

Wrap up – Opportunities

SOSC22 – Perugia, 2/12/2022

Davide Salomoni (<u>davide@infn.it</u>)



Opportunity areas

- Collaborations with INFN related to computing topics mentioned at SOSC22 are possible in technical areas such as:
 - System administration
 - (Advanced) cloud computing infrastructures, including integration with HPC, GPU, FPGA, adaptation of multi-disciplinary use cases
 - User support
 - Middleware development
 - Security / data protection
 - Scientific communication
- The detail of the technologies involved in these opportunities is too long to list. Have a look at topics covered at SOSC22 for a sample.



Targets and forms

- The opportunities we refer to here are targeted to (not exhaustive list):
 - People:
 - Masters' Degree Students (MD thesis)
 - Post-graduates (including post-docs)
 - Organizations:
 - Academia and research organizations
 - Industry
- These collaborations may take different forms (not exhaustive list):
 - For people:
 - Research associates (also in its more recent forms, such as "research contracts")
 - Temporary contracts, beyond the "research contracts" form
 - Permanent positions (as a *possible* outcome of some NRRP projects, see later)
 - For organizations:
 - Scientific collaborations
 - Technology transfer projects



Locations

- In principle, most INFN sites are eligible to host one of the collaboration opportunities mentioned here.
- *In practice,* check with the organizers of this school in case a specific location interests you.
- Most of the positions typically involve some travelling, in Italy and abroad, and require flexibility with working hours.



Education and recruitment process

- As a rule, most STEM degrees are fine.
- We value intelligence, curiosity, desire to learn and contribute to the advancement of science.
 - INFN is strongly mission-focused and is consistently ranked among the best scientific organizations worldwide. We want to keep it like that, and possibly improve.
 - And, as Steve Jobs said: "it doesn't make sense to hire smart people and tell them what to do; we hire smart people so they can tell us what to do."
- The recruitment process is based on public, competitive exams.

Where to look and contacts



- The INFN institutional site for job offers is https://jobs.dsi.infn.it
- However, opportunities may also be offered before they are published in that page, depending on dynamic conditions such as new projects, technology transfer agreements, dissertation topics, etc. Therefore, in case you are interested in the areas mentioned in this presentation and you don't see a specific job opening in the site above already, we suggest you contact the SOSC22 organizers at sosc22-pc@lists.infn.it.
 - This of course is especially true if you think that your own organization may be interested in collaborating with INFN.

Examples of collaboration opportunities

- INFN has many ongoing initiatives. Here I provide just some examples of externally funded projects, which offer several kinds of collaboration opportunities.
 - I will not mention opportunities related to established experiments funded with "regular" funds – such as, for instance, grants in the context of a certain physics collaboration.
- Opportunities can be divided by the overall framework:
 - European projects
 - National projects
 - Local projects



European projects

- INFN participates to many EU projects involving computing technologies. These projects normally let us hire young people to perform some of the related activities.
- Let me mention here InterTwin, "An Interdisciplinary Digital Twin Engine for Science". In this project, we work on, among other things:
 - A Federated compute infrastructure, covering HTC, HPC, Quantum
 - A Federated data infrastructure
 - Federation services and policies
 - An Al-based resource orchestrator



NRRP and INFN Cloud



- For what regards the **NRRP** (*), our reference project is ICSC, the Italian Research Center on High-Performance Computing, Big Data and Quantum Computing.
- We have already mentioned and used INFN Cloud multiple times here at SOSC22. In the context of the NRRP projects, ICSC included, INFN Cloud will become the "National DataCloud", a unique, distributed set of Cloud-native resources and services offered to academia, research and industry.

(*) NRRP = National Recovery and Resilience Program = Piano Nazionale di Ripresa e Resilienza (PNRR)



Federazione Cloud

Modello: integrazione «debole» di cloud indipendenti

In uso sulla cloud INFN

Tutti i centri INFN adotteranno il modello nell'arco del progetto ICSC

Estensione a CINECA e ad altri provider del centro nazionale (CMCC, ...) Attività già iniziate nell'ambito del POR-FESR SUPER dell'Emilia Romagna Federazione integrata dei sistemi Tier1 e dei sistemi Europei di EuroHPC mediante rete ad iperconnetività



Italian Research Center on High-Performance Computing, Big Data and Quantum Computing

Proof of concept previsto a M8

Cloud national infrastructure for supercomputing. Hub & Spoke organization:

10 vertical spokes for technology developments and software applications



ICSC Kick-off Meeting Bologna, 25-26/11/2022



MUR

INFN

ICSC

Big Data

Centro Nazionale HPC,

e Quantum Computing



ICSC, some details





Infrastructure



Data Lake



SOSC22 - Perugia, 2/12/2022

11



Architectural pillar: the DataCloud PaaS

- The DataCloud PaaS is rooted on:
 - 1. A distributed resource orchestration framework
 - 2. A standard—based **federated solution for identity** access management (INDIGO-IAM)

Paas Secure Provider (RP) Constraint Co

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

Examples of INFN Cloud services: on-demand HTCondor



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





Multiple ways to ingest and process data are possible. For example, to handle sensitive data (e.g., in the nation-wide Health Big Data project, or in NRRP projects), we plan to support these options, which would be a subset of those offered by DataCloud as a whole:

- 1. Central harvesting of data generated remotely
- **Edge-level anonymization**, followed by 2. central ingestion and analysis of data
- 3. Edge-level feature extraction, followed by central ingestion and analysis of features
- 4. Federated learning based on edgelevel training, followed by publishing of the trained methods and by inference performed either centrally or at other edge locations.

https://www.physicamedica.com/article/S1120-1797(21)00320-3/fulltext

The DataCloud internal structure







INFN



ICSC: Innovation grants

SCOPES

- A) Fostering technology scale-up and transfer
- B) Supporting new start-ups and spin-offs
- C) Addressing skill gaps
- D) Creating ISCS community and promoting entrepreneurial culture

MODALITIES

- Exploitation plans
- Call for ideas and business plans
- Contests and challenges
- Innovation grants

ACTIVITIES

- Deployment of demonstrators
- Scale-up grants
- Proof fo concepts
- Pilot applications
- Pre-seed funds
- Life long learning
- Training
- Industrial PhD projects

NRRP: DigitAl lifelong pRevEntion (DARE)

- We observe massive digital transformations, with multiple processes of integrating digital technology and data into all areas of everyday life, including health.
 - DARE will enhance the tools and knowledge that allow us to exploit the potential of data to define, monitor, and even predict health trajectories for the sake of health promotion and prevention.
- Digital prevention = health promotion and prevention actions enabled by digital technologies, which have the potential to significantly improve the speed and accuracy of key public health functions such as forecasting, surveillance, early detection of and response to acute and chronic/complex diseases and, more generally, for health-related conditions through the lifespan.
- How? Working with both health data and health-related data (citizen-generated data), such as sociodemographic data, telecommunications data, and weather data.



 These data might also include personal data that are not directly health-related (e.g., location data, customer shopping data, or social data collected through smartphones or self- tracking devices).





National project: Health Big Data

Financial Contribution:	Duration:
55 million €	10 years
Provided by the Ministry of Health	Start: 16 December 2019

51 IRCCS, 4 networks (pediatric, oncology, cardiology, neuroscience) INFN is the Technology Provider of the project. Clear synergies with NRRP projects.

- Creation or enhancement of:
 - A set of local IT platforms in each participating IRCCS to ensure extraction, integration and interoperability of clinical and scientific data.
 - A federated IT platform, ensuring connectivity between IRCCS and advanced analysis of shared data.
- Handling of vastly heterogeneous data:
 - omics (genomics, transcriptomics, proteomics, metabolomics), clinical (electronic medical record and patient follow-up data), imaging and radiomics; data provided by the patient.



Local project: INFN-IRCCS Sant'Orsola

- A joint research agreement with the following objectives:
 - Develop secure applications for genomic data
 - Develop GPU-based solutions for genomic analysis methods
 - Create a federated and integrated cloud platforms for omics data
 - Adapt genomic pipelines to cloud and data lake architectures based on microservices
 - Integrate omics data and other clinical data, such as Electronic Medical Records (EMR)



Conclusions



- There are plenty of opportunities to work with INFN in technology / computing areas. They are directed mostly to graduates in STEM disciplines and cover a vast number of state-of-the-art topics.
- For instance, through just NRRP projects, INFN is looking to hire O(50) temporary positions in these areas very soon. Several of them will eventually be converted to permanent positions.
 - A first bunch of competitive exams should be published in the coming weeks.
- In addition, there are many other possible collaborations through projects at the international, national and local levels. Collaborations with other institutions and industry are also foreseen.
- For any inquiries about these opportunities, send email to <u>sosc22-pc@lists.infn.it</u>.

The End

Thank you for attending SOSC22 and making it a success.

Enjoy your studies and work, do not lose focus or motivation, and always remember to have fun!



