PaNOSC Portal Architecture

16th June, 2020

Presenter: J. Hall on behalf of PaNOSC WP4

Authors: J. Hall (ILL), S. Caunt (ILL), W. Turner (ILL)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

Overview

- What is the PaNOSC Portal?
- Common analysis services goals
- Why are we developing a common portal?
- Portal workflow
- Challenges
- Architecture
- First iteration
- Facility deployment
- Status and Roadmap

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

What is the PaNOSC Portal?

- **The objective of this work package is to make data analysis services available through cloud hosted services and on the EOSC.**
  - Build on experience and user feedback from VISA and CalipsoPlus
  - Simplify access to scientific software and tools
  - Promote collaborative and reproducible scientific analysis
- **Development of a common remote analysis application deployed at each site.**
  - Provide Remote Desktops and Jupyter Notebooks
  - FAIR access to scientific data (public and embargoed)
  - Each facility implements a connector providing access to their compute and data infrastructure
- **Provide a common API to manage cloud resources on existing compute infrastructure.**
  - A user will be able to select any facility and start remotely analysing their data via a single interface
- **Tie into services provided by other work packages**
  - FAIR Data API (WP3)
  - AAI / data access and transfer (WP6)
  - Learning resources (WP8)

panosc

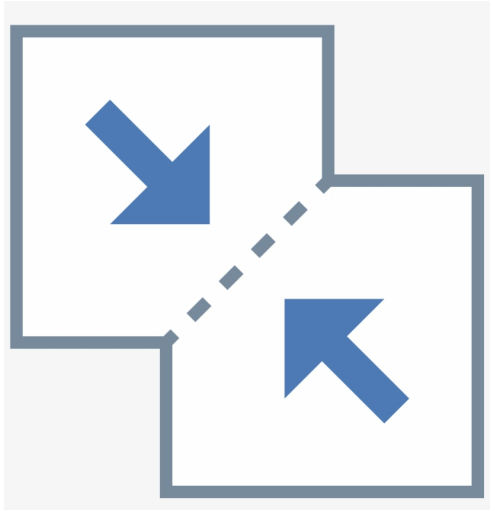
Common analysis services goals

- Provides seamless access to data and remote analysis tools
- Common user interface and experience
- Searching user data and open data across all facilities
- Create a cloud resource for a selected dataset
  - Tailored to the data analysis requirements
  - Configurable environments
- Collaborating on analysis and sharing the results
  - Common workflows for analysing well defined datasets
  - Sharing Remote Desktop sessions and Jupyter Notebook
  - Facilitating reproducibility
- Quotas to manage the distribution of cloud resources
- Common authentication using UmbrellaID

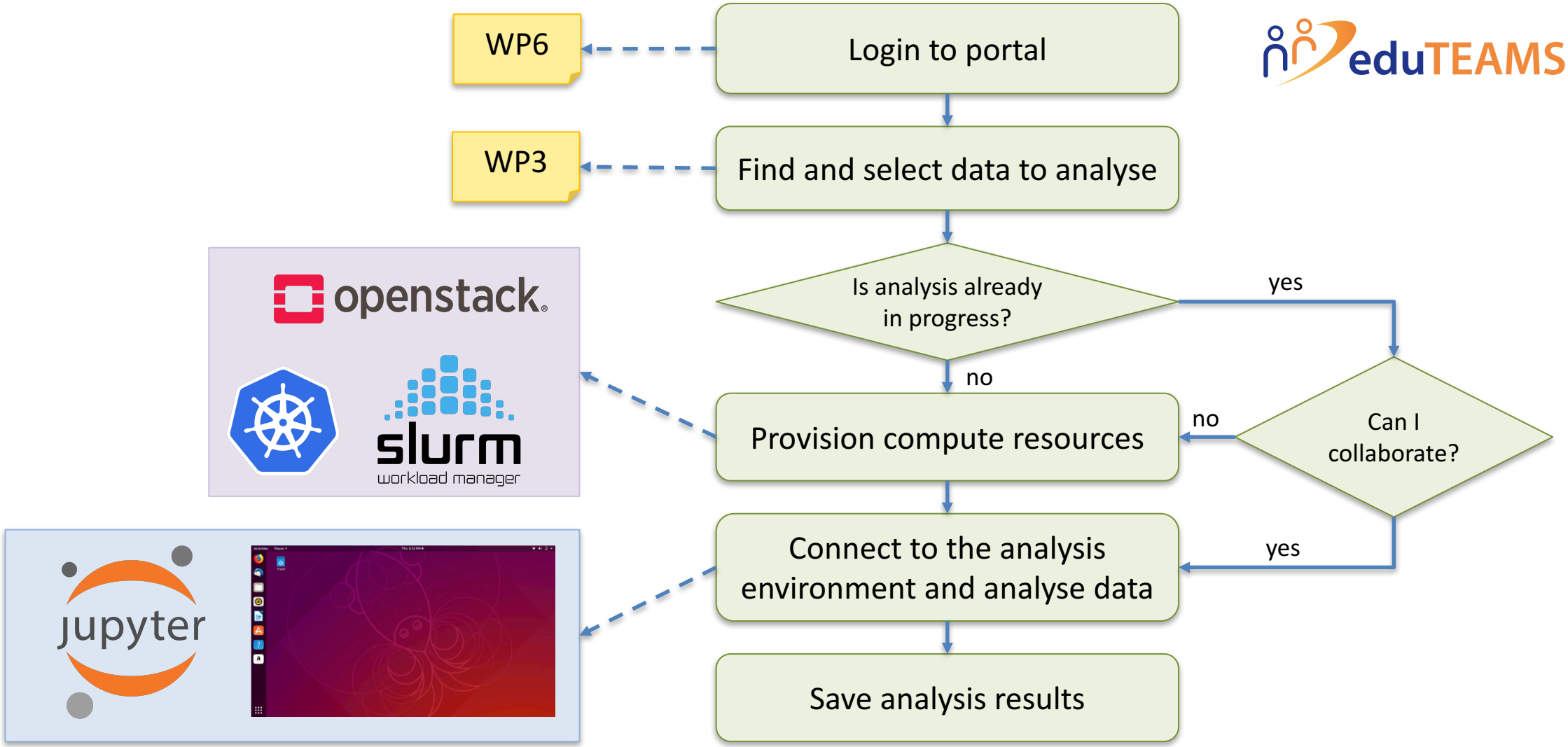
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

Why are we developing a *common* portal?

- Same user experience irrespective of the facility or where the data is stored
- Single point of entry for PaN users to access and analyse their data
- Merge existing solutions into a common user experience
  - User interface and design
  - Authentication
  - Search and access datasets
    - Open or embargoed
  - Re-use of common software and tools



Portal workflow



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

panosc

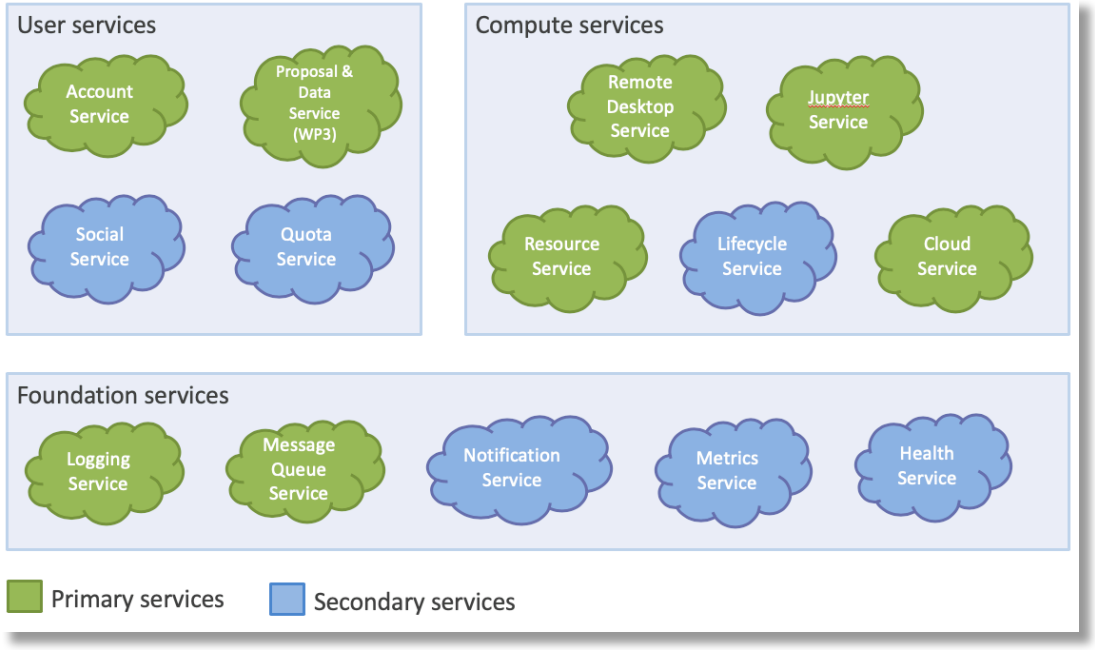
Challenges

- **Each facility has different infrastructure, services and needs**
  - Data analysis survey carried out to get a better understanding
  - Use cases and specification validated by all partners
- **Distributed development across facilities**
  - Different skill sets and availabilities
- **Different levels of maturity of existing solutions**
  - Data policy, DOIs, Remote Desktops, Jupyter Notebooks, NeXUS.
- **Deployment**
  - Different levels of experience

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

Architecture

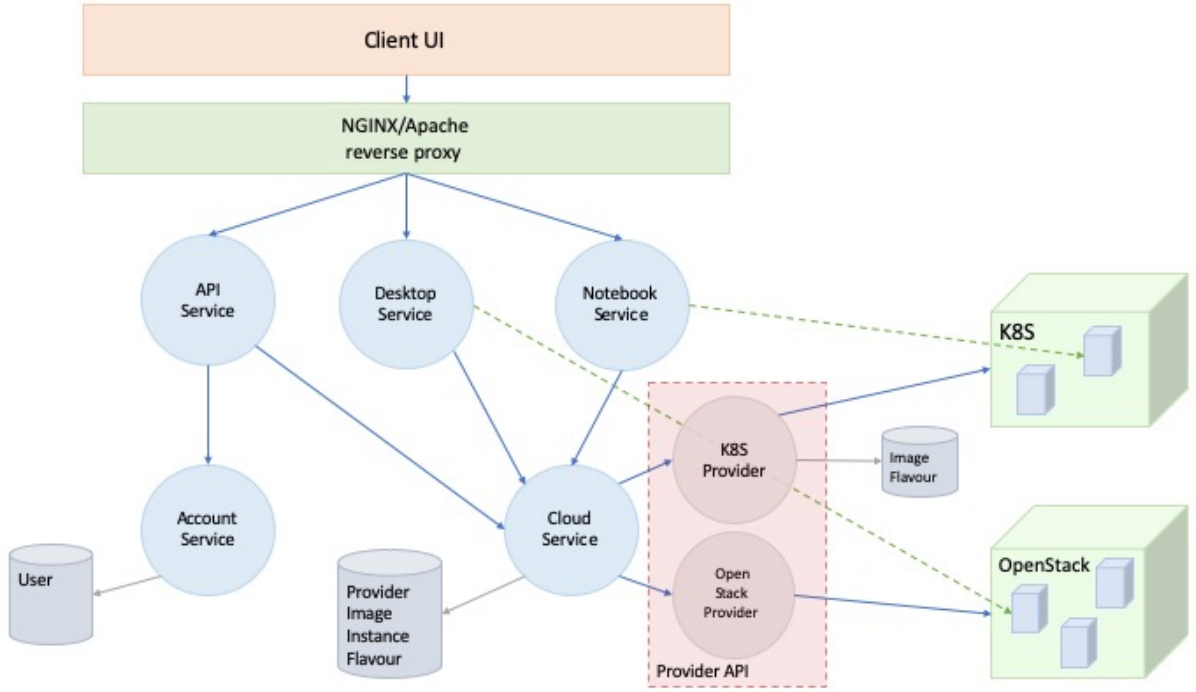
- **Built on a micro-service architecture**
  - Easy to have multiple developers developing separate functionalities (shared responsibilities)
  - Well-defined API for each module/service
    - Enable site-specific implementations
  - Less dependence on a single language/technology (allows flexibility)
- **Identified *core* and *secondary* services**
  - Analysis of the use cases and feature prioritisation
- **Iterative development**
  - Enables quick feedback
  - Reactive to changes
- **Documentation publically [available]**



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

First iteration

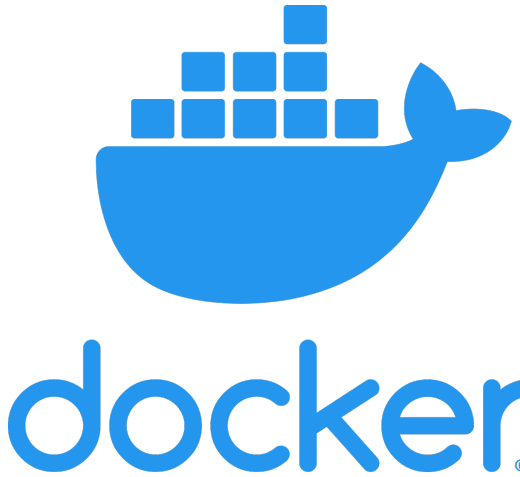
- **Management of compute resources**
  - Implementation of Kubernetes *cloud provider*.
- **Validation of architecture core services**
  - Deployed at partner sites
- **User authentication (OpenID connect)**
- **Obtain feedback**
  - Encountered problems / constraints / edge cases
- **Micro-services developing use NodeJS / Java**
- **All source code is Open Source and made available on GitHub.**

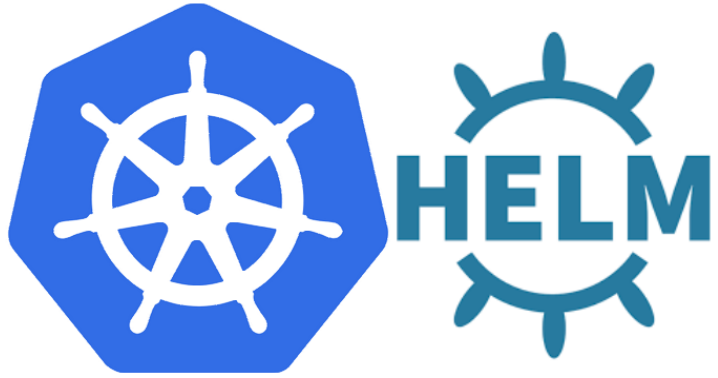


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

Facility deployment

- **Automated the deployment of the portal**
  - Use Kubernetes to manage the orchestration of the micro-services
- **Packaged as HELM charts**
  - Configurable for site specific needs
  - Easy to upgrade
  - First stable production release
  - Open sourced on GitHub
- **Demo version available**
  - Single command, no configuration
  - Dummy environment
  - Allows for simple testing of latest features





This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

Status and Roadmap

- **First development iteration completed**
- **Validated the architecture**
  - Feedback from partners
- **Next stage is to integrate the frontend and backend developments**
- **Looking at integrating**
  - FAIR Data API of WP3
  - Data Access from WP6
- **Development of other micro-services**
  - SLURM and OpenStack Cloud Providers (priority)
  - Messaging / Notification service
  - Quota service
  - Metrics service
- Demonstrator deployed at each site: November 2021
- Obtain user feedback after beta testing / validation: May 2022

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852

Thank you

wp4@panosc.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823852