Connecting Europe Facility 2014-2020

AQMO

Air Quality and MObility
Grant Agreement Number: INEA/CEF/ICT/A2017/1566962
2017-FR-IA-0176

D6.1

AQMO platform integration
Final

Version: 1.0
Author(s): François Bodin (UR1), Laurent Morin (UR1), Benjamin Depardon (UCit), Yiannis Georgiou (Ryax), Sylvie Perdriel (AmpliSIM), Philippe Segers (GENCI)

Date: 26 / 11 / 2020

Project and Deliverable Information Sheet

AQMO Project	**Project Ref. №:** INEA/CEF/ICT/A2017/1566962	
	Project Title: Air Quality and mobility	
	Project Web Site: http://aqmo.irisa.fr/	
	Deliverable ID: <D6.1>	
	Dissemination Level: PU *	**Contractual Date of Delivery:** 31 / 10 / 2020
		Actual Date of Delivery: 26 / 11 / 2020
	EC Project Officer: Mark VELLA MUSKAT	
Authorship	**Written by:**	François Bodin (UR1)
	Contributors:	Benjamin Depardon (UCit), Yiannis Georgiou (Ryax), Laurent Morin (UR1), Sylvie Perdriel (AmpliSIM) Philippe Segers (GENCI)
	Reviewed by:	Corentin Lefèvre (Néovia Innovation)
	Approved by:	

* - The dissemination level are indicated as follows: **PU** – Public, **CO** – Confidential, only for members of the consortium (including the Commission Services) **CL** – Classified, as referred to in Commission Decision 2991/844/EC.

Document Status Sheet

Version	Date	Status	Comments
0.1	31/10/2020	Draft V1	
0.2	18/11/2020	Draft V2	
1.0	26/11/2020	Final version	

References and Applicable Documents

List all external documents referenced in this document

[01] AQMO deliverable D2.1 Report on the preliminary sensor platform
[02] AQMO deliverable D3.1 Report on the Numerical Simulation Software
[03] AQMO deliverable D4.1 Architecture solution for HPC as a service
[04] AQMO deliverable D5.1 Workflow Management System
[05] AQMO deliverable D5.2 Report on the analysis of the visualization tools
[06] AQMO deliverable D6.3 Data Portals
[07] RUDI http://rudi.datarennes.fr/
[08] Soulhac, L., Salizzoni, P., Cierco, F.-X. et Perkins, R. J., 2011. The model SIRANE for atmospheric urban pollutant dispersion: PART I: presentation of the model. Atmospheric Environment. Volume 45, Issue 39, 7379-7395, (http://air.ec-lyon.fr/SIRANE/)
[09] IDRIS http://www.idris.fr/eng/info/missions-eng.html
[10] AWS https://aws.amazon.com/
[11] MQTT https://mqtt.org/
[12] Kubernetes https://kubernetes.io/
[13] CCME https://ucit.fr/index.php/ccme/
[14] Dataset of AQMO on the French Open Data portal https://www.data.gouv.fr/fr/datasets/mesures-des-niveaux-de-particules-fines-par-capteurs-mobiles/
[15] Dataset of AQMO on the European Open Data portal https://www.europeandataportal.eu/data/datasets/5f9691b1ffc619881d6b0bb1?locale=en

List of Acronyms and Abbreviations

Below is an extensive the List of Acronyms used in previous deliverables. Please add additional ones specific to this deliverable and delete unrelated ones.

ACB	AmpliSIM Connection Backend
AMWS	AmpliSIM Modelling Web Service
AQMO	Air Quality and MObility
AWS	Amazon Web Service
CCME	Cloud Cluster Made Easy
CLI	Command-Line Interface
CU	Central Unit
EC	European Commission
EU	European Union
HPC	High Performance Computing
IT	Information Technology
MoSP	Mobile Sensor Platform

MQTT	Message Queuing Telemetry Transport
OSM	Open Street Map
PM	Particulate Matter
RMOS	Resource Management and Orchestration System
RSS	Rich Site Summary
SME	Small and Medium Enterprise
SOAP	Simple Object Access Protocol
URL	Uniform Resource Locator
WMS	Workflow Management System
WP	Work Package

List of Figures

Executive Summary

This deliverable presents the integration of the AQMO infrastructure components. The following components have been integrated in a common platform:

1) Mobile sensor platform: this encompasses a set of sensors that are connected wirelessly with a central unit that is in charge of storing, processing and communicating the collected data.
2) Primary server: this component is in charge of storing the data before their long-term storage. It also provides support for that data visualization and platform monitoring capabilities.
3) Open data repository: this part is related to the Rennes Métropolis Open Data initiative currently developed by the RUDI project. The Rennes Métropolis Open Data is linked with the French Open Data portal (see [06] et [14]) and the European Data Portal (see [06] and [15]).
4) Simulation as a service: This component provides SIRANE simulation access as a service. SIRANE is the numerical model for pollutant dispersion.
5) Resources management: The element of the platform aims at dealing with computing resources.
6) HPC resources: This is the resource provided by the IDRIS supercomputing center.
7) Cloud resources: In the case of AQMO we use HPC clusters deployed on Amazon Web Services (AWS) through UCit’s software CCME (Cloud Cluster Made Easy)
8) SDN/VPN: Many devices in AQMO are enrolled in a Software Define Network that provide the interconnection with encryption capabilities between the devices and the primary server.

The platform is now operational and buses are collecting data everyday. Numerical simulation based on the SIRANE numerical model can be used as a service using a REST API.

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

Introduction

The air quality monitoring and its improvement in an urban area requires the full cooperation of local administrations, citizens, and industrial partners. The AQMO project objective is to provide an end-to-end platform able to collect the air pollution measurements, perform numerical simulation of pollutant dispersion and provide access to these informations in a convenient and timely manner contributing to the creation of a numerical twin of the city. Two critical aspects of this end-to-end platform are resources management and data logistic.

A critical challenge for AQMO is the integration of numerous heterogeneous hardware and software components. The approach chosen has been to couple very loosely the pieces when this is not impacting performance. The goal of the approach is to allow for maximum reuse of the components outside AQMO, still preserving a coherence in the platform.

The initial objective of 20 sensor equipped buses has been reduced to 10 after a careful analysis from Keolis. The area coverage objective could be achieved by selecting the right bus lines and taking into account that a bus is deployed on different lines from one day to another.

The map in Figure 1 shows the current coverage with 8 buses in november 2020.

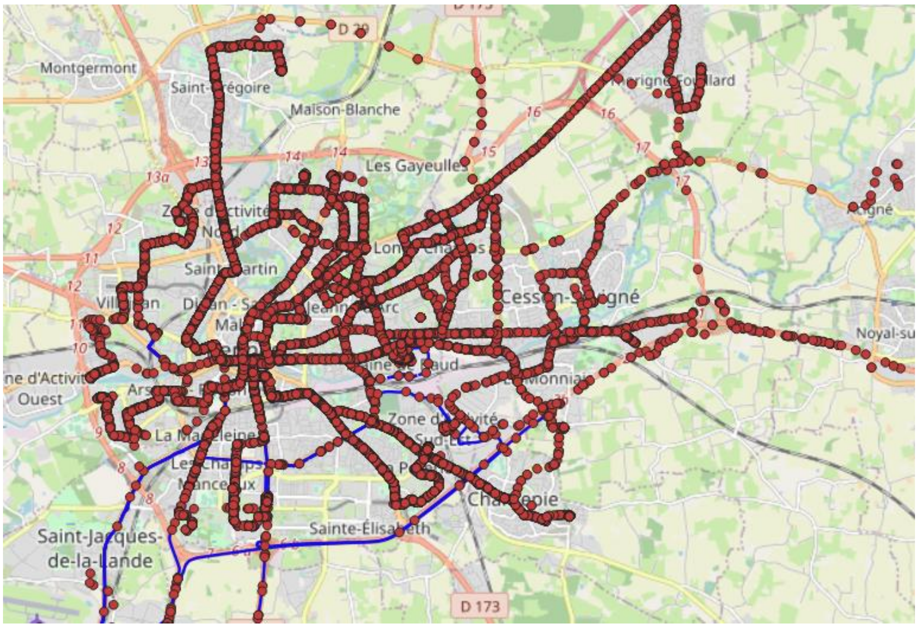

Figure 1: In red, measurement coverage with the first 8 buses equipped with AQMO sensors (two installations have been delayed because of the Covid-19 lockdown).

The document is organized as follows. Section 1 provides an overview of the components of the platform. The next sections describe each component of the platform. They are followed by the integration strategy we have implemented. We conclude with the current status of the platform.

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

1. Platform overview

Figure 2 gives an overview of the technical components of the platform.

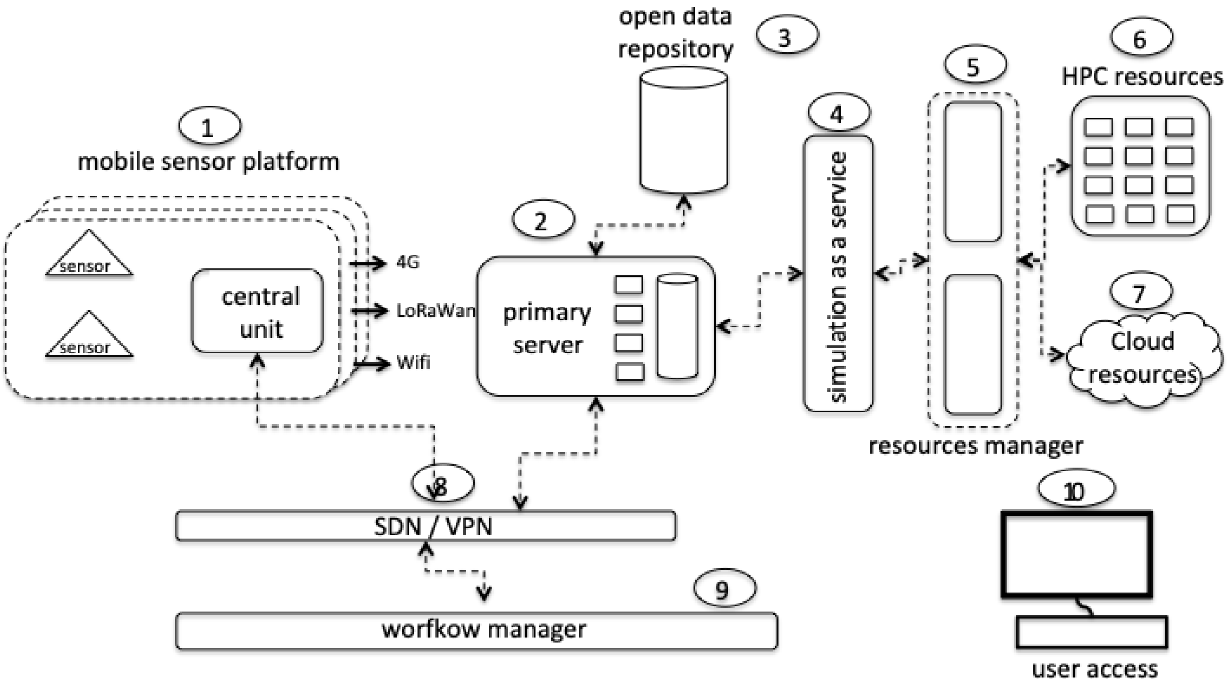

Figure 2: AQMO platform overview.

The components in Figure 2 are the following:

1) Mobile sensor platform: this encompasses a set of sensors that are connected wirelessly with a central unit that is in charge of storing, processing and communicating the collected data.
2) Primary server: this component is in charge of storing the data before their long-term storage. It also provides support for that data visualization (see [05]) and platform monitoring capabilities. This part is related to the “Fog” topic.
3) Open data repository: this part is related to the Rennes Métropolis Open Data initiative currently developed by the RUDI project [07].
4) Simulation as a service: This component provides SIRANE [08] simulation access as a service. SIRANE is the numerical model for pollutant dispersion.
5) Resources management: The element of the platform aims at dealing with computing resources.
6) HPC resources: This is the resource provided by the IDRIS supercomputing center [09].
7) Cloud resources: In the case of AQMO we use HPC clusters deployed on Amazon Web Services [10] (AWS) through UCit’s software CCME [13] (Cloud Cluster Made Easy)
8) SDN/VPN: Many devices in AQMO are enrolled in a Software Define Network that provide the interconnection with encryption capabilities between the devices and the primary server.

9) Workflow manager: This software platform is in charge of orchestrating the different automations and data analytics workflows (pipelines) taking place at the edge (buses), fog (primary server) or HPC/Cloud resources..
10) User access: User accesses provide monitoring and data visualisation [05], etc.

The next sections provide a short description of the components.

2. Mobile sensor platform and primary server

The platform used two types of sensor carrier to achieve measurement mobility. The first type is buses from the STAR transportation network (https://www.star.fr/). The buses allow us to cover more areas of the metropolis. The second type is drones that can be used to measure the vertical gradient of pollutant (in our case particulate matter - PM).

The remainder of this section is divided in three parts. The first paragraph shows the buses install, the second presents the drone install. The last part gives an overview of the digital IoT components. More details can be found in AQMO deliverable D2.1 Report on the preliminary sensor platform.

2.1 Buses

Figure 3 shows the setup of the sensor (see Figure 7) on the bus body. The sensor communicates with the central unit (see Figure 6) using a Wifi connection.

Figure 3: Sensor on the roof of a bus.

2.2 Drones

Figure 4 shows the installation of two sensors on a DJI Matrice 100 drone. Figure 5 presents the first measurements obtained using this setup. As it can be seen, the PM10 measurements tend to vary a lot compared to the PM2.5. More studies are necessary to understand these numbers.

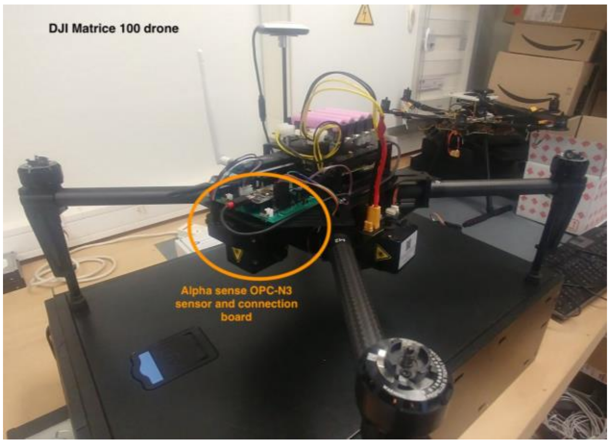

Figure 4: DJI Matrice 100 equipped with an OPC-N3 sensor and data management and communication workload.

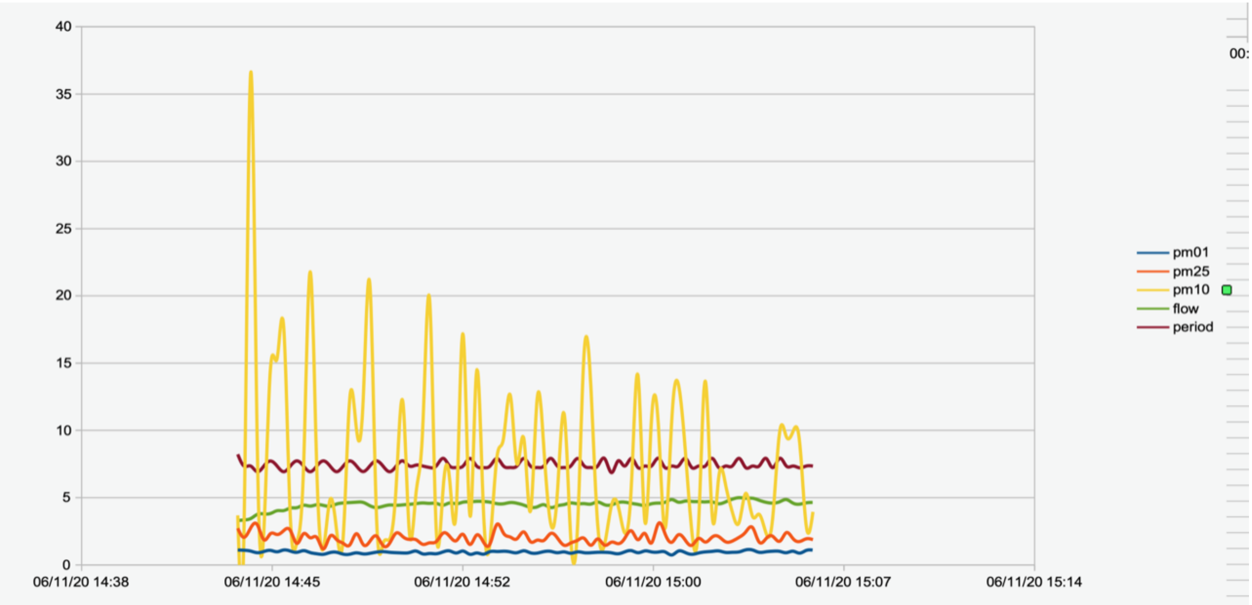

Figure 5: First PM measurements with a drone.

2.3 IoT components

In the centre of the MoSp (mobile sensor platform), the CU (Central Unit) is the heart of the system and it centralizes the data storage and processing of all measurements operated by the platform. It also provides a reliable time-stamping and a global positioning mechanism. The data processing workflow, eventually complemented with some data analytic tasks, is operated under its control. It is also exclusively in charge of all communications and interactions with the global AQMO platform. All sensor modules are satellite micro-systems connected directly and exclusively to the CUs. A sensor module must be connected to the central unit via a fully specified hardware and software API. All sensors are gathered in a Software Define Network (SDN) and are inter-connected globally using VPNs.

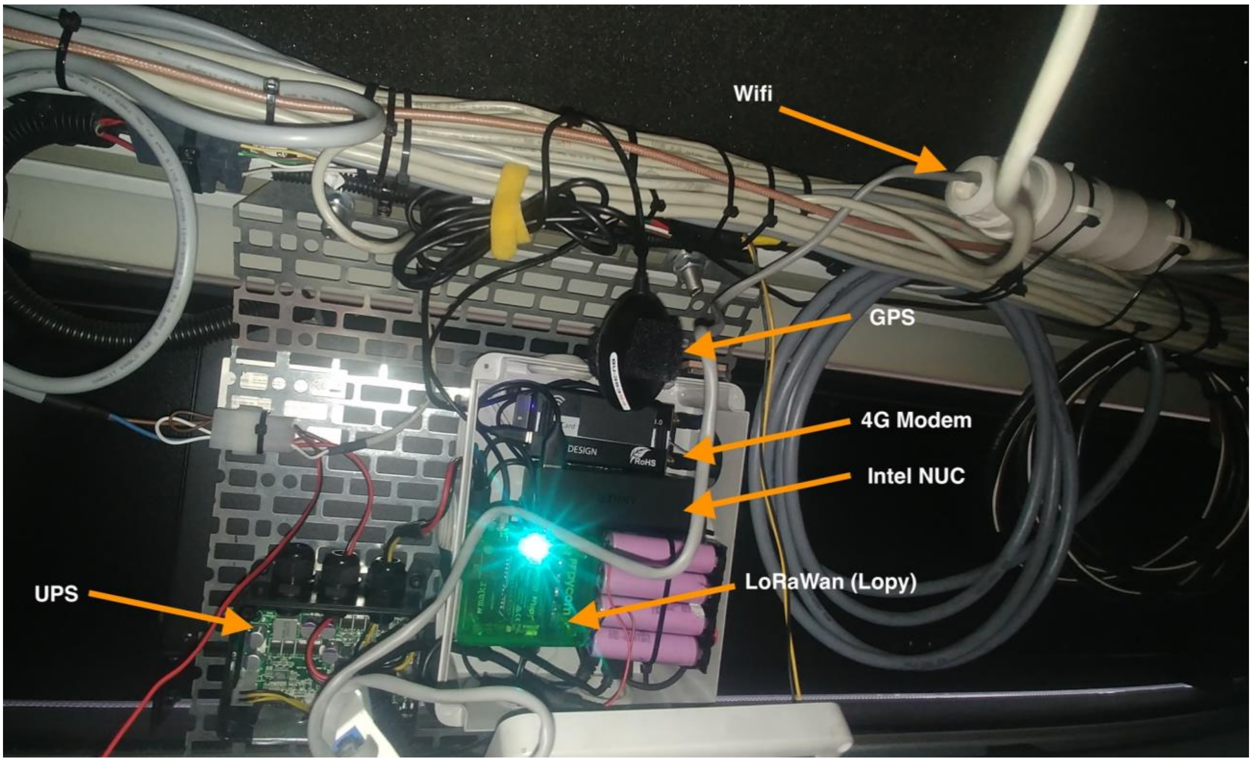

Figure 6: Central unit setup inside a bus (first prototype).

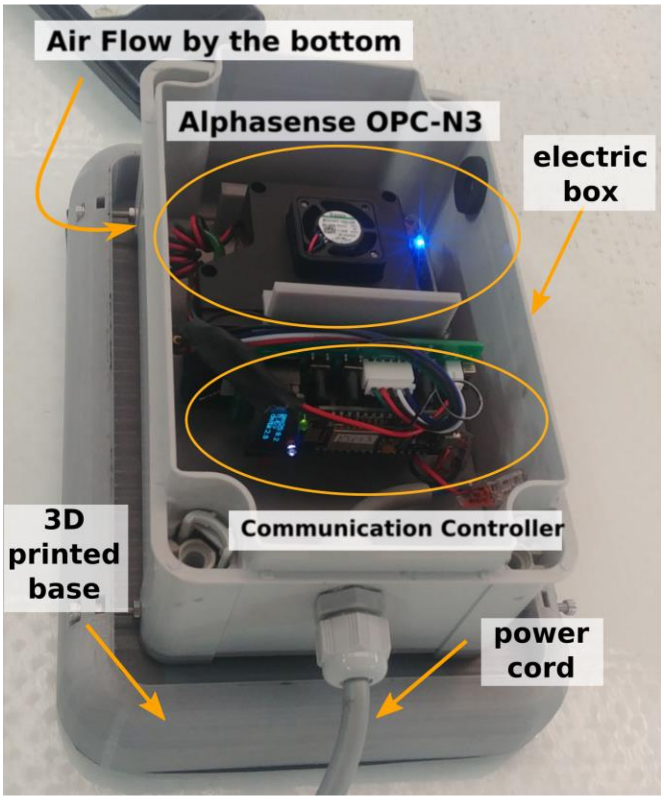

Figure 7: Sensor setup.

2.4 Sensor visualization

The sensor visualization is provided using a map. The interface provides a way to visualize the coverage and the localisation of mobile sensor platforms, and a basic interface for data visualization (the list of sensors). The real-time localization map interface is shown in Figure 8 This interface also provides a REST API to access the data. The interface and API are described in the document “AQMO Sensor Platform, The Sensor Data Interface”. More advanced and data-oriented visualization are under development in order to address the need of different data end-users: maintenance and monitoring, pollution estimation, search over historic measurements.

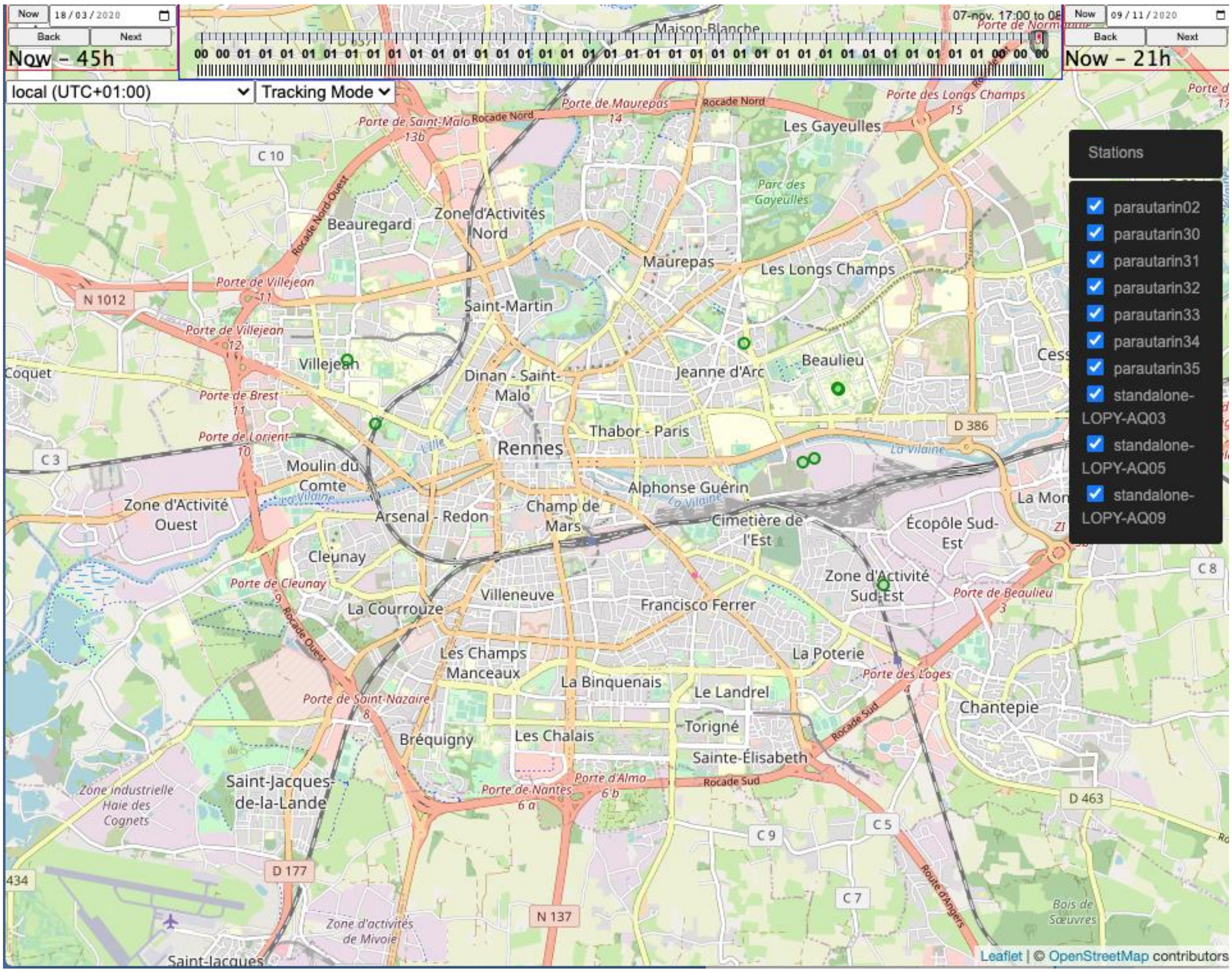

Figure 8: Real time map displaying sensor positioning. Green dots indicate where the sensors are in quasi real-time (few seconds latency).

3. Interoperability with the RUDI Open Data Repository

The AQMO platform is interoperable with the RUDI data portal [07] currently being developed by Rennes Metropoles. The data portal is a federation of producers that are conformant to an interface contract. The architecture of RUDI is illustrated in Figure 9.

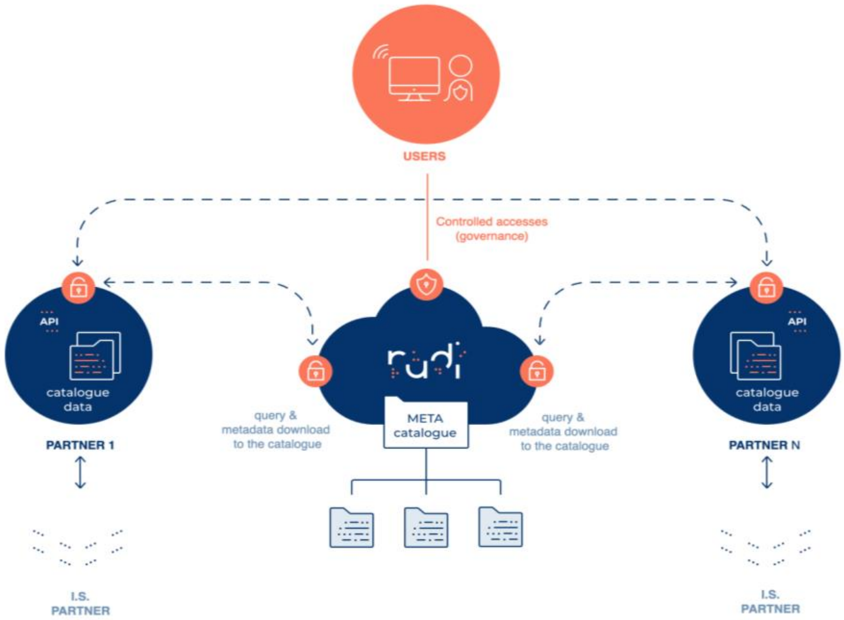

Figure 9: RUDI data portal (FR), Source: RUDI.

RUDI provides an API to publish data to the catalog of the data portal. The AQMO data sets will be published using this API. The first RUDI system will be available mid-2021.

4. HPC resources management

This component allows the delivery of HPC as a Service to end-users and third-party services through "plug-and-play" services. This platform is designed to comply with AmpliSIM's use-cases and requirements (job and data management, types of services, API compatibility etc.) – AQMO's simulation service provider – and CNRS' security constraints and requirements to access their Supercomputer at IDRIS – AQMO's resource provider.
Access to the computing platform is done through web services APIs, allowing any authorized software tool to run air quality simulations and access generated data. The "HPC services" are provided through the integration of an HPC Portal (NICE EnginFrame) that abstracts the submission of HPC jobs to an HPC cluster by embedding them in web services. These Web Services are accessible through SOAP APIs and a CLI (used by the workflow management system).

On top of this abstraction, provided by the APIs, the platform also embeds a prediction tool that will help users in the selection of the better computing platform depending on their nature (HPC centre or Public Cloud), and submit their jobs more easily – ideally in a transparent way – while taking into account their operating constraints (cost, time to result, access to data…). This tool currently estimates the execution time, waiting time and time to result of the jobs on the available platforms, prior to their submission - allowing to select the "best" target platform based on selected criteria.

Access to HPC resources in the public cloud is done through the dynamic creation of HPC Clusters by CCME, which relies on AWS' APIs. CCME provides a CLI that is used by the workflow management system.

Finally, the API Gateway depicted in Figure 10 is currently the APIs of the Workflow Management System, that allows the integration with external tools such as AmpliSIM.

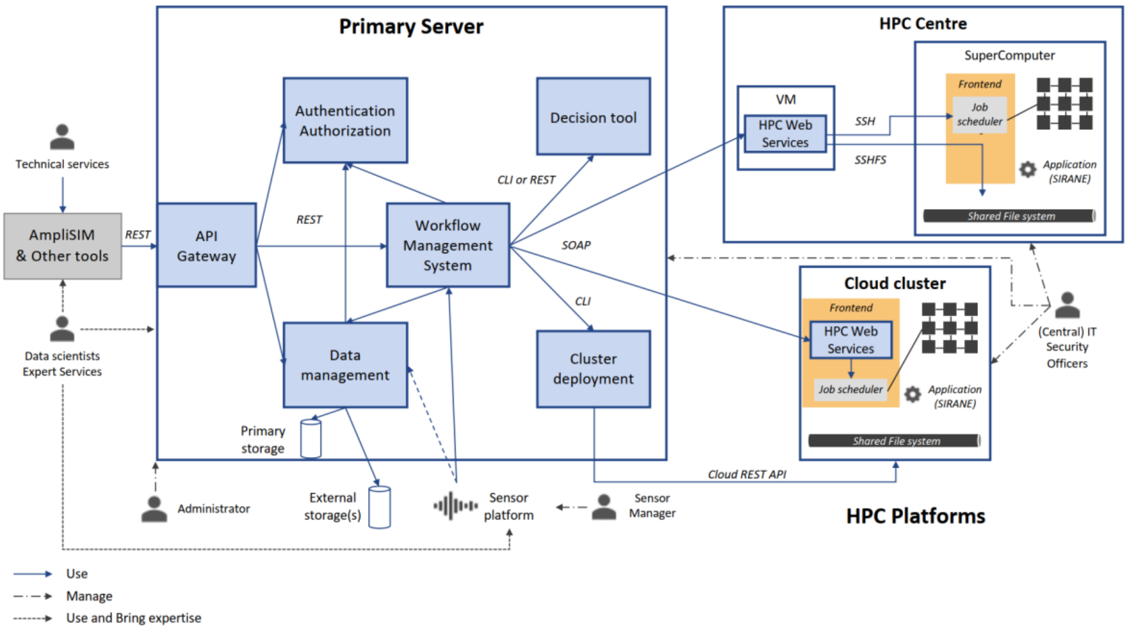

Figure 10: APIs of the workflow management system.

More details can be found in AQMO deliverable D4.1 Architecture solution for HPC as a service.

5. Simulation as a service

This component implements a set of functions to allow for simulations within a city when targeting local scale on-demand situations. Due to constraints regarding the urban setup to be modelled, the spatial resolution and the modelling speed, the SIRANE model was chosen and was integrated within the AmpliSIM Modelling Web Service (AMWS). This integration allowed for multiple scenario analysis. Specific developments were also performed to introduce an additional modelling capability within AMWS: this capability allows running simulations incorporating forecasts repeatedly in time. Hence, on a regular basis, for instance daily, simulation results are updated using the last measurements in the field and the new meteorological forecasts for the days to come.

To support this use of AWMS, both forecast simulations and urban modelling, additional data were incorporated within AWMS, namely building data and large-scale (meso-scale) meteorological forecasts. When available, open source data was chosen, for instance OpenStreetMap (OSM) data (see for instance Figure 11).

Figure 11: Concentration map obtained from SIRANE outputs and post processed through ATL. Map data for concentration can also be integrated in third party tools using AWMS API.

Transparent deployment of calculation from AWMS onto the supercomputer cluster from IDRIS has been performed using secured connexions between AmpliSIM Connection Backend (ACB) and the supercomputer. The deployment is performed by creating on-the-fly submission shells incorporating multi steps submissions for model and post-processing. Post-processing relies on multi level tiled data that allows for light and fast navigation in the simulation results. Results can be visualized directly with AMWS using its cartographic web interface (see Fig. 4) or indirectly through Application Program Interface (API) requests to AWMS.

This components has two main interfaces :

1) An API to run simulations. The API is based on a replication of a simulation that specifies all basic data. Then the parameters can be altered to run new simulations with different parameters.
2) AmpliSIM Connection Backend (ACB) connected to the resources used.

More details can be found in AQMO deliverable D3.1 Report on the Numerical Simulation Software and the AmpliSIM-API documentation.

6. Workflow manager

The Workflow Management System (WMS) is responsible for automating the orchestration of task collections upon computational resources. The WMS usually relies upon a Resource Management and Orchestration System (RMOS) which will handle the difficult job of actually executing the tasks on the resources. The specificities of AQMO platform such as the usage of a hybrid distributed computing infrastructure, the need for lightweight and low-power systems software stack for the edge layer computational resources (within buses and drones), the need for stream processing and that of secure network transfers widens the scope of traditional workflow management and orchestration systems.

In AQMO, the combined jobs of WMS and RMOS are performed by the proprietary solution Ryax. It is based on Kubernetes orchestrator and resource management. In particular the lightweight K3S distribution for optimal usage on the Edge computing nature of the project. In addition the software defined networking solution of ZerotierOne is used to cover the needs of security, authentication, flexibility and simplicity for the layer of network management. Figure 12 illustrates the internal architecture of the proposed solution.

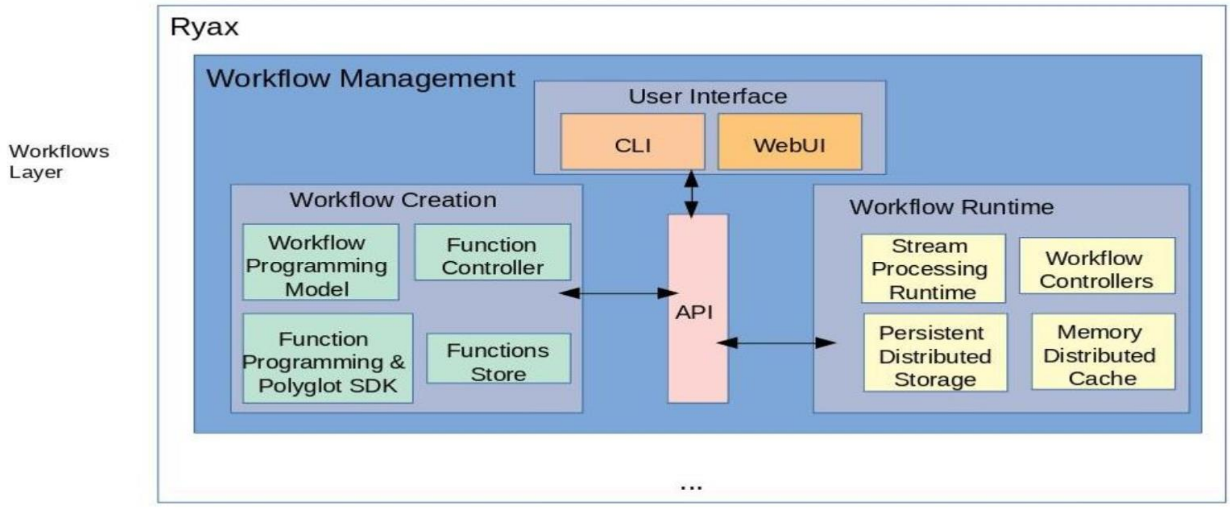

Figure 12: Ryax Workflow Management internal architecture.

In particular, the Workflow Management System is installed on the central units, of each bus or drone, where it orchestrates workflows related to treatment of data coming from sensors; along with the primary server, where it orchestrates pipelines related to data transferring, among others. Even if it does not need to be installed on the HPC or Cloud side of resources it automates the execution of simulations on them through a fine integration with the different decision and interface tools of the HPC/Cloud layer. More details can be found in AQMO deliverable D5.1 Workflow Management System.

7. Integration strategy

The logical structure of the platform follows usual practice. This is illustrated in Figure 13. The connection between the layers are ensured in two ways. A set of APIs have been defined and the relevant components are enrolled into a Software Defined Network (SDN) that offers a first level of security.

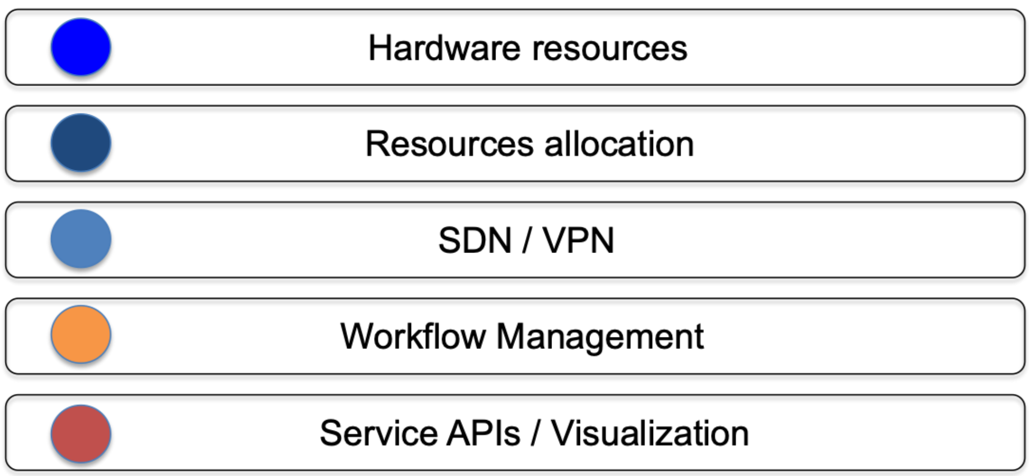

Figure 13: AQMO platform logical layers.

The integration strategy we have applied in AQMO follows two main criteria:

1) Flexible enough to account for the heterogeneity of the platform components;
2) Allow for an independent reuse of each component.

To ensure these constraints we have relied on an usual approach based on the definition of API and the deployment of standard communication protocols. Figure 14 shows where these API/protocols are set in the platform. Each API/interface corresponds to a well-defined functional perimeter. As a consequence, each component can be exploited independently of each other. Figure 14 also shows the mapping with the logical layers (given Figure 13). For instance, the mobile sensor platform is not in the perimeter of the resource allocation as it is manually installed and included in the platform contrary to other servers that can be allocated on demand.

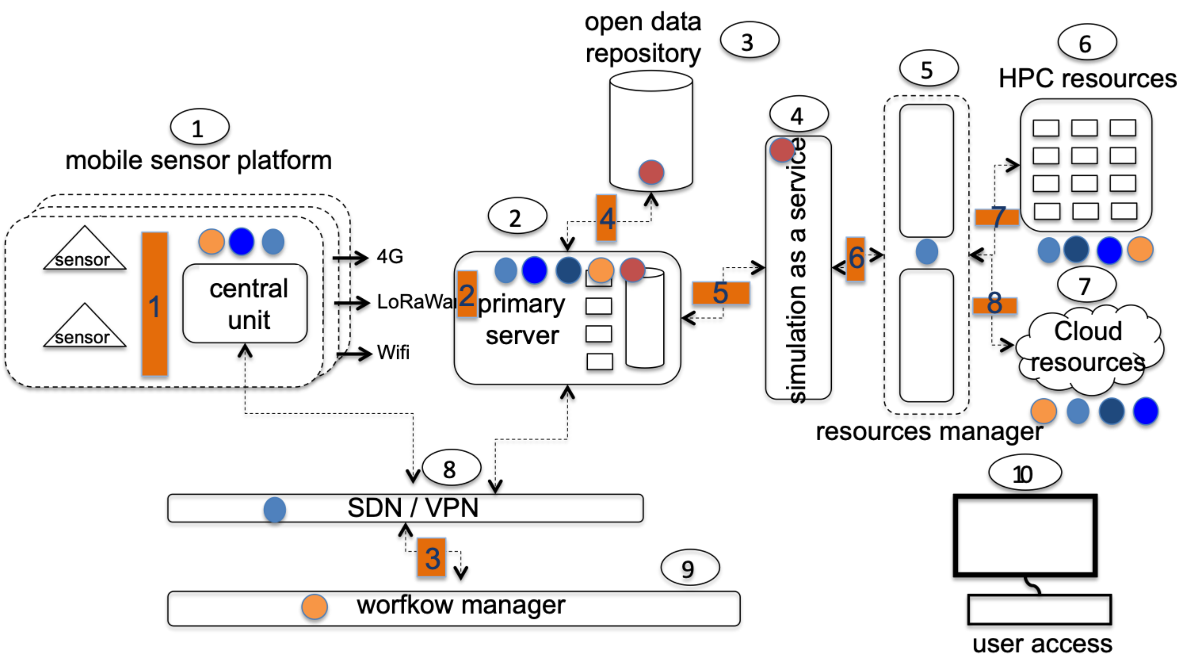

Figure 14: Platform API (in orange) and logical platform layer relevance.

The APIs are the following:

1) Sensors communicate to the central unit using the MQTT communication protocol over Wifi.
2) The primary server communicates between the devices.
3) The workflow manager acts on top of a SDN. It is based on the Kubernetes technology [11]. It orchestrates the data analytics workflows on both the central units and the primary server and it automates the executions on HPC/Cloud resources.
4) The integration of the data sets generated by the platform are published on the open-data portal using the RUDI producer API. This API is currently being specified [06].
5) Simulation as a service provides an API to drive the configuration and launching of a simulation. The configuration of the simulation allows to change the source of

pollutant in order to allow "what-if" scenarios to be explored. It is implemented over the *https* protocol.

6) This interface allows access to computing resources to execute the numerical simulation. Selection of target computing resources is done through the Decision Tool's API.
7) This interface provides access to the IDRIS supercomputing center computing resources. It is based on "robot account", with limited numbers of actions, only the ones mandatory for launching the simulation, according to security policy of the HPC centre.
8) This interface to access Cloud resources is provided by the CCME infrastructure.

Conclusion

The integration process has been one of the most challenging for the AQMO project. The heterogeneity of the components ranging from microcontrollers to supercomputers as well as multimodal communication have made the task particularly difficult. The number of technologies used is also numerous however every time this has been possible we have relied on well established industry standards. Security was also taken into acount at each step of the integragion.

The platform is one of the rare end-to-end platforms that performs data logistic from sensors to simulations. The platform is operational and there are currently 8 (with a target of 10) buses capturing air quality data running in Rennes. At this time, data have been collected for 24 months, starting with two sensors installed at the end of 2018, providing some elements of a digital twin of a smart city.

Current effort is focusing on the sustainability of the platform.

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