CINECA
2021 - 2022
Dear colleagues,

the previous annual report summarizing the activities of the Cineca HPC department for the 2020-21 period focused on the united response of the scientific community to the COVID pandemic. A year later, the focus has shifted to the intense work that is facing us as early as the next twelve months to enhance the economic resources that have been mobilized with the Next Generation Europe action to recover from the effects of the pandemic and to support the priorities shared by the Member States of the European Union and by the international community.

In particular, we need to confront the ecological and energy transitions necessary to mitigate climate change, especially development sustainability and digital innovation. At a national level, Next Generation EU funding from the National Recovery and Resilience Plan (PNRR) has been structured into six missions: Digitization, innovation, competitiveness, culture and tourism; Green revolution and ecological transition; Infrastructures for sustainable mobility; Education and research; Inclusion and cohesion and Health. Cineca takes part in the fourth mission, Education and research, which is managed by the Ministry of University and Research (MUR) to which Cineca belongs. In this mission, specific implementation tools were identified: key technologies, research infrastructures, extended public-private partnerships and territorial ecosystems and Cineca is involved in funded projects which belong to all the fields in these sectors. The first is the project for the establishment of the National HPC, Big Data and Quantum Computing Foundation: Cineca is one of the promoters and founders of this center which is funded by the key technologies mechanism. HPC and its holistic ecosystem is one of the five key technologies defined by the MUR and associated with this action is also that relating to the financing of research infrastructures and Cineca is involved as an implementing body in some projects funded by this method. Overall, the economic resources assigned to Cineca for the enhancement of the national HPC infrastructure and the increase in staff for specialist support will be over 80 million euros.

In the field of key technologies, Cineca also participates as a founding member in the National Center for the Biodiversity. In actions financed with the extended partnerships mechanism, we participate in those relating to the Growing Resilient, Inclusive and Sustainable area and to the Translational Medicine area. As regards Ecosystems for territorial innovation, we are involved in the project relating to the territory of the Emilia Romagna Region for the area of Sustainability of Economic Development. Overall, the projects that see Cineca as a core partner have been funded with approximately one billion euros, of which approximately 10% has been directly assigned to Cineca.

This confirms on the one hand the centrality of the roles of HPC, ML and AI and of big data processing as enabling assets for advanced research and technological innovation, on the other hand it is recognition of the quality and contribution that Cineca can provide.

Naturally, the activity carried out over the last twelve months has been particularly characterized by the installation of the pre-exascale system of the Leonardo project, which will start production at the beginning of 2023. The Leonardo system is the most important project for the next period and we are already committed to upgrading the computing system to extend the general purpose partition with the integration of a partition configured with microprocessors based on HBM technology and the booster partition configured with the latest accelerators. Furthermore, the integration of a quantum computing system is already foreseen, and this will complete the tier0 infrastructure for the scale-out computing.

The evolution of the computing systems and the national HPC infrastructure, integrated and federated into the European HPC system must also include an investment in the staff of the high-level support team. To achieve this there will be an action which will see an increase in the next six months of at least fifty new structured figures in the department dedicated to this activity, to the national and European scientific communities, and to the flagship projects both at international levels such as Destination Earth and Copernicus, and at a national level, particularly the National Center for HPC, Big Data and Quantum Computing.
2021 IN NUMBERS
2 HPC Infrastructures
4 Leonardo
7 ICEI
8 Users’ statistics
14 Resource allocation: projects, agreements, collaborations
15 PRACE
17 Iscra
18 EUROFusion
22 Agreements
23 Training HPC
26 Staff
27 Events
33 HPC H-index
36 SCIENTIFIC OVERVIEW
38 LIFE SCIENCES
39 ORCHESTRA
41 ARNO
42 EARTH AND ENVIRONMENTAL SCIENCES
43 SEBASTIEN
44 Digital technology enablement for Destination Earth
BIG DATA AND ARTIFICIAL INTELLIGENCE
Ocean, Supercomputers and AI in action
Day-ahead hail prediction to prevent losses in agriculture

MATERIAL SCIENCES
MaX: HPC and HTC meet for Material Sciences

QUANTUM COMPUTING
Cineca quantum computing perspective

HUMANITIES AND SOCIAL SCIENCES, DIGITAL ECOSYSTEMS
The great cycle: the immersive experience of TUBKE Monumental
DARE: a digital environment for urban regeneration
DARE: IMC platform and the management of 3D digital objects

 COMPUTING PROVISIONING SERVICES

INDUSTRIES
Bringing HPC-fueled innovation to the EU industry EUROCC

NATIONAL RECOVERY AND RESILIENCE PLAN
High Performance Computing projects supported by the National Recovery and Resilience Plan – Next Generation EU
2021
IN NUMBERS
CINECA HPC is the main Center for scientific computing in Italy. It runs a large computing infrastructure and make it available to Italian and European Researchers, as well as to important Italian Companies within a programme for supporting national industrial competitiveness.

Here we present data about usage and users, projects, events and classes offered in the field of High Performance Computing.

At the end of this section we present our staff, the true resource that makes everything work.
Starting from March 2021, the HPC system Galileo was taken off the production and substituted with a new and more performant structure called Galileo 100, co-funded by the European ICEI (Interactive Computing e-Infrastructure) project and made available on August 2021. Engineered by DELL, this new infrastructure is based on PowerEdge processors and NVIDIA GPU accelerators and features 554 computing nodes, subdivided into 340 standard nodes, also called "thin nodes", 180 data processing nodes, also called "fat nodes" and 34 GPU nodes interconnected by a 100Gbs Infiniband interconnection and dedicated to visualization purposes, thus also called "viz nodes". Beyond these nodes, the infrastructure is equipped with 77 computing servers OpenStack for Cloud Computing since the Cloud Infrastructure, hosted on Meucci, was enhanced and moved onto this partition. The whole architecture is interconnected by a Mellanox Infiniband 100GbE internal network and it is supported by a 20 PB active storage, accessible from both the cloud and compute partition of Galileo 100, 1PB Ceph storage (full NVMe/SSD) for the cloud partition and 720 TB fast storage. Galileo100 is mainly used for special High-End projects for the technical and industrial HPC computing, as well as for dedicated Climate and Weather projects, reaching a peak performance of 3.53 TFlop/s per single node and 2PFlop/s in total. The Tier-0 cluster Marconi100 and Marconi-A3 remained unchanged in the 2021. Furthermore, on January 2021, a new specialised system for AI, based on NVIDIA DGX, has been made available for users. It is composed by 3 nodes and features Dual AMD ROME 7742 processors, accelerated by 8x NVIDIA A100 GPUs per node. Finally, in addition to aforementioned systems, other infrastructures are present and dedicated to special uses or to specific communities. These includes:

- **GATEWAY**: 28 nodes (Skylake Intel Xeon 8160) with 28 cores each, connected with a OmniPath network. It is a dedicated interactive cluster for EUROfusion.

- **CNAF**: 238 Broadwel nodes (Intel Xeon E5-2697) connected with a custom internal network and linked with a high performance link with the Bologna INFN computing room. The cluster is dedicated to INFN users for WLCG program.
<table>
<thead>
<tr>
<th>Supercomputer</th>
<th>Resources (Core Hours)</th>
<th>Performance capacity (Exaflops)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Galileo</td>
<td>96,771,456</td>
<td>14.516</td>
</tr>
<tr>
<td>Marconi A3</td>
<td>1,340,490,240</td>
<td>321.718</td>
</tr>
<tr>
<td>Marconi 100</td>
<td>274,589,184</td>
<td>988.521</td>
</tr>
<tr>
<td>Galileo100</td>
<td>58,715,136</td>
<td>15.545</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,770,566,016</strong></td>
<td><strong>1,340.299</strong></td>
</tr>
</tbody>
</table>
The rapid growth of the amount of data coming from diverse sources, including numerical models and sensors, requires more and more computing resources, able to process data in a timely fashion. Furthermore, the launch of new leading Italian and European initiatives, which face increasingly advanced computational challenges, are based on the availability of these technologies. Leonardo will be one of the most powerful high-performance computing (HPC) systems worldwide, capable of executing more than 250 Petaflops, i.e., more than 250 quadrillion (or 1015) operations per second. HPC systems of Leonardo class are key enabling technologies, aiming to raise the competitiveness of Italian and European research and paving the way to the forthcoming exaflopera in Europe. Leonardo can be considered a significant step forward in promoting European research in the field of computational sciences. The goal is to strengthen the European presence in high-performance computing, a strategic asset aiming to advance the technological growth of the member states of the Union. Cineca has a proven history of providing the most powerful supercomputers in the world. With the support of national and regional government, Cineca was designated by the EuroHPC initiative to host and operate Leonardo, one of the three recently procured precursors of Exascale systems. Leonardo will be installed at the newly built data center located in the Bologna Big Data Technopole beside the ECMWF premises. The Technopole can therefore be considered one of the main European hubs for computing and data processing. Leonardo features almost 5000 computing servers and over 100 PB of storage, all mounted on 155 Atos Sequana X2000 racks. Computing nodes are powered by state-of-the-art technology. The accelerated partition is constituted of 3456 nodes each equipped with 4 customized NVIDIA A100 GPUs. The CPU partition features next-gen Intel Sapphire Rapids processors and 512 GB of DDR5 RAM. All 5000 computing nodes are interconnected via an NVIDIA HDR200 based network with a Dragonfly+ topology to achieve maximum throughput and minimal latency. The storage is based on two DDN appliances targeting capacity and high IOPS. An important consideration is the energy efficiency of the system and its integration in the data center. The IT power consumption of the system is expected to reach over 6 MW, peaking at 9MW during burst runs. It is therefore especially important to reduce at the minimum any source of energy consumption other than IT. All the computing part of the system is under direct liquid cooling (cold plate) able to extract over 95% of the dissipated heat. The cooling water in the racks operates at a temperature range of 36-46 degrees and is provided through 4 independent dry cooling systems. This is the main reason for the estimated power usage effectiveness of Leonardo in the order of 1.1, classifying it as an extremely energy efficient system. Leonardo is expected to be operational from Q1 2023.
Figure 1: high-level logical overview of Leonardo supercomputer architecture
Figure 2: Leonardo Cineca
Interactive Computing E-Infrastructure (ICEI) is a project aiming to provide a set of composable services to be used to build the Human Brain Project and other scientific community driven digital platforms. In other words, the partners of the project (Cineca, BSC, CEA, CSCS and Juelich) have agreed on a framework to provide computing and storage services with common features. For example: cloud services based on the Openstack suite, or object storage for archival data repository services.

To achieve this goal the project partners have taken responsibility for a series of procurement procedures for R&D services and hardware. Cineca has taken the leading roles in the provisioning of two R&D services and one hardware system. The latter has led to the procurement and installation of Galileo100. This system started the operations in July 2021 and production in the following September. Given the goal of the project, Galileo100 was the first Cineca HPC system providing multiple services thorough a convergent (tightly integrated) infrastructure. In fact, 5 main services, each powered by different hardware infrastructure, can be identified:

- Scalable computing (SCC). Galileo100 is provided with 340 "thin" nodes to host small-to-mid size parallel jobs. Each node is equipped with: 2x CPU 8260 Intel Cascade Lake 24 cores, 384 GB RAM DDR4 2933MT/s and 480 GB SSD local disk

- Interactive computing (IAC). Galileo100 offers on top of the same configuration of SCC, 34 nodes with 2 NVIDIA V100 GPUs and 180 nodes with 3 TB of Intel persistent memory. This partition is particularly suited for use cases that require large memory servers and intensive I/O workloads, as well as GPU resources.

- Cloud services (Cloud). Galileo100 provides 77 computing servers configured as the SCC partition but with double the size of RAM memory (768 instead of 384 GB). To service the Cloud, the system features 12 OpenStack storage nodes providing a total of 720 TB of dedicated CEPH storage full flash (NVMe/SSD) for high IOPS. This storage will host VM root disks, disk volumes and VM snapshots.

- Active data repository (ACD). Galileo100 features over 20 PB of online storage with adequate performance for hosting simulation data

- Archival data repository (ARD). Galileo100 offers object storage with S3 API functionalities for long term data storing.

Lastly, Cineca took the lead to procure on behalf of the project partners an Interactive computing service. This will be based on the JupyterHub application suite adapted by E4 and co-designed for the specifics of the ICEI systems. The result of this R&D project will be installed on the Galileo100 system exploiting the IAC partition and ported on the future systems provided by Cineca.
Users Statistics
Elda Rebol, Massimiliano Guarneri
Cineca

On our HPC infrastructure, the active users at the end of 2021 are 4295, 468 more than the year before. A large part of the users are males (78%), 43% of them are aged between 31 and 45 years old while 35% are less than 35 years old.

73% of the total number of users works for an Italian institution, clustered in the Lombardia region (21%), Emilia Romagna region (18%) and Lazio and Friuli (12-15% each), while going to city level, Milan, Rome, Bologna and Trieste, in the order, have a user’s percentage going from 18% to 11%.

The more represented foreign countries are Germany with a 5% and United Kingdom, France, Spain and China with 3%, mainly working for universities and public or non-profit organizations.

Our computational resources are used by scientists within all scientific disciplines, the most represented three are Chemistry and Material Science (26%), Engineering and Computational Fluid Dynamics (15%), Condensed Matter Physics (13%).

Resource
Figure 5: classification of users' institutions
Figure 6: geographical distribution of international users
Figure 7: geographical distribution of Italian users
<table>
<thead>
<tr>
<th>Discipline</th>
<th>Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humanities and Social Sciences</td>
<td>72</td>
</tr>
<tr>
<td>AI and Machine Learning</td>
<td>55</td>
</tr>
<tr>
<td>Mathematics</td>
<td>79</td>
</tr>
<tr>
<td>Particle Physics</td>
<td>175</td>
</tr>
<tr>
<td>Life Science - Bio Informatics</td>
<td>208</td>
</tr>
<tr>
<td>Earth and Climate Sciences</td>
<td>251</td>
</tr>
<tr>
<td>Life Science - Computational Biology</td>
<td>268</td>
</tr>
<tr>
<td>Nuclear Fusion</td>
<td>394</td>
</tr>
<tr>
<td>Computational Engineering</td>
<td>421</td>
</tr>
<tr>
<td>Astrophysics and Plasma</td>
<td>428</td>
</tr>
<tr>
<td>Computational Chemistry</td>
<td>460</td>
</tr>
<tr>
<td>Condensed Matter</td>
<td>485</td>
</tr>
<tr>
<td>Computational Fluidodynamics</td>
<td>518</td>
</tr>
<tr>
<td>Other</td>
<td>522</td>
</tr>
</tbody>
</table>

Figure 8: distribution of users by disciplines
A very significant part of the allocations on Cineca’s system are based on a peer review selection in order to ensure the highest scientific value of the selected projects. In fact, Italian and European researchers take advantage of the HPC clusters in Cineca for performing their computational research, applying for “computational projects” which allow for access to the HPC facilities. PRACE and ISCRA programs allocate the majority of the resources on Marconi100 and Galileo100 and they are both based on the peer-review mechanism at the European and national level respectively. On the other hand, Institutions which want to use HPC infrastructures can sign an “Agreements” with Cineca as well as the “Industrial” users. Generally, a project is characterized by a starting date, an end date, a science domain, a budget (in terms of core-h/GPU hours) and a PI (Principal Investigator) with several collaborators. Along with these wide classes of projects, there are dedicated partitions on Marconi A3 and Galileo100 for weather and biogeochemical forecast activities and industrial projects while some special partitions of Marconi100 and Marconi A3 are dedicated to the EUROfusion consortium. For these partitions, the allocation is made autonomously by the dedicated committee. Considering the whole Cineca HPC infrastructure during 2022 542 million core-h have been made available out of a total allocation of 657 million core hours. This means that the systems have been overallocated at 121.2% - this is due to the huge request for computing resources by the scientific community. The most significant part of the resources in terms of core hours has been dedicated to agreements (63%), as well as 25% to PRACE, 25% to ISCRA and 2% to industrial projects.
PRACE
Partnership for Advanced Computing in Europe
Massimiliano Guarasi, Elda Rossi, Gabriella Scipione
Cineca

During 2021, two PRACE calls have been awarded on Marconi100: call 22 and call 23. For the regular calls a total of 23 projects have been allocated on Marconi100; 11 of those have been assigned to Italian PIs. According to the PRACE-2 agreement, as in the previous calls, 5% of the PRACE dedicated computational resources have been awarded to Center of Excellence (23 projects). The two calls distributed a total of 15,104,495 GPU hours (corresponding to 120,835,945 Local Core Hours or 1,329,195.000 PRACE Cumulative Core Hours), as reported in the following table.

<table>
<thead>
<tr>
<th>Accepted</th>
<th>Call 22</th>
<th>Call 23</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 April, 2021 to Thursday, 31 March, 2022</td>
<td>1 October, 2021 to Friday, 30 September, 2022</td>
</tr>
<tr>
<td>Regular call</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Reserved to CoE</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Total projects for call</td>
<td>24</td>
<td>22</td>
</tr>
<tr>
<td>Total accepted projects</td>
<td></td>
<td>46</td>
</tr>
<tr>
<td>Marconi 100 Node Hours / Local Core Hours</td>
<td>59,965.909</td>
<td>60,870.036</td>
</tr>
</tbody>
</table>
Iscra
Italian SuperComputing Resource Allocation
Paola Alberigo
Cineca

ISCRA (Italian Supercomputing Resource Allocation) gives computational resources and technical support to researchers affiliated to Italian institutions, in order to support the Italian research. It provides HPC resources through Class C projects (for code testing and pre-production) and Class B projects (full production). IscraC projects are evaluated and selected on a continuous basis, IscraB projects twice a year, on the basis of a scientific peer review procedure. In 2021 Cineca provided to the Italian research community 107M core-h on Galileo100 and Marconi100. In particular two IscraB calls were launched as reported in the table below (call 23 and 24). Since 2021 in addition to HPC resources also Cloud, DGX and Quantum Simulator resources are available through the ISCRA programme.

<table>
<thead>
<tr>
<th>Accepted</th>
<th>34</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rejected</td>
<td>36</td>
<td>31</td>
</tr>
<tr>
<td><strong>Total number submitted projects</strong></td>
<td><strong>141</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Total accepted projects</strong></td>
<td><strong>74</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Call 23B</th>
<th>Call 24B</th>
<th>Local Core Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGX</td>
<td>40.000</td>
<td>4.664.480</td>
</tr>
<tr>
<td>GALILEO100</td>
<td>2.104.998</td>
<td>31.151.760</td>
</tr>
<tr>
<td>MARCONI100</td>
<td>37.354.272</td>
<td>32.981.152</td>
</tr>
<tr>
<td>MARCONI A2</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supercomputing is an important aspect of nuclear fusion research as it plays a crucial role in the modelling of the plasma and materials, validating the experimental results of fusion devices and designing the next-generation fusion reactor.

The implementation of the EUROfusion Consortium programme has benefited in 2021 from dedicated high performance computing resources provided through the HPC Marconi-Fusion and the Gateway operated by Cineca in Bologna. As foreseen in the Project Implementing Agreement of Marconi-Fusion, the operation of the systems has been successfully monitored and supervised by the HPC Project Committee, with the support of the Operation Committee and the Ticket Committee. Availability of the resources (most important key performance indicator) was good and compliant with the Service Level Agreement for all partitions throughout the year with about \( \geq 21 \text{ Mnode-h} \) of usage for about \( \geq 24 \text{ Mnode-h} \) of availability of Marconi Fusion Skylake partition.

The EUROfusion computing projects running on HPC systems at Cineca are selected by the Allocation Committee on the basis of their relevant contributions to the following topics:
- Plasma turbulence and related transport processes
- Fast particle physics
- Non-Linear and/or Extended MHD
- Edge physics
- Heating and Current Drive
- Reactor Materials
- Reactor Technology.

The following table shows the computing time allocated in the six annual cycles since the HPC for Fusion services were provided at Cineca.

**Plasma Physics**

The plasma physics turbulence is still the topic requiring a large computing time as a key issue for future burning plasma experiments, such as ITER. Multiphysics HPC simulations are undergoing to Verification and Validation phases by means integrated plasma models for the Energetic Particles (EP) transport due to mesoscale Alfvénic instabilities primarily excited by EPs and EP coupling with microturbulence and macroscopic MHD modes mostly driven by thermal plasmas (Taimourzadeh, 2019). Initial V&V studies of the linear gyrokinetic simulations of Reversed Shear Alfvén Eigenmodes (RSAEs) in tokamaks are still carrying out using a gyrokinetic particle codes. Good agreement in RSAE frequency, growth rate, and mode structure have been obtained among these simulations, and between simulation results and experimental measurements. The successful linear V&V lends some degree of confidence to nonlinear gyrokinetic simulations that provide new kinetic insights on nonlinear Alfvén eigenmode dynamics and EP transport, and help the construction of reduced EP transport models which are needed for fast parameter...
scans, shot-to-shot analysis, and optimization of ITER experiments.

One of the major challenges in magnetic confinement thermonuclear fusion research concerns the confinement of the energetic particles (EPs) produced by fusion reactions and/or by additional heating systems. In such experiments, EPs can resonantly interact with the shear Alfvén waves. In the frame of the EUROfusion 2019–2020 Enabling Research project 'multi-scale energetic particle transport in fusion devices' (MET), a detailed benchmark activity has been undertaken among few of the state-of-the-art codes available to study the self-consistent interaction of an EP population with the shear Alfvén waves. In Vlad,2021 linear studies of EP driven modes with toroidal mode number n = 1 are presented, in real magnetic equilibria and in regimes of interest for the forthcoming generation devices (e.g. ITER, JT-60SA, DTT). The codes considered are HYMAGYC, developed in ENEA Frascati, MEGA, and ORB5, the first two being hybrid MHD-gyrokinetic codes (bulk plasma is represented by MHD equations, while the EP species is treated using the gyrokinetic formalism), the third being a global electromagnetic gyrokinetic code.

The following figure shows the results of the three codes for the two reference cases (peaked on-axis and off-axis EP density profiles): in both cases, for all the three codes considered, a mode driven by the EPs is observed. As an example, let's consider, first, the results for the peaked on-axis EP density profile (left column): after an initial transient phase of the simulations, both HYMAGYC and MEGA observe, as the most unstable one, a mode located around s ~ 0.4, with a dominant m = 2 poloidal Fourier component for the electrostatic field $\varphi$, and located around $\omega = -130$ kHz in frequency, i.e. with a frequency slightly lower (in absolute value) than the lower shear Alfvén continuous spectrum. These observations, together with the fact that its radial location is quite close to the minimum $q$ value, suggest that this mode is a so-called RSAE. On the other side, ORB5, for the same parameters, observes an externally localized TAE (Toroidal Alfvén Eigenmode) with frequency $\omega = -200$ kHz. This discrepancy between HYMAGYC and MEGA on one side, and ORB5 on the other, has been reconciled by considering the dependence of the EP driven modes on the EP density: indeed, all the three codes observe, for low EP density values, a dominant externally localized TAE; this mode, as the EP density is raised, becomes, eventually, sub dominant to a stronger RSAE. The occurrence of the ‘cross-over’ between RSAE and TAE depends strongly on the different damping experienced by the codes, which reflects in considerable variations of the net growth rates observed for this particular case (peaked on-axis EP density profile). The results for the peaked-off axis EP density profile (right column), on the contrary, show a remarkable agreement among the three codes, obtaining all of them a TAE localized close to the magnetic axis, where the EP density profile is characterized by positive radial gradient.
The HYMAGYC code, with a careful analysis of the frequency power spectra, is able to recognize also a sub-dominant, more external mode, characterized by a negative frequency $\omega \approx -200$ kHz, and located within the toroidal gap, at the radial position $s \approx 0.7$ as depicted in the following figure.

**Reactor Technology**

The computing resources allocated in reactor technology has grown on the last allocation cycles. The neutronics analyses performed with Monte Carlo codes are fundamental in nuclear reactor design. Detailed neutronics analyses have been performed with MCNP5 Monte Carlo code aimed at evaluating the DEMO divertor nuclear responses to provide inputs for thermo-hydraulic and thermo-mechanical analyses and to assess its shielding performances. A heterogeneous neutronics model of DEMO divertor has been developed and integrated in the DEMO MCNP model with a Water Cooled Lithium Lead (WCLL) blanket. The model is shown in the figure below. (Figure 12)

Neutron flux, nuclear heating, cumulated damage and helium production have been calculated in all divertor components (plasma facing components, cassette body, liner,
reflector plates and supports) as well as on the fixation systems, rails, in-vessel coil and vacuum vessel. Three-dimensional distributions of a) Irradiation damage (dpa/FPY) and b) helium production in divertor EUROFER are shown in the plots below. The results of these analyses provided important outcomes for the optimization of the design of the DEMO divertor.

**Integrated Tokamak Modelling**

ENEA and Cineca jointly provide to manage the Gateway infrastructure as platform for code development and data analysis for tokamak experiments by means an integrated modelling approach. Over the years, ITER has developed The Integrated Modelling and Analysis Suite (IMAS): a scientific software infrastructure which orchestrates the collective development and execution of integrated plasma codes and plasma applications describing fusion operations in tokamak experiments such as ITER. Currently, the IMAS suite is available on ITER HPC, EUROfusion Gateway HPC at Cineca and on the HPC infrastructures of some of the ITER collaborators (JET, etc...). Plug-in for open source data analysis and visualization are provided to interface IMAS in several models including plasma equilibrium and Scrape Off Layer in operative conditions of a tokamak as shown in the following figure. (Figure 13)

**Recent papers:**


Agreements

Maurizio Cremonesi
Cineca

Cineca computing platforms may be also accessed by agreements for collaborations. Several Universities and Research Institutions decided to begin or continue to support their researchers through agreements that reserve an amount of computing and storage resources without the need of submitting for selection a project proposal. These agreements grant access to resources from one to a few years and are particularly suited for training and test projects as well as experimental activities. The research and academic institutions with active agreements in 2021 are listed below. A total of 356 Mcore-h have been granted to agreement projects. The largest piece of allocation has been reserved by INFN with almost 144 Mcore-h, behind SISSA with 74 Mcore-h and ICTP that booked a total of 45 Mcore-h. Milano Bicocca among the other Universities reserved the largest slice with 38 Mcore-h. INAF also holds an important agreement with 27 Mcore-h in 2021. The institutions in the field of Life Science research all together sum up almost 4 Mcore-h grant. The other Universities and Polytechnics sum up a global 14 Mcore-h.

During year 2021 more than 20 agreements with Universities and Research institutions were active.

<table>
<thead>
<tr>
<th>Institution</th>
<th>Award resources (Mcore-h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFN (Istituto Nazionale Fisica Nucleare)</td>
<td>143.65</td>
</tr>
<tr>
<td>SISSA (Scuola Internazionale Superiore di Studi Avanzati)</td>
<td>73.73</td>
</tr>
<tr>
<td>ICTP (International Centre for Theoretical Physics)</td>
<td>45.42</td>
</tr>
<tr>
<td>Università degli Studi di Milano Bicocca</td>
<td>38.17</td>
</tr>
<tr>
<td>INAF (Istituto Nazionale di Astrofisica)</td>
<td>29.25</td>
</tr>
<tr>
<td>OGS (Istituto Nazionale di Oceanografia e Geofisica Sperimentale)</td>
<td>4.28</td>
</tr>
<tr>
<td>Regione Emilia-Romagna</td>
<td>3.33</td>
</tr>
<tr>
<td>Università degli Studi di Milano</td>
<td>3.21</td>
</tr>
<tr>
<td>Università degli Studi di Trieste</td>
<td>3.17</td>
</tr>
<tr>
<td>Università degli Studi dell’Aquila</td>
<td>2.50</td>
</tr>
<tr>
<td>Politecnico di Milano</td>
<td>2.45</td>
</tr>
<tr>
<td>Elixir (Distributed Infrastructure for Biological Data)</td>
<td>2.25</td>
</tr>
<tr>
<td>Università degli Studi di Genova</td>
<td>1.22</td>
</tr>
<tr>
<td>Istituto G. Ronzoni</td>
<td>1.19</td>
</tr>
<tr>
<td>Telethon</td>
<td>0.99</td>
</tr>
<tr>
<td>IIT (Istituto Italiano di Tecnologia)</td>
<td>0.92</td>
</tr>
<tr>
<td>Università di Bologna</td>
<td>0.84</td>
</tr>
<tr>
<td>OspPBG (Ospedale Pediatrico Bambini Gesù)</td>
<td>0.71</td>
</tr>
<tr>
<td>Università degli Studi di Brescia</td>
<td>0.49</td>
</tr>
<tr>
<td>Università degli Studi di Bergamo</td>
<td>0.26</td>
</tr>
<tr>
<td>INMI (Istituto Nazionale Malattie Infettive Spallanzani)</td>
<td>0.25</td>
</tr>
<tr>
<td>Politecnico di Torino</td>
<td>0.21</td>
</tr>
</tbody>
</table>
Training has always been a hallmark of the support activities carried on by Cineca for the Italian research community and users. Starting from 2012 it was extended to European researchers, with Cineca being recognised as a PRACE Training Center (PTC) in HPC.

In 2020 the pandemic emergency came so unexpectedly so it took some time to convert live events in online events.

In the pre-pandemic times, the courses were spread over three Cineca sites (Bologna, Milan, Rome) so for almost each courses we had three editions.

In 2021 teachers had more time to prepare the teaching material in the online mode so we were able to increased the number of students accepted. Furthermore we increased the number of students accepted, opening the course to the largest number of people admitting some “official students” with full access to theory and practicals and the rest in the waiting list as Auditors allowing them to access to the whole course excluding just from the support during practicals.

Furthermore that way of teaching allowed easily the recording of the sessions. They were published on the courses web site, building a rich library of teaching videos, many of them held in English language, making them available for the people worldwide.

The forcing in this new way of teaching increased our competence in terms of organization and in using learning platforms.

Students expressed in many fashions the fact that we fully reached our goals in converting process and organization.

Many comments in the surveys expressed their likelihood in following the course being at home, not having to plan their travels, with no expenses and with the possibilities to join the class at the very last minute.

During 2021 we reached these numbers: 16 online courses, and 4 schools. More than 50 people in the department were involved in teaching, for a total of 84 days of lessons and highly specialised training. In total, more than 725 Italian and European researchers took advantage of our training program, with high satisfaction levels (8,8/10), as shown by the surveys we regularly administer.

In addition to the Cineca HPC Department classes, our experts were involved in partnership with external societies or universities, giving their support in organization or teaching in events like Masters and Doctorate schools, workshops, tutorials about HPC and Big data topics.

Other important initiatives: Summer of HPC (a PRACE initiative offering young students from all over Europe to work with HPC teams); European Researchers Night (Cineca is the coordinator of Society, a European project organising the Night in Bologna and surroundings in 2018-2019); HPC-Europa3 (we are the coordinator of the project offering transnational access to researchers visiting HPC centers).

The Training activity is regulated by two specific procedures that refers to the ISO standard: ISO 27001:2013 “Information Technology - Security techniques - Information security management system”.
725 Students from all over Europe
84 Days of lectures
50 Teachers
16 Classes
6 Collaborations in events
4 Schools of Advanced Studies
2022 Plan

Due to the pandemic, the training events in Cineca until December 2021 were all scheduled ONLINE since, even with controlled access, the events are assumed to be very risky.

In these first days of 2022, due to the large vaccination program, the curve of pandemic is rapidly decreasing putting some hope for getting back to the in classroom events.

These in person events have the big advantage to allow people to connect and possibly to establish new job-connection and new collaborative projects.

On the other hand, we notice a strong increase of the number of students interested to follow courses, maybe for the big advantage of the possibility to follow the lessons remotely without arrange their travel and accomodations.

In agreement with the teaching staff, in 2022 at least some courses will back in the classrooms, but thanks also to the big experience made during these two years, we will go on planning ONLINE events and some hybrid events for giving people the best of both worlds.

Our effort is constantly focused in increasing the number of courses and schools developed by Cineca HPC Department, teaching the most new technologies available and trying to admit the largest number of students worldwide is possible. For doing this we will try to admit as much students as possible, giving them access to tutors support and, for each course, a reasonable numbers of auditors giving them access at least to an email support service. For some courses we will open online forums so the students will be able to discuss and support each others with the tutoring of our experts. We will try to record and put online large part of the lessons, tentatively in English language, making them feasible.

Furthermore we are experimenting new ways of interaction between people remotely connected like closed rooms of discussions with few students.

Figure 14: numbers in training
STAFF

In the HPC department of Cineca work about 98 people, distributed among the three premises: 77 in Bologna, 7 in Milan and 14 in Rome.

Several new colleagues were enrolled in 2021 and 2022: we wish them a very fruitful and interesting job.

Andrea Giunchi
I completed my bachelor’s, master’s and Ph.D. degrees in Industrial Chemistry at University of Bologna. I worked in the field of Computational Chemistry with particular emphasis on periodic DFT calculations based on plane waves applied to organic crystal systems. In Cineca I work as HPC User Support for ENI, installing software and assisting users on ENI clusters.

Laura Bellantani
I obtained my PhD in Physics and Nanosciences at the University of Modena and Reggio Emilia. My postgraduate and postdoc research involved the theoretical study and numerical simulation of quantum logic gates in condensed-matter systems. I am working in the HPC Department of Cineca on the porting and optimization of codes, mainly focused on Material Science, and on the support of profiling tools.

Davide Giosue Lippolisi
I got my master’s degree in Theoretical Physics at the University of Bologna, Alma Mater Studiorum with a thesis in stochastic modeling and statistical analysis of cytokines metabolic cycles. He researched in medical Image analysis at Ospedale S.M.N of Reggio Emilia. In Cineca he’s a desktop and web developer for scientific applications. Furthermore, he manages DevOps operations on Kubernetes.

Fabio Di Sante
I got my PhD in “Earth Science and Fluid Mechanics” at the University of Trieste. I have nine years of experience in working with hydrological and climatological models. At Cineca I work in the Data Management group as support to weather forecasting chains and as high-level user support to weather and climatological applications, models and tools.

Giorgia Frumenzio
I got my master in “Molecular Biology” at the University of Parma. In my research, I studied the dynamics of several proteins from humans and SARS-COV2 by using computational techniques like Molecular Dynamics. In Cineca, I work as HPC specialist providing support for the optimization of scientific applications.

Lucia Rodríguez Muñoz
I earned a Bachelor’s Degree in Physics, and a Master’s Degree and a PhD in Astrophysics at the Universidad Complutense de Madrid (Spain). Then, I worked for 5 years as a researcher at the Università degli Studi di Padova. The purpose of my research was to improve our knowledge on galaxy evolution. To this aim, I analyzed multiwavelength data obtained with some of the most powerful telescopes in the world to study the properties of distant galaxies, and subsequently, derive their most likely evolutionary paths. Currently, I work as an HPC Data Engineer in the HPC Data Management group at Cineca.

Elisa Alberti
Elisa Alberti graduated in Political Sciences from the University of Bologna (Italy) in 2004 and achieved a specialisation in European Funds from the University of Padova (Italy) in 2010. She is a Project Manager with a decade experience in European Funded projects, in proposal writing and in financial and administrative management, under different EU funded programmes, such as FP7, H2020, LIFE, H2020.

She is currently member of PMO (Project Management Office) of the Cineca’s HPC Department.

Francesco Finelli
I got a master’s degree in Physics at the University of Rome “La Sapienza” (theoretical Physics, with a focus on classical dynamics and plasmas) and a Ph.D. in Physics at the University of Pisa. My research focused on numerical simulation of magnetic reconnection in space plasma, both to evaluate models and to generate data to train and test algorithms for automatic detection of reconnection in turbulence. In Cineca I work in HPC User Support.

Maria Montagna
I got a master degree in Physics at La Sapienza University of Rome, and a PhD in Engineering and Physical-Mathematical Modelling at the University of L’Aquila. The main interests of my research activity were focused on material science and biophysics (Organic/Inorganic Interactions in Biolinspired Systems, Catalytic Complexes, Hydration Processes) using classical and ab-initio molecular dynamics simulations and theoretical chemistry calculations. In Cineca I am part of HPC User Support working on setting and managing of HPC production environment.

Federico Tesser
I got my PhD at the Research Center INRIA Bordeaux - South West, resolving in parallel the Poisson equation on some hierarchical and superimposed grids, in Python. After that, I came back to Italy, where my occupation for almost 3 years has been to be a technologist at the Politecnico di Torino, working on inter/intra node optimization and performance engineering.
Events

The European Researchers’ Night of Society project
Simona Caraccioli
Cineca

In 2021, 30 pre-events (24 onsite or in mixed mode and 6 fully online) were organized in the month preceding the ERN. The relatively good Covid situation in Italy allowed us to organize the majority of pre-events onsite, although the participants were limited in number by health safety rules. Most pre-events and the main ERN event were included in the summer event program promoted by the Municipality of Bologna “Bologna Estate”. This gave good visibility to the project and contributed positively to the awareness campaign.

Researchers from the Society consortium made themselves available to give face-to-face or online lectures in the schools. The rich program of available lessons (a total of 74 for all school levels) was advertised directly to teachers and also through the channels of the Regional School Office of the Ministry of Education. Schools in Italy start mid-September and this leaves too little time to plan lessons before ERN. Furthermore, this year, schools had to face quite challenging problems for returning students to attendance after last year’s remote teaching period. For these reasons most lessons have been scheduled to take place in the period after ERN, preferably by the end of October, but even beyond.

SOCIETY Rinascimento organized activities and events for the European Researchers’ Night in 2021 in seven towns in the Emilia Romagna Region, Italy: Bologna, Ravenna, Forlì, Cesena, Rimini, Predappio, and, for the first time, Ferrara. Given the difficulty to predict the evolution of the pandemic and of the corresponding restrictions, the consortium decided from the very beginning to opt for a dual approach, organizing activities and events both online and in the streets.

The SOCIETY Rinascimento project organised a mixed online/onsite event, featuring 37.5% of activities in presence and 62.5% online. A total number of 384 researchers were involved, 10 of whom benefited from MSCA funds (Table 1), while 91 researchers were supported by other EU schemes.

For Bologna, the Night in person was organized downtown, along via Zamboni, at the heart of the academic district. The road was closed with identifiable entry points at both ends, and a single exit, to allow an ordered visiting experience. Access was only granted to holders of the EU digital Covid certificate released upon vaccination against COVID-19.

Stands were organized underneath the portico’s arcades that recently became part of the UNESCO heritage. Each stand featured a desk for materials on display and a monitor. The total number of stands (33) is smaller than in previous editions. This was necessary to ensure sufficient space around each activity and also because some of the activities were proposed online.

Online activities were available on the virtual platform of the SOCIETY Rinascimento project, accessible through the website http://nottediericercatori-society.eu/. Contents of the virtual stands will remain available to visitors for one year, until September 2022.

The platform hosted the live streaming of the inauguration ceremony, thanks to the support of our stakeholder Lepida TV. The possibility of broadcasting further live events held in presence was discussed among the partners but eventually discarded, to focus the visitors’ attention to the many virtual stands held by the researchers.

We estimated that approximately 4700 people attended all events. Considering the pandemic limitations, we registered a good onsite attendance in Bologna (approximately 2000 participants), Rimini (354), Ferrara (330), Ravenna (231), forlì (144). Online activities were followed through the sharewert platform (723 registered users) or through social media channels such as youtube and Fb (965). The web page where the scientific magazine Sapere was made available for download had more than 800 accesses, both during the Night (183) and afterwards.
Figures 15, 16: European Researchers’ Night in Bologna

Figures 17, 18, 19: European Researchers’ Night in Rome
Events

The European Researchers’ Night of Net project

Cineca

The NET-sciNeE Together project organized its second edition of the European Researchers’ Night, an initiative promoted by the European Commission since 2005 that involves thousands of researchers and research institutions every year in all European countries.

The topic was the challenges given by climate change. The goal was to create opportunities for researchers and citizens to meet and spread scientific culture and knowledge of the research professions in an informal and stimulating context. NET is a partnership of all the major local (Lazio) Research Institutes and Universities, such as CNR (coordinator of the project), Cineca, Ispra, INAF, INGV, ENEA, University of Rome Sapienza, University Tor Vergata, University of Tuscia and Telematic University UniNettuno. Cineca in this 2021 edition coordinated the Impact Assessment. The impact assessment campaign was designed to contribute towards achieving the objectives of assessing the impact of the NIGHT event, using qualitative and quantitative parameters, and of identifying what kinds of activities are most successful in attaining the European Researchers’ Night main aims.

The project had an ambitious program of a total of 70 pre-events with the participation of more than 2,800 people to promote theResearchers Night.

Apart from Rome, where most of the activities took place, the pre-events were in many cities involved in NET project such as Livorno, Vulcano, La Spezia, Viterbo, Cerveteri, Tarquinia, Chioggia, Ozzano, Brasimone.

The pre-events were conducted in the most diverse contexts to bring researchers into closer contact with citizens, outside the traditional context of scientific research, promoting the interaction.

NET scientific treks achieved great success with 36 events organized in many locations throughout the national territory: naturalistic treks or in urban parks, urban walks visiting churches and finding links between art and astronomy, trekking on the volcanoes of the Aeolian islands, were an opportunity to discuss scientific disciplines such as botany, biology, geology, astronomy, archaeology and more.

NET 2021 promoted the 7 “scientific aperitif” format bringing researchers in very informal contexts in the late afternoon, in close contact with an audience of 40-50 people. The locations were associations – Casetta Rossa and Fusolab 2.0 - in two districts of Rome, one in the centre and the other in a suburban area, both frequented by young people who engaged in social activities.

Another important pre-event was “Talenti per la scienza” at the Talenti Park Arena, lasting 2 days, in which hands-on activities for children, demonstration stands, scientific talks and show conferences were organized, with the presence of researchers of almost all NET partners.

NET 2021’s most prestigious pre-event fea-
tured the 2019 Nobel Prize in Physics Michel Mayor. Professor Mayor was a project guest for three days in Rome at Sapienza University of Rome.

The presence of Prof. Mayor was an opportunity to explore the themes of space research and extraterrestrial life, considering that Mayor discovered the first exoplanet orbiting a solar-type star in our galaxy, the Milky Way.

The activities carried out for the NET 2021 Night included a program dedicated to students of Italian primary and high schools mainly from the Lazio region with the participation of various institutes from other Italian regions. The activities for schools were launched with an initiative called "Pillole di scienza-Pills of Science (Did you know that ...)", short videos dedicated to a specific scientific theme. The videos were produced by each Partner on some themes which were then discussed in subsequent lessons. Many activities were converted into online activities.

The in person and online activities for the schools took place mainly in the period from 20 September to 25 September 2021, during the week of the Night with an extension until November 2021. The activities dedicated to schools can be summarized in 108 events. 53 events over 108 were carried out in presence (32 at partner sites, research centres and event venues of the Researchers’ Night and 21 directly at the School sites). The remaining 55 events were online events (39 live events with the researcher connected to the classes and 16 on demand, always available on the internet).

The NET Consortium organized the final event principally in the centre of Rome (Testaccio) inside the “The City of the Other Economy - CAE” which is one of the first spaces in Europe entirely dedicated to the promotion of the other economy. CAE is a very big outdoor place where thousands of people could be hosted even according to the strict anti-covid rules that were forced in Italy in September 2021. The final event was organised for two nights 24 and 25 September (Friday and Saturday) from 18:00 to 24:00. In addition to this big event, those partners who could host the general public at their sites arranged also other events. Moreover, some online activities have been done, such as updated virtual tours for almost all institutions of the consortium to open virtually their scientific laboratories that were not accessible during the pandemic. It was a way to visit the environment of laboratories in which the researchers carry out their research daily and to permit them to discover, even from home, the tools and technologies used.

Researchers involved in the final event inside CAE welcomed about 4000 visitors with a wide offer of activities and about 30 stands. Many activities were designed for the families and general audience offering hands-on activities, games, science demonstrations, informal conversations with researchers and science entertainment. The participation in the other cities was about 1500 attendees. More than half of the visitors attended for the first time events related to the European researchers Night. The events had a great impact on the participants, who declared, overall (more than 97% in all questionnaires), to be interested in attending similar events. These numbers are significant to describe the satisfaction and the effectiveness of the NET events.
Events

Summer of HPC 2021
Maddalena Guarreal
Cineca

Also in 2021, Cineca participated to PRACE Summer of HPC program. Summer of HPC (SoHPC) is an initiative funded by the European Union through the PRACE project, which aims to allow University students from all over Europe to have a study and work experience abroad at major European computer centres. The selected students work under the supervision of a mentor (generally a researcher) and a company tutor (coming from the computer centre that will host them), on projects related to High Performance Computing (HPC) and scientific visualization using calculation resources provided by PRACE.

PRACE Summer of HPC aims to increase the interest of new generations in HPC and related technologies by disseminating the experiences of program participants via social media. Since the pandemic affect also the possibility to go abroad on 2021, the internship had to take place remotely. Cineca hosted four students working on two separate projects. Particularly David Mulero Perez and Sepideh Shamsizadeh worked under the supervision of Prof. Andrea Bartolini of the University of Bologna on a project titled “Visualisation and Anomaly Detection of A Supercomputer”.

The other project was coordinated by Prof. Silvia Ceramicola of National Institute of Oceanography and Experimental Geophysics (OGS). The two students that work on this project, Mario Gaimann and Raska Soemantroro won also the 2021 SoHPC Best Performance Award, with the project “Automated Classification for Mapping submarine structures by Artificial Intelligence strategies”. The award ceremony took place during the EuroHPC Summit Week 2022 / PRACEdays22 on March 24th, 2022 in Paris. The two students worked in collaboration with OGS, NOC and Cineca, as well as the DPC-MaGIC Project.

The reasons for the award were: “Mario and Raska worked on a particularly challenging project and on their own developed an automated tool to recognize seabed structures. They worked very hard on this project, researching the various AI methods available to them and deciding on which methods were best to use. Furthermore, through their own initiative they researched, identified and learned various technologies and tools to use for their project. The outcomes of their work far exceeded the initial expectations of the project. Indeed, they will aim to publish their results through funding via the HPC-Europa3 scheme.”

Figure 20: the winners of the award
HPC H-index

Source: Web of Science Database: Web of Science Core Collection
Search: (FT=Cineca or FO=Cineca) and (PY=2021-2022)
Date: 03.11.2022

Results found:
1028

Number of Citations:
3626

Average Citations per Item:
3.53

H-index:
20
Figures 21, 22: number of publications and citations from year 2008 to 2022 (2022 only January-October)

Figure 23: research areas of the publications
Figure 24: main affiliations of the authors

Figure 25: journals in which Cineca users published most frequently

Figure 26: nationalities of the organizations of authors
SCIENTIFIC OVERVIEW
Cineca is involved in a wide scientific field thanks to EU projects and collaborations. From artificial intelligence to urgent computing, bioinformatics and digital humanities, HPC is used to support them all.
LIFE SCIENCES
ORCHESTRA
Connecting European Cohorts to Increase Common and Effective Response to SARS-CoV-2 Pandemic
Gabriella Scipione, Chiara Della Casa, Elisa Rossi
Cineca

ORCHESTRA is a three-year international research project aimed at tackling the coronavirus pandemic, led by University of Verona and involving 26 partners (which extends to 37 partners when considering the wider network) from 15 countries. The project, started in December 2020 with a budget close to 20 million euro, is funded by the European Union's Horizon 2020 research and innovation programme. The ORCHESTRA project aims to respond to this need through the creation of a new pan-European cohort built on existing and new large-scale population cohorts in European and non-European countries.

37 partners
From 15 European non-EU countries

Objectives

- to develop evidence-based recommendations for effective prevention, protection, and optimized treatment of COVID-19 patients (including long-term consequences) with a special focus on ‘at risk’ population, including healthcare workers and fragile individuals
- to assess impact of environmental factors, socio-economic determinants, lifestyle and confinement measures on the spread of COVID-19
- to provide knowledge on the efficacy of vaccines against SARS-CoV-2
- to provide a model for responsiveness for future pandemic outbreaks.

In this context, the integration and analysis of the very heterogeneous characteristics of SARS-CoV-2 health data coming from many different sources such as electronic health records, retrospective and prospective patient registries, and related ‘omics’ data, are key drivers for the evolution from evidence-based medicine towards precision medicine and precision public health.

Cineca in the project is the coordinator of the WP7 “Data Management” work package, in charge of the development of the platform that will collect data from the various cohorts involved. Together with other leading HPC European Centers (HLRS - Germany and CINES - France), Cineca will be responsible for the design and deployment of a federated architecture based on three layers: National Data Providers, National Hubs (NH) and the centralized Orchestra Portal.

WP 7: Data Management
Leader: CINECA
Co-Leader: CINES / USTUTT

An extremely challenging task within the ORCHESTRA project is the collection of data from the different European and international countries. This is the central task of Work Package and is tied to the various legal aspects of the individual sites involved. Data is to be collected centrally and decentered with the help of supercomputing and without the possibility of drawing conclusions about sensitive information.

Figure 27: ORCHESTRA partners map
Starting from the definition of the general architecture, CINECA deployed a first solution to manage sensitive personal data in the Italian National Hub (NH) in compliance with the GDPR and the Italian regulations.

Data are ingested from the national data providers, mainly from Universities and Hospitals located in Italy, using a dedicated instance of REDCap provided in the NH infrastructure.

The process includes managing and storing pseudonymized data in the National Hub (NH) based on a cloud infrastructure with typical repository functions. Before opening the NH to real world data collection, some actions have been carried out at Cineca:

- signature of the “appointment of data processor”, a data sharing agreement between each data controller (National Data Provider) and data processor (CINECA),
- adoption of all technical and organisational measures to minimize the risk,
- implementation of a Vulnerability Assessment and Penetration Testing (VAPT)
- execution of a detailed internal audit, with internal DPO officers,
- Preparation of a document to support Data Protection Impact Assessment (DPIA).

Additionally, for all the use cases in which standard data sharing procedures are not applicable by definition or due to any kind of legal restrictions (e.g. informed consent to share data for research purposes not given for retrospective cohorts), alternatives such as Federated Learning are explored.

Considering the European context, in Europe there is an increasing demand for regulated data usage in HPC data centres with respect to GDPR regulations. The HPC cloud infrastructure released for the Orchestra project can act as a model to manage future public health threats.

Figure 28: ORCHESTRA objectives
Starting from 1987, ARNO Observatory routinely collects and integrates NHS administrative data for each single patient, to support governance activities in the healthcare (expenditure control and appropriateness of drug prescription) of LHUs and to study Patient Care Pathways.

ARNO Observatory allows to highlight cohorts of population with chronic disease that need to be followed, evaluated, and studied. Diabetes is expected to become one of the world’s most common diseases within the next twenty-five years with high impact on National Health System resources, it is a condition with high prevalence of comorbidities, and which involves high quantity of resources. The aim of ARNO Diabetes Observatory, in collaboration with SID (Società Italiana di Diabetologia), is to assess the impact of resources of patients with diabetes, compared to people without diabetes, in order to analyze prevalence and incidence of this disease and evaluate the burden of disease in terms of impact of new drugs, comorbidity, complications and costs. ARNO Diabetes has become one of the reference sources for the scientific community and periodically Cineca publishes reports to describe the patient profile and use of healthcare resources (Figure 1).

In 2021, new papers related to initial treatment and to incidence of diabetes have been published, respectively to:
1) investigate diabetes treatment initiation and continuation in the next sixth months in newly diagnosed Italian subjects
2) assess incidence of diabetes in Italy from several million residents, highlighting differences in rates in men and women across decades of age and Incident rates of insulin or noninsulin treated subjects

Figures 29-33 reports on ARNO Diabetes

EARTH AND ENVIRONMENTAL SCIENCES
Climate change has an increasingly important impact on livestock farms, influencing production, health and fertility of animals. At the same time, farms are sources of greenhouse gases, at the root of global warming. It is therefore necessary to develop strategies that help the entire livestock sector to make the best decisions to adapt to these changes and mitigate their effects.

SEBASTIEN is a new project co-funded by the Connecting Europe Facility of the European Union, and it aims to deliver a Decision Support System (DSS) in the form of a package of services hosted in an open and interoperable web Portal. The main goal of the project is to increase the efficiency and sustainability of the Italian livestock system and support breeders and operators in the animal production chain with tools that can help them make targeted and informed choices, but also avoid economic losses and provide technical support to companies.

SEBASTIEN started at the beginning of 2022 and is coordinated by CMCC Foundation (https://www.cmcc.it/); it relies on Cineca and CMCC infrastructure for monitoring data exploitation and on climate simulations already produced by the Highlander project coordinated by Cineca. The applications developed by SEBASTIEN will use large amounts of public data, of different nature, harmonized and analysed using the most modern analysis techniques such as artificial intelligence methods.

The SEBASTIEN project will develop sentinel indicators by integrating environmental, meteorological, reproductive and production data:
- In situ sensors that allow measuring parameters related to animal body conditions.
- Historical and future climate simulations at unprecedented spatial details over Italy, to be combined into indicators of environmental comfort/wellbeing conditions.
- Satellite images integrated for multi-dimensional analyses of feed availability.

Four free applications will be developed which will allow:
1. estimating the adaptation of different livestock species and breeds to changing environmental conditions
2. providing a forecast system that allows farmers to take actions to fight dangerous environmental conditions
3. guiding farmers in the choice of the best pastures
4. monitoring the presence and development of parasites and pathogens.

The indicators will be tailored to different categories of end users, to allow breeders, companies, government bodies and other operators in the agri-food-livestock sector to be able to face their daily needs and work problems quickly and efficiently.
Digital technology enablement for Destination Earth

Technology Partnership Lead for Destination Earth
Compute Department
Thomas Geenen
ECMWF

ECMWF (the European Centre for Medium-Range Weather Forecasts) is one of three entrusted entities tasked with delivering Destination Earth (DestinE). DestinE is an ambitious initiative of the European Commission to develop a highly accurate digital twin, or replica, of Earth. This will allow users of all levels to better explore natural and human activity and to develop and test a range of scenarios and potential mitigation strategies. For example, it will better enable policy and decision makers to anticipate and mitigate the effects of extreme weather events and climate change; potentially saving lives and alleviating economic consequences.

Bringing together scientific and industrial excellence from across Europe, DestinE will contribute to revolutionising the European capability to monitor and predict our changing planet, complementing existing national and European efforts such as those provided by the national meteorological services and the Copernicus Services. Destination Earth will also support the European Commission’s Digital Strategy and the Green Deal priority actions on climate change, biodiversity and deforestation.

High-performance computing (HPC) and big data handling have been a key enabler for weather and climate prediction for decades. The recent evolution of computing architectures has stimulated a wide range of programmes that aim to prepare the operational weather prediction infrastructures for future technologies. These programmes increasingly realise that it will take more than progress in HPC to fulfil future extremes prediction and in particular climate adaptation targets as addressed by DestinE. This is where the need to exploit the opportunities offered by the entire range of digital technologies (Figure 1) comes into play, and where new ways of co-developing Earth system and computational science, and co-developing DestinE’s digital ecosystem need to be found.

Cineca has been selected to support ECMWF on this journey, leveraging decades of experience in HPC and code optimization and as one of the leading HPC centres in Europe, deploying the first generation pre-exascale systems in Europe. Code optimization and cloud experts from Cineca will work closely with ECMWF developers and join the workstreams tasked to deliver technical enablers that builds the evolving architectural runway supporting the digital twin deployments. Since the ultimate ambition of DestinE is to run digital twins across all EuroHPC systems, a large focus will be in portability of applications and services to all (pre) exascale systems. We will also leverage the extended experience of Cineca in community management and training to onboard the DestinE community and associated consortia on the EuroHPC ecosystem.
Figure 34: DestinEs extreme-scale capabilities critically depend on the seamless interplay with many components of the digital continuum. Cineca is well positioned to both work on the capabilities side, scaling simulations and models to exascale system size, and on several components of the digital continuum.
BIG DATA
AND
ARTIFICIAL INTELLIGENCE
AI models engineered on Cineca HPC infrastructure are a core element of the services that Gmatics developed as part of the DAMAS – Data-driven Model for the Analysis of Sea-state experiment funded by the EUHubs4Data project (https://euhubs4data.eu/experiments/damas/). The experiment successfully reduced the execution time and increased the resolution and the accuracy of sea circulation models and height of waves prediction, providing as a result:

- AI models that replicate the ENEA MITO and WAVE physical forecast models and produce forecasts over the next 5 days with hourly temporal resolution,
- The nowcast chains for wave and circulation for the next 12 hours,
- Hindcast pipelines for the historical analysis of wave and circulation sea parameters and
- A graphical user environment 122,586 core hours have been consumed and 44 TB of data have been processed in order to reduce the AI model runtime for the whole Mediterranean Sea to 15 minutes (the experiment final target was 30 minutes), increase the geospatial resolution to 1x1 Km, achieve a point accuracy with a Root Mean Square Error between DAMAS and ENEA results below 5% for each variable (total peak and mean period, total mean wave direction, sea surface wave significant height, sea surface height, surface salinity, surface temperature, surface wind u-component and v-component speed) and increase the areal accuracy above 90% for each variable.

These results were achieved during the 9 months duration of the experiment by first matching the performance of the ENEA model and then improving the geospatial and temporal resolution, which required the fine tuning of the AI model using satellite data, ensembling techniques and transfer learning techniques.

DAMAS was funded by the EUHubs4Data project (https://euhubs4data.eu/) that promotes the collaboration between Data Driven Innovation Hubs in EU through the

Figure 35: DAMAS dashboard showing forecasts on two variables (wave direction and wave height) at a specific location of the Mediterranean sea
definition of a common catalogue of datasets, data services and training services. Towards this goal, SMEs are invited to propose experiments that make use of the EUHubs4Data federated catalogue for the development of innovative products or advanced services and are granted access to funding through the implementation of open calls. Cineca is one of the participating Innovation Hubs providing services to the SMEs and oversees the training programme and the project Data Management Plan.

A Federation of Data Driven Innovation Hubs is being established in order to be operational during the project and beyond as the European reference for data driven innovation and experimentation, fostering collaboration between data driven initiatives in Europe, federating solutions in a global common catalogue of data services, and sharing data in a cross-border and cross-sector way.
AI Day—ahead hail prediction to prevent losses in agriculture
Bhaskar Agarwal, Michele Bottazzi, Giorgio Pedrazzi, Gian Franco Marras, Roberta Turra
Cineca

Every year, almost a third of all food produced is lost or discarded, resulting in a 940 billion economic loss. This comes at a time when it is estimated that 50 percent more food will be required over the next 20-30 years. The increased prevalence of extreme weather events due to climate change has heightened the need to build support decision systems that can assist farmers in mitigating agricultural losses.

Precision Agriculture is a technology-enabled, data-driven approach to farming management that watches, measures, and analyses the demands of fields and crops to help optimize agricultural production while reducing waste and expenses. The EU Cybele project (www.cybele-project.eu) aims at generating innovation and creating value in the domains of agri-food by implementing such Precision Agriculture demonstrators.

Our task within the project was to develop a hail prediction model. Hail is a rare phenomenon, very difficult to predict. We treated the problem as a binary classification to predict whether it will hail or not at a given hour and location, a day in advance. The focus was on the Valencia region, and an enclosing area of around 200,000 square kilometers in Spain was selected for training the model. We used the weather forecast and the derived climate instability indices as input, amounting to a total of 6 TB. The location and timestamp of 100 hail events in Spain was available as truth data for the period of 2015-2019. Owing to the imbalance in the problem, we decided to spread a hail event both temporally (an hour in the past and the future) and spatially (a radius of up to 10 kms). This led to a more physical and complete target data set. We employed 3 different gradient boosting algorithms to create an ensemble model and worked with two KPIs. While we optimised the model based on the F1-score, we assessed the Recall and 'Corrected Recall' of the test set. The Corrected Recall, a completely new KPI, was designed for the project based on synthetic target data creation. After assessing the needs of the stakeholders, it was clear that the hail events most critical to them were the summer months. Thus, we trained and optimised the model in these months from 2015-2017 and tested on the same months in 2018-2019.

Our model returned a lower than 20 % false positive rate, and a 50-100% Corrected Recall of hail events. The final model was then ingested in containerised demonstrator designed for direct use by the farmers.
Figure 36: model diagram CYBELE project
MaX: HPC and HTC meet for Material Sciences

Materials at Exascale
Mariella Ippolito, Fabio Affinito
Cineca

This year will see the conclusion of the second MaX project funded under the Horizon2020 programme of the European Commission. During the last years, thanks to MaX, Cineca took part in an activity aimed at making the codes for electronic structure calculations ready for the supercomputers of next generation, including the incoming Leonardo machine. In this framework, Cineca was particularly involved in two different actions. The first one was about the porting and the optimization of the Quantum ESPRESSO code on the GPUs accelerators. In collaboration with the SISSA team (where the code was born), we actively participated to transform the code so that it could run on accelerated systems, like Marconi100. This allows the material scientists to exploit at the best the most modern computational architectures, getting advantage from the computing power offered by the NVIDIA accelerators. The second action, which has seen CINECA as leader, is the coordination of the user support in the MaX project. This role permitted to help the community of users to fix problems related to execution of complex applications on HPC systems. The role of the user support for material science applications will be furtherly enhanced in the next years, also beyond the MaX project, with the implementation of the application support in the High-Level Support Team of the Leonardo system. Looking ahead at the next future, the availability of large-scale infrastructures, such as Leonardo, will open new possibilities for many different scientific domains, including material sciences. As a matter of fact, we expect that besides a single application that is able to run on a very large number of computing nodes, we will see a different model of exploitation, where a multiplicity of application running on few nodes will interact together for the resolution of a physical problem. In other words, we expect that HPC, High-Performance Computing (a few applications with large scale-out), will be tightly working with HTC, High-Throughput computing (a large number of applications with small scale-out). The management of complex workflows is really challenging from a technical point of view, because it requires a high robustness of the computing infrastructure including an efficient management scheduling system. The experience and the expertise gathered in these years with MaX will permit us to get ready to prepare the new challenges posed by the advent of Leonardo and to provide the material scientists with new complex tools to discover the materials of our future.

MAX has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 824143.
MAX offers a materials design ecosystem for running high-throughput calculations with automated data storage in graph databases.

**AiiDA**

as the workflow engine

- 61 registered plugin packages
- ~110 workflows
- 21% of calculations and workflows plugins supported directly by MAX

Results obtained with AiiDA workflows can be easily shared thanks to its close integration with the Materials Cloud, where data can be uploaded, visualized and analyzed.

www.aiida.net

**MATERIALS CLOUD**

as the web dissemination platform

Materials Cloud is recommended as a data repository for Materials Science by Nature’s Scientific Data and by the EU Commission’s Open Research Europe journals:

www.materialscloud.org

- 22,145,443 crystal structures
- 7,521,504 structures with associated DFT SCF calculations
- 2,965,252 data and calculation nodes with full AiiDA provenance
- 674,778 reproducible calculations with full AiiDA provenance

**AiiDALAB**

as the cloud simulation environment

AiiDALAB provides the user with a dedicated and intuitive simulation environment directly in the cloud or on remote or local resources. Here the user can run and manage AiiDA workflows with tailored lightweight web applications, all from the browser:

www.materialscloud.org/work/aiidalab

31 May 2021

WWW.MAX-CENTRE.EU
QUANTUM COMPUTING
Without a doubt, the concrete realization of the first quantum computers represents one of the most important scientific events of this new millennium. Theorized at the beginning of the 1980s, for over 40 years scientists all over the world have been studying how to make devices capable of exploiting the laws of quantum mechanics to increase our computing power. After 20 years of studies, it was finally possible to create some small prototypes of a functioning quantum computer in the laboratory: for the real breakthrough, however, it is necessary to wait another twenty years, to arrive at the present day. Today quantum computers are gradually approaching the world of supercomputing. Although far from the science fiction models of universal quantum computers theorized in the 1980s, the computers we have today are beginning to be suitable to be used effectively as accelerators for computing, as well as being irreplaceable scientific research tools.

A quantum computer is a device which is completely different to a traditional computer. Perhaps the only thing in common is the fact that, given an input, it manages to complete calculations in order to produce an output. But the ways in which it manages to manipulate the input data in order to return a result are completely different to those used by a classic computer. The quantum computer, in fact, bases its computation on a basic logical unit that in some sense can be considered an evolution of the bit, called qubit. A qubit (short for quantum bit) is basically a bit capable of exploiting the laws of quantum mechanics to assume superposition states between the two classical states 0 and 1, effectively offering the theorists of quantum algorithms the possibility of exploiting effects impossible to play with a classic computer.

The power of quantum computers is precisely this: given their particular nature very close to the laws of quantum mechanics, they offer developers the possibility of writing very advanced algorithms, with some "extra gears" that allow them, in some cases, to be enormously better performing than their classic counterparts. I say "in some cases" because it is absolutely not certain that a quantum algorithm is necessarily better than a classical algorithm: it depends on the problem you want to try to solve. This statement also gives me the opportunity to underline a very specific concept: a quantum computer is not a more powerful classical computer, but rather a different device. A quantum computer will never replace the classical computer: although there are several algorithms capable of guaranteeing even exponential speed-ups compared to classical algorithms, there are also many other fields of application where the classical computer is more performing than a quantum computer.

Quantum computing is the science that explains how to program a quantum computer; given the deep differences between classical and quantum computers, quantum computing also has profound differences with classical computing. Learning to program and to use the new quantum devices in the best possible way and as soon as possible is strategically very important, especially for a supercomputing center.

It is now clear that quantum computing is a branch of computer science closely linked to high performance computing, especially if we look closely at this historical period. In fact, in recent years we are about to reach what scientists call the era of NISQ computers (Noisy Intermediate-Scale Quantum computers). A NISQ computer, exactly as its acronym says, is a quantum computer of intermediate scale (whose chipset varies between a hundred and a thousand qubits) and noisy (i.e. the quality of the qubits is not yet sufficient to run the most famous algorithms, and more demanding in terms of hardware). Its importance lies in the fact that, according to a large part of the world scientific community, it will be one of the first devices with which it will be possible to reach the first forms of quantum advantage.
(i.e. the moment in which a quantum computer will be able to perform more faster than classic supercomputers).

More specifically: a NISQ computer will be able to achieve quantum advantage thanks to its use in synergy with current supercomputers. Put simply, we think of using NISQ computers as accelerators for computing, in the same way as GPUs are used today. We at Cineca have been monitoring the evolution of quantum computers since 2018. Over the years we have studied and tested quantum technologies on the market, as well as having studied quantum computing also by exploiting our internal computing resources to create quantum computer emulators. In 2021 we founded the Quantum Computing Lab, of which I am the coordinator, and since then we have worked hard to try to offer our Italian universities and research centers all the help we can to understand and exploit this new technology. At the moment we can divide our commitment on three fronts:

Collaboration with European and Italian supercomputing centers and universities for the development of quantum computing: the European scientific community is very active from the point of view of software development for quantum computing. We at Cineca collaborate with other European players on various projects financed by the European community and by Italian internal resources, such as the PNRR. In particular, it is worth mentioning the HPCQS consortium, of which we are an active part, which was born in response to a call from the EuroHPC JU for the construction of a quantum computer network capable of supporting European supercomputers. From the Italian side it is worth mentioning our participation in at least two partnerships concerning the development of quantum software and financed by the PNRR. Not to mention all the other projects that we have submitted and of which we are waiting to know the outcome.

In addition to the participation in the European HPCQS project, Cineca is also one of the founders of the EuroQCS (European Quantum Computers and Simulators) initiative. The EuroQCS initiative was born in February 2022, with the drafting of a whitepaper with the same name in collaboration with the major European supercomputing centers. The whitepaper focuses the reader’s attention on the dual problem of integration between HPC and QC and on the need to be able to access the largest number of QCs available on the market today. In fact, unlike traditional classical computers, the choice of technology with which to build the qubits of quantum computers greatly influences their performance (and we don’t know yet which is the best one!). The EuroQCS initiative became even more concrete in October 2022, when the ranking for the acquisition of a QC through EuroHPC JU was made known. Italy, represented by Cineca, participated and won under the name of EuroQCS-Italy and together with other hosting entities such as France, Spain and Poland as well as partners such as Germany, Slovenia and Ireland, officially gave birth to the EuroQCS consortium, preparing to give the possibility for all European researchers to be able to use in a simple and effective way all the main quantum technologies currently on the market.

Teaching and Outreach: in addition to actively working on the development of quantum computing software, we at Cineca also take care of the creation of schools and dissemination events to teach future generations of quantum programmers and interested researchers how to correctly use these new technologies. We started in 2018 with the first edition of the very popular HPCQC conference (now in its fourth edition); between 2020 and 2021 we created 2 editions of the “Introduction to Quantum Computing” school and one edition of the “Practical Quantum Computing School”, as well as countless participations in many dissemination events.

Support to Italian research: to complete the development and teaching activities, we also offer the concrete possibility to Italian universities and research centers to use via the cloud various different technologies of quantum computers, offered totally free of charge through the project, ISCRA.

Figure 37: QuantumLab website
https://www.quantumcomputinglab.cineca.it/
HPC and Quantum Computing
A winning combination
HUMANITIES AND SOCIAL SCIENCES, DIGITAL ECOSYSTEMS
The Great Circle – The immersive experience of Tübke Monumental
Luisa Mantovani-Löffler, Markus Löffler
Kunst Kraftwerk, Leipzig, Germany

"TÜBKE MONUMENTAL" is a digital art installation presented by the Kunst Kraftwerk in Leipzig, Germany. It is based on the world heritage artwork: THE EARLY BOURGEOIS REVOLUTION IN GERMANY.

The painting, located in the Panorama Museum in Bad Frankenhausen, Germany, is one of the largest panorama oil on canvas paintings in the world. It spans 14 metres by 123 metres with more than 3000 characters. It was created between 1976 to 1987 by the Leipzig painter Werner Tübke (1929-2004), recognized as one of the most important East German artist.

The project has three components. First, a high-resolution digitalisation of this artwork has been undertaken, resulting in a stitched unique 300 Gigapixel image. It is made explorable on a large touch screen, usable for research and digital art mediation. Visitors can easily navigate through the painting to explore every detail.

Second, to explore the personality of Werner Tübke, five eye witnesses were invited to report their encounters with the artist. They provide a deep insight into his mentality and mission.

Third, the key part of the project is the 27-min immersive multimedia installation named „THE GREAT CIRCLE“ created by Franz Fischnaller in close cooperation with VisITLab/Cineca (Antonella Guidazzoli, Maria Chiara Liguori, Silvano Imboden e Daniele De Luca) and the American composer Steve Bryson. Franz Fischnaller has developed the storyboard in 12 scenes, elaborating the painting together with the VisITLab/Cineca team in a creative combination of in-between the 2D gigapixel images and 3D visualisation.

It was a challenge to select 250 figures of the painting and rearrange them according to the storyboard. Besides modelling, texturing, animating, lighting, technical issues were addressed.

VisITLab/Cineca had a big contribution in (e.g. the 3D reconstruction of the landscape, the complex fluid animation of a flooding, the tower of Babylon which rotates and breaks into pieces or the devil to operate his human trap).

The installation provides a unique emotional and multisensory introduction into the painting whose theme is the ever repeating history of failing utopias and human incapability to create a social and just society.

Tübke Monumental was opened to the public in March. It had an enthusiastic media and audience response. This has been praised as setting new standards in immersive art perception. It will be open until end of 2022.
Figure 38: the Great Circle (ph. Luca Migliore)
Figure 39: another perspective of the visual installation (ph. Luca Migliore)
Figure 40: keyvisual of the installation (ph. Luca Migliore)
Figure 41: during the vernissage at Kunstkraftwerk (ph. Luca Migliore)
DARE: a digital environment for urban regeneration

Alessandro Bucci
Municipality of Ravenna - DARE management team

Between 2021 and 2022 Cineca has given great contributions to the UIA European Project DARE - Digital Environment for Collaborative Alliances to Regenerate Urban Ecosystem in middle-sized cities. At the heart of the project there is the aim to provide a digital environment that makes data accessible, understandable, and useful, to enable citizens to become actors in the ongoing Urban Regeneration Process of the Darsena neighbourhood in Ravenna, Italy. The digital environment is composed of 3 parts: a Data Management Platform (DMP), a Content Management System (CMS), and a web interface that functions as a portal (www.darsenaravenna.it). In this framework, Cineca is responsible for the development of the DMP, i.e. the very foundation of the digital environment. In fact, the DMP is the digital space where data from the entire ecosystem of the Darsena is collected and elaborated, and is designed to receive and retrieve data from various sources, from sensors to archives. This wide variety of data is here catalogued and made available for subsequent processing to assess, for example, the level of the Quality of Life in the area. The results are then elaborated and rendered to be displayed online through the web interface. For example, the DMP today receives data from sensors installed inside two public apartment buildings and a school, collecting information on energy performance, indoor and outdoor air quality and thermo-hygrometric conditions: this provides insights to citizens and policy makers on which aspects need intervention and what goals to aim for in terms of increased well being and energy saving.

Through an authentication function, authorised users can progressively define which data sources are active and select the parameters for information retrieval, as well as view their history. A system of Rest APIs, developed with the OpenAPI (Swagger) standard, is used to make the data available to be visualised online.

To grant the durability and the active use of the DMP beyond the project, a pivotal element is to build the necessary competences within the local Public Administration, which owns and manages the digital environment. To do so, Cineca on one hand is designing a capacity building program enabling a selected group of members of the IT department, in particular of GIS-Geographic Information System Office, to operate the DMP autonomously and, on the other, is simplifying the use of the DMP, to make the data management increasingly user-friendly.

Urban Innovative Actions is an Initiative of the European Union promoting pilot projects in the field of sustainable urban development.
Figure 42: DARE project diagram
DARE–IMC platform and the management of 3D digital objects
Elisa Borghi, Antonella Guidazzoli, Maria Chiara Liguori
Cineca

DARE Darsena Ravenna project - https://www.darsenaravenna.it/ - aims to demonstrate the effectiveness of an innovative digital-based and citizen-centred governance approach, to facilitate, support and speed up the implementation and evaluation of the regeneration process of the Darsena, a neighbourhood in the Italian city of Ravenna. As part of the project and its effort in involving the citizens, a selection of 3D reconstructions related to the history of the area was created by FrameLAB, the multimedia and digital storytelling laboratory of the University of Bologna, rising the need of managing them in the DARE-IMC platform, a customised version of IMediaCities- https://www.imedicities.eu/ for the management and metadata enrichment of audio visual contents.

The set up of an easy-to-use set of metadata for managing the 3D assets was, therefore, a prominent feature. The attention was directed towards metadata standards used by digital libraries or systems such as Europeana, the Italian Archival National System, Carare 2.0, 3D-Icons, the Dublin Core, VRA CORE or SIGECweb. The final selection of metadata built a set deemed suitable to keep into account the peculiarities of 3D models without being too burdensome, accounting for example for levels of detail, scientific and artistic validation or, besides georeferencing, the height on the sea level, in order to convey information about the correct positioning of the reconstruction with respect to the historical digital terrain model. Besides the management of 3D assets, the DARE-IMC platform provides a viewer, in order to appreciate the reconstruction without downloading the zip file, whenever it is made available. The 3D models can be realised in the preferred format, but by providing an additional .glb version, the viewer can show the reconstruction directly in the catalogue. The API from the platform enables the use of the DARE-IMC platform contents on the DARE Approdo website.

The approaching end of the project has led to the uniformisation of the aesthetics along the indications of the graphic identity of the project, with interventions both on the DARE-IMC interface and the Virtual Gallery set up, the tool available in the platform for setting up personalised exhibitions starting queries on the audio visual contents. After the conclusion of the project an editorial board will keep the platform and the website updated.
Darsena verde. La Tattica 2022-2023

La Darsena che sarà? Per noi comincia dalla Tattica Verde, che sviluppa un quartiere da vivere.

Figure 43: DARE Website https://www.darsenanaravenna.it/

Figure 44: DARE-IMC platform, the visualisation of the 3D models in the catalogue
COMPUTING PROVISIONING SERVICES
INDUSTRIES
Bringing HPC-fueled innovation to the EU industry

Claudio Arlandini
Cineca

More and more industries look at Cineca as a trusted innovation partner. In our mission to bring innovation through technologies like HPC, AI and Big Data to industries, particularly to Italian SMEs where the technology gap with respect their EU counterparts is larger, we adopted the best practices recommended by the European Commission for the Digital Innovation Hubs, and we are at the center of a growing innovation actors network (European Digital Innovation Hubs, Competence Centres for Industry 4.0, Industrial Associations,...).

One of the pillars of our action, together with training, networking and access to state-of-the-art infrastructures, is the so-called “test-before-invest” service. Test-before-invest means giving SMEs the unique opportunity of experimenting with new digital technologies – software, hardware, business models – to understand new opportunities and return on investments, also including demonstration facilities and piloting. Every innovation-introducing project in a company is associated with a risk. Trying a specific innovation in a controlled environment, without the necessity of upfront investments, and with the time and effort of industry technicians co-funded by the EU or other funding agencies means reducing this risk. At the end, the company will be able to make informed choices, and, most important, without “dangerous” tying up with specific vendors.

For Cineca, test-before-invest activities means first of all supporting Italian companies in EU projects-funded Proof of Concepts (PoC) experiments, mostly selected through competitive open calls. The support starts well before the beginning of a PoC. When an interesting funding opportunity is recognized, it is disseminated through Cineca long-standing partners, like for instance the BI-REX Competence Centre for Big Data, which also act as a first filter individuating the SMEs that most best target the opportunity. Then the SME is helped in writing a strong proposal, an important step since the competition is usually fierce and SMEs rarely have internal experiences and skill sets available for dealing with such kind of research and innovation projects. Winners are then supported in their PoC experiment with computational resources and most important the competences to transform the tested technologies into competitive advantages.

We will present here selected examples from two ongoing EU projects that see Cineca as a core partner: FF4EuroHPC and EUHubs4Data.

FF4EuroHPC
The FF4EuroHPC Project is the third edition of the successful Fortissimo project series. It started at the beginning of September 2020 and will last three years. This project is funded by the European Commission through the Horizon 2020 Programme. It might be considered the successor of projects Fortissimo (July 2013 - December 2016) and Fortissimo2 (November 2015 - December 2018).

The aim, like in the previous Fortissimo projects, is to strengthen the global competitiveness of European small and medium enterprises (SMEs), by facilitating the access to and the use of high-performance computing-related technologies. Through two open calls, ninety-two SMEs have the opportunity to submit innovative projects experiments to demonstrate the benefits of HPC cloud-based advanced modelling, simulation and data analytics.

The project partners support the participating SMEs in the planning, implementation and realisation of their proof of concept (PoC). When the experiment is successfully concluded, a success story raises to demonstrate the benefits of the use of HPC technologies to those involved in the value chain, from the end-user to the HPC-infrastructure provider. The success stories should inspire the industry community in investing in R&D to develop innovative ICT solutions for business benefits. Among the others, these should be the main benefits for SMEs:
• saving time for product/service development and decrease time to market;
• saving costs by optimising data processing tasks;
• higher levels of innovation and more patents;
• more specialised, tailor-made products and services;
• cooperative engagement between Higher Education Institutions and businesses inspires state-of-the-art solutions;
• gaining new knowledge and know-how by R&D experts.

Cineca and its partners propelled Italian SMEs to an unprecedented success rate, with 5 selected experiments over 16 in the first open call, and 8 over 26 in the second one. These experiments cover a wide range of technologies and industrial areas, from new materials for aerospace to fintech and medicine. We present here a couple of noteworthy experiments.

**Catching plastic waste in the Mediterranean with AI**

Figure 45: project diagram

Protecting seas and oceans against the litter is becoming a global concern and there is a growing need worldwide for more efficient, clean and autonomous technologies to identify and collect marine detritus, especially plastics, in a systematic and repetitive way, combining aerial drones individuating plastic waste and marine drones to collect them. Green Tech Solution Srl (GTS) is an agile and innovative start-up that operates in the field of environmental protection and blue growth economy. The use of HPC makes it possible to tackle a computational problem that GTS met during its service for recovery plastic litter in sea. optimizing the plastic litter recovery strategy forecasting the position of hundreds of detritus floating in the sea with suitable accuracy in space and time.

The experiment is fundamental to drive into the next phase the collaboration of the unmanned systems as it requires >250,000 hours of deep learning which is impossible under conventional computational systems but possible thanks to Cineca. The limited forecasting capability of the future position of detritus thus is limiting the efficiency of recovery of the whole system. The experiment aims to overcome this limitation and targets to improve the current Deep Learning approach to 1. Identify and classify marine litters in terms of dimensions and materials (PET, PPT, Biological); 2. Predict the possible trajectories of classified waste over a longer time; 3. Search the “best” trajectory to collect as much waste as possible under constraints.

It is expected that this system will be able to reduce the cost for public administrations to clean up the Mediterranean coast of at least 70%, without involving the use of polluting diesel-fueled boats.

The partners are:
• End User: Green Tech Solution SRL
• Domain Experts: Università degli Studi di Napoli Parthenope and BI-REX - Big Data Innovation & Research Excellence
• HPC Provider: Cineca

For more info: https://www.if4eurohpc.eu/en/experiments/2021070911193805/hpcebased_navigation_system_for_marine_litter_hunting
Saving children with congenital heart diseases with HPC

Congenital heart diseases (CHDs) account for nearly one-third of all congenital birth defects. Without the ability to alter the prevalence of CHD, interventions and resources must be focused to improve survival and quality of life. The Modified Blalock Taussig Shunt (mBTS) is a common palliative operation on cyanotic heart diseases, but it is associated with significant mortality, the most threatening complications being over-shunting and shunt thrombosis. The experiment aims at setting up an application to support surgical planning with advanced numerical means, exposed in an interactive and effective manner. This application uses patient-specific computationally intensive models running in a reliable and cost-efficient way on a cloud-based HPC environment. The experiment aims at building up an affordable decision support web application that, thanks to the medical digital twin (MDT), allows surgeons to approach the mBTS medical intervention at best. The tool generates the ROM of a patient-specific vascular district in which the shunt implantation is geometrically parameterized. The ROM generates pre-defined CFD results of different configurations of the vasculature with shunt implantation. With a dedicated User Interface, the medical staff will be able to inspect the MDT of the patient in an interactive way, to finalize the decisions on surgical intervention.

The partners are:
- End User: InSilicoTrials Technologies SpA
- Technology expert, ISV: RBF Morph Srl
- Application expert, End-User: Fondazione Toscana Gabriele Monasterio
- Technology expert: RINA Consulting
- HPC expert, Host Centre: Cineca

For more info: https://www.ff4eurohpc.eu/en/experiments/2021070910512579/cloudbased_hpc_platform_to_support_systemicpulmonary_shunting_procedures

Figure 46: project workflow

EUHubs4Data

The objective of EUHubs4Data is to create a federation for Big Data cross-border experimentation and innovation, providing a complete pan-European catalogue of data
sources and services to foster data driven innovation at local and regional level. To do
dthis, its members, most of them being recognised as i-Spaces (DIHs) of the Big Data Value
Association, provide tools, assets and methods to European companies and society to
benefit from digital opportunities coming from data driven solutions and models. The
project aims to establish a European federation of Big Data Digital Innovation Hubs based
on the premise “global catalogue, local offer”. To achieve this, EUHubs4Data must attract,
support and engage end users from local and regional ecosystems through a common
European catalogue of federated datasets and services. EUHubs4Data brings together
relevant initiatives to create this common catalogue, capitalizing the work done by all those
initiatives, and improving the efficient use of data across the EU.
One key way to implement this vision is the selection of cross-border 12-months data-
experiments that include one European SMEs and two or more EUHubs4Data partners.
Selected through three rounds of competitive open calls.
Cineca is now supporting four experiments selected in the first two open calls (the final one
will open in September 2022). We present here one of them.

Providing nowcasting services of sea conditions thanks
to HPC and AI

Off-shore industry operations security and Cities/Regions coastal area planning and
monitoring require fast and accurate sea-state forecasts. All the services now available
cover the sea areas only up to 10 km of distance from the coast. The current scenario for
sea state analysis and forecast is based on physical models and the two main operational
services over the Mediterranean Sea are run by the Copernicus Marine Environment
Monitoring Service (CMEMS) and the Italian National Agency for New Technologies, Energy
and Sustainable Economic Development (ENEA).
The experiment aims to develop an AI based model for sea-state analysis and forecast,
providing results aligned with the physical model currently used at national level, including
e.g. a sea circulation forecast daily service over the next day and a sea circulation daily
forecast over the next 5 days.
The advantages of adopting a AI-based method are:
• Much faster data processing and generation of results,
• Improved accuracy by ensembling various results (from different initialization parameters),
• Improved geo-spatial resolution (1x1km) and coastal coverage (up to 1 km from the coast)
also by using Earth Observation (EO) satellite data (e.g., Sentinel-1 and Sentinel-3) and
meteorological data (Mistral from Cineca),
• Wave-Circulation Nowcasting over the next 12 hours, with 30 minutes temporal resolution
and 1x1Km geo-spatial resolution,
• Easy evolution of the model and tuning to different geographic areas.

The partners are:
• End User: GMATICS srl
• Supporting DIHs: Cineca, Data Cycle Hub
The current technology transfer project for industry is Euro-CC (started in May 2021 and to be completed in September 2022) and is the main project for the industry in progress.

Thanks to European funding, the EuroCC-Italy Competence Center, a program for the development and promotion of High Performance Computing tools (HPC), advanced simulation, high performance data analysis (HPDA) and intelligence artificial (AI) in order to promote innovation in industry and the public sector a national level.

The Long-term goals of the EuroCC-Italy Competence Center are:

- to support the national ecosystem fostering the collaboration between research teams and institutes, HPC computing centre and corporates in order to support SMEs in innovation projects.
- to become an open service platform aggregating competences and new institutions to support innovation and collaboration with industry in the continuum of HPC/HPDA/AI technologies and methodologies.
- to attract young researchers and professionals, support the dialogue between generations in an innovation friendly environment.

To reach these goals EuroCC Italy relies on the competences and infrastructures available in Cineca, as well as the competences and networks of third parties, such as Associazione Big Data, (ABD), a scientific network in computational sciences, BI-REX (a network of industries and industry associations, which deliver services for innovation capable of reaching SMEs in the manufacturing sector), corporates like Eni, Dompé and Leonardo, which have a long-lasting experience in using and applying HPC, HPDA and AI in their research and business processes. In the Competence Center (CC) corporates have the role of supporting innovations in SMEs which belong either to their supply and value chain, or to the green economy national and European ecosystem, through concrete application examples and needs. The challenging objectives of this first project year were:

- to set up working groups effectively involving all third parties, monitor their progress and ensure timing and quality of the project deliverables;
- to develop a first level of training on HPC, HPDA and AI through the collaboration with third parties, which will be added to existing training programmes for the industry;
- to set up and launch the calls for Proof-of-Concept (PoC) projects involving SMEs and innovative start-ups;
- to map the training activities in HPC/HPDA/AI at the national level, to reach an overview of the research teams working in the field, and identify those that are already cooperating with the industry;
- to create awareness on EuroCC-Italy and its programme, on relevant HPC technologies, on Quantum Computing, and on the success stories showing business impact of HPC, HPDA and AI based innovation, in SMEs and corporates.

In particular, the Competence Center intends to convey knowledge from the world of research into targeted training and technology transfer actions, in order to respond to concrete needs of the national production fabric, thanks also to coordinated intervention of trade associations.

During 2021, 11 training courses and 6 events were organized with the aim of:

- Share knowledge about HPC, HPDA, AI and their use;
- Provide high technology skills to industry to create innovative products;
- Create business-to-business (B2B) events where industries could meet, locally and nationally, organizations involved in technology transfer linked to HPC, HPDA and AI.

The training courses were organized face to face in seven different cities: Bologna, Genoa, La Spezia, Milan, Turin, Trieste and Modena.
National Recovery and Resilience Plan
High Performance Computing projects supported by the National Recovery and Resilience Plan – Next Generation EU

Maurizio Ortali
Cineca

The EU's long-term budget, coupled with Next Generation EU (NGEU), the temporary instrument designed to boost the recovery, form the largest stimulus package ever financed in Europe. A total of €2.018 trillion are helping rebuild a post-COVID-19 Europe. In this context of extraordinary funding, Italy outlined a detailed National Recovery and Resilience Plan – PNRR (Plano Nazionale di Ripresa e Resilienza), a large package encompassing 6 policy areas – “missions” (M1: digitization, innovation, competitiveness, culture; M2: green revolution and ecological transition M3: infrastructure for sustainable mobility; M4: education and research; M5: inclusion and cohesion; M6: healthcare) and 16 components. Among the different intervention areas, of particular interest for large-scale research infrastructure is the Mission 4 Component 2 called “from Research to business” with the aim to strengthen and encourage the dissemination of innovative models for basic and applied research, conducted in synergy between universities and businesses.

In this context, during 2022 Cineca has participated to several PNRR Research initiatives that have been financed by the Italian Ministry of Research (MUR), with the aim to further develop and exploit its large scale computational platforms:

1. National Center for High Performance Computing, Big Data and Quantum Computing (ICSC)
2. National Biodiversity Future Center (NBFC)
3. Regional Ecosystem for Sustainable Transition in Emilia-Romagna (ECOSISTER)
4. Enlarged Partnership GRINS (Growing Resilient, INclusive and Sustainable)
5. D-3-4-Health initiative (Digital Driven Diagnostics, prognostics and therapeutics for sustainable Health care)

1) National Research Center for HPC, Big data and Quantum Computing (ICSC)

The National Research Center HPC, Big data and Quantum Computing, provides a fundamental opportunity for the national system at the scientific, industrial and economic level to face present and future scientific and social challenges, strengthening and expanding existing skills and infrastructural resources. The National Center will be 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 affiliated institutions, as well as through open tenders. The Hub will also implement all activities in education and training, entrepreneurship, knowledge transfer, policy and awareness raising.

The Hub and the Spokes are made up of universities, research bodies, private and public operators. The proposed National Center includes a transversal Spoke (Spoke 0 “Supercomputing Cloud Infrastructure”) and 10 thematic spokes.

The National Center, coordinated by National Institute of Nuclear Physics (INFN), has two main objectives:

1) create a national computing / computing infrastructure, similar to a Datalake, grouping the existing High Performance Computing (HPC), High Throughput Computing (HTC), Big Data and network infrastructures, with new targeted resources, activated through internal funding, in order to provide the scientific and industrial communities with a flexible and uniform Cloud interface;
2) create an attractive ecosystem around the infrastructure that supports the academic world and the industrial system, favoring the exploitation of IT resources and the development of new advanced computing technologies.

Cineca in the Center is clearly a key actor, being among the founders of the Hub, Spoke Leader of the Infrastructure component and Affiliated party (partner) in two other Spokes (Future HPC and Quantum Computing). The Grant assigned to Cineca across the 3 year lifespan of the project will significantly empower the Cineca HPC Department personnel and complement and expand new HPC and Quantum innovative components.

Figure 47: funding members
2) National Biodiversity Future Center

The National Research Center (CNR) coordinates the National Biodiversity Future Center (NBFC) with the aim of systematizing Italian research and innovation excellence on biodiversity and delivering to Italy, after three years of work, an institution to safeguard, develop and enhance biodiversity that is unique in its gender and sustainable over time. In particular, the NBFC intends to carry out operational interventions aimed at:

1) Monitoring, preserving and restoring biodiversity in the marine, terrestrial and urban ecosystems of the Peninsula. Thanks to modern technologies, expressed in terms of Key Enabling Technologies (biotechnologies, artificial intelligence and digitization, technologies for life sciences) it will be possible to increase knowledge relating to the heritage of biodiversity and to map distribution, value and peculiarities in the habitats of terrestrial ecosystems, freshwater, transitional, marine and urban. This knowledge and technologies will be the basis of intervention programs for biodiversity conservation and restoration also thanks to the development of increasingly realistic predictive models, the design of new early warning technologies and the definition of tools to support functional biodiversity and ecosystem resilience.

2) Enhance biodiversity and make it a central element on which to base sustainable development. The study of the different levels of biodiversity and technological innovations will make it possible to identify new products, processes (Bioprospecting) and solutions (Nature Based Solutions - NBSs) capable on the one hand of implementing and protecting the different living forms and on the other of using this diversity, to obtain new drugs, foods, more efficient and sustainable materials. These interventions will have direct and indirect effects on the well-being of the person, for example, by reducing temperatures and pollution, increasing the tourist value and use of degraded and highly man-made areas, favoring the relationship between people and improving the living environment.

In the context of this National Center (project duration: 3 years), Cineca will be a key actor being one of the founders of the Hub and will participate in the spoke 7 with the aim of providing the Center with skills for the use of computing infrastructures and HPC.

3) ECOSISTER - Ecosystem for sustainable transition of Emilia-Romagna

As defined in the PNRR call for proposals, the regional “Innovation Ecosystems” are networks of state and non-state universities, Research publics, local public bodies and other highly qualified public and private entities aimed to intervene in areas of technological specialization consistent with the industrial and research vocations of the reference territory, promoting and strengthening the collaboration between the research system, the production system and local institutions. To support his strong and highly competitive production system, coordinated by University of Bologna and Art-ER, some Universities and Public Research Bodies of Emilia-Romagna decided to build an integrated innovative ecosystem, in association with Universities, Research Centers and other territorial players. In order to maintain a leadership role in the international context and remain anchored to the vocations that characterize the Emilia-Romagna region, the project intends to support the ecological transition of the regional economic and social system through a process that transversally involves all sectors, technologies and skills by combining digital transition and sustainability with the work and well-being of people and the protection of the environment, in line with the objectives of the Pact for Work and Climate, and integrating with regional, national and European programs. In the 36 months of the lifespan of the project, Cineca will participate in the activities of Spoke 6, with a role of trait-d'union and consolidation with respect to the position of Emilia-Romagna as National HPC Champion by aggregating its cutting-edge skills in high performance Computing, Data Science and Technology, to support the ecological transition.
4) GRINS – Extended Partnership

The GRINS (Growing Resilient, INclusive and Sustainable) Consortium is an extended partnership coordinated by University of Bologna- Department of Economics - that intends to develop AMELIA (dAta plattFOrM for the transfEr of knowldge and statistical Analysis), an Online Data Platform giving access to high quality data and instruments for data analysis for a wide range of applications. It will offer tools to support fundamental and applied research for firms and households and for policy analysis and evaluation of the actions of public administrations. The GRINS AMELIA will create knowledge and transfer it to private and public actors as well as to the national system of research according to the guidelines and principles of the PNRR. The project is designed following the priorities set by the Italian National Research Plan (PNR) and in strict adherence with the fundamental underline goals that inspire the whole EU-NRRP action: favouring resilient, inclusive, and sustainable growth.

In the 36 months of the lifespan of the project, Cineca will participate in the activities of Spoke 0, with the aim to support the development of the ICT and HPC component needed to the project, by aggregating its cutting-edge skills in high performance Computing, Data Science and Technology.

5) D-3-4-Health Initiative

D-3-4-Health project, coordinated by Sapienza University (Rome) aims at enabling new technologies for data collection and analysis in order to provide personalized medicine. This approach involves the prevention, the diagnosis and therapy of oncological, neuropsychiatric and metabolic pathologies through the creation of a “digital” and a “biological twin” of patients.

Cineca will participate in two spokes of the project and the CINECA HPC and cloud infrastructure will be a key asset for the proposal, as the appropriate usage of HPC architectures has already expressed high potentiality in the field of personalized medicine where urgent computing is critical. To this extent, CINECA can contribute leveraging the experience acquired in several European projects with GDPR compliant procedures for data protection and data management and supplying HPC or Cloud infrastructure to host pseudonymized or anonymous data.

In addition to these new starting projects, Cineca has also supported the National Institute of Oceanography and Applied Geophysics (OGS) on a PNR proposal submitted to the MUR in the field of “strengthening and creation of research infrastructures” (Terabit project), in continuity with the past activities the PRACE-Italy.

Finally, Cineca has defined the contractual process and launched with the Central Institute for the Digitization of Heritage (IC-DP) of the Ministry of Culture (MIC) the activity envisaged for Cineca in Measure 1.1 Platforms and digital strategies for access to heritage cultural, belonging to measure M1C3 of the PNRR: Culture and Tourism.