Service
Catalogue

EGI: advanced computing for research

eGi

> EGI's mission is to create and deliver open solutions for science and research infrastructures by federating digital capabilities, resources and expertise between communities and across national boundaries.

www.egi.eu

EGI: advanced computing for research

EGI delivers advanced computing services to support scientists, multinational projects and research infrastructures. EGI is also committed to support business and innovation.

The EGI services can be requested by everyone involved in research and innovation from both the public and private sector.

The EGI Federation brings together over 300 data and computing sites spread across Europe and worldwide. It is governed by the EGI Council.

To request a service, fill the form on our website: go.egi.eu/request

Map of Services

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

Storage and Data

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

Archive Storage - Back-up your data for the long term and future use in a secure environment

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

Training

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

Training infrastructure - Dedicated computing and storage for training and education

Cloud Compute

Run virtual machines on-demand with complete control over computing resources

Cloud Compute offers guaranteed computational resources in a secure and isolated environment with standard API access, without the overhead of managing physical servers.

With Cloud Compute you can:

- Execute compute- and data-intensive workloads (both batch and interactive)

- Host long-running services (e.g. web servers, databases or applications servers)

- Create disposable testing and development environments scaled to your needs

- Select CPU, memory, disk configurations to fit your requirements

- Manage your resources in a flexible way with integrated monitoring and accounting capabilities

Cloud Compute is provided by the EGI Federated Cloud sites

Cloud Container Compute

Run Docker containers in a lightweight virtualised environment

Cloud Container Compute gives you the ability to deploy and scale Docker containers on-demand. It offers guaranteed computational resources in a secure and isolated environment with standard API access, without the overhead of managing the operating system.

The result is improved performance, ideal for development work.

Cloud Container Compute is provided by the EGI Federated Cloud sites

Cloud Container Compute offers:

- On-demand provisioning

- Lightweight environment for maximised performance

- Standard interface to deploy on multiple service providers

- Interoperable and transparent

- Seamless development and operations environments

What happens when a human cell meets Salmonella?

Konrad Förstner and colleagues, from the University of Würzburg in Germany, used **Cloud Compute** to run READemption, a pipeline for the computational analysis of RNA sequencing data.

Cloud Compute helped them to handle computational demand peaks when new data sets arrived and that sped up the whole process significantly.

Their results were published in *Nature* (doi:10.1038/nature16547) and show that a small piece of RNA has a say on both the infection process and the immune response.

"The EGI Federated Cloud accelerated our analysis dramatically and helped us to focus on scientific solutions and results, instead of resource management and system administration."

K. Förstner, University of Würzburg

Research Communities using Cloud Compute

The EXTraS project is harvesting 13 years of data collected on-board the ESA's X-ray space observatory XMM-Newton. The project is using Cloud Compute to implement four lines of analysis with ad-hoc software pipelines.

The DRIHM project is prototyping an e-infrastructure to simulate extreme hydro-meteorological events such as flash flooding.

The National Bioinformatics Infrastructure of Sweden uses Cloud Compute to provide bioinformatics tools to their researchers, including high-profile tools to predict 3D protein structures, for example. So far, more than 6,700 unique users in 73 countries have made the most of these resources.

High-Throughput Compute

Execute thousands of computational tasks to analyse large datasets

EGI offers more than 650,000 cores of installed capacity, supporting about 1.6 million computing jobs per day.

The EGI High-Throughput Compute infrastructure is over 10 years old and has helped, for example, CERN scientists to analyse the data from millions of experiments at the Large Hadron Collider. This collaborative work opened the way to the discovery of the Higgs boson and the Nobel Prize attributed in 2013 to Peter Higgs and François Englert.

High-Throughput Compute is provided by the distributed network of computing centres of the EGI Federation.

Main features:

• More processing capacity over long periods of time – faster results

• Shared resources among users – collaborative research

• With monitoring and accounting tools with information about availability and resource consumption

• With workload and data management tools to manage all computational tasks

HADDOCK - computational tools for structural biologists

HADDOCK is a web portal that offers computational tools for structural biologists via a user-friendly interface. It is an integrative approach used to model the structure of complexes of proteins and other biomolecules.

The thousands of computational tasks operated in the back-end of HADDOCK are powered by **High-Throughput Compute.**

So far, HADDOCK has processed more than 130,000 submissions from over 7,500 scientists, which translates into about 8 million High-Throughput Compute jobs per year on the EGI infrastructure.

The research performed thanks to HADDOCK has resulted in hundreds of peer reviewed publications.

More examples of what is possible to do with High-Throughput Compute: go.egi. eu/cst

The EGI Services are provided by the EGI Federation

- **EGI Council participants:** national e-infrastructure providers and international research organisations (CERN and EMBL)

- **Integrated e-infrastructure providers**

+650,000
Cores of compute capacity
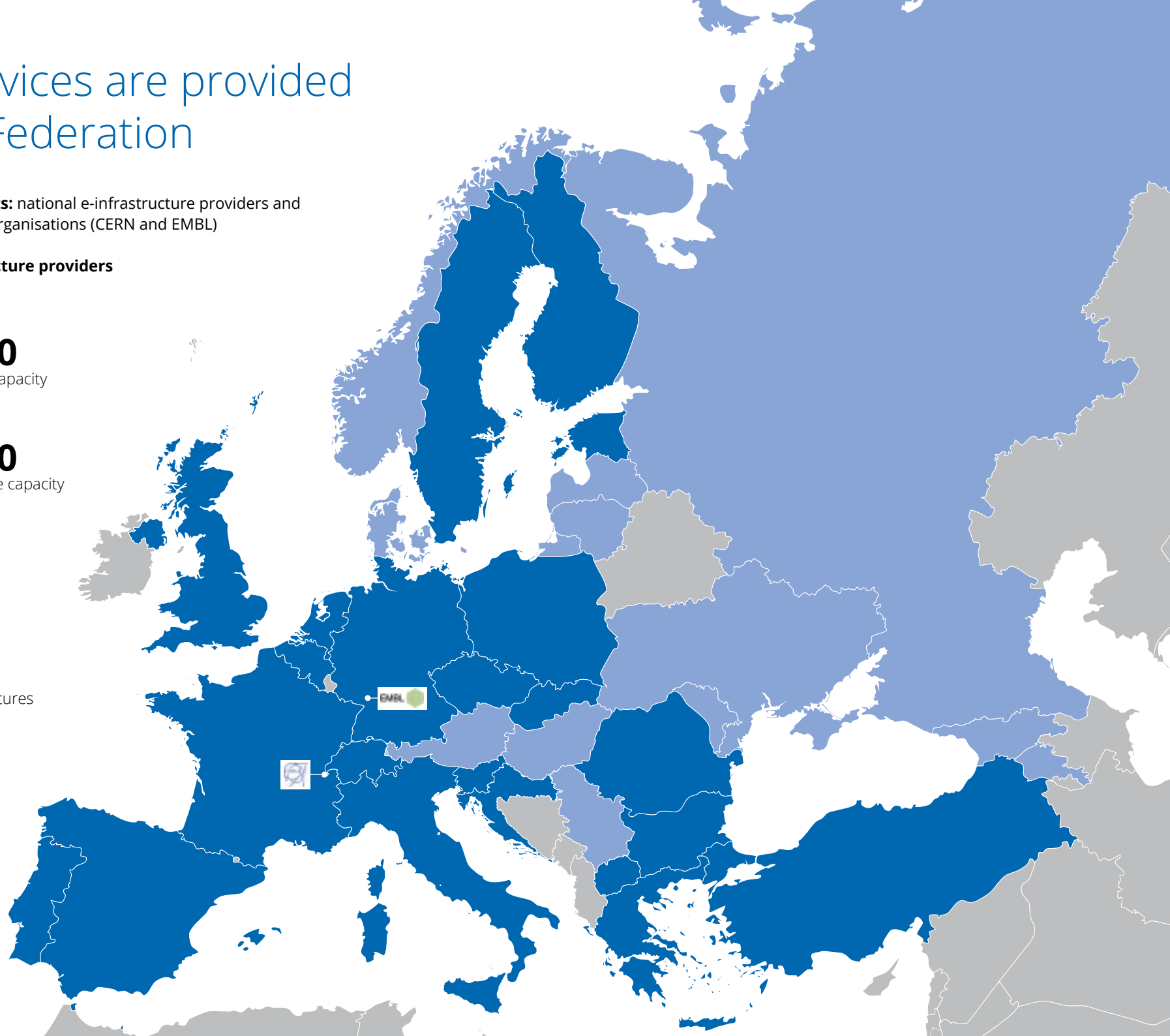
+500,000
Tetabytes of storage capacity

+45,000
Users

15
Research Infrastructures integrated with EGI

EMBL

Numbers and membership information correct as of September 2016

Online Storage

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

Online Storage allows you to store data and share it across distributed teams. Your data can be accessed through different standard protocols and can be replicated across different providers to increase fault-tolerance.

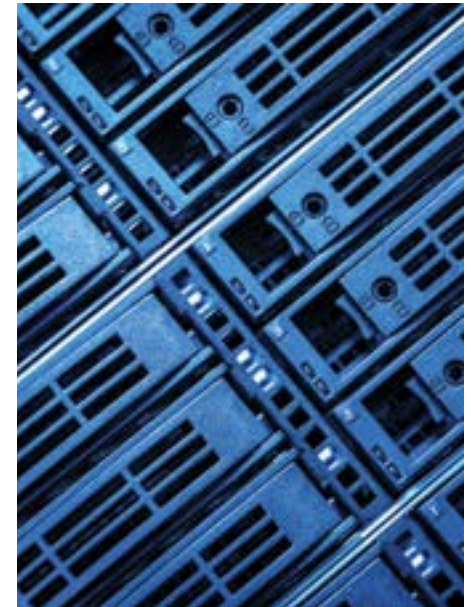Online Storage gives you complete control over the data you share and with whom.

Online Storage is provided by the distributed network of computing centres of the EGI Federation.

With Online Storage you can:

• Assign global identifiers to files

• Access highly-scalable storage from anywhere

• Control the data you share

• Organise your data using a flexible hierarchical structure

Archive Storage

Back-up your data for the long term and future use

Archive Storage allows you to store large amounts of data in a secure environment freeing up your online storage resources.

The data on the Archive Storage can be replicated across several storage sites, thanks to the adoption of interoperable open standards. The service is optimised for infrequent access.

Archive Storage is provided by the distributed network of computing centres of the EGI Federation.

With Archive Storage you can:

• Store data for long-term retention

• Store large amounts of data

• Free up your online storage

Data Transfer

Transfer large sets of data from one place to another

Data Transfer allows you to move any type of data files asynchronously from one place to another. The service includes dedicated interfaces to display statistics of on-going transfers and manage network resources.

Data Transfer is ideal to move large amounts of files or very large files. The Data Transfer service has mechanisms to ensure automatic retry in case of failure.

Main features:

• Ideal for very large files

• Able to handle large amounts of files

• Transfer process with automatic retry

Data Transfer is provided by the distributed network of computing centres of the EGI Federation.

Online storage and data transfer at unprecedented scales

ATLAS control room at CERN

The Worldwide LHC Computing Grid (WLCG) is a global collaboration of more than 170 computing centres set up to provide the computing resources needed to store, distribute and analyse the data generated by the Large Hadron Collider (LHC) at CERN.

During 2016, WLCG transferred on average 80 Petabytes of data per month, with peaks at 96 Petabytes during summer. This corresponds to more than 1 billion files per month transferred to thousands of particle physicists working across the world.

"WLCG must manage hundreds of petabytes of data, with more than 10PB of new data being added each month to the LHC runs. Without these, the analysis of the data from the LHC would be almost impossible."

Ian Bird, WLCG project leader

Training Infrastructure

Dedicated computing and storage for training and education

The Training Infrastructure is a set of cloud-based computing and storage resources dedicated to onsite tutorials, online training courses or as a platform for self-paced learning.

Trainers can deploy custom virtual machine images as the training environment for the students. The virtual machines can be customised according to specific needs and the community can benefit from the easy deployment and easy reuse of course materials.

Main features:

• High-quality computing and storage environment

• Easy-to-use, on-demand access

• Allows easy deployment of courses and reuse

The Training Infrastructure is provided by the EGI Federated Cloud sites

FitSM Training

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

With FitSM Training you will learn the fundamentals of IT service management and how to implement the FitSM standard in your organisation.

FitSM is a lightweight standard for IT service management. It brings order and traceability through a common conceptual and process model setting out realistic requirements.

The training programme is certified in three levels (Foundation, Advanced, Expert) and is fully accredited by TÜV-SÜD.

"I learned how to implement the FitSM in an IT organization and gain from the benefits the framework provides for efficient service management."

P.Weber, Karlsruhe Institute of Technology

*FitSM Training is provided by the EGI Foundation team
Certification is provided by TÜV-SÜD*

Interested in requesting
EGI Services for your work?

Read more about the services in our website
egi.eu/services

To request a service, please fill the online form
go.egi.eu/request

Our User Support Team will process your
request and contact you afterwards
to discuss implementation.

To contact the user support team,
please send email to **support@egi.eu**

www.egi.eu

In EGI's vision, researchers from all disciplines have easy, integrated and open access to the advanced computing capabilities, resources and expertise they need to collaborate in data/compute-intensive challenges.