# EGI: advanced computing for research in Europe... and beyond!

Peter Solagna Senior Operations Manager EGI Foundation

AENEAS kick off meeting 28 February 2017







EGI Overview - AENEAS kick off meeting 28-2-2017



# **EGI: Advanced Computing for Research**

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.

# 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



680,000

Cores

15



258,000 Terabytes of online storage



~40,000~200 Virtual Organizations

Users



Research Infrastructures collaborating



EMB.

integrated carbon observation system

WLCG Modelaide LHC Commu

w©-nmr

ISC/M

BBMRI Biotranking and Biomolecular

DARIAH-EU Digital Research Enviro

atch

ally.

el



# **EGI Membership**

**IFIN-HH** 

https://www.egi.eu/about/egi-foundation/

- Major national e-Infrastructures: 22 NGIs
- EIROs: CERN and EMBL-EBI
- EGI Foundation
  - GA :.... Gauß-Allianz (ERICs) belspo arnes grnet IICT Institute of Information and Communication Technologies <br/>
    <br/>
    <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/>
     <br/> SNIC ULAKBİM INFN **CESNE** # EENet SURF CSIC **EMBL** CYFRONET CSC CERN .

 $\frown \star$ 



# **International Partnerships**





Open Science Grid USA

# Sir

## Africa and Arabia Council for Scientific and

Canada

Industrial Research, South Africa



## *Latin America* Universida de Federal do Rio de Janeiro

26



*China* Inst. Of HEP Chinese Academy of Sciences



*India* Centre for Development of Advanced Comp.



Asia Pacific Region Academia Sinica at Taiwan



Ukrainian National Grid

EGI Overview - AENEAS kick off meeting 28-2-2017



# **EGI Federation, 2016 QR3** The largest distributed compute e-Infra worldwide





EGI Overview - AENEAS kick off meeting 28-2-2017



# **Services Catalogue**

### 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



#### Online Storage >

**Storage and Data** 

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



## http://go.egi.eu/ServiceCatalogue

EGI Overview - AENEAS kick off meeting 28-2-2017



# **Services Catalogue**

### 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



#### Online Storage >

**Storage and Data** 

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 >

Helpdesk

Service Monitoring

Monitor the performance of IT services

Transfer large sets of data from one place to another



Your point of contact to ask for support at EGI

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

**Configuration Database** Manage the configuration information of federated einfrastructure assets and their functional relations

**Collaboration Tools** IT tools for better coordination

### 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

EGI Overview - AENEAS kick

# Two (of the many) communities using Cloud Compute





The DRIHM project is prototyping an e-infrastructure to simulate extreme hydrometeorological events such as ash ooding. The National Bioinformatics Infrastructure of Sweden uses Cloud Compute to provide bioinformatics tools to their researchers, including 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



# Two (of the many) communities using HTC



The LHC experiments in 2016 produced and analysed 80 Petabytes of data per month.

WLCG is the largest resource provider and service consumer of the EGI Federation

## HADDOCK

A web portal offering tools for structural biologists

Used to model the structure of proteins and other molecules.

So far, HADDOCK processed + 130,000 submissions from over 7,500 scientists.

# Thank you for your attention.





www.egi.eu







# **E-Infrastructure services enable the Open Science Vision**











- On-demand provisioning
- Lightweight environment for maximized performance
- Standard interface to deploy on multiple service providers

- Reduce time to production by removing friction between development and operations environments
- Interoperable and transparent









# Analyze large datasets by executing large numbers (thousands) of computational tasks

- Access to high-quality computing resources
- Integrated monitoring and accounting tools to provide information about the availability and resource consumption
- Workload and data management tools to manage all computational tasks

- Large amounts of processing capacity over long periods of time
- Faster results for your research
- Shared resources among users, enabling collaborative research









- Assign global identifiers to files
- Access highly-scalable storage from anywhere
- Control the data you share
- Organise your data using a exible hierarchical structure

- Highly scalable storage system accessible from anywhere
- Easily share data
- Access through different interfaces











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

- Store data for long-term retention
- Store large amount of data
- Free up your online storage

- Stores large amounts of data
- Long-term retention
- Reliable and interoperable









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

- Move research data fast
- Specialized analytics of on-going transfers
- User interface to manage transfer and network resources

- Ideal for very large files
- Able to handle large amounts of files
- Transfer process with automatic retry





# Astronomy/Astrophysics/Astro-particle physics projects and RIs in EGI

ARGO, AUGER, CTA, KM3NeT, LHCb, LOFAR, Large Synoptic Survey Telescope/LSST, PAMELA, ESA Planck Mission, XENON etc.



<sup>- -</sup> auger - - lofar - -- magic - - - vo.cta.in2p3.fr - - - xenon.biggrid.nl

# Supporting international research communities and thematic services





Structural Biology Distribution of users (2016, QR3)

- > 2700 users
- 81 countries

(credits: A. Bonvin, WeNMR)

## Installed compute capacity trends 2011-2016





# **Services Catalogue**

### 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



#### Online Storage >

**Storage and Data** 

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



## http://go.egi.eu/ServiceCatalogue

EGI Overview - AENEAS kick off meeting 28-2-2017

# **Cloud Compute**



EGI-Engage



- On-demand provisioning
- Full control over computing resources
- Standard interface to deploy on multiple service providers

- Execute compute- and data-intensive workloads, including GPGPU computing in the cloud
- Host long-running services
- Create disposable testing and development environments
- Select virtual machine configurations and application environments
- Manage your Cloud Compute resources



# **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 hydrometeorological events such as ash ooding. The National Bioinformatics Infrastructure of Sweden uses Cloud Compute to provide bioinformatics tools to their researchers, including 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



# **EGI Federated Cloud**

- System of cloud infrastructures
- Standard user interfaces
  - Clouds and their interconnections are based on open standards, open technologies
  - Based on OCCI/OGF and OpenStack
- Harmonised operational behaviour
- Value proposition: distributed cloud computing for analysis of distributed large datasets





EGI Overview - AENEAS kick off meeting 28-2-2017



# **Cloud Container Compute**



# Run Docker containers within isolated userspace with no overhead

- On-demand provisioning
- Lightweight environment for maximized performance
- Standard interface to deploy on multiple service providers

- Reduce time to production by removing friction between development and operations environments
- Interoperable and transparent







EGI-Engage



# Analyze large datasets by executing large numbers (thousands) of computational tasks

- Access to high-quality computing resources
- Integrated monitoring and accounting tools to provide information about the availability and resource consumption
- Workload and data management tools to manage all computational tasks

- Large amounts of processing capacity over long periods of time
- Faster results for your research
- Shared resources among users, enabling collaborative research









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

- Assign global identifiers to files
- Access highly-scalable storage from anywhere
- Control the data you share
- Organise your data using a exible hierarchical structure

- Highly scalable storage system accessible from anywhere
- Easily share data
- Access through different interfaces













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

- Store data for long-term retention
- Store large amount of data
- Free up your online storage

- Stores large amounts of data
- Long-term retention
- Reliable and interoperable







**EGI-Engage** 





## **Transfer large sets of data from one place to another**

- Move research data fast
- Specialized analytics of on-going transfers
- User interface to manage transfer and network resources

- Ideal for very large files
- Able to handle large amounts of files
- Transfer process with automatic retry

