

Updating a production Galaxy platform : success and gotchas



CONFIGURATION

- 1 serveur Proxmox de virtualisation (24 CPU, 256 Go RAM)
 - VM de développement (2 CPU, 4 Go RAM)
 - VM de production (4 CPU, 30 Go RAM)
- Cluster SGE (différentes files dont une dédiée aux formations)
- Serveur frontal Apache avec *RewriteRule* (attention au *cookie-path*)
 - *attention au paramètre cookie-path*
- Authentification externe (LDAP via Apache)
- 3 *web handlers*, 3 *job handlers* avec *threadpoolworker* à 5

REINSTALLATION DE GALAXY ET MIGRATION DES DONNEES

- Mail aux utilisateurs d'annonce de migration + relance :
 - export des historiques et workflows
- Mise en place d'une VM de développement avec réinstallation complète de Galaxy + copie de la configuration, des outils et des banques :
 - test des outils manuellement avec création d'historiques en vue d'une mise en place de tests fonctionnels
- Mise en place d'une VM de production (copie de la VM de développement) :
 - modification de la configuration
- Réouverture avec conservation des anciens portails
 - mail aux utilisateurs

RESTAURATION DES DONNEES

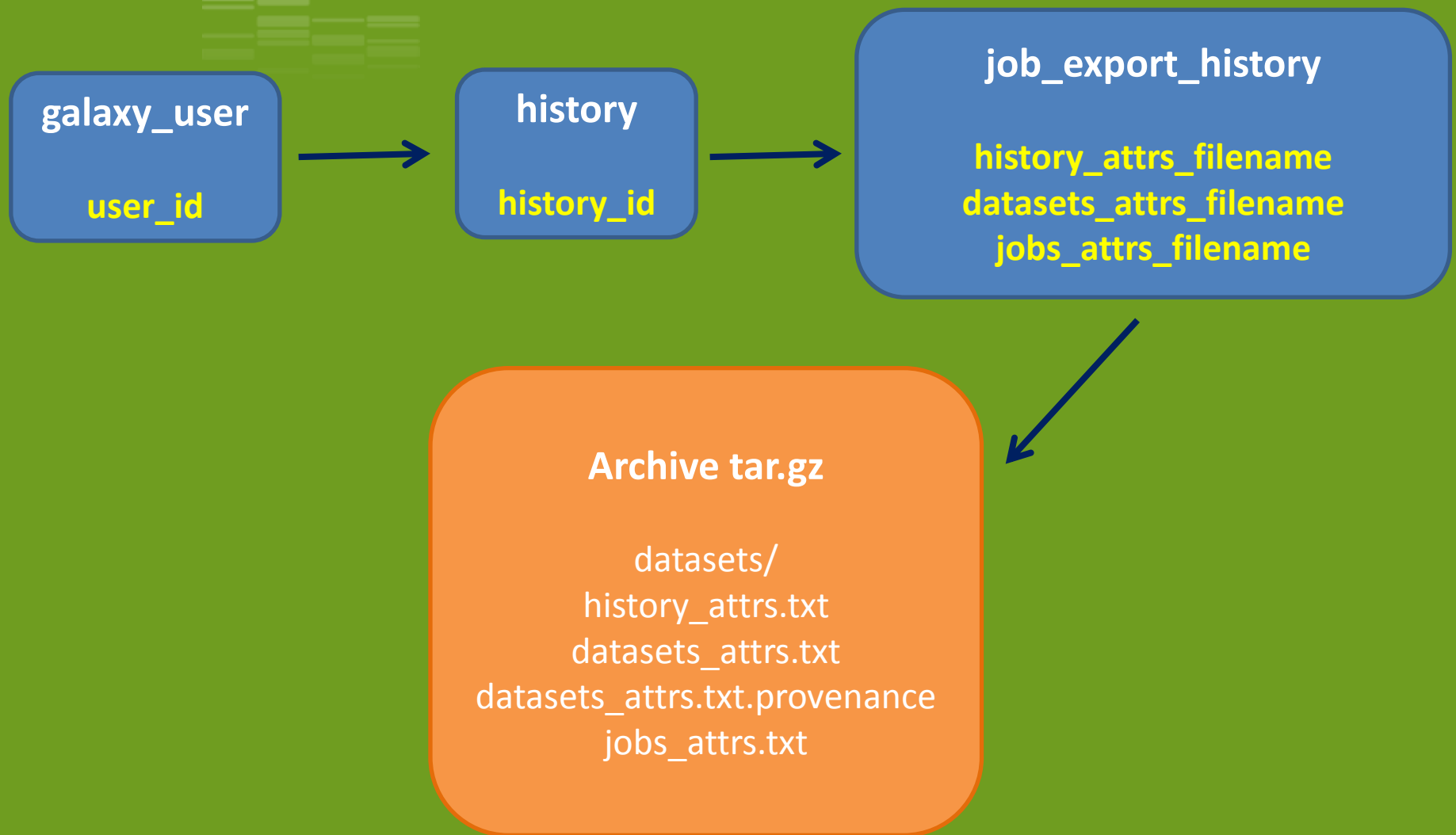
- Utilisation de la librairie Bioblend :
 - pour les workflows : *export_workflow_to_local_path*
 - pour les historiques : *export_history* et *download_history*
 - absence de message dans les logs
 - impossibilité d'export via le navigateur web et via Bioblend
 - clé d'export non générée par *export_history* et nécessaire à *download_history*
- Tests effectués sans succès :
 - retrait des web et *job handlers*
 - retrait du système d'authentification via le LDAP
 - augmentation du *TimeOut* sous Apache
 - passage par l'URL de la machine et non plus celle du proxy

PISTES DE RESOLUTION

Archive tar.gz

datasets/
history_attrs.txt
datasets_attrs.txt
datasets_attrs.txt.provenance
jobs_attrs.txt

PISTES DE RESOLUTION



PREMIER CAS

- **datasets_attrs.txt** : **uuid** présent pour chaque *dataset*
 - recherche dans la base de données dans la **table datasets**
 - récupération du **dataset_id**
 - recherche du dataset dans le **répertoire database/files** sur le serveur via **dataset_[dataset_id].dat**
 - copie du *dataset* dans le **répertoire datasets**
 - création d'une **archive tar.gz** à partir des fichiers et du répertoire *datasets*

Ceci est fonctionnel et nous avons récupéré deux historiques de cette manière.

DEUXIEME CAS

- **datasets_attrs.txt** : absence d'uuid mais présence de l'emplacement du *dataset* sur le serveur dans **file_name** :
 - copie du *dataset* dans le répertoire datasets
 - modification du fichier datasets_attrs.txt
 - création d'une **archive tar.gz** à partir des fichiers et du répertoire datasets

Ceci n'est actuellement pas fonctionnel !!! Nous sommes encore en train de travailler sur cet aspect.

CONCLUSION

- Points de vigilance :
 - fichiers de configuration
 - différences notables entre les différentes versions
- Démarches et scripts mis à disposition :
 - <http://migale.jouy.inra.fr/?q=fr/galaxy-historique>