

# OpenBioMaps data management service for biological sciences and biodiversity conservation

<b>Short description</b>	<p>This EAP case would like to create a service with EOSC that allows multiple users to run tasks that are above the level of a PC through the same interface. In fact, it would like to develop a "service in service" - specifically for projects that collect nature conservation and biodiversity data.</p> <p>To serve these diverse tasks this case needs a fully configurable VM which let us deploy our service interface (API) which will be available in the OpenBioMaps Network and provide computation capacity access to the involved projects.</p>
<b>Type of community</b>	Others (EAP)
<b>Community contact</b>	Dr. Miklós Bán ( <a href="mailto:banm@vocs.unideb.hu">banm@vocs.unideb.hu</a> )
<b>Interviewer</b>	<a href="#">Miguel Caballer</a>
<b>Date of interview</b>	
<b>Meetings</b>	
<b>Supporters</b>	

- [User stories](#)
- [Use cases](#)
- [Requirements](#)
  - [Technical Requirements](#)
  - [Capacity Requirements](#)

## User stories

We would like to create a service with EOSC that allows multiple users to run tasks that are above the level of a PC through the same interface. In fact, we would like to develop a "service in service" - specifically for projects that collect nature conservation and biodiversity data.

To serve these diverse tasks we need a fully configurable VM which let us deploy our service interface (API) which will be available in the OpenBioMaps Network and provide computation capacity access to the involved projects.

According to our recent experiences in our PC based local computational cluster, the number of processors is the most important in these ecological analyzes. A "typical" analysis is now running at an acceptable rate on 16 threads. The parallel computing requirements of image analysis can be much higher, and GPU usage can be interesting there. Some analyzes, for example, genetic analyzes or larger spatial analyzes require a lot of memory.

No.	User stories
US1	As a service owner, I need a fully configurable VM which let us deploy our service interface
US2	As a service owner, I want to be able to deploy the service in an automated way using the Cloud
US3	As a service owner, I would be interested in using GPUs.

## Use cases



#### Instruction

The OpenBiomaps system is deployed as a single VM on EGI Cloud Compute. A single VM will attach two pre-created volumes for persistence with a total size of 2TB. Services are launched using docker compose.

It will be evaluated the usage of the following services:

- B2FIND
- EGI DataHub
- B2DROP

Technical Plan:

[https://drive.google.com/file/d/1t4JA53-GmY2\\_nc4dfFSU6iPEfr7bsHAm/view?usp=sharing](https://drive.google.com/file/d/1t4JA53-GmY2_nc4dfFSU6iPEfr7bsHAm/view?usp=sharing)

Step	Description of action	Dependency on 3rd party services (EOSC-hub or other)
UC1		
UC2	...	
...		


## Requirements

### Technical Requirements



#### Instruction

- Requirement number: Use numbers RQ1, RQ2, RQ3, ...
- Requirement title: Use a short but descriptive title. Use the same title in the Jira ticket 'Summary' field
- Link to requirement JIRA ticket: Open a ticket in <this JIRA queue <https://jira.eosc-hub.eu/projects/EOSCWP10/issues/EOSCWP10-4?filter=allopenissues>> (click on 'CREATE' button in the middle-top of JIRA)
- Source use case: Refer back to the use cases above (UC1, 2, ...)

Requirement number	Requirement title	Link to Requirement JIRA ticket	Source Use Case
Example	EOSC-hub to provide an FTS data transfer service		UC1
RQ1			
RQ2			

## Capacity Requirements

EOSC-hub services	Amount of requested resources	Time period
EGI Cloud Compute	1 VM instances linux based: 48 CPU per VM, 96GB RAM per VM, 20 GB disk per VM, Ports open: 22, 80, 9880	Until end of EAP (end of 2020 as of now)
EGI Object Storage	1 x 2TB	Until end of EAP (end of 2020 as of now)
Infrastructure Manager	Orchestrate deployment	Until end of EAP (end of 2020 as of now)