ODYSSEA

Operating a network of integrated observatory systems in the Mediterranean Sea

Project Deliverable Report

Deliverable Number:6.5
Deliverable Title: Platform Prototype
Author(s):N.Granier
Work Package Number: 6
Work Package Title: Platform Development, Operation and Maintenance

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727277

ODYSSEA Project Information	
Project full title	ODYSSEA: Operating a network of integrated observatory systems in the Mediterranean Sea
Project acronym	ODYSSEA
Grant agreement number	727277
Project coordinator	Georgios Sylaios, DUTH
Project start date and duration	1st June 2017, 54 months
Project website	http://odysseaplatform.eu/

Deliverable Information	
Work package number	6
Work package title	Platform Development, Operation and Maintenance
Deliverable number	6.5
Deliverable title	ODYSSEA Platform prototype
Description	Description of the Marinomica platform and its future functionalities.
Lead beneficiary	CLS
Lead Author(s)	Nicolas Granier (CLS)
Contributor(s)	Simon Keeble
Revision number	V1.0
Revision Date	31-08-2020
Status (Final (F), Draft (D), Revised Draft (RV))	F

Dissemination level (Public (PU), Restricted to other program participants (PP), Restricted to a group specified by the consortium (RE), Confidential for consortium members only (CO))	PU

Document History			
Revision	**Date**	**Modification**	**Author**
0.1	26/8/2020	Initial draft	N. Granier
1.0	31/8/2020	Final draft	G. Sylaios

Approvals				
	Name	**Organisation**	**Date**	**Signature (initials)**
Coordinator	G. Sylaios	DUTH	31-08-2020	GS
WP Leaders	N.Granier	CLS	31-08-2020	NG

PROPRIETARY RIGHTS STATEMENT

Table of Contents

Table of Figures

1 Executive Summary

This document describes the content of the Marinomica platform, presently accessible from the URL https://odyssea.groupcls.com/ .

It covers the following topics:

- Software architecture aspects
- IT architecture aspects
- The operational aspects of the platform
- Next steps and future functionalities

2 Introduction

The ODYSSEA platform is currently being re-branded for its long-term future. The brand is “Marinomica” and a full communication and marketing plan is currently in development. Through Marinomica, the project creates its own unique brand and identity that is separate from the currently EU-funded project ODYSSEA. In this document, the platform will henceforth be referred to by its new name, Marinomica.

The Marinomica platform is a distributed platform offering several services to end-users to manage data centred on the Mediterranean.

The main services offered offer the end-user many possibilities, including:

- Easy discovery of data and services offered from the Marinomica catalogue;
- Visualisation of gridded and InSitu data and their metadata;
- Download of raw and processed data;
- Production of a set of data “ODYSSEA made” (Algorithms provided by the WP7 and integrated in a structure developed for the platform);
- Launch of on-the-fly algorithm processing;
- Retrieve views of interest based on a profile on a dashboard.

The platform architecture has been designed from the beginning to be open and interoperable. All functionality within the platform respect OGC protocols and thus allow the interoperability of each service, so machine-to-machine access is possible and standard.

The Marinomica platform has been designed with scalability in mind to allow the user base, data volumes, products, and services to expand and contract as required. The architecture technologies used, such as Kubernetes and Docker, will allow easy expansion of the platform, as demand dictates, by the initialisation of additional instances of the services and reduction during quieter times. The platform also has the necessary flexibility to respond to changing needs, the addition of new data, new treatments and migration to any cloud DIAS.

To date, the main functionalities and components have been developed and are available online. The platform is entering a phase of consolidation of the operation of the various services and improvement of

ergonomics. The latter part of the project development will be devoted to supplying the platform with new data from the project's own algorithms, models, and sensors as well as additional data from external data providers.

3 Overall Technical Architecture

To meet the need for scalability, since this is one of the main objectives, the platform components have a high degree of independence and are able to communicate asynchronously with each other.

FIGURE 1 : SOFTWARE ARCHITECTURE

Data from external suppliers are ingested and stored in the system. Sometimes only a gateway to the supplier's service is created. The choice is made as to the usefulness of the data and how frequently it will be used within the platform.

The Data Collection component is in charge of downloading and formatting of the gridded and in-situ data. This component automatically accesses the different services of external providers to repatriate the data. It is able to query via ftp, SOS servers and WCS:

- Manage the external data source registry;
- Check input datasets consistency;
- Standardise datasets.

This component is managed by the operating team. When a new dataset is added to the Marinomica catalog, the team configures the downloading of new data at the desired frequency with the desired nomenclature and many other specificities.

The Pre-Processing component is responsible for the internal formatting of the data.

For gridded data:

- It converts dates, coverage;
- It manages projections;
- It is able to convert grib into NetCDF CF;
- It stores the files in a tree structure specific to the project so that the other components can access the data.

For InSitu data:

- It converts the input data into a NetCDF format internal to the ODYSSEA project;
- It pushes the data to the new format in a PostgreSQL database that will be read and exploited by the SOS server.

The Post-Processing component is an innovative development within the platform. It is able to integrate algorithms in a standardised way to orchestrate them and multiply the instances of them in case of a heavy workload.

This component uses the WPS protocol to standardize the launching of treatments and therefore allowing new algorithms to be added simply. Utilising the K8S technology, it can multiply instances and manage processing in asynchronous mode. This component offers a significant added value for the future, when adding new “ODYSSEA Made” services and products. Today it supports daily exultation of the TRIX algorithm, the first “ODYSSEA Made” algorithm.

The Product Factory is a deployed component which is based on the same functionality as Post-Processing, i.e. the use of the WPS protocol, Docker and K8S technologies. This component will allow the user to launch from the front end its own processing with their own settings and thus recover their own products generated on-the-fly.

This component is ready but there are currently no algorithms deployed which use this functionality. It is expected that such algorithms will come online during the next project period.

The Catalogue is now standard on all data distribution platforms. Here the ‘Catalogue’ references to both “ODYSSEA made” products but also products available via other CSW protocol compatible portals, such as the CMEMS portal.

The Visualization component is a WMS server deployed to generate layers on-the-fly, allowing the user access to visualize maps, and related data elements, in an interactive way from the front end. Users can vary map colour palettes for the min / max projections and gives the dataset values at a given point.

The Download component gathers several tools that allow access to the gridded and insitu data. Data access is based on the standard protocols for WCS data for gridded data and SOS data for sensor data. The front end simplifies the launch of a data download, but the data access remain OGC compliant, allowing interoperability.

The User Management component is in charge of access rights management is deployed in a partitioned environment in order to comply with the GDPR regulations. This component requires some additional effort to allow the management of user profiles, password updates and some other functionalities. These evolutions will be part of the efforts to be made in the coming months.

The Front-End component is the component which is the heart of the platform. It is this component that will allow the user to access the data to the indicator to the information in a simple and obvious way. Following extensive user testing and evaluation, the User Experience (UX) is currently being appraised and a subsequent deployment is expected to significantly improve the overall system usability.

4 Physical Architecture

A comparison was made between the cost of purchasing/building a platform and purchasing services from a cloud hosting provider.

As a result of this comparison and the progress of DIAS development at the time the decision was made, it was decided that it would be more cost effective to purchase hardware and deploy the platform at CLS initially.

As a result, the platform is deployed deployed on a K8S cluster hosted at CLS. At the end of the project, the Marinomica platform can either remain hosted on the CLS hardware infrastructure, if required, or be deployed to a cloud service provider, such as Copernicus Data and Information Access Services (DIAS). All services developed will be compatible for easy deployment on a cloud platform.

Kubernetes cluster is built on multi-site infrastructure, allowing Business Continuity in case of datacenter failure. Kubernetes core components are spread across 3 sites, with compute nodes across 2 sites. Shared resources available for k8s workloads are: 48CPU and 378GB RAM Data are stored on Netapp FullS SD metrocluster system. Kubernetes version is currently v1.15.9, with an upgrade to v1.18 planned before next winter.

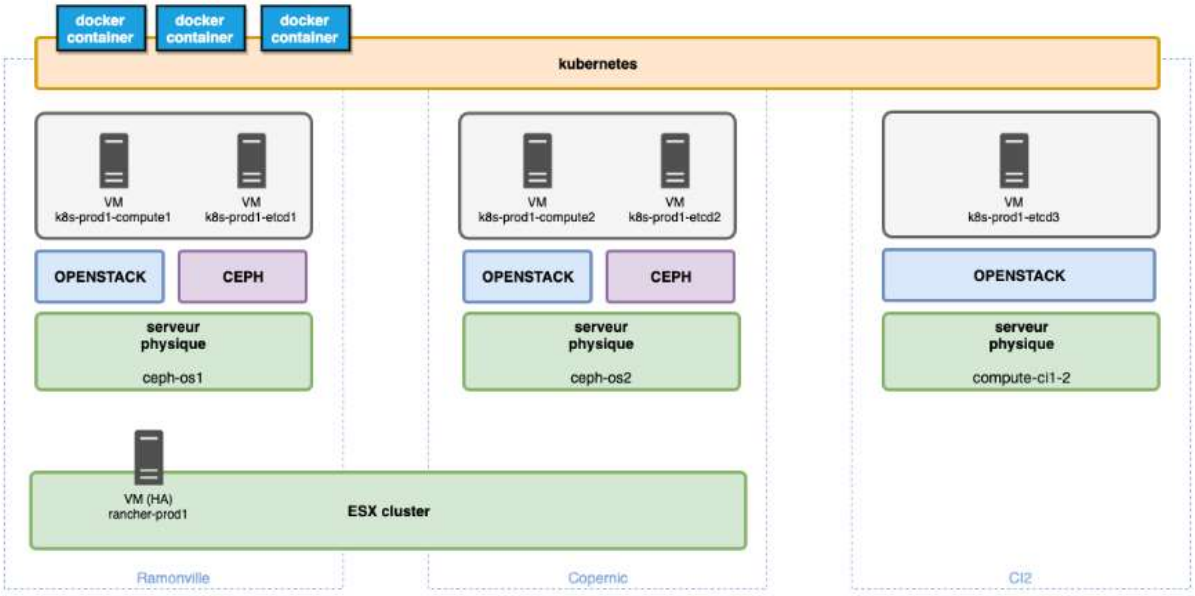


FIGURE 2 : K8S CLUSTER DESIGN

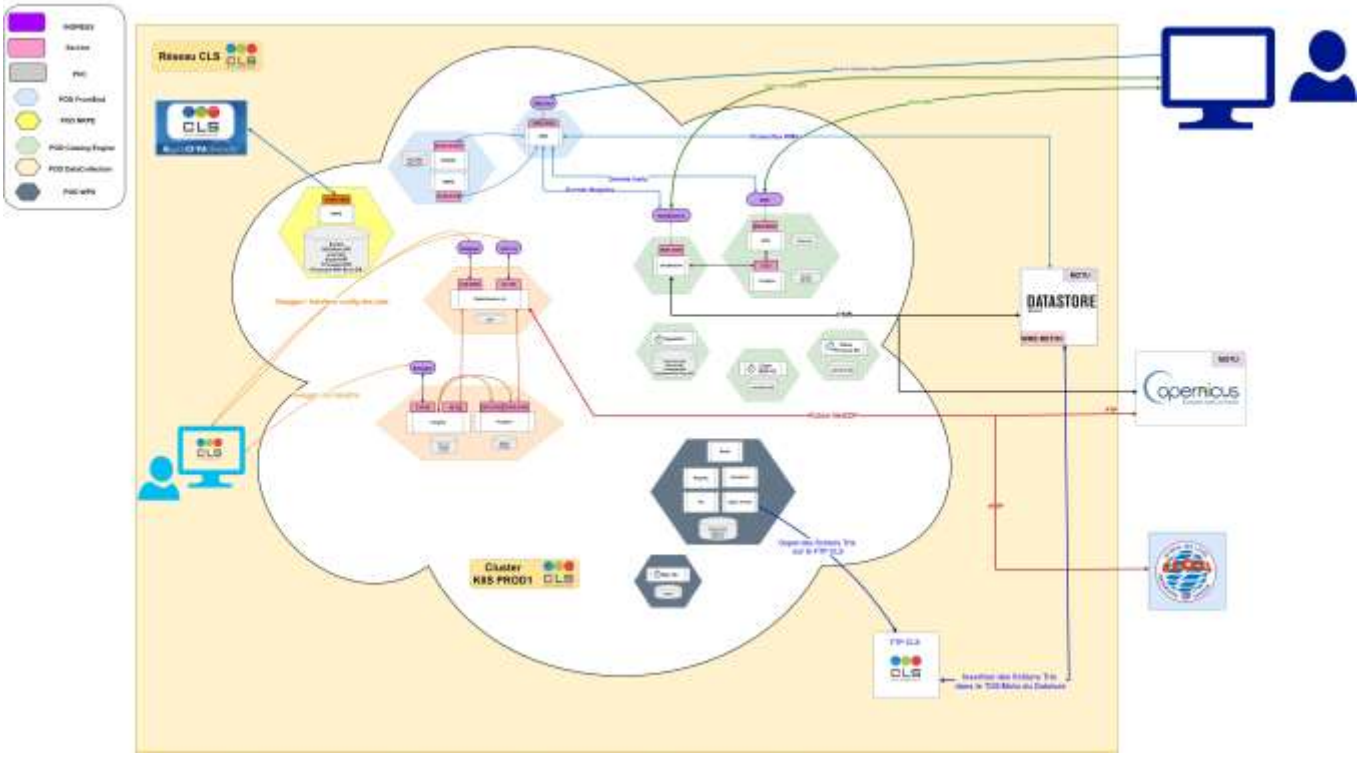
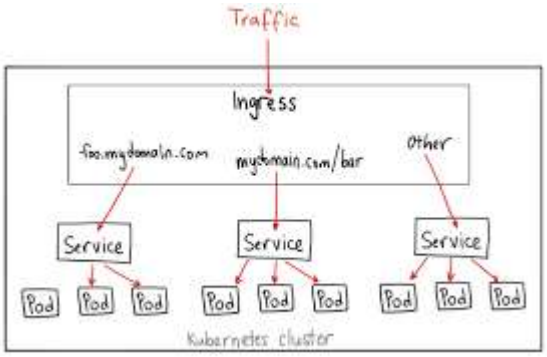


FIGURE 3 : PHYSICAL ARCHITECTURE

Physical Architecture:

- **POD**: These are the smallest deployable IT units that can be created and managed in Kubernetes. A pod usually contains a docker image, but it is possible to have several of them. This is the case in Marinomica for some components such as the catalogue-engine.
- **Service**: This is internal to the cluster (most of the time), they allow the pods to communicate with each other.
- **Ingress**: This allows access to k8s services from outside the cluster (they can also be internal to CLS networks such as data collection or external platform URL) - HTTP links
- **PVC**: Storage volume (BDD, folder needed to persist)

The diagram below summarises the main points:



- **POD Frontend**: Front end app + MongoDB/Redis
- **POD NRPE**: POD which allows Nagios checks
- **POD Catalogue engine**: SOS Server + GeoNetwork + CronJob ingestion
- **POD Data Collection**: Swagger + Hangfire + UI DataCollection
- **POD WPS**: Component developed for Post Processing and the Product Factory that facilitates the generation of data such as TRIX for example.

5 Exploitation team organisation

CLS operation Team:

At CLS, a dedicated team is in charge of ensuring the operation of the platform. The Marinomica operational team is composed of:

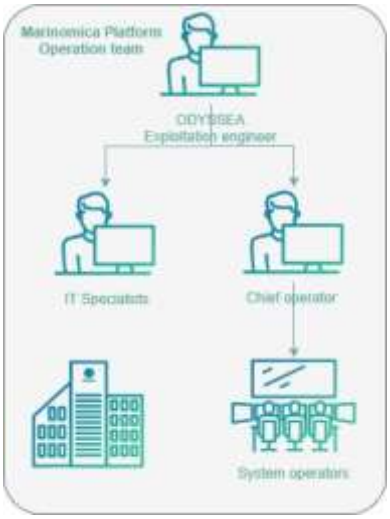


FIGURE 4 : OPERATION TEAM ORGANIZATION

- An exploitation engineers responsible to deploy the application, supervise it, but also being in direct link with Marinomica development teams (Edisoft, Hidromod, Blue Lobster).

- Backup exploitation engineers responsible to take over the application incident when the main exploitation engineer is not available.

- Operators are responsible for supervising the Marinomica platform on working days/hours and investigate at first level support stage using procedures developed by the exploitation engineers. All operational procedures are documented in Confluence to ensure procedures are followed correctly but also allowing for easy adaptation as lessons are learned.

- IT specialists in charge of all hardware aspects.

FIGURE 5 : SCREEN SHOT OF CONFLUENCES (OPERATIONAL PROCEDURES)

Any issues which are not resolvable at first level support are documented by the generation of a Jira ticket and referred to the appropriate engineer responsible for second level support. The Jira ticket ensures full traceability of the Marinomica issue.

A third level of support, if it is needed, may be engaged. This will be the originating developer (ODYSSEA project partner).

Deployment:

When a new release of a specific Marinomica component is delivered into the GitLab platform https://gitshare.cls.fr/odyssea. The exploitation engineer will deploy this package into the Quality Control environment, test it and if any issue is detected he will directly contact the development teams which will investigate based on the information provided. Once everything is tested, the new release is finally deployed into the production environment.

6 Marinomica platform prototype, features available

The functionalities currently available on the Marinomica platform are described in the user manual. Please check deliverable D8.1 “Report on Platform functionalities” http://odysseaplatform.eu/download/deliverables/D8_1_ODYSSEA_Report-on-functionality-assessment.pdf

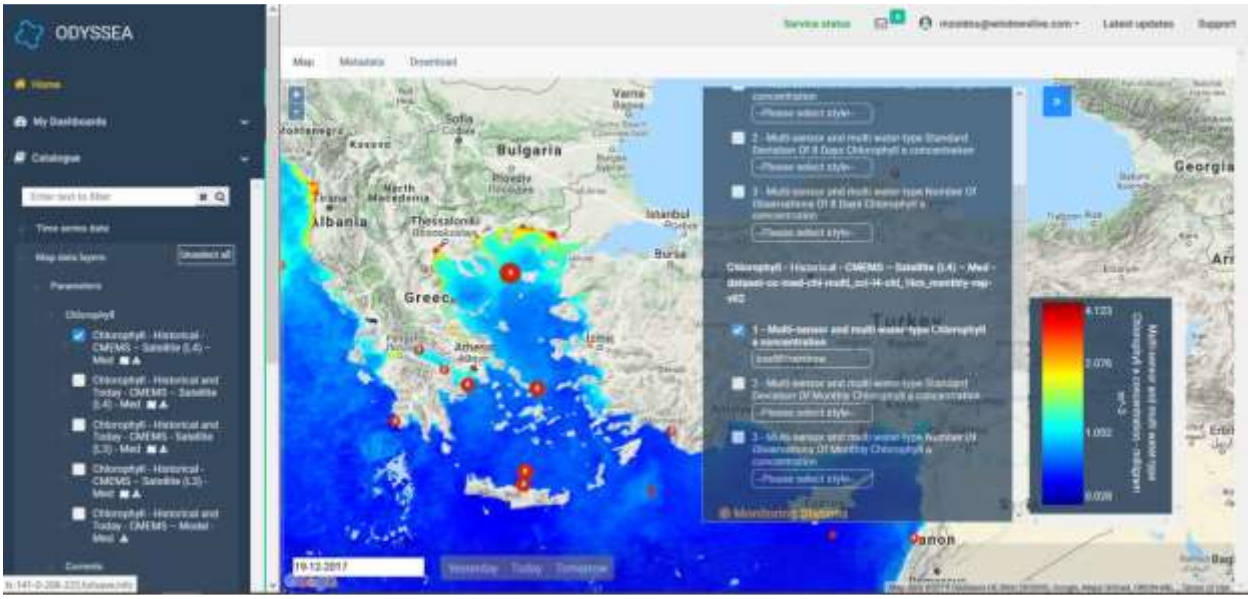


FIGURE 6 : SCREENSHOT OF MARINOMICA PLATFORM

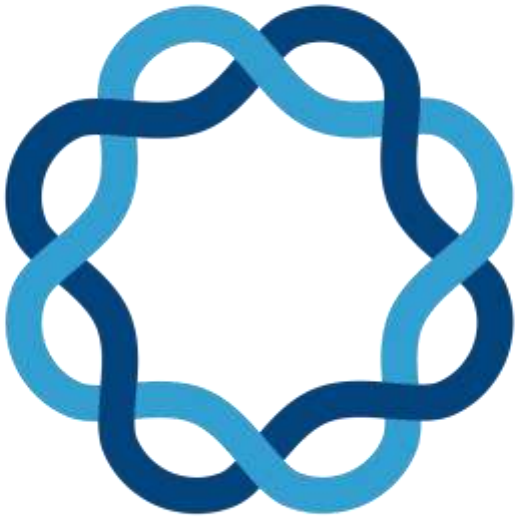
7 Marinomica platform prototype, next steps

Next steps over the next eighteen months include:

- Implement the branding change to Marinomica and deploy on the domain www.marinomica.com.
- Integrate new data from external providers
- Integrate new algorithms provided by WP7
- Integrate the outputs of all observatories
- Integrate future sensor data from the various upcoming missions
- Improve user management Authentication/Authorisation
- Correct anomalies
- Improve interface ergonomics
- Propose predefined dashboards according to user profiles

8 References

Marinomica Platform - User Guide EN (PDF)



ODYSSEA