

Présentation projet stratégique OSUG workflow

Y. Rogez¹, V. Bourdon¹³, R. Cailletaud³, S. Douté¹, F.
Malbet³, F. Roch³⁴, F. Thollard², V. Visan³

3 Octobre 2025

¹ IPAG, ² ISTerre, ³ OSUG, ⁴ Gricad

- Projet Stratégique OSUG (mixte workflow / GPU)
- Pour la partie workflow : un ingénieur à temps plein (Vlad Visan) + un stagiaire
- Piloté par Y. Rogez, géré à l'OSUG

Objectifs principaux (pour nous)

- gdm-sar avec plus de contrôle (avoir la main sur les étapes)
- pouvoir suivre le workflow par le système et sur le cluster
- faire tourner les jobs en tant qu'utilisateur
- indexer les métadonnées des jobs
- être compatible multiples clusters/ordonanceurs
- être open source
- avoir une version gratuite fonctionnelle²
- pouvoir lancer des workflows planifiés, récurrents
- pouvoir déployer des workflows dynamiques

²Un peu refroidi par des augmentations de prix récentes de certains logiciels

M1 à MPD.

Objectifs secondaires

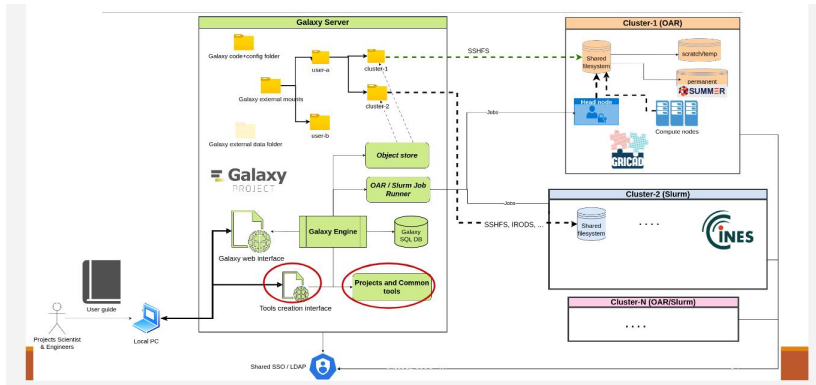
- être compatible GAI-DATA-Terra
- pouvoir faire de la visu et du check “à la volée”
- pouvoir partager les workflows
- être cyber-sécuré
- avoir un coût d'administration système raisonnable (maintenance, MàJ)

Choix de la solution technique

- Etude de différentes solutions (18).
- Short liste de 8,
- choix entre galaxy et airflow \Rightarrow galaxy

<https://ipag-wiki.osug.fr/spaces/CALCOSUG/pages/222419953/Diagrammes+-+Veille+comparative+finale>

Architecture du projet



Avancement du projet (1/2)

Avancement : Etat fonctionnel : runs qui tournent sur le cluster, premiers workflows implémentés (cf démo).

Retards :

- nécessité de changer notre architecture (imposé par Gricad, cyber)
- migration dahu-oar3 / kraken
- gestion de fichier "locaux" pas uniformément géré dans la code de galaxy
- ...

Reste à faire :

- finaliser la gestion des collections
- finaliser la gestion de parties admin (déploiement, machine de prod/test, ...)
- améliorer la gestion des workflows dynamiques (e.g. ressources calculées dynamiquement, saut d'étape, construction d'une étape à la volée)
- améliorer la gestion des petits jobs
- finaliser le couplage avec AMI (indexation des méta-données).
- permettre l'accès aux logs pendant les jobs
- contribution Galaxy
- finaliser docs (utilisateurs / admin)

DEMOS :

live : <https://galaxy.osug.fr>
pré enregistrée : <https://formater.osug.fr/>
(merci @Vlad)