EPJ Web of Conferences **295**, 10003 (2024) https://doi.org/10.1051/epjconf/202429510003
CHEP 2023

ICSC: The Italian National Research Centre on HPC, Big Data and Quantum computing

Claudio Grandi[1*], *Diego* Bettoni[2], *Tommaso* Boccali[3], *Gianpaolo* Carlino[4], *Daniele* Cesini[5], *Luca* dell'Agnello[5], *Giacinto* Donvito[6], *Davide* Salomoni[5§], and *Antonio* Zoccoli[1,7]

[1]INFN Bologna, Viale Carlo Berti-Pichat 6/2, 40127, Bologna, Italy
[2]INFN Ferrara, via Saragat 1 - I 44122 Ferrara, Italy
[3]INFN Pisa, Largo B. Pontecorvo, 3 – 56127 Pisa, Italy
[4]INFN Napoli, Complesso Universitario di Monte S. Angelo, ed. 6 - Via Cintia, 80126, Napoli, Italy
[5]INFN CNAF, Viale Carlo Berti-Pichat 6/2, 40127, Bologna, Italy
[6]INFN Bari, Via E. Orabona 4 - 70125 - Bari, Italy
[7]Università degli Studi di Bologna, DIFA, Via Irnerio 46, 40126, Bologna, Italy
[§]Now at: Fondazione ICSC, Via Magnanelli 2, 40033, Casalecchio di Reno, Italy

Abstract. ICSC ("Italian Center for SuperComputing") is one of the five Italian National Centres created within the framework of the NextGenerationEU funding by the European Commission. The aim of ICSC, designed and approved through 2022 and eventually started in September 2022, is to create the national digital infrastructure for research and innovation, leveraging existing HPC, HTC and Big Data infrastructures and evolving towards a cloud data-lake model. It will be available to the scientific and industrial communities through flexible and uniform cloud web interfaces and will be relying on a high-level support team; as such, it will 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. The ICSC IT infrastructure is built upon existing scientific digital infrastructures as provided by the major national players: GARR, the Italian NREN, provides the network infrastructure, whose capacity will be upgraded to multiples of Tbps; CINECA hosts Leonardo, one of the world largest HPC systems, with a power of over 250 Pflops, to be further increased and complemented with a quantum computer; INFN contributes with its distributed Big Data cloud infrastructure, built in the last decades to respond to the needs of the HEP community. On top of the IT infrastructure, several thematic activities will be funded and will focus on the development of tools and applications in several research domains. Of particular relevance to this audience are the activities on "Fundamental Research and Space Economy" and "Astrophysics and Cosmos Observations", strictly aligned with the INFN and HEP core activities. Finally, two technological research activities will foster research on "Future HPC and Big Data" and "Quantum Computing".

* Corresponding author: Claudio.Grandi@bo.infn.it

1 Scenario and motivations

In the next years an unprecedented amount of data will be produced by scientific, industrial, and institutional actors, with an increasing challenge when extracting social and economic value from this data explosion. In this context, supercomputing, numerical simulation, artificial intelligence, high-performance data analytics and big data management will be essential and strategic for understanding and responding to grand societal challenges and in stimulating a people-centred process of sustainable growth and human development, allowing academia, industry, and institutions to develop services and discoveries.

Since many years US, China and Japan are making great strides in these frontiers; Europe started slightly later and in the last years has defined a clear strategy. Among the main pillars are: the EuroHPC Joint Undertaking [1], the European Open Science Cloud (EOSC)[2], the European Processor Initiative (EPI) [3] and the Quantum Flagship [4], according to the European Data Strategy and the European Approach to AI. EU Member States have made significant investments in European Petascale and pre-exascale infrastructures. EuroHPC has put exascale supercomputers on the roadmap, as well as Quantum Computing. Within this framework Italy is investing in HPC co-funding with EuroHPC, Leonardo, a pre-exascale world-class Tier-0 system, which will offer highly competitive HPC services to Italian and European public and private researchers.

After the COVID-19 pandemic, the European Union has started the NextGenerationEU program, aiming for a fast recover from the emergency, and at the same time transforming the European economies and societies. Out of the total 806.9 G€ funding, 191.5 G€ have been assigned to Italy[†]. In Italy, 30.88 G€ are devoted to research and education, out of which 11.44 G€ are for the initiative "From Research to Business" and 1.58 G€ for the initiative "Research Infrastructures".

To follow up on these actions, in 2022 a project to implement the Italian National strategy for HPC and Big Data has been prepared, addressing the following points:

- build a world-class supercomputing cloud infrastructure to store, manage and analyse all the data produced, in combination with hardware general purpose cloud infrastructure;
- set up centres of excellence with high level teams of experts to develop domain applications;
- set up strong links between the scientific community and the industrial system, (addressing both corporations that need HPC and SMEs that need easy, fast and scalable services in a non-HPC DevOps infrastructure);
- train young scientists as well as managers to become experts in these fields;
- implement organic and structural measures for innovation and for increasing the Technology Readiness Level (TRL), also for SMEs without specific HPC skills and resources;
- build capacity by sharing expertise within the broader scientific community and strengthening the areas where Italy is lagging behind;
- assess and maximize the social and economic impact of the initiative at local, national and EU level with special attention to ethical implications;
- set-up a group of experts with the specific task of analysing and monitoring the ethical implications of big-data use and management.

The *Italian National Research Centre on HPC, Big Data and Quantum computing* (ICSC) [5] has been funded in the "From Research to Business" initiative with 320 M€. In addition, the synergic project *Terabit Network for Research and Academic Big Data in Italy*

[†] NextGenerationEU is also known in Italy as *Piano Nazionale di Ripresa e Resilienza* (PNRR)

(TeRABIT) [6] has been funded with additional 41.0 M€ in the "Research Infrastructures" initiative. ICSC aims at creating the national digital infrastructure for research and innovation, starting from the existing HPC, HTC and Big Data infrastructures and evolving towards a cloud data-lake model accessible by the scientific and industrial communities. This is obtained using flexible and uniform cloud web interfaces, relying on a high-level support team, forming a globally attractive ecosystem based on strategic public-private partnerships to fully exploit top level digital infrastructure for scientific and technical computing and promoting the development of new computing technologies.

2 Organization of the ICSC project

ICSC started in September 2022 and the implementation phase will last 36 months. It is structured according to the hub and spoke model: the hub is responsible for the validation and management of the research program, whose activities are elaborated and implemented by the spokes and their affiliate institutions, as well as through open calls. The hub implements all the activities on education and training, entrepreneurship, knowledge transfer, policy, and outreach, and coordinates a transversal research group on Societal Implication and Impact. Hub and spokes consist of Universities, Research Institutions and private and public operators. ICSC includes one cross spoke, Spoke 0 ("Supercomputing Cloud Infrastructure"), and 10 thematic Spokes, as shown in Fig. 1 and described in the following.

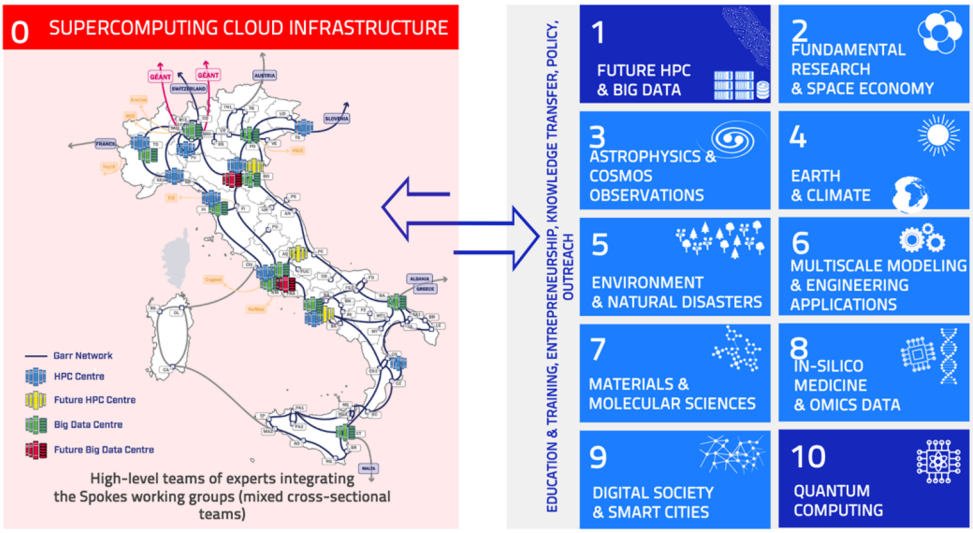

Fig. 1. The eleven spokes of the Italian National Research Centre on HPC, Big Data and Quantum computing (ICSC)

The Spoke 0 is built upon a supercomputing cloud infrastructure that includes resources from the major national players: the network provided by the Italian NREN GARR, the HPC systems provided by CINECA and the HTC/Big Data infrastructure provided by INFN. The 10 thematic spokes involve all the relevant institutions and national experts in their respective fields, and they will provide an ecosystem to guide and support the scientific communities and the industrial system, including SMEs and Public Administration. The Spokes from 2 to 9 develop advanced applications based on the infrastructures made available and maintained by Spoke 0. These Spokes represent the scientific domains coherent with the PNRR guidelines and chosen for the relevance with the respect to the ICSC's objectives. Spokes 1 and 10 are focused on the development of high potential future technologies: advanced

codesign of high-performance and high-throughput hardware and software systems, and development of expertise on hardware and software quantum computing, respectively.

In addition to the Universities and Research Institutions, ICSC involves several Italian companies aiming to establish a synergy between the scientific community and the industrial system to the advantage of national research and productive systems. ICSC will contribute to make Italy competitive in the economic challenges of the coming years, supporting the production system in making the most of the data collected along the specific processes, and master the hardware and software technologies connected to them, providing the right combination between computing power and cost effective, quick and easy accessibility to data and application services. Dedicated targeted activities to increase the TRL to market levels are being developed. Finally, a basic element of the ICSC programme is to form a new cohort of computing savvy researchers and Ph.D. laureates, able to steer the transfer methods and technological solutions from academia to the productive sectors, mastering the most up-to-date DevOps, DataOps and MLOps framework. More generally, the programme is aimed at building capacity to strengthen Italian competitiveness and guaranteeing long-term sustainability.

The entities participating to ICSC are 25 universities, 12 research institutes and 14 strategic private companies, as per the limits imposed by the call. In total, about 1500 staff are shared by the partners, more than 250 new fixed term personnel and more than 250 PhD will be hired. Part of the 320 M€ will be used for open calls (32 M€) and for innovation grants (32 M€).

The ICSC headquarters will be located in Bologna, in the new area known as the *Big Data Technopole*. In this area are already operating the data centre of the European Centre for Medium-Range Weather Forecasts (ECMWF), the new pre-exascale HPC system Leonardo operated by CINECA and will soon the INFN-CNAF Tier-1 computing centre, that will be migrated in fall 2023. Other research institutions will also move their offices.

3 TeRABIT

The TeRABIT project aims to upgrade three existing Italian digital research infrastructures: the GARR-X (now GARR-T) network infrastructure by GARR, the Prace-Italy HPC system by CINECA and OGS and the HPC-BD-AI distributed Big Data infrastructure by INFN. Not by chance the three infrastructures are owned by the same partners of ICSC Spoke-0. Indeed, the two projects are built to be completely synergic, even if with some specificities. Network interventions by GARR will be in different regions, in particular TeRABIT foresees the upgrade of the connection to Sardinia, where is one of the sites nominated to host the Einstein Telescope gravitational wave observatory. CINECA will upgrade the Galileo 100 Tier-1 system, and INFN will implement small heterogeneous HPC systems, the so called HPC bubbles, to address use cases not addressed by ICSC. TeRABIT has been funded with 41 M€, started on January 2023 and the implementation phase will last 30 months.

In the following, the ICSC infrastructure will be presented without making any distinction among these two projects.

4 The supercomputing cloud infrastructure

4.1 Network

GARR, the Italian Research & Education Network, is deploying the new GARR-T infrastructure able to scale to multiples of Terabits per second (Tbps). GARR-T(erabit) is the new generation network of GARR which is seamlessly evolving from the current

infrastructure to an innovative network based on the state of the art of transmission technologies and will be future proof. The main advantage of this new network architecture is the capability to adapt to the increasing requirements of users in terms of capacity, capillarity, and services. A larger and pervasive use of automation will be one of the major characteristics of the new network that will thus be more dependable and measurable thanks to sophisticated monitoring systems.
Technically GARR-T is a fibre optic network (currently comprising more than 16.000 km) owned in IRU (Indefeasible Right of Use) using a high-performance transmission system based on Open Line System technologies with an orchestration based on software and hardware under the full control of GARR, able to scale to several tens of Terabits.
The new infrastructure will follow state-of-the-art paradigms that will allow to manage the network in a flexible way, giving access to the lower optical level also to users not on the same premises of the PoP (Point of Presence). It will also implement new solutions in terms of security and Edge Computing.
GARR-T has been designed and is being implemented to fulfil the needs of Research and University in Italy for next decade in complete synergy with the evolution of the European Research and Education Network GEANT aiming to build a common and open optical network landscape in Europe. As part of this process, the international connectivity between the Italian and European HPC centres is being increased to address the challenges of the next decades.

4.2 HPC systems

The European HPC (EuroHPC) Joint Undertaking co-funded three pre-Exascale systems in Europe. One of them, Leonardo, is hosted in the Bologna Technopole, managed by CINECA and started its operations in late 2022. In order to further develop and provide access in Italy to a world leading computing service and data infrastructure - by the means of high-end supercomputers indispensable to run the most demanding and strategic applications - as part of the ICSC actions the operational performance of Leonardo is being further increased. The system will be upgraded with an investment, economical equivalent to the 30% of the initial investment of 120 M€, as also foreseen in the strategic plan of EuroHPC, which will co-fund such an upgrade. This will provide new functionalities to address the evolution of Italian research system. Moreover, the integration on the Leonardo platform of a quantum computing physical system, is foreseen, to combine digital and quantum technology in advanced use and scientific cases, that also in line with the strategic plan of EuroHPC. CINECA, besides the deployment and the exploitation of the Tier-0 HPC national supercomputer system, will manage and deploy also the main Tier-1 HPC systems at national level. Galileo 100, the CINECA Tier-1 system already in production in the Casalecchio data centre will be integrated in the National Centre HPC infrastructure; one more Tier-1 system will be installed in the Bologna Technopole, and one in Naples. The storage system HPC system infrastructure will be also upgraded as well, moving from the current capacity in the order of 100 PB to the level of some hundreds of Petabytes of online capacity.

4.3 Big Data federated cloud

INFN is operating a system of ten data centres spread over Italy, able to manage and elaborate the data coming from several (about 50 currently) scientific experiments, including those at the CERN LHC collider. By 2025, the computing and storage capacity will be increased by at least a factor 2 to offer services also to research domains beyond those directly connected to INFN.

Besides the main data centre, that will be migrated to the new Bologna Technopole in fall 2023, the facilities of all other data centres will be upgraded to be able to sustain the increased load foreseen in the next decade. Two new data centres are being built in INFN-LNGS, focused on natural and anthropic disaster mitigation, and in INFN-LNF, focused on space economy. The processing capacity will be enhanced to increase the offer for all scientific domains, thus facilitating the development and migration of HPC optimized scientific applications toward state-of-the-art, federated and scalable solutions. In order to fulfil the needs of as many domains as possible, small heterogeneous systems (e.g., conventional, accelerated, FPGA, ...) for both R&D and production purposes are being procured. Furthermore, systems to manage data at different level of confidentiality, e.g., sensible data coming from medical applications, genomics, etc. are being created.

The overall ICSC architecture is based on federation services enabling the aggregation at multiple levels (IaaS, PaaS or SaaS) of existing data centres managed by the partners of the National Centre through open-source middleware. The architectural model is dynamic, so that both the resources and the application domains can be volatile. Figure 2 shows the proposed layered architecture.

The high-level federated architecture will be compatible with EuroHPC and the European Open Science Cloud (EOSC). It will be based on a clear separation between the physical and logical layers. At the logical level, all the resources are seen as part of an Italian research data-lake. Each scientific domain will have a unique entry point to its own data-lake that is connected to the top-level, national entry point.

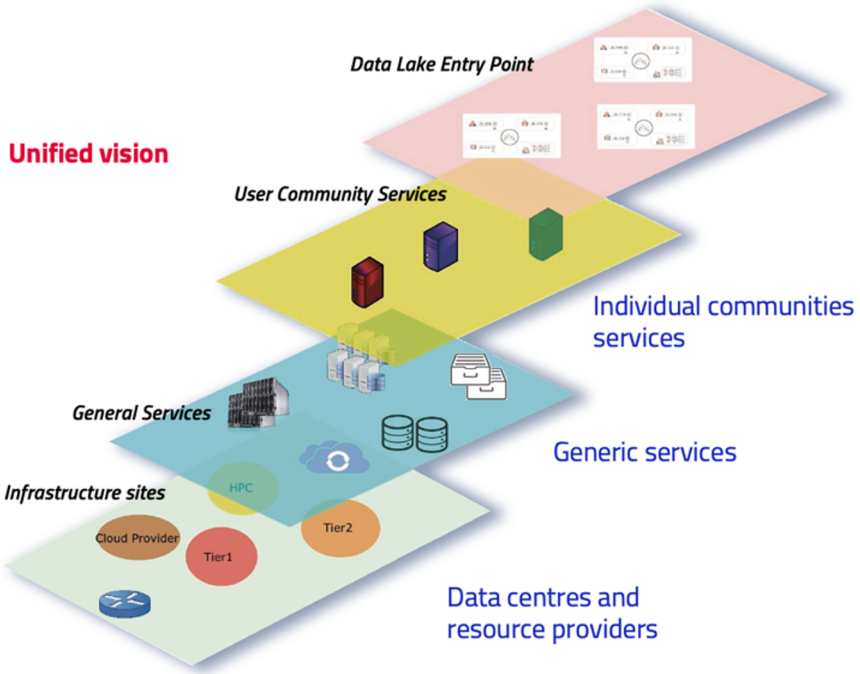

Fig. 2. Layered architecture of the ICSC data-lake

Each physical data centre can support one or more application domains, by publishing specific capabilities and running application-specific services. The middleware will make it possible to create, evolve and share a service portfolio composed of both core components and possibly reusable application-based, composed services. Each application domain data-

lake is defined as the sum of the services (e.g., portals, software services, resources) available for that domain. A dynamic match-making mechanism connects the physical and logical domains allowing to fulfil the high-level requests coming from end users. More details about the middleware infrastructure may be found in [7].

5 Applications development

This paper focused on the impact of the infrastructure, but ICSC makes strong investments also on applications development and technological research. Of particular relevance to the HEP field are Spoke 2 "Fundamental research & space economy" and Spoke 3 "Astrophysics & cosmos observations".

Activities address the sustainability of the computing methodologies used in the research field in the next future. Key directions are:

1. Economic sustainability, where the computing infrastructure (including resources, codes and algorithms) needed to extract knowledge from the collected / produced data can be procured within a reasonable cost envelope. According to the currently adopted standard criterion adopted for the computing of high-energy physics experiments, economic sustainability means that the yearly budget for computing must at most be constant at today's level. Solutions such as the use of heterogeneous computing with accelerators promise to decrease the cost per operation, while data-lake infrastructure designs point towards a reduction of the total storage deployment.
2. Environmental sustainability. This includes many aspects:
    a. The design of algorithms with a better cost per operation. This includes brand new solutions that may be designed, but also general strategies like a more pervasive utilization of Machine Learning technologies, which are known to lower the processing time in situations such as heavily combinatorial algorithms.
    b. The use of heterogeneous technologies with a better operation/Joule cost. This includes the use of mobile-derived CPUs like ARM, known to perform at least 3x better energetically, and the use of GPGPU solutions, which scale similarly by using a more parallel architecture.
3. Scientific sustainability, in which the algorithms increase in complexity when analysing data from more and more precise experiments. This is crucial to the advancement of the scientific domains and implies that the resource needed will increase faster than a naive extrapolation from data rates would suggest.

Training and dissemination programs are part of the activities.

6 Summary and outlook

The NextGenerationEU program funds are a unique opportunity to upgrade the Italian computing infrastructure for research, and to make a step forward towards a collaborative and fruitful collaboration with the national productive system.

The Italian Ministry for University and Research (MUR) asked INFN to drive the preparation of the project, and INFN has now a leading role in its implementation. This represents an important acknowledgement of the expertise built in the past years by INFN and by the HEP community at large, in the field.

The strong collaboration among the main players (Network, HPC and HTC/Big Data) will improve the efficiency and the effectiveness of the infrastructure. Besides the increased availability of resources, the evolution of the INFN infrastructure towards a more open

system, able to address the needs of a wider research community, will improve the flexibility and resilience of the system, also for the HEP community itself.

The increased number of staff building their careers on computing will help covering areas in our field that are currently understaffed.

The work presented in this paper has been funded by the NextGenerationEU European initiative through the Italian Ministry of University and Research, PNRR Mission 4, Component 2 - ICSC: Investment 1.4, Project code CN00000013 - CUP I53C21000340006; TeRABIT: Investment 3.1, Project code IR0000022 - CUP I53C21000370006.
This paper includes material prepared together with the INFN partners in ICSC Spoke-0 and TeRABIT (CINECA and GARR).

References

1. EuroHPC: https://eurohpc-ju.europa.eu
2. EOSC: https://eosc.eu
3. EPI: https://www.european-processor-initiative.eu/
4. QT: https://qt.eu
5. ICSC: https://www.supercomputing-icsc.it/en/icsc-home/
6. TeRABIT: https://www.terabit-project.it/ (in Italian)
7. F. Fanzago et al. "INFN and the evolution of distributed scientific computing in Italy", Proceedings of the CHEP 2023 conference (to be published)