EGI Cloud Compute

Introducing EGI Federation and EGI Cloud Compute

Baptiste Grenier

Service Delivery and Information Security officer

2024/06/04

FranceGrilles technical workshop

TLP: WHITE Public

Outline

- About EGI: resources providers, users and service
- A look at EGI Cloud Compute
- The French contribution: LETHE project as an example
- Joining Cloud Compute: benefits and requirements
- Contributing to and integrating with EOSC

Section 1

About EGI

Vision

All researchers have seamless access to services, resources and expertise to collaborate and conduct world-class research and innovation

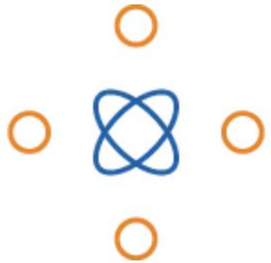
Mission of the EGI Federation

Deliver open solutions
for advanced computing and data analytics
in research and innovation

Mission of the EGI Foundation

Enable the EGI Federation
to serve international research and innovation together

Our structure

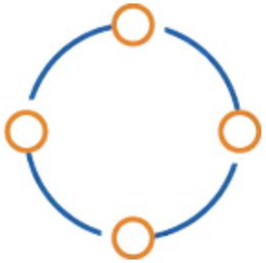


EGI Federation

EGI is a federation of computing and storage resource providers united by a mission of delivering advanced computing and data analytics services for research and innovation.

EGI Foundation

EGI Foundation is a not-for-profit organisation created to coordinate and develop the EGI infrastructure and engage diverse users of our broad service portfolio.



EGI community

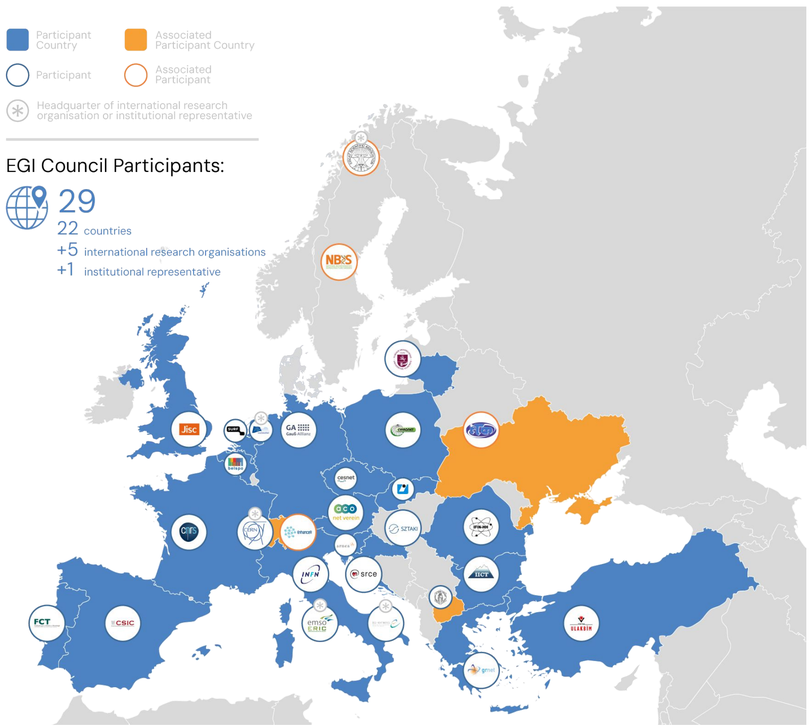
The EGI community is a community of researchers, developers, funders, technologists, dreamers and do-ers: anyone with a stake in advanced computing for research.

EGI in a Nutshell

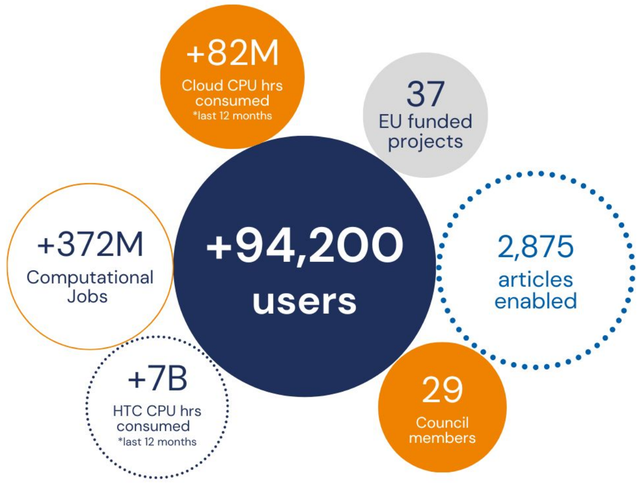
www.egi.eu

EGI Federation

A European flagship digital infrastructure for data-intensive scientific computing



EGI in numbers[1]



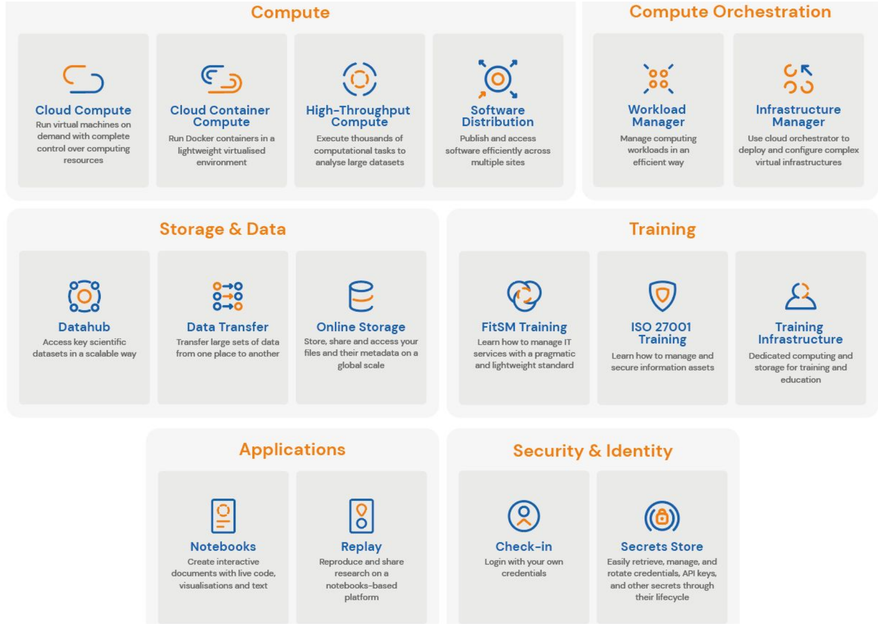
Why a federation?

- Support science at international scale
- Build an hyperscale facility for research
- Invest nationally, access globally
- Bring computing to the data

EGI Services

EGI delivers advanced computing services to support scientists, international projects, research infrastructures and businesses.

EGI services for research



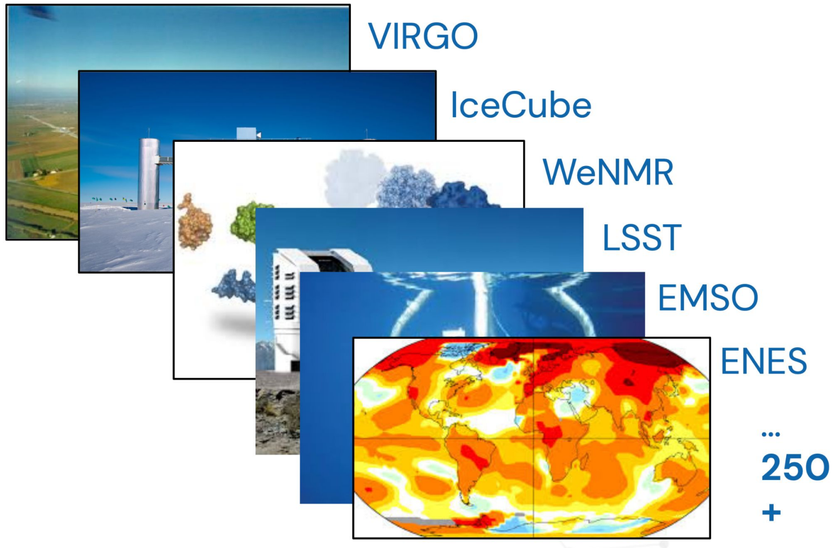
1 the key numbers are correct as of April 2024.

The EGI Federation is an international e-infrastructure…

We provide advanced computing and data analytics for research and innovation

2010

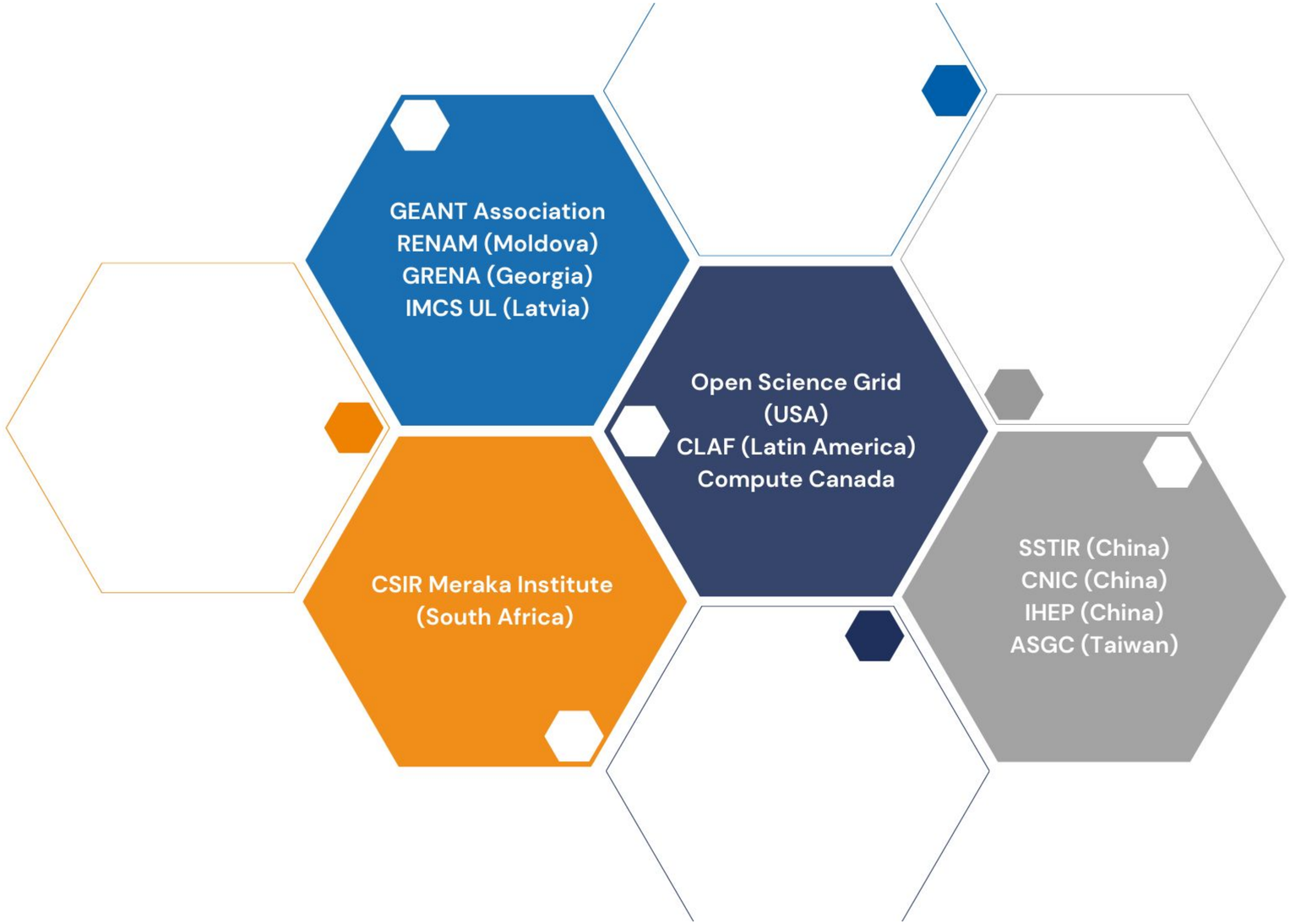
From the Large Hadron Collider compute grid…



2024

….to a multidisciplinary, global, open science infrastructure

… with International collaborations



Our target groups

Research sector

Research communities and research infrastructures

International research projects and research collaborations

Small international groups and individual researchers

Private sector

Small medium enterprises

Industry

Public authorities & policy makers

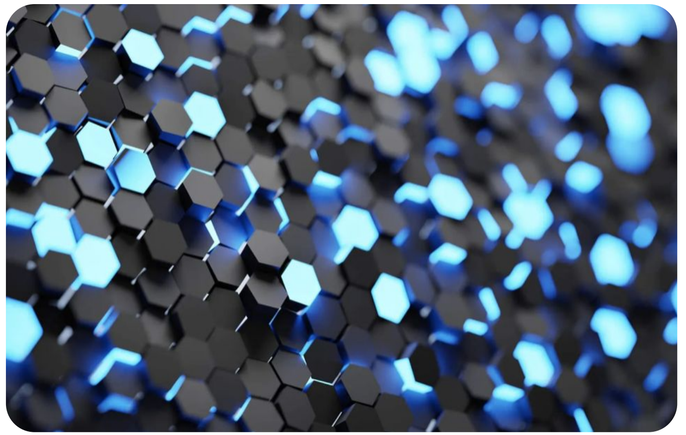
Public Authorities

Policymakers

Our Services

Services for Research

Large-scale computing and data analytics services are helping scientists to accelerate their research.



Services for Federation

Our internal services provide tools for coordination of the EGI Federation, improving how we work together.

Services for Business

We help companies to exploit services and resources for compute- and data-intensive research and innovation.

Services for research

Compute

Cloud Compute
Run virtual machines on demand with complete control over computing resources

Cloud Container Compute
Run Docker containers in a lightweight virtualised environment

High-Throughput Compute
Execute thousands of computational tasks to analyse large datasets

Software Distribution
Publish and access software efficiently across multiple sites

Compute Orchestration

Workload Manager
Manage computing workloads in an efficient way

Infrastructure Manager
Use cloud orchestrator to deploy and configure complex virtual infrastructures

Applications

Notebooks
Create interactive documents with live code, visualisations and text

Replay
Reproduce and share research on a notebooks-based platform

Storage & Data

Datahub
Access key scientific datasets in a scalable way

Data Transfer
Transfer large sets of data from one place to another

Online Storage
Store, share and access your files and their metadata on a global scale

Security & Identity

Check-in
Login with your own credentials

Secrets Store
Easily retrieve, manage, and rotate credentials, API keys, and other secrets through their lifecycle

Training

FitSM Training
Learn how to manage IT services with a pragmatic and lightweight standard

ISO 27001 Training
Learn how to manage and secure information assets

Training Infrastructure
Dedicated computing and storage for training and education

Services for Federation

Services for business

EGI DIH supporting SMEs and Industry with Digital Transformation

Test before invest

Technical support for companies and SMEs to test and validate services and solutions before making the investment

Find funding

Knowledge and technical assistance through training and consultancy

Networking and community building

Interaction between companies and EGI community opens new market opportunities

Skills and training

Funding opportunities and investment mechanisms to support sustainable innovation

Read about our services for business: https://www.egi.eu/services/business/

Our users

95,000
Total number of users

+10.200
New users in 2023

Top 5 cloud communities

WeNMR	NBIS	Biomed	BioISI	ENVRI
41K	21K	1.5K	1K	967

By number of registered users

Top HTC community

atlas, cms, alice, lhcb, belle, virgo

Essential partners and the largest adopters

Research infrastructures (RI) and research communities

13
new scientific communities

23
RIs on ESFRI roadmap

49
RIs using our services

1
new RIs engaged in 2023

EGI Federation annual report 2023

An overview of the key results of 2023

- The EGI Federation annual report 2023 is now available online
  - https://www.egi.eu/publication/annual-report-2023/
  - https://zenodo.org/records/11393415

Delve deeper into the EGI Federation's accomplishments in 2023 by downloading the full Annual Report. The report provides a comprehensive overview of our activities, achievements, and future goals.

www.egi.eu
#EGI2024

EGI

Registration Open

2024

Early Bird until 31/07

go.egi.eu/EGI2024

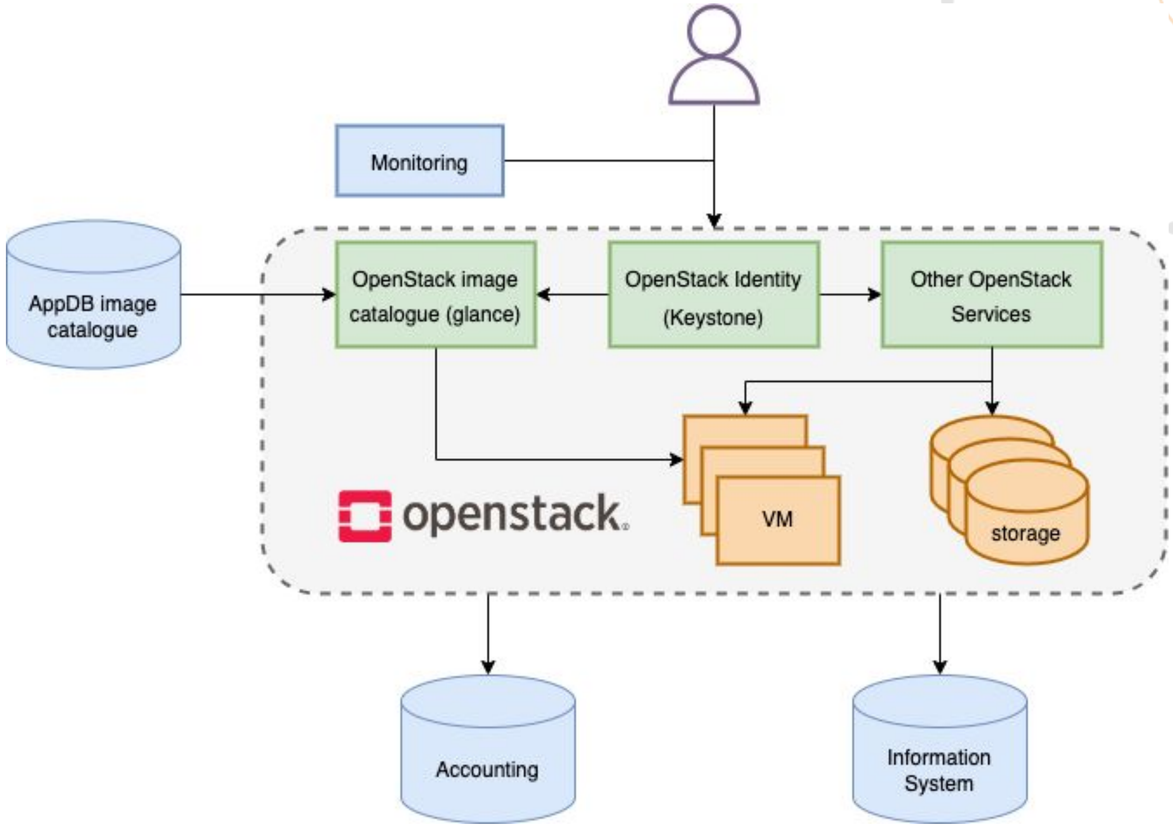
Section 2

EGI Cloud Compute

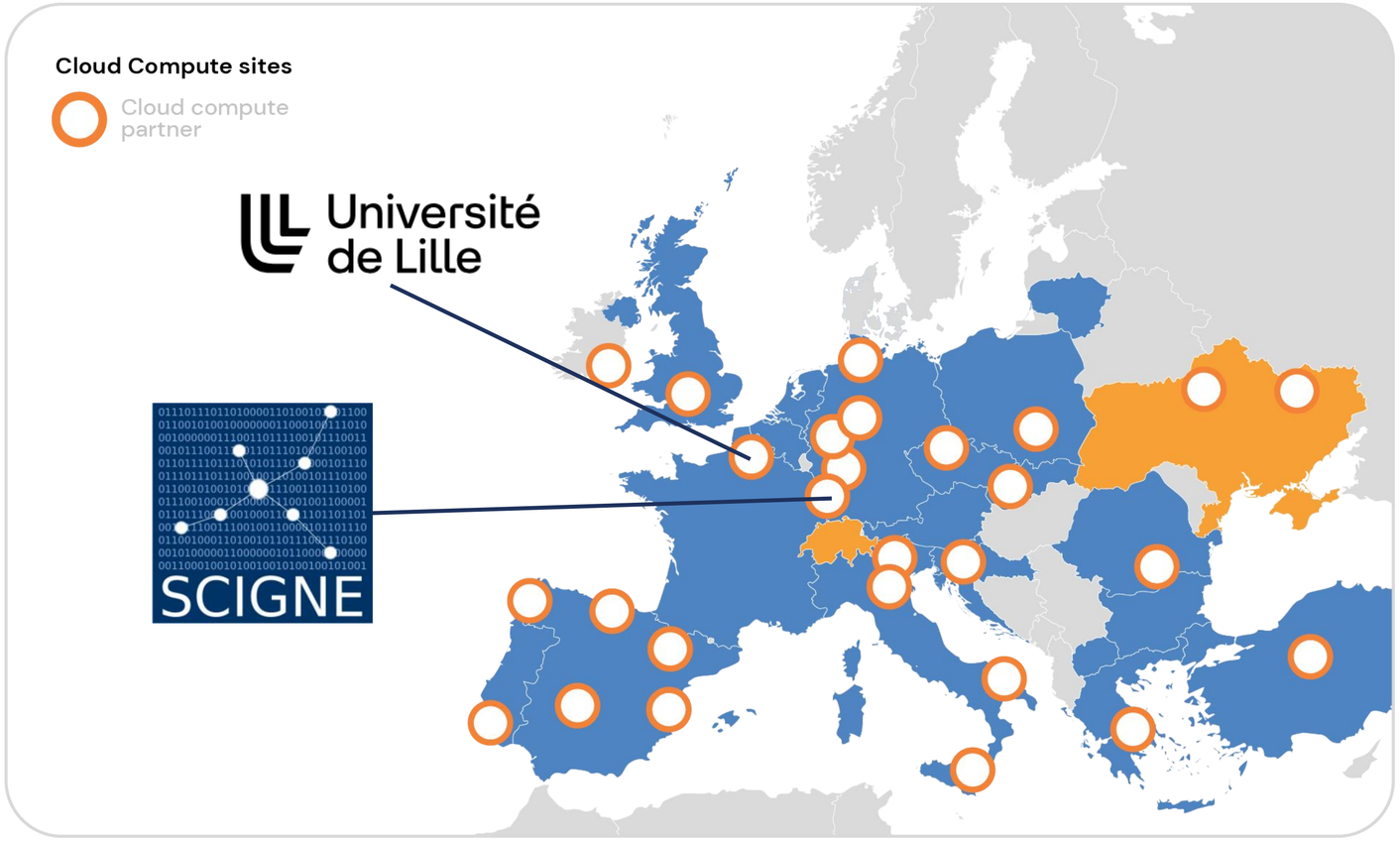
Cloud Compute

A distributed Infrastructure as a Service (IaaS) powered by federated providers

- **Run workloads as VMs** - flexible and customisable configurations, including GPUs
- Common **VM image catalogue** - powered by AppDB
- **Dynamic DNS** for assignment of memorable hostnames
- **fecloudclient** for interaction via CLI and automation
- Supported by **30 OpenStack providers** across Europe and beyond

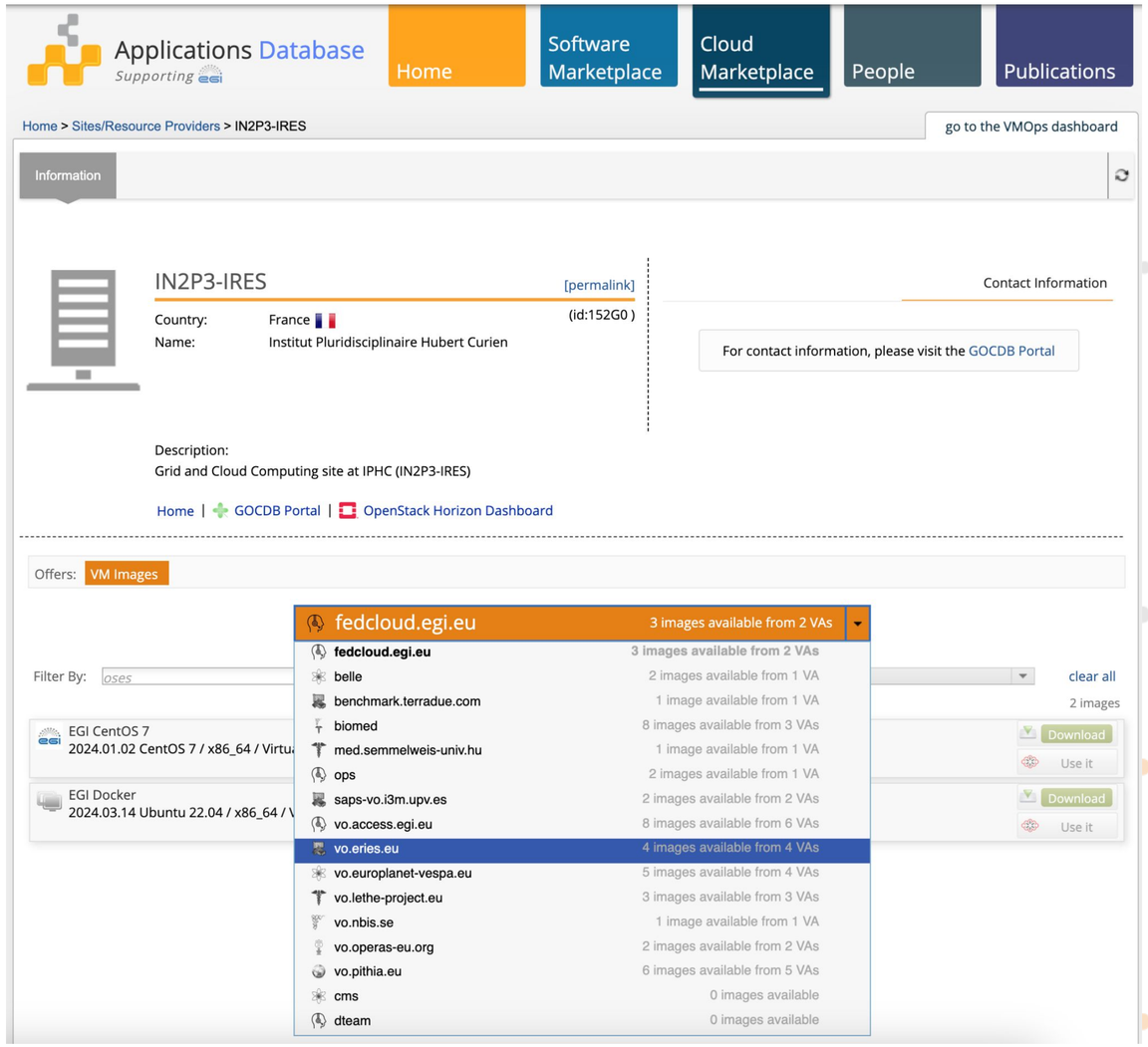


Cloud Compute providers



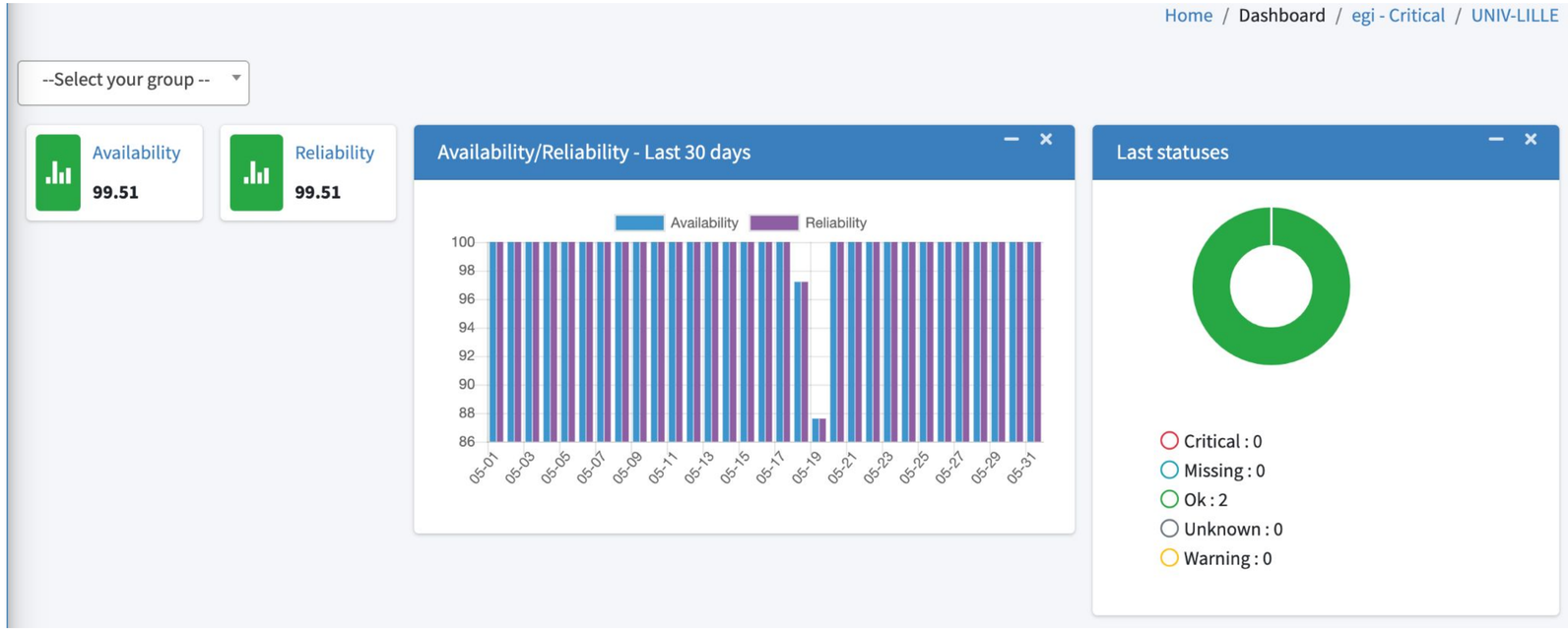
Discovery

- **Central discovery with AppDB**
  - Web UI + REST/GraphQL APIs
  - Easily understand which providers support a given community and with which kind of resources
  - Relying on the GLUE2.1 OGF Standard
  - Integrated with the VM image management



Monitoring

- **Check 1: service is reachable**
  - Just tries to connect to the HTTPS endpoint
- **Check 2: starts a VM**
  - Needs Check-in, AppDB Sync to work
  - Tests from a user's perspective using the "ops" VO
- **Check 3: accounting freshness**
  - Checks if accounting records were sent recently
- **Check 4: discovery freshness**
  - Checks if the AppDB has recent information about the site



Accounting

Collects usage information for every provider / community

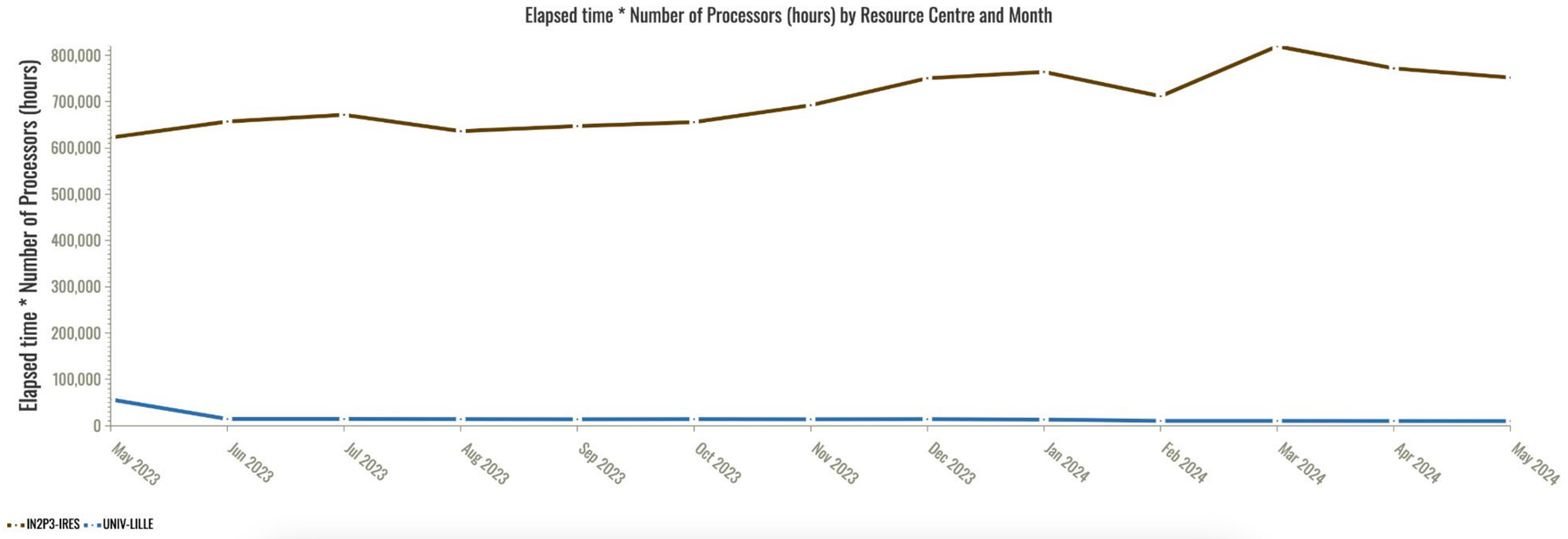
NGI_FRANCE — Elapsed time * Number of Processors (hours) by Resource Centre and Month (All VOs)

Resource Centre	May 2023	Jun 2023	Jul 2023	Aug 2023	Sep 2023	Oct 2023	Nov 2023	Dec 2023	Jan 2024	Feb 2024
IN2P3-IRES	622,233	656,925	671,407	636,002	646,955	655,542	692,035	750,365	764,071	711,603
UNIV-LILLE	56,618	14,413	14,429	14,150	13,693	14,152	13,700	14,160	13,393	10,466
Total	678,852	671,338	685,836	650,151	660,648	669,694	705,735	764,525	777,464	722,068
Percent	7.25%	7.17%	7.33%	6.95%	7.06%	7.16%	7.54%	8.17%	8.31%	7.71%

1 - 2 of 2 results ‹ 1 › Number of rows per page

Download JSON Data / Download CSV Data

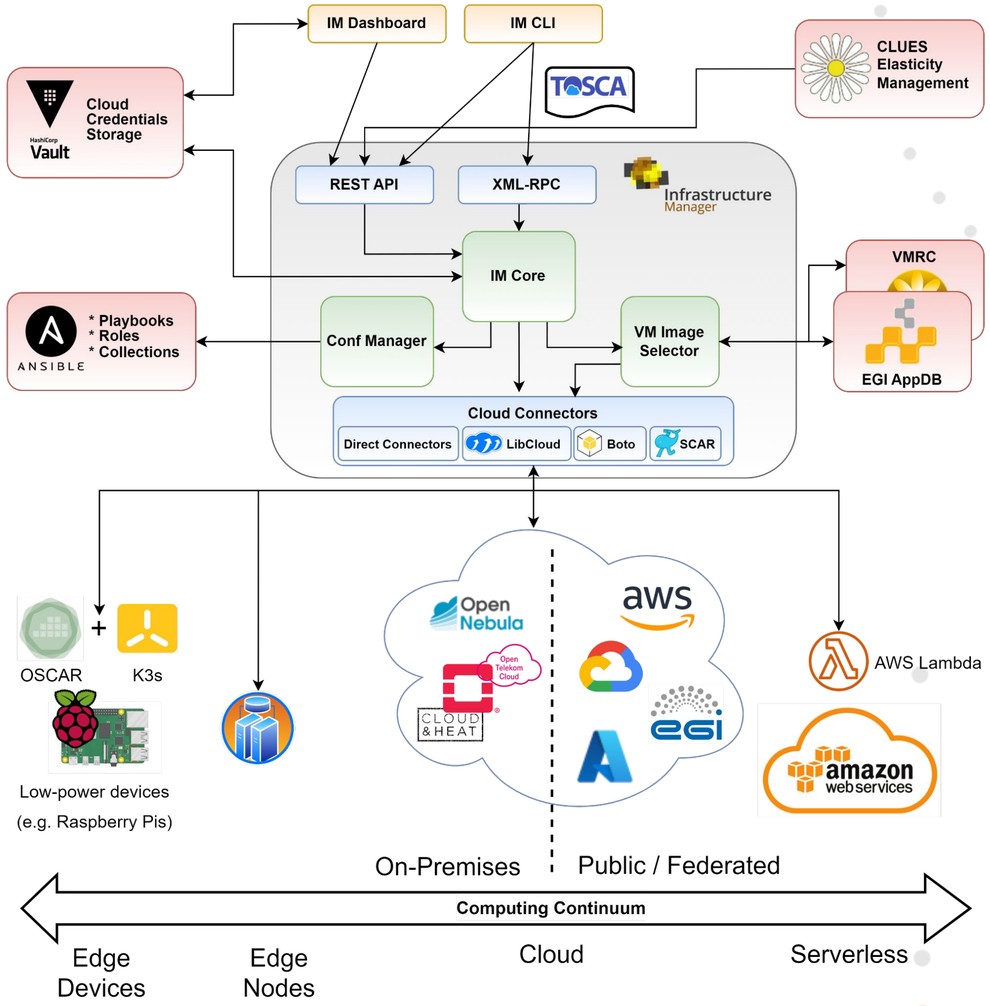
The information in the previous table is also shown in the following graph.



Infrastructure Manager

IM deploys virtual infrastructures on Cloud

- **Automates** the deployment, configuration, software installation, monitoring and update of virtual infrastructures
- **Infrastructure as Code (IaC)** using **RADL/TOSCA** for infrastructure description & **Ansible** for contextualization
- Automated elasticity with CLUES
- Wide variety of back-ends from edge to serverless, making applications **Cloud agnostic**



EGI Federated Cloud User Groups

46

Active Service Level Agreements using capacities from 33 federated research clouds from 17 countries[1]

19.3%

increase in Cloud CPU/h consumption in 2023.

vo.ai4eosc.eu **6M** Cloud CPU/h

The resources offered by the EGI Infrastructure are used to support piloting activities in the context of the AI4EOSC EC-funded project

[1] EGI Cloud Federation consists of 33 sites, 31 certified and 2 undergoing certification

Top 5 Cloud compute user communities based on CPUh consumption:

perla-pv.ro **4.1M** Cloud CPU/h

In 2023, the allocated computational resources were used to perform ab initio DFT calculations for band alignments between the perovskite layer and electron transporter layers and numerical analysis of bias stress test in order to anticipate the PSC degradation. In addition, two papers concerning iodine migration were published.

vo.pangeo.eu **4.6M** Cloud CPU/h

See success stories table below

vo.access.egi.eu **2.7M** Cloud CPU/h

Running piloting activities in the EGI Infrastructure

vo.deltares.nl **1,8M** Cloud CPU/h

The EGI infrastructure contributed to further speed up the multi-threaded scaling of the distributed hydrological models. Over the last year the use case pilot switched to Docker and moved from Python to Julia to use the available CPU more efficiently and shorten the duration of the job execution. Now, run times are 2 to 11 times faster, depending on the chosen routing scheme. The improved performance achieved allows the use of distributed hydrological models in large-scale hydrological forecasting and climate-change applications, which is currently often limited to lumped models. Two publications were submitted.

Section 3

The French contribution

EGI: figures for France in 2022

+3400 service users

In 2022, +3400 researchers from French institutions used the services provided by the EGI Federation

30 Supported communities

In 2022, the French infrastructure supported 30 research communities in the following disciplines: Agriculture, Climate Research, Health and Medicine, Linguistics, Physics

+900 publications

The research communities, projects and scientific collaborations from France supported by the EGI led to more than 900 peer-reviewed scientific publications

Projects

French partners participate in 20 collaboration projects + EGI-ACE, iMagine and interTwin

626 M CPU-hours

5.7 M Unified Middleware Distribution

17 HTC Data Centres

EGI Federated Cloud member

Resource Providers supporting the EGI VO SLAs framework

Overview of French resource providers

13 SLAs agreed to support Customers' needs:

- WeNMR, Biomed, VESPA, Prominence/Fusion, PITHIA, OPERAS-ERIC, NBIS, LETHE, EMPHASIS, BioISI, BELLE-II, EISCAT_3D, EuroScienceGateway

Resource providers:

- IN2P3-IRES (9): BELLE-II, BioISI, BioMed, EMPHASIS, LETHE, OPERAS, PITHIA, NBIS, VESPA
- UNIV-LILLE (1): Prominence
- CREATIS-INRA-LYON (1): Biomed
- GRIF (1): Biomed
- IN2P3-CPPM (1): Biomed, WeNMR
- CNRS (3): EISCAT_3D, EuroScienceGateway, WeNMR, Biomed

Total **HTC CPU/h** delivered in total in 2023: 509,686,809

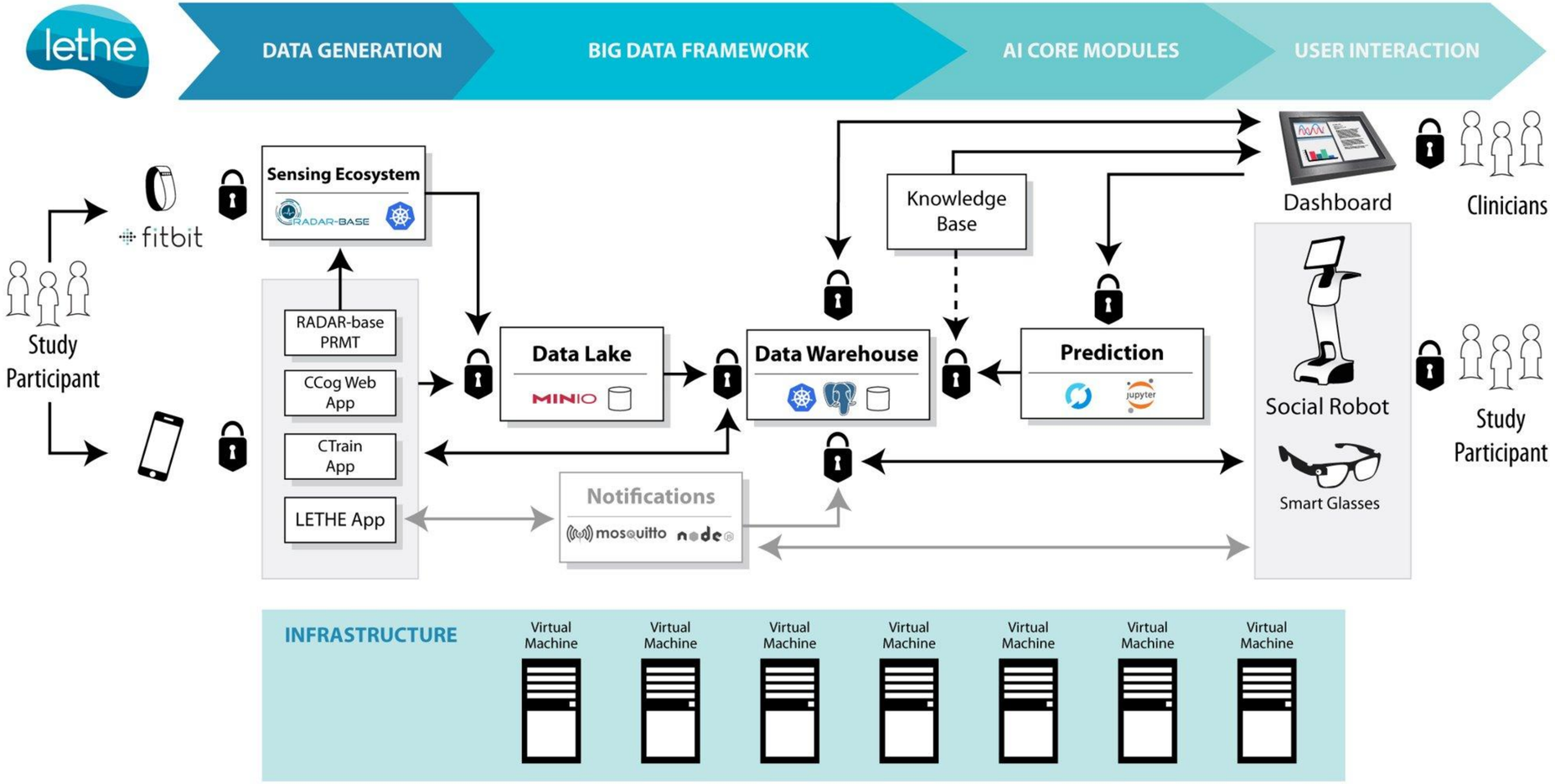
Total **Cloud CPU/h** delivered in total in 2023: 6,640,351

LETHE project resources usage

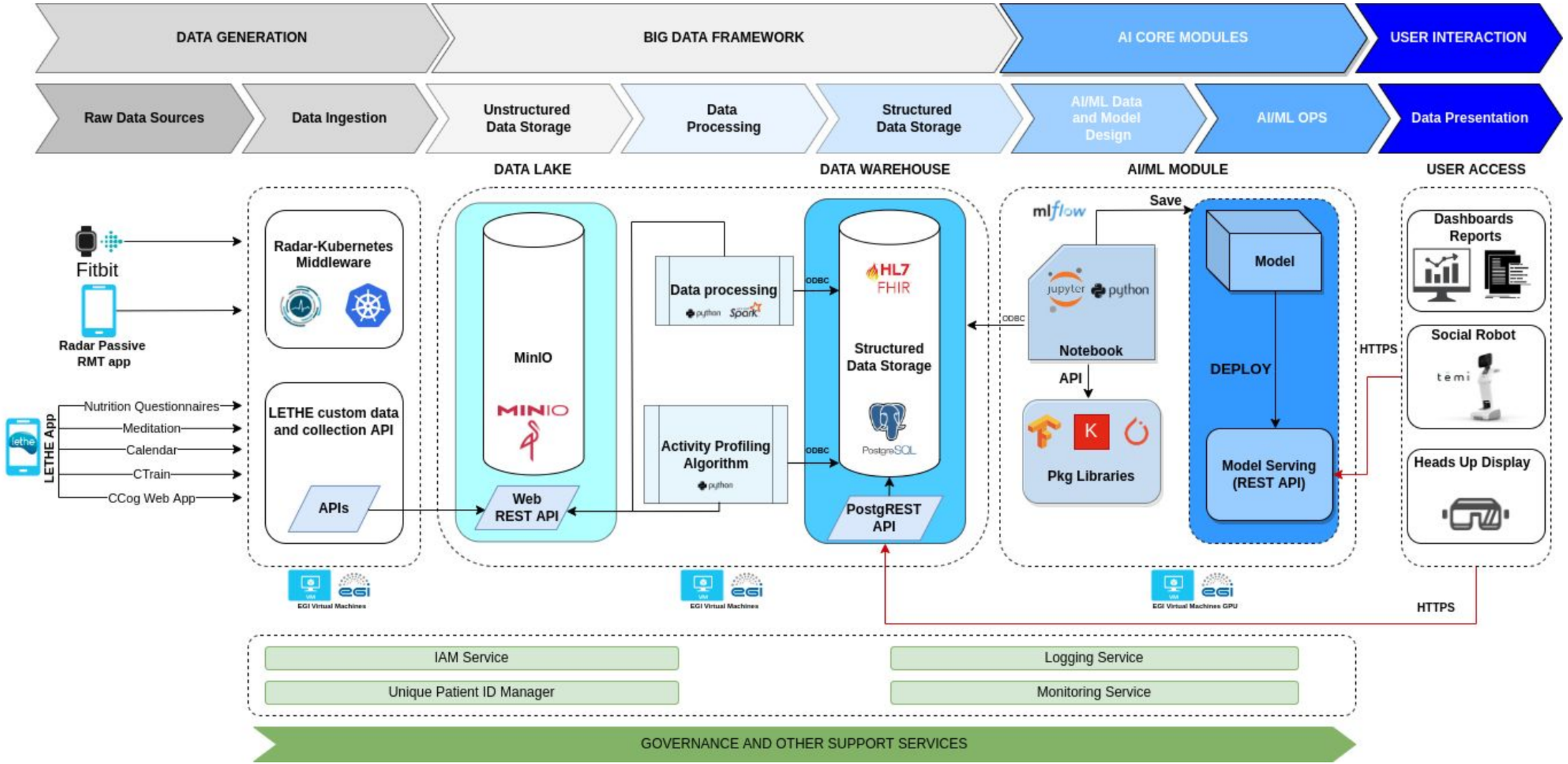
The LETHE project

Digital Cognitive Biomarkers: https://www.lethe-project.eu/

A personalized prediction and intervention model for early detection and reduction of risk factors causing dementia, based on AI and distributed Machine Learning



LETHE Phase II architecture

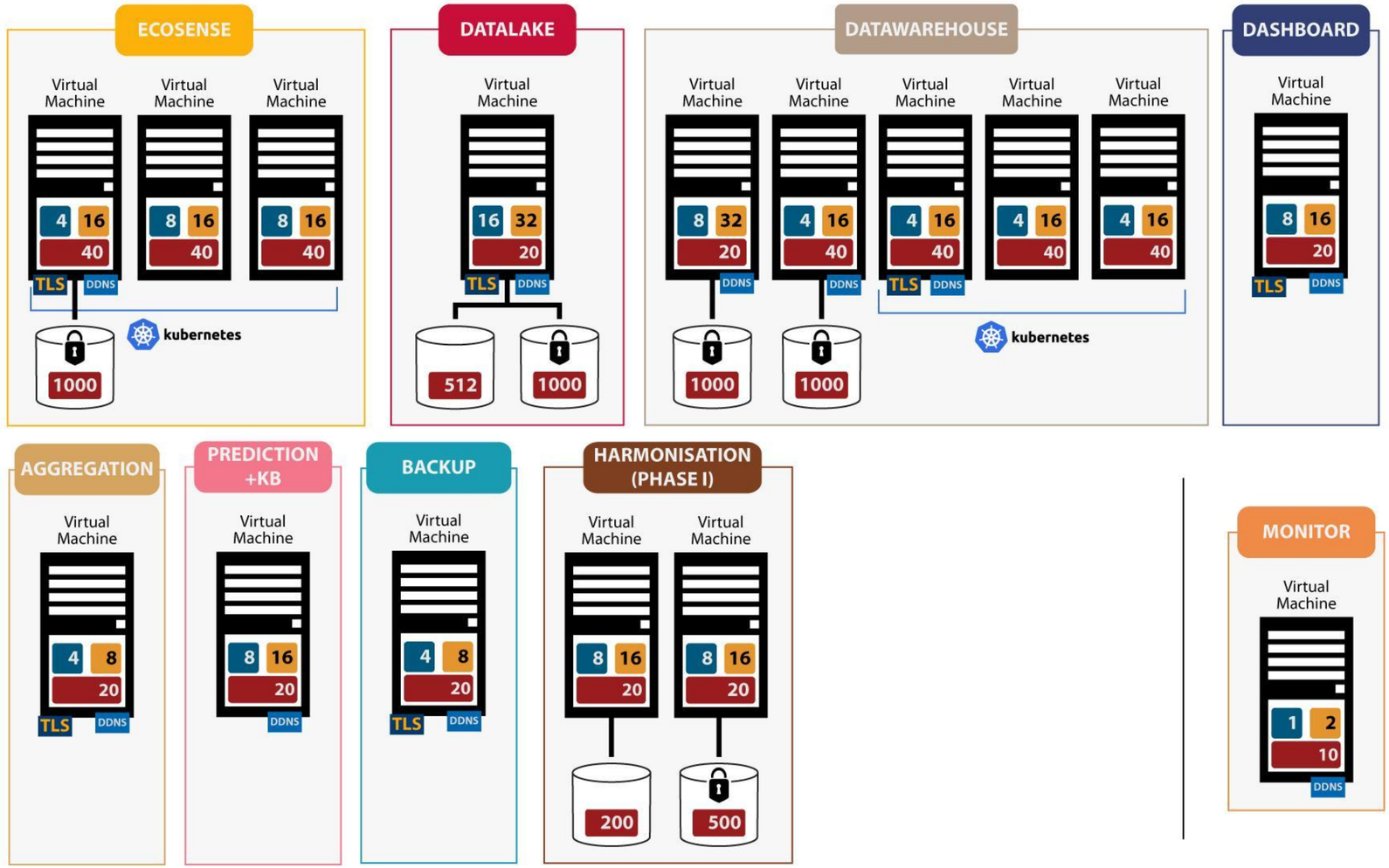


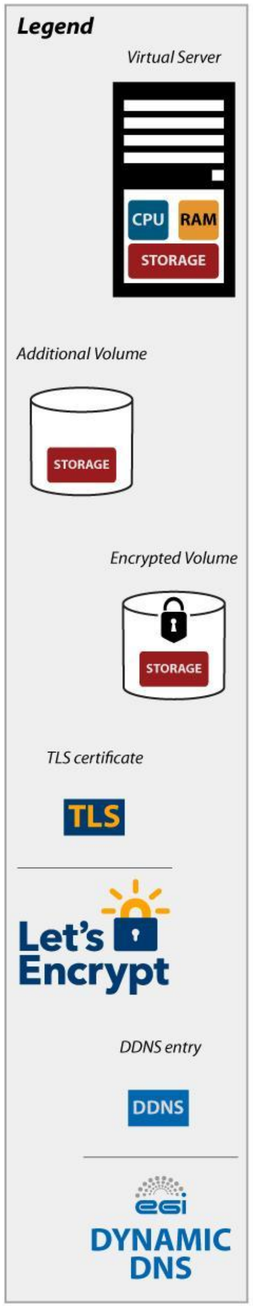
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101017405

LETHE Phase II infrastructure

Supported by IN2P3-IRES

16 VMs **89** vCPUs **234** GB RAM **18** volumes **12.5** TB **11** IPs





Process used to identify the cloud provider for LETHE

One of the possible way to be involved in and funded by a project

- **EGI Foundation involved in a EU Project proposal to provide cloud resources**
  - **Got funding for cloud resources provisioning as well as for other activities**
  - **No specific cloud provider was directly involved in the proposal**
    - For some projects, resources providers are identified and directly involved as partners
- **Identification of the architecture and requirements**
  - **With the support of EGI, the LETHE project identified and formalised the architecture and requirements**
- **Identification of cloud providers able to support the project**
  - **EGI Foundation launched a call for support addressed to Cloud Compute providers**
- **Selecting and funding a cloud provider**
  - **A provider able to meet the requirements got selected**
  - **The provider got funded by EGI as a sub contractor to provide the resources**
  - **Agreements got put in place (SLA and DPA with LETHE, OLA and subcontractor agreement with cloud provider)**
- **Supporting the LETHE user community**
  - **Support by EGI Foundation and cloud provider personnel**

Section 4

Joining EGI Cloud Compute: benefits and requirements

The benefits of joining EGI Cloud Compute

- **Funding aspects**
  - Complementary funding
  - Usage of unused resources (possibly already funded)
- **Supporting science and research**
  - Reaching new user communities
  - Involvement in international 1st class research-supporting projects
- **Exposure**
  - International exposure and cross borders activities
- **Maturity**
  - A framework, process and tools mant to increase the maturity of the service delivery
  - Sharing of experience with providers from other countries having similar needs and concerns

Requirements derived from user needs

Requirements on Cloud sites to meet user requirements

- **Using common and well known interfaces**
  - **OIDC authentication and Authorisation via EGI Check-in, leveraging organisational identities**
  - **Management using standard OpenStack API and CLI**
  - **Infrastructure as Code**
    - With TOSCA, RADL via IM; with Terraform or other tools able to leverage OpenStack APIs
- **Providing support on using the services**
  - **Documentation: global and possibly complemented by site-specific documentation**
  - **Individual support (like done by EGI CST team)**
- **Providing a clear and agreed quality of service (like via SLA)**
  - **Provide clear information about the back-up implementation of data**
- **Providing access to GPGPU**

Requirements derived from user needs #2

Requirements on Cloud sites to meet user requirements

- **Integration in cross-borders deployments**
- **Being able to invoice access for Pay4use use-cases (typical for business)**
- **Having conducted a Risk Assessment covering SACM**
  - Service unavailable due to hardware, software or network failure, human error
  - Loss of data due to hardware, software or network failure, human error
  - Lack of personal to maintain and operate the service
  - Major disruption, security incident, DDoS attack,...
- **Taking security aspects in consideration**
  - Security monitoring by the sites and infrastructure
  - Documented Technical and Organisational Measures in place to secure the operations
  - Data Center certification

How to join EGI Cloud Compute

Subtitle

1. **Registration and Certification**
   - PROC09 Resource Centre Registration and Certification
   - It ensures provider meets all EGI policies and all needed information (e.g. security contacts) is available
2. **OpenStack Technical integration**
   - Configure Check-in for Authentication and Authorization
   - Configure generation and sending of accounting records
   - Allow monitoring of the OpenStack from EGI monitoring
   - Enable access to selected communities

Full documentation at https://docs.egi.eu/providers/cloud-compute/

Section 5

Contributing to and integrating with EOSC

EGI contribution to EOSC

- **EGI as a contributor to the EOSC EU Node**
  - EGI and partners deliver and operate several services within the EOSC EU Node, including AWM , Monitoring, Accounting, Helpdesk, Security coordination and Notebooks

- **EGI as technology provider for EOSC**
  - Core components of EOSC developed by EGI: Monitoring, Accounting, Helpdesk, Application Workload Manager (AWM)
  - EOSC-Beyond project will further evolve the EOSC core and expand it with new features and services

- **EGI Federation aspires to be part of the EOSC Federation as a Node**
  - Transnational access to international use cases and to services operated within the other EOSC Nodes
  - Capacity contributed from selected compute-data centres of the EGI federation.

EGI Integration with EOSC EU Node

See more on https://zenodo.org/records/11128540

- **Scenario 1: Expand EOSC EU Node capacity**
  - Allow EOSC EU Node to serve a wider set of communities and to satisfy more demanding requirements with capacity from EGI Federated Providers
  - Application Workflow Management (AWM) component of EOSC EU Node (based on IM) as an integration point
  - EOSC EU Node acts as business channel towards users.
- **Scenario 2. Aggregate user demand from EGI**
  - Proxy requests to EOSC EU node to increase its utilisation
  - Enable co-located processing and analysis of research data hosted by the EOSC EU Node, and to integrate it with services from EGI Federation.
  - EGI contributes to AWM recipes deployable on EOSC EU Node
  - EGI provides the business channel towards users.

Thank you

Baptiste Grenier

+31 (0)627 860 852

baptiste.grenier@egi.eu

www.egi.eu

This work is partially funded by the EU research and innovation programme