EPJ Web of Conferences **295**, 08009 (2024) https://doi.org/10.1051/epjconf/202429508009
CHEP 2023

INFN Cloud

Users and Projects Support, Training and Communication

Francesco Sinisi[1,*], *Ahmad* Alkhansa[1], *Alessandro* Costantini[1], *Cristina* Duma[1], *Federica* Fanzago[3], *Daniele* Lattanzio[1], *Alessandro* Pascolini[1], *Stefano* Stalio[2], *Emidio* Giorgio[4], *Carmelo* Pellegrino[1], and *Davide* Salomoni[1]

[1]INFN CNAF
[2]INFN LNGS - Laboratori Nazionali Gran Sasso
[3]INFN Padova
[4]INFN LNS - Laboratori Nazionali del Sud

Abstract. Having a long tradition in state-of-the-art distributed IT technologies, in the last couple of years INFN made available to its users "INFN Cloud": a cloud infrastructure and related services portfolio dedicated to the scientific communities supported by INFN.
Given the distributed nature of the infrastructure as well as the considerable number of technical solutions provided to the INFN users, it is important to have a reliable user support service aimed to properly interact both with INFN Cloud users and administrators.
As an added value, proper training activities have been defined and differentiated to different types of users and the training courses are integrated with a rich set of user guides and technical documentation.
In this article, an overview of the INFN Cloud, its evolution to DataCloud project, and the support and training activities will be provided and presented.

1 Introduction

Since several decades, INFN runs and supports the largest research and academic distributed infrastructure in Italy. In this context, INFN Cloud provides a set of cloud computing services (§ 3) for the INFN community. These are cloud services that support scientific computing, software development, training and the extension of local computing and storage resources: from data analysis to the solution of computational problems, from services for data management and storage to the development of ready-to-use machine learning technologies, up to the possibility of creating virtual and personalized environments, both for individual users and for scientific collaborations.

INFN Cloud is the result of a history of excellence spanning several decades, which has led INFN to develop, manage and support a infrastructure for scientific research spread across the entire national territory, with a large national data center, CNAF[1], and 9 other large data centers, all interconnected at high speed through the GARR Consortium network[2], to meet the needs of several dozen international collaborations in physics and many other scientific fields. In particular, INFN-CNAF is one of the Worldwide LHC Computing Grid (WLCG)

*e-mail: francesco.sinisi@cnaf.infn.it

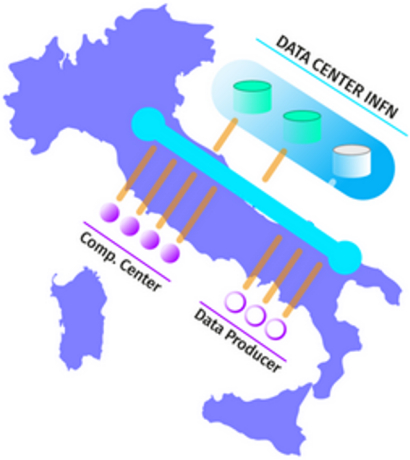

Figure 1. INFN distributed architecture.

Tier-1 data centers. It provides support in terms of computing, networking, storage resources and services also to a wide variety of scientific collaborations, ranging from physics to bio-informatics and industrial engineering.

In the following sections an overview of the INFN Cloud, its evolution to DataCloud project, and the support and training activities are provided and presented. In particular, we will deal both with the activities aimed at supporting users during the adoption of the proper cloud services most suitable for their needs, and the improvements of the INFN Cloud portfolio. In the latter case, we are constantly collecting new requirements that can be converted in further and reliable solutions to be offered to the communities.

2 INFN Cloud - Organization and Architecture

INFN Cloud is internally organized into 7 Work Packages (WPs), where people belonging to different INFN units are actively collaborating to the project. The WPs are here listed:

- WP1: Operations
- WP2: Documentation, User Support, Communication and Training
- WP3: Resources, Data Lake and Sustainability
- WP4: Security and Policies
- WP5: Middle-ware and New Services
- WP6: Research and Development, Test-beds, Use Cases
- WP7: Integrated Systems Management and Legal Compliance

On top of the WPs there is the INFN Cloud Project Management Board (PMB), composed by the INFN Cloud Coordinator and by the WP Leaders, who control and organize the high level activities in the project. The PMB is reporting to the INFN Computing Coordination Committees (C3SN) and to the INFN President.

The INFN Cloud infrastructure is based on the *core backbone*. It is composed by CNAF, the national center of INFN (National Institute for Nuclear Physics) dedicated to Research and Development on Information and Communication Technologies, with ReCaS-Bari[3] data center. The backbone is interconnected at high speed (~ 10 Gbps) through the GARR Consortium network (see Figure 1 for more details). This basic infrastructure is completed with other Server Providers, the number of which is continually growing, thus allowing an increase in the computational resources offered by INFN.

INFN Cloud backbone hosts the Core Services, such as the dashboard, the orchestrator, which chooses the most suitable site to allocate resources, monitoring and accounting servers, storage servers and their replica, and other services typically reserved for special purpose tasks. The backbone, as well as other federated Server Providers, can also host community-specific, reserving part of their computing and storage capacity, as well as catch-all user services.

The INFN Cloud infrastructure is strictly regulated by the *Rules of Participation* and to be part of it, a Cloud Service Provider must obtain the approval by the INFN Cloud project management board. In case of special needs, the INFN Cloud infrastructure can transparently use other public or private Cloud resources to augment its capacity or solutions. The access to the INFN Cloud services is reserved to INFN personnel or people that established formal collaborations with the Institute, such as research associates from Universities or other research centers.

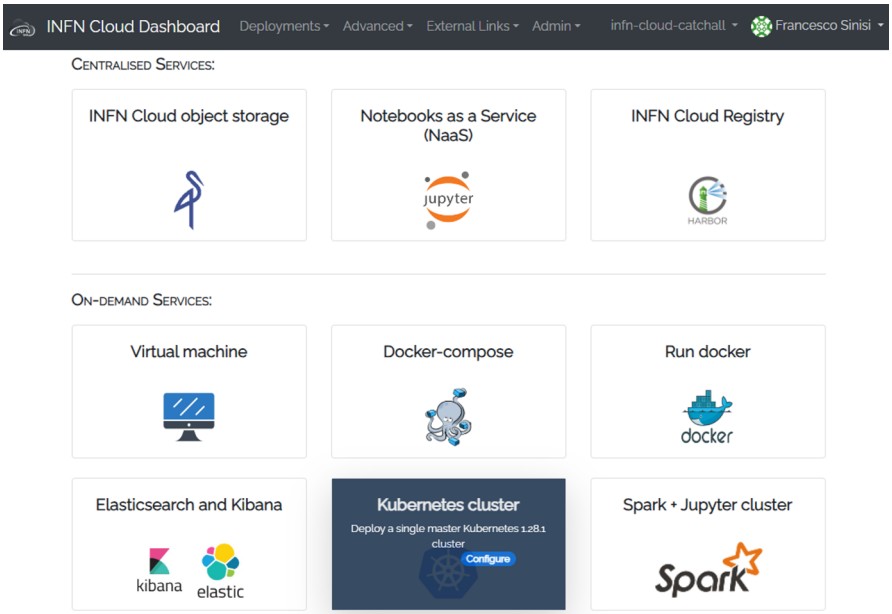

Figure 2. INFN Cloud PaaS Services made available to the user, depending on the INDIGO-IAM group the user belongs to.

Authentication and authorization within INFN Cloud is enforced through the INDIGO-IAM[4] (Identity and Access Management) federated solution, currently maintained and developed by INFN and fully compliant with European Open Science Cloud (EOSC) and industry standards (OpenID Connect, OAuth2). From an administrative prospective, each user can be associated to different groups, each one allowing the access to different set of services

(look at the top right part of the Figure 2). INDIGO-IAM, therefore, allows the management of the authorization policy for all the authenticated users within INFN Cloud.

3 INFN Cloud - Service Catalogue

The INFN Cloud Service Catalogue, as from Figure 3, is composed by different macro services:

- Compute Services: A list of services that enable a specific cloud technology (Docker, Docker Compose, Kubernetes);
- Analytics: A collection of ad-hoc solutions for analytic purpose (Elasticsearch and Kibana, Spark, Jupyter);
- Machine Learning: List of ready-to-use Machine Learning services (custom Jupyter environment);
- Data Services: Data management and storage services (OwnCloud, NextCloud, Object Storage);
- Scientific Community Customizations: Customized environments.

Each macro service groups a set of services that can be easily deployed by users via the INFN Cloud Dashboard, a Graphical User Interface that supports the user on the configuration and on the management of the deployed services (Figure 2).

The INFN Cloud, thus, offers to its users a comprehensive number of services, that can be deployed and hosted on the Cloud infrastructure. Each user-oriented service can use the IaaS, PaaS and SaaS (Infrastructure, Platform and Software as a Service, respectively) components to cope with the needs of the user. As we will see in the sub-chapter § 4.1, the INFN Cloud Service Catalogue is defined upon clear users' requirements. It is based on modular, open source solutions and can be easily extended either by the INFN Cloud support team or directly by end users. In fact, new services and features are continuously added to the dashboard, without forgetting the continuous maintenance and updating of those already present.

4 INFN Cloud - User Support

Given the distributed nature of the infrastructure, and the complexity of the services supported, which require specific technical solutions, at least some of them, it is crucial to have a reliable support network based on specific expertise and assisted by a set of services aimed at tracking the interactions between the users and the INFN Cloud experts in order to improve the user experience and facilitate the use of the resources by providing explanations and documentation to them.

4.1 User support organisation

The user support is organised in a multi-level structure where the *first level* (*L1*) is responsible for managing users registration requests, new use-cases enrollment, guiding the use of the services available in INFN Cloud portfolio; the *second level* (*L2*), instead, provides support for those requests or problems that require higher privileges or specific expertise.

To get access to the INFN Cloud services, some requirements must be met by the users. Among those requirements, the user must have at least a digital identity in the INFN registry to be able to use services at the SaaS level. For PaaS services, the user must also gain a nomination as system administrator. This is required because each PaaS service instantiated

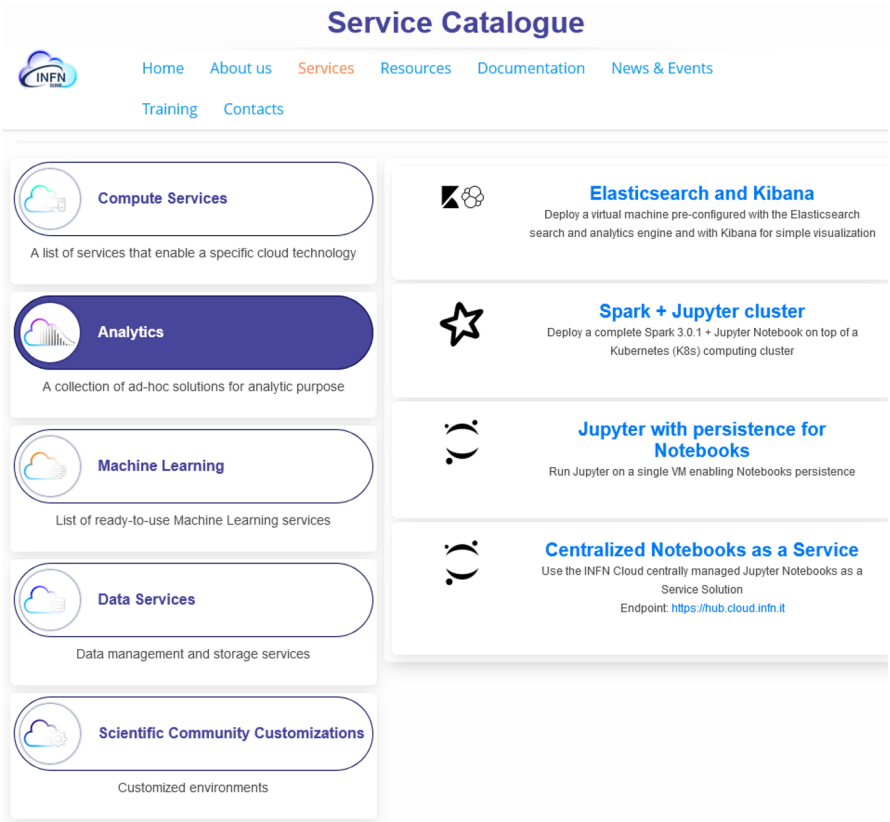

Figure 3. INFN Cloud Service Catalogue.

via the INFN Cloud dashboard launches at least one Virtual Machine with an associated public IP, that can be used to access the machine itself and, in some cases, to expose the applications running on them. To avoid legal issues caused by possible security breaches, each PaaS user must be properly educated and trained on security risks, mitigations and procedures. It is one of the duties of the user-support team, in particular at the L1, to enforce the control over the users' possession of such requirements.

4.2 User support Tools

The user-support activity is tracked on a ticketing system based on the Atlassian tool Jira Service Desk[5], that provides a number of services aimed at handling user requests as registration or training requests, incidents, technical support and general tips. The user's request generates a ticket that is managed by a member of the support team, who will assist the user.

Moreover, the ticketing system can be used also for statistical analysis by tracking the number of tickets opened and resolved over a period or check trends. Figure 4 and Figure 5 show such statistical information corresponding to the last year of activity.

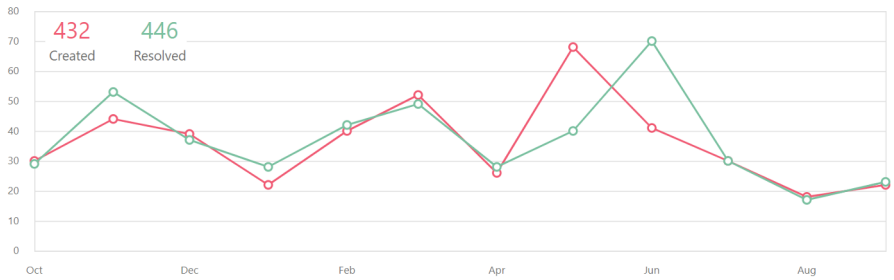

Figure 4. Number of tickets opened and resolved over the last year.

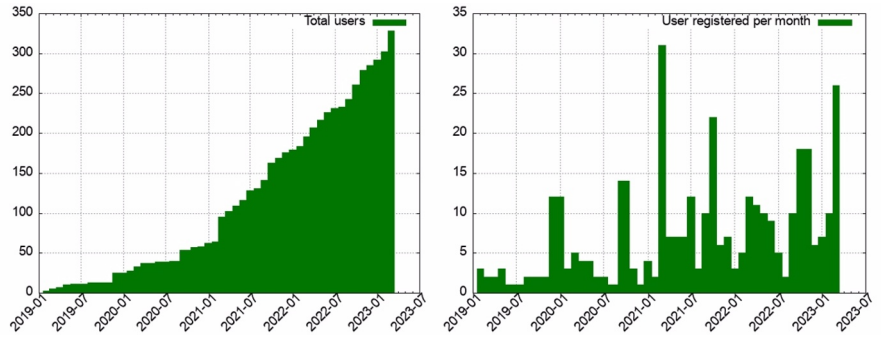

Figure 5. Trend of subscriptions to INFN Cloud in cumulative (left side) and frequency (right side) form, since it has been operational.

5 INFN Cloud - Training and Guides

Training is a milestone activity carried out by the DataCloud WP2. It is provided at different levels, both to users and administrators. It is useful to users to make them comfortable with INFN Cloud solutions, while improving the skills and technical knowledge of administrators. The list of training events can be consulted at https://www.cloud.infn.it/training/[8].

The training for users, in fact, is specifically devoted to cover different learning needs, to promote correct and efficient usage of hardware and software resources, and to enhancing internal understanding of INFN Cloud details. For this reasons, two types of training events have been individuated: internal and user training. The first, delivered from developers or system integrators, is directed to the user support team, as well as other INFN Cloud staff, and aimed at improving knowledge on newly implemented products and the correct way to configure and use them. The second is targeted to new or potential INFN Cloud users, with the aim of illustrating the products available, their potentialities and the correct way to use them, going behind the simple products usage but also touching arguments such as responsibilities and legal duties. Moreover, a more advanced level of user training has been provided for those already experienced, showing possible implementations of their scientific workflows using the tools offered by INFN Cloud. The most common or used use cases are proposed during the training, along with possible implementations.

As an added value, the training courses are complemented with a rich set of user guides and technical documentation, accessible online at https://guides.cloud.infn.it[9], to ease the use and the adoption of the services made available through the INFN Cloud PaaS. All the

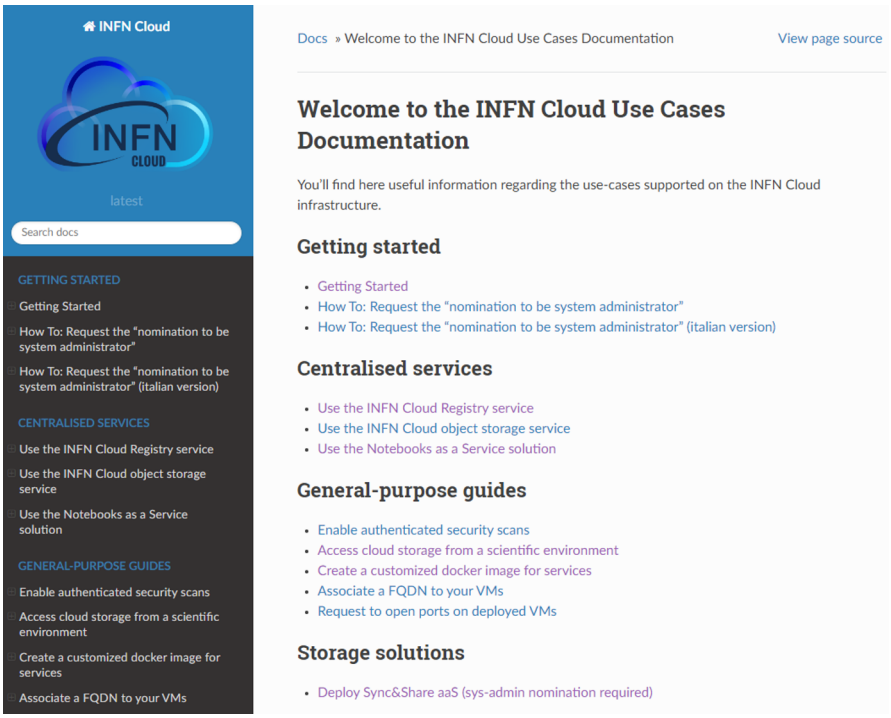

Figure 6. Screenshot of the INFN Cloud guide. On the left part of the general index, on the right an extract from a user guide.

guides made publicly available undergo a process of continuous review. Figure 6 shows screenshots related to the user guide: a screenshot of the general index is shown on the left side; an extract from the guide aimed at deploying the MATLAB Notebook PaaS service is shown on the right side.

The tool adopted for the public documentation is the well known Read the Docs[10] deployed as a Core service within INFN Cloud. Read the Docs make use of the RST (reStructuredText) markup language. The main advantage of using Read the Docs is that it can be integrated into a Git repository, allowing to take advantage of the typical Git tools such as version control and branches, and the possibility to use continuous integration automation.

6 Conclusions

Exploiting the PaaS paradigm allows researchers to focus less on the implementation details of their IT services and tools, and more on their research work. To achieve this important goal, the INFN has built, and is continually refining and expanding, the INFN Cloud. As already mentioned at the beginning of the article, the evolution of INFN Cloud is called DataCloud: an higher number of sites, providing one order of magnitude more computing resources with respect to the current setup, will join the federation; an expansion of computational and human resources for each site, also thanks to the help of European funding NRRP (National Recovery and Resilience Plan) will also take place; an increase in the number and size of projects, specifically linked to NRRP and having scientific-computing use-cases both

HPC (High-Performance Computing) and HTC (High-Throughput Computing), that want to take advantage of this infrastructure.

In the DataCloud project, not only does the Documentation, User Support, Communication and Training work package represent the meeting point between the scientific communities and the infrastructure managers, but also takes care of users and groups on-boarding, education and training. In order to make available to the users the infrastructure and the related cloud-native services to carry out research activities, the INFN has designed, deployed and is maintaining the INFN Cloud infrastructure and the related PaaS and SaaS services running on it. Part of the objectives are, in fact, also to facilitate and speed up the activities of those research communities making available new and reliable cloud-native technologies and give the needed support to use them in an effective way.

References

[1] CNAF
https://www.cnaf.infn.it/en/ - last seen Sept 2023
[2] GARR
https://www.garr.it/en/ - last seen Sept 2023
[3] ReCaS Bari
https://www.recas-bari.it/index.php/en/ - last seen Sept 2023
[4] INDIGO-IAM repository
https://github.com/indigo-iam/iam - last seen Sept 2023
[5] Jira Service Desk
https://www.atlassian.com/software/jira/service-management/features/service-desk - last seen Sept 2023
[6] Atlassian
https://www.atlassian.com/ - last seen Sept 2023
[7] INFN Cloud Service Desk
https://servicedesk.infn.it/servicedesk/customer/portal/50 - last seen Sept 2023
[8] INFN Cloud Training
https://www.cloud.infn.it/training/ - last seen Sept 2023
[9] INFN Cloud Guide
https://guides.cloud.infn.it/ - last seen Sept 2023
[10] Read the Docs
https://docs.readthedocs.io/en/stable/ - last seen Sept 2023