HYPATIA's Resource Management & Access Policy

Version 1.0 (17/01/2022)

Authors

Name	Affiliation	Role
Thanasis Vergoulis	ATHENA RC	RAC member (coordinator)
Martin Reczko	BSRC A. Fleming	RAC member
Alexandros Dimopoulos	BSRC A. Fleming	RAC member
Kostis Zagganas	ATHENA RC	Consultant (technician)
Theodore Dalamagas	ATHENA RC	RAC member
Evangelos Pafilis	IMBBC HCMR	RAC member
Stelios Sartzetakis	GRNET	RAC member
Fotis Psomopoulos	CERTH	RAC member
Aristotelis Chatziioannou	BRFAA	RAC member
Elias Manolakos	NKUA	RAC member
Maria Klapa	FORTH	Consultant (tech. coordinator)

Table of Contents

Introduction

The Infrastructure

HYPATIA[1] is a Cloud infrastructure developed to support (offering computational resources) a broad spectrum of bioinformatic services, which are or will be provided to the members of the ELIXIR-GR community and to the broader community of life scientists in Greece and abroad, as well. It is currently free of use (see Section "Eligibility" for details). The infrastructure is named after Hypatia (Ὑπατία), a Hellenistic Neoplatonist philosopher, astronomer, and mathematician, who lived in Alexandria, Egypt, then part of the Eastern Roman Empire (born circa 350-370AD; died 415AD).

Under the hood, HYPATIA is comprised of the following *physical machines* (or nodes):

Node type	Characteristics	Quantity
Basic nodes	2 CPUs x 14 cores/CPU, 2.6GHz, 19MB L3 cache, 512GB DDR4 RAM, 2 x 480GB SSD	32
Hefty nodes	2 CPUs x 24 cores/CPU, 2.1GHz, 19MB L3 cache, 1024GB DDR4 RAM, 2 x 600GB HDD	2
GPU nodes	2 CPUs x 14 cores/CPU, 2.6GHz, 19MB L3 cache, 768GB DDR4 RAM, 2 GPU (16GB HBM2, 14TFLOPS single-precision, 7TFLOPS double precision), 2 x 480GB SSD	3
I/O nodes	2 CPUs x 14 cores/CPU, 2.6GHz, 19MB L3 cache, 512GB DDR4 RAM, 2 x 480GB SSD	8
Infrastructure nodes	2 CPUs x 14 cores/chip, 2.6GHz, 19MB L3 cache, 192GB DDR4 RAM, 2 x 480GB SSD	9

Basic nodes are used for simple computational needs and to implement a Ceph-based storage layer; Hefty nodes are more powerful machines (more cores & memory) that are used for resource-intensive computations; GPU nodes are machines including graphic accelerators, which can be used to speed-up several computation types; I/O nodes are machines that can better serve computations that involve many disk I/O operations; infrastructure nodes are machines dedicated to orchestrate the aforementioned compute nodes.

Regarding HYPATIA's storage, apart from the local storage of each node, the system also offers a Ceph-based storage backend (replication=3) consisting of 12 x 10TB HDD disks per Basic node (i.e., ~1PB available storage). Moreover, each I/O node contains 2 x 1.9TB SSD disks leveraged for the implementation of a bcache index.

[1] HYPATIA: https://hypatia.athenarc.gr/

Resource Allocation in HYPATIA

HYPATIA's computational resources are allocated for predetermined time periods to particular user-created projects (see Section "Supported Project Types" for more details on the supported types of projects), following the processes and restrictions described in the current document. In general, users are able to submit project requests through a dedicated Web interface (see Section "Web Interfaces"). The members of a specialised committee, called RAC (see Section "The Resource Allocation Committee (RAC)"), are responsible (among other responsibilities) to examine the user requests through a dedicated interface (see Section "Web Interfaces") and to decide about their acceptance, modification, or rejections. Notable exceptions are the automatically accepted projects (see Section "Project Review Process"); each user is able to create a particular number of automatically accepted projects.

The Resource Allocation Committee (RAC)

The members of the Resource Allocation Committee (RAC) are responsible to:

- Make decisions about the existing management & access policy of HYPATIA's computational resources and determine the respective procedures to be applied. These decisions should then be approved by the ELIXIR-GR board; the policies are subject to modifications by later RAC meetings, following the same process.
- Propose the addition or removal of RAC members. The updated RAC composition should be then approved by the ELIXIR-GR board (see below).
- Participate in the project reviewing process (see Section "Project Review Process").

One of the RAC members serves as the *RAC coordinator*. The RAC coordinator is the current Compute Platform coordinator of ELIXIR-GR.

Suggestions for RAC's composition modifications can be made anytime. These modifications should be approved by a subsequent ELIXIR-GR board meeting.

Access Policy

Objectives

The objectives of the described access policy can be summarised as follows:

- Optimise the utilisation/distribution of the available computational resources of HYPATIA.
- Maximise the throughput of served user projects and keep all users satisfied, implementing those architectures which match these objectives.
- Guarantee the elastic use of the resources allocating them on demand and releasing them when they are not needed, when possible.
- Facilitate (a) hosting of bioinformatics services on HYPATIA, (b) executing of bioinformatics software products and/or workflows on HYPATIA, (c) reproducing computational experiments.
- Monitor, review and evaluate the performance of HYPATIA and examine/suggest corrective measures for its constant improvement.

Eligibility

Access rights to HYPATIA will be granted to members of organisations participating in the ELIXIR-GR community, faculty of research & academic institutions in Greece and abroad, and employees of companies. Each individual wanting to have access to HYPATIA should have an ELIXIR-AAI account (see Section "Web Interfaces") that will give them access to the Project Request Interface (PRI) and to the Resource Management Interface (RMI) (Section "Web Interfaces").

Currently, making a bronze HYPATIA account is possible for any individual involved in an activity related to life sciences without any other restriction. Furthermore, at this phase, the use of HYPATIA is free and no fees apply to any user type (even for the industrial users). However, every individual using the infrastructure should explicitly acknowledge this fact in any relevant webpage, presentation, and publication as described in 'Acknowledge usage of Hypatia' below. Moreover, currently, CPU-time and resource usage are being monitored (see Section "Monitoring HYPATIA") but not monetised.

It should be noted that any individual who is a RAC member will also have access to the Request Administration Interface (RAI) (Section "Web Interfaces") using the same credentials. Since the composition of the RAC changes over time, each individual has access to the RAI only during their term of office as a RAC member.

Finally, based on the position of each individual in their organisation, the corresponding HYPATIA user will be classified in one of the following user types:

- Bronze user: This type of user is the one with the less privileges and the most limitations. It is the default upon user registration. It can be used for testing or educational purposes during training events (indicatively).
- Silver user: This type of user is suitable for regular lab members. Any bronze user can ask to be upgraded to a silver user by sending an email to hypatia@athenarc.gr. The request will be forwarded to one RAC member that will decide if the upgrade should be made. The RAC may also agree to downgrade a user, if particular conditions apply (e.g., in case of improper use of resources).
- Gold user: This type of user is for core ELIXIR-GR members or principal investigators that have been explicitly accepted by the ELIXIR-GR board, and it comes with extended quotas and permissions. Again, conditions for downgrade may also apply.

It should be noted that, apart from user types, there are also some user roles (e.g., administrator, RAC member) that can be assigned to users of any type. Finally, it should be noted that user upgrades are done with email requests to the administration team and are accepted by the ELIXIR-GR head of node.

Acknowledge usage of Hypatia

When the work carried out using HYPATIA's resources leads to publications, acknowledgement of ELIXIR-GR support is requested, and a copy of each publication should be emailed to ELIXIR-GR (admin@elixir-greece.org). In addition to those possible publications, the user is also encouraged to provide any material with dissemination purposes in case it is requested by ELIXIR-GR. Users must quote and acknowledge the use of HYPATIA resources in all publications related to their production and development projects as follows:

"This work was supported by ELIXIR-GR HYPATIA cloud infrastructure."

"24/7 services" projects hosting services should display HYPATIA's small logo[2] on their web pages mentioning 'powered by HYPATIA'.

Supported Project Types

HYPATIA's users are able to make requests for the following types of projects:

1) "24/7 services" projects. This type of project is suitable for hosting 24/7 services, following the Virtual Private Server (VPS) model. The focus is on relatively lightweight services (e.g., Web server, API endpoint, database); computationally intensive projects should be served by "on-demand batch computations" or "on-demand computation machine" projects (see below).

[2] Logo: https://hypatia.athenarc.gr/img/layouts/hypatia-logo-v1.png

2) "On-demand batch computations" projects. This type of project is suitable for batches of computational tasks to be executed. Each computational task could involve the execution of a particular software product or of a workflow that combines many software products. Two main ways to execute tasks are provided: (a) using pre-uploaded Docker-based and CWL-described software packages or workflows (tasks are created via a relevant Web UI), and (b) using Jupyter servers.

3) "On-demand computation machines" projects. This type of project is suitable for (short-term) computational experiments that are not containerised. A VM having particular characteristics is provided to the user for a short-term period. Project duration can be extended (after review); currently there is no limit in the number of renewals, but acceptance depends on whether other users have requested the resources and the total time already allocated to each project.

4) "Storage volumes" projects. This type of project is suitable for creating storage volumes that can be attached to VMs that belong to "24/7 services" or "On-demand computation machines" projects. In the future, it is planned for HYPATIA to also offer cold storage projects (that make use of tape storage). This functionality relies on external infrastructure (tape storage services provided by GRNET), hence it will be enabled when the respective connection will be agreed and ready to use.

Server Pools

Note on SSD storage: due to a design decision by Canonical, the additional SSD drives on I/O nodes were used to create a cache, which is used to make volume storage faster and thus they are not available for persistent data storage. However, each node (basic & I/O) offers ~300 GB of ephemeral SSD storage, available on the server for VM disks. The issue with this approach is that if a VM gets moved to another server or the server fails the data gets lost. Thus, users should always keep a backup of their data when using VMs of type I/O.

24/7 services pool:
- 14 basic machines (CPU overcommitment ratio: 1/8)
- 2 I/O machines (1/8)

On-demand batch computation pool:
- 12 basic machines (1/1)
- 3 I/O machines (1/1)
- 1 GPU machine
- 1 hefty machine (1/1)

On-demand computation machines pool:
- 6 basic machines (1/1)
- 3 I/O machines (1/1)
- 2 GPU machines

- 1 hefty machine (1/1)

Available VM flavors

24/7 service projects:

- XSmall (2 vcores / 4 GB RAM / 40GB volume storage)
- XSmall.io (2 vcores / 4 GB RAM / 40 GB ephem. SSD storage)
- Small (4 vcores / 8 GB RAM / 40GB volume storage)
- Small.io (4 vcores / 8 GB RAM / 40 GB ephem. SSD storage)
- Medium (4 vcores / 16 GB RAM / 40GB volume storage)
- Medium.io (4 vcores / 16 GB RAM / 40 GB ephem. SSD storage)
- Large (8 vcores / 16 GB RAM / 40GB volume storage)
- Large.io (8 vcores / 16 GB RAM / 40 GB ephem. SSD storage)
- XLarge (8 vcores / 32 GB RAM / 40GB volume storage)
- XLarge.io (8 vcores / 432 GB RAM / 40 GB ephem. SSD storage)

On-demand computation machines:

- Monster (96 cores / 987 GB RAM / 40GB volume storage)
- Little monster (56 cores / 484 GB RAM / 40 GB ephem. SSD storage)
- Piccolo monster (28 cores / 242 GB RAM / 40GB volume storage)
- Tiny monster (14 cores / 121 GB RAM / 40 GB volume storage)
- Super tiny monster (7 cores / 60 GB RAM / volume storage)
- GPU monster (4 cores / 16 GB RAM / 40 GB volume storage)
- I/O monster (56 cores / 484 GB RAM / 300 GB ephem. SSD storage)
- Piccolo I/O monster (28 cores / 242 GB RAM / 150 GB ephem. SSD storage)

On-demand batch computation available VM flavors:

- 1 x Monster (96 cores / 987 GB RAM / 100GB volume storage)
- 3 x Little monster (56 cores / 484 GB RAM / 40 GB ephem. SSD storage)
- 7 x Piccolo monster (28 cores / 242 GB RAM / 40GB volume storage)
- 14 x Tiny monster (14 cores / 121 GB RAM / 40 GB volume storage)
- 24 x Super tiny monster (7 cores / 60 GB RAM / volume storage)
- 2 x GPU monster (4 cores / 16 GB RAM / 40 GB volume storage)

Project Review Process

In general. Each user has a limited number of Automatically Accepted Projects (AAPs), i.e., projects that can be accepted without review: bronze users may create up to 1 AAP, silver users up to 3 AAPs, gold users up to 3 AAPs. Apart of the AAPs, all the other projects of a user should undergo a review process by RAC members; each application should receive the judgment of 2 reviewers; in case that these 2 members disagree about the acceptance or rejection of the request, a third RAC member will be invited to also review the request. The reviewers for each application are selected on a volunteer basis and in a way that avoids conflicts.

Review Criteria. Reviewers evaluate the merits of each request according the following review criteria:

1) Appropriateness of methodology. Does the request describe appropriate tools, methods, and approaches for addressing the relevant objectives?
2) Compliance with resource limits. Is the request compliant with the determined resource limits for the particular type of project and type of user?
3) Efficient use of resources. Has the request identified appropriate resources to undertake the plan with the proposed methodology? Will those resources be used efficiently and reasonably?
4) Reasonable past use of resources (when applicable). In the case of requesting modifications or time extension for an existing project, the RAC members should also consider if the hitherto use of resources was reasonable. The reviewers can utilise the resource usage recorded by the infrastructure monitoring tools (see the "Monitoring HYPATIA" section), the user visits to the webpage of the service[3] (if applicable), and the number of citations of all relevant publications (if applicable).
5) The current load of the infrastructure. Will it be possible for the infrastructure to serve a project of the determined requirements at the particular time point (based on the current load of the infrastructure)?

Users submitting requests should provide justification, suitable to the nature of the request, that addresses these criteria and the reviewers apply the criteria in the context of the request. This will be made by providing correct and adequate data in the forms of the PRI (see Section "Web Interfaces").

Response Time. The reviewers should respond to any project creation request within 2 weeks after its submission and to any project modification request within 2 weeks after its submission.

Conflicts. Since an individual could simultaneously be a RAC member and an HYPATIA user, conflicts of interest may occur. To avoid conflicts, users declare their conflicts (and the type of conflict) during the initialisation of their profile and can later edit them. Requests made by an individual cannot be reviewed by themselves or by a conflicting RAC member.

Acceptance/Rejection. In case of acceptance, the user receives a message that confirms that the project has been accepted. It is also possible that the reviewers propose a modification in the requested resources and ask the submitter for a modified request to be made. In case of rejection, a kind message, including the reasons of rejection (determined by the RAC reviewers), is forwarded to the user that made the request. The users can provide feedback to the reviewers if they believe that this feedback could change the decision. This action re-opens the request until the reviewers (the same as in the first stage) provide their final decision. The creator of a rejected request can ask for re-evaluation (based on provided feedback) only once. After the final rejection, the user has the option to resubmit a request for the same project (maybe slightly modified) after 3 months.

[3] For example, getting data from Matomo.

Resource Project Request Lifecycle

In the following paragraphs we present a resource project request lifecycle.

Project creation. The user creates a project request using the PRI interface (see Section "Web Interfaces"). If the user has exceeded the limit for AAPs (see Section "Project Review Process"), then two RAC members will review the project request through the RAI interface (see Section Web Interfaces). The RAC members decide the acceptance or rejection of the project, based on the criteria described in Section "Project Review Process". It will also be possible for the reviewers to propose modifications to the requested resources. In that case, if the user that created the request agrees with that modification, the project will then be accepted with the modified resources. After the acceptance of a project request, the user can manage the allocated resources using the RMI interface (see Section Web Interfaces). For more information about rejected projects see Section "Project Review Process".

Project modification. A user owning an active project can request the modification of its resources (or of its recorded metadata) through the PRI (Section Web Interfaces). The modification request is also processed by the RAC and it follows a similar process to the one described for the project creation case.

End of project duration. After the end of a project, all HYPATIA resources allocated to the project become automatically released. Of course, the owners of the project get several email notifications as the end of project approaches and they can request a time period extension through a project modification request.

Resource Limits (per User and Project Type)

In this section we present the resource limits per user and project type.

"24/7 services" projects

User Type	**Upper limits**
Bronze	1 project (1 AAP) 1 VM 4 cores, 8GB RAM 6 months
Silver	Unlimited projects (3 AAP) 1 VM 4 cores, 16GB RAM 12 months
Gold	Unlimited projects (3 AAP) 1 VM 8 cores, 32GB RAM 24 months

“On-demand batch computations” projects

User Type	Upper limits
Bronze	1 project (1 AAP) 500 m/job, 500 jobs, 2 cores/jobs, 16 GB/job, 1 jupyter server 6 months
Silver	Unlimited projects (3 AAP) 1000 m/job, 1000 jobs, 16 cores/job, 64 GB/job, 1 jupyter server 12 months
Gold	Unlimited projects (3 AAP) 1000 m/job, 5000 jobs, 32 cores/job, 980GB/job (max avail.), 1 jupyter server 24 months

“On-demand computations machines” projects

User Type	Upper limits
Test	-
Silver	-
Gold	Unlimited projects (0 AA) 1 VM Any of the available VM flavors. 3 months

“Storage Volumes” projects

User Type	Upper limits
Bronze	1 project (1 AAP) 1 unit 250GB 6 months
Silver	Unlimited projects (3 AAP) 1 unit 500GB storage 12 months
Gold	Unlimited projects (3 AAP) 1 unit 10TB storage 24 months

Web Interfaces

Project Request Interface (PRI).
This interface is used to submit project requests for one of the three types of projects. The main information required from the user (common to all types of projects) consists:

1. The name of the project,
2. The project end date
3. The project members

Depending on the project type other information is required, like the VM configuration or the batch projects resources. More details can be found on HYPATIA's help page.

Request Administration Interface (RAI).
Moderators: The moderators' RAI consists of an interface to view the resources of a request, approve or reject a request (if a decision has not been already reached) and modify a request (after communicating with the user). Moderators are also able to change which type of e-mail notifications they would like to receive from the system.
Administrators: The administrators' RAI includes the ability to see statistics regarding projects, view all currently active projects, send a mass notification to all users, view and delete all active VMs (as well as VMs deleted in the past), administer users and finally alter the platform configuration.

Resource Management Interface (RMI).
On-demand batch computations: Users can access the resources on HYPATIA-COMPUTE, where they can execute containerized software, CWL workflows in a distributed environment and create Jupyter servers and notebooks.
24/7 Service and on-demand computation machines: The participants of a project can create their VM using a public SSH key and then perform other operations like reboot, start and shutdown. They are also able to delete their VM. More details can be found on HYPATIA.

Authentication & Authorisation

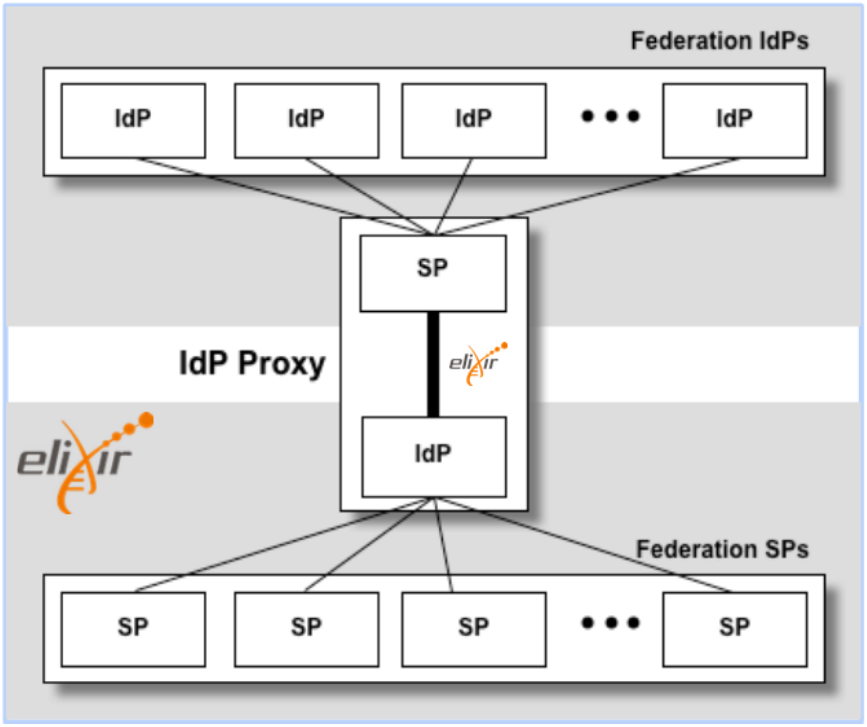
Authentication. Authentication to all described Web interfaces is based on the exchange of information between two types of machines:

a. Identity Providers (IdPs): the entities (e.g. academic or research institutions) authenticating their users and certifying their identities. Furthermore, they may provide Service Providers with the users' personal data in order to facilitate user authorization.
b. Service Providers (SPs): the entities providing services to the community. They may receive individual users' personal data, under consent, with the purpose of authorization and providing personalized services.

In contrast to older practices, where SPs stored the login credentials in order to authenticate users, the register or login function is delegated through redirection to the IdP, which authenticates the user via a secure interface. After the authentication succeeds, the IdP redirects

the user to the SP while sending along personal information about the user to the SP. The advantage of this approach is that it provides a centralized login service instead of multiple login credentials on different SPs. All communication is carried out using encryption algorithms and based on specific protocols, like OAuth2, SAML, OpenID, OpenID Connect, etc.

The authentication service that supports this architecture and is used for authentication in HYPATIA is the ELIXIR-Europe AAI. Currently the ELIXIR-Europe AAI supports two kinds of protocols, OpenID Connect and SAML. The structure of the ELIXIR AAI can be seen in the following figure:



We can see here that the ELIXIR AAI consists of an SP and an IdP which are interconnected. Whenever a user is required to login on a federation SP, they are redirected to the ELIXIR IdP, which redirects the user to their own Institutional IdP (federation IdP) to login. After the user uses their Institutional Interface to login, they are redirected again to the ELIXIR SP and their data is sent to the ELIXIR AAI. Finally, the ELIXIR AAI redirects the user (along with the required information) through the ELIXIR IdP to the federation SP from which the user initiated the login process.

Moreover, in order to facilitate login for Temporal users and users not having academic login (e.g., industrial users), a custom IdP server will be created, specifically for this function.

Authorisation. All users have access to PRI (Section "Web Interfaces") to submit project requests and modifications and to the RMI (Section "Web Interfaces") to manage the allocated resources to the active, accepted projects they own or participate in. If the user is also marked as a RAC member, they will have access to RAI, as well (Section "Web Interfaces").

Monitoring HYPATIA

Monitoring the usage of HYPATIA's resources is very important, since It is a way to infer the importance of the infrastructure and it will help the (possible) monetisation of resources for industrial users in the future.

The infrastructure takes advantage of all logging and monitoring mechanisms provided by the installed on the infrastructure software stack (e.g., OpenStack, Kubernetes) to record a great variety of relevant information. This information is aggregated to produce real-time reports that are accessible through the RAI interface (see Section "Web Interfaces").

Support

On-line Help

HYPATIA comes with an online documentation (accessible through the “Help” menu item) that describes the most important functionalities of the platform. Moreover, there is a ticketing mechanism, which can be used to provide feedback (bug reports & new feature suggestions) to the developing team.

Training Activities

HYPATIA’s users have the opportunity to attend training events (webinars and seminars) about the usage of the platform. In addition, training events (webinars, seminars etc) will be organised to help users understand containers and workflow description languages, which are central for on-demand computation batch projects. These events will contain hands-on sessions to help users transform their software components into container images and create CWL files based on their used workflows. Also a task force to help with the containerization of valuable ELIXIR-GR software products and with their description in CWL will be formed.

Federation with Other Infrastructures

Part of the resources could be temporarily allocated to test federated computing technologies with other ELIXIR Cloud infrastructures; also resources of other federated Clouds could be available to HYPATIA’s users in the future, leveraging similar functionalities.

GRNET is able to set up VPN links when requested by administrators of organizations members of ELIXIR-GR hosting compute - storage facilities at their premises. IP VPNs can immediately set up over the acces link of the organization, and if the data transfer and communications needs exceed the available bandwidth additional optical links could be requested to set up. Dedicated ELIXIR VPN links will guarantee security by traffic isolation and the deployment of distributed containerized applications that are scalable, able to support the community in a flexible manner.