



# ELIXIR Greek Node Cloud Infrastructure

Stelios Sartzetakis, GRNET



[www.elixir-greece.org](http://www.elixir-greece.org)



 [grnet@health](mailto:grnet@health)

## GRNET's role @ a glance

- National Infrastructures for Research and Technology – GRNET State-owned company (est. 1998)
- Operates under the auspices of the Greek Ministry of Digital Governance.
- Provides advanced e-Infrastructures & Digital Services to the Greek academic, research institutions, public administration, culture and health organizations:
  - National and International connectivity;
  - e-Infrastructures & Digital Services include: networks, computing, storage, data centres
- It is the main infrastructure/service enabler for Open Science in Greece and leader in a coordinated effort for development of electronic infrastructures and services in Southeast Europe and the wider region.



# GRNET: integrated environment of cutting-edge technologies

## - e-Infrastructures – services. Contributing to the:

- Ministry of Digital Governance's goal to achieve Digital Transformation for the country, aiming to upgrade and improve the overall citizen's experience in receiving public services;
- empowerment of the research community and the growth of research activity in the country;
- enhancement of the educational processes and improvement of the educational experience;
- setting up a more extroverted Greek Research and Education community by offering equal participation opportunities for its members in ICT global actions;
- modernization of the Health sector and the improvement of the overall citizen's health service provision experience;
- shaping of new models of cooperation amongst public agencies, research and education communities, citizens and businesses;
- optimization in the use of Greek State resources in the sectors of Public Administration, Education, Research, Health and Culture.



## GRNET in numbers

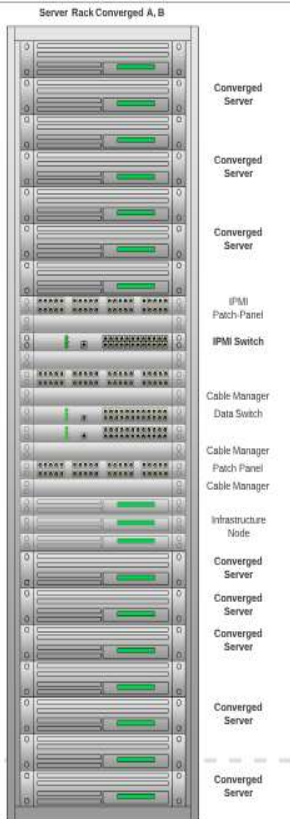
- 10000Km dark fiber backbone, 6 Datacenters in 3 distant geographic locations – 2MW, 1Tbps network capacity, 10s Pbytes storage
- 2000 servers, 10s of thousands virtual machines, 10 thousand active users
- HPC with more than 1000 projects and more than 40 thousand core years (last 5y)
- ESA Copernicus mirror site – 20Gbps traffic
- Runs Greek Internet Exchange (GR-IX) node (as trusted third party) – peering all Greek ISPs and SPs up to multiple 10Gbps (open, neutral)
- Fiber optic connections to 31+ hospitals + Services – imaging backup service, eduroam/edugain access to publications

GRNET role in ELIXIR-GR (national node of the [ESFRI European RI ELIXIR](#) / distributed paneuropean bioinformatics infrastructure):

**cloud co-design, hosting, operations and support**



# ELIXIR Greek Node Cloud Infrastructure



## 32 converged nodes

28 cores / 2.6 GHz / 512GB RAM

## 2 “fat” nodes

48 cores / 2.1 GHz / 1,024GB RAM

## 3 GPU-enabled nodes

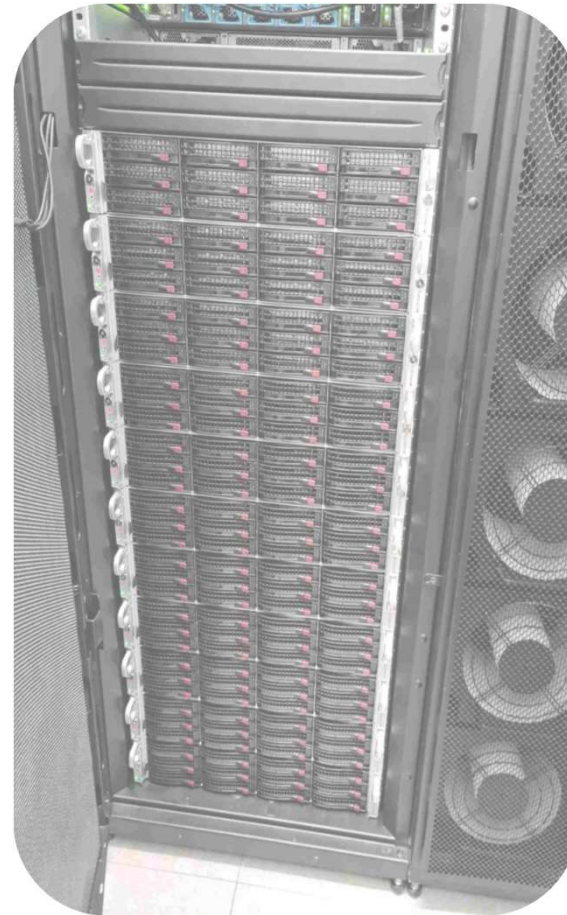
28 cores / 2.6 GHz / 768GB RAM / 2GPUs

## 8 I/O nodes

## 9 infrastructure nodes

28 cores / 2.6 GHz / 512GB RAM

**1,552 cores in total**



# Access to EG-CI: through three types of projects

<https://egci-beta.imsi.athenarc.gr/>

## 1. 24/7 services

- Providing virtual private servers (VMs) for 24/7 services

edugain open Authentication Access  
Authorization based on ELIXIR AAI



## 2. On-demand computation

- Supporting batch execution of software products

## 3. Cold storage *[Q1 2021]*

- Providing tape storage to support dataset backup

Cloud Infrastructure Manager (CLIMA)

<https://github.com/athenarc/clima>

On-demand, distributed execution of containerized  
CWL workflows (SCHeMa):

<https://github.com/athenarc/schema>



**24/7 service**



**on-demand  
computation**




**cold storage**





# Main interface

 EG - CI

DashboardUser optionsModerator optionsAdmin optionsHelp ▾Logout (vergoulis) 🔔 0 ▾

Dashboard

New projectProject requestsGold user ?













List of projects

Create new project

History of project requests


Redirection to resources management subsystem

Active projects (4)

Project	Type	Owner	Expires in	
egci-webinar	☰	You	69 days	 Update  Details  Access
CAST	☰	You	32 days	 Update  Details  Access
genome_assembly_1	☰	kostis_zagganas@elixir-europe.org	68 days	 Update  Details  Access
bufet-diploma-thesis	⚡	kostis_zagganas@elixir-europe.org	81 days	 Update  Details  Access

Expired Projects (6) ▾

# On-demand software catalogue

 Software W Admin Options Logout (vergoulis) 0

+ New image+ Existing image+ Image Request

NEW FEATURE:  
Image request

Software Name	Version	Image	Uploader	
Augur ?	translate	dockerHub: nextstrain/base	kostis_zagganas	
BUFET ?	1.0	localImage: bufet-1.0:latest	kostis_zagganas	
NAMD ?	2.12	dockerHub: researchcomputing/namd_212_multicore_osgvo-el6	kostis_zagganas	
PEMA ?	v.1.3	dockerHub: hariszaf/pema:v.1.3	hariszaf	
StochMPI ?	latest	dockerHub: spyroukostas/mpi-ssa-sbml-simulator:latest	spyroukostas	
biowardrobe-FASTQC ?	0.11.5	dockerHub: biowardrobe2/fastqc:v0.11.5	kostis_zagganas	

Working Project:  
bufet-diploma-thesis  
Remaining Jobs: 9776  
[Create new project in EG-CI](#)



### New job (BUFET v.1.0)

Input/Output directory ?

#### Arguments

miRNA-Gene interactions file		
Output file name		
miRNA query File		
Ontology file		
Number of random miRNA groups		
Number of threads		
Synonyms file		
miRanda free energy		
miRanda score		
Species (human/mouse)		
Alternative form interactions File (1/0)		
Alternative form ontology File(1/0)		



# On-demand workflow catalogue

The screenshot displays the SCHEMA SCHEDULER interface. At the top, there is a navigation bar with tabs for Software, Workflows, Data, Job history, Help, Admin Options, and Logout (vergoulis). Below the navigation bar, there is a section for "Resources from project: bufet-diploma-thesis (9776 remaining jobs)" with a "+ New project" button. The main content area shows a table of workflows with columns for Software Name, Version, and Uploader. The workflows listed are GWAS-workflow, Hashsplitter, MSA, and RNA-seq-workflow, all with version 1.0 and uploaded by kostis\_zagganas. To the right of each workflow row are buttons for Run, Edit, and Delete. A modal window titled "Run workflow (GWAS-workflow v.1.0)" is open, showing an "Output directory" field with "Select" and "Clear" buttons, and an "Arguments" section with fields for "metadata" and "variants", each with "Add" and "Remove" buttons. At the bottom of the modal are "Run" and "Run example" buttons. A blue arrow points from the "Run" button of the GWAS-workflow row to the modal window.

**SCHEMA SCHEDULER**

Software Workflows Data Job history Help Admin Options Logout (vergoulis) 0

+ New workflow

Resources from project: bufet-diploma-thesis (9776 remaining jobs) + New project

Software Name	Version	Uploader	
GWAS-workflow ?	1.0	kostis_zagganas	Run Edit Delete
Hashsplitter ?	1.0	kostis_zagganas	Run Edit Delete
MSA ?			Run Edit Delete
RNA-seq-workflow ?			Run Edit Delete

### Run workflow (GWAS-workflow v.1.0)

Output directory ?

Select Clear

### Arguments

metadata variants

Run Run example

# Using international standards

- EG-CI's SCHeMa implements:
  - GA4GH's **TES** specification for simple job execution
  - GA4GH's **WES** specification for workflow execution
- By providing **TES & WES endpoints** we make it very easy to federate our resources with those of other ELIXIR Clouds
  - We test this feature using GRNET's resources during various events.

## Task Execution Service (TES)

Every compute environment has a different API for the batch execution of tasks. For example, each of the three major cloud vendors provides this service, but using completely different APIs. By providing a common interface that abstracts over their differences, compute engines can quickly move from one compute system to the next.

## Contributors



WORK STREAMS

Cloud, Discovery, Data Security



DRIVER PROJECTS

TBD

## workflow-execution-service-schemas



**Global Alliance**  
for Genomics & Health  
Collaborate. Innovate. Accelerate.

## Workflow Execution Service (WES) API

develop branch status: build passing failed 1

The **Global Alliance for Genomics and Health** is an international coalition, formed to enable the sharing of genomic and clinical data.

## Cloud Work Stream

The **Cloud Work Stream** helps the genomics and health communities take full advantage of modern cloud environments. Our initial focus is on "bringing the algorithms to the data", by creating standards for defining, sharing, and executing portable workflows.

We work with platform development partners and industry leaders to develop standards that will facilitate interoperability.

## What is WES?

The Workflow Execution Service API describes a standard programmatic way to run and manage workflows. Having this standard API supported by multiple execution engines will let people run the same workflow using various execution platforms running on various clouds/environments. Key features include:

- ability to request a workflow run using CWL or WDL
- ability to parameterize that workflow using a JSON schema
- ability to get information about running workflows

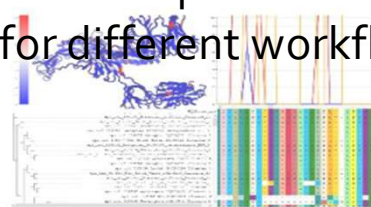
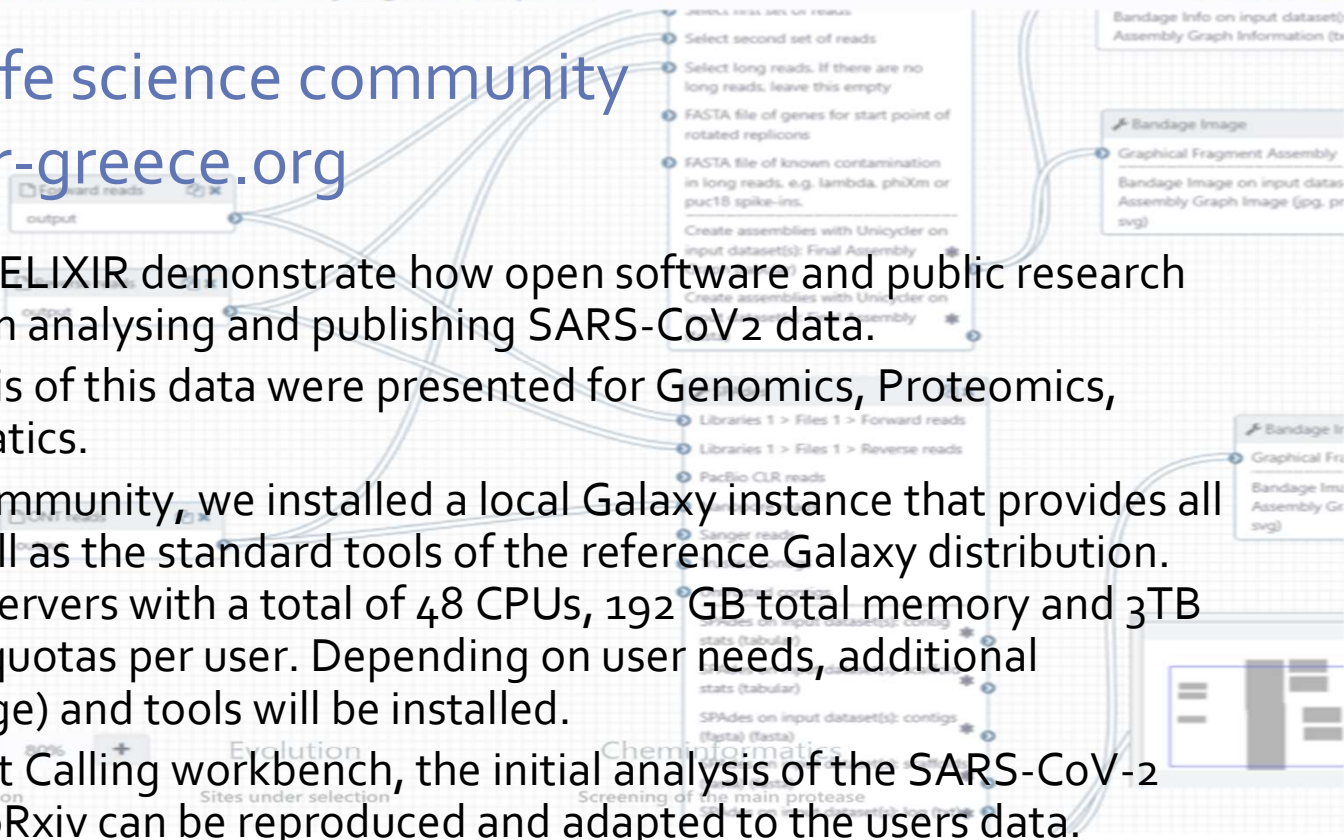


# Galaxy for the Greek life science community

<https://usegalaxy.elixir-greece.org>

- The Galaxy Community and ELIXIR demonstrate how open software and public research infrastructures can be used in analysing and publishing SARS-CoV2 data.
- Best practices for the analysis of this data were presented for Genomics, Proteomics, Evolution, and Cheminformatics.
- For the Greek life science community, we installed a local Galaxy instance that provides all tools used in webinars as well as the standard tools of the reference Galaxy distribution. The instance operates on 3 servers with a total of 48 CPUs, 192 GB total memory and 3TB disk space, initially without quotas per user. Depending on user needs, additional resources (CPU, RAM, storage) and tools will be installed.
- Within the Genomics/Variant Calling workflow, the initial analysis of the SARS-CoV-2 genome, as published on bioRxiv can be reproduced and adapted to the users data.
- Comprehensive training material for different workflows in Galaxy is available at <https://training.galaxyproject.org>

COVID-19 - Genomics [2] Assembly of genome sequence



grnet@health

elixir  
GREECE

## Local European Genome Archive



- latest local EGA release deployed
- In close contact with the Local EGA development teams (Spain, Nordic countries)
- Running on: OpenStack / Kubernetes / Docker Swarm cluster
- Scheduled to become available by end of year
- ELIXIR GR Local EGA implementation:
  - GDPR compliant
  - Two backups at different physical locations
  - On GRNET fiber network backbone
  - Isolated from other services/users
  - ISO/IEC 27001:2013 certified environment
  - Integrate with ELIXIR Beacon 2
  - Established contact with national actions:



Flagship Action For Epidemiological Study Of SARS-CoV-2





**Thank you!**  
**stelios@grnet.gr**

**FOLLOW US  
STAY INFORMED**

[www.grnet.gr](http://www.grnet.gr)  
info@grnet.gr

- grnet.gr
- grnet\_gr
- grnet sa
- grnet edyte
- grnet.gr

 grnet@health