Project Title	Prototype of HPC/Data Infrastructure for On-demand Services
Project Acronym	PHIDIAS
Grant Agreement No.	INEA/CEF/ICT/A2018/1810854
Start Date of Project	01.09.2019
Duration of Project	36 Months
Project Website	www.phidias-hpc.eu

D2.4 Definition of resources needed on the existing infrastructure and relevant of specific acquisition to implement WP4, WP5, and WP6 services

Work Package	**WP2 – Computer & storage workflows**
Lead Author (Org)	**Penalva Jean-Christophe (CINES)**
Contributing Author(s) (Org)	**Cazenave Nicolas (CINES), Charpentier Johanne (CINES), Penalva Jean-Christophe (CINES), Desconnets Jean-Christophe (IRD), Sudre Joel (CNRS)**
Due Date	**01.02.2020**
Date	**12.05.2020**
Version	**V0.9**

Dissemination Level

X	PU: Public
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Co-financed by the Connecting Europe Facility of the European Union

The PHIDIAS project has received funding from the European Union's Connecting Europe Facility under grant agreement n° INEA/CEF/ICT/A2018/1810854.

Versioning and contribution history

Version	Date	Author	Notes
0.1	12.01.2020	Penalva Jean-Christophe (CINES)	TOC and V0.1
0.2	28.01.2020	Cazenave Nicolas (CINES) Charpentier Johanne (CINES) Penalva Jean-Christophe (CINES)	Review + addon
0.3	31.01.2020	Cazenave Nicolas (CINES)	
0.4	10.03.2020	Lepaytre Marion (CINES)	Layout
0.5	24.03.2020	Jean-Christophe Desconnets	Review
0.6	10.04.2020	Penalva Jean-Christophe (CINES)	Review
0.7	11.04.2020	Sudre Joél (LEGOS)	Review
0.8	15.04.2020	Charles Troupin (ULiège)	Review
0.9	12.05.2020	Corentin Lefèvre (Neovia)	Final edition

The PHIDIAS project has received funding from the European Union's Connecting Europe Facility under grant agreement n° INEA/CEF/ICT/A2018/1810854.

Table of Contents

List of Figures

The PHIDIAS project has received funding from the European Union's Connecting Europe Facility under grant agreement n° INEA/CEF/ICT/A2018/1810854.

Terminology

Terminology/Acronym	Description
API	Application Program Interface
CINES	Centre Informatique National de l'Enseignement Supérieur
CSA	Coordination and Support Action
DoA	Description of Action
EC	European Commission
FAIR	Findable Accessible Interoperable Reusable
GA	Grant Agreement to the project
HPC	High-Performance Computing
HPDA	High Performance Data Analytics
IRODS	Integrated Rule-Oriented Data System
KPI	Key Performance Indicator
OGC	Open Geospatial Consortium
PHIDIAS	Prototype of HPC/Data Infrastructure for On-demand Services
SRB	**S**torage **R**esource **B**roker
WP	Work Package

The PHIDIAS project has received funding from the European Union's Connecting Europe Facility under grant agreement n° INEA/CEF/ICT/A2018/1810854.

Executive Summary

This document is the definition of the architecture need for the PHIDIAS project.

The aim is to define a clear architecture between partners and in the CINES' data centre.

First, we will define the services expected during the life of the project, then we will be able to define the necessary infrastructure for the CINES.

The challenge is to adress a lot of **distributed, heterogeneous and multimodal** data around producers (satelittes, in situ ...). Some of the data are already stored, and some can be transfered in real time from satellites.

With this document we can see all the relevant investments at CINES to open the services .

1 Introduction

The PHIDIAS project addresses a big challenge for Earth sciences. Today, many datasets are disseminated and researchers are not always aware about their existence, their level of quality and more generally how to access and reuse them for analysis, in combination with their own datasets to provide new information and knowledge .

The rationales of the PHIDIAS project are the following :

- Pluridisciplinarity or interdisciplinary research requires cooperative analysis and depends on data belonging to multiple stakeholders or disciplinary communities managed in silo, structured using different data formats and heterogenous data models, governed by a variety of repositories and use policies.
- Need to satisfy the data production concerns of the owners while producing interesting research outcomes by analysing those data.
- Make all these datasets available to researchers by making them respectful of the FAIR (Findability, Accessibility, Interoperability, and Reusability) principles.

The goals of PHIDIAS are mainly to :

- Open access to data wherever it is using an original data semantic interoperability strategy,
- Reconcile data and processing
- Minimize data movement
- Homogenize access to data
- Provide access to a global catalogue to provide an unified data description (metadata).

These objectives require coordination between the partners and the infrastructure.

1.1 Open access to data wherever it is

Today data is scattered in the network. Some producers keep the data, some data centres aggregate some data, some laboratory accumulate data from their researchers.
The data are not always accessible, either because of access rights or because of a lack of means of dissemination or lack of effort to organise and share the data through FAIR principles.
PHIDIAS wants to facilitate access to the data, wherever it is located.

1.2 Reconcile data and processing

PHIDIAS will facilitate processing.
Distribute the treatments to various locations and not only one per treatment.

1.3 Minimize data movement

The aim is to reduce the movements.
Every time data has to be copied from one site to another, it costs time, energy and money.
The goal is to reduce unnecessary copying.
The data can be distributed across the different PHIDIAS sites, and stored locally at the most relevant location for future processing.

1.4 Homogenize access to data

Researchers want to have access to common and open data format and share data model. This will facilitate the treatments.

1.5 Provide access to a global catalogue

Catalogue is an important component of the project. One catalogue to reference all the data and their location. This catalogue references in a unified way all the data and makes it possible to discover them.

2 Services

PHIDIAS project strives to provide many services. Most of them will be decentralized. We need to define a clear architecture.

The aim will be to access, via secure environment, to :

- Data (for upload and download),
- Data and processing chain,
- Tools allowing one to execute predefined or new treatments on data sets.

The data from the different communities that the PHIDIAS project wishes to federate are of strategic interest. It is therefore essential that the access to these data within the project is done through secure protocols, and this, whatever the type of access desired (data upload, download, catalogue access, job submission or job tracking).

Access to data should provide the ability to upload or download data securely.

- Once the architecture is well defined, and all the partners are properly integrated into this architecture, it will be possible to propose a unified and global vision of all the data disseminated in the different project sites, via a catalogue that will serve as a reference.
- The partners will have to define together the needs of the different communities in terms of data processing. The proposed tools will have to meet these needs, and will make it possible to define the processing that will be applied to the data (either in batch mode, or on data as they are produced).

The processing on the data can be define on three axis:

- **First axis**, for the interactive processing, through direct connection on the HPC machine, or through notebooks, or with a dedicated application.
- **Second axis**, processing on demand, batches of jobs. This can be used for big runs, like HPDA or HPC. The job will be executed in asynchronous mode.
- **Third axis**, event driven computing type. Producers of data can send their data just in time that they are produced (every time that a satellite send its data for example). So, the data centre must trigger and run jobs every time that data appears. The workflow must never stop, and in emergency case, it can keep data, and run as soon as the

compute part comes back to life (it can appear during maintenance phase for example).

3 Architecture

The architecture of the solution consists of different blocks. Each one with a dedicated function.
We can see all the futures services in the Figure 1 :

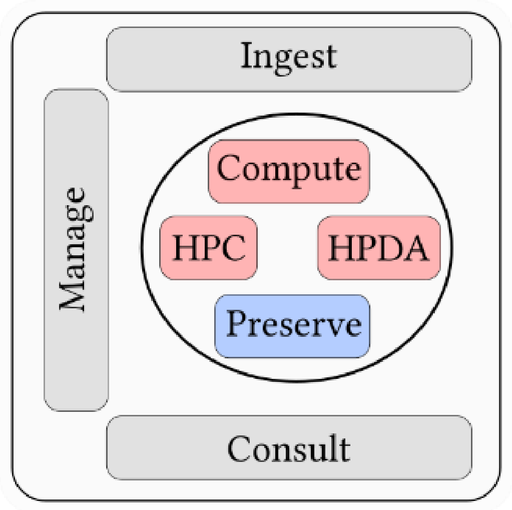

Figure 1 - Global services

- The Ingestion part, will use IRODS protocol to open the solution to others sites.
- The Consult part will be use to open access to some data or some results from previous treatment through web interface.
- The Manage part is used to pilot the solution, both from an operational and accounting point of view.

The four central part of the solution are services offered by the CINES :
- Compute, to crunch some small piece of data interactively
- HPC, to make some big compute on a lot of data
- HPDA to filter or prepare data before the run in HPC environment
- Preserve, for long-term preservation of the previous results, or data from other data centres.

The figure below displays an example of the final architecture, with all services deployed.

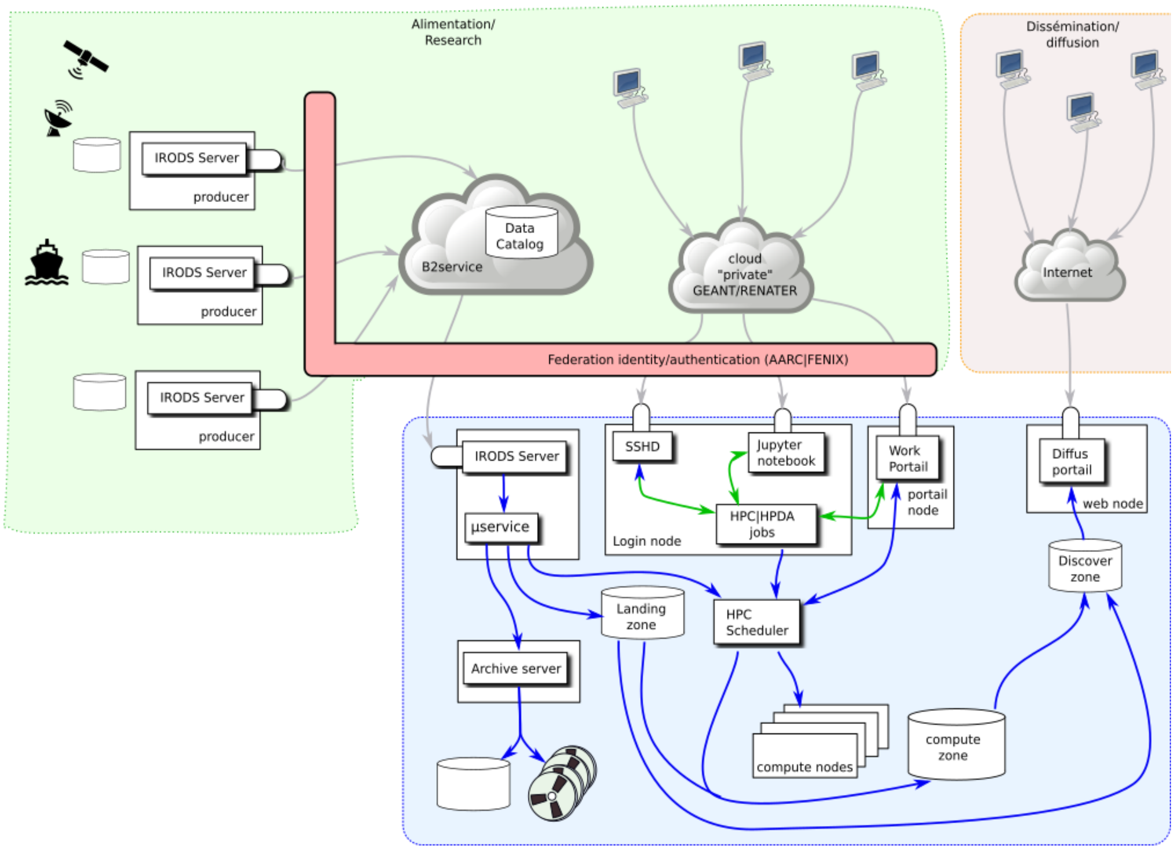

Figure 2 - Global architecture

This architecture contains three main zones:

The Green zone

That's the "secure zone". To interact in this zone, you have to be identified and authenticated. This can be made through an identity federation (with tool like AARC, or FENIX, Keycloak or RENATER identity federation).

In this zone you can find the producers of data, the storage, and the researchers involved in the project.

You can have access to a global catalogue of all the data in the different sites.

You can have access to the data processing by multiples ways :

- Directly with a connection in the data centre and with basic use of shell,
- Through notebook interface
- Through web portal interface
- There's a last method adapt on on-fly data. The data comes directly from producers and are computed to discover significant phenomena

The Grey zone :

It's the internet, with all the public who wants to discover data. The data are accessed through a web portal, some can be downloaded, and some results from previous computing too. There's no identification control, nor authentication. But all the searches will be logged.

The Blue zone :
Blue Zone represents the data centre, where it can be possible to upload data and submit jobs (through the interfaces). The data centre offers a long-term preservation facility for part of the data. The partners will decide which part will be preserved.

Volume of data
In term of data volumes, we can define different steps in the project. First, we're going to implement a prototype to open access through a SRB (**Storage R**esource **B**roker, probably IRODS). For this SRB, we need some local storage in CINES, around 100 Tb at the beginning.
This storage will be cut in three parts :

- The landing zone: it is used to store the files from others sites. In this zone, the data will not remain for a very long time (months to the max). The goal is to move data from this state to a more robust storage (in the archive, or in the discovery zone).
- The discovery zone : is used for the public web. It doesn't need to be backed-up (the data is already on others sites, or in the long term preservation).
- The compute zone : is used only for the processing. The data can stay in this zone for days, if the treatments need it.

Storage
Resource Broker provides a uniform interface to heterogeneous computer data storage resources over a network. As part of this, it implements a logical namespace (distinct from physical file names) and maintains metadata on data-objects (files), users, groups, resources, collections, and other items in an SRB metadata catalog (ICAT in IRODS) stored in a relational database (as PostgreSQL by default, or in Elasticsearch base). The metadata can be queried to locate files based on attributes as well as by name. Attributed, can be among the geographical data of the satellites.

4 Workflows

We can follow different use cases in the PHIDIAS architecture:

4.1 – Data Workflow
All the use cases are piloted by an API, or an application. This application define all the steps, from the data source to where the final results will be saved and displayed.
For example, an application can define that all the data in a specific collection have to be run through an HPC computing program, then the results will be pushed to a web service.
Or, all the data that come from a satellite service must be computed to detect artefacts, then the results will be saved to be discovered in the catalogue.
The data can be automatically pushed in a long-term preservation service from their original collection.

4.2 – Use case 1 : Data -> compute one time -> publish

The producer put the data in a collection in its IRODS server. Automatically, the data is copied from the producer IRODS to the CINES IRODS, then copied to the compute zone. After this copy, a job is running on the data, at the end the job will copy the result into a specific collection. All the data in this collection will be pushed to the discovery zone.

4.3 – Use case 2 : Data -> long-term preservation

The producer puts the data in a collection in his IRODS server. Automatically, the data is copied from the IRODS of the producer to the IRODS CINES, then the data is ingested in the long-term preservation solution at CINES.

4.4 – Use case 3 : Streaming data -> compute one time

The producer generates data all day long in their IRODS. Automatically, the data is copied from the IRODS of the producer to the CINES's IRODS then copied to the compute zone to run jobs on it. In case of specific discovery, the result will be put back in a collection.

5 Security

5.1 Goals

Safety is at the heart of all projects.
PHIDIAS does not deviate from this rule, and will make every effort to ensure different levels of security.
In storage spaces, during transfers, during remote accesses.
For this, we have determined zones. Each zone covers a community with rights.

5.2 Zones

5.2.1 Zone 1

The first zone covers the producers and consumers of data.
In this area we find the elements that allow the sharing of resources. Within the project it will be IRODS servers.
Only the sites referenced in the project can have IRODS servers integrated in the IRODS zone.
The communication between IRODS servers is done through a pool of referenced addresses and selected ports.
All the communications in and out of sites are filtered, the servers can only communicate between themselves.
The access to the data in an IRODS server in each site can only be done after an identification/authentication.
This identification/authentication will be managed through a federation. The interest of this federation is to facilitate the global addition or removal of users authorized to access the data.
Reference administrators will be on each site to accomplish this task.

5.2.1 Zone 2

The second zone covers the processing.

In this zone the data will be retrieved (and sometimes delivered) from the IRODS field. These data will be able to pass through different processing.

These processes can be controlled in three different ways:

- Via a person connected to a machine in this area, who will launch jobs (of HPC and/or HPDA calculation).
- From a notebook. But access to the notebook must still be preceded by an identification/authentication protocol.
- From a web application. Access to this application must also comply with the security rules explained above.

The data in this area can be moved, copied or even deleted.

In this area, there are several visible storage spaces:

- That of the IRODS area.
- The one used during processing.

5.2.3 Zone 3

The third zone, the most open, is the one of public access from the Internet.

This area will be reserved for public access to PHIDIAS project data.

Partners may already have a portal installed on their sites, and the new portal to CINES will allow access to some of the data that will have been processed at the centre.

The choice of the data to be accessible will be made by the partners responsible for the data and processing.

The storage space reserved for this dissemination will be accessible from the web only for consultation. The only modifications that can be made to it will come from the computing zone.

The service may be provided by standard API used by the different data producers (OPeNDAP, Dataverse, OGC data services) .

Conclusion

The construction of an *ab initio* infrastructure is not an easy task, taking into account the different constraints of the project.

The first step is to gather the needs of all stakeholders, which could be difficult especially while foreseen to provide innovative services.

The next step is to build a sustainable solution, with the technologies of the moment and that fits within the budget.

As it stands, this report does not yet provide a complete view of the work that has already been done. The compilation of requirements, still in progress, and the simultaneous construction of an infrastructure will take time. We have given ourselves a few months to build a solution that offers useful services to the community.

We will be in a position to post an initial solution by mid-2020.