

EGI Conference 2024
30 Sept - 4 Oct 2024 - Lecce (Italy)

The evolution of INFN's Cloud Platform

**IMPROVEMENTS IN ORCHESTRATION
AND USER EXPERIENCE**

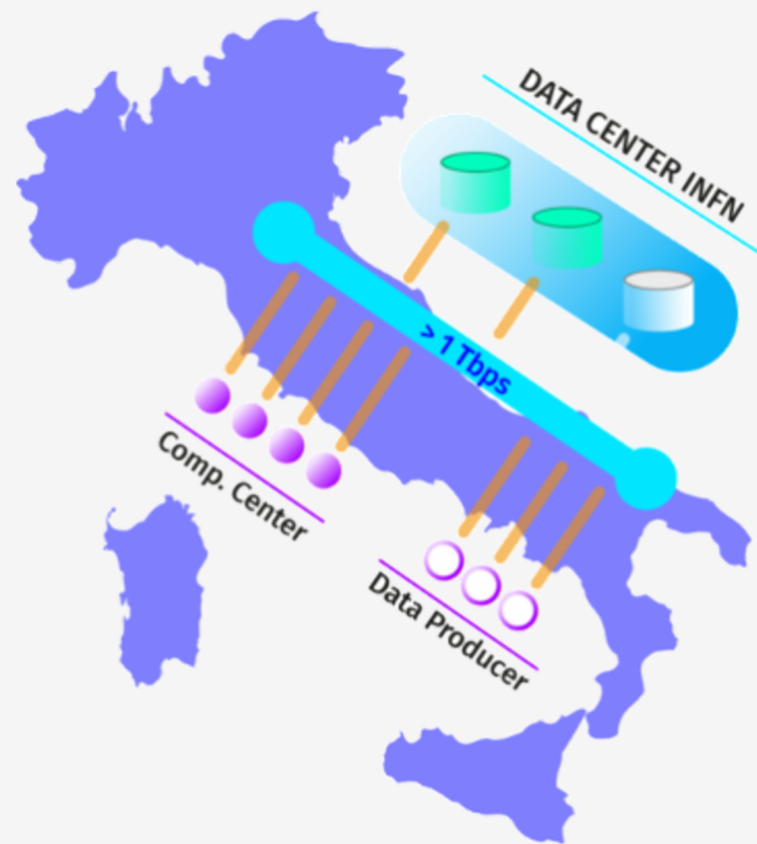
Marica Antonacci (INFN Bari)



INFN DataCloud MISSION

- Address the **challenges** faced by researchers in accessing and utilizing **distributed compute and storage resources**
- Offer a full set of **high-level services** to scientific communities
- Focus on **training** and on assisting user communities to migrate their use cases onto the cloud

INFN Cloud ARCHITECTURE



Federation

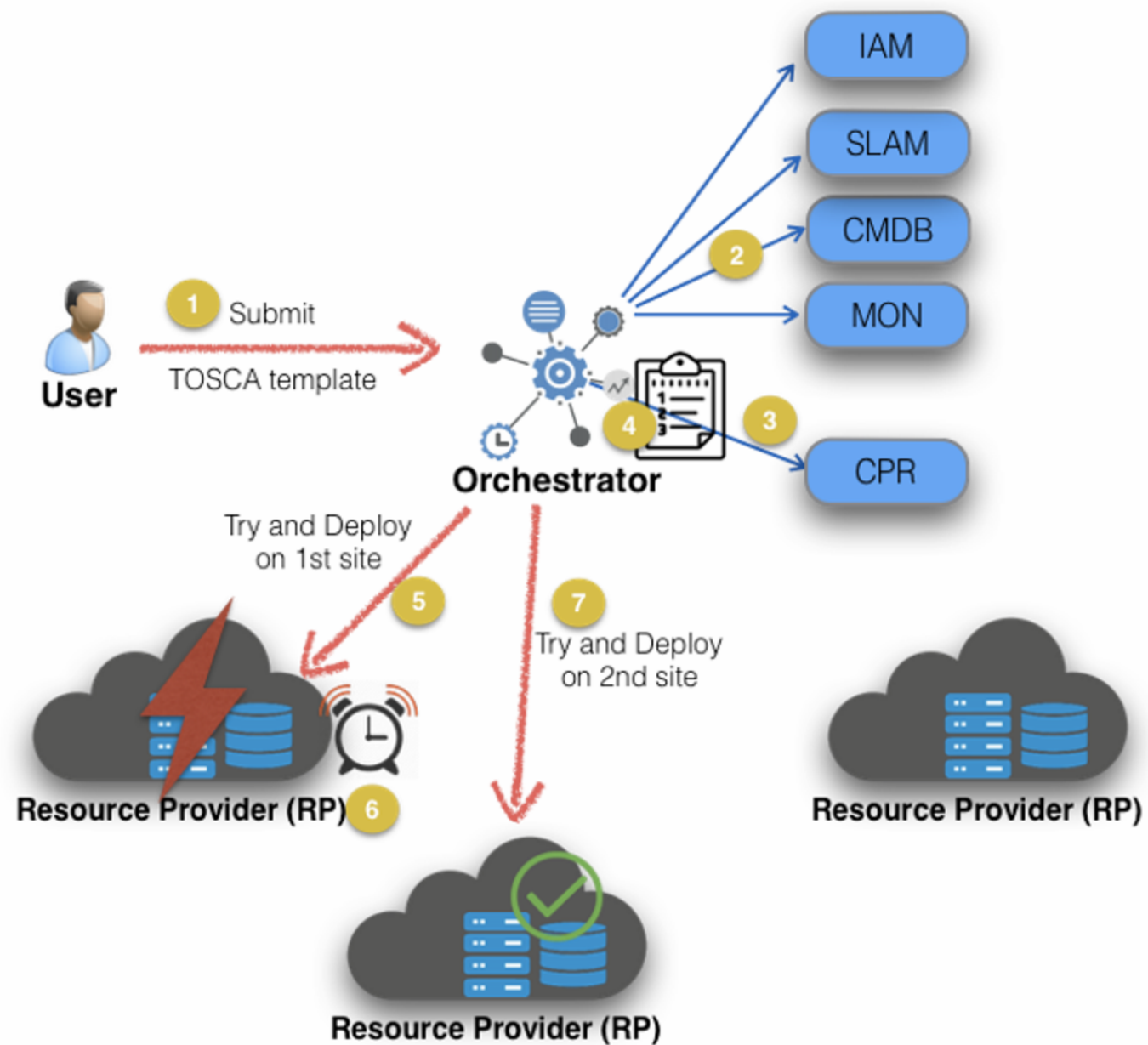
of existing Cloud Infrastructures for both compute and data

Dynamic Orchestration

of resources through the INDIGO PaaS Orchestration System

Consistent AuthN/AuthZ

across all Cloud layers using OpenID-Connect/OAuth2



THE FEDERATION MIDDLEWARE

The INFN Cloud platform's federation middleware is built on the **INDIGO PaaS Orchestration** system, a suite of open-source microservices that create an abstraction layer, providing seamless access to compute and storage resources across diverse and heterogeneous providers.



Infrastructure as Code paradigm

PLATFORM EVOLUTION OVERVIEW

INFN's cloud platform has continuously evolved to meet the increasing demands of scientific research. However, as the platform expanded, the **need for modernization** became evident, driven by three key factors:



Security

The platform must replace obsolete and vulnerable software components to address modern security challenges. This involves developing new, more secure components that can withstand emerging threats.



Robustness and Scalability

The effective management of a federation of heterogeneous and distributed providers and services requires modern and robust components that can efficiently scale as the number of users and services increases.



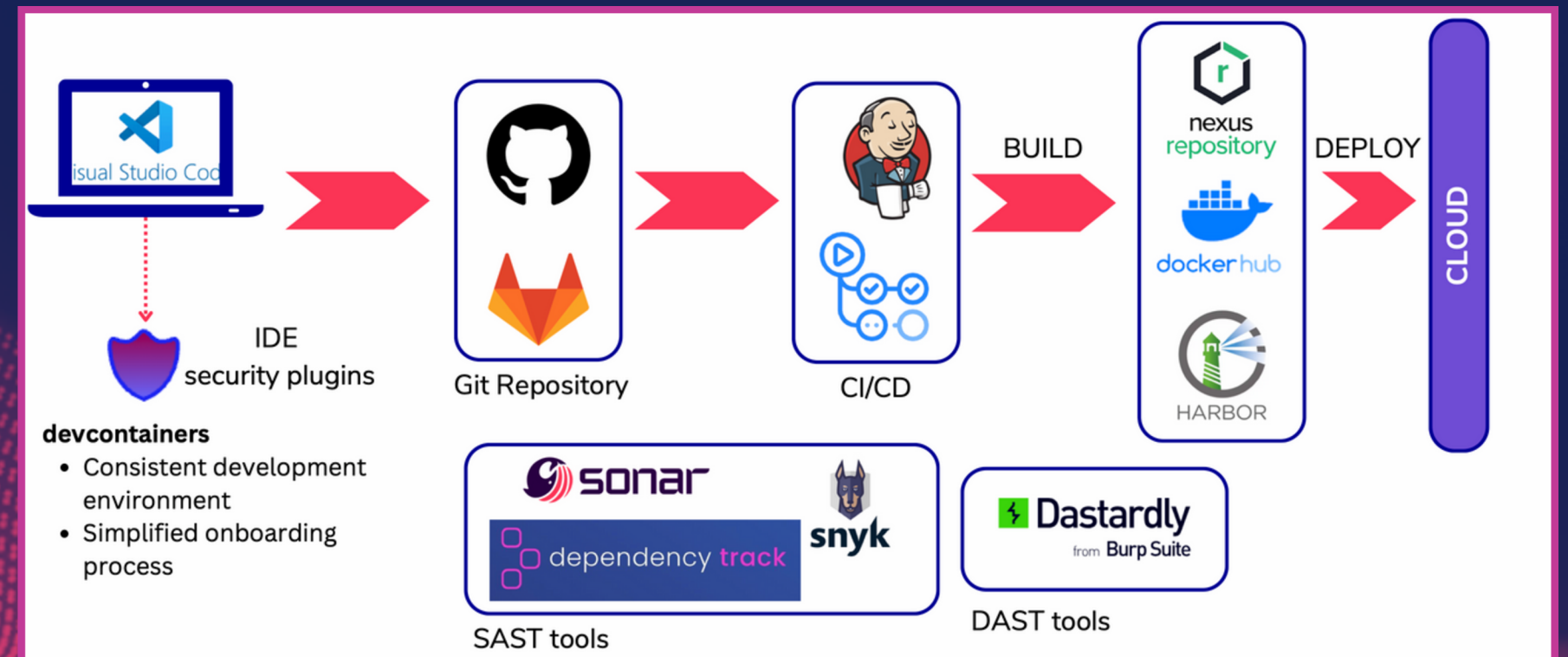
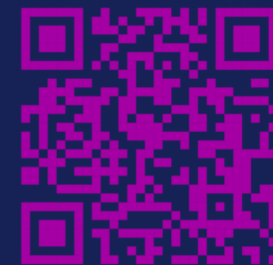
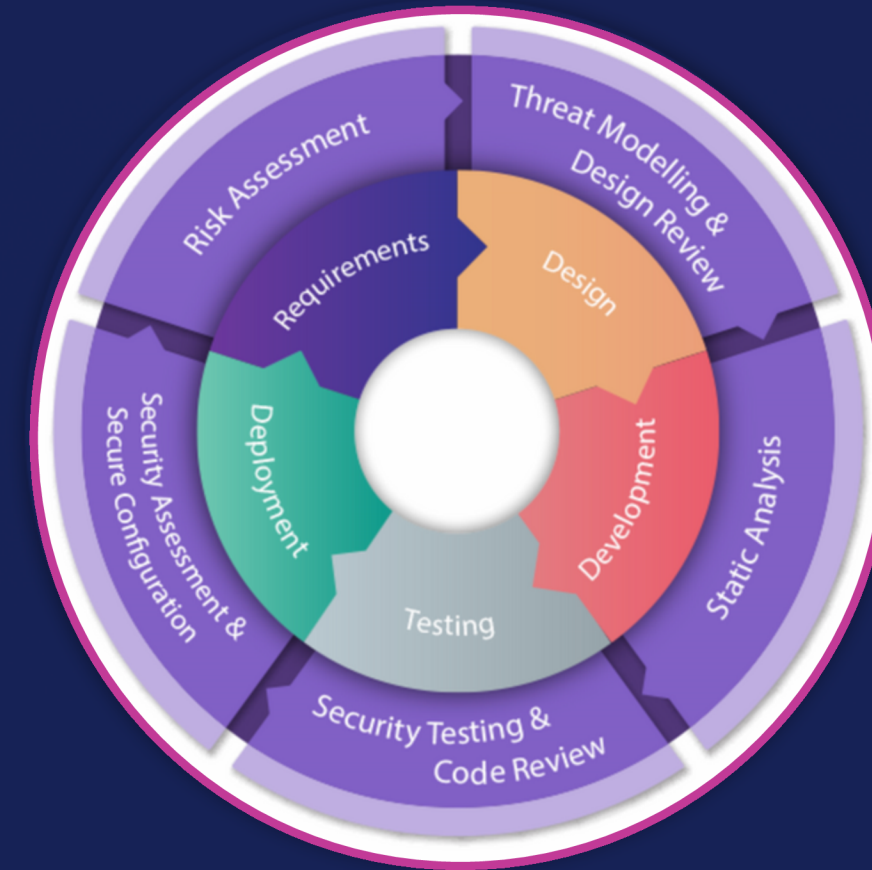
User Experience

It is crucial to provide an optimal user experience. Tools like the web dashboard simplify the deployment of complex services, hiding the intricate details of the underlying system and making it accessible to end users.

DevSecOps APPROACH

- We are dedicated to strengthening the **security posture** of the software developed within the DataCloud project.
- This commitment is driven by our extensive involvement in distributed computing infrastructures and various research initiatives, including projects handling **sensitive and confidential data**, such as those in the healthcare sector.
- We are transitioning from a DevOps to a **DevSecOps** approach, integrating automated security measures throughout the software development lifecycle.

SECURE SOFTWARE DEVELOPMENT CYCLE



FEDERATION REGISTRY

Within the PaaS Orchestration System architecture, a key micro-service is responsible for maintaining the **federation's information system**, which stores detailed data on resource providers and their capabilities. This system is crucial for enabling **matchmaking** between user deployment requests and available resources.

Current SYSTEM



- Relies on **outdated, unmaintained** components.
- Lacks **standardization in APIs**, creating inconsistencies and integration challenges with newer systems and services.



New SOLUTION

- A state-of-the-art web application using **FastAPI**, secured by **OpenID-Connect/OAuth2**.
- **neo4j** graph database replaces CouchDB, offering **improved data organization** and flexibility in handling diverse cloud providers.

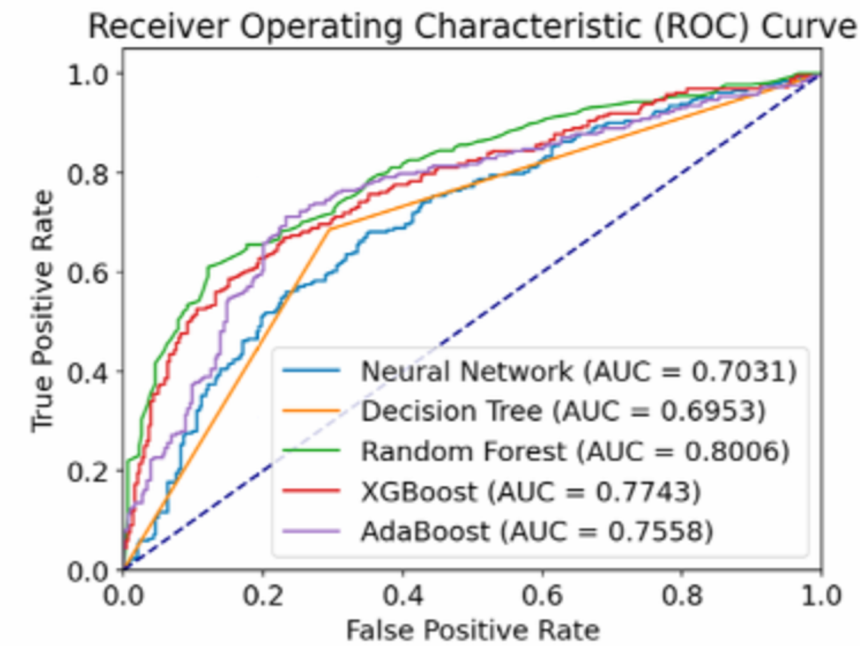
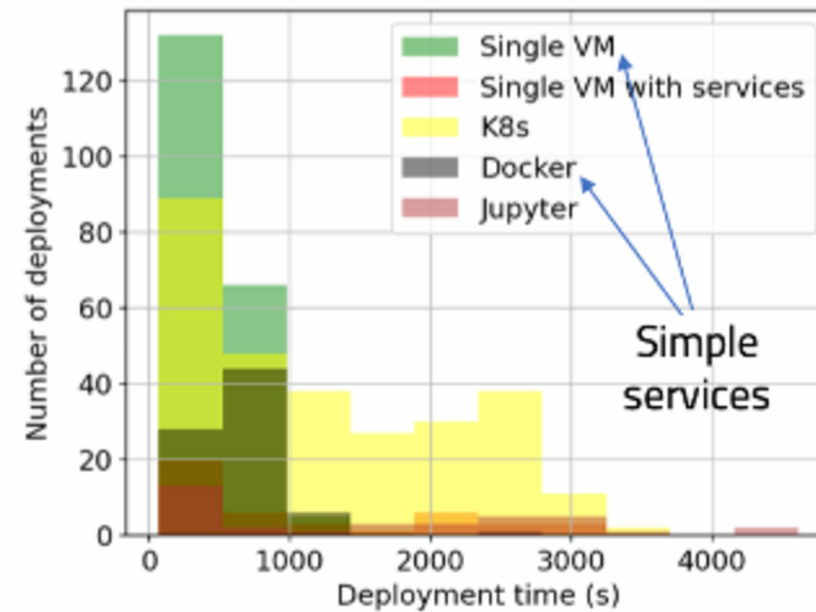
ADVANTAGES

- Better scalability and performance
- Simplified updates for new provider types
- Independence from outdated software
- Standardized APIs to ensure smooth integration with other services



Leveraging AI

FOR SMARTER ORCHESTRATION



Preparatory Work

- Identified significant metrics and data sources.
- Prepared datasets to thoroughly analyze the problem.

We are introducing AI-based techniques to **improve provider selection** and **resource allocation** in our orchestration system.

By incorporating artificial intelligence, we can make provider choices more **dynamic** and **efficient**, optimizing deployments across the cloud federation.

- **Predictive Model:** Forecasts deployment success by analyzing key metrics, improving provider selection.
- **Regression Model:** Predicts deployment times based on past data, optimizing scheduling and resource use.

Dashboard REDESIGN

The **INFN Cloud Dashboard** serves as the primary access point for both **on-demand** and **centralized services**, offering users a **graphical representation of TOSCA templates** to browse the service catalog and easily deploy available resources.

To make the platform more accessible and intuitive, we have implemented significant improvements to the user interface and overall usability.

The screenshot displays the 'ON-DEMAND SERVICES' section of the dashboard. A dark blue sidebar on the left contains navigation links: DEPLOYMENTS, ADVANCED, EXTERNAL LINKS, and ADMIN. The main content area lists six services, each with an icon, a title, a brief description, and a 'CONFIGURE' button:

- Virtual machine**: Launch a compute node getting the IP and SSH credentials to access via ssh.
- Docker compose**: Deploy a virtual machine with docker engine and docker compose pre-installed. Optionally run a docker compose file fetched from the specified URL.
- Run docker**: Run a docker container.
- INDIGO IAM as a Service**: The on-demand deployment service for the INDIGO IAM provides a quick and easy way for organizations to deploy their own instance of the INDIGO IAM, which is an open-source Identity and Access Management system.
- Elasticsearch and Kibana**: Deploy a virtual machine pre-configured with the Elasticsearch search and analytics engine and with Kibana for simple visualization of data with charts and graphs in Elasticsearch.
- Kubernetes cluster**: Deploy a single master Kubernetes cluster.

At the bottom of the sidebar, there are links for 'Settings' and 'Help', and a user profile for 'Marica Antonacci'.

The screenshot shows the 'SSH keys management' section of the dashboard. It features a user profile for 'Marica Antonacci' (admins/beta-testers) at the top. Below the profile, there is a section for 'SSH keys management' with a HashiCorp Vault logo. The text explains that SSH keys allow for a secure connection between a computer and a virtual server. There is a text input field for 'UPLOAD SSH PUBLIC KEY' and an 'UPLOAD' button. Below that, there is a section for 'CREATE NEW KEY PAIR' with a '+ CREATE NEW SSH KEY PAIR' button.

The screenshot displays the 'My deployments' section of the dashboard. It includes a 'Refresh' button and a 'New deployment +' button. Below these is a search bar and a 'Show 10 entries' dropdown. The main content is a table with the following data:

DESCRIPTION	DEPLOYMENT IDENTIFIER	STATUS	CREATION TIME	DEPLOYED AT	ACTIONS
test	11ef6f81-8f57-e13d-8583-22533e954eeb	CREATE_COMPLETE	2024-09-10 14:32:00	BACKBONE-CNAF	Details
test	11ef5e0b-0a6a-fe9b-b005-3a4acf66fdef	CREATE_COMPLETE	2024-08-19 09:11:00	BACKBONE-CNAF	Details
k8s	11ef2a2f-67b6-61e4-ad50-22533e954eeb	UPDATE_FAILED	2024-06-14 09:20:00	BACKBONE-CNAF	Details
docker build	11eefefe-ef53-f8f2-8be4-56fce75e0bfa	CREATE_COMPLETE	2024-04-20 10:15:00	RECAS-BARI	Details
k8s lisa	11ee9a9a-dfcd-c750-9c93-0242ec4ec63f	UPDATE_COMPLETE	2023-12-14 16:07:00	BACKBONE-BARI	Details

At the bottom, it shows 'Showing 1 to 5 of 5 entries' and navigation buttons for 'Previous', '1', and 'Next'.

Dashboard ☰

Kubernetes cluster STEP 2/3

DEPLOYMENT DESCRIPTION (0/50)

Description

CONFIGURATION ADVANCED

ADMIN TOKEN

Enter your password

Password token for accessing Grafana dashboard

NUMBER OF NODES

1

Number of K8s node VMs

PORTS

+ Add rule

Ports to open on the K8s master VM

MASTER FLAVOR

--Select--

Number of vCPUs and memory size of the K8s master VM

NODE FLAVOR

--Select--

Number of vCPUs and Memory Size of each K8s node VM

CONTINUE →

← Back

CANCEL ⓧ

Dashboard ☰

Kubernetes cluster STEP 3/3

✓ **CHECK DATA**

DEPLOYMENT DESCRIPTION: k8s

ADMIN TOKEN: ⓧ

MASTER FLAVOR: large: 4 VCPUs, 8 GB RAM

NODE FLAVOR: large: 4 VCPUs, 8 GB RAM

NUMBER OF NODES: 1

SUBMIT ↗

← Back

CANCEL ⓧ

Dashboard ☰

Virtual Nodes

Show 10 entries

NAME	STATUS
k8s-mast 39b1732e-7 fa163ed94b	STARTED
k8s-node 38a75a66-7 fa163ed94b	STARTED

Showing 1 to 2 of 2 entries

Previous **1** Next

Add Virtual Nodes ✕

⚠ Please note that the "NUMBER OF NODES" field specifies the total number of nodes you want in your cluster after the addition, not the number of new nodes to be created. For example, if you currently have 3 nodes and want to increase to 5, you should enter 5.

NUMBER OF NODES

1

Number of K8s node VMs

NODE FLAVOR

--Select--

Number of vCPUs and Memory Size of each K8s node VM

Force update

Trigger an update even if no changes are detected.

ADD +

CANCEL ⓧ

Manage Ports 11ef7bea-c2b9-b9f5-8ecf-0242c687447b (0fad0e6e-49a2-4cc8-ba09-a5dd65209f92) ← Back **Add Port** +

Show 10 entries Search:

DIRECTION	ETHER TYPE	IP PROTOCOL	PORT RANGE	REMOTE IP PREFIX	DESCRIPTION	ACTIONS
Ingress	IPv4	UDP	Any	-	-	Delete
Ingress	IPv4	TCP	Any	-	-	Delete
Ingress	IPv4	TCP	22	0.0.0.0/0	-	Delete
Ingress	IPv4	TCP	8443	0.0.0.0/0	-	Delete
Egress	IPv6	Any	Any	-	-	
Egress	IPv4	Any	Any	-	-	

Showing 1 to 6 of 6 entries

Add Port ✕

► How to add a port? Brief explanation

RULE*

Custom TCP Rule

DESCRIPTION (0/255)

DIRECTION*

Ingress

OPEN PORT*

Port

PORT*

80

CIDR*


0.0.0.0/0

CANCEL ⓧ **ADD PORT** +

Dashboard NEW FUNCTIONALITIES


Service Catalog

Fully-managed centralised services




INFN Cloud object storage 🔖
the centrally managed service based on Ceph Rados-Gateway

[GO TO SERVICE →](#)




INFN Cloud Registry 🔖
The centrally managed service INFN Cloud Registry, based on Harbor

[GO TO SERVICE →](#)



INFN-Cloud monitoring 🔖
the INFN-Cloud monitoring service


[GO TO SERVICE →](#)



Notebooks as a Service (NaaS) 🔖
Jupyter Notebooks as a Service
Scope: admins/beta-testers


[GO TO SERVICE →](#)

Self-managed on-demand services




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

[CONFIGURE →](#)




Docker compose 🔖 🔖
Deploy a virtual machine with docker engine and docker compose pre-installed. Optionally run a docker compose file fetched from the specified URL.

[CONFIGURE →](#)




Run docker 🔖 🔖
Run a docker container

[CONFIGURE →](#)




INDIGO IAM as a Service 🔖 🔖
The on-demand deployment service for the INDIGO IAM provides a quick and easy way for organizations to deploy their own instance of the INDIGO IAM, which is an open-source Identity and Access Management system...

[CONFIGURE →](#)




Elasticsearch and Kibana 🔖 🔖
Deploy a virtual machine pre-configured with the Elasticsearch search and analytics engine and with Kibana for simple visualization of data with charts and graphs in Elasticsearch

[CONFIGURE →](#)




Kubernetes cluster 🔖 🔖
Deploy a single master Kubernetes cluster

[CONFIGURE →](#)




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

[CONFIGURE →](#)




HTCondor mini 🔖 🔖
Deploy HTCondor mini, a technology preview of an all-in-one ("minicondor") HTCondor. This type of install is useful for testing and experimentation.

[CONFIGURE →](#)




HTCondor cluster 🔖 🔖
Deploy a complete HTCondor cluster

[CONFIGURE →](#)




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

[CONFIGURE →](#)



Jupyter + Matlab (with persistence for Notebooks) 🔖 🔖
Run Jupyter on a single VM enabling Notebooks persistence and Matlab integration

[CONFIGURE →](#)



Sync&ShareaaS 🔖 🔖
The INFN-Cloud Sync&ShareaaS is based on popular storage solutions such as ownCloud¹ and Nextcloud². INFN-Cloud users have full control over the configuration parameters of their Cloud...

[CONFIGURE →](#)

CONCLUSIONS

INFN's cloud platform has undergone significant improvements, driven by the need for enhanced security, scalability, and user experience:

- **Replacing obsolete components** with modern, scalable solutions, such as the Federation-Registry, ensures robustness and long-term sustainability.
- The adoption of **AI-driven** models for provider selection and resource optimization has improved **orchestration** efficiency and reduced failures
- A focus on the **user experience**, with a revamped dashboard and simplified service deployment, makes the platform more accessible and secure.

THANK YOU

Do you have any question?



cloud-support@infn.it



www.cloud.infn.it

