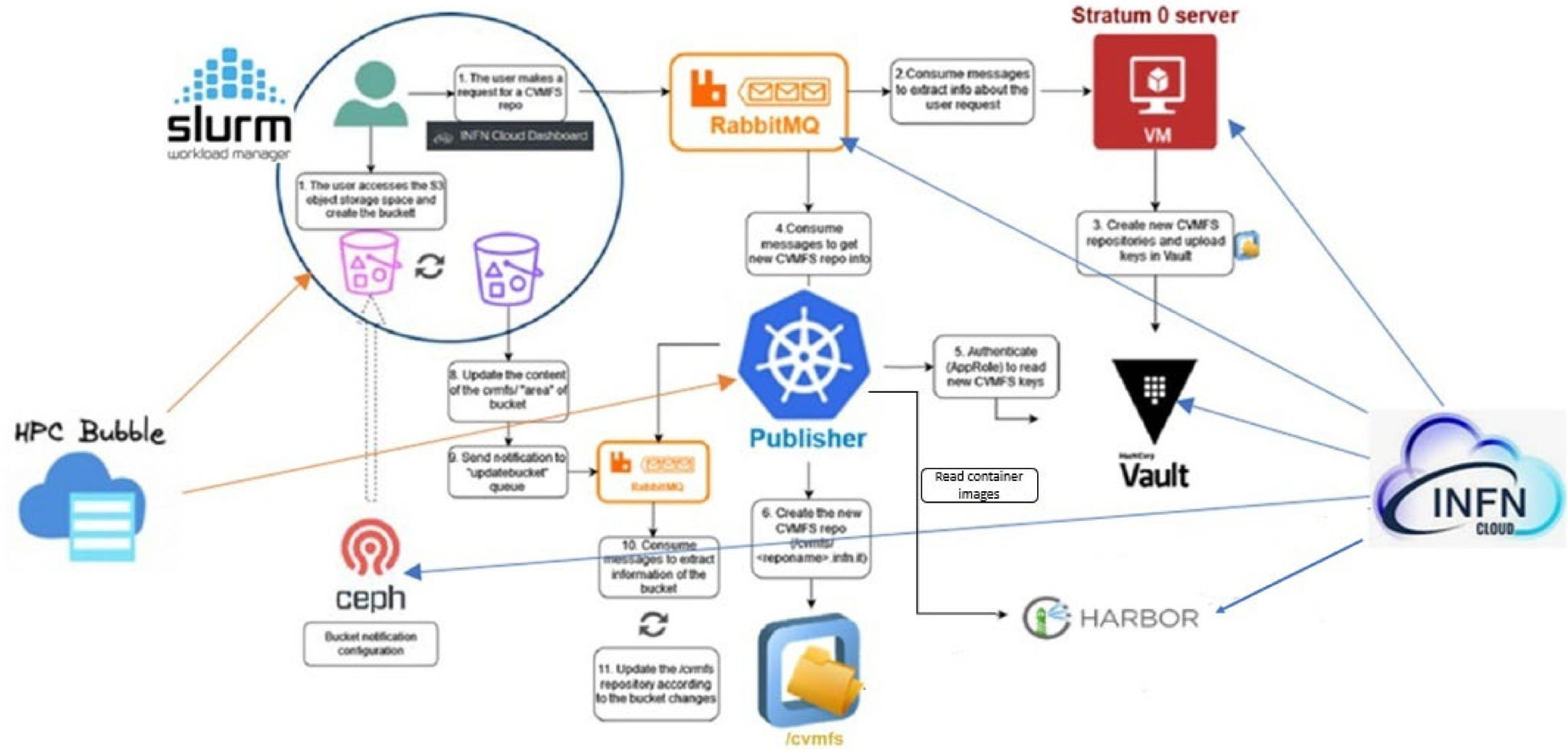
Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

Software Distribution on INFN DataCloud: a TERABIT use case.

Francesca Del Corso – INFN Sez. Perugia

The project focused on the development of a cloud-native, open-source software management service using the INFN DataCloud on-premises infrastructure. The goal is to create a scalable and resilient system that can efficiently distribute user and group software in high-demand environments, while remaining easy to manage and transparent for end users.



Workflow overview

- The user **requests** a CVMFS repository (personal or group) via the **INFN Cloud dashboard**.
- The request is send to **RabbitMQ** and is elaborated in order to create the repository.
- Once created, the relative **keys** are published in a **Vault system.**
- The user accesses the **S3 object storage** space and creates a **bucket** (personal or group).
- He **uploads** what he want to **distribute** in a specific area of the bucket named *cvmfs.*
- The S3 bucket service system sends a message to RabbitMQ so that the system get **notified** and can **synchronize** the content of the correspondent CVMFS repository.
- At this point, the user can access the **CVMFS client** in **read** mode to the **distributed** software.
- The user can **populate** his repository by installing a **CVMFS publisher** by using his repository keys.

Objectives and added value

The project leverages TERABIT resources to simulate large-scale and real-world scenarios. Stress tests involve high data transfer volumes and thousands of concurrent synchronization requests between S3-backed buckets and CVMFS repositories.

The **INFN HPC Bubble** in Bari could provide the ideal TERABIT infrastructure:

- to scale as needed
- to ensuring the availability and core INFN DataCloud services accessibility (CVMFS server, Vault, RabbitMQ, RadosGW, Harbor).

A **Kubernetes** instance inside the HPC bubble will run the Docker containers deployed for the synchronization app,

A **Slurm** workload manager inside the HPC bubble will run jobs simulating the S3 user operations.