



Finanziato
dall'Unione europea
NextGenerationEU



D3.2.5

TIER-1 SYSTEM OPENED FOR PRE-PRODUCTION

FIRST RELEASE

JUNE 2025

Mirko Cestari, Giuseppa Muscianisi (CINECA)
Giorgio Bolzon, Stefano Salon (OGS)

Eccetto dove diversamente specificato, il contenuto di questo sito è
concesso in licenza con Attribuzione Creative Commons - 4.0 internazionale.



Executive Summary

This document reports on state-of-the-art of the services available to users and provided through the PRACE-Italy Tier-1 HPC system (GAIA). While the system is still to date considered in pre-production and not open to all the project requests (for example via ICSC portal¹, ISCRA calls² and other resource access channels), most of the services that will be provided are already being documented³ or can be easily described.

¹ <https://www.supercomputing-icsc.it/en/icsc-home/>

² <https://www.hpc.cineca.it/hpc-access/access-cineca-resources/iscra-projects/>

³ https://docs.hpc.cineca.it/cloud/general/general_info.html

Table of Content

1. TIER-1 SYSTEM OBJECTIVES	1
1.1. Introduction	1
1.2. High-Performance Computing Cloud Model	3
1.3. Main use cases	4
2. TIER-1 SYSTEM CLOUD SERVICES	6
2.1. OpenStack Overview	6
2.2. Cloud Specifics	8
2.3. ISO27001 best practices	8
2.4. Tutorials and repositories	9
3. CONCLUSION	9

List of Figures

Figure 1 - High-level architecture of GAIA.	2
Figure 2 - Monitoring of the montly network trafic of the Tecnapol data centre. the maximum network I/O registered is well below the total capacity of 100+100 Gb/s.	5
Figure 3 - High-level OpenStack architecture.	6
Figure 4 - The OpenStack Landscape	7

1. TIER-1 SYSTEM OBJECTIVES

1.1. Introduction

The system high level architecture is reported in **Errore. L'origine riferimento non è stata trovata.**

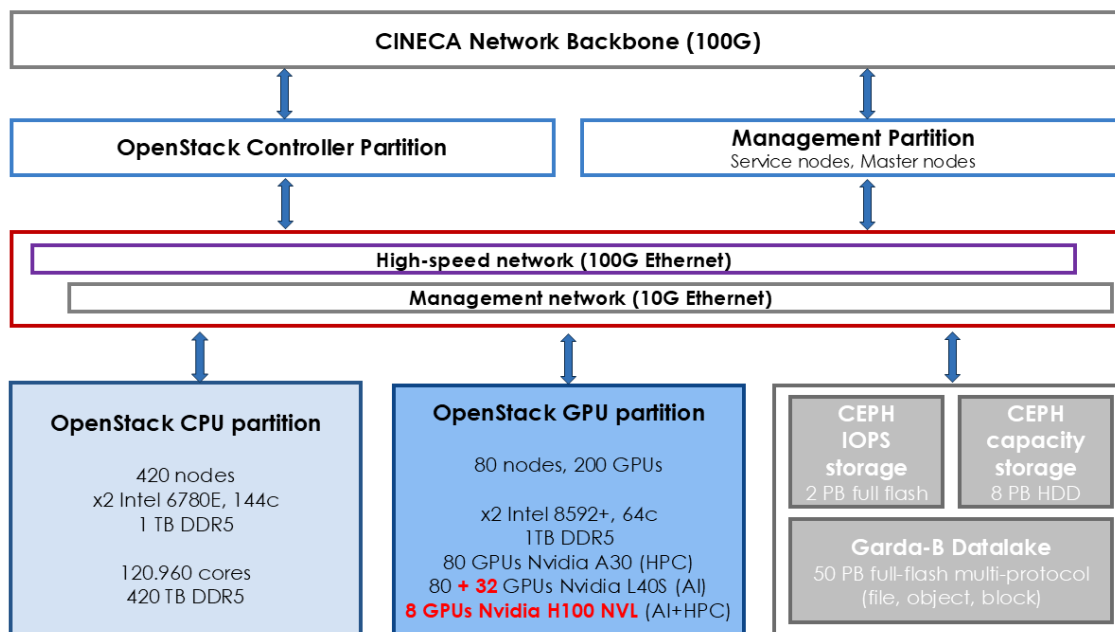


Figure 1 - High-level architecture of GAIA.

The CPU partition features standard dual socket server nodes (2x CPUs) based on Intel Xeon 6780E, for a total of 288 physical cores per node and 1 TB DDR5 RAM 6400MT/s.

The GPU partition features CPUs Intel Xeon 8592+, for a total of 128 cores and 1 TB DDR5 RAM 4800MT/s. Moreover, this partition offers high-end NVIDIA GPUs both for AI and 3d rendering workloads (L40s) and for scientific use cases requiring double precision FP64 arithmetic (A30 and H100). Thanks to the contract modifications, more GPU capacity was acquired – highlighted in red in Figure 1 – for a total of:

- 80 GPUs NVIDIA A30
- 112 GPUs NVIDIA L40s
- 8 GPUS NVIDIA H100 NVL

The storage is composed of two main tiers:

- CEPH
 - CEPH high ops partition, based on full flash solution
 - CEPH high capacity, based on HDD drives.
- GARDA-B data lake
 - Full flash, multi-protocol solution

While the CEPH tier will be mainly serving the virtual infrastructure (OpenStack cloud) by providing block storage for virtual machines and storage for internal services, the data lake will be transversal to the cloud and the other supercomputing infrastructure thanks to the NFS and S3 protocols. The name "GARDA" is a reference for the largest lake in Italy and "B" stands for Bologna, since an equivalent infrastructure will be installed in the CINECA data centre of Naples to provide a redundant storage infrastructure.

The high-speed data network is based on 100Gbps/400Gbs Dell appliances and is deployed through a specific software defined network application to manage all the network devices. The management network is used for system deployment and configuration (OS provisioning, bare metal configuration, and so on) and is also featuring Dell appliances.

1.2. High-Performance Computing Cloud Model

The Tier-1 HPC system just presented is well positioned to be featured as a HPC, or High-Performance Computing, cloud system since it integrates high-performance computing resources and capabilities with cloud computing infrastructure. It combines the computational power and scalability of traditional HPC systems with the flexibility and on-demand nature of cloud services.

HPC Cloud users can exploit the computational power of HPC systems with the added benefit of extreme flexibility. Users can organize their pool of computational resources how they see fit for their personal workflow, creating a single virtual machine that harnesses the full capability of all the resources at once, or smaller virtual machines that can communicate and interact with each other.

More in details, with an HPC Cloud system the users can:

- decide how to exploit computational and storage resources, creating their personal virtual infrastructure (**flexibility**)
- manage their virtual machines, for example by deciding the operative system and software environments (**flexibility**)
- scale the computational resources based on their needs allowing to handle varying workloads efficiently (**scalability**)
- optimize the use of the resources without idle time (**resources optimization**)
- access the resources remotely (**accessibility**)

Cloud computing resources can be provided using different service models:

- **Infrastructure as a Service (IaaS):** Provides virtualized computing resources over the internet. Users can deploy those resources how they see fit.
- **Platform as a Service (PaaS):** Offers a platform allowing developers to build, deploy, and manage applications without worrying about the underlying infrastructure.
- **Software as a Service (SaaS):** Delivers software applications over the internet, accessible via a web browser, without the need for local installation or maintenance.

Following the extensive experience gained with ADA cloud infrastructure in CINECA⁴, at the beginning it is planned to provide resources using the IaaS model⁵.

⁴ <https://docs.hpc.cineca.it/cloud/systems/ada.html#ada-card>

⁵ https://docs.hpc.cineca.it/cloud/general/cineca_cloud_model.html#cineca-hpc-cloud-model

1.3. Main use cases

The main objectives and use cases that can be hosted on the Tier-1 HPC cloud infrastructure are listed as follows.

General-Purpose Infrastructure-as-a-Service

The cloud must be able to host data processing and analysis services, supporting typical IaaS use cases. The service must accommodate use cases requiring a combination of performance and flexibility. Cloud users are free to install and use any software according to the Acceptable Use Policies of CINECA.

Cloud's native support for automation (Infrastructure-as-code) will be adopted. CINECA hosts an internal Gitlab service⁶, which can be used to create platform deployments driven by CI/CD⁷ where the status of the IT platform and their installation are always tracked and versioned, for maximum reliability and flexibility. In practice, git software versioning is leveraged to support collaborative work on IT infrastructure deployment and evolution adopting standards such as Ansible and Terraform/OpenTofu. The goal is to write all the necessary code to automate the installation of the IT platforms and govern its status relying on the integration of the git repository tool (in this case gitlab) and its CI/CD services with the Cloud APIs (in this case OpenStack). This allows to standardize IT platform deployment and evolution, while providing templates that can be easily extended by simply cloning a git repository.

Windows software images are not provided, due to licensing restrictions.

An important requirement for general purpose IaaS is to support instance backup services, which should either happen automatically or triggered on-demand by the user.

HPC Mini Clusters

The cloud must be able to host HPC mini-clusters with adequate performance. These should include the following features:

- A filesystem shared by the login and compute nodes of the cluster.
- Provisioning of GPU compute resources.
- Complete middleware software stack of numerical libraries and communication libraries.
- High performance networking, including native RDMA capabilities.
- Placement of VMs. Note, directing the placement of VMs to specific hosts would normally require admin privileges.

Hosted Web Services

Hosted web services might include virtual lab portals and user-friendly platforms for interactive dataset presentation and exploration. As an example, the Tier-1 HPC system would host platforms for use cases such as weather and ocean forecast maps and genomics data.

⁶ <https://gitlab.hpc.cineca.it>

⁷ CI/CD means Continuous Integration and Continuous Deployment.

Data management services should be accessible from both internet and intranet and capable of receiving and transmitting data.

Currently, the data centre Internet connectivity (via GARR, the national research and education network) is 100+100 Gb/s. this bandwidth can be further extended reaching up to 1 Tb/s. The current network activity can be monitored here: https://gins.garr.it/home_statistics.php, and produces plots of this kind:

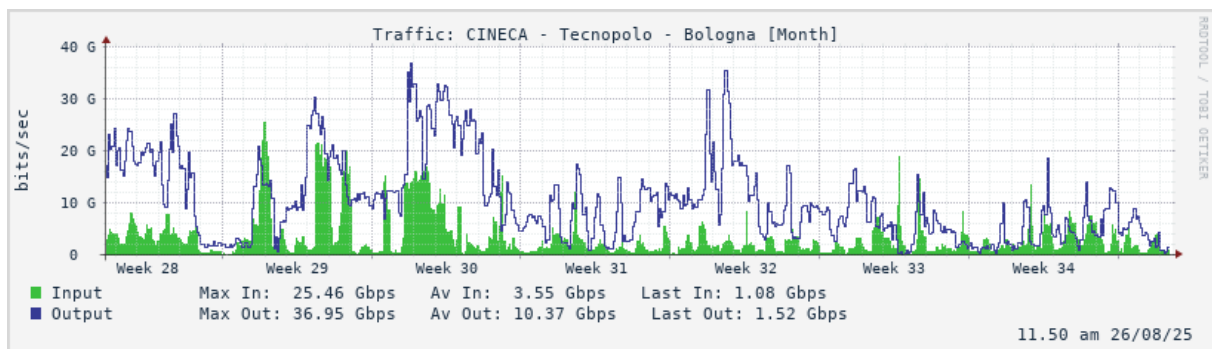


Figure 2 - Monitoring of the monthly network traffic of the Tecnopol data centre. the maximum network I/O registered is well below the total capacity of 100+100 Gb/s.

Bridging HPC infrastructure

Instances hosted on the cloud should be able to act as frontends to workloads running on the other HPC systems present in the CINECA ecosystem.

Frontend interfaces could include Jupyter, Dask and Slurm.

Trusted Research Environments

The cloud system would support the hosting of research workloads processing sensitive and controlled-access data. This will require the infrastructure implementation to meet key security measures, including:

1. Encrypted communication on the cloud system control plane networks.
2. Support for encrypted storage volumes.
3. Support for encrypted filesystems (both in transit and at rest).

Kubernetes and Containerised Workflows

The Tier-1 HPC system would support the flexible and automated deployment via Kubernetes of containerised workflows. The users can deploy Kubernetes clusters using both the orchestrator provided by the system or by a custom ad-hoc fashion way, using tools such as Rancher, RKE2, OKD and Juju Charms.

2. TIER-1 SYSTEM CLOUD SERVICES

2.1. OpenStack Overview

The Tier-1 HPC infrastructure will be based on OpenStack⁸, an open-source cloud operating system that controls large pools of computing, storage and networking resources through a data centre, all managed and provisioned through API with common authentication mechanisms, see the high-level schema in Figure 3.

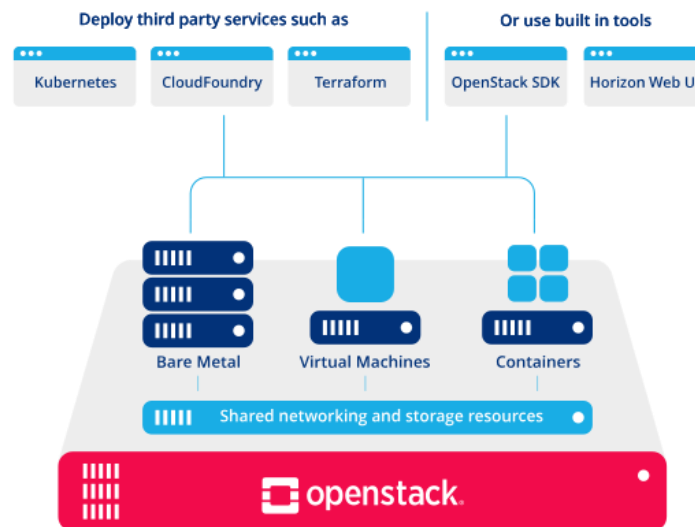


Figure 3 - High-level OpenStack architecture⁹

OpenStack is a flexible tool that has a modular structure: it is split up in different components that can be installed and activated independently to address every user's need. Figure 3 shows the high-level overview of the OpenStack landscape that allows to see both the role of each service and how they work together.

⁸ <https://www.openstack.org/>

⁹ <https://www.openstack.org/software/>

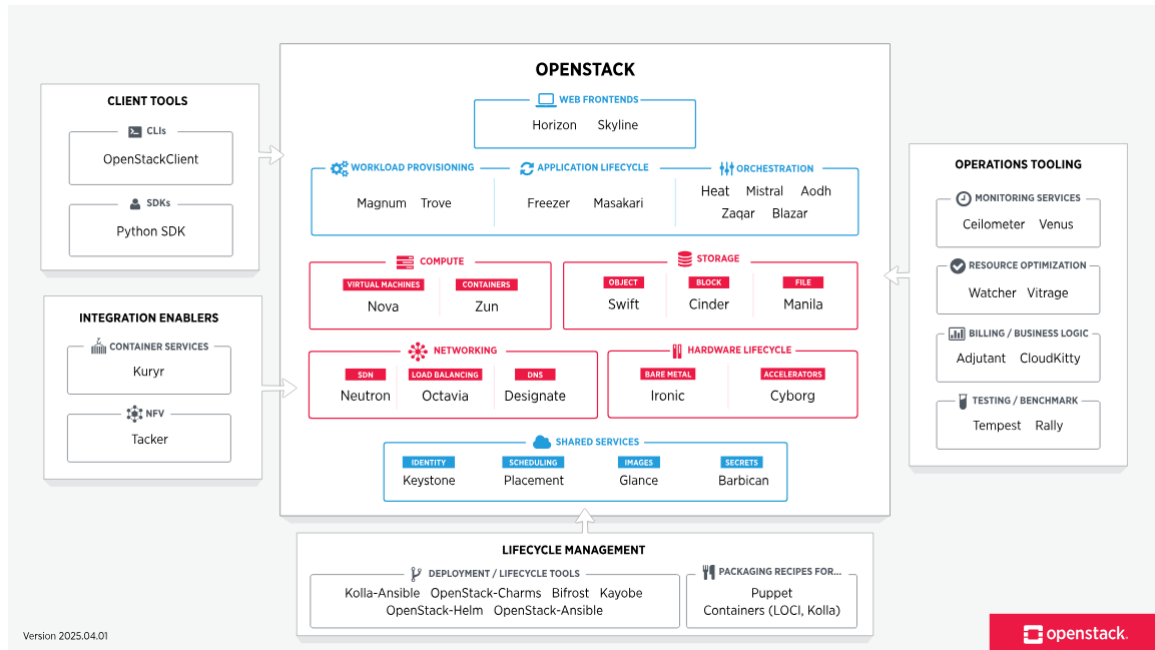


Figure 4 - The OpenStack Landscape¹⁰

To address the use cases presented previously, the core services fundamental for OpenStack to function are all available in CINECA cloud infrastructures:

- Nova: it handles the CPU virtualization and the creation and management of virtual machines;
- Cinder: it handles block storage and volumes;
- Neutron: it handles networking within the resources of a project and with external networks;
- Glance: it handles virtual machine images and snapshots;
- Horizon: it handles the OpenStack Dashboard, which provides a web-based user interface to OpenStack;
- Keystone: it handles identity, providing authentication and authorization services for all OpenStack services.

Aside from the above-mentioned core services, on the Tier-1 HPC system additional services are available:

- Barbican, Key Manager Service to support data encryption use cases
- Manila, Shared File System Service among the virtual machines
- Octavia, for load balancer as a service (LBaaS) deployment
- Trove, for database as a service (DBaaS) management

Additional services are also available:

- Kayobe and Bifrost, for the deployment and lifecycle of the system,

¹⁰ <https://www.openstack.org/software/>

- Rally, Adjutant and CloudKitty for billing, testing and benchmarking purposes.

To manage the resources assigned to a project, the users can interact with OpenStack using in-house tools like Horizon Web Dashboard and the Command Line Interface (CLI) or using Infrastructure as a Code tools. The users can interact with the Tier-1 HPC cloud infrastructure by the Horizon dashboard or by the OpenStack command line with the CINECA IdP configured for authentication. Infrastructure-as-Code (IaC) can also be used to automate the deployment of cloud resources with the use of tools like Terraform and Ansible.

2.2. Cloud Specifics

In the Tier-1 HPC system both the test infrastructure and the production are deployed. The latter partition is the only available to end users. The test infrastructure is a subsystem of the production infrastructure, it is accessible only for CINECA staff and it is intended for testing and validation purposes, as new features or bug-fix are checked before realising them in the production.

Different Linux OS cloud-images are provided to the users, both Debian-based as Ubuntu and RHEL-based as Rocky Linux. Anyway, the users are free to add their own custom images, or different versions of the previous one, on the bases of their needs and requests.

Different pools of flavour are also defined, as a combination of virtual CPUs, RAM and root disk, to instantiate the VMs on the different hypervisors present in the system, based on the use-cases to address. So, there are flavours dedicated to VMs able to execute:

- CPU-only computation, hosted on the named "OpenStack CPU-partition"
- CPU-only computations requiring a huge amount of RAM but not the use of the GPU accelerators, hosted on the named "OpenStack GPU-partition"
- GPUs computation, hosted on the named "OpenStack GPU-partition", able to access the Nvidia GPU L40s, A30 or H100nvl via PCI passthrough. The flavours able to leverage the L40s are intended for data visualisation and rendering, those for the A30 and H100NVL are intended for offloading numerically-intensive floating-point computation.

2.3. ISO27001 best practices

Additionally, the Tier-1 HPC infrastructure design must comply with ISO27001 best practices, mainly:

- Overall system's log centralized collection (currently on the cluster's management dedicated storage, accessible only by the administrator). Six months retention is foreseen, moreover log files must be backed-up.
- OpenStack log centralized collection. Currently two months "online" retained on the OpenSearch dedicated nodes storage. Daily "index segments" dumped and backed up to be retained 6 months and analysed (in case) offline in separated cluster
- Possibility to trace and retain every single project's activity, especially access logs, to be able to provide them to the tenants, on-demand.

- Infrastructure (as a code) backup: resiliency
- Infrastructure as a code change tracking and change proofing Auditing.
- Systems compromise detection (runtime integrity checking and measurement).
- Access control and restriction:
 - network access restriction (ssh ACLS, firewall, L2 segregation)
 - Admin only users to cluster compute and admin nodes (LDAP, ssh ACLS, PAM)
 - OpenStack API network access control (Firewall)
 - OpenStack API L7 firewall analysis
 - Lateral movement assessment and restriction inside cluster
 - Support users access restriction (bastion/VPN, dedicated host/VM, 2FA).
- Vulnerability assessment: Tools and Patching policies
- Projects backup policies: Tools to automate Volumes backup

2.4. Tutorials and repositories

Different step-by-step tutorials and GitLab repositories are available to the users¹¹, to allow them to both easily learn how to use the system through the Horizon dashboard and automatically deploy and setup the granted resources in their own OpenStack project.

The step-by-step tutorials concern mainly in the usage of the Horizon dashboard.

The GitLab repositories are a collection of Terraform/OpenTofu/Ansible scripts allowing, as an example, to deploy VMs leveraging or not the GPUs, or to deploy a RKE2 Kubernetes cluster, or set up a HPC Slurm mini-cluster infrastructure.

Furthermore, in the frame of WP5 training activities, a webinar for the TeRABIT users will be organized as soon as the system will be available to users (in September 2025).

3. CONCLUSION

The main hardware and software features of the PRACE-Italy Tier-1 system (GAIA) was presented. Their combinations allow to provide a rich and fully featured HPC Cloud system. After the review of the HPC Cloud model, the main use cases that can be hosted on the system were presented. The cloud operating software chosen to allow virtualization in the Tier-1 HPC infrastructure is OpenStack. A brief explanation of its main features was exposed together with few specifics for the Tier-1 HPC cluster. It was described how the users can access the granted resources and how they can easily and automatically deploy and set-up them to address their own needs and use cases.

More information, practical guides and tutorials can be found the documentation repository:

https://docs.hpc.cineca.it/cloud/general/general_info.html

¹¹ https://docs.hpc.cineca.it/cloud/tutorials/index_tutorials_and_repos.html