



Lynda Feddak <feddaklynda00@gmail.com>

Rendez-vous du projet "Développement d'outils d'analyse IA et utilisation de la plateforme Galaxy-Ecology Europe pour le partage et l'exploitation de données d'imagerie sous-marine"

Yvan Le Bras <yvan.le-bras@mnhn.fr>

10 novembre 2023 à 11:51

À : Lynda Feddak <feddaklynda00@gmail.com>

Cc : nadine le bris <nadine.le_bris@obs-banyuls.fr>, islem kobbi <islem.kobbi@etu.sorbonne-universite.fr>, shivamshan sivanesan <shivamshan.sivanesan@etu.sorbonne-universite.fr>, sofiane ouali <sofiane.ouali@etu.sorbonne-universite.fr>

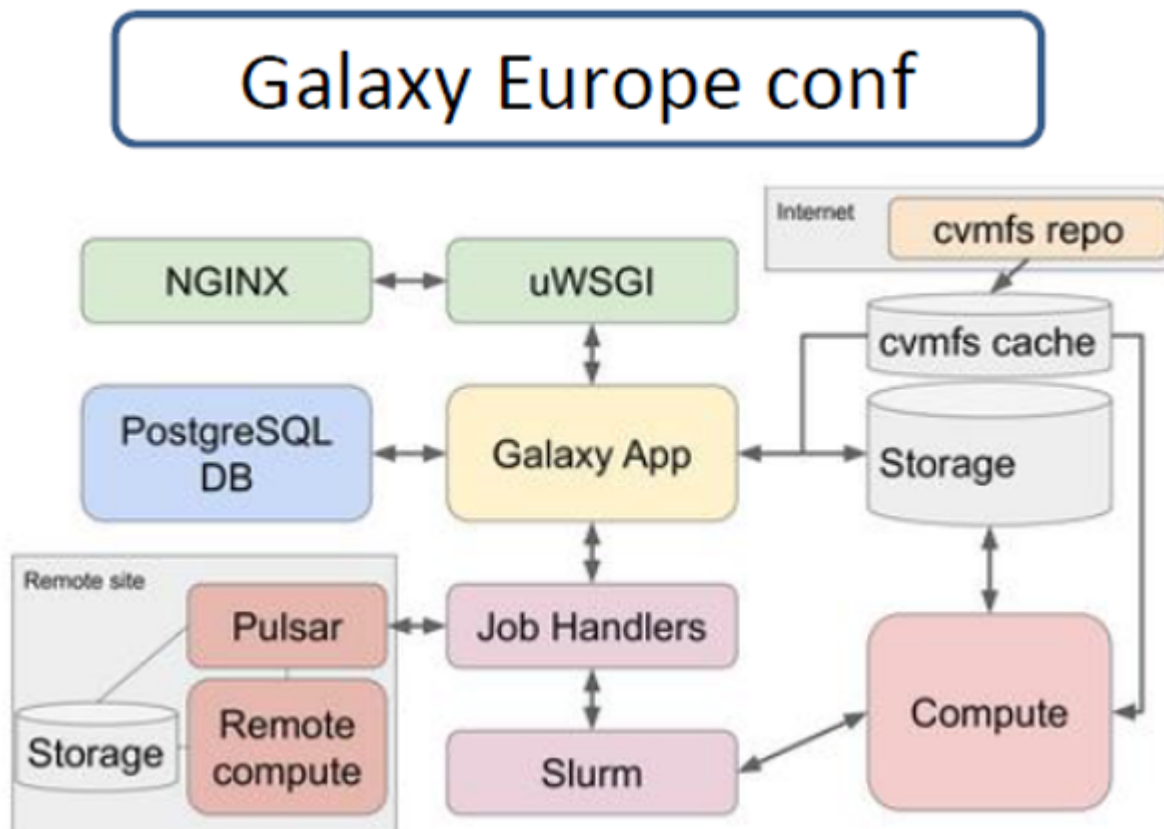
Bonjour Lynda, bonjour à tous,

Désolé du temps de réponse.

Pour le côté authentification des utilisateurs, sur la plateforme européenne Galaxy-E, <https://ecology.usegalaxy.eu/>, il y a une gestion des utilisateurs, je vous invite à tester la création de compte et l'utilisation de la plateforme pour mieux vous rendre compte de ces aspects. On peut ensuite créer des groupes d'utilisateurs, là j'ai moins d'expérience mais je peux me renseigner. Cela permet de pouvoir gérer des droits d'accès à des données notamment. Après, en tant qu'utilisateur Galaxy, on peut de base choisir de partager ses données ou ses historiques analytiques en mode restreint à quelques utilisateurs ou large à tous les utilisateurs de l'instance Galaxy.

=> Ne pas hésiter à me décrire plus finement ce qu'il faudrait gérer en terme d'autorisation d'accès.

Pour l'architecture logicielle, de manière basique vous pouvez vous référer au schéma fait par les collègues de artbio concernant Galaxy https://artbio.github.io/startbio/Run-Galaxy/Galaxy_architecture/ et un schéma dédié de l'architecture Galaxy Europe qu'on utilise ci-dessous



Ce qu'il faut voir de notre côté, c'est que un outil Galaxy = un fichier xml qui permet de décrire l'outil, ses dépendances (noms des recettes conda nécessaires à l'exécution du script), la ligne de commande à lancer (incluant le script devant être lancé, peu importe le langage), les entrées (données, paramètres) et sorties attendues, les tests unitaires et une section "help" décrivant ce que fait l'outil. Derrière ce fichier xml, lorsque l'outil va être lancé par un utilisateur, il va être exécuté dans un container dédié. Il y a 2 types d'outils, les outils "classiques" qui lance un script (R, Python, Bash, ...) et les outils "interactifs" qui vont avoir leur propre interface graphique (Jupyter notebook, RStudio, Application R SHiny, Application java, ...). dans le cadre du projet actuel, il se dégage l'intérêt de pouvoir intégrer VIAME dans Galaxy. VIAME étant un outil interactif, il faudra pouvoir créer le fichier xml qui va bien, et avoir une image docker pour le déployer. Ici, il y a un premier travail à faire, car, si j'ai bien vu, VIAME est mis à dispos via un docker-compose <https://kitware.github.io/dive/Deployment-Docker-Compose/> et il faudra proposer une version "docker unique" pour Galaxy. Ensuite, il faudra définir comment VIAME doit interagir avec Galaxy, s'il faut partager entre Galaxy et VIAME des ressources particulières de stockage et/ou calcul notamment. Enfin, sur le côté utilisation, cela veut dire que chaque utilisateur de VIAME va lancer une instance de VIAME (un container), faire ce qu'il a à faire, puis quitter VIAME ce qui arrêtera et supprimera le container. Le côté "persistance" entre exécution d'instance de VIAME devra donc passer par des données qui seront localisées dans l'espace personnel utilisateur Galaxy (nommé historique) et/ou dans un espace commun à plusieurs utilisateurs (nommé "bibliothèque de données", plus compliqué potentiellement mais à bien évaluer selon les besoins). Ces "données" assurant la persistance entre exécutions de VIAME (par le même utilisateur ou différents utilisateurs) pourront contenir si pertinent des infos sur les paramètres VIAME.

=> matériels de formation pour développer u noutil Galaxay : <https://training.galaxyproject.org/training-material/topics/dev/> (section tool development surtout)

=> matériels de formation pour tout ce qui touche Galaxy : <https://training.galaxyproject.org/training-material/>

Ne pas hésiter à revenir vers moi pour toute question, tout commentaire.

Amicalement,

Yvan

De: "Lynda Feddak" <feddaklynda00@gmail.com>

À: "nadine le bris" <nadine.le_bris@obs-banyuls.fr>, "Yvan LE BRAS" <yvan.le-bras@mnhn.fr>

Cc: "islem kobbi" <islem.kobbi@etu.sorbonne-universite.fr>, "shivamshan sivanesan"

<shivamshan.sivanesan@etu.sorbonne-universite.fr>, "sofiane ouali"

<sofiane.ouali@etu.sorbonne-universite.fr>

Envoyé: Jeudi 2 Novembre 2023 19:25:13

Objet: Fwd: Rendez-vous du projet "Développement d'outils d'analyse IA et utilisation de la plateforme Galaxy-Ecology Europe pour le partage et l'exploitation de données d'imagerie sous-marine"

[Texte des messages précédents masqué]

[Texte des messages précédents masqué]