EOSC ENVRI DevOps framework -- Progress report

Zhiming Zhao[1], Andrea Manzi[2]

[1]University of Amsterdam ENVRI-FAIR WP7 leader

[2]EGI foundation

ENVRI-FAIR has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824068

Time line

- Use case 1: workflow automation
- Use case 2: Jupyter notebook for data intensive science
- Use case 3: EOSC DevOps framework

Phase 1	Phase 2	Phase 3	Phase 4
Describe the key use case: scenario, scope, KPI, steps etc. **Get familiar with the EOSC services**, following training and practices from the other projects etc. Get the requested resource provisioned Setup OLA with resource providers and agreement with Jelastic.	**DevOps pipeline configured**, including Git, automated testing, integration and **deployment demonstrate** in at least via two service development	**Demonstrate the initial version of the workflow from ENVRI-FAIR**, with automated workflow execution in Cloud; **Demonstrate the other common data services** identified in the ENVRI communities (optional)	**Exploit** the results to the development activities in ENVRI-FAIR sub domain **Sustain** the development by finding new opportunities, e.g., new EOSC projects etc.

Resource provisioning

1. EGI OLAs agreed with the resource providers (CESGA and INFN-CATANIA) and cloud resources provisioned

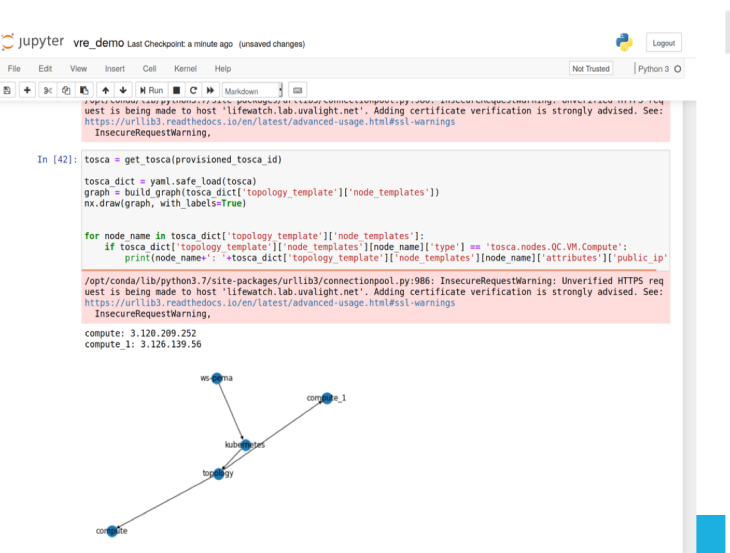
1. Installation of the Jelastic platform performed by Jelastic engineers on CESGA VMs
   a. Admin interface: https://jca.j.fedcloud.eu/
   b. User interface: https://app.j.fedcloud.eu/

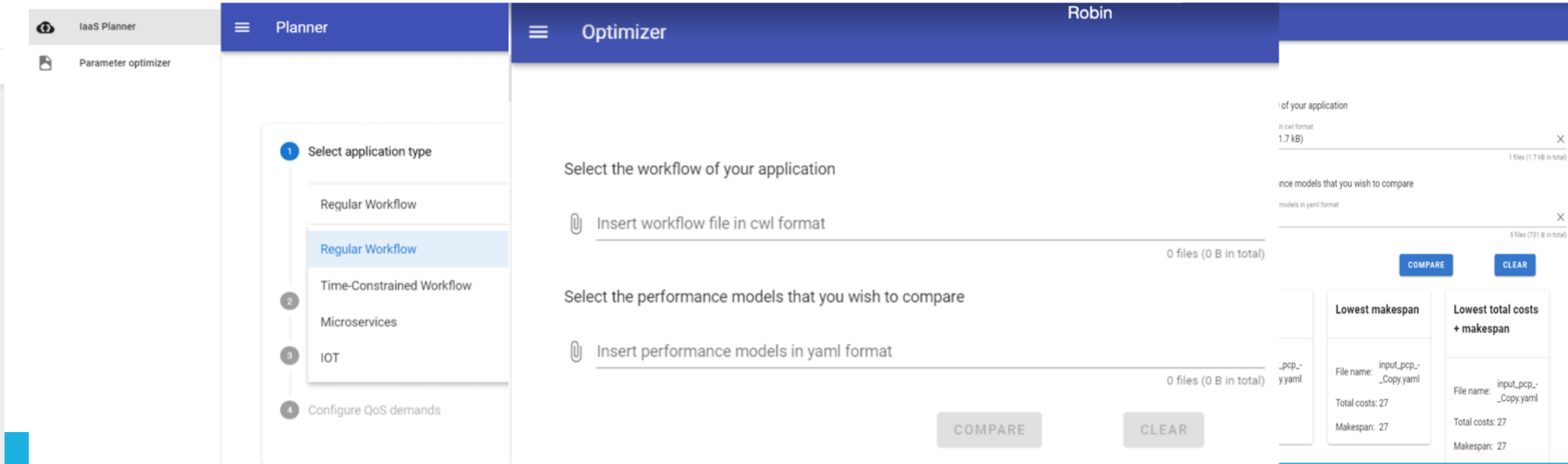
1. Onedata Oneprovider for EGI DataHub deployed at CESGA with 10TB storage (ENVRI-FAIR space)
   a. https://oneprovider-envri.datahub.egi.eu

1. EGI community Notebooks instance deployed at INFN-CATANIA K8s cluster (access via EGI Checkin)
   a. https://envri-notebooks.fedcloud-tf.fedcloud.eu

Activities during the last phase

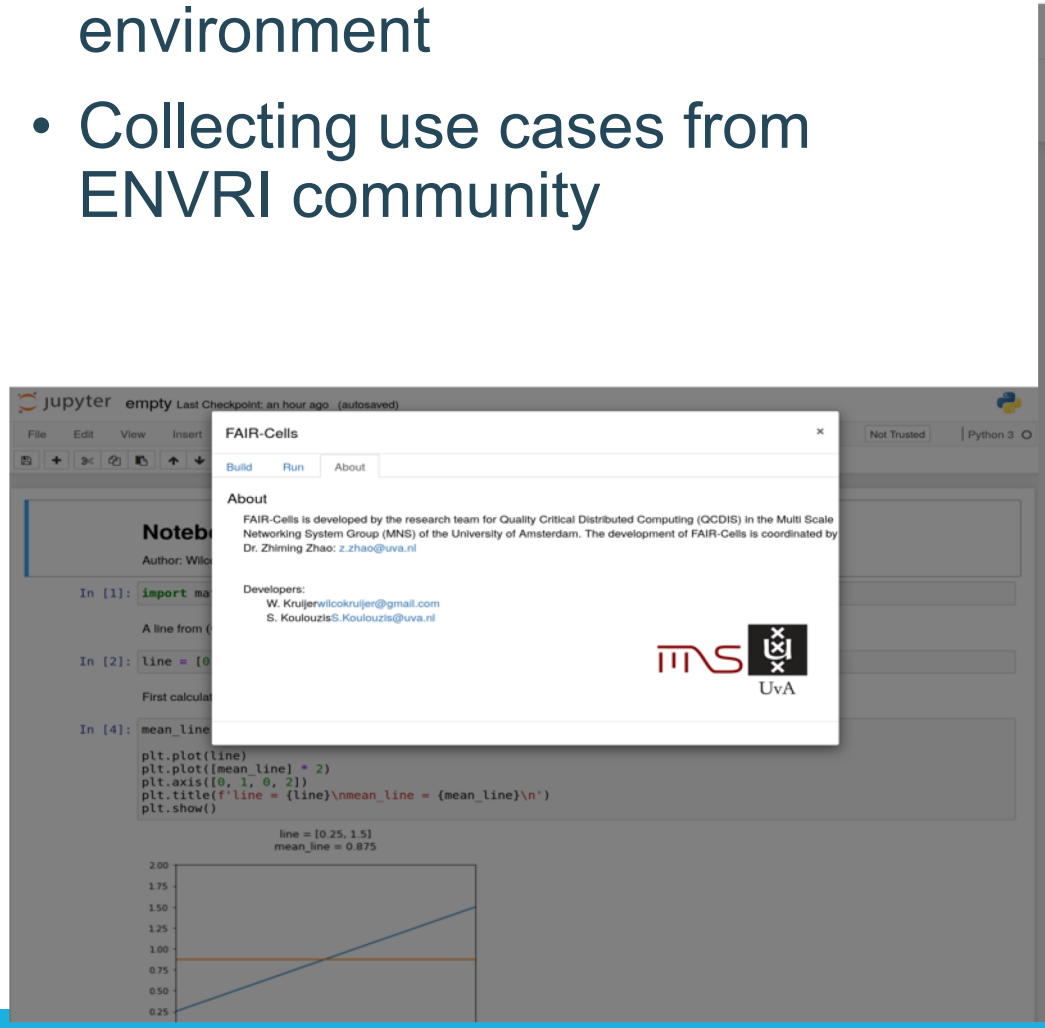
- Get familiar with Jelastic
- Integrate the tools developed in the previous phase
  - IaaS automation engine with OpenStack interface, and authentication interface
  - Work on the use case of ecology Lidar data processing

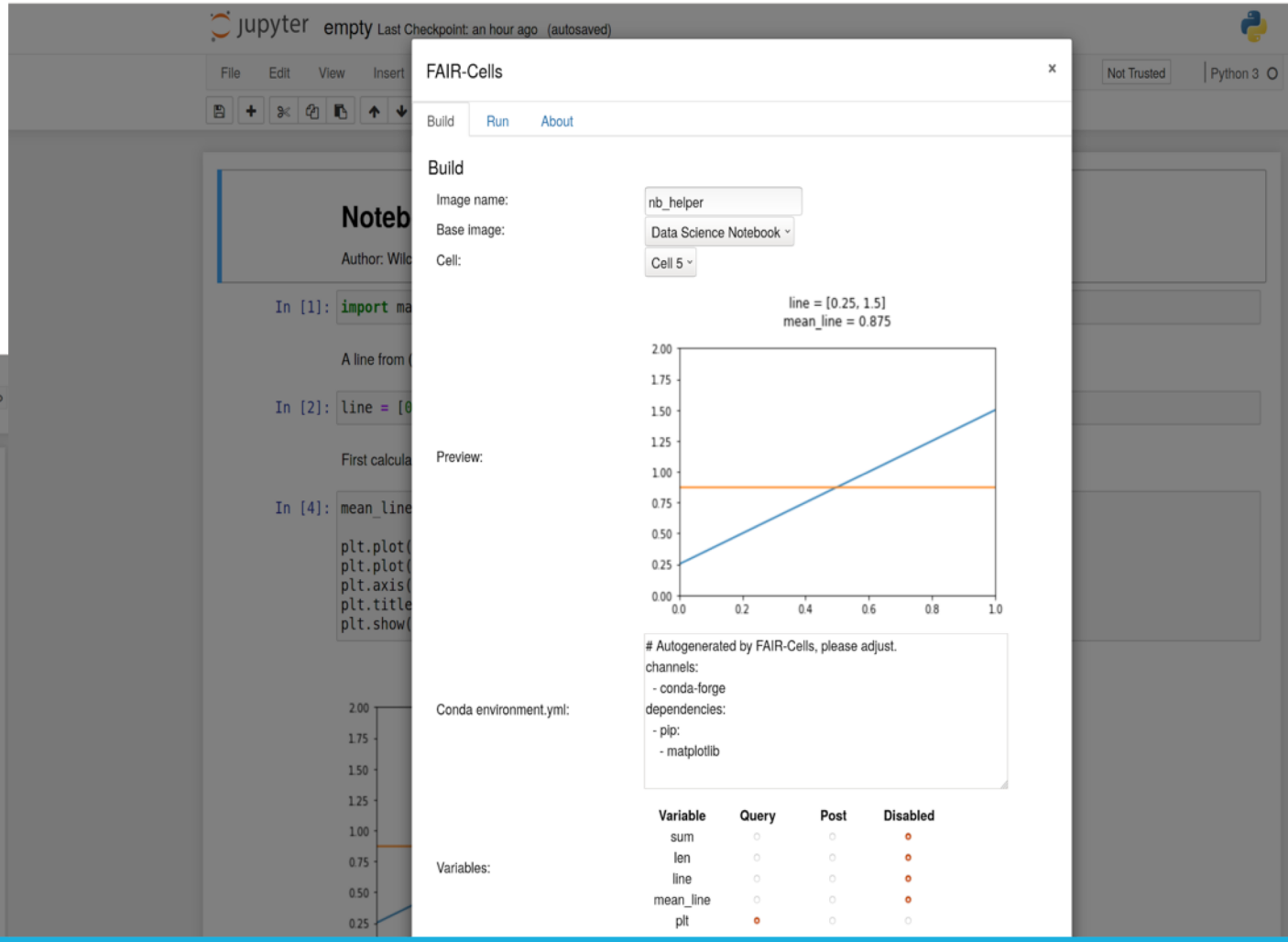




Running Jupyter in distributed cloud

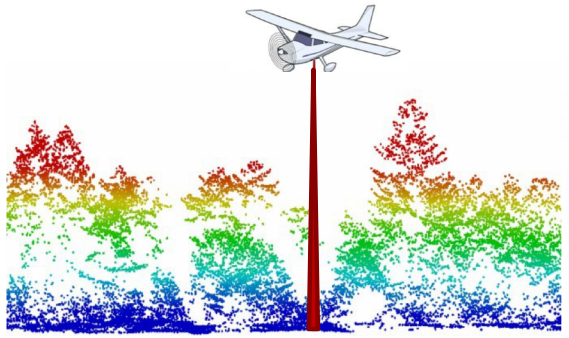
- FAIR-Cells: Customize Jupyter environment
- Collecting use cases from ENVRI community

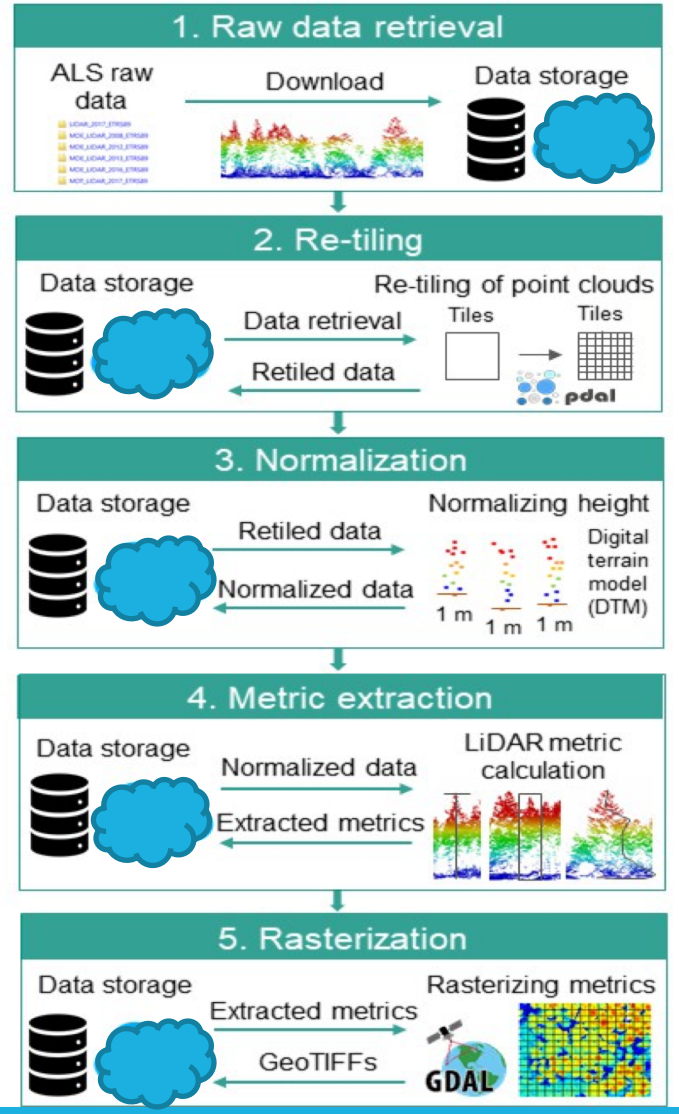




Processing Lidar data on Cloud

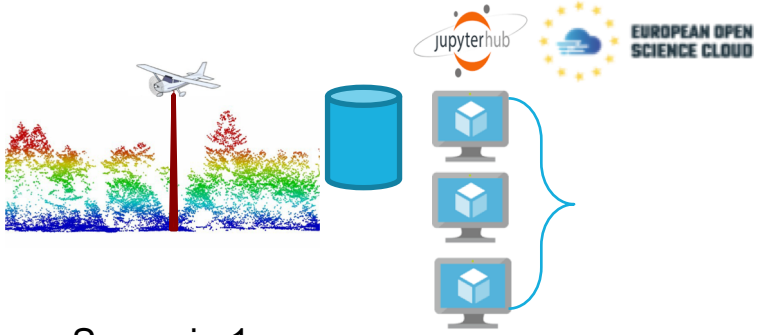
- Running the current Jupyter notebook with bigger data sources on larger infrastructure
- Laserchichek is a program to process the Lidar data via 5 steps
- The program will be executed on
  - Jupyter Hub, and
  - On distributed EOSC VMs using the FAIR –CELLs tool developed



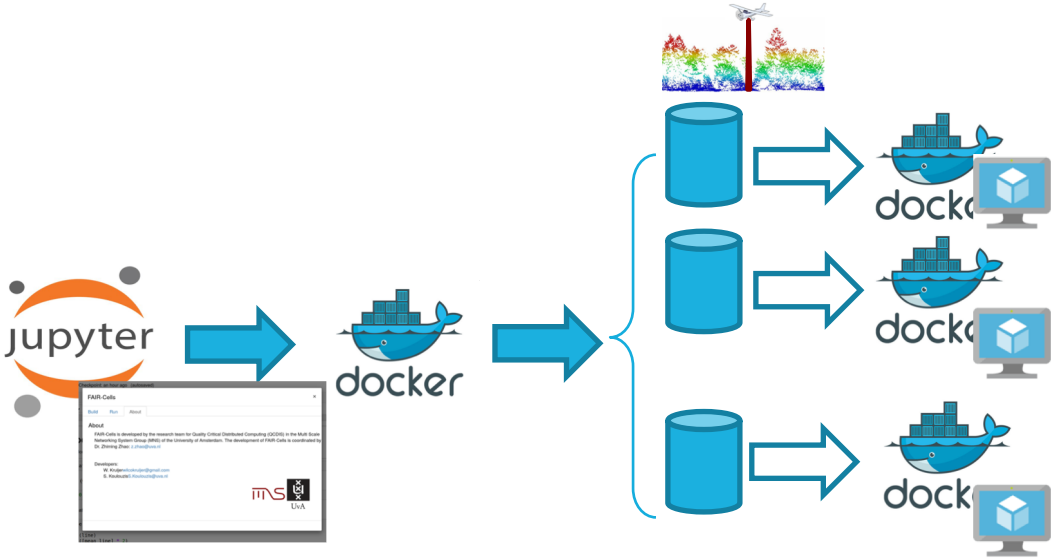


Demonstrator

- Scenario 1: running the laserchilcheck on the JupyterHub
- Scenario 2: containeralizing the code using FAIR-CELLs, and then distributed containers on different VMs
- Submitted an abstract to the EGI conference.



Scenario 1



Scenario 2

EUROPEAN OPEN SCIENCE CLOUD

ENVRI FAIR

DevOps pipeline comparison

Action:

1. Establish DevOps pipeline for a software we developed in ENVRI
2. Compare Jalestic and other DevOps options (Azure and Open source choices)
3. Create a practice report

Summary

Phase 1
Describe the key use case: scenario, scope, KPI, steps etc.
Get familiar with the EOSC services, following training and practices from the other projects etc.
Get the requested resource provisioned
Setup OLA with resource providers and agreement with Jelastic.

Phase 2
DevOps pipeline configured, including Git, automated testing, integration and **deployment**
demonstrate in at least via two service development

Phase 3
Demonstrate the initial version of the workflow from ENVRI-FAIR, with automated workflow execution in Cloud;
Demonstrate the other common data services identified in the ENVRI communities (optional)

Phase 4

Exploit the results to the development activities in ENVRI-FAIR sub domain
Sustain the development by finding new opportunities, e.g., new EOSC projects etc.

Next phase

- Demonstrator: scaling LidarData notebook to large scale data and infrastructure (by Nov.)
  - From notebook to REST services/containers
  - Describing distributed workflows
  - Automated the execution using EOSC IaaS
- Demonstrator 2: DevOps frameworks demonstrator (by Dec).
  - Demonstrator 3.1: demonstrate the automation of testing, integration, deployment
- Exploitation to ENVRI-FAIR in the last three months.

Sustainability

1. The use case is jointly with the LifeWatch ERIC VL innovation center; the output will be taken by LifeWatch ERIC to further develop
2. Exploitation of the results to the ENVRI communities (via the common development plan of RIs in ENVRI).
3. Getting support from other ongoing relevant projects, e.g. ARTICONF, BlueCloud and CLARIFY
4. Other opportunities

EOSC Earlier Adopter Program

www.envri.eu

envri.eu/envri-fair

@ENVRIcomm

ENVRI community

facebook.com/ENVRIcomm