

VESPA-Cloud

Principal Investigators: Baptiste Cecconi baptiste.cecconi@obspm.fr
Shepherd: [Baptiste Grenier](#)

Entry in the community requirement database: [VESPA](#)

About the pilot

Description of supported work

VESPA (Virtual European Solar and Planetary Access) is a mature project, with 50 VESPA providers distributing open access datasets throughout the world (EU, Japan, USA). In October 2019, the current number of data products available within the VESPA network reaches 18.3 millions (among which 5 millions products from the ESA/PSA, Planetary Science Archive).

The VESPA team is supported by the Europlanet-RI-2024 project (started on Feb 1st 2020 for 48 months, H2020 grant agreement No 871149).

Each VESPA provider (institutes, scientific teams...) is hosting and maintaining a server (physical or virtualized) with the same software distribution (DaCHS, Data Centre Helper Suite), which implements the interoperability layers (from IVOA, International Virtual Observatory Alliance, and VESPA) and following FAIR principles. Each server hosts a table of standardized metadata with URLs to data files or data services. Data files can be hosted by the VESPA provider team, or in an external archive (e.g., ESA/PSA - Planetary Science Archive).

The VESPA architecture relies on the assumption that data provider's servers are up and running continuously. The VESPA network is distributed but not redundant. For small teams with little or no IT support is available locally, the services are down regularly. We thus need a more stable and manageable platform for hosting those services. The EOSC-hub "cloud container compute" service would solve this problem.

We propose to use the EOSC infrastructure to host VESPA provider's servers (through a controlled deployment environment with git-managed containers). The open-source DaCHS framework is developed for Debian distribution. A docker containerization will be used to facilitate the framework deployment on other Linux environments.

Objectives

The VESPA providers would be able to:

- order a VM with the VO framework installed,
- configure the server for their science application,
- manage the server packages with the VESPA team,
- update the content and the metadata.

The VM has a fixed public DNS and public web http interfaces (with astronomy interoperability protocol access points). The VM will be registered in the Astronomy Virtual Observatory Registry, and thus will be reachable with any IVOA tools. The services can then be used by the final users within their science workflows.

General

This EAP is also quite innovative by involving GEANT providing eduTEAMS as the VESPA community AAI together with EGI Check-in and EUDAT B2ACCESS as e-infrastructures AAI.

Team

Participant	Role	Name and Surname
Observatoire de Paris	PI	Baptiste Cecconi baptiste.cecconi@obspm.fr
EGI Foundation	Shepherd	Baptiste Grenier
	Technical support	
CESNET	Resources provider	Mirek Ruda
IN2P3	Resources provider	Jerome Pansanel
GÉANT	Resources provider	Christos Kanellopoulos
	Technical support	
EUDAT - DKRZ	Resources provider (B2FIND Service)	Claudia Martens
	Technical support	
EUDAT - MPCDF	Resources provider (B2SAFE service)	Beatriz Sanchez Bribian
	Technical support	Johannes Reetz
EUDAT - Juelich	Resources provider (B2ACCESS AAI proxy for B2SAFE)	Sander Apweiler
	Technical support	

GRNET	Resources provider	Nicolas Liampotis
	Technical support	

Technical Plan

The full technical plan can be found here:

Work planned for Q1	<ol style="list-style-type: none"> 1. Having access to the VM at the sites 2. Validating access to storage from VMs <ol style="list-style-type: none"> a. iRODS b. Object Storage 3. Being able to manually deploy the full stack on a VM <ol style="list-style-type: none"> a. deployment of containers from git-managed repository
Work planned for Q2	<ol style="list-style-type: none"> 1. Automate VM deployment and management (cloud-init) + puppet or equivalent 2. Allow the VESPA Hub teams to deploy services by themselves NB: There are three VESPA Hubs: <ol style="list-style-type: none"> a. Observatoire de Paris (Paris, France), b. INAF/OATS (Trieste, Italy), c. Heidelberg Univ. (Heidelberg, Germany) 3. Test harvesting of metadata by B2FIND (DaCHS exposing OAI-PMH endpoint) 4. Configure community on the eduTEAMS Community AAI Service.
Work planned for Q3	<ol style="list-style-type: none"> 1. Document service deployment for data providers external to VESPA community 2. Having a VM template in appdb to have VM available at all sites 3. Having access to group management in eduTEAMS Community AAI: implement OAuth and group authorization in DaCHS
Work planned for Q4	<ol style="list-style-type: none"> 1. Document the process required to deploy the service following an EOSC order, providing their SSH keys 2. Onboard the service to have it recorded and orderable in the EOSC marketplace 3. Study technical follow up <ol style="list-style-type: none"> a. Doing computing on demand on batch resources using UWS b. Study how to deploy an Elasticsearch solution in EOSC c. Explore integration with Zenodo d. Explore usage of INDIGO PaaS 4. Explore sustainability options after EOSC-hub <ol style="list-style-type: none"> a. Discuss economic models allowing to provide the services b. Prepare agreements to continue to operate the services deployed during the EAP

EOSC services and providers

Providers

- EGI: SLA and OLA: <https://documents.egi.eu/public/ShowDocument?docid=3598>
 - CESNET
 - IN2P3
- EUDAT
 - MPCDF
 - DKRZ
- GÉANT

Services

- EGI cloud compute (VMs)
- EGI Dynamic DNS update for domain update
- EGI AppDB for VM template management
- EGI Check-in for access to EGI Services
- EUDAT B2ACCESS to access EUDAT services
- EGI Object Storage
- EOSC Monitoring
- EOSC Marketplace to publish service
- EUDAT B2SAFE for storage
- EUDAT B2FIND for discovery
- INDIGO PaaS for automated deployment (evaluation)
- Zenodo for DOI/PID (evaluation)
- eduTEAMS Community AAI Service for community membership management (users, groups, roles)

