Finanziato dall'Unione europea
NextGenerationEU

D2.4.2

DELIVERY AND TESTING OF USER AUTONOMOUS MANAGEMENT ENDPOINTS

FEBRUARY 2026

Fabio Farina (GARR)
Massimo Carboni (GARR)
Claudia Battista (GARR)

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

Executive Summary

The activity has successfully integrated automation and observability to manage complex multi-technology domains scientific use cases, reaching 100% implementation for domain-specific orchestration workflows. The optical domain is managed via a specialized Workflow Orchestrator (WFO), while the packet domain utilizes the PACMAN framework to reduce massive component update times from months to days. This infrastructure is supported by a comprehensive Observability Portal that provides real-time telemetry and inventory data.
This framework is extensible and can assist a user utilizing multiple management domains. It has been validated in the federated use case UrgentSHAKE, targeting real-time seismology simulations, using the three RIs and with the collaboration of ICSC.

Table of Content

1 THE GARR-T AUTOMATION 4
2 STATUS 6
2.1 Observability Portal 6
2.2 Automation and Orchestration Portals 8
2.3 Federated observability for UrgentSHAKE 10
3 STATUS AND FURTHER DEVELOPMENT BEYOND THE PROJECT 14
4 REFERENCES 15

List of Figures

Figure 1 – Optical network automation use-case 5
Figure 2 – Packet network automation use-case 5
Figure 3 - Autonomous management endpoints high-level architecture block diagram for GARR-T 6
Figure 4 - Landing page for monitoring information 7
Figure 5 – Traffic summary dashboards 7
Figure 6 – Dashboard detail with summarised information on power consumption 8
Figure 7 – PACMAN configuration snippet 10
Figure 8 – WFO user interface with active optical resources 10
Figure 9: high level view of UrgentSHAKE use of the RIs and ICSC 11
Figure 10 – Details on telemetry probes layout for SHAKE federated monitoring 12
Figure 11 – UrgentSHAKE monitoring dashboard pilot job detail 13

List of Tables

Table 1 – Development progress 14

1 THE GARR-T AUTOMATION

The GARR-T network integrates automation as a key component to manage the increasing complexity of multi-technology domain and federated scientific use cases. By reducing the need for human intervention, GARR-T aligns with modern ICT operational standards, ensuring technological continuity and efficient management of the upgrade of the network infrastructure. The main automation driver is the need for controlling network services life-cycle with their monitoring, including telemetry, logging, and alerting.
Modern network operations rely on automation and observability to manage the complexity of today's infrastructures and distributed services. Automation enhances dynamic provisioning and uniform configuration updates, mainly as a result of DevOps methodologies. Complementing this, observability provides deep insight into system health through the correlation of logs, metrics, and traces, allowing for proactive anomaly detection and reduced resolution times. This shift addresses several needs:

- Efficiency and Reliability: Automation moves from slow, error-prone manual management to faster configurations and proactive fault correction to reduce downtime.
- Scalability and Security: the management of large-scale networks through uniform rule application and DevOps-inspired methodologies ensures consistency across all devices.
- Deep Insight: Observability goes beyond simple monitoring by correlating logs, metrics, and traces, allowing for early anomaly detection and a lower Mean Time To Repair (MTTR).

These pillars enable orchestration, which offer advantages as

- Speed and Autonomy: Users can directly request or modify network resources without waiting for manual intervention.
- Simplification: Intuitive interfaces hide the underlying infrastructure complexity.
- Error Reduction: Predefined workflows prevent common manual configuration mistakes.
- Control and Governance: All actions remain regulated by strict authorization policies.

To streamline the development of autonomous management endpoints, the automation activity focused on two guiding scenarios on the connection of HPC and Clouds.

Use-Case 1: Optical Network Automation

This scenario focused on the lifecycle management of point-to-point optical circuits. Authorized operators have been able to request for new services, change transport capacity for existing ones, and where needed decommission circuits. All by using a unique workflow management tool, paired with an additional endpoint to real-time monitoring data and automated alerts reporting when the actual service status deviates from the expected operational parameters.
Figure 1 shows how a purely optical network can be configured between HPC nodes.

Use-Case 2: Packet Network Automation

Mirroring the optical scenario, this use-case enables operators to agile instantiation of Layer 3 network services using a native DevOps strategy. Operators now have the ability to deploy overlay networks on top of the packet infrastructure, monitor resource health in real-time, and perform massive network configuration updates.
Figure 2 shows how a packet-based network can be configured between HPC nodes.

Figure 3 shows the block diagram of the software stack for autonomous management endpoints as implemented in GARR-T.
The information sources of all underlying infrastructures are channelled into the observability pillar. Here, the data are aggregated and structured to create a unified view of the overall homogeneous current state, providing a common data source for automatic device inventory, real-time telemetry, logging correlation, and alerting propagation. The programmability and orchestration blocks use the policies and requests to identify and correct any deviations between the observed and the expected states.

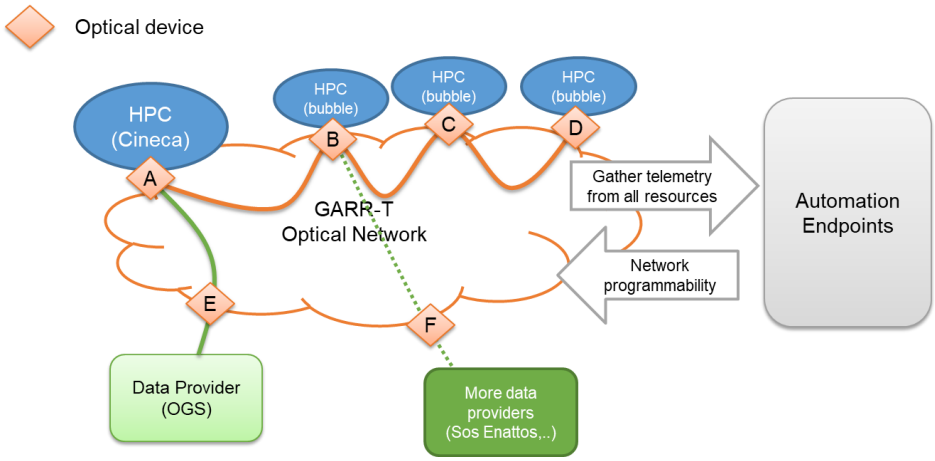


Figure 1 – Optical network automation use-case

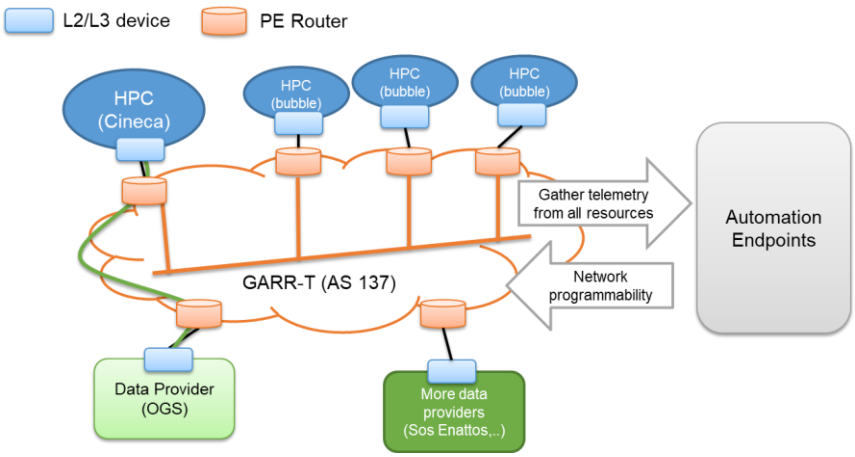


Figure 2 – Packet network automation use-case

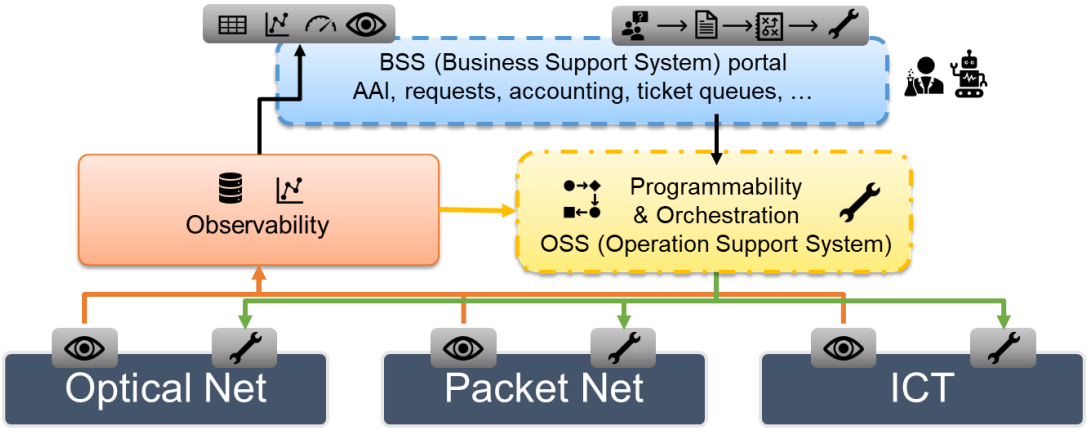


Figure 3 - Autonomous management endpoints high-level architecture block diagram for GARR-T

2 STATUS

2.1 Observability Portal

The observability portal has been implemented integrating several tools, each selected on data nature and update frequency. Netbox is the chosen open-source asset management, providing a daily-updated representation of physical and organizational resources. A TIG stack (Telegraf, InfluxDB, and Grafana) used to collect, process, and visualize streaming telemetry and metrics. Additionally, OpenSearch and ClickHouse are integrated as data lakes to correlate logs and traffic flows. Finally, access to this data was secured via federated authentication to Grafana, the core of the Observability portal, allowing for differentiated visibility levels tailored to the needs of infrastructure operators and general research users. In the initial deployment phase, the Netbox inventory was activated as the authoritative "Single Source of Truth," utilizing custom scripts and agents for automatic resource discovery. Netbox is configured to provide REST and GraphQL APIs, allowing other components of this portal to consume data regarding sites, devices, and inventory without duplication. Simultaneously, the telemetry infrastructure was established by deploying ClickHouse and OpenSearch data lakes as containerized workloads within a Kubernetes cluster to manage logs and network flows. Time-series databases are also containerized to handle device metrics and reporting aggregates, with resource allocation specifically sized to ensure the sustainable processing of high-volume incoming data streams.

The Observability portal is available at https://stats.garr.net. Some high-level status dashboards are publicly available with no need of authentication, more detailed views are protected with federated authentication and RBAC authorization, based on users' home organization. Here GARR users can access more than 40 dashboards providing information on the status of their services, collecting the views according to different perspectives to support the roles of the users. Categories include: high-level views on the services' traffic for planning and reporting purposes, low-level details on the devices for services' sanity assessment, power consumption information aggregated

on different criteria to support sustainability measurements.

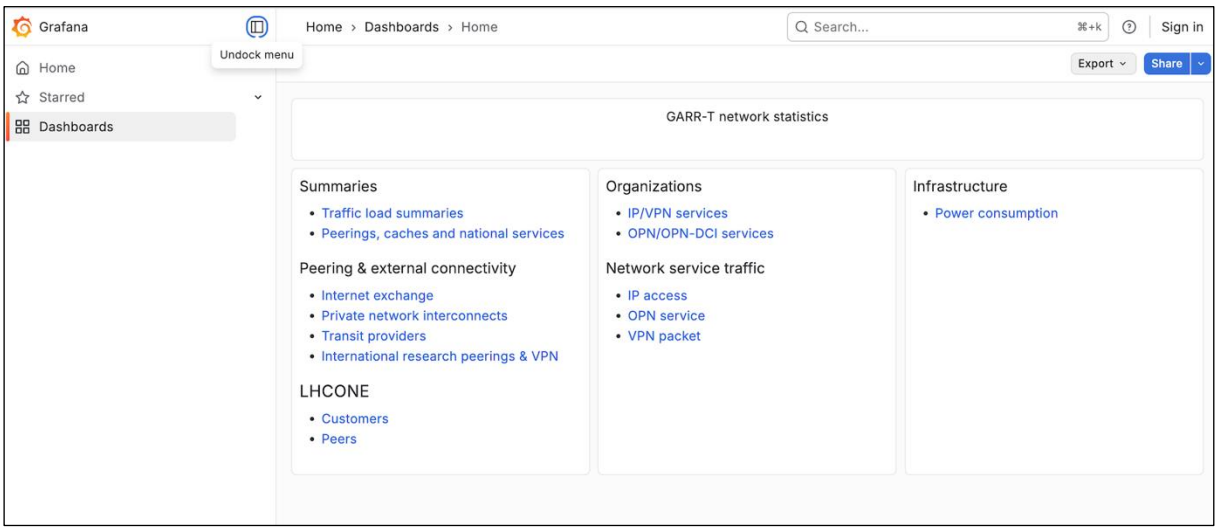


Figure 4 - Landing page for monitoring information



Figure 5 – Traffic summary dashboards



Figure 6 – Dashboard detail with summarised information on power consumption

2.2 Automation and Orchestration Portals

Programmability has also been implemented through the integration of available open-source tools, supplemented by specific solutions developed by the Research Network community and project-specific tools to fulfil the use-case requirements. More specifically, distinct automation tools were integrated for each technology subdomain to overcome the complexity of inherently different resource representation between the optical and the packet domains. Optical network automation leveraged device APIs to program both services and device operating system upgrades through a workflow engine. Packet network leveraged the PACMAN framework—an Ansible-based IT automation tool specifically designed for configuring GARR-T routers and IP services.

Packet network configuration automation

The PACMAN automation framework, based on the Ansible, has been developed to abstract and standardise the network resources configuration, letting technical users to schedule the automation for the creation, the change and the decommissioning of IP services on the backbone routers. Also the upgrade of monolithic and modular devices is supported, simplifying and accelerating wide management campaigns. Thanks to PACMAN automation, the time needed for massive routers updates have been reduced from months to days. GARR NOC adopted PACMAN for on-boarding and configuring the TeRABIT devices in a smooth, standard, homogeneous activity.

PACMAN is a code-based solution, close to the so-called GitOps philosophy where changes in the network are mediated through code repositories, and only after approval Continuous Integration and Continuous Deployment automation pipelines apply the new configuration on the network. Under this philosophy, NOC operators interact with PACMAN and the network by adding specifications and policies on the internal Git repositories that host the services' definitions. Then the continuous integration / continuous delivery chains alter the running the configurations on the devices.

For urgent intervention PACMAN supports also a more user-friendly interface through a layer called AWX. AWX is an open-source service that integrates Ansible playbooks with workflows, generating GUIs for parameters insertion and automation processes triggering. AWX transparently exposes workflows through REST APIs. PACMAN and AWX have shown In this way, automation workflows become usable by other services without the need to develop custom code. Interactions on AWX are mediated by authentication and authorization, and workflows can also introduce verification steps that limit the execution only after authorizations, before changes are applied to the infrastructures.

Figure 7 reports a PACMAN configuration fragment showing how complex network setup can be minimised.

Optical network service automation

GARR decided to specialize the usage of the WorkFlow Orchestrator (WFO), supported by the GÉANT European backbone, for the operation of optical transport devices and optical network services. The WorkFlow Orchestrator was originally developed by the Dutch and American research and Education networks, SURF and ESnet. Today WFO is the management platform for most of the operation in GÉANT. WFO is a framework that offers libraries to create orchestration processes, automatically generates GUIs and APIs starting from the descriptions of the modelled resources data structures. Resources can be composite, depending on other lower-level resources. WFO actions are protected via authentication and authorization, making actions executables only to users with the right roles. WFO exposes REST API that are callable from third party software for further integration. Each entity modelled in WFO is paired sets of actions defining the workflows that implement the change of configurations on the devices. These workflows that GARR developed on WFO enable the creation, the inspection, the modification and the decommissioning of optical light-paths and transport services on the GARR-T and TeRABIT devices. The adoption of WFO enabled the prompt activation of all the light-paths the build the traffic matrix among the TeRABIT infrastructures.

backbone.yml 619 B

```yaml
bb_intf:
  ae24:
    bb_ip_v4: 185.191.180.58/31
    bb_ip_v6: 2001:760:ffff:ffaa:0:0:180:58/127
    description: rs1.bo01 - [Lepida]
    lacp:
      xe-0/0/2:0: ae24 - link 1 - rs1.bo01 - [Lepida]
      xe-0/0/2:1: ae24 - link 2 - rs1.bo01 - [Lepida]
    target_address: 185.191.180.59
  et-0/0/0:
    bb_ip_v4: 185.191.180.60/31
    bb_ip_v6: 2001:760:ffff:ffaa:0:0:180:60/127
    description: rs1.bo01 - [f22c1]
    target_address: 185.191.180.61
  xe-0/0/2:3:
    bb_ip_v4: 185.191.181.202/31
    bb_ip_v6: 2001:760:ffff:ffbb:0:0:181:202/127
    description: rl2.b003 - [locale]
    target_address: 185.191.181.203
```

Figure 7 – PACMAN configuration snippet

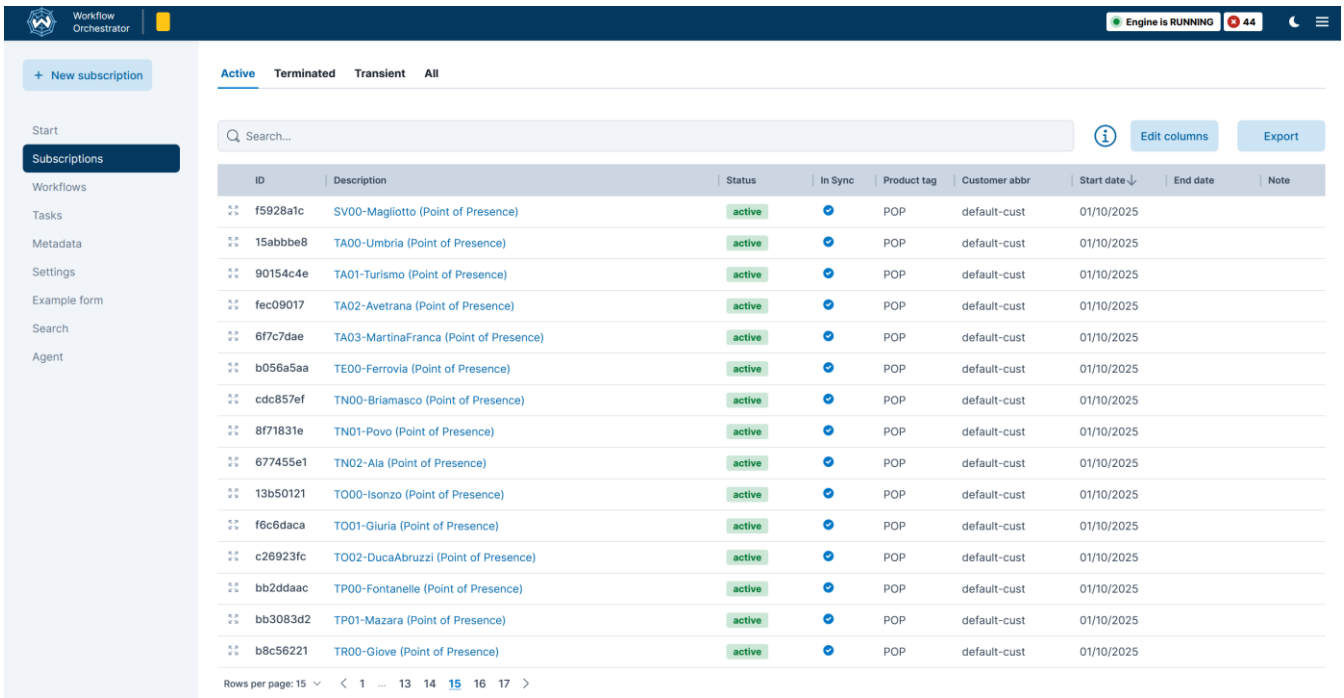


Figure 8 – WFO user interface with active optical resources

2.3 Federated observability for UrgentSHAKE

The UrgentSHAKE [1][2] use case has been identified as the best candidate to validate the federated multi-domain observability model. UrgentSHAKE needs to urgently start HPC workflows to run seismology computations to assess damage in the north-east Italian regions, after an event is detected. The system received data from a large set of sensors in the regions. UrgentSHAKE runs multi-node parallel code chains, with data staging actions. The simulation and analysis flows require

first results in less than one hours and more detailed ones in few hours to match the requirements for an effective search and rescue action.
To immediately gather access to HPC resources, the system will submit in parallel to all the available HPC systems (PRACE (Galileo100), EuroHPC node Leonardo and at the ReCAS INFN HPC Bubbles). Access to Leonardo has been temporary agreed for developments with the ICSC projects
The executions are triggered from a dedicated virtual machine connected to the HPC queueing systems. This entry point is the best candidate for locating the federated measurement probes of the execution status. In addition, the VM is a demarcation point between each infrastructure operator point of view, that should remain constrained to the facilities, and the use-case user point of view that can safely be propagated to external federated data lakes with no risks of exposing sensible information. From a user perspective, it is essential to be aware that the overall chain of resources needed to complete the urgent processing, is always responsive and available.
A high-level view of the UrgentSHAKE use of the HPC systems is in Figure 9.

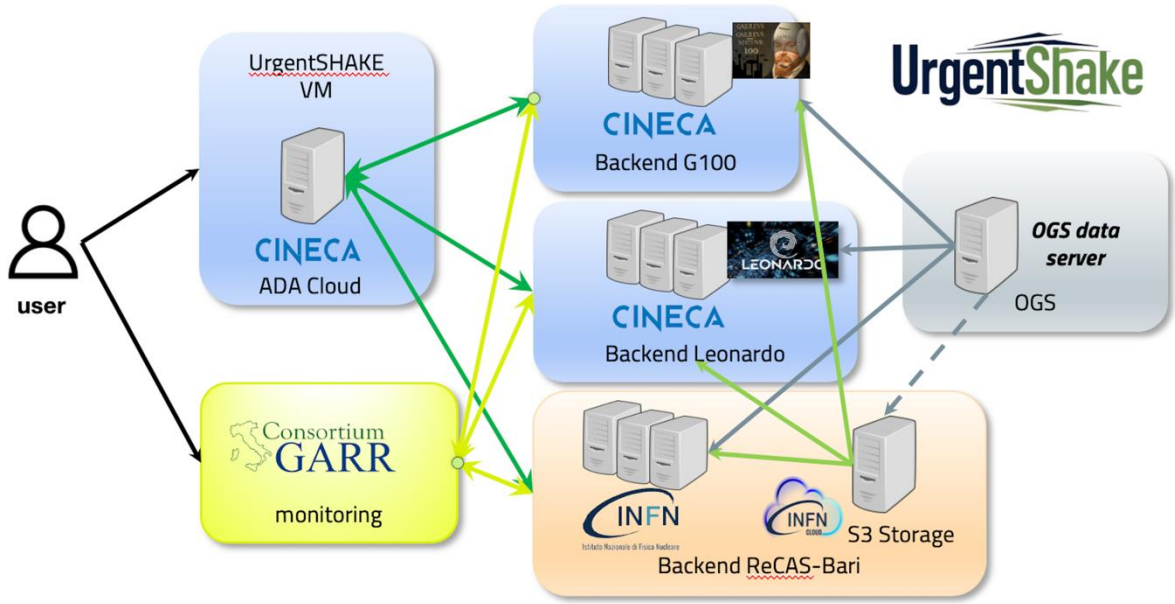


Figure 9: high level view of UrgentSHAKE use of the RIs and ICSC

Under these requirements, a data flow for the federated monitoring has been implemented as a three steps process:

- Information gathering and normalization - standardization block of the shared data format
- Data persistence and processing block
- Query, visualization, alarming block

The high-level architecture for the UrgentSHAKE observability use-case is shown in Figure 10.

Only relevant metrics for researchers have been collected by developing ad hoc probe

configurations. Custom Python scripts have been introduced where the available information was not available in natively machine-consumable formats, producing JSON structures that the data lakes can consume.

The collected measurement series have been organised as dashboard in the Observability portal. The following Figure 11 reports the metrics collected from Galileo100 and Leonardo HPC clusters.

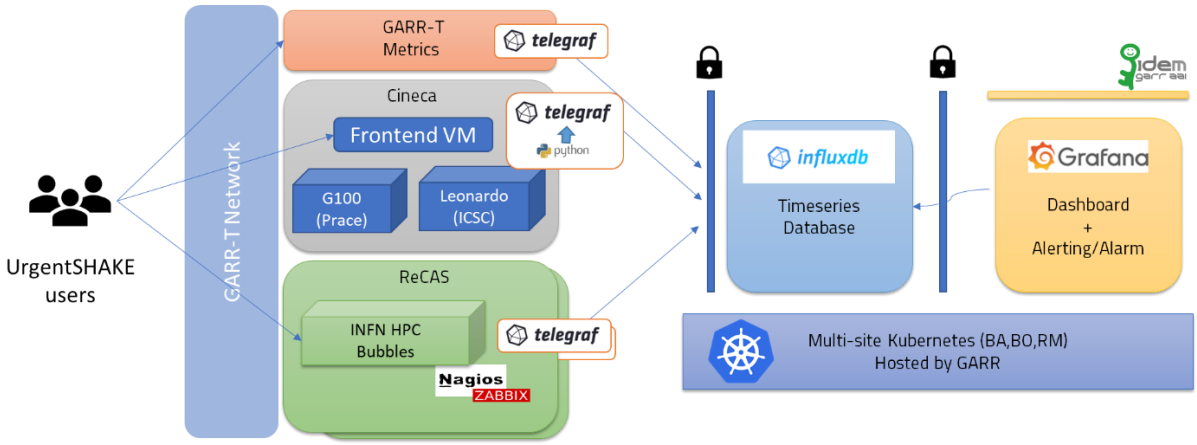


Figure 10 – Details on telemetry probes layout for SHAKE federated monitoring

The identified essential minimum metrics are:

- Network metrics
  - Bidirectional reachability between the UrgentSHAKE Virtual Machine and the InfluxDB data lakes
- Infrastructure metrics
  - File system status as a Red-Green synoptic showing that storage location gathering the simulation data are readable
  - Status of the HPC job submission queues, reporting #queued jobs vs #running jobs. This ensures that the computing facilities complete short test executions as expected, being ready to process real urgent requests when needed
- Application metrics
  - Verify SSH login from the use-case virtual machine to the job submission backends
  - Successful submission to queue systems as standard, unprivileged user
    Track completed jobs execution time and their return code, again to ensure that the use-case software can run as expected

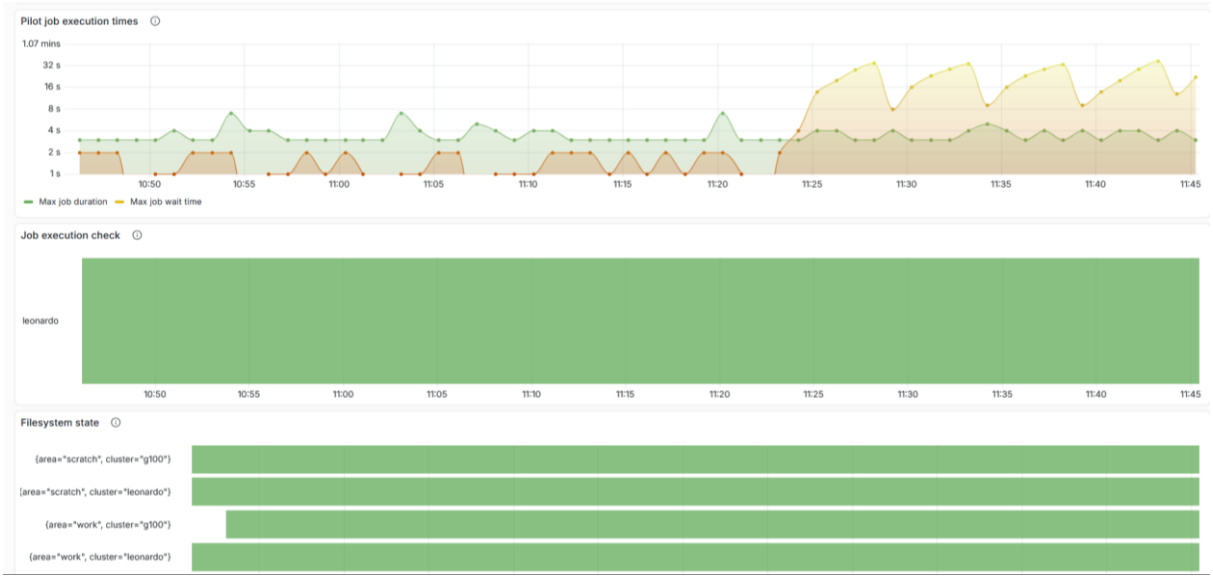


Figure 11 – UrgentSHAKE monitoring dashboard pilot job detail

3 STATUS AND FURTHER DEVELOPMENT BEYOND THE PROJECT

Progress on the implementation of autonomous management endpoints is reported in Table 1.

Component development	Previous Deliverable	Current Status
Orchestration Workflows: Software tools for nodes programmability	68%	100%
Observability User-Portal: automatic inventory, streaming telemetry and unified logging	93%	100%
Provisioning & Management Portal	37%	50%

Table 1 – Development progress

The objective for creating endpoints for Observability and Automation has been achieved, with separate orchestration tools for the optical and packet networking service. The ambitious goal of creating a unified provisioning and management portal for the non-technical end-users is under testing phase, using a realistic use case. The unification effort requires a deeper integration of the business aspects among the infrastructures involved in the project, which is not mature yet and requires further testing.

Converging domain-specific orchestrators into a unified BSS/WFO architecture requires a sophisticated abstraction layer to harmonize the disparate data models of the optical and packet layers. The primary challenge lies in bridging the gap between physical-layer constraints (wavelengths, attenuation) and logical-layer dynamics (routing, multiplexing) across multi-institution domains, transforming a raw infrastructure into a single, programmable ecosystem. Furthermore, implementing FCAPS at the network slice level demands granular visibility and strict isolation for OPN and VPN services within a shared infrastructure. The complexity resides in automating high-level business requests via a self-service portal, while maintaining real-time path computation and cross-layer event correlation, ensuring that increased agility does not compromise the carrier-grade reliability of the GARR network. Although the objective of delivering a fully integrated BSS orchestration self-service portal has not yet been reached, the ongoing results in use-case modelling, operational workflows, and the supporting data model contribute to a significant progress.

GARR will pursue the implementation effort after the end of the TeRABIT project. The development strategy is to keep moving towards a unified management architecture for the infrastructure. The next steps will focus on:

- Orchestration Convergence: GARR will continue internal development to converge both optical and data domains into a single WFO-based orchestration tool.

- Business Support System Integration & FCAPS: Following the redefinition of the GARR Information System (BSS), FCAPS functionalities will be will introduced at the network slice level.
- Advanced Service Delivery: This evolution will enable the provisioning of Private Network Services—specifically OPN (Optical Private Networks) and VPN (Virtual Private Networks)—within the GARR infrastructure.

These developments will facilitate offering services to users and to other Research Infrastructures, with the technical capacity to offer the service more adequate to users' needs.

4 REFERENCES

[1] Zuccolo, E., Bolzon, G., Pitari, F., Rodríguez Muñoz, L., Monsalvo Franco, I. E., Scaini, C., Poggi, V., Smerzini, C., and Salon, S.: UrgentSHAKE: An HPC System for Near Real-Time Physics-Based Ground Shaking Simulations to Support Emergency Response Efforts , EGU General Assembly 2025, Vienna, Austria, 27 Apr–2 May 2025, EGU25-9353, https://doi.org/10.5194/egusphere-egu25-9353, 2025. https://ricerca.ogs.it/handle/20.500.14083/42843

[2] Elisa Zuccolo, Giorgio Bolzon, Fabio Pitari, Lucía Rodríguez Muñoz, Chiara Scaini, Manuela Vanini, Valerio Poggi, Stefano Salon; Advancing Rapid Response to Earthquakes with Tiered Physics-Based Ground-Shaking Simulations: The UrgentSHAKE System. *Seismological Research Letters* 2025;; 96 (5): 2995–3011. doi: https://doi.org/10.1785/0220240472

[3] GEANT: the European Research and Education backbone http://www.geant.org

[4] ESnet: the Department of Energy's dedicated science network https://www.es.net