INDIGO PaaS Orchestrator to coordinates the process of creating deployments on the IaaS platforms made available by the federated providers.

+5 FTE

- And, of course, other products developed by other centers
- **Question: Should we consider using more and more commercial products?**

The Times They Are A-Changin'

- Besides the (slow?) shift from Grid to Cloud paradigm
- Integration of non-Grid resources
  - Use of HPC centers: no more a test!
  - CINECA is providing (pledged) resources from Leonardo for:
    - CSN4 use cases (booster partition with GPUs)
    - Tier1 users (standard x86 processors + opportunistic computing on booster)
  - In Germany HPC centers will provide CPU instead of Tier2s
- Use of on-x86 architectures
  - Not only GPUs
  - ARM and RISC-V nodes already available in some WLCG sites
  - At CNAF a small installation (4 nodes with Ampere)
    - ~50% more efficient in terms of HS06/W
    - Used so far only from Atlas, Alice, CMS
- At Tecnopolo will be also available a Quantum computer and (probably) an AI machine
  - On Leonardo several quantum simulators already available
  - Neutral Atoms Analog Quantum Simulator (140 qubits) – Q2 2025
  - Superconducting Digital Quantum Computer (54 qubits) – Q1 2025

Not plug and play: some real effort needed to make it usable!
E.g., no direct WAN access, no common AAI
In general, every HPC center needs a specific integration effort

About quantum computers….

IBM Q may 2016
5 qubits

IBM Q may 2017
17 qubits

IBM Q oct 2017
50 qubits (labo)

IBM Q System One jan 2019
20 qubits

Rochester sept 2019
53 qubits

Nature paper on quantum utility with QEM june 2023

Eagle nov 2021
127 qubits

System Two 2023

Osprey may 2023
433 qubits
(retired in 2023)

2017 2018 2019 2020 2021 2022 2023 2024

IBM oct 2017
HPC 56 qubits simulator

IBM Q System One jan 2020
28 qubits

Hummingbird 2020
65 qubits

Condor 2023
1,121 qubits (retired)

Heron r1 2023
133 qubits

Heron r2 2024

… and quantum computing

From a talk by M. Grossi at 2nd INFN workshop on Quantum Computing

Conclusion

- Complexity & learning theory mostly gives us insights into **worst-case behavior**

  → ML: Learning theory predicted deep neural networks to not be trainable

  → Optimization: The travelling salesperson problem is NP-complete. An instance with 85900 cities was solved in 2006. Exponential complexity does not imply infeasibility
- **Benchmarking** can help us to understand the behavior on specific instances
- We need to make a comparison of **computational cost** - may lead to poly advantages!
- Change the goal: quantum advantage will be unlikely in many cases **BUT** we can identify promising paths for **hybrid computational advantages**
- We can train the loss on a classical device, and sample on quantum (GENERATIVE MODELs)

  → larger devices for high-quality data?
- What's the role of data?
- Community goal is bridging the gap between near-term and fault-tolerant quantum machine learning

See: https://agenda.infn.it/event/42801/timetable/?view=standard

A common goal

- To address some of these issues, but also due to the arrival of other collaborations/communities with different needs in the playground, there is an evolution towards new models
- Efforts to build a common infrastructure including all the ingredients (i.e., data centers, middleware, operation models, computing models) through some initiative/projects
- SPECTRUM project (HEP, Radio Astronomy) working to harmonize several technical aspects
  - AAI
  - Data management
  - Workflow Management
  - Software tools
- JENA (HEP/ECFA, NuPEC, APPEC) is expected to provide inputs to the European Strategy for Particle Physics Update
  - Spectrum will give input to JENA

(No) conclusions

- As stated at the beginning of this talk, it is quite impossible to describe a realistic scenario on computing in a time horizon longer than a few years.
- On the other hand, naive extrapolations for the experiments after HL-LHC do not show alarming scenarios, besides the costs (hw and staff!) and these probably can be addressed exploiting new opportunities (new architectures, shared infrastructure,..)
- The strategy is a good opportunity to address these issues and study possible solutions