



Open Forecast - Hardware and Software Architecture

© The Open Forecast Project

www.open-forecast.eu

@OpenForecastEU



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

Project	Open Forecast
Action number	2017-DE-IA-0170
Report title	Open Forecast Use Cases
Report number	OF_R5.1
Date of publication	31.05.2021
Revision	Rev_01
Lead Partner	GWDG
Authors	Sven Bingert
Dissemination level	PU ¹

1

PU = Public

PP = Restricted to other programme participants (including the Commission Services).

RE = Restricted to a group specified by the consortium (including the Commission Services).

CO = Confidential, only for members of the consortium (including the Commission Services)

Executive Summary

This report describes the software and hardware components used to derive and run the Open Forecast System as the result of the action. Mainly, the hardware used is located at one of the two sides HLRS and GWDG, while the focus of HLRS is the providing High Performance Computing Resources (HPC) and GWDG providing cloud resources for running services of the Open Forecast Service.

1 Table of Contents

Executive Summary.....	2
List of Figures.....	4
List of Tables.....	5
2 Introduction.....	6
2.1 Project Overview.....	Error! Bookmark not defined.
2.2 Project Goals.....	Error! Bookmark not defined.
2.3 Purpose of this Report.....	Error! Bookmark not defined.
3 Hardware Architecture.....	6
3.1 HRLS Hardware Architecture.....	Error! Bookmark not defined.
3.2 GWDG Hardware Architecture.....	6
3.2.1 Compute power.....	6
3.2.2 Storage.....	7
4 Software Architecture.....	8
4.1 Software APIs.....	Error! Bookmark not defined.
5 Acknowledgements.....	10
6 References.....	Error! Bookmark not defined.
Appendix A: Place the title of appendix here.....	Error! Bookmark not defined.

List of Figures

Figure 1: Generic view on the Open Forecast System	6
Figure 2: Storage architecture and the services accessing the storage. In green the different HPC systems and depicted in grey are the cloud services.	8
Figure 3: Open Forecast System workflow. It shows the different components and their interaction.	10

List of Tables

Table 1: List of services running in the cloud infrastructure 7

2 Introduction

The Open Forecast System consists of different sub-service. This report lists the underlying hardware architecture and the software components used.

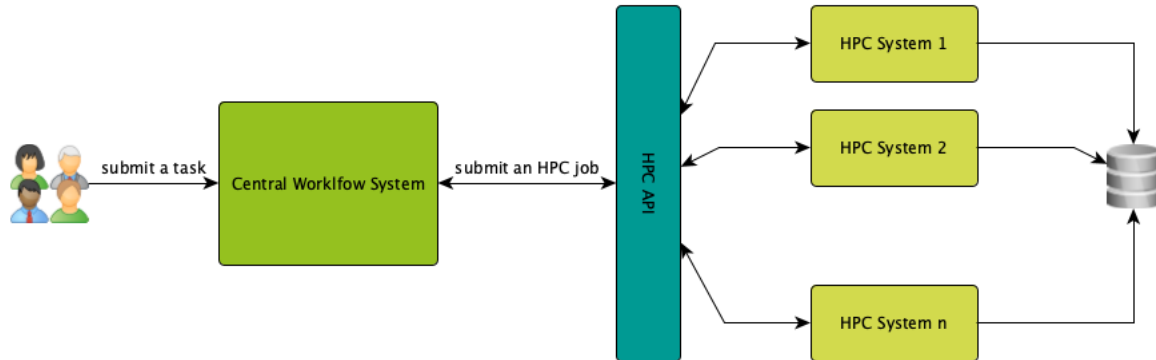


Figure 1: Generic flow on the Open Forecast System

3 Hardware Architecture

The Open Forecast use cases are deployed and executed on the infrastructure of the two sides High-Performance Computing Center Stuttgart (HLRS) and Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG). The HPC intensive task are executed on the HLRS infrastructures while the service development and service provisioning are done at GWDG.

3.1 Hardware Architecture

3.1.1 Cloud Compute power

For the development and provision of the Open Forecast related services the GWDG offers a cloud infrastructure² based on OpenStack³. With a functional account a number of virtual machines (VMs) can be configured and used. The configuration includes the number of CPUs, RAM and disk space but also placing public keys and setting the firewall rules. The GWDG offers a web UI for the configuration. Table 1 lists the VMs required for the Open Forecast. Each service has second VM of the same configuration for development purposes. A staging instance is not yet established but could be easily integrated.

3.1.2 HPC Compute Power

For the execution of the use-case the HPC resources of the HLRS are used primarily. A full description of the HLRS Systems can be found here: <https://www.hlrs.de/systems/> For testing purposes and minor tasks also the GWDG HPC resources were used. A description can be found here: <https://www.gwdg.de/hpc/> A more detailed description on the use of the HPC resources for the Open Forecast use-cases can be found the Open Forecast HPC report.

² <https://www.gwdg.de/server-services/gwdg-cloud-server>

³ <https://www.openstack.org>

Table 1: List of services running in the cloud infrastructure

Name	CPU cores	RAM (GB)	DISK (GB)	Description
Sen2Agri	16	32	200	Service for downloading and processing satellite data using Sen2Agri ⁴ software
Flow	4	8	80	Workflow engine based on flowable ⁵
GeoServer	1	16	60	Running the Geoserver to provide the data
HPC-API	2	8	40	Running the HPC-API to connect the workflow engine with the HPC resources
Proxy	4	4	40	A proxy server to easily switch between the productive and the development server
SensorAPI	4	4	40	Service for collecting the data from the sensor.community project, providing an API for data download, and storing sensor data in a graph data base

3.1.3 Storage

The data produced in the two use cases are stored in the GWDG infrastructure. The data is assigned to a functional account. The storage infrastructure is based on StorNext and the automatic backups are configured using the IBM Tivoli Storage Manager (TSM). The storage can then be mounted by the virtual machines of the cloud infrastructure. This requires to use IP addresses in a specific IP network range for VMs. Figure 2 depicts the storage architecture used in the Open Forecast project.

⁴ <http://www.esa-sen2agri.org>

⁵ <https://flowable.com/open-source/>

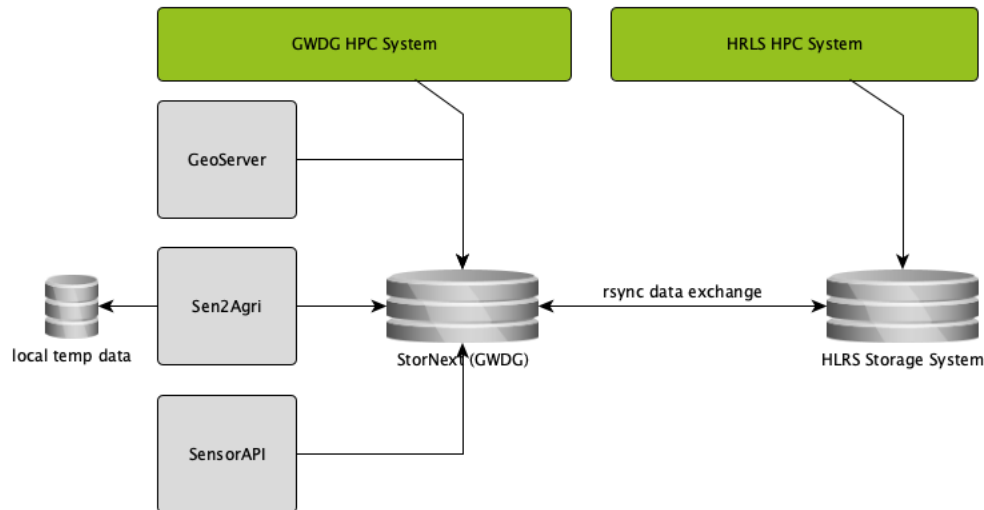


Figure 2: Storage architecture and the services accessing the storage. In green the different HPC systems and depicted in grey are the cloud services.

4 Software Architecture

The following software are chosen to deliver the services of the Open Forecast System.

4.1 Flowable

Flowable⁶ is used as the central workflow system (c.f. Figure 1). It allows to define different pre- and post-processing steps and to integrate HPC resources to be used for the execution.

4.1.1 Flowable Interfaces

- Flowable has a large set of possible interfaces. The communication with the HPC system is achieved using the “http”-task available in the workflow design. In this task a dedicated REST call to the API is configured based on the use-case requirements.
- Each workflow usually includes steps where the user or administrator is informed via email.
- For the design of a workflow, flowable offers a Web-Interface. Each flowable Web-Interface requires authentication with a local user management.
- For the execution of a workflow, flowable offers a Web-Interface too. Using this interface, the user can enter workflow specific information or can check the current status of the workflow. The status is nicely visualised by highlighting the current task in the workflow diagram.
- Workflow definitions can be downloaded as XML files for sharing and archiving purposes.
- Flowable also offers a complex REST interfaces with a multitude of end points for configuration and management of workflows and their execution.

⁶ <https://flowable.com/open-source/> (last visited 31.05.2021)

4.2 Sen2Agri

Sen2Agri is used to retrieve satellite data for post-processing.

4.2.1 Sen2Agri interfaces

Sen2Agri offers a web interface for the configuration of sites and products to be downloaded and pre-processed. The resulting data will be stored on the Open Forecast Storage (c.f. Figure 1).

4.3 Geoserver

Geoserver is used to provide data processed in the Open Forecast System. It reads data from the Open Forecast Storage.

4.3.1 Geoserver interfaces

Geoserver standardised interfaces for retrieving geo-spatial data, e.g. WMS. It also provides a REST interface for the configuration of the different services.

4.4 HPCSera

HPCSera is the middle layer between the workflow engine and the HPC systems (c.f. Figure 1). It is used to abstract the complex mechanisms of user-login, compiling and submitting jobs to the specific queuing system of each HPC system.

4.4.1 HPCSera Interfaces

HPCSera offers a REST interface for both submitting a job by the workflow engine and for pulling a job by the HPC system. The REST interface is described here <http://hpc-api.open-forecast.eu>.

4.5 SensorAPI

The SensorAPI is used to provide the data mirrored from the sensor.community network. In the Open Forecast System the data is post-processed and the new data product is also available via the SensorAPI.

4.5.1 SensorAPI Interfaces

The API reads the data directly from the Open Forecast time series data base (influxdb) and provides a REST interface for retrieving the data. The REST interface is described here: <https://sensordata.open-forecast.eu/docs/>.

4.6 GitHub

GitHub is used for managing code developed for the Open Forecast System. It is also used to design, store and provide the use-case specific containers to be submitted to the HPC queuing system. Each container contains a set of instructions. A use case can be defined using one or more containers.

4.6.1 Github Interfaces

Figure 3 depicts the Open Forecast System workflow including a gitlab repository. The job definitions sent via HPCSera define a dedicated container image to be pulled from the Gitlab

Container Registry. The Container is managed using GitLab continuous integration and continuous delivery (CI/CD) functionality.

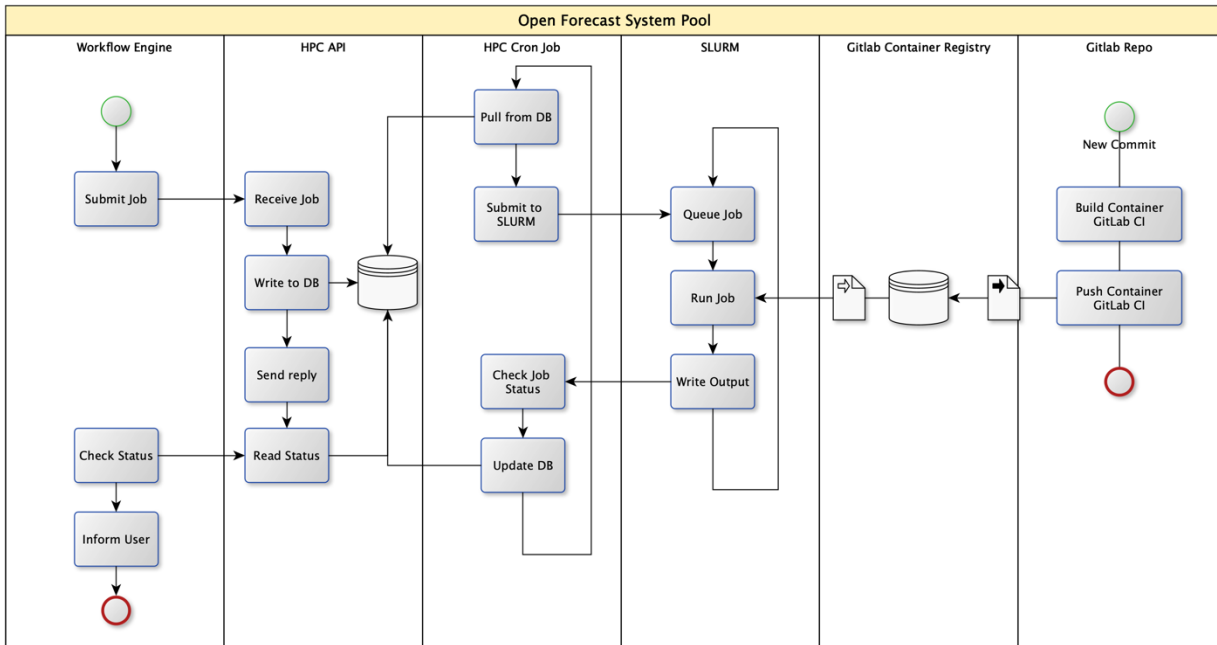


Figure 3: Open Forecast System workflow. It shows the different components and their interaction.

5 Acknowledgements

The work carried out by the Open Forecast project is co-financed by the Connecting Europe Facility of the European Union under action number 2017-DE-IA-0170.