



SESAME

Capgemini
Pascal PORCAR
Olivier DUGAST
ArianeGroup
David LARROQUE
Mikael AFONSO

Object: Release note for infrastructure deployed in cloud sandbox

30 July 2020

Dear Madam, Sir

Inside SESAME project, Work Package 2 consists in providing a data platform allowing to:

- Ingest data
- Secure data
- Process data and run algorithms
- Visualize data
- Export / extract data
- Secure platform accesses
- Scale
- Collaborate

That data platform is planned to serve the 2 already identified use cases and more to come as the project progresses.

A 3-step approach has been defined to enable the target data platform construction:

1. A Proof of Concept

Based on Public Cloud (virtualization based on shared hardware), this first step includes main components and allows to demonstrate several security concepts / solutions and to incubate different concepts.

This first platform provide infrastructure deployment automation.

As ArianeGroup security requirements are no fully covered, this platform doesn't contain any ArianeGroup data

2. Prototype 1

Based on Private Cloud (virtualization based on dedicated hardware), this second step will include security configuration fitting cyber security requirements. This second platform will host ArianeGroup Data ingested via batch only.

3. Prototype 2

Evolution of Prototype 1 platform, this third step moves the data platform towards an



Project funded by European Union

This project has received funding from the European Union's Horizon 2020 research and innovation framework programme under grant agreement No 821875. [Click here for more info](#)

industrialized platform including Big Data components including the ability to host large volume of data, to manage various types of data and manage real time data ingestion.

First step has been achieved.

This first version addresses the target features the following way:

- Ingest data
Data can be ingested via file load onto file system using *Webdav* protocol and **NextCloud** for user, S3 protocol and **Minio** for application
Data can be stored via different types of databases (Influx DB for TimeSeries, MongoDB and PostgreSQL for Geolocalisation) + **Grafana** and **Chronograf** for datavisualization solutions.
- Secure data
Users can access to the platform using unique login and password
KeyCLOAK solution has been installed to provide Single Sign On (SSO) feature. Logs and traces can be accessed using **ElasticSearch** and **Kibana**
Users accesses will be filtered and limited to the Sesame Project team using network and firewall technologies.
- Process data modelize and run algorithms
Dataiku has been installed
Connection with Predict has been implemented as well
- Visualize data
Dataiku is also used as data visualization solution
- Export / extract data
- Other components for transversal features have been implemented:
Onedev : git repository to store and share algorithm source code
Redmine : project management to plan work and work the plan
JSPWiki : Wiki to share documentation
Wekan : Ticketing solution to share and prioritize tasks

To support these functional features, technical components have been installed on Public Cloud platform:

- OpenSuse
- VM Vsphere + NSX
- **Kubernetes** as orchestrator **and Docker** as containerization solution
- **Clam AV** as antivirus
- **Veeam** as backup solution
- **Apache** as reverse proxy

WP2 team



Project funded by European Union

This project has received funding from the European Union's Horizon 2020 research and innovation framework programme under grant agreement No 821875. [Click here for more info](#)

Sesame Project
www.sesame.com



Project funded by European Union

This project has received funding from the European Union's Horizon 2020 research and innovation framework programme under grant agreement No 821875. [Click here for more info](#)