Demonstration: Big Data Analytics for agricultural monitoring using Copernicus Sentinels and EU open data sets

🔗 eosc-hub.eu

🐦 @EOSC_eu

Dissemination level: Public
Disclosing Party: Project consortium
Recipient Party: European Commission

Scientific case: BDA for AgMon

- The EU's Copernicus Sentinels produce 10 PB data per year.
- Have become essential Earth Observation sensors for monitoring land, oceans and atmosphere.
- Full, free and open license. Focus on 10m resolution Sentinel-1 & -2.
- Urgent need to ensure uptake beyond the traditional "Space" sector (ECA special report 07/2021)
- Big Data Analytics required to support applied science "at scale"
- EOSC Early Adaptor Project to demonstrate advantages of "cloudification"
- Agriculture context: "signature" extraction for agricultural parcels
- Demonstrate mutual benefits in agronomics and public use domains

m Sentinel-2 @ DIAS
p Sentinel-1 @ DIAS
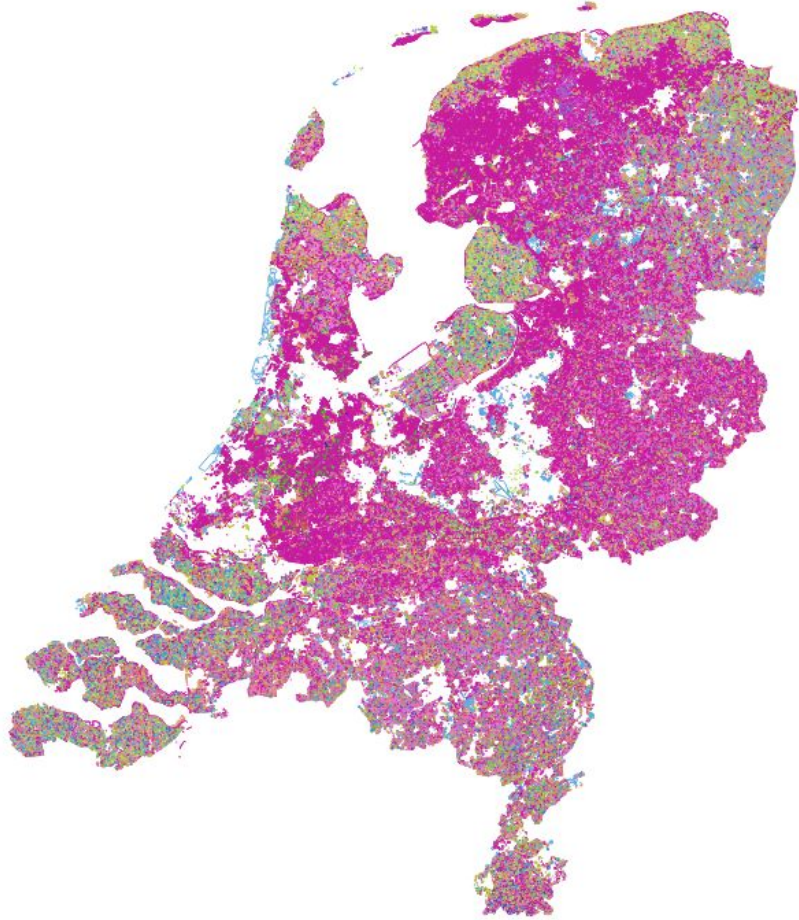
n BRP 2019 @ pdok.nl

n ~ 770,000 parcels /yr
m ~ 4000 S2 granules/yr
p ~ 1200 S1 scenes/yr
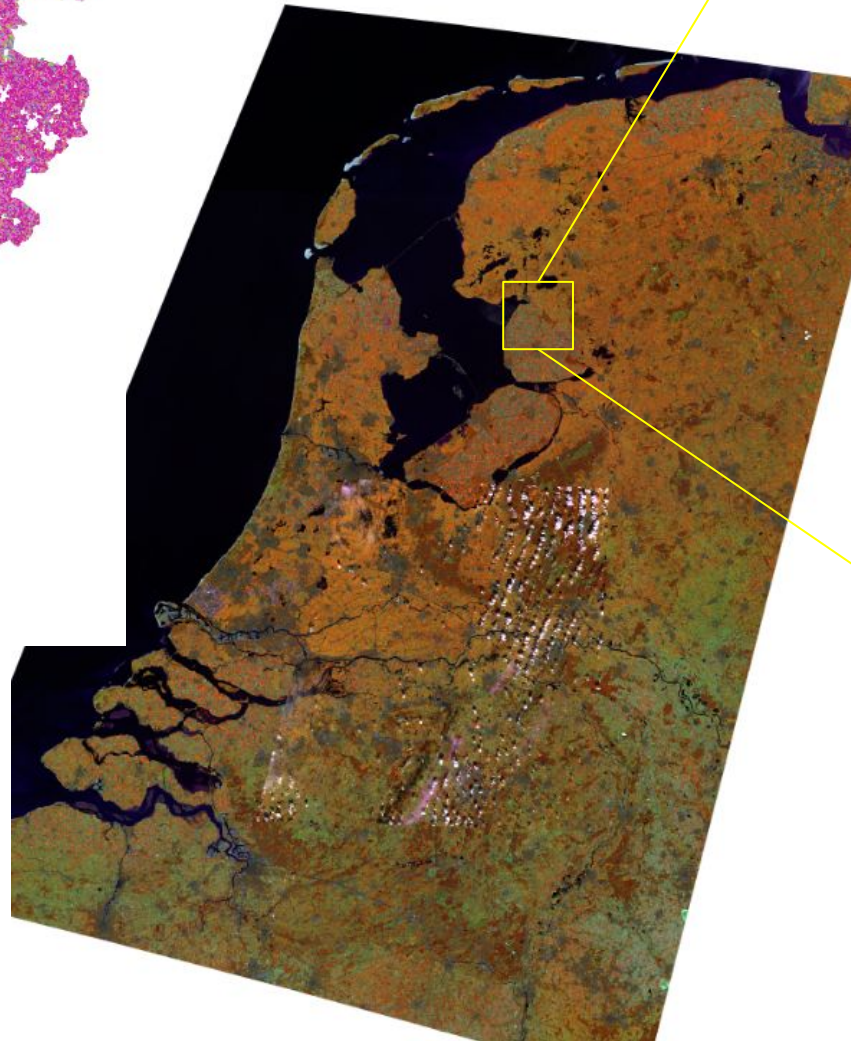
n-m, n-p*2 spatial time series
for Sentinel-1, -2 CARD
for b bands (b=14 (S2), 2 (S1))
x 100 for whole EU

Scientific platform

- Required functionalities based on CAP "Checks by Monitoring".
- Testing with Member States on Copernicus DIAS platforms.
- Demonstrate expanded use cases in a federated set up.
- To facilitate exposure to science data use, cross-domain collaboration
- Stimulate scaling of scientific data use to national, European
- Integrate latest developments in parallel software and hardware
- Code released as open source: github.com/ec-jrc/cbm

Integrations with EOSC-hub

Service	Provider	Description of the integrations	Allocated ICT resources (cloud, storage, etc.)
CloudFerro Infrastructure	CloudFerro*	Integrates access to multiple Petabyte S3 store of Copernicus Sentinel data with cloud compute	24 vCPUs, 32 GB RAM, 100 GB HDD
EGI Cloud Compute	EGI	Server backend setup on performant cloud resources with large data storage (CEPH based HDD + S3)	CESNET: 16 vCPUs, 64 GB RAM, 100 GB HDD 32 vCPUs, 128 GB RAM, 200 GB HDD Database server, 4 vCPUs, 16 GB RAM, 1 TB HDD 20 TB S3 storage
EODC Data Catalogue Service	EODC*	Querying and processing of Copernicus Sentinel-1 data and compute solutions for ARD	8 vCPUs, 16 GM RAM, 100 GB HDD

*CloudFerro and EODC are part of the EOSCHub EO pillar
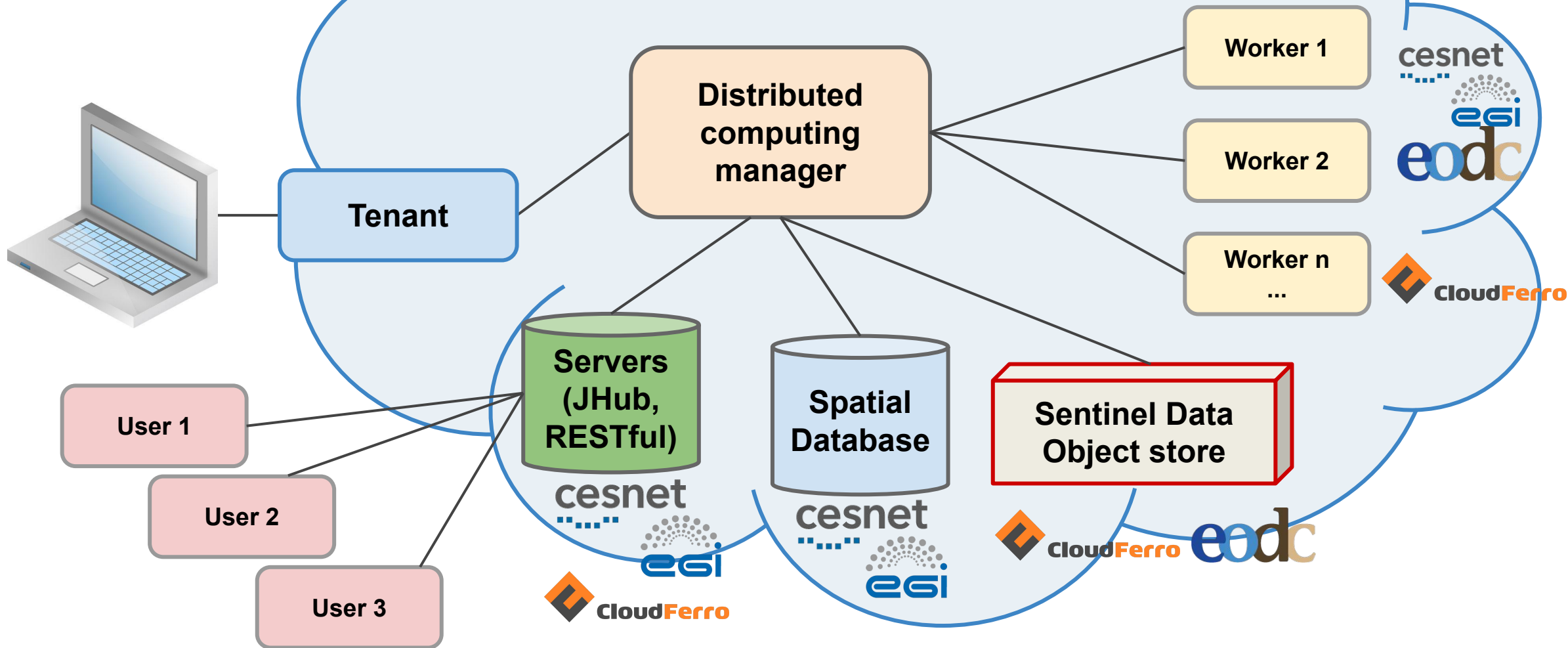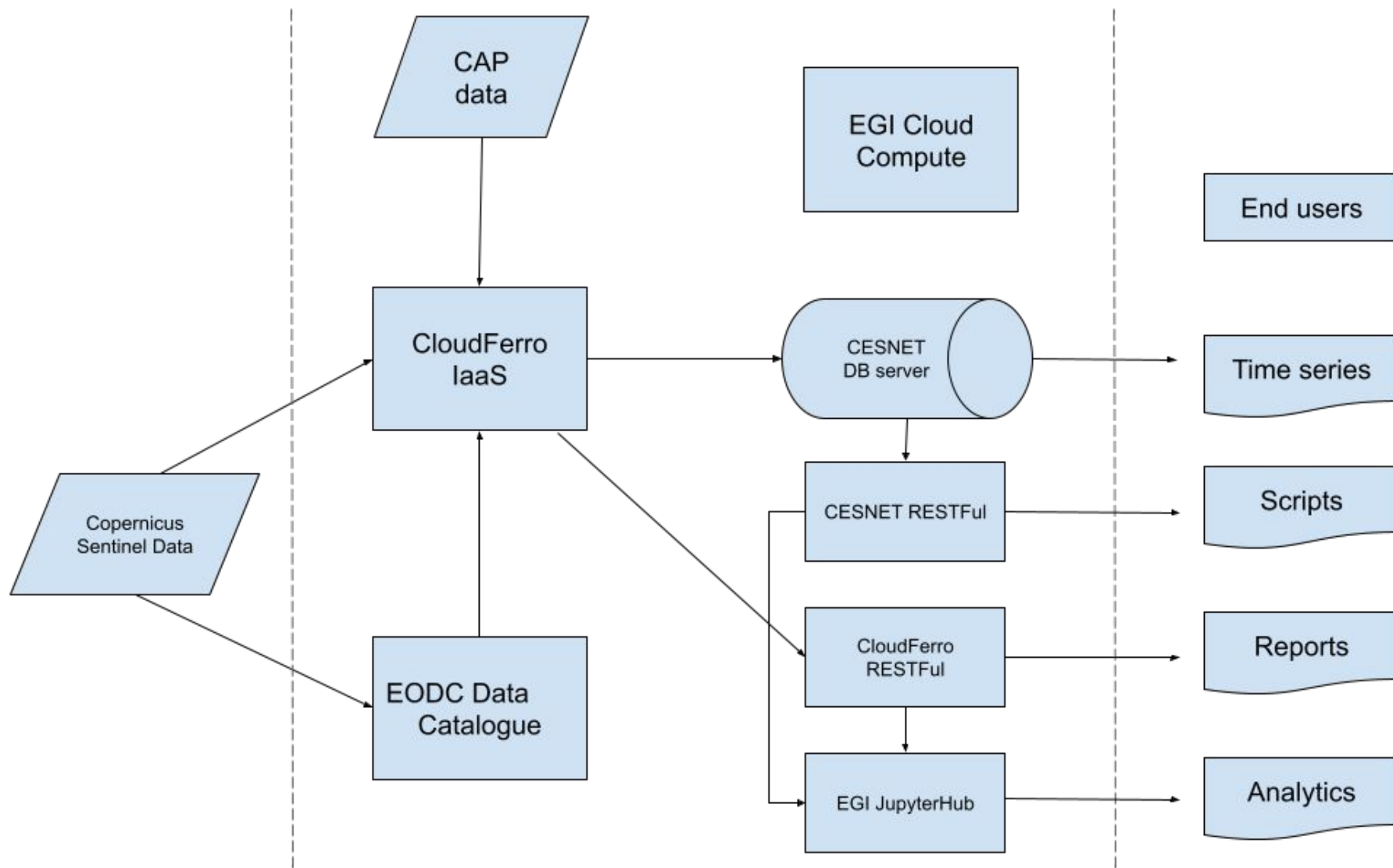
- A federated European compute infrastructure provides a realistic alternative to the US "hyperscalers" [in the EO domain].

- Federation facilitates "pluggable" modular evolution, fully based on open source, with advanced HW solutions (e.g. GPU, HPC) and SW (hyperscaling, ML). Educational efforts needed.

- Huge untapped potential in science - public - private use. Requires a a long term perspective. EOSC + euroHPC + European Data Spaces

Introduce the demonstration

- We demonstrate a "front-end" workflow on JupyterHub.
- Extracts time series, image chips from the "back-end".
- Support interactive analysis of agricultural practice at parcel level
- But also basis for non-interactive script-based analysis
- Including higher order processing (e.g. segmentation, ML)
- Some components work globally, others require pre-processing

**Thank you
for your attention!**

_____

Questions?

EOSC-hub

🔗 eosc-hub.eu   🐦 @EOSC_eu