www.egi.eu
@EGI_eInfra

AI and ML landscape and platforms and services by EGI

12.3.2021

Ville Tenhunen, EGI Foundation

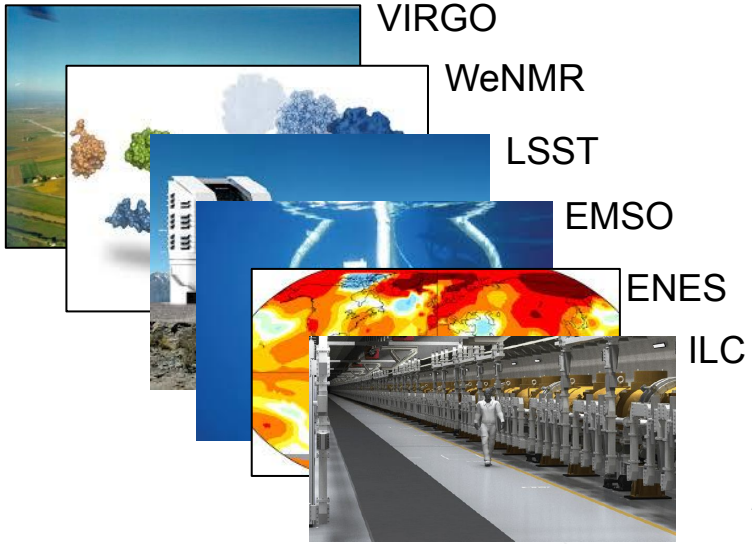
The work of the EGI Foundation
is partly funded by the European Commission
under H2020 Framework Programme

Agenda

- EGI Federation
- AI and ML Landscape in Europe
- EGI Services
- Extending the EGI offer with AI/ML services
- Projects: EGI-ACE and others
- EGI AI and ML Working Group

The EGI Federation is an international e-infrastructure that provides advanced computing and data analytics for research and innovation.

2010

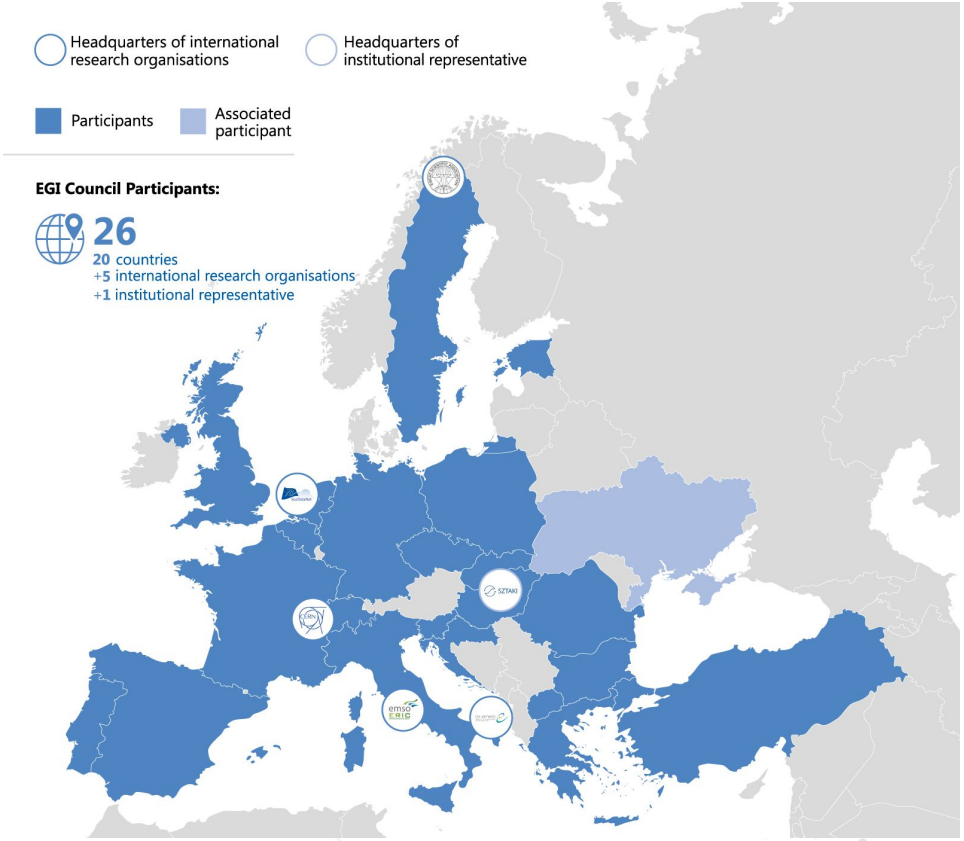


From the high-energy physics compute grid (WLCG)

To a multi-disciplinary, multi-technology infrastructure

EGI Council participants

Belgium	BELSPO's website	Bulgaria	IICT's website	Croatia	SRCE's website
Czech Republic	Metacentrum's website	Estonia	EEnet's website	France	France Grilles' website
Greece	GRNET's website	Germany	NGI-DE's website	Republic of North Macedonia	MARGI's website
Italy	INFN's website	Netherlands	SURF's website	Poland	Cyfronet's website
Portugal	FCT's website	Romania	IFIN-HH's website	Slovakia	IISAS' website
Slovenia	ARNES's website	Spain	CSIC's website	Switzerland	EnhanceR's website
Turkey	ULAKBIM's website	United Kingdom	Jisc's website	International	CERN's website
International	EISCAT's website	International	EMSO's website	International	IS-ENES website
International	SeaDataNet website	Organisation	SZTAKI's website		



www.egi.eu/about/egi-council

https://www.egi.eu/services/

EGI Service Catalogue – Value proposition 1

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

Workload Manager
Manage computing workloads in an efficient way

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

Applications

- SaaS: CHIPSTER, NAMD, ECAS
- PaaS: VMOps, EC3, FGSG

Applications on Demand
Use online applications for your data & compute intensive research

Notebooks
Create interactive documents with live code, visualisations and text

Security

Check-in
Login with your own credentials

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

ISO 9001 Certified Quality Management System www.tuv-sud.com/ms-cert

ISO 20000 Certified IT Service Management System www.tuv-sud.com/ms-cert

https://docs.egi.eu Service documentations

EGI HTC Federation & Cloud federation

HTC providers

Cloud providers

>1 Million computing cores

> 900 PB disk & tape

2,915 service end-points

Capacity is allocated through Virtual Organisations:
https://operations-portal.egi.eu/vo/

EGI Internal Portfolio - Value proposition 2

See all the services offered for EGI participants:
https://www.egi.eu/internal-services/

Coordination

Operations Coordination and Support
Coordinate activities to ensure seamless operations

Community Coordination
A joint approach to user engagement

ITSM Coordination
Ensures professional service management for EGI IT services

Technical Coordination
Progress and innovation through collaboration

Strategy and Policy Development
One federation, one vision, one strategy

Project Management and Planning
A joint approach to planning and management

Security Coordination
Enhance local security for a safer global infrastructure

Communications
Share your successes at a larger scale

Operations

Marketplace BETA
Expose your services to a broader audience

Accounting
Track and report the usage of your services

Collaboration Tools
IT tools for better coordination

Configuration Database
Manage the configuration information of federated e-infrastructure assets and their functional relations

Operational Tools
Integrate resources and operations in a federated ecosystem

Helpdesk
Your point of contact to ask for support at EGI

Service Monitoring
Monitor the performance of IT services

Validated Software and Repository
Benefit from a repository of high-quality software validated for the EGI infrastructure

Security

Check-in
Login with your own credentials

Attribute Management
Manage memberships and groups in communities and virtual organisations

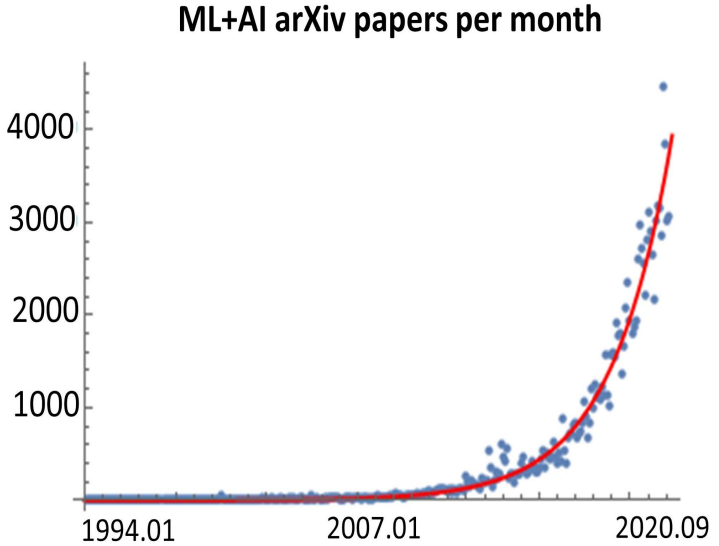
ISO 9001
Certified Quality Management System
www.tuv-sud.com/ms-cert

ISO 20000
Certified IT Service Management System
www.tuv-sud.com/ms-cert

Digital transformation in science

'I wanted to point out that almost everything about science is changing because of the impact of information technology. Experimental, theoretical, and computational science are all being affected by the data deluge, and a fourth, "data-intensive" science paradigm is emerging.'

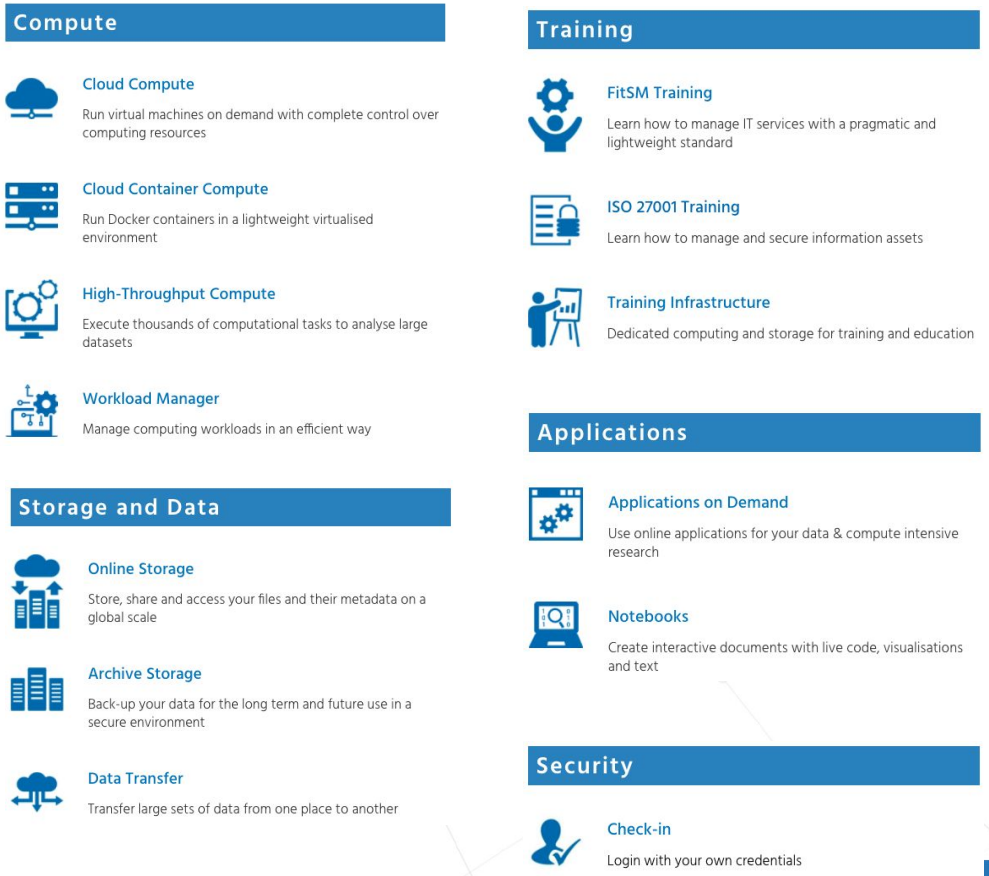
(Turing Award winner Jim Gray in Tony Hey et. al (ed.); The Fourth Paradigm: Data-intensive Scientific Discovery, 2009)



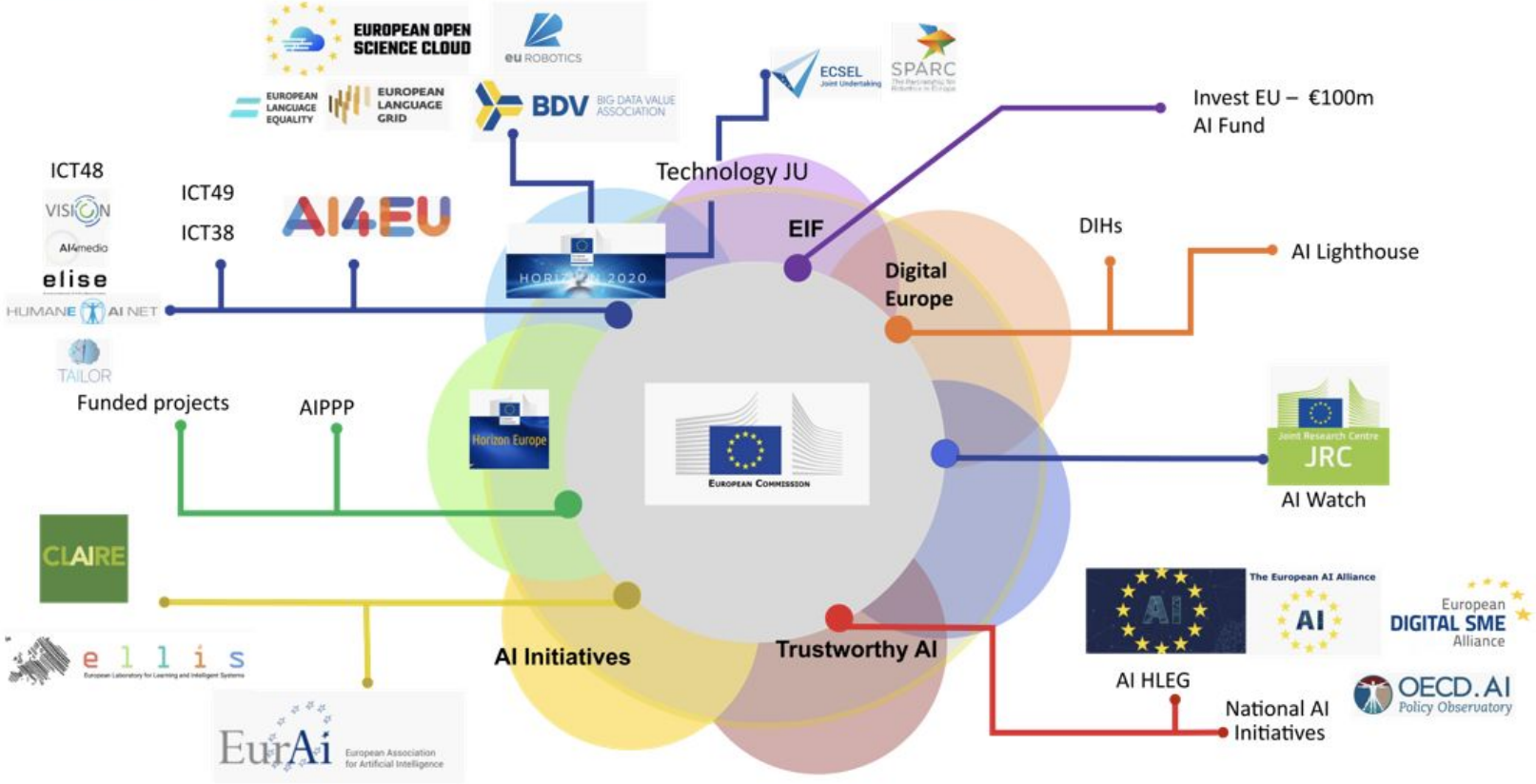
https://twitter.com/MarioKrenn6240/status/1314622995139264517?s=20

Extending the EGI offer with AI/ML services

- **Provide adequate resources** to deploy ML/AI algorithms and platforms (GPUs, storage, etc.)
- **Offer relevant datasets** as a service for ML/AI platforms and algorithms
- **Operate relevant ML/AI platforms** for researchers
- Offer a wide **library of ML/AI analytics** for researchers (models, algorithms)
- **AI/ML Competence Center** for user support and training



European high level AI landscape



Map by Gabriel Gonzalez Castañe, University College Cork

Projects using ML/AI where the EGI Foundation is involved

EGI-ACE
EGI-ACE delivers the EOSC Compute Platform and contributes to the EOSC Data Commons through a federation of cloud compute and storage facilities, PaaS services and data spaces with analytics tools and federated access services. EGI-ACE also delivers to EOSC ML/AI services, like DEEP Hybrid DataCloud and SZTAKI AI solution.

AI4PublicPolicy
The platform will provide **policy development and management** functionalities based on AI, ML, DL, NLP and chatbots.

Big Data for Next geneRation enerGy
Modular big data analytics & AI technologies to improve **decision making in energy**

DIGITbrain
Concept of digital twin to implement the **smart Manufacturing** as a Service

LETHE
A project designed to **prevent cognitive decline in an ageing population** at an early time point by a person centric solution

EOSC Future
Add AI/ML capabilities to the EOSC Portal to **simplify and personalise data and service discovery**.

EUHubs4Data
An instrument for data-driven cross-border experimentation and innovation

SoBigData++
Multi-disciplinary research infrastructure for **big data analytics in social sciences**

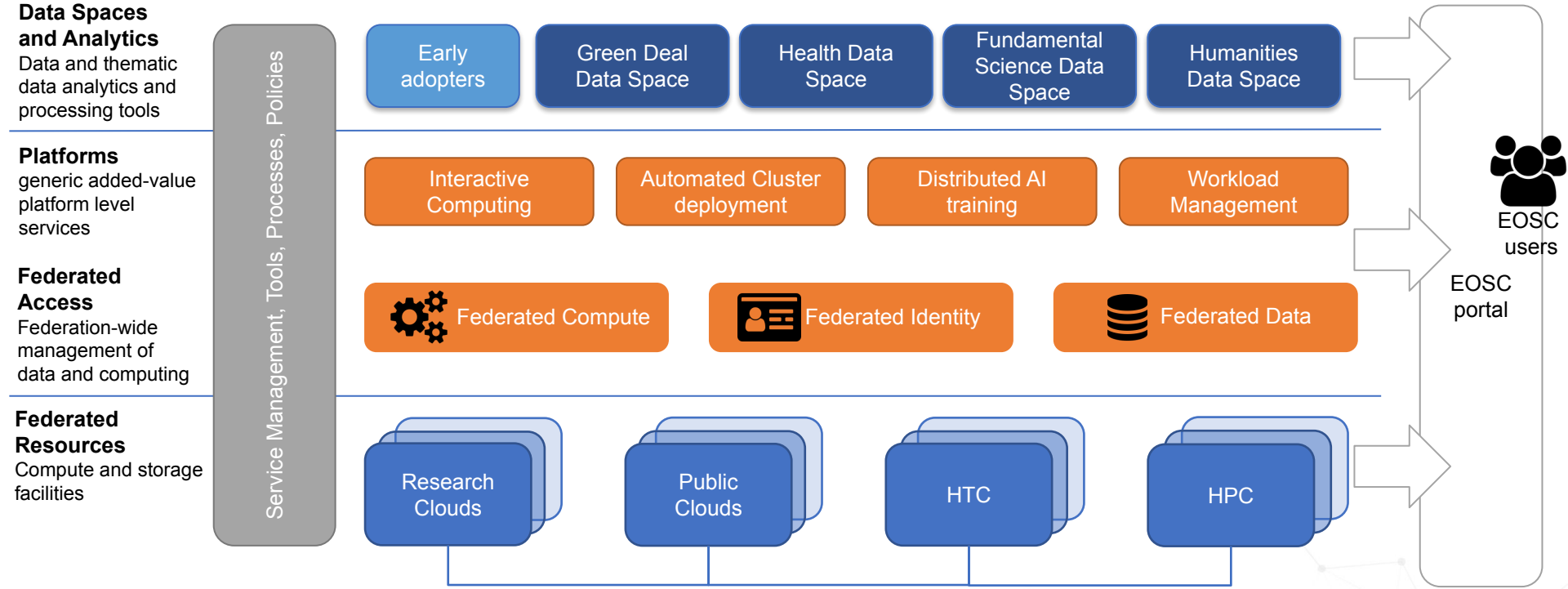
StairwAI
The Artificial Intelligence on-demand **platform for low-tech SMEs**

EGI-ACE factsheet

Implement the **Compute Platform of the European Open Science Cloud** and contribute to the **EOSC Data Commons** by delivering integrated computing, platforms, data spaces and tools as an integrated solution that is **aligned with major European cloud federation projects and HPC initiatives**.

- **Duration** 30 months (Jan 2021 - July 2023)
- **Consortium**
  - EGI Foundation as coordinator
  - 33 participating partners, 23 third-parties
- **Budget**
  - 12m EUR (8m EUR from EC, 4m EUR from partners)
  - 57% of the budget is for service provisioning

EGI-ACE concept and methodology: Tiered service architecture



Infrastructure layer – Cloud, HTC, HPC

For users and communities within the project and for new EOSC users

- Increasing capacity (Cloud and HTC):
  - **Phase 1 (month 1-10)** ⇒ 13.8 Million CPU hours/**41,300 GPU hours**/7.5 PB/month
  - **Phase 2 (month 11-20)** ⇒ 27.6 Million CPU hours/**83,000 GPU hours**/15 PB/month
  - **Phase 3 (month 21-30)** ⇒ 41.4 Million CPU hours/**124,000 GPU hours**/23 PB/month
- **Mix of funding mechanisms:**
  - Virtual Access:

Diverse locations AND compute flavors

  - Local grants enabling policy-based access (EGI Cloud Federation + new partners):

Moldova, Latvia, Georgia, Hungary, Austria, USA, China, South Africa

  - Pay-for-use:
  - Pilot HPC sites:

Upcoming EGI-ACE call for use cases

The call invites international research communities and national communities with international relevance to request services for:

- Large-scale data processing, scientific analysis, visualisation
- Hosting of large datasets and related analytics services in the cloud for EOSC users
- Federate and make accessible community-specific compute services in EOSC

The services are sponsored for users by the European Commission and by national funding agencies.

The call is open during 2021 and 2022 and submissions can be made any time.

There are 3 cut-off dates per year, each followed by an evaluation of the submitted applications. Next cut-off date: 31st of March

Further questions: egi-ace-call@mailman.egi.eu

Using GPUs in EGI Federated Cloud

GPUs resources are available on selected providers of the EGI Cloud.

Prerequisites: An account in EGI CheckIn, membership of a VO with GPU supported and have an application using GPUs

Basic usage:

- Create a VM with GPU flavor and image
- Login and execute apps on the VM
- Delete VM when done and release resources for others

IISAS, Slovakia

- IISAS-GPUCloud
- IISAS-Nebula

IFCA-CSIC, Spain

- IFCA-LCG2

LIP-INCD, Portugal

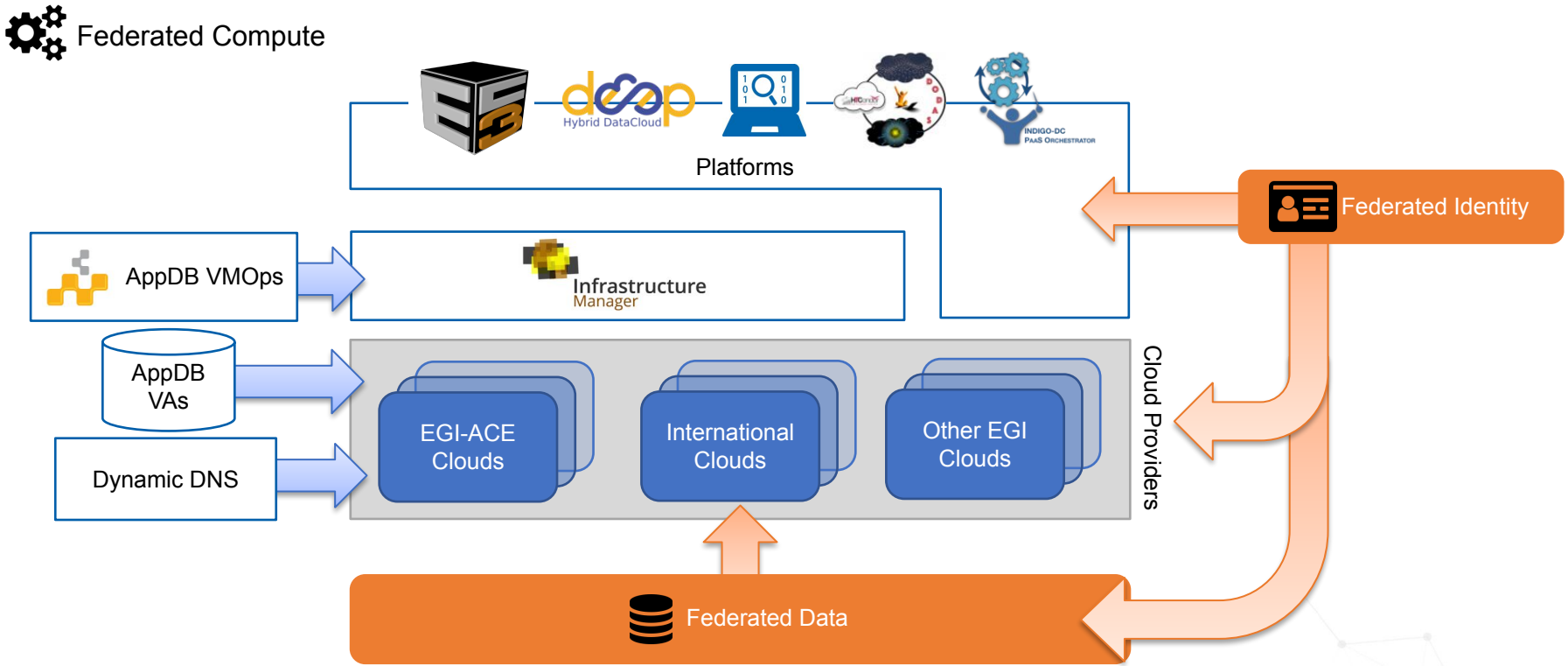
- NCG-INGRID-PT

CESNET, Czech republic

- CESNET-MC

More information: https://docs.egi.eu/users/cloud-compute/gpgpu/

Expanding EGI Cloud with EGI-ACE



DEEP as a Service

Global objective: Promote the use of intensive computing services by different research communities and areas, and their support by the corresponding e-infrastructure providers and open source projects.

DEEP as a Service (DEEPaaS): Automatic deployment of catalog modules to exploit their functionality. More: https://deepaas.deep-hybrid-datacloud.eu

DEEP Open Catalog: Sharing knowledge, models and applications between users. More: https://marketplace.deep-hybrid-datacloud.eu

DEEP training facility: 1-click development and training environments in Cloud and HPC resources. More: https://train.deep-hybrid-datacloud.eu

DEEPaaS API: Exposing models through a common and standard-based API. More: https://deepaas.readthedocs.io

SZTAKI AI

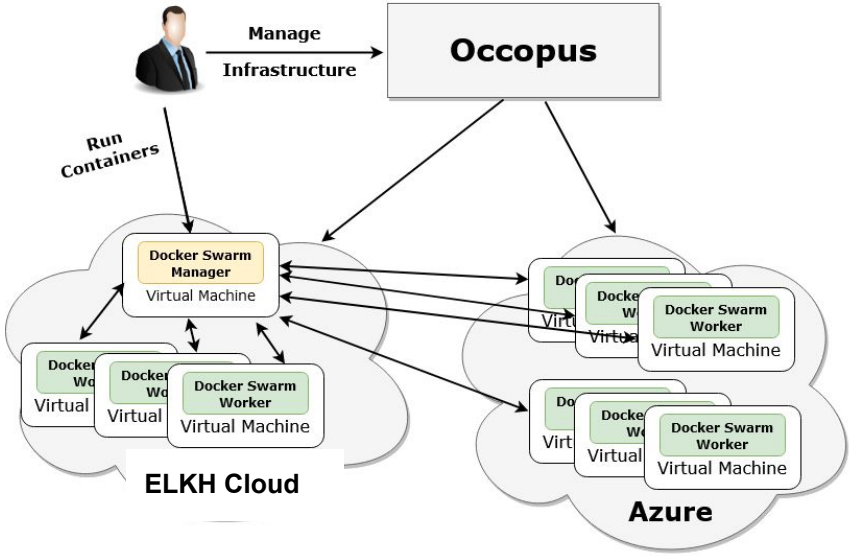
The ELKH Cloud project provides:

- Reference architectures for building typical, frequently used e-infrastructures (e.g. Kubernetes cluster, Hadoop cluster, Spark cluster, Tensorflow with GPU, etc.)
- Direct user support for building non-tipical e-infrastructures and creating a corresponding reference architecture for these e-infrastructures
- A repository for storing these reference architectures. Any ELKH Cloud user can access and use these stored reference architecture

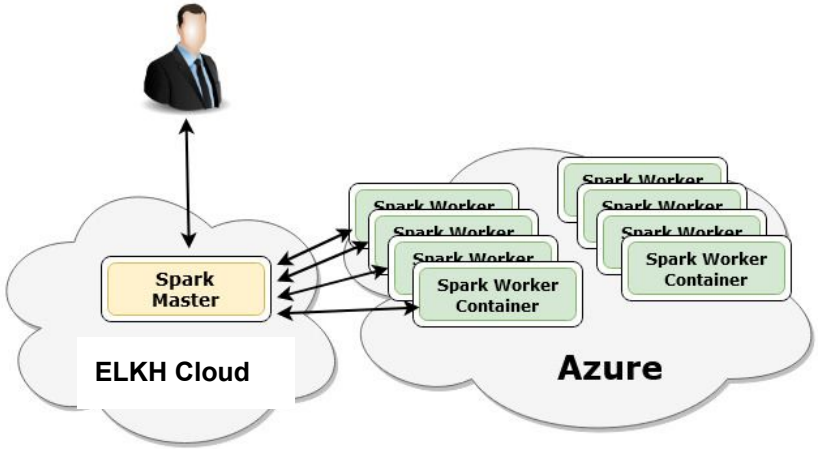
- Occopus cloud orchestrator
- JupyterLab
- Docker-Swarm cluster
- Kubernetes cluster
- Apache Hadoop cluster
- Apache Spark cluster with RStudio
- Apache Spark cluster with Python
- TensorFlow with Jupyter Notebook
- TensorFlow with Jupyter Notebook and GPU usage
- DataAvenue for large volume data transfer
- gUSE/WS-PGRADE Science gateway framework for HTC and workflow applications
- Flowbster workflow system for large volume data processing
- CQueue cluster for serving VM based HTC applications
- MiCADO for building and running scalable cloud applications

SZTAKI AI examples

Deploying docker-swarm cluster in a hibrid cloud (ELKH Cloud + MS Azure Cloud)



Deploying Apache Spark cluster on hybrid cloud (ELKH Cloud + MS Azure Cloud)



EGI AI and ML Working Group

Planned activities

- Perform a technology scouting to identify available AI/ML technical solutions
- Gather requirements from major user communities
- Identify gaps of the current solutions
- Discuss and plan enhancements for solution providers;
- Identify and engage potential resource providers willing to support the use cases

Membership

- Relevant user communities → to bring requirements
- Technology providers → PaaS solutions for ML/AI
- e-Infra providers → bringing compute and storage resources

Major European providers expected to join → AI4EU, Deep ML platforms, etc..

Interested to join as a technology provider or user communities?

Conclusions

- EGI is the pan-European e-infrastructure for researchers (External services) and for national infrastructures (Internal services)
- The EGI portfolios are expanding for AI/ML
  - GPUs and platforms via the EGI-ACE H2020 project
  - Applications and datasets via 8 H2020 projects
  - A WG on AI/ML is under setup

- **Check EGI-ACE services in the EOSC Marketplace**
- **Respond to EGI-ACE upcoming call for use cases**
- **Join the EGI AI/ML WG**

www.egi.eu
@EGI_eInfra

Thank you!

ville.tenhunen@egi.eu

The work of the EGI Foundation
is partly funded by the European Commission
under H2020 Framework Programme