

Expanding the capacity and capabilities of an Earth Observation application by means of the European Open Science Cloud

19/09/2022

Amanda Calatrava (UPV), Thiago Emmanuel Pereira (UFCG), Francisco Brasileiro (UFCG), Ignacio Blanquer (UPV)

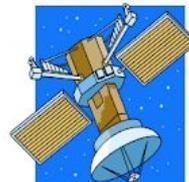
In this lightning talk we will present you...

1. What is SAPS?
2. Technical Description: Architecture
3. SAPS in the EOSC ecosystem



1. What is SAPS?

- **SAPS** (*SEB Automated Processing Service*) is a service to estimate Evapotranspiration (ET) and other environmental data that can be applied on water management and the analysis of the evolution of forest masses and crops.
- **Objective:** provide wider access to knowledge on the impact of human and environmental actions on vegetations, leading better forest management and analysis of risks.
- **User profile:** Researchers in Agriculture Engineering and Environment.
 - Few users that produce results that may be relevant for many users



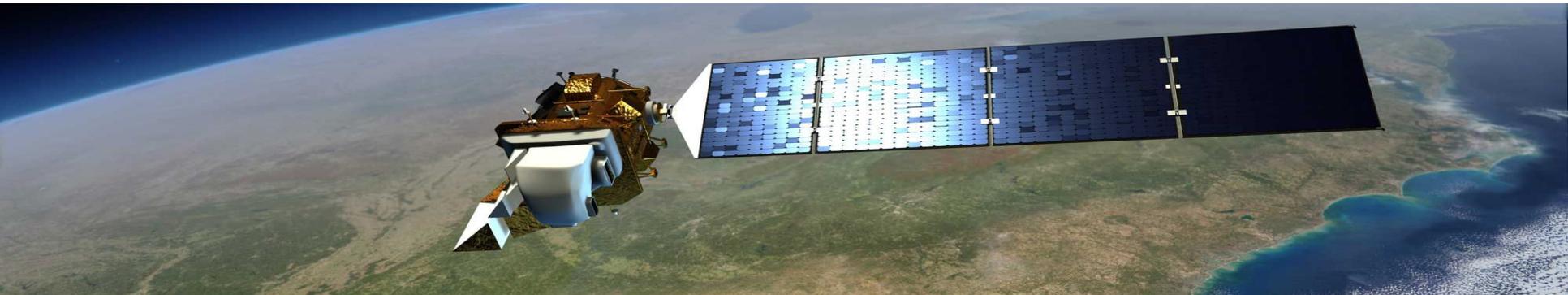
2. Technical Description: Architecture

Workload Characterization

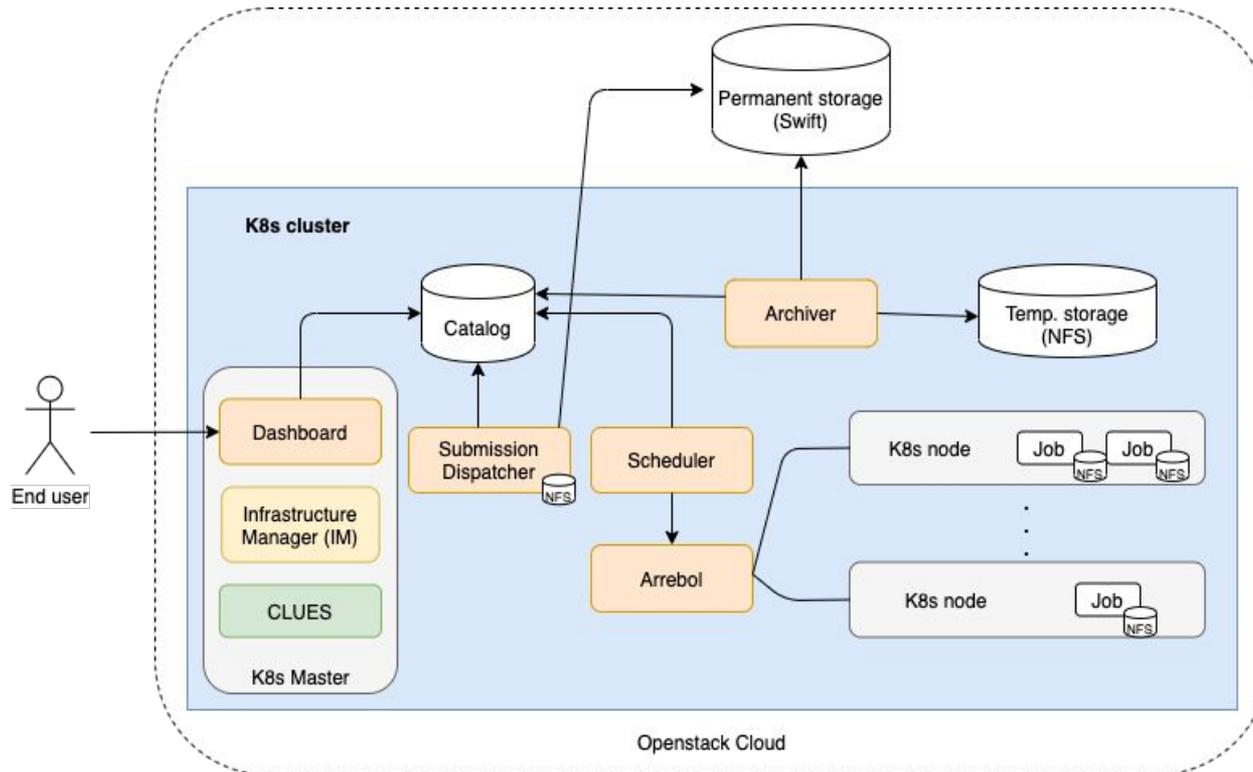
- SAPS uses **containers** to facilitate the deployment of customizable versions of evapotranspiration processing algorithms.
- The application type is mainly **Sequential/High-Throughput**, limited by CPU and RAM.
- Users submit batches of jobs typically for a single scene and a period of 40 years, which leads to a batch of about 800 jobs.
- Each job computes evapotranspiration and other vegetation indexes of an entire Landsat scene (170km x 185km, 30-meter resolution) for a particular date.
- A job represents a **three-step sequential workflow**:
 1. Input download,
 2. Input preprocessing
 3. Evapotranspiration estimation.

Data Sources

- It requires integrating access to the [LANDSAT](#) data repository.
 - 9 very high spatial-resolution NetCDF files of 12 GB each.
 - Historical information regarding various climate and vegetation indicators for a time range spanning from 1984 to 1995.
- Meteorological information and elevation data are downloaded from mirror servers of [NCEI](#) and [CS](#) services managed by UFCG.
- Each job of the application requires to download around 0.5GB of data.



Architecture



- **7 main components.**
- **Kubernetes** as the LRMS.
- Dashboard needs to be accessible from outside.
- Temporary and permanent **storage** solutions are needed.

Gaps and Bottlenecks

- The service is limited in terms of computing, storage resources and data access.
- It requires:
 - Larger-scale deployment
 - Integrated AAI
 - Standardisation of services
 - Improved scalability
 - FAIR management of produced data



3. SAPS in the EOSS ecosystem

Projects involved

- SAPS is one of the 10 thematic services of the **EOSC Synergy** project, in the field of Earth Observation
 - <https://www.eosc-synergy.eu/supporting-science/saps/>
- SAPS is supported by an **EGI-ACE** grant from the 4th Call:
 - A total of 44 vCPU, 168 GB of RAM and 108 TB of object storage, divided into a 4vCPU cores, 8GB RAM front-end node and 10 worker nodes with 4vCPU and 16GB of RAM.
 - The resources are provided by IN2P3-IRES to the members of the VO saps-vo.i3m.upv.es, accessible at https://aai.egi.eu/registry/co_petitions/start/coef:163



EOSC Services

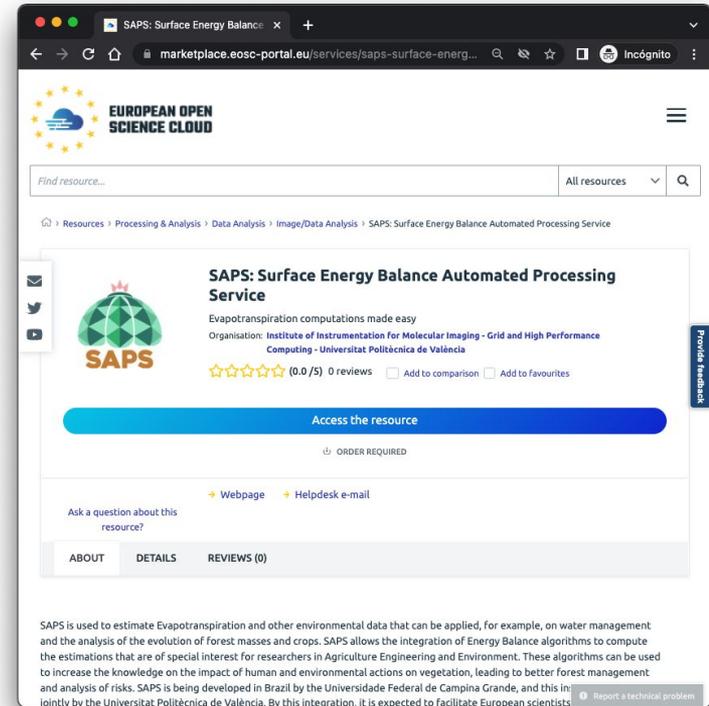
- **EC3 and IM tools:** both are services used by SAPS to deploy and configure a Kubernetes cluster automatically with SAPS running on it. Also, EC3 is used to manage the elasticity of the K8s cluster automatically. These tools facilitate the deployment and management of SAPS service.
- **EOSC computing resources:** through EC3 and IM, the SAPS service is deployed on top of a virtual elastic K8s cluster, that may rely on EOSC federated cloud computing resources or in on-premises solutions like Openstack.
- **EGI Check-in:** the SAPS dashboard has been integrated with EGI Check-in. So, to access a SAPS cluster, you should identify yourself with EGI Check-in service.

BREAKING NEWS

Evaluating EGI DataHub in combination with B2Handle and B2FIND.

SAPS in the EOSC Marketplace

- Marketplace endpoint:
<https://marketplace.eosc-portal.eu/services/saps-surface-energy-balance-automated-processing-service>
- Service endpoint:
<http://saps.vm.fedcloud.eosc-synergy.eu/>
- You can also deploy your own SAPS instance with EC3 using the recipes in here:
https://github.com/ufcg-isd/saps-docker/tree/master/ec3_recipe



The screenshot shows a web browser window displaying the EOSC Marketplace page for the 'SAPS: Surface Energy Balance Automated Processing Service'. The page features the European Open Science Cloud logo, a search bar, and a breadcrumb trail: Resources > Processing & Analysis > Data Analysis > Image/Data Analysis > SAPS: Surface Energy Balance Automated Processing Service. The service card includes a logo for SAPS, a description: 'Evapotranspiration computations made easy', the organization: 'Institute of Instrumentation for Molecular Imaging - Grid and High Performance Computing - Universitat Politècnica de València', a 0.0/5 star rating, and an 'Access the resource' button. Below the button, it indicates 'ORDER REQUIRED' and provides links for 'Webpage' and 'Helpdesk e-mail'. At the bottom, there are tabs for 'ABOUT', 'DETAILS', and 'REVIEWS (0)'. A footer note states: 'SAPS is used to estimate Evapotranspiration and other environmental data that can be applied, for example, on water management and the analysis of the evolution of forest masses and crops. SAPS allows the integration of Energy Balance algorithms to compute the estimations that are of special interest for researchers in Agriculture Engineering and Environment. These algorithms can be used to increase the knowledge on the impact of human and environmental actions on vegetation, leading to better forest management and analysis of risks. SAPS is being developed in Brazil by the Universidade Federal de Campina Grande, and this is jointly by the Universitat Politècnica de València. By this integration, it is expected to facilitate European scientists.'

Questions & Contact

Amanda Calatrava Arroyo

amcaar@i3m.upv.es

Instituto de Instrumentación para la
Imagen Molecular (I3M)

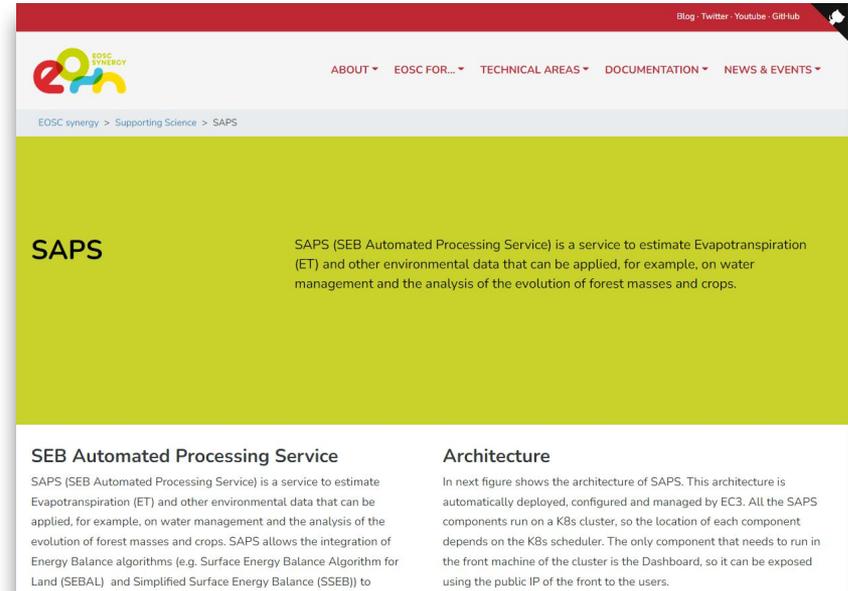
Universitat Politècnica de València



www.eosc-synergy.eu



@EOSC_synergy



The screenshot shows a web page for SAPS (SEB Automated Processing Service) on the EOSC Synergy website. The page has a red header with navigation links: ABOUT, EOSC FOR..., TECHNICAL AREAS, DOCUMENTATION, and NEWS & EVENTS. Below the header, the breadcrumb trail reads "EOSC synergy > Supporting Science > SAPS". The main content area has a green background and features the title "SAPS" followed by a description: "SAPS (SEB Automated Processing Service) is a service to estimate Evapotranspiration (ET) and other environmental data that can be applied, for example, on water management and the analysis of the evolution of forest masses and crops." Below this, there are two columns of text. The left column is titled "SEB Automated Processing Service" and describes the service's function: "SAPS (SEB Automated Processing Service) is a service to estimate Evapotranspiration (ET) and other environmental data that can be applied, for example, on water management and the analysis of the evolution of forest masses and crops. SAPS allows the integration of Energy Balance algorithms (e.g. Surface Energy Balance Algorithm for Land (SEBAL) and Simplified Surface Energy Balance (SSEB)) to". The right column is titled "Architecture" and states: "In next figure shows the architecture of SAPS. This architecture is automatically deployed, configured and managed by EC3. All the SAPS components run on a K8s cluster, so the location of each component depends on the K8s scheduler. The only component that needs to run in the front machine of the cluster is the Dashboard, so it can be exposed using the public IP of the front to the users."

Want to see more?

Watch a demo in:

<https://www.youtube.com/watch?v=gXfNlmsRSrc&index=10>

Try the training course with EC3:

<https://moodle.learn.eosc-synergy.eu/course/view.php?id=14§ion=6>

Documentation:

<https://github.com/ufcg-lsd/saps-engine/tree/docs/sphinx>

