INFN

Scientific Computing at INFN (and beyond)

1951 2021 infn

70 YEARS OF RESEARCH SHAPING THE FUTURE

IBERGRID 2022 – Faro, 10-13/10/2022

Davide Salomoni (davide@infn.it)

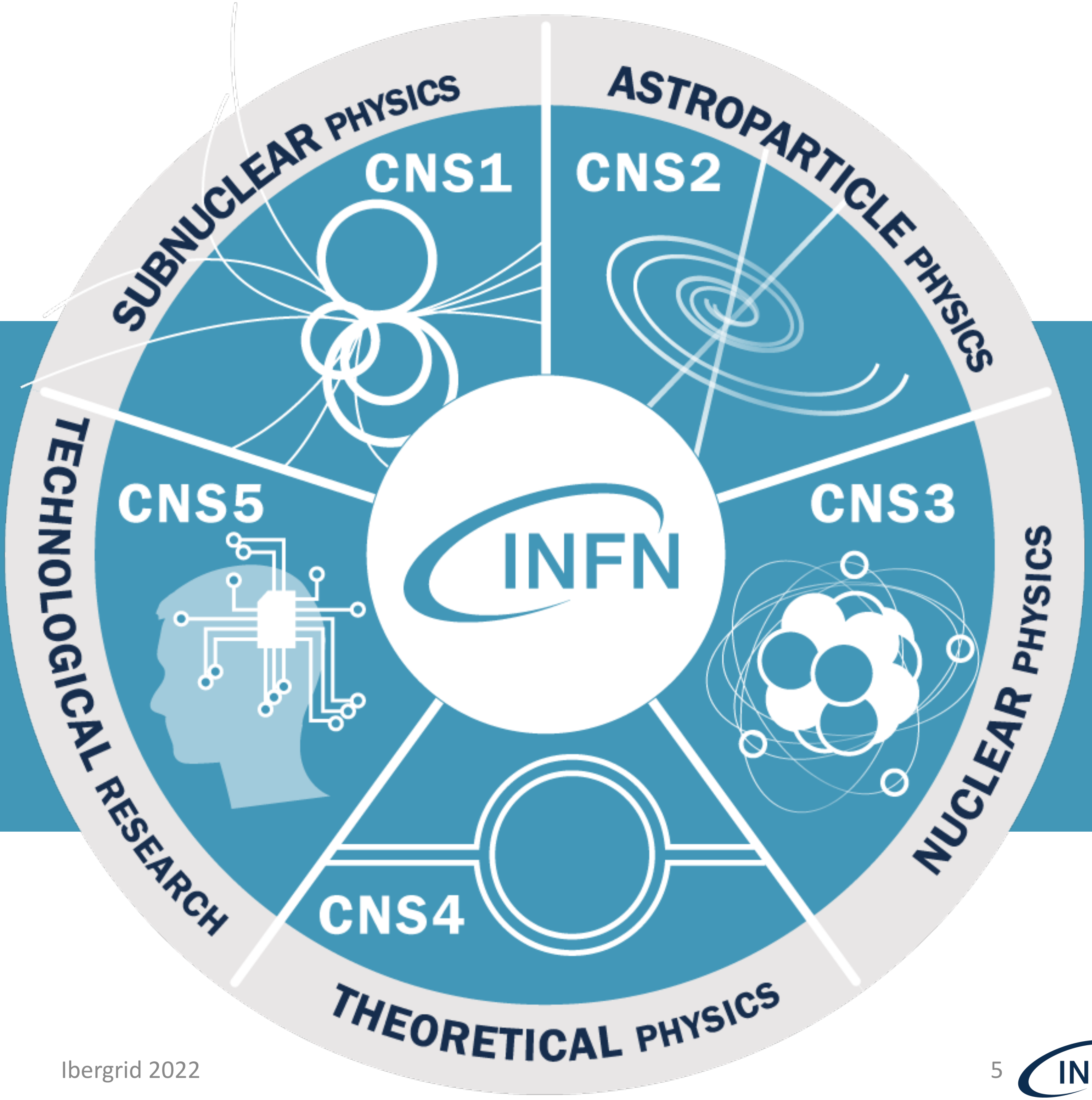
Table of Contents

About myself

- Director of Technology at the Italian National Institute for Nuclear Physics (INFN) with some 30 years of experience in distributed computing for science.
  - Based in Bologna at CNAF, the INFN National Center for Data Processing and Computing Technology Research.
  - Coordinator of the INFN-wide computing & storage infrastructure.
  - Member of the INFN Computing management board.
  - https://www.linkedin.com/in/davidesalomoni/, email: davide@infn.it
- Adjunct Professor at the University of Bologna for the PhD program in "Data Science and Computation" and for the MD in Bioinformatics for the courses "Infrastructures for Big Data Processing" and "Biomedical Data Bases".
  - https://www.unibo.it/sitoweb/d.salomoni/en

INFN and Computing

The **5 research lines**
and the INFN National Scientific Committees



The INFN **Facilities**

■ 4 National Laboratories

■ 20 Divisions

■ 6 Associated groups

■ 3 National Centres and Schools

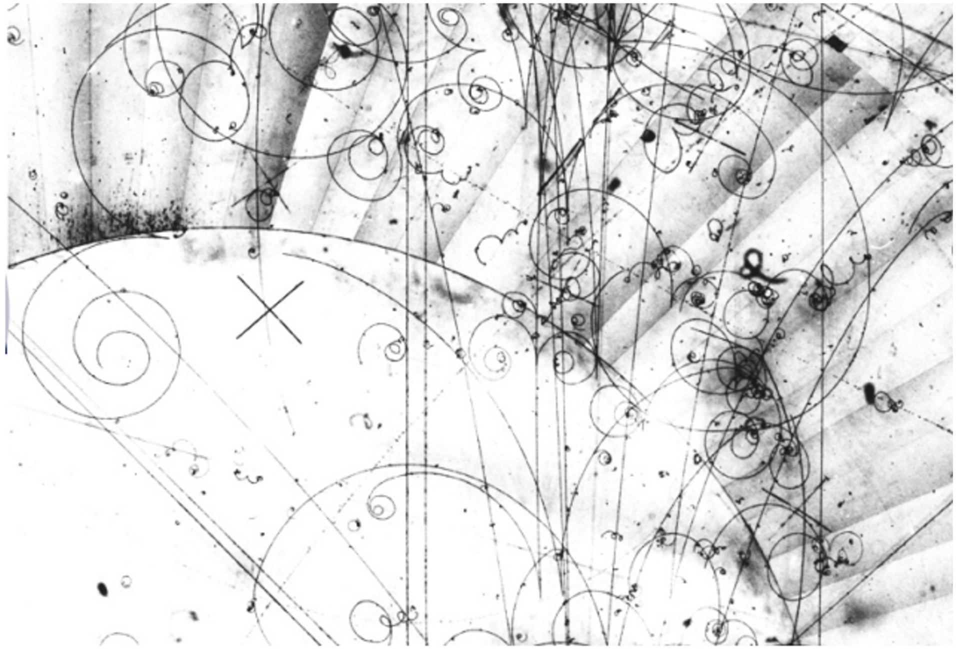
■ 1 International consortia

INFN

A brief history of computing @ INFN

1962

CNAF ("*Centro Nazionale Analisi Fotogrammi*") was founded, dedicated to what was at the time the most technologically challenging analysis method: bubble chambers images

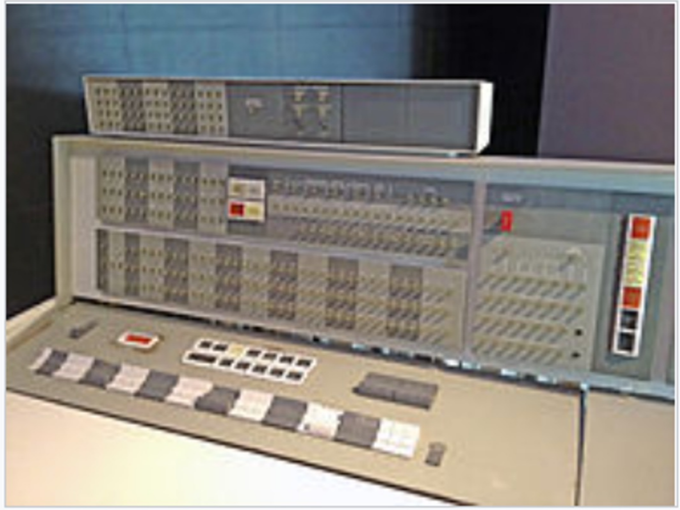


The launch: the first director of CNAF Massimo Masetti (on the left) speaking with Luigi Gui (in the center) at the CNAF inaugural ceremony

1962 → **The 60s**

The number of films to be analyzed increased to 100k's per year, introducing the need for more automation in the analyses.

"Computers" appeared at CNAF: IBM7094 and later IBM 360/44



IBM 7094 operator's console showing additional index register displays in a distinctive extra box on top. Note "Multiple Tag Mode" light in the top center.

~ 200 kflops

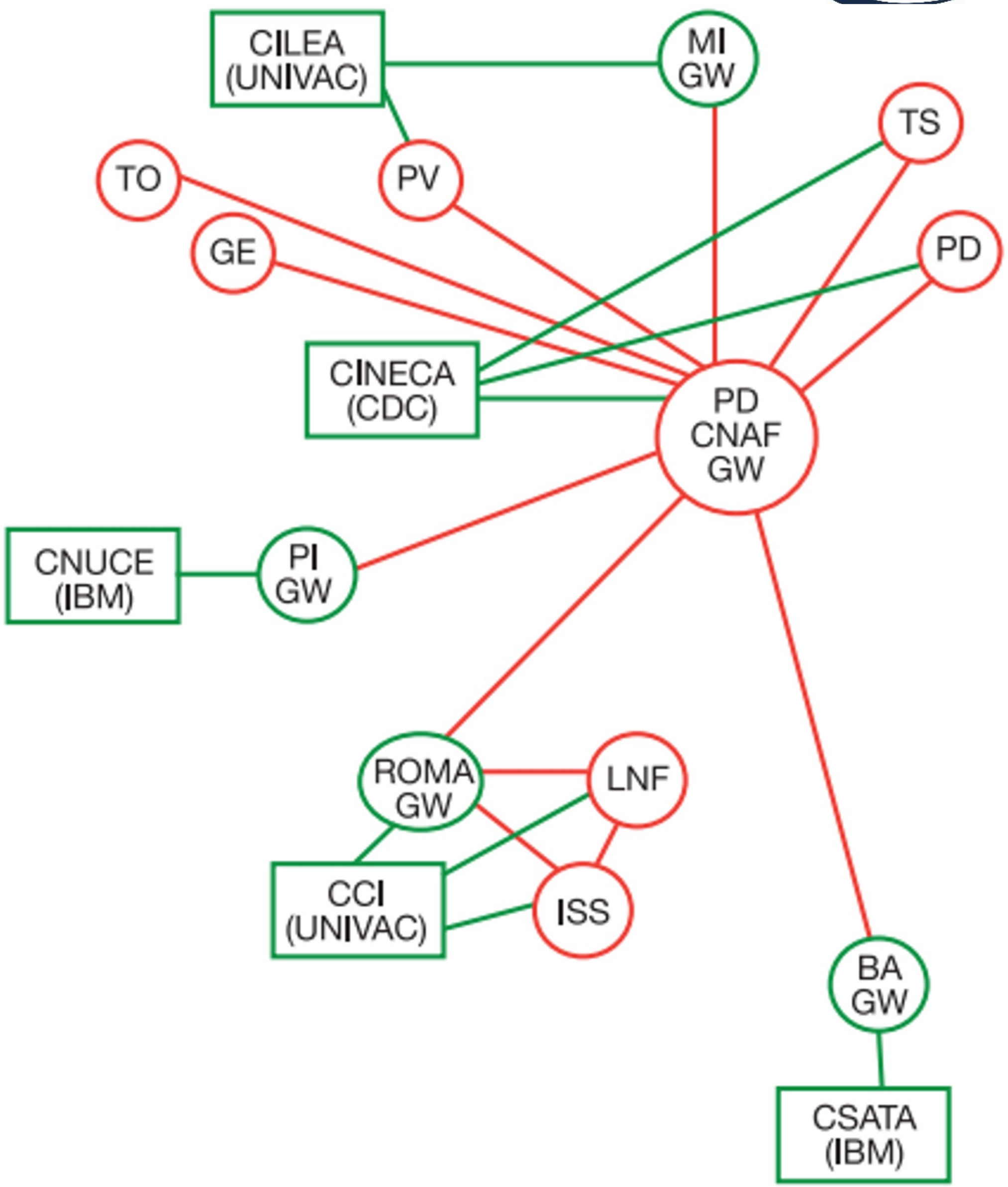
IBM System/360 Model 44

System/360 Model 44 front panel

Manufacturer	International Business Machines Corporation (IBM)
Product family	System/360
Release date	August 16, 1965
Discontinued	September 23, 1973
Memory	32–256 KB Core

1962 | The 60s | The 70s-80s

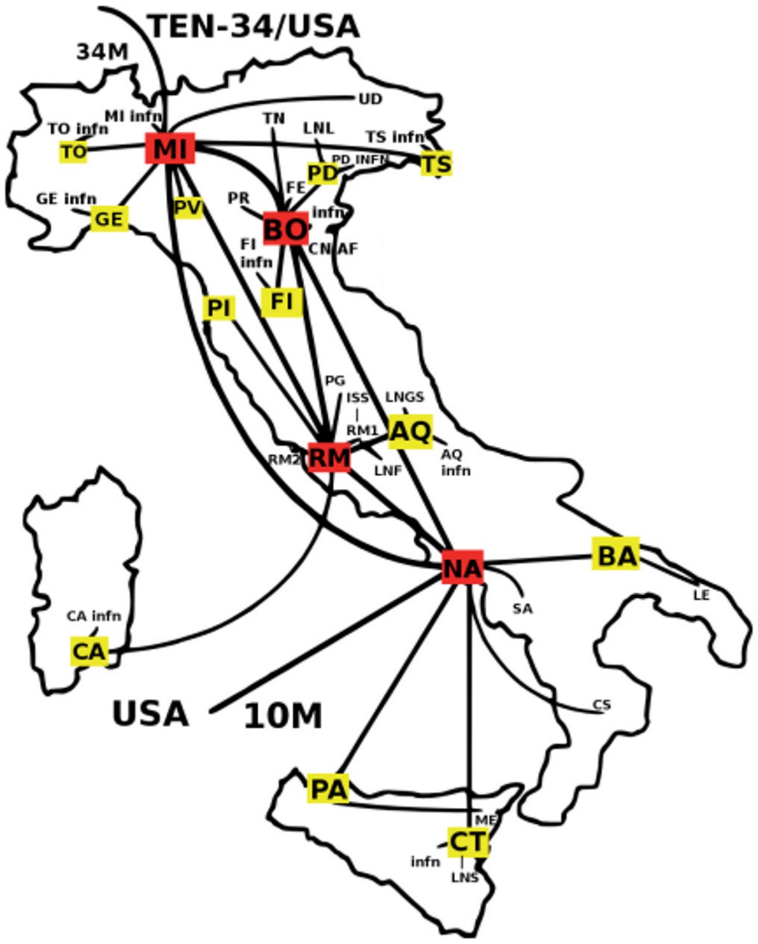
Computers became more and more popular among physicists; due to the distributed nature of INFN, they were installed in different structures and mostly handled independently.
From the need to allow intercommunication, the <u>INFNet project</u> was started, using dial-up connections. <u>CNAF</u>, with its technology-related mission, became the central node of this effort.
In the early 80s, a connection was built to CERN (via CERNet), first for direct access and later to FNAL.
At the end of the 80s, INFNet was topping 64 kbit/s.



1962 → The 60s → The 70s-80s → 1988-1995

Remote access to computers was quickly becoming a need in other scientific domains; GARR was pioneering 2 Mbps connections by 1998, starting with a CERN-CNAF link and later with connections to CINECA, Rome and Milan. This is the infrastructure that handled LEP, TeVatron, SLC computing.

Over time, that became the backbone of research networking in Italy, which reached 34 Mbps by 1995. Still today, research networking is handled in Italy by GARR.
By that time, we were in the planning of the “LHC” era, and it was clear that Computing would have been a major effort for HEP and for INFN specifically. CNAF was again having a central role for INFN Computing.



1962 → The 60s → The 70s-80s → 1988-1995 → 2000-2005

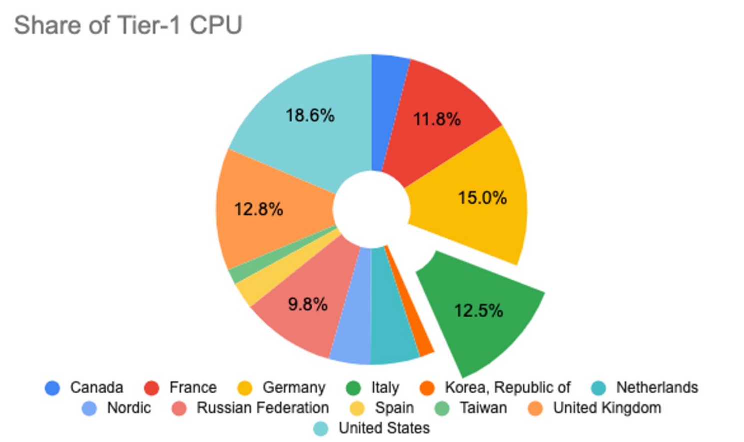
Ten main international centers were selected to host the Worldwide LHC Computing Grid (WLCG):

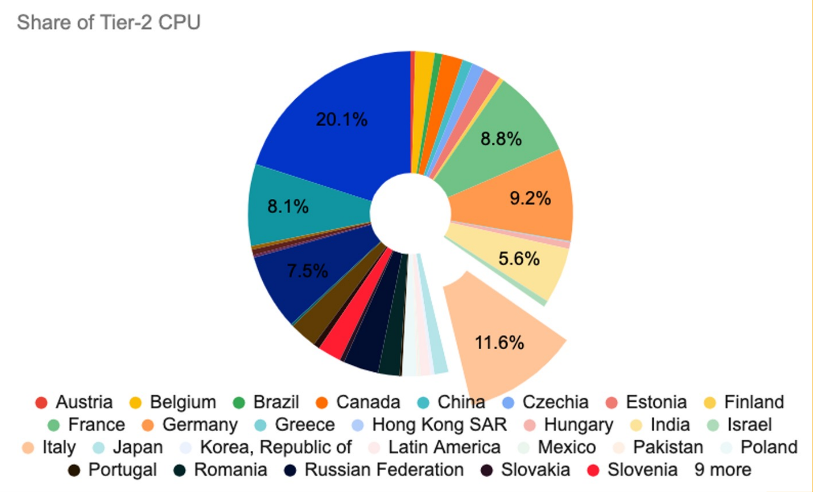
- In Italy, this was the Tier-1 at CNAF (red in the picture)
- 9 additional “Tier-2” centers were added, at LNL, LNF, Turin, Milan, Pisa, Rome, Naples, Bari, Catania (yellow in the picture)
- Then came the GRID, the Cloud, …

- All these centers are still operational, even if their size has increased ~100x since then, and their interconnectivity now reaches multiples of 100 Gbps, thanks to the GARR-X network.
- Collectively, our distributed infrastructure currently offers about 140,000 CPU cores, 120PB of enterprise-level disk space, 100PB of tape storage.

WHERE ARE WE TODAY?

- INFN Distributed Computing federation delivers the LHC experiments O(7-20)% of their computing budgets
- ”non LHC” (VIRGO, Astro, Nuclear, …) is ~ 10-20% of the total
- Sites are of top quality among their peers, and have worked uninterruptedly for the last ~15 years
- In many cases, the infrastructures are close to those deployed in 2005 or slightly after
  - They could be bigger (CNAF uses ~1 MW, more recent centers are leaning towards ~10 MW)
  - They could be “greener”
    - Free cooling, direct cooling, …
  - They could use more recent technologies, hardware and software
    - Go towards a national cloud, implement a datalake model for storage, …





In the meantime, computing-related projects and collaborations grew…

- The development of **Scientific Computing within INFN** was originally driven by the needs of its own theoretical and experimental communities. However, being at the forefront of computing in research seeded **many projects with a much broader scope**.
- The **key overall driver** was always to let our users **effectively exploit all available resources and technologies**.

"Preparing the GRID"

DataGRID, eGee Enabling Grids for E-sciencE, LCG, EGI-Engage, EGI-InSPIRE

"Preparing the Cloud"

EOSC-hub, EOSC Future, interTwin, Skills4EOSC, INDIGO - DataCloud, eXtreme DataCloud, Hybrid DataCloud, EOSCpilot The European Open Science Cloud for Research Pilot Project, EOSC-Pillar, DICE, HELIX NEBULA THESCIENCECLOUD, EGI-ACE

EUROPEAN OPEN SCIENCE CLOUD, EuroHPC Joint Undertaking

"Expanding beyond HEP"

ESCAPE, Asterics, EOSC Future, C-SCALE, elixir ITALY, IoTwins, EXSCALATE 4COV, gaia-x, LifeWatch ERIC, HARMONY ALLIANCE

Computing challenges (as we see them)

The overall framework of computing challenges

- The next generation of High-Energy Physics (HEP) and of many other experiments presents **unprecedented needs for computing**, apparently close to “impossible”.
- They are not far away: for instance, the “High-Luminosity LHC” experiments at CERN are only a few years ahead.
- In general, we see **two competing trends**:
  - “**Moore's law** is a term used to refer to the observation made by Gordon Moore in 1965 that the number of transistors in a dense integrated circuit (IC) doubles about every two years” – there are similar “laws” for storage, networks etc → if you wait long enough, every computing need will become economically affordable.
  - **Experiment + physics complexity**: every generation of experiments will collect more data, core complex events, more detailed snapshots.
- Who wins? Usually, we refer to “flat budget” as the situation where the needs increase with the same “slope” of technology → this leads to a theoretical constant amount of money per year.
  - This is sort of accepted by the Funding Agencies …

A back-of-the-envelope estimate of storage needs

- We can use a simplified model for “a detector”:
  - It “takes a picture” of a collision event every 25 ns (@ 40 MHz)
  - It has ~ O(100) Million acquisition channels (10x for the detectors to come)
  - Assuming 1 channel = 1 byte, the raw data rate would be →
- **40e6 ev/s * 100e6 byte/ev = 4 PB/s**
  - 4 PB/s in 5 years would amount to 120 ZB (ZettaBytes)
- A “storage problem” is automatic, given the need to investigate rare events with a high precision.

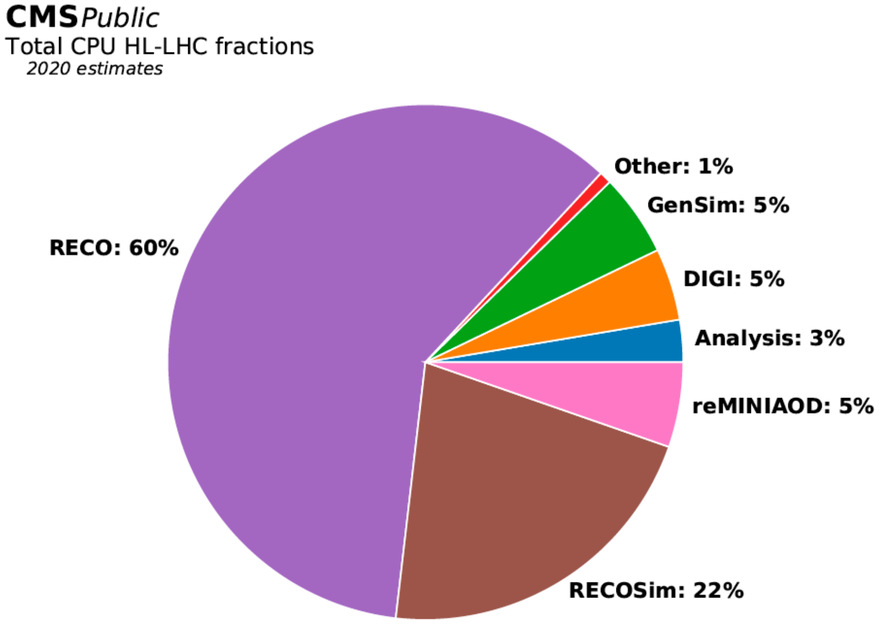
How about CPUs?

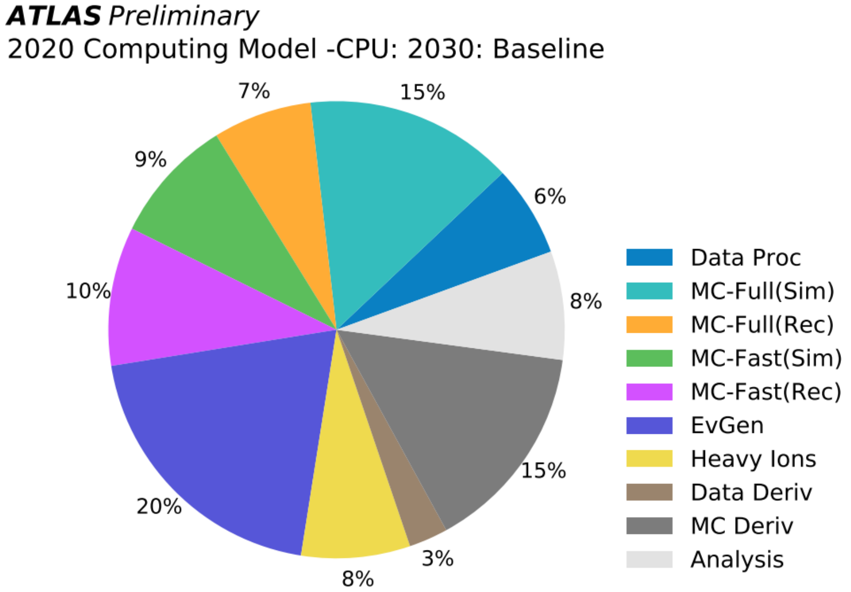
- Besides storage needs, **it turns out that CPU power is also a problem.**
- Where do we spend CPU time in current HEP experiments?
- Broad brush list:
  - Interpretation of RAW detector signals into physics objects ("**Reconstruction**")
  - Statistical studies of the reconstructed events ("**Analysis**")
  - Simulation of the physics processes ("**Generators**"), the detector response ("**Simulation**"), the electronics ("**Digitization**)"

Where do we spend CPU time?

- Different experiments have different shares in the CPU utilization, but in general **simulation** (from partons to electronic signals) and **reconstruction** (from electronic signals to “physics objects” like jets, leptons, etc.) are the most time-consuming activities.

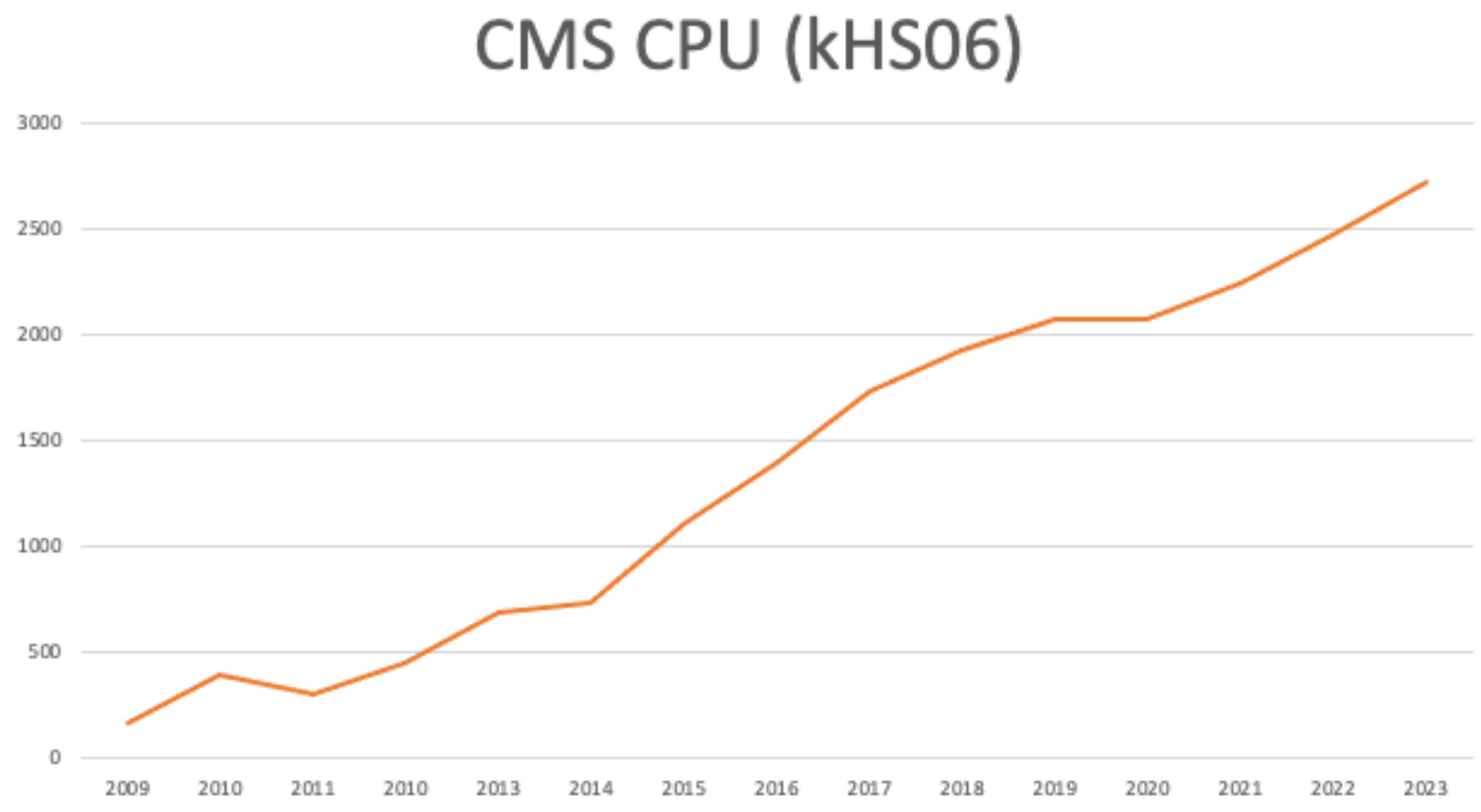
- As a rule of thumb, # of simulated events > # of collected events





How did computing for LHC (experiments at CERN) evolve?

- Pre-LHC to Run 1 (CMS CPU) – 2x
- Another 10x was 2005-2009
- All with
  - The same sites
  - The same infrastructure



Storage and CPU drive the trigger rate

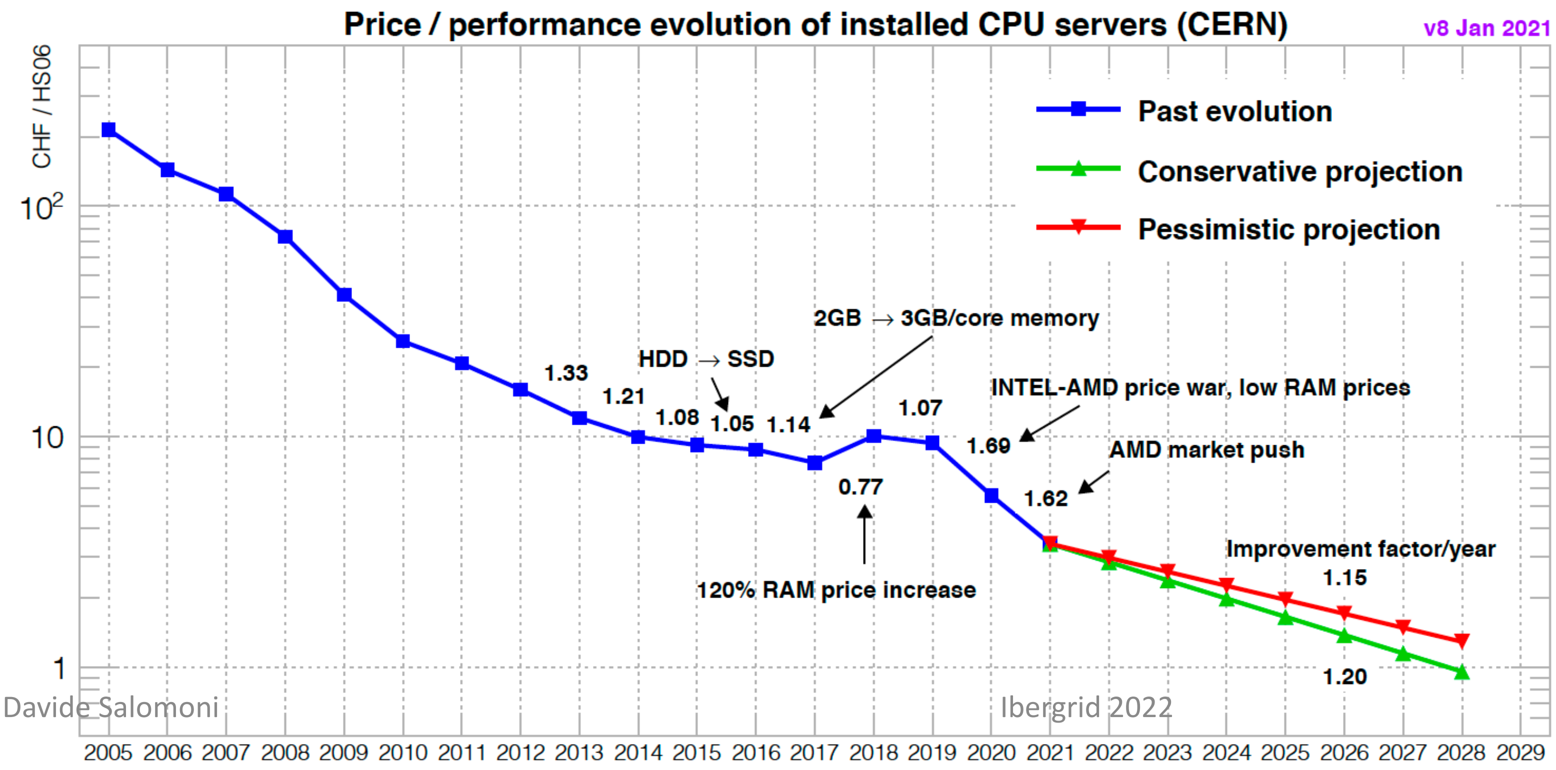
- In an ideal world, all the 40 MHz 25 ns snapshots (i.e., events) would be saved and analyzed, leading to the 120 ZB in 5 years computed above.
- In practice, a much lower rate can be saved for $$ reasons; years of studies have defined the “minimum” possible rate that still preserves physics capabilities, at least for the most important physics channels.
- In the end, it is a tension between what you can afford and what you would like to collect; LHC history (CMS-ATLAS) shows that:
  - Run-1 (2010-2012) : 100-500 Hz (out of the 40 MHz)
  - Run-2 (2015-2018) : ~ 1 kHz
  - Run-3 (2022-2024) : 1-2 kHz
  - Run-4 (2027+, see later) : > 5 kHz

What is then the limiting factor?

- Apart from some limits on the electronics *("I cannot dispatch more than X consecutive triggers")*, **the real limit** on the numbers and type of events collected by HEP and other experiments is **Computing**, and on its turn the **amount of money** one can dedicate to that.
- This can also be seen as a reversed process: I know what I can spend on computing → I know how many events I can collect → I know what type of physics I can do.
- **Therefore, any R&D, new idea or solution which allows to reduce the costs of Computing is very visible and increases the physics potential of the experiments. <u>Ultimately, this is the reason why INFN is very interested in computing technologies and in related innovative solutions</u>.**

How is technology going?

- In the glory days of 2000-2010, we were able to get year over year a +40% of performance for the same price ($\sqrt{2}$ ~ 1.41) on the type of technology we were interested in (linux boxes with x86 CPUs)
  - This made the CPU increase shown on a previous slide “harmless”: $(\sqrt{2})^{10}$ (2005-2015) = 32x
  - What about more recent times?



2005-2010: -60%/y
2010-2015: -20%/y
2015-2020: -10%y
2020-2022: some recovery seen, extrapolations are currently at -15/-20%/y

Extrapolations for High-Luminosity LHC

- A 20% yearly increase in performance from 2021 to ~2028 gives us a speed-up factor from technology advancements of $(1.2)^7$ ~**3.6x**.
- Back-of-the-envelope linear estimates foresee increased needs for computing at High-Luminosity LHC of ~**75x** (box on the right).
- Therefore, **we miss a factor 20x**!
- How can we cope with this?
  - Get 20x more money from the Funding Agencies (unrealistic, the ballpark figure would be >> 1 BEur/y)
  - Find ways to reduce the needs…

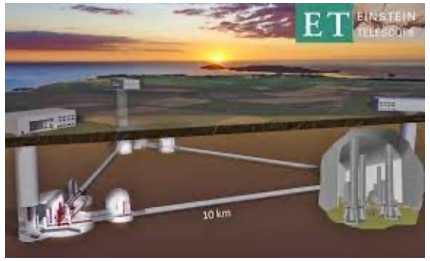
1 LHC Experiment ~2020: ~200.000 CPU Cores; ~200 PB disk; ~350 PB tape

Using the same computing model

1 HL-LHC Experiment ~2028: ~15M CPU Cores; ~15 EB disk; ~26 EB tape

What about other scientific domains?

- **DUNE**: ProtoDUNE in 2019 collected 3 GB/s (same as CMS at the same time); real DUNE expected 80x at the end of the 2020s.
- **SKA**: up to 2 PB/day (CMS ~3) , 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.
- **ET** aiming at being ~10% of a HL-LHC experiment.

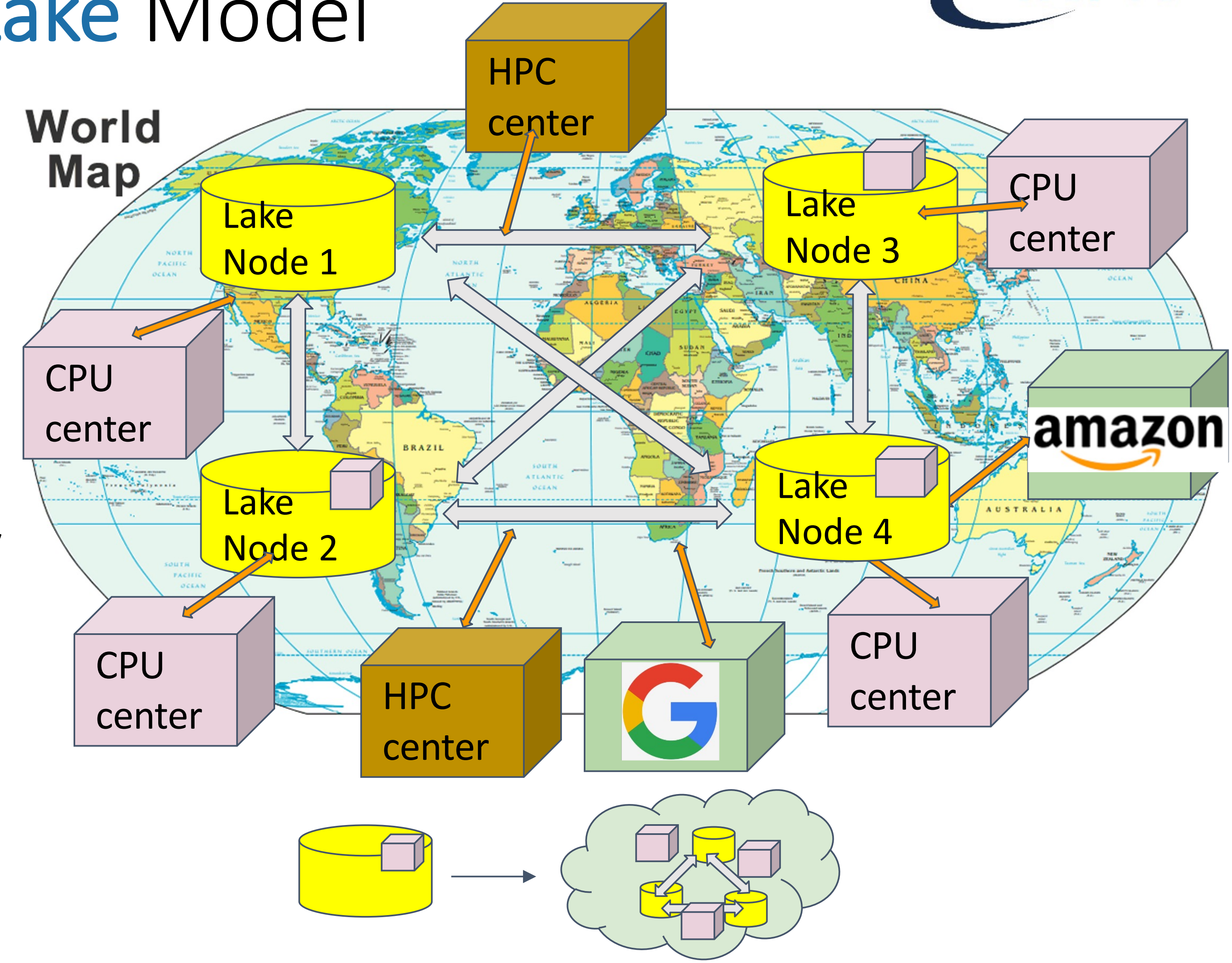


How can we cope with these levels of requests?

- We need some strategies to reduce *at least* the increase from 20x to ~1x. We can for instance consider:
  1. **Infrastructural changes**
  2. **Technological changes**
  3. Physics / Science #1: change the analysis model.
  4. Physics / Science #2: reduce the scientific reach (for example increasing trigger thresholds).
     - Not even considered here … it is the “desperation move” if we fail with everything else.
  5. Something unexpected...
- Let’s discuss a bit about 1 and 2.

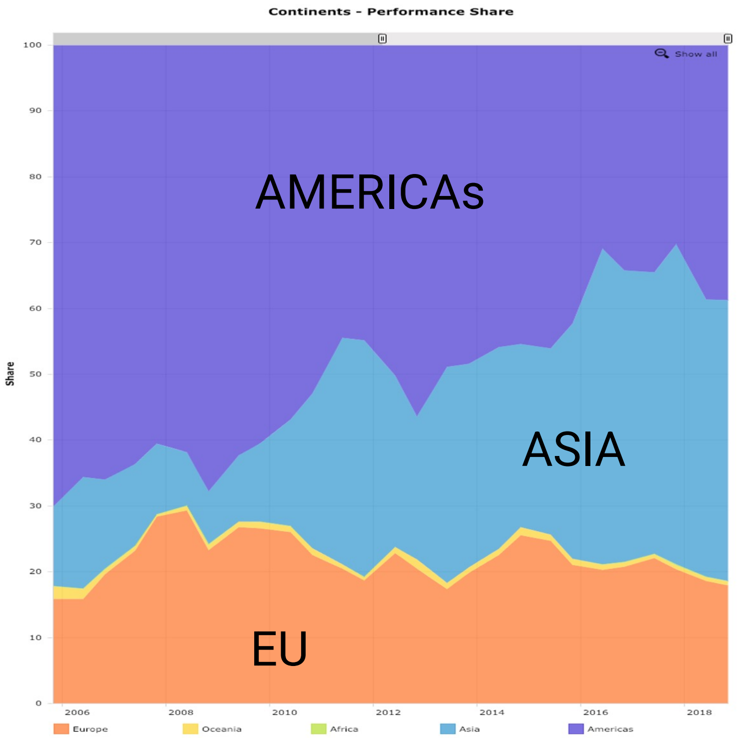
The Cloud / Data Lake Model

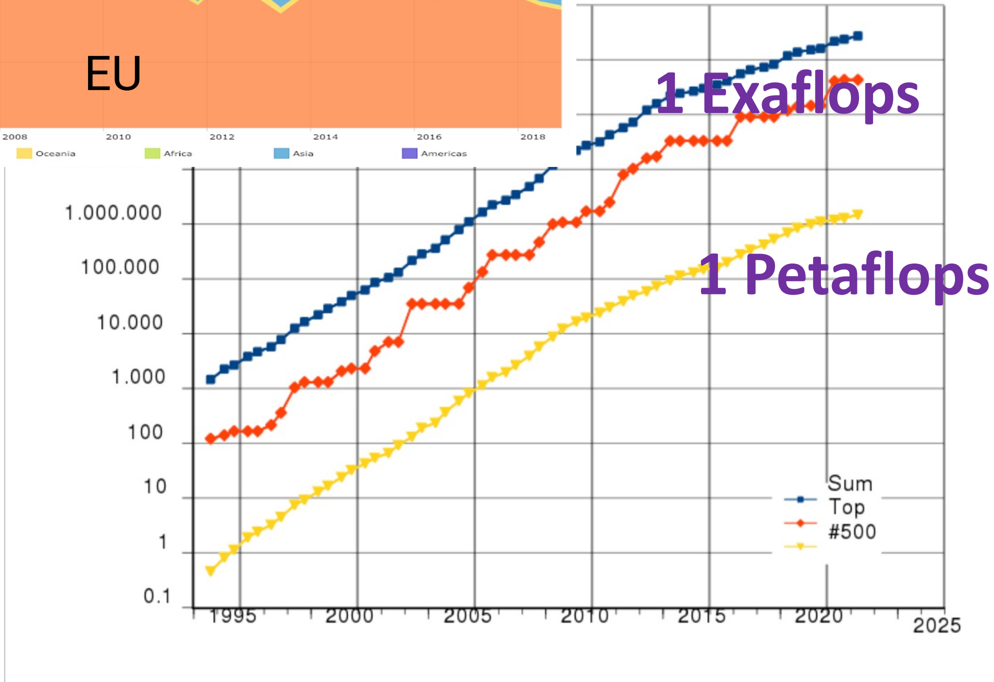
- Keep the real value from the experiments safe
  - **(RAW) data and a solid baseline of CPU in <u>owned and stable</u> sites**
  - Allow multiple CPU resources to join, even temporarily
    - Eventually choosing the cheapest at any moment
  - **Solid networking**: use caches / streaming to access data
- Reduce requirements for computing resources
  - Commercial Clouds
  - Other sciences' resources
    - SKA, CTA, Dune, Genomics, ...
  - Use HPC systems *if possible*



Supercomputing (HPC)

- The world is full of Supercomputers. Why ?
  - Real scientific use cases
    - Lattice QCD, Climate modeling, ...
  - Industrial showcase ("Country XY is technologically capable")
    - And hence not 100% utilized: opportunities for smart users. *Can we be one of them?*
- Many not trivial problems to solve:
  - **Data access** (access, bandwidth, ...)
  - **Accelerator Technology** (KNL, GPU, FPGA, TPU, ???, ...)
  - **Submission of tasks** (MPI vs Batch systems vs proprietary systems)
  - **Node configurations** (low RAM/Disk, ...)
  - **Not-too-open environments** (OS, ...)
- Some hints of global slowing down, but not for top systems where the "HPC war" is on
- 1 Petaflops = 10^{15} floating point operations per second
- 1 Exaflops = 10^{18} floating point operations per second





HPC integration @ INFN

- Experimentation with HPC systems (CINECA, mostly) started in 2019 for LHC-like workflows, with excellent results using 4 PRACE grants
  - Demonstration of the capability to execute LHC workflows on HPC systems
  - Demonstration of the capability to execute workflows on non-standard architectures (Power9)
  - Demonstration of the capability to circumvent HPC security via user-level tools ☺

Enabling CMS Experiment to the utilization of multiple hardware architectures -- a Power9 Testbed at CINECA

T.Boccali[1], A.Malta Rodrigues[2], D.Spiga[3], M.Mascheroni[4] for the CMS Collaboration

[1] INFN Sezione di Pisa, [2] University of Nebraska-Lincoln, [3] INFN Sezione di Perugia, [4] University of California San Diego

Extension of the INFN Tier-1 on a HPC system

Tommaso Boccali[1], *Stefano* Dal Pra[2], *Daniele* Spiga[3], *Diego* Ciangottini[3], *Stefano* Zani[2], *Concezio* Bozzi[4], *Alessandro* De Salvo[5], *Andrea* Valassi[6], *Francesco* Noferini[7], *Luca* dell'Agnello[2], *Federico* Stagni[6], *Alessandra* Doria[8], *Daniele* Bonacorsi[9]

First experiences with a portable analysis infrastructure for LHC at INFN

Diego Ciangottini[1], Tommaso Boccali[2], Andrea Ceccanti[3], Daniele Spiga[1], Davide Salomoni[3], Tommaso Tedeschi[1], and Mirco Tracolli[1]

Integration of CINECA-PRACE into CMS Computing

Diego Ciangottini
Daniele Spiga
Mirko Mariotti
Tommaso Boccali
(for the INFN team - see last page)

First tests of Job Distribution on PowerPC

Daniele Spiga
Marco Mascheroni
Tommaso Boccali

However, HPC => technological changes

- As we have seen, the unitary cost of our computing (the CPUs) *does not decrease fast enough*; this is true for OUR computing (Linux PCs with Intel x86 architecture), but not necessarily for ALL computing.
- GPGPUs and FPGAs and even more ASICs are better at metrics such as Operations/$$, while ARM and Power8/9 architectures are better at Operations/Joule.
  - … but in most cases using fruitfully at least the first category needs a complete rewriting of the code – **which costs a lot!**

Table 6. SLOCCount measured lines of source code for ATLAS and CMS.

Experiment Type	Source Lines of code (SLOC)	Development effort (person-years)	Total estimated cost to develop
ATLAS	5.5M	1630	220 M$
CMS	4.8M	1490	200 M$

As a reference:
- the Linux Kernel is: 15M SLOC, 4800 FTEy, 650M$ (3x CMS)
- Geant4 is: 1.2M SLOC, 330 FTEy, 45 M$ (1/4x CMS)

And that's only the "core code"…

- We rely on many externals (Geant4 is an external, ROOT is an external, Pythia is an external), which greatly inflate the total code size.
- This (in unreadable fonts) is the list of externals for a typical CMS release:

alpgen qd **root**_cxxdefaults sockets catch2 **gcc-ccompiler gcc-cxxcompiler gcc-f77compiler** mpfr cmsswdata codechecker csctrackfinderemulati cuda-stubs cuda-gcc-support cvs2git dablooms db6 dmtcp doxygen eigen fastjet-contrib fastjet-contrib-archi gcc-analyzer-ccompile gcc-analyzer-cxxcompi gcc-atomic gcc-checker-plugin gcc-plugin gdb geant4-parfullcms geant4data py2-numpy openloops git glibc glimpse gmake gnuplot gosam gosamcontrib hdf5 igprof intel-license ittnotify lapack lcov libffi libxslt llvm md5 openblas ofast-flag openmpi **professor** py2-sympy py2-absl-py py2-appdirs py2-argparse py2-asn1crypto py2-atomicwrites py2-attrs py2-autopep8 py2-avro py2-awkward py2-backcall py2-backports py2-backports-functooccj py2-backports_abc py2-beautifulsoup4 py2-bleach py2-bokeh py2-bottleneck py2-cachetools py2-certifi py2-cffi py2-chardet py2-click py2-climate py2-colorama py2-contextlib2 py2-cryptography py2-cx-oracle py2-cycler py2-cython py2-dablooms py2-decorator py2-defusedxml py2-docopt py2-downhill py2-dxr py2-entrypoints py2-enum34 py2-flake8 py2-flawfinder py2-fs py2-funcsigs py2-functools32 py2-future py2-futures py2-gast py2-gitdb2 py2-gitpython py2-google-common py2-googlepackages py2-grpcio py2-h5py py2-h5py-cache py2-hep_ml py2-histbook py2-histogrammar py2-html5lib py2-hyperas py2-hyperopt py2-idna py2-ipaddress py2-ipykernel py2-ipython py2-ipython_genutils py2-ipywidgets py2-jedi py2-jinja2 py2-jsonpickle py2-jsonschema py2-jupyter py2-jupyter_client py2-jupyter_console py2-jupyter_core py2-keras py2-keras-application py2-keras-preprocessi py2-kiwisolver py2-lint py2-lizard py2-llvmlite py2-lxml py2-lz4 py2-markdown py2-markupsafe py2-matplotlib py2-mccabe py2-mistune py2-mock py2-more-itertools py2-mpld3 py2-mpmath py2-nbconvert py2-nbdime py2-nbformat py2-networkx py2-neurolab py2-nose py2-nose-parameterize py2-notebook py2-numba py2-numexpr py2-oamap py2-onnx py2-ordereddict py2-packaging py2-pandas py2-pandocfilters py2-parsimonious py2-parso py2-pathlib2 py2-pbr py2-pexpect py2-pickleshare py2-pillow py2-pip py2-pkgconfig py2-plac py2-pluggy py2-ply py2-prettytable py2-prometheus_client py2-prompt_toolkit py2-protobuf py2-prwlock py2-psutil py2-ptyprocess py2-py py2-pyasn1 py2-pyasn1-modules py2-pybind11 py2-pybrain py2-pycodestyle py2-pycparser py2-pycurl py2-pydot py2-pyflakes py2-pygithub py2-pygments py2-pymongo py2-pyopenssl py2-pyparsing py2-pysqlite py2-pytest py2-python-cjson py2-python-dateutil py2-python-ldap py2-pytz py2-pyyaml py2-pyzmq py2-qtconsole py2-rep py2-repoze-lru py2-requests py2-root_numpy py2-root_pandas py2-rootpy py2-scandir py2-schema py2-scikit-learn py2-scipy py2-seaborn py2-send2trash py2-setuptools py2-simplegeneric py2-singledispatch py2-six py2-smmap2 py2-soupsieve py2-sqlalchemy py2-stevedore py2-subprocess32 py2-tables py2-tensorflow py2-terminado py2-testpath py2-theanets py2-theano py2-thriftpy py2-tornado py2-tqdm py2-traitlets py2-typing py2-typing_extensions py2-uncertainties py2-uproot py2-uproot-methods py2-urllib3 py2-virtualenv py2-virtualenv-clone py2-wcwidth py2-webencodings py2-werkzeug py2-wheel py2-widgetsnbextensio py2-xgboost py2-xrootdpyfs pydata pyminuit2 pyqt python-paths python_tools rootglew scons sloccount tcmalloc tcmalloc_minimal tensopy2-virtualenvwrapperrflow tinyxml2 xtl blackhat boost boost_header python bz2lib cascade_headers ccache-ccompiler ccache-cxxcompiler ccache-f77compiler zlib gmp photos_headers pythia6_headers openssl clhep clhepheader cppunit cuda curl libxml2 dcap root_interface xz xerces-c vecgeom_interface hepmc_headers distcc-ccompiler distcc-cxxcompiler distcc-f77compiler dpm expat fastjet fftjet fftw3 freetype gbl gdbm gsl giflib google-benchmark libjpeg-turbo hector heppdt **madgraph5amcatnlo** llvm-cxxcompiler jemalloc jimmy_headers ktjet libhepml libuuid **llvm-ccompiler** llvm-f77compiler meschach mxnet-predict numpy-c-api x11 oracle pacparser yoda protobuf **python3** qd_f_main sqlite sigcpp tauola_headers **tbb tensorflow-framework** tensorflow-runtime tensorflow-xla_compil0-pafccj3 toprex_headers utm valgrind vdt_headers xrootd xtensor boost_system boost_iostreams boost_serialization boost_program_options boost_python boost_regex boost_signals boost_test cascade yaml-cpp photos pythia6 pcre cub **cuda-api-wrappers** cuda-cublas cuda-cufft cuda-curand cuda-cusolver cuda-cusparse cuda-npp cuda-nvgraph cuda-nvjpeg cuda-nvml cuda-nvrtc das_client vecgeom hepmc frontier_client google-benchmark-main libpng iwyu-cxxcompiler libtiff libungif llvm-analyzer-ccompil llvm-analyzer-cxxcomp mcdb opengl openldap oracleocci pyclang qtbase sip starlight tauola tensorflow-c tensorflow-cc tkonlinesw toprex vdt boost_chrono boost_filesystem boost_mpi cgal lhapdf classlib davix rootcling **geant4core** photospp geant4static graphviz lwtnn millepede qt3support rivet tkonlineswdb cgalimageio herwig rootmathcore rootrio **pythia8** geant4vis thepeg pyquen qt rootrint rootrflx rootsmatrix rootx11 sherpa charybdis rootthread dire tauolapp **geant4** geneva herwigpp jimmy qtdesigner rootgeom rootxmlio vincia rootcore evtgen roothistmatrix rootmath rootxml rootphysics rootgpad rootfoam rootspectrum root rootminuit rootgraphics rootgui rootinteractive roothtml rootminuit2 dd4hep-core roofitcore mctester professor2 rooteg rootgeompainter rootrgl rootged rootguihtml rootmlp rootpy **dd4hep** dd4hep-geant4 roofit rooteve roottmva roostats rootpymva histfactory coral

- Note that **gcc** is there! CMS ships its own compiler, so dependency on the Linux host is only at the level of glibc.

So: what are the main drivers for the next 10 years of scientific computing for INFN?

1. Infrastructure

- **Renew infrastructures** to be ready for the High Luminosity-LHC (HL-LHC) era, up to ~2035 or more
- **Use more compact computing** (from today's ~20 kW/rack to 80 or more)
- **Lower the PUE** (power usage effectiveness), be greener
- **Extend and expand networking** for a future-proof infrastructure

2. Hardware, Software, Services

- **Foster and simplify the utilization of more viable technologies** (€/task or J/task), like GPUs, FPGA, down to Quantum when available
- **Be more efficient, elastic and resilient**
  - Pervasive use of **geographically distributed storage** ("the Datalake")
  - Abstract from physical machines, and form a **national pool of resources and high-level services** ("the Cloud")
  - **Extend elastically to external providers** such as traditional HPC or to other cloud providers ("dynamic federations")

The evolution of Infrastructure and of Higher-level Services must proceed together

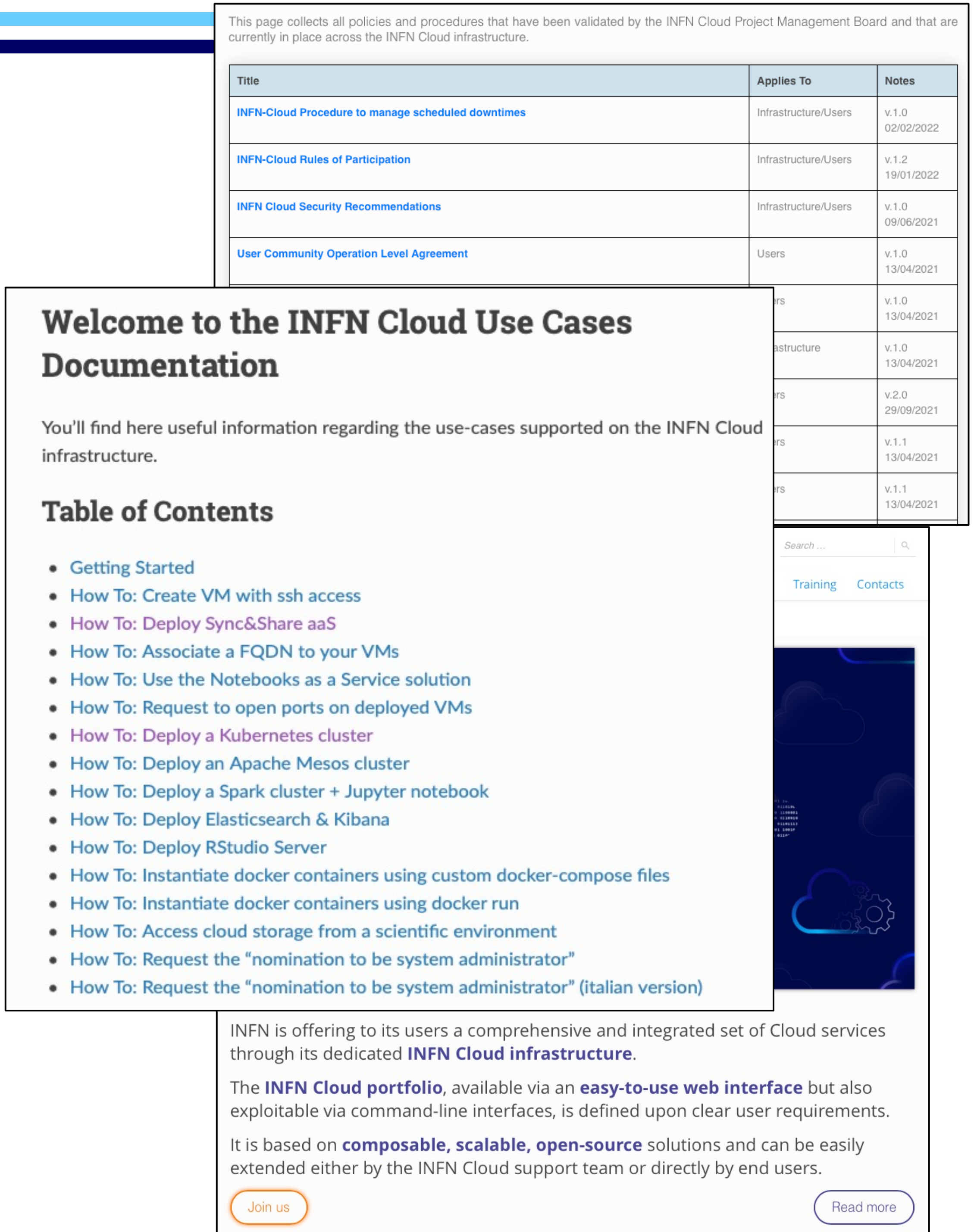
A founding milestone for us: INFN Cloud

INFN Cloud, https://www.cloud.infn.it/

- In **production** since March 2021.
- The **initial seed** of a National Datalake for research and beyond, building on (existing|renewed|new) e-Infrastructures.
- The **base of the evolution** of the INFN Distributed Computing vision.
- Built on a **thin middleware layer** running on top of *federated clouds*, decoupling physical and logical views via a **service composition** mechanism.
- The **INFN foundation** for the NRRP computing-related initiatives (more on this later).

What is INFN Cloud

- A production-quality set of resources and solutions providing:
  - A **core backbone**, with ancillary and special-purpose services.
  - A **multi-site, federated Cloud infrastructure**.
    - INFN Cloud can transparently federate INFN sites as well as public or private Clouds (e.g.: AWS, Google Compute Cloud, Microsoft Azure, and others)
  - A **customizable portfolio of services** accessible via web interfaces, terminal or API.
  - A **fully distributed organization for the support and management** of both infrastructure and services.
  - A set of **rules that define access resources and policies**, according to INFN, national and European laws.



This page collects all policies and procedures that have been validated by the INFN Cloud Project Management Board and that are currently in place across the INFN Cloud infrastructure.

Title	Applies To	Notes
INFN-Cloud Procedure to manage scheduled downtimes	Infrastructure/Users	v.1.0 02/02/2022
INFN-Cloud Rules of Participation	Infrastructure/Users	v.1.2 19/01/2022
INFN Cloud Security Recommendations	Infrastructure/Users	v.1.0 09/06/2021
User Community Operation Level Agreement	Users	v.1.0 13/04/2021
	rs	v.1.0 13/04/2021
	astructure	v.1.0 13/04/2021
	rs	v.2.0 29/09/2021
	rs	v.1.1 13/04/2021
	rs	v.1.1 13/04/2021

Welcome to the INFN Cloud Use Cases Documentation

You'll find here useful information regarding the use-cases supported on the INFN Cloud infrastructure.

Table of Contents

Search ...

Training Contacts

INFN is offering to its users a comprehensive and integrated set of Cloud services through its dedicated **INFN Cloud infrastructure**.

The **INFN Cloud portfolio**, available via an **easy-to-use web interface** but also exploitable via command-line interfaces, is defined upon clear user requirements.

It is based on **composable, scalable, open-source** solutions and can be easily extended either by the INFN Cloud support team or directly by end users.

Join us Read more

Excerpt from the INFN Cloud portfolio

Compute Services
A list of services that enable a specific cloud technology

Analytics
A collection of ad-hoc solutions for analytic purpose

Machine Learning
List of ready-to-use Machine Learning services

Data Services
Data management and storage services

Scientific Community Customizations
Customized environments

Virtual Machine
Launch a compute node getting the IP and SSH credentials to access via ssh

Docker-compose
Run a docker compose file fetched from the specified URL

Apache Mesos cluster
Apache Mesos abstracts CPU, memory, storage, and other compute resources away from machines (physical or virtual)

Kubernetes cluster
Deploy a single master Kubernetes 1.17.0 cluster

Elasticsearch and Kibana
Deploy a virtual machine pre-configured with the Elasticsearch search and analytics engine and with Kibana for simple visualization

Spark + Jupyter cluster
Deploy a complete Spark 3.0.1 + Jupyter Notebook on top of a Kubernetes (K8s) computing cluster

Jupyter with persistence for Notebooks
Run Jupyter on a single VM enabling Notebooks persistence

RStudio
RStudio is an integrated development environment (IDE) for R.

Jupyter with persistence for Notebooks
Run Jupyter on a single VM enabling Notebooks persistence

Working Station for Machine Learning INFN (ML_INFN)
Run a single VM with all the ML-INFN environment exposing both ssh access and Jupyter

Sync&Share aaS
The INFN Cloud Sync & Share as a Service is based on the popular ownCloud storage solution.

Object Storage
The INFN Cloud Object Storage as a Service.

Secure storage: MINIO HashiCorp

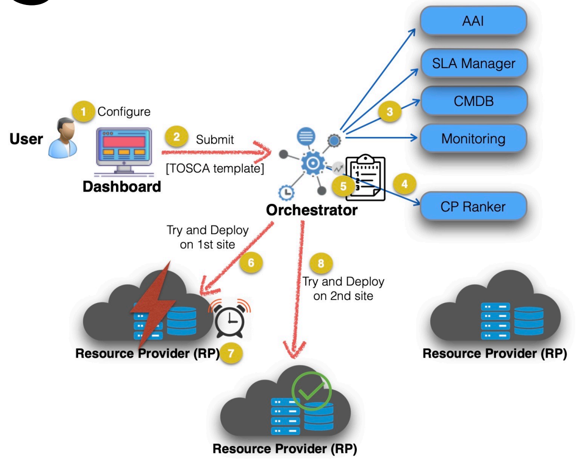
In-memory data store: redis

Secure backup: IBM Spectrum Protect

PaaS Orchestrator: Indigo IM

Selectable storage QoS levels: fast (SSD), normal (HDD), archive (tape-backed), remote replicas

Architectural pillar: the INDIGO PaaS



- The INDIGO-DataCloud PaaS is rooted on:
  1. A **distributed resource orchestration** framework
  2. A standard—based **federated solution for identity access management** (INDIGO-IAM)
- In practice:
  - Following authentication, a user requests a service via a Dashboard, APIs, or a CLI.
  - The PaaS Orchestrator is contacted, and a series of ancillary services get involved (e.g. AAI, SLA Manager, Provider Ranker, Monitoring).
  - A deployment of the required service is scheduled and eventually delivered on one of the federated resource providers.
  - All services are described through an **Infrastructure as Code** paradigm, via a combination of TOSCA templates (to model an application stack), Ansible roles (to manage the automated configuration of virtual environments), and Docker containers (to encapsulate high-level application software and runtime).

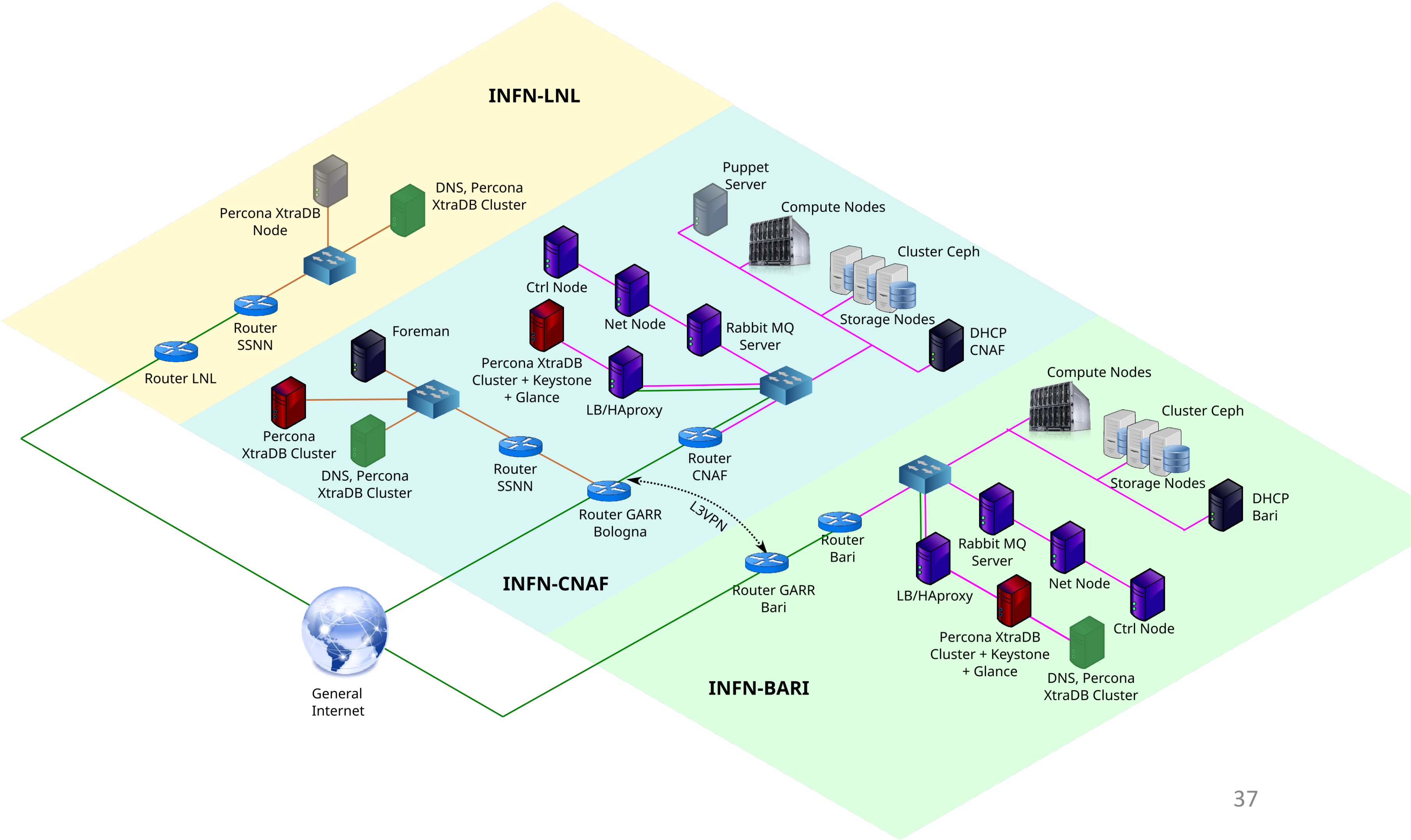
The INFN Cloud Backbone

The **INFN Cloud Backbone** is a multi-site cloud infrastructure running both **INFN Cloud Core Services** and **some user-level resources.**

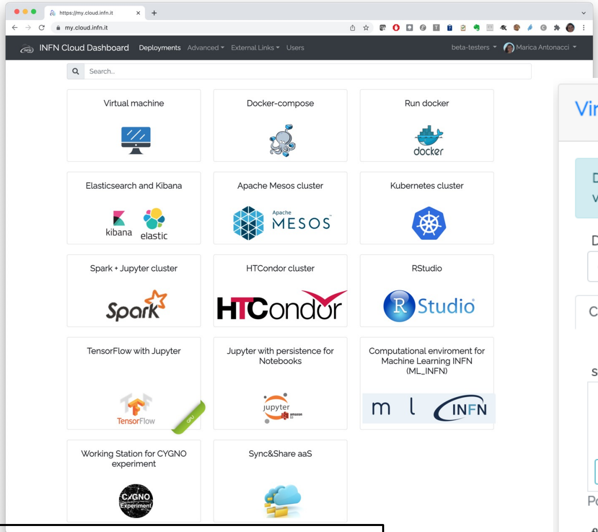
The topology of the INFN Cloud Backbone simplifies the implementation of **geographic HA or failover** for its Core Services.

The backbone infrastructure is managed by the INFN Cloud Team with **extensive use of infrastructure automation tools**.

Federated clouds connect directly to the backbone.

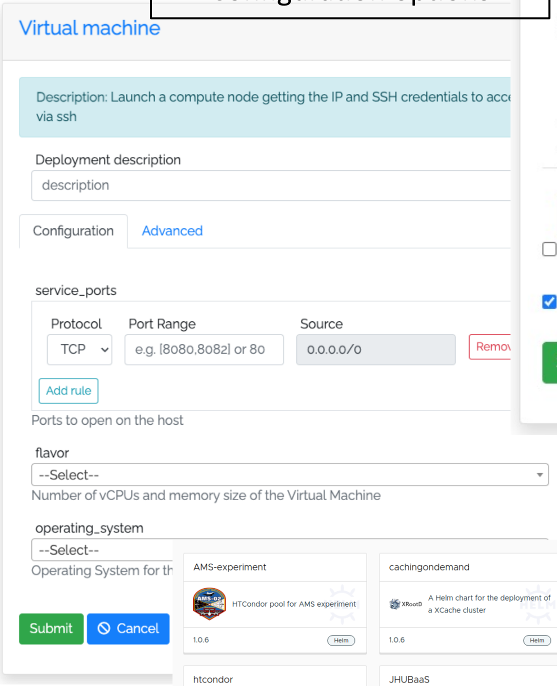


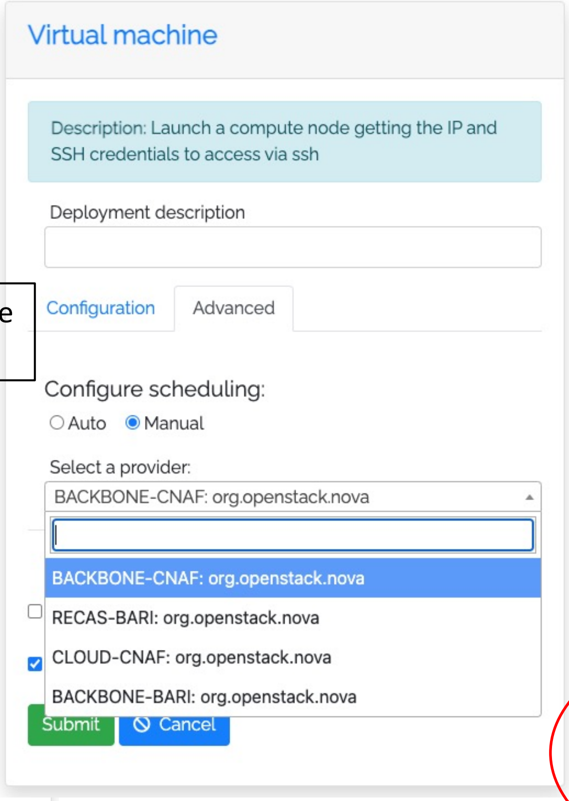
Status, Dashboards

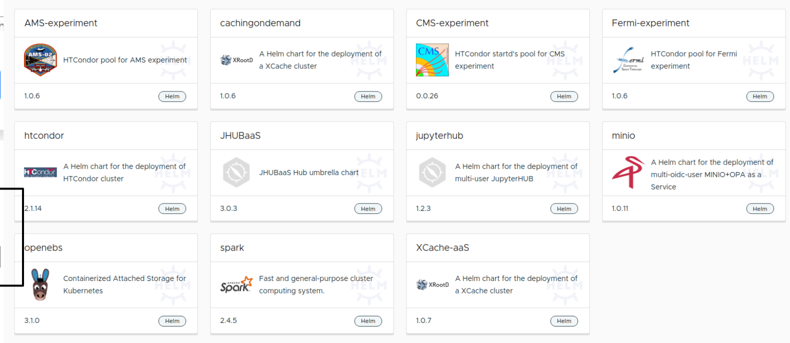


Per-user/per-group
View of the main dashboard

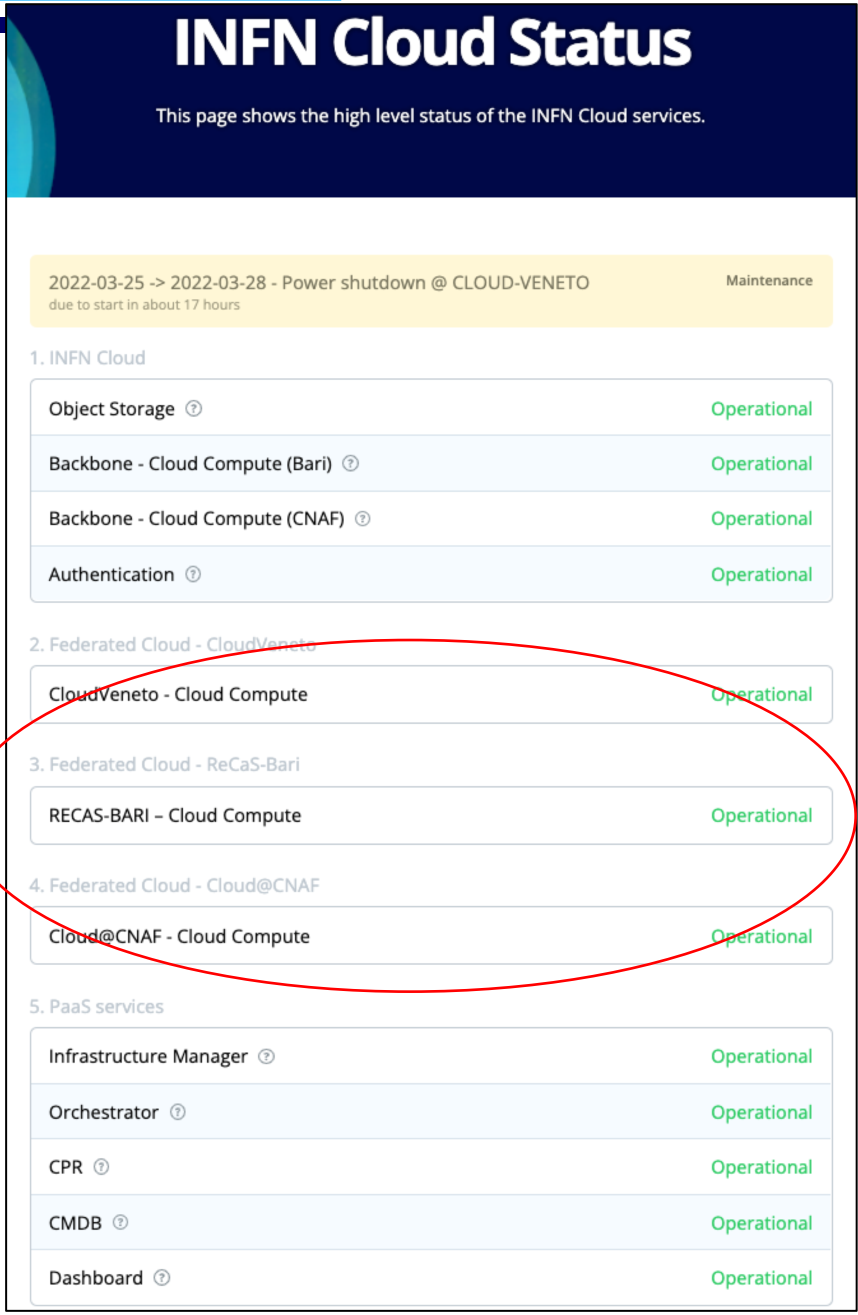
Basic/Advanced per-service
Configuration options





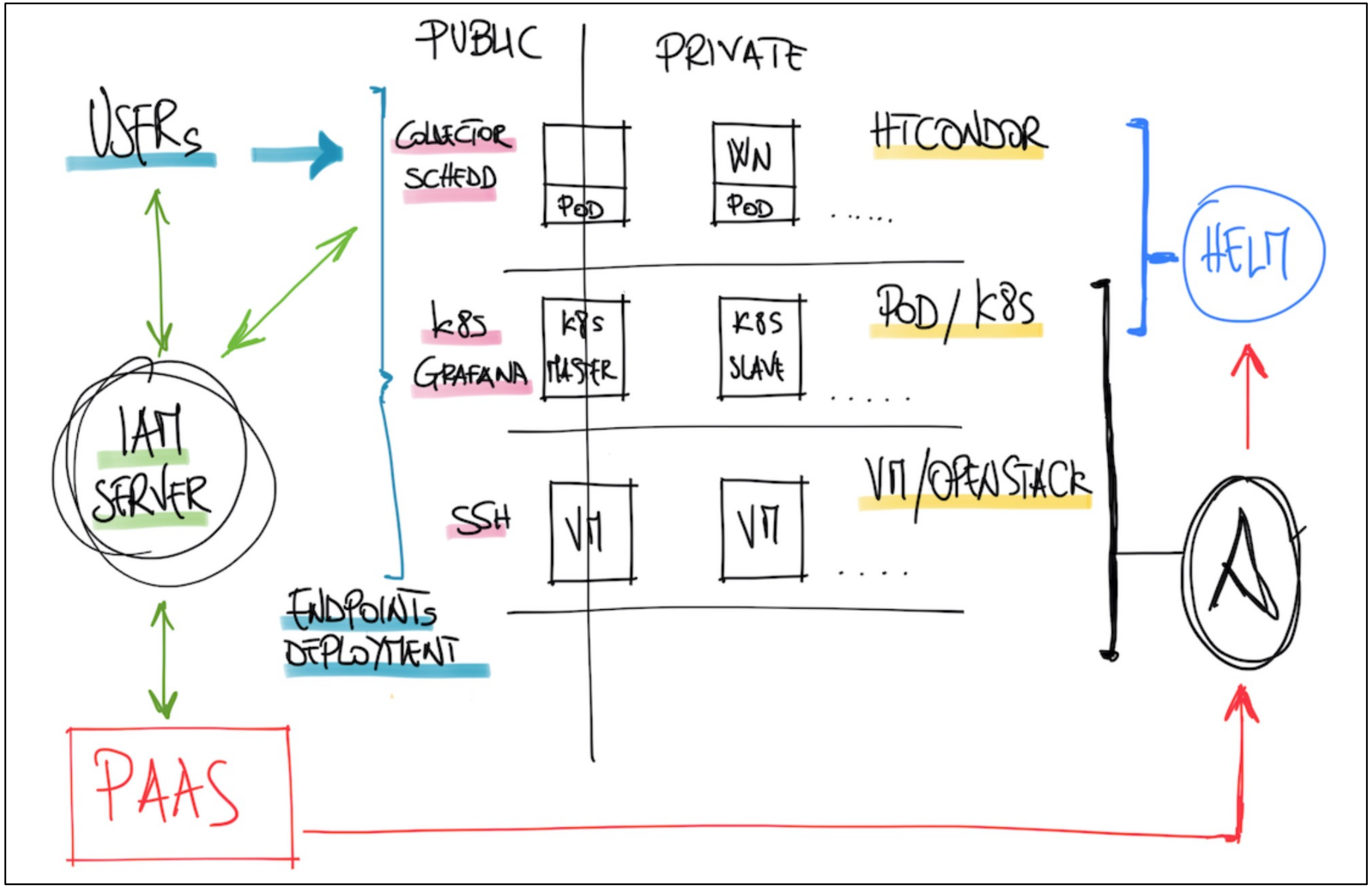


Kubeapp-level
Service dashboard



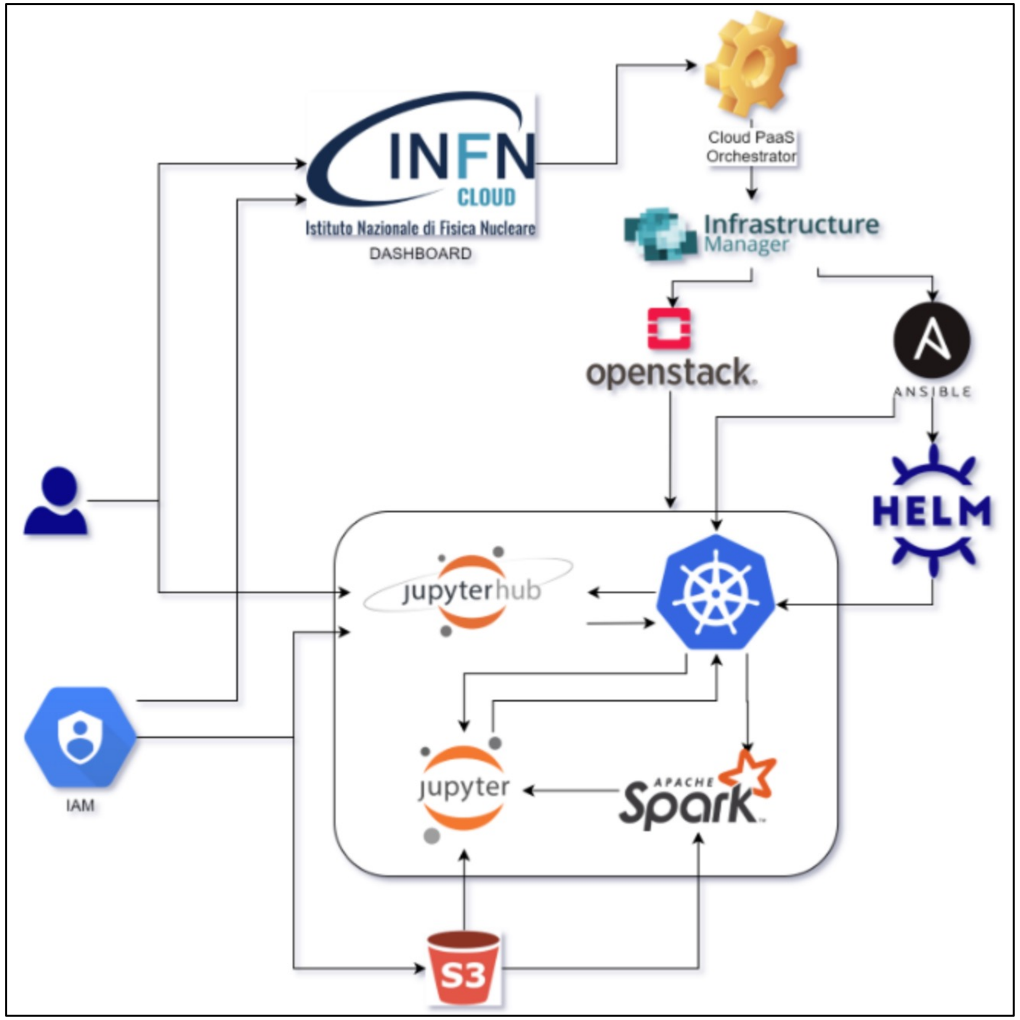
Examples of INFN Cloud services: on-demand HTCondor

- This service instantiates a k8s cluster which is then used to automatically deploy an HTCondor cluster.
- HTCondor services are deployed using dedicated PODs.



Examples of INFN Cloud services: on-demand Spark + Jupyter

- This service creates a Spark cluster on-demand, based on Kubernetes as resource manager and on JupyterHub for the user interface.
- The JupyterHub application, the Jupyter notebooks launched from it, the Spark driver and executors are all Docker containers orchestrated by a Kubernetes cluster deployed on a federated cloud (the box in the picture on the right). The components outside the box implement the workflow to instantiate the service.



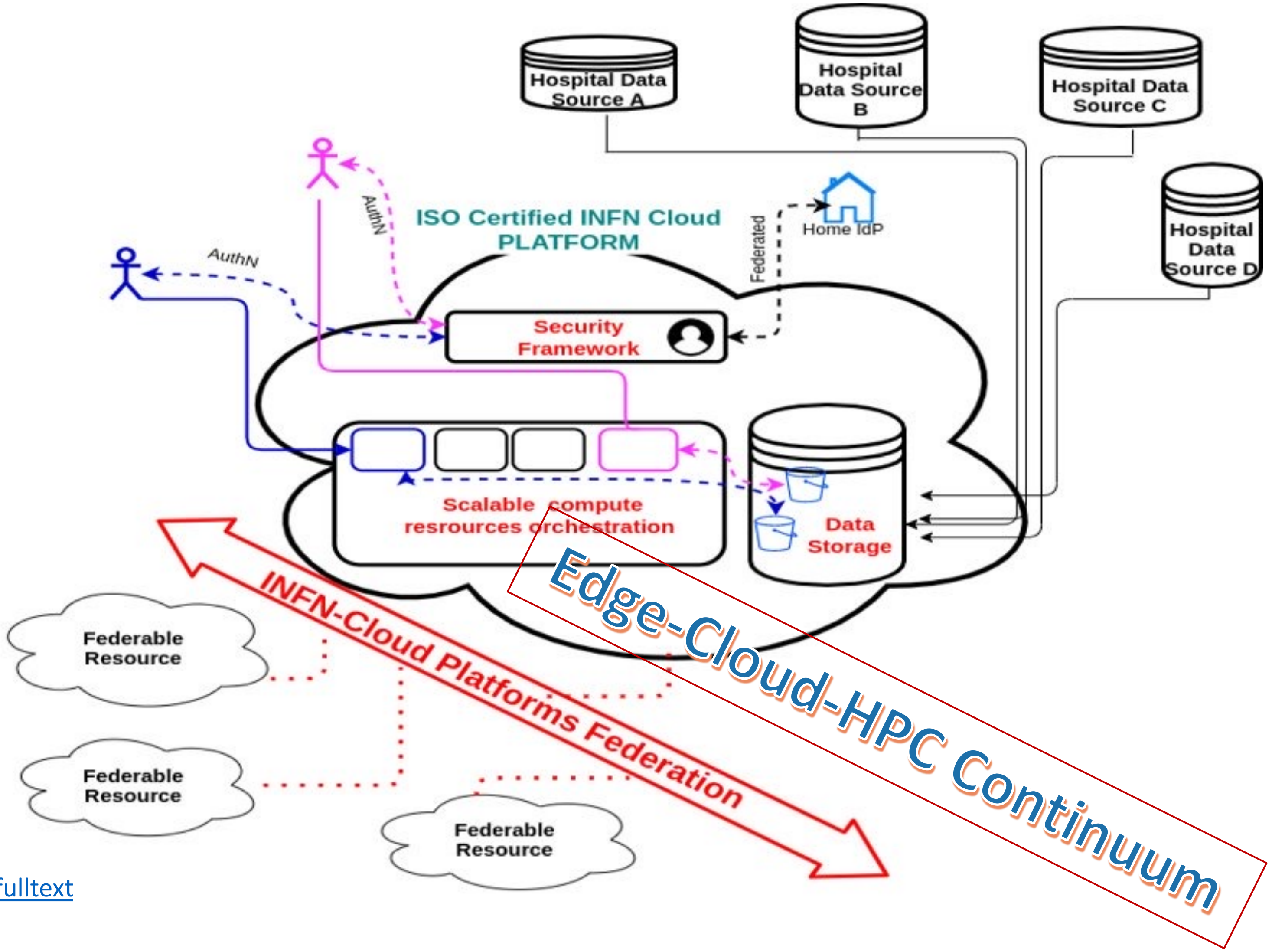
Examples of INFN Cloud services: managed services

- Several solutions are also provided by INFN Cloud as “centrally managed services”. Among them:
  - Harbor as open-source registry
  - GitLab as DevOps platform
  - NextCloud as collaboration platform
  - Jupyter Notebook as a Service
  - Vault as identity-based secrets and encryption management system
  - CVMFS as software distribution service
  - Wazuh as security monitoring platform
  - … and others.

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:

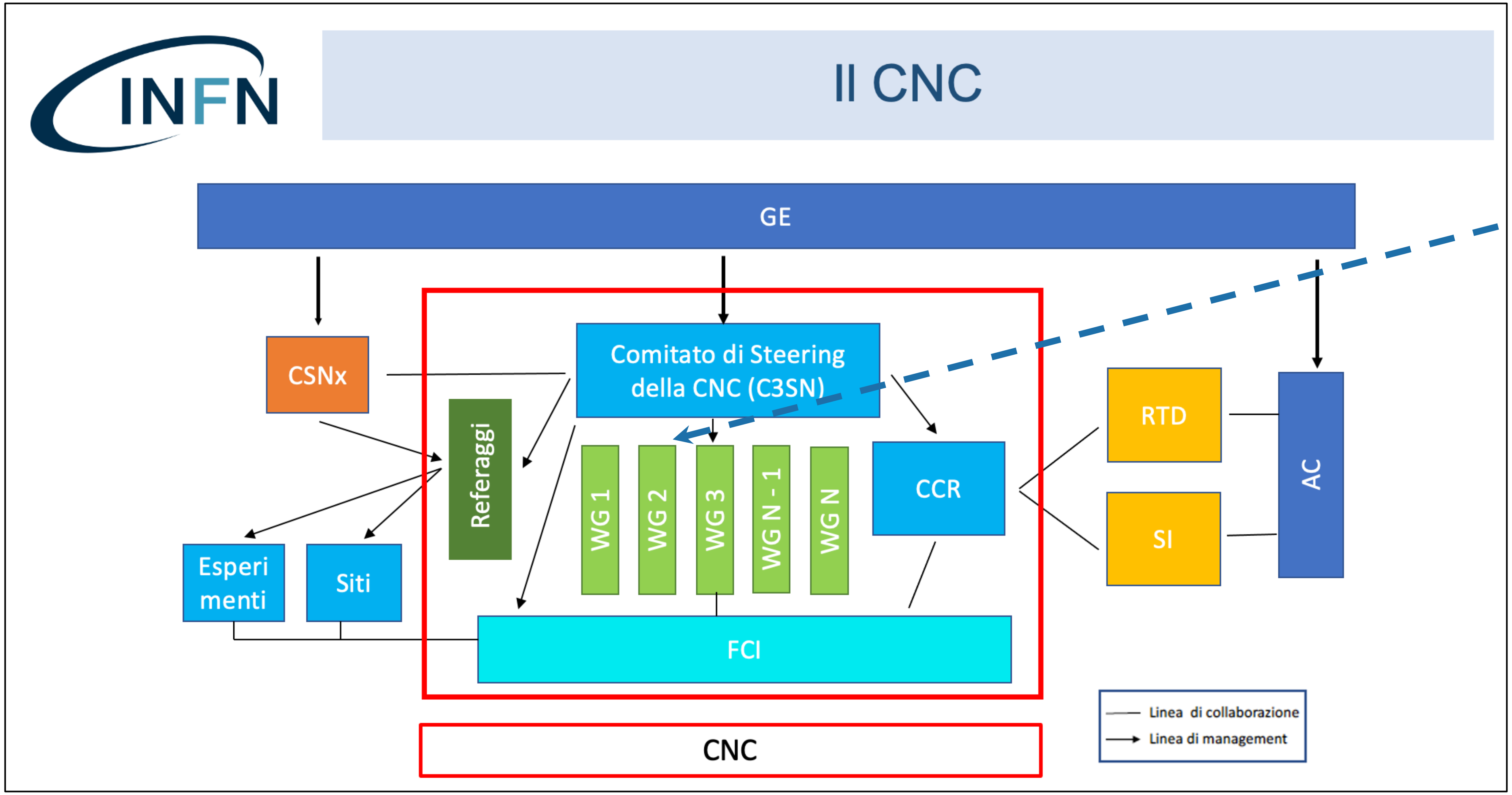
1. **Central harvesting** of data generated remotely
2. **Edge-level anonymization**, followed by central ingestion and analysis of data
3. **Edge-level feature extraction**, followed by central ingestion and analysis of features
4. **Federated learning** based on edge-level training, followed by publishing 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

Preparing for the future: the new INFN Computing Management Structure

The New INFN Computing Structure



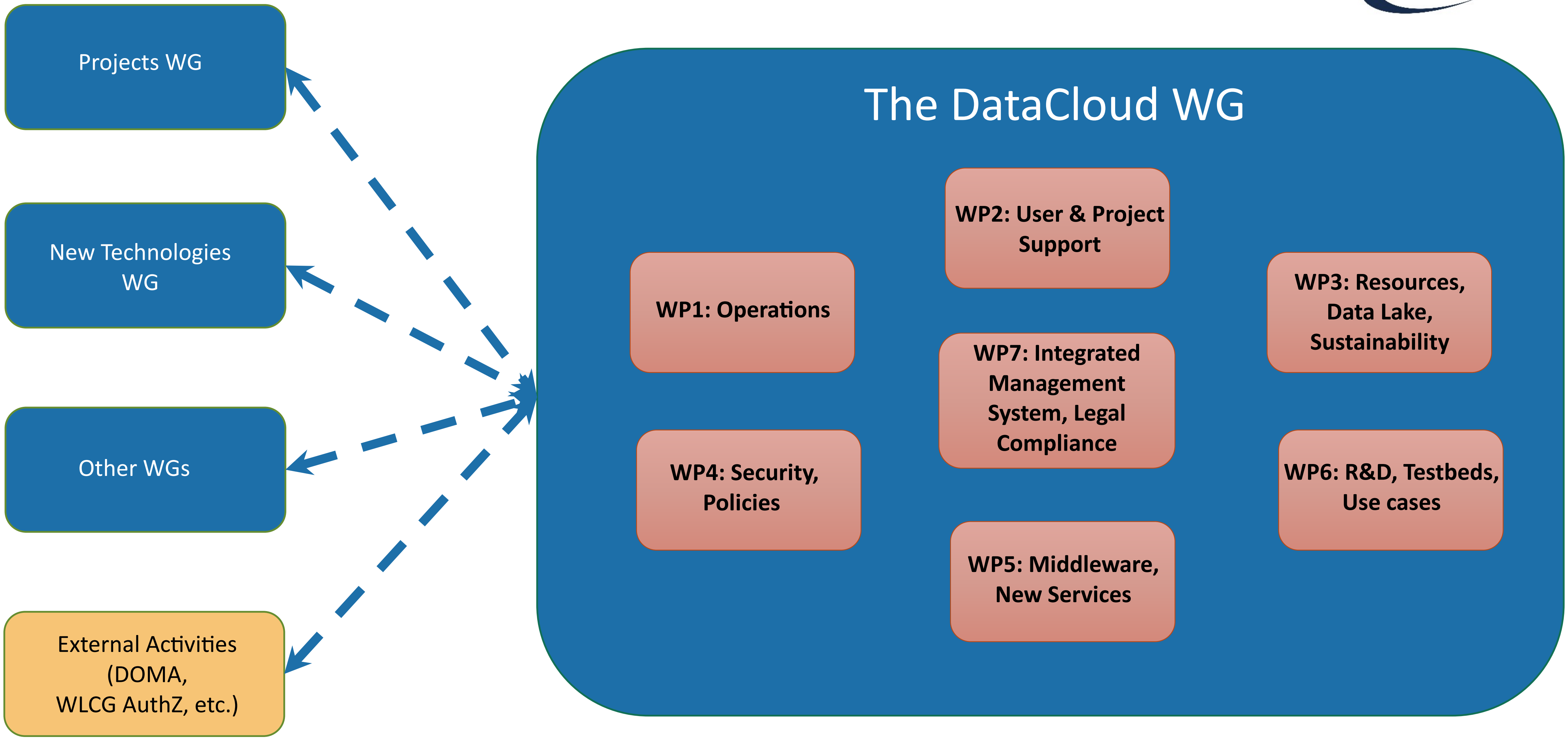
One of these WGs is the **DataCloud WG**, extending and evolving the existing INFN Cloud project.

Datalake +
INFN Cloud
⬇
DataCloud

The mandate of the DataCloud WG

- The DataCloud WG deals with **several core activities related to computing @ INFN**:
  - Development, implementation & management of the INFN Datalake architecture.
  - Development of ISO-Certified solutions. These have been seeing strong demand, thanks also to activities linked to the National Recovery and Resilience Plan (NRRP) and to various collaborations.
  - Support to users and to the management and operation of INFN Tier-x sites.
  - Development of new services.
- A key point of this WG is the **integration** between the traditional, WLCG-like Tier-x infrastructure and the “Cloud Native” model (currently represented by INFN Cloud).
  - **Integration of what?** → Of resources, methods, people, solutions.

The DataCloud WG structure



Opportunities:
The Bologna Technopole

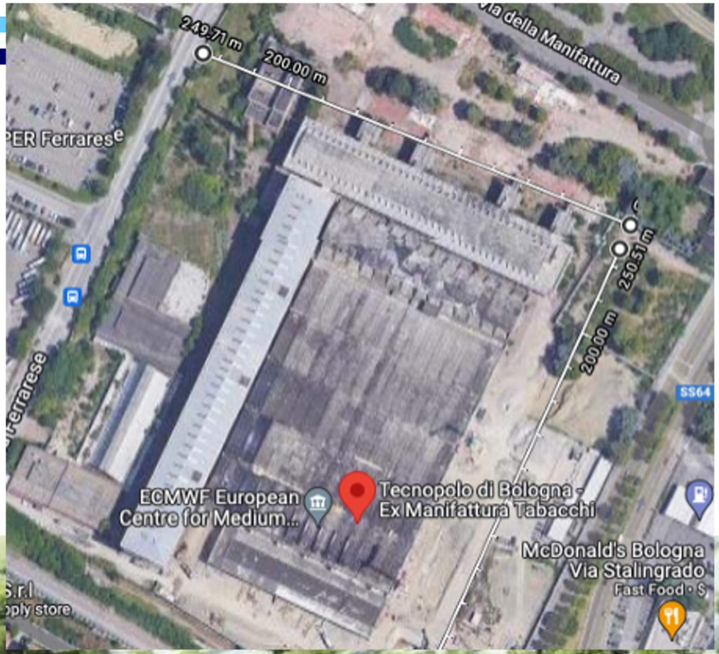
The "Big Data Valley"

- In **2017**, Bologna won a bid to host the "*European Centre for Medium-Range Weather Forecasts*" (ECMWF)

- The Emilia-Romagna region decided to repurpose the "*Manifattura Tabacchi **" area in Bologna to host a technology district, for ECMWF <u>and more</u>: the Tecnopolo

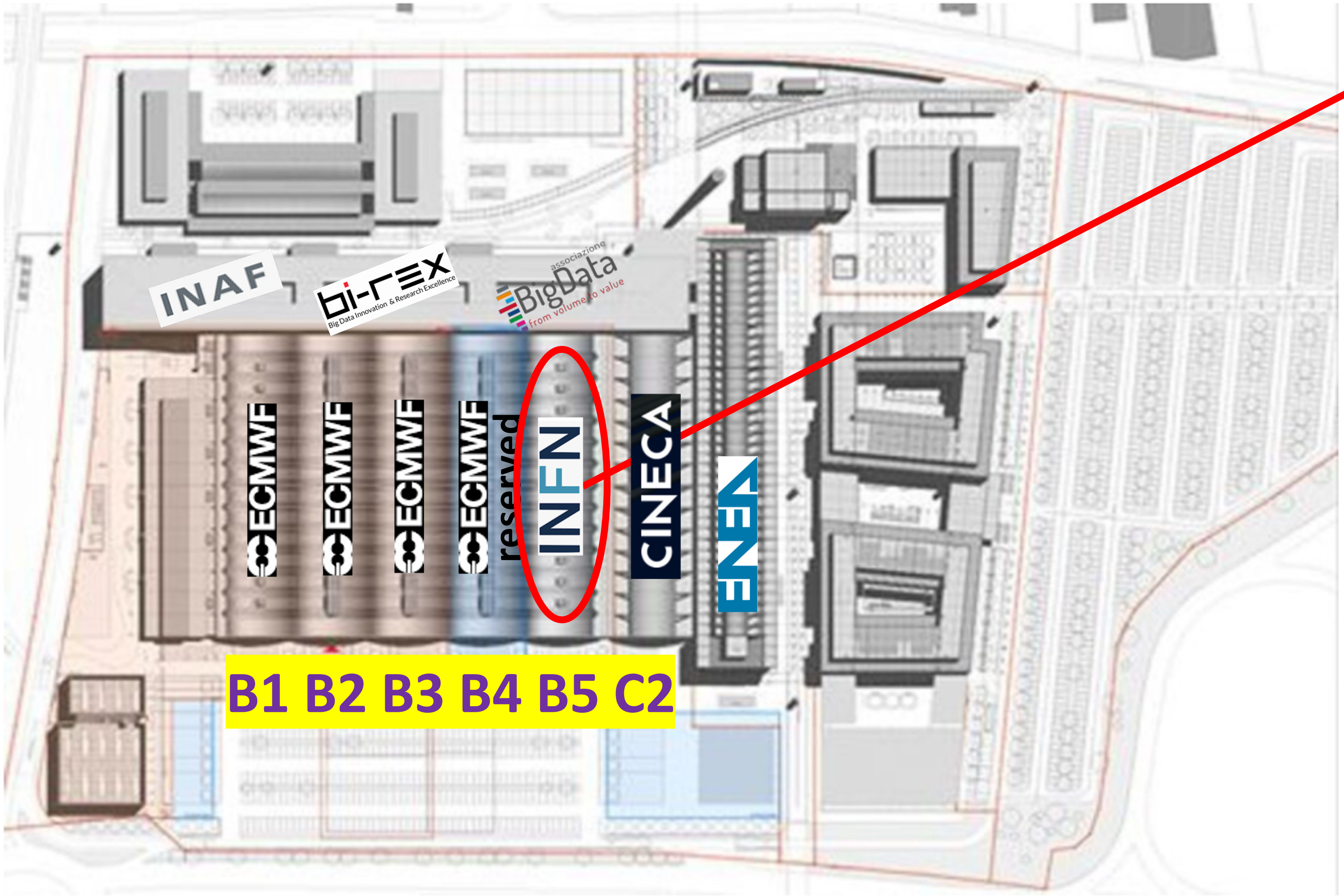
(*a former tobacco factory)

Roughly 250x250 m^2



How it will be

What will be at the Tecnopolo?



Each of the 6 “botti” (barrels) is ~5000m^2 of usable IT space

Same architect and design of the ”Sala Nervi” in the Vatican

The INFN + CINECA project

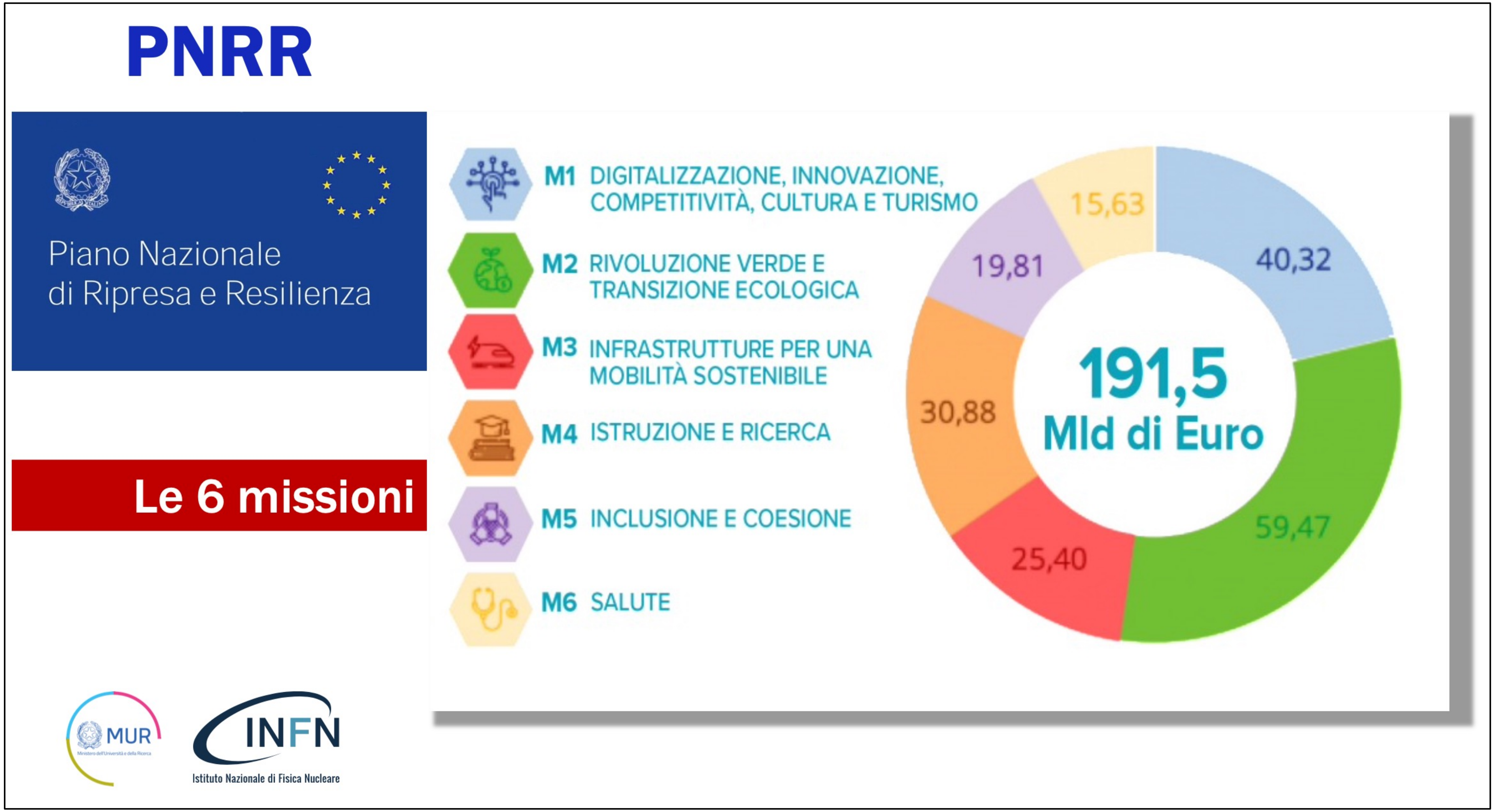
- ECMWF is up and running!

See the “Datacenter Session” on 12/10 for more details

- The CINECA (“C2”) and INFN (“B5”) barrels are expected to be ready by:
  - ~October 2022 (CINECA)
  - ~mid 2023 (INFN)
- Two phases are expected:
  - **Phase-1 (2023-2025)**: Leonardo + CNAF data center relocated. Total **13 MW.**
  - **Phase-2 (2025+)**: infrastructure up to **23 MW** ready for post-exascale and for the next generation of scientific experiments.

Opportunities: The Italian National Recovery and Resilience Plan (NRRP)

The NRRP Budget



Some NRRP projects already approved and led by INFN…

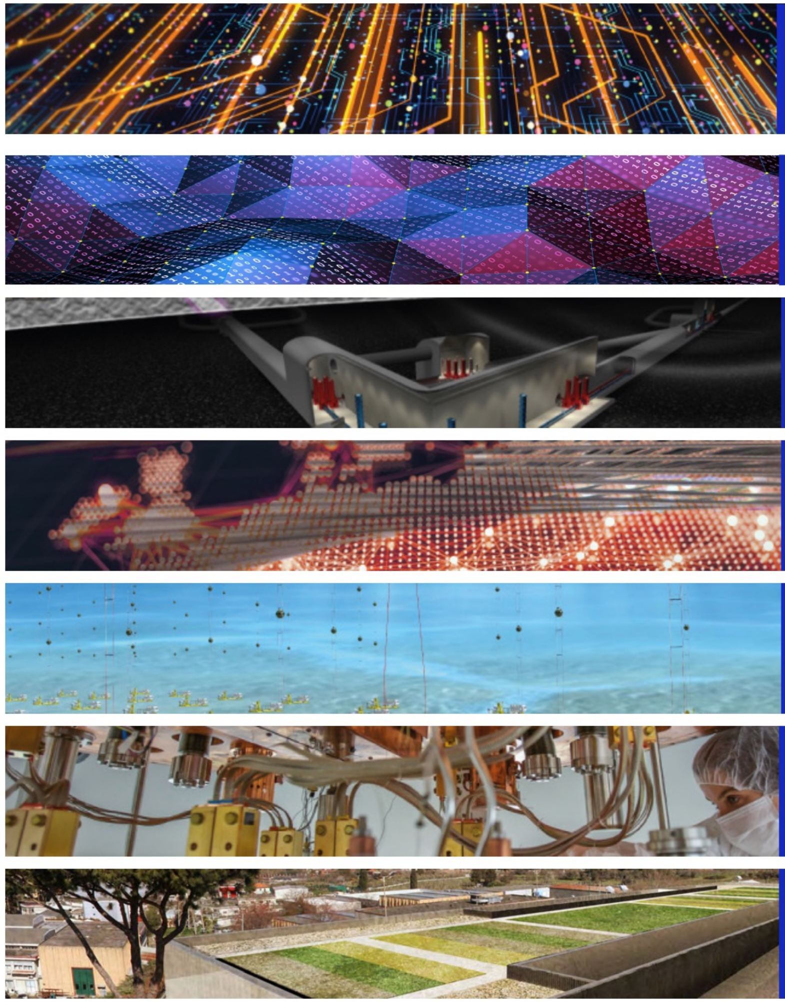
PNRR

Piano Nazionale
di Ripresa e Resilienza

MUR

INFN
Istituto Nazionale di Fisica Nucleare

Tutti i progetti approvati CN , IR, Ecositemi e partenariati estesi



ICSC
National Centre for HPC, Big Data and Quantum Computing

TeRABIT

ETIC
Einstein Telescope Infrastructure Consortium

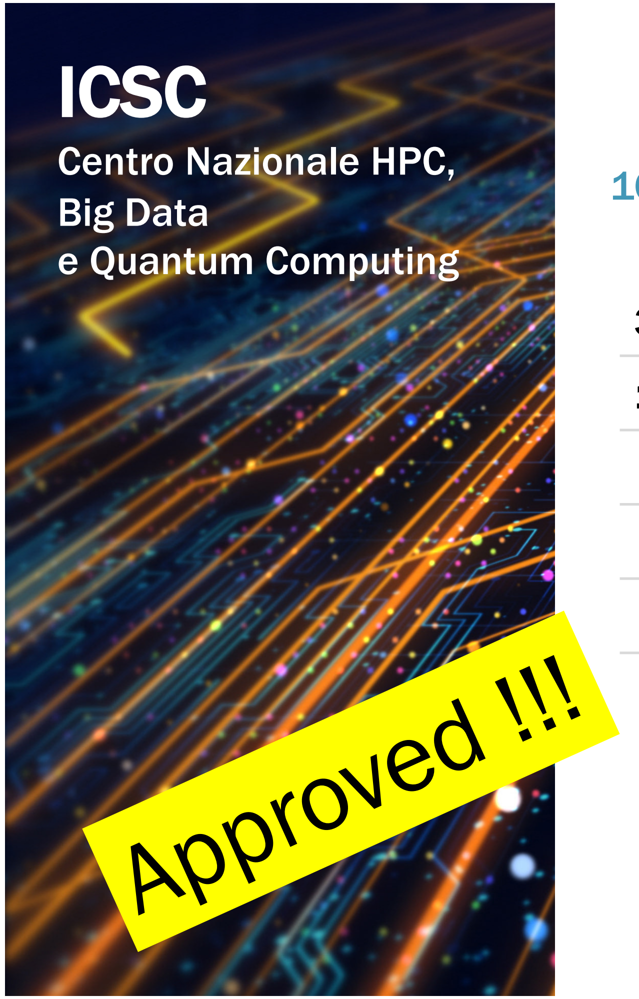
IRIS
Innovative Research Infrastructure on applied Superconductivity

KM3NeT

LNGS: Gran Sasso National Laboratory upgrade

EuPPS
LNF


INFN
Istituto Nazionale di Fisica Nucleare
ICSC
Centro Nazionale di Ricerca in HPC,
Big Data and Quantum Computing
Centro Nazionale HPC,
Big Data e Quantum Computing



Piano Nazionale di Ripresa e Resilienza · MUR · INFN Istituto Nazionale di Fisica Nucleare

Cloud national infrastructure for supercomputing.
Hub & Spoke organization:
10 vertical spokes for technology developments and software applications

320 + 41 M€ Total Budget

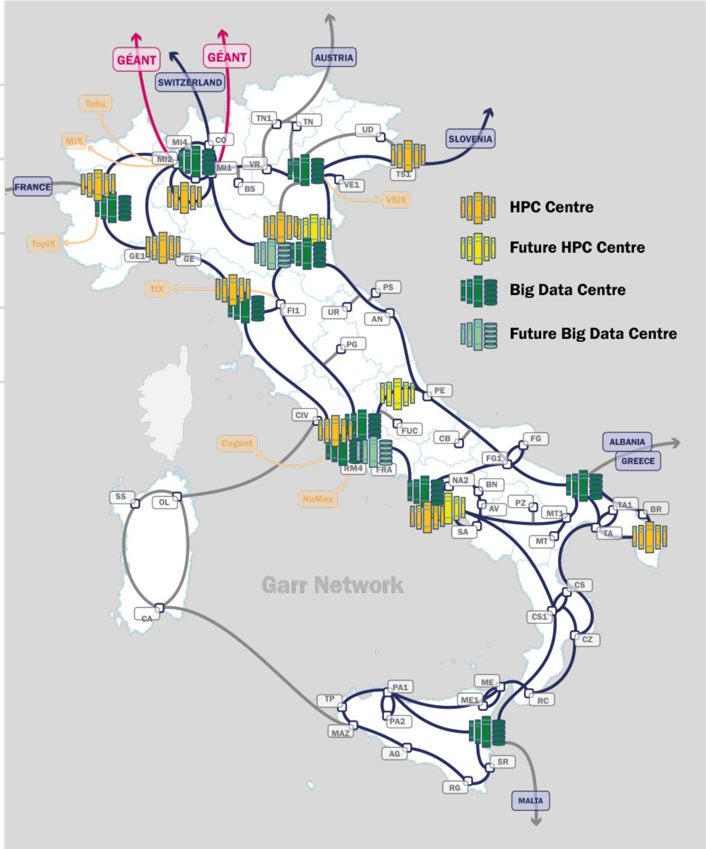
139 M€ Cloud Infrastructure

32 M€ Open Call

32 M€ Innovation & TT

42% investment South Regions

34	MUR Universities and Research institutions
15	Private Companies
1575	Researchers and engineers
250	New Temporary positions
250	New PhD
40 %	Female



Public Research Institutions
Founding Members:
a pervasive initiative throughout Italy

Enti Nazionali

INFN Istituto Nazionale di Fisica Nucleare
CINECA
ENEA
Consiglio Nazionale delle Ricerche
INAF Istituto Nazionale di Astrofisica
Istituto Nazionale di Geofisica e Vulcanologia
Consortium GARR

Solo HUB

UNIMORE Università degli Studi di Modena e Reggio Emilia
Università di Parma
OGS

Private companies
Founding members:
strategic players for digital transformation

LEONARDO · Terna Driving Energy · sogei · eni · autostrade per l'italia · ENGINEERING · FINCANTIERI · HUMANITAS RESEARCH HOSPITAL · FERROVIE DELLO STATO ITALIANE · ThalesAlenia Space · INTESA SANPAOLO · UnipolSai ASSICURAZIONI

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 DEVELOPEMENTE

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

The ICSC will include ten **thematic Spokes** and one ***Infrastructure*** **spoke**

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

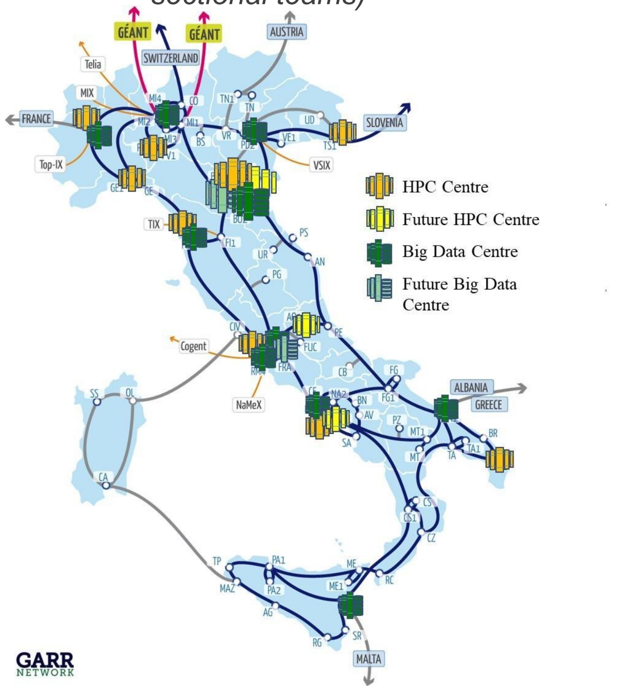
8 IN-SILICO MEDICINE & OMICS DATA

9 DIGITAL SOCIETY & SMART CITIES

10 QUANTUM COMPUTING

0 SUPERCOMPUTING CLOUD INFRASTRUCTURE

equipped with high-level teams of experts integrating the Spokes working groups (mixed cross-sectional teams)



PNRR Diego Bettoni 10

ICSC: Main figures over the next 3 years

1575
Personnel shared by partners

200+
New researchers

250+
New PhDs

32 ME
Open calls

Sustainable & Well balanced (territory, gender, age, size)

32 ME
Innovation grants

PNRR Diego Bettoni 11

NRRP: TeRABIT

- Complementing ICSC, TeRABIT (**Te**rabit network for **R**esearch and **A**cademic **B**ig data in **IT**aly) is a 41M€ NRRP approved project to create a distributed, hyper-networked, hybrid Cloud-HPC computing environment.
- This will be done by integrating and upgrading three leading digital Research Infrastructures: GARR-X, PRACE-Italy and HPC-BD-AI (the INFN computing infrastructure).
- Specifically, INFN targets TeRABIT to expand INFN Cloud for **the creation of distributed "HPC Bubbles"**.
  - The aim is to realize a **scalable open Edge-Cloud Continuum, integrating AI technologies**. Through this architecture, we will be able to dynamically and efficiently ingest, process and re-process data at multiple locations ("keep computing close to data production – whenever possible and sensible").

Opportunities: Health-related Initiatives

EPIC Cloud

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

- Motivation: the GDPR states that Clinical and medical data (in particular genomic) is personal data (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 the **EPIC Cloud**: a reference Cloud implementation for the treatment of sensitive data at INFN.

See the presentation on EPIC on 11/10 for more details

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.

ACC / Health Big Data

Alleanza Contro il Cancro

The National Oncology Network founded in 2002 by the Ministry of Health, joined by 51 IRCCS, ISS, AIFA, INFN and Politecnico di Milano and several patients' associations to perform translational research in the field of cancer research.

- Genomic pseudonymized data
- GDPR applies
- Italian Data Protection Authority rules apply

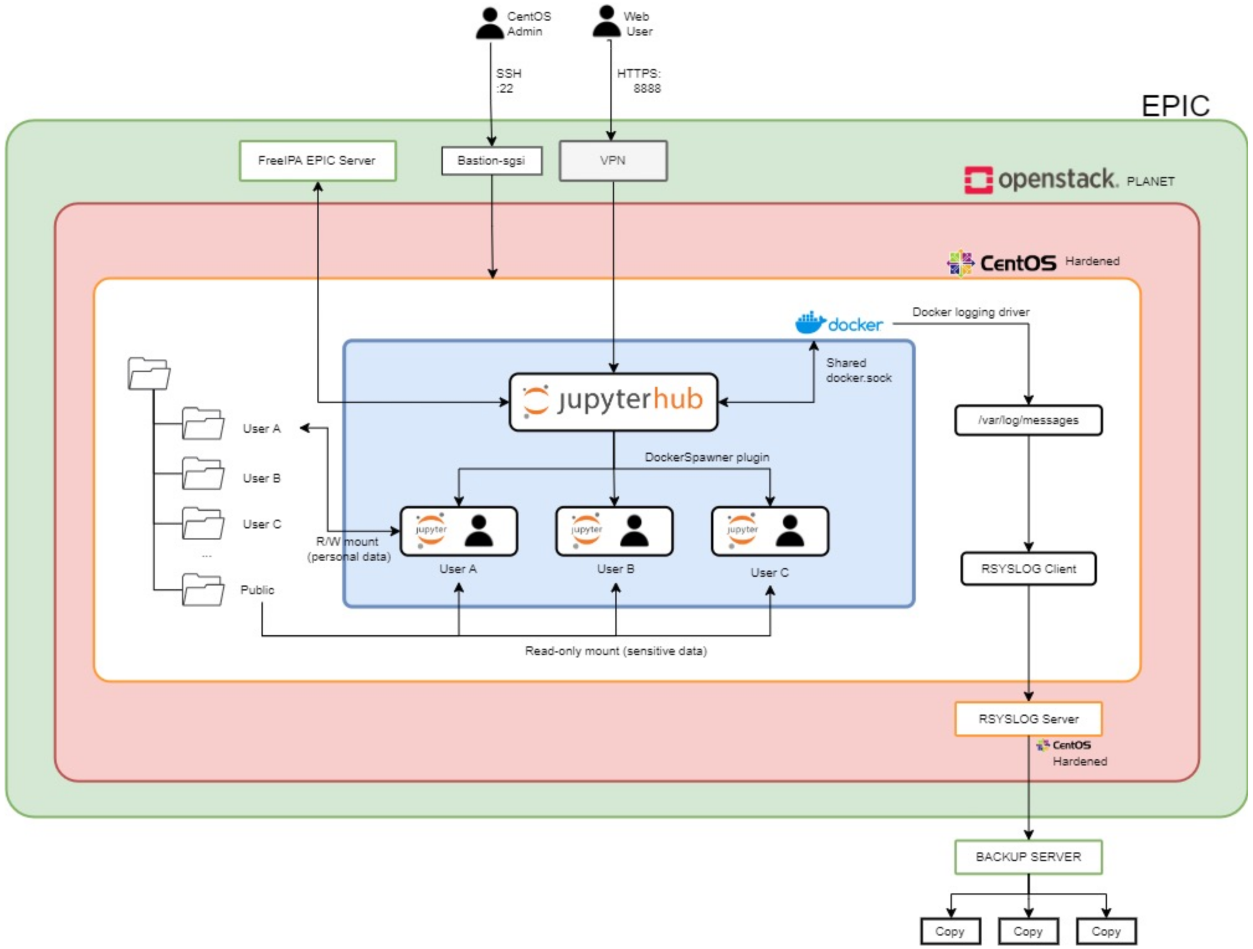
https://www.alleanzacontroilcancro.it/en/

PLANET

Pollution Lake Analysis for Effective Therapy

An INFN-funded research initiative, aiming to implement an observational study to assess a possible statistical association between environmental pollution and Covid-19 infections, symptoms and course.

- Data:
  - Pseudonymized clinical data (Covid-19 and Electronic Health Records from several hospitals)
  - Atmospheric data, population density, urban vs rural environment, mobility, socio-economic conditions
- Regulated by GDPR, Italian Data Protection Authority and by a convention between INFN and the Italian Istituto Superiore di Sanità.
- Container-based platform
- Data processing through a JupyterHub service
- 2FA for both analysis and infrastructure access



INFN-AIFM collaboration

ASSOCIAZIONE ITALIANA di FISICA MEDICA e SANITARIA
La Fisica al servizio della salute

INFN

To support the development and large-scale validation of AI-based tools for medical applications a joint effort has been conducted by the Medical Physics and High Energy Physics communities.

- Design of a dedicated computing infrastructure.
- The goal is to support the development and extensive validation of AI-based tools for **precision medicine** in the field of diagnostic and therapeutics.

Physica Medica xxx (xxxx) 1–11

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

ELSEVIER

Invited commentary

Enhancing the impact of Artificial Intelligence in Medicine: A joint AIFM-INFN Italian initiative for a dedicated cloud-based computing infrastructure

Alessandra Retico[a], Michele Avanzo[b], Tommaso Boccali[a], Daniele Bonacorsi[c,d], Francesca Botta[e], Giacomo Cuttone[f], Barbara Martelli[g], Davide Salomoni[g], Daniele Spiga[h], Annalisa Trianni[i], Michele Stasi[j], Mauro Iori[k,*], Cinzia Talamonti[l,m]

ARTICLE INFO

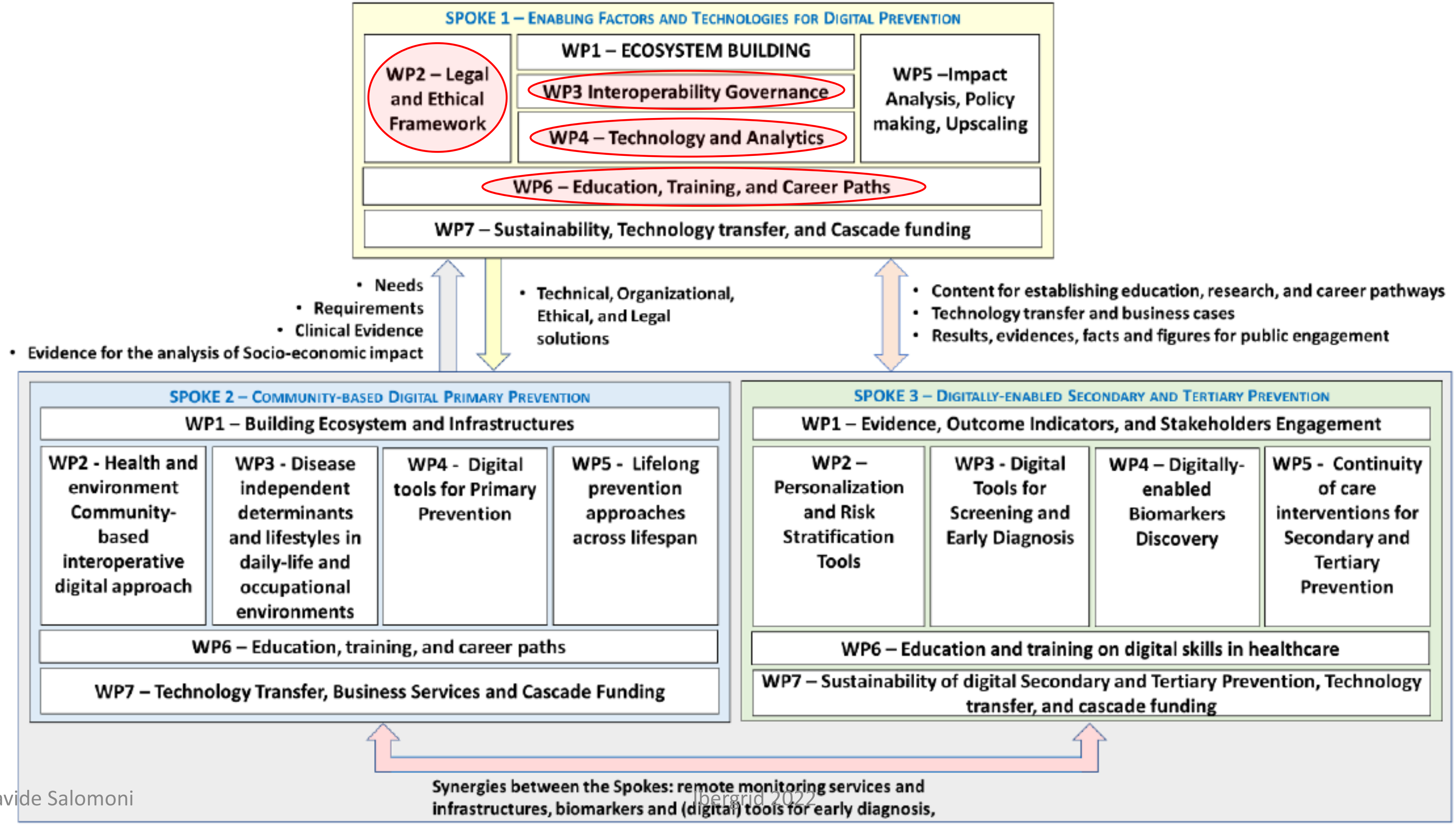
Keywords:
Artificial intelligence
Decision support systems
Computing infrastructure
Distributed learning

ABSTRACT

Artificial Intelligence (AI) techniques have been implemented in the field of Medical Imaging for more than forty years. Medical Physicists, Clinicians and Computer Scientists have been collaborating since the beginning to realize software solutions to enhance the informative content of medical images, including AI-based support systems for image interpretation. Despite the recent massive progress in this field due to the current emphasis on Radiomics, Machine Learning and Deep Learning, there are still some barriers to overcome before these tools are fully integrated into the clinical workflows to finally enable a precision medicine approach to patients' care. Nowadays, as Medical Imaging has entered the Big Data era, innovative solutions to efficiently deal with huge amounts of data and to exploit large and distributed computing resources are urgently needed. In the framework of a collaboration agreement between the Italian Association of Medical Physicists (AIFM) and the National Institute for Nuclear Physics (INFN), we propose a model of an intensive computing infrastructure, especially suited for training AI models, equipped with secure storage systems, compliant with data protection regulation, which will accelerate the development and extensive validation of AI-based solutions in the Medical Imaging field of research. This solution can be developed and made operational by Physicists and Computer Scientists working on complementary fields of research in Physics, such as High Energy Physics and Medical Physics, who have all the

http://doi.org/10.1016/j.ejmp.2021.10.005

NRRP: DigitAl lifelong pRevEntion (DARE)



Final considerations

Some recurring keywords

- **Integration** between the "traditional" infrastructural and the "Cloud Native" models.
  - But integration of *what*? Of resources, data sources, methods, people, solutions.
- **Continuum** between edge, cloud and HPC, in the recognition that "one size (even if big, or *Exascale*) does not fit all".
- **Simplification** for what regards [secure] data access, use and reuse. No matter where the resources are, no matter what the vertical domain is.
- **Continuous learning, innovation and adaptation**, if we want to profit from opportunities such as those offered by NRRP funds and by the increasing number of scientific collaborations.

Conclusions: the Background

1. INFN is **revisiting and expanding its computing infrastructure, management structure and expertise** to tackle the challenges expected in the next 10+ years in several scientific domains.

2. We have **identified many initiatives and projects** that, at multiple levels, are in line with the INFN mission to serve science. The Bologna Technopole and NRRP funds are precious opportunities here.

Conclusions: the Foreground

3. Our overall technological approach is to try and **abstract from where resources are, leveraging *aaS* models** to build a scalable [trans-]National federated structure integrating know-how, people, hardware, solutions.

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

Conclusions: the Synthesis

5. However, all these tasks cannot easily be performed by a single country, let alone a single institution. **We need to keep on working collaboratively, sharing know-how and ideas**. Initiatives such as Ibergrid are instrumental in facilitating this (many thanks to the organizers!).

Also at Ibergrid 2022

- Barbara Martelli, **“EPIC Cloud: a secure, GDPR-compliant, open-source cloud platform for life-science applications”** → 11/10, 13:30-14:00
- Session on **INDIGO-DataCloud** → 11/10, 13:30-15:30
- Davide Salomoni, **“The INFN infrastructure at the Bologna Technopole”** → 12/10, 11:18-11:27

- For more information on these slides: davide@infn.it.