2ème Webinar du NNCR Cluster

6 juin 2019

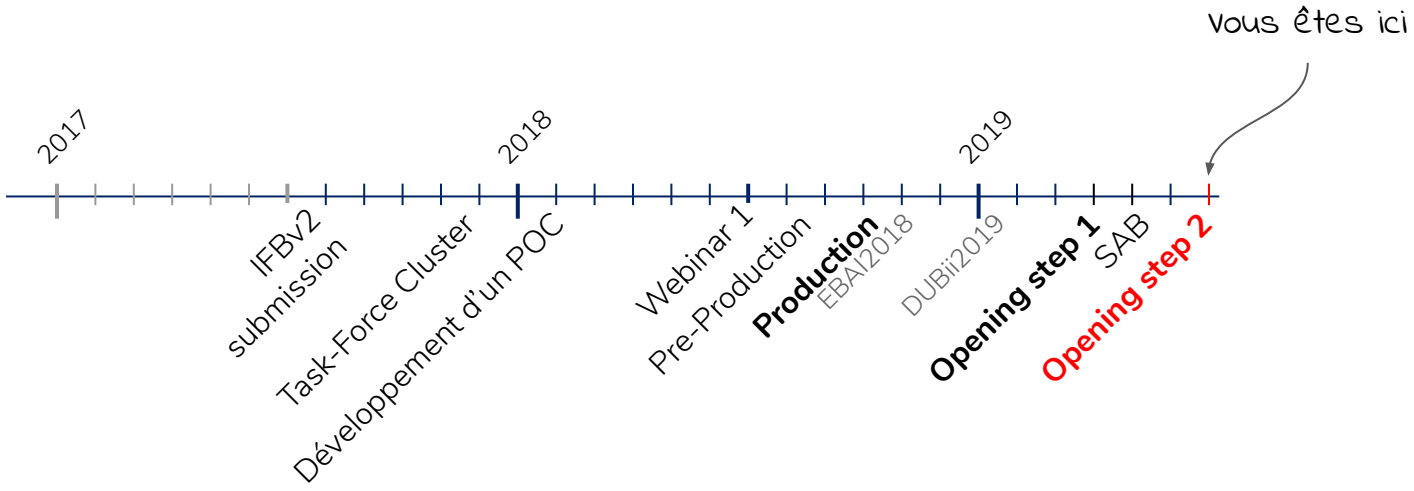
Les speakers du jour

Julien Seiler
IGBMC, Strasbourg
Responsable Cluster Infra

Gildas Le Corguillé
Station Biologique de Roscoff
Responsable Cluster Software

David Benaben
CBIB, Bordeaux
Administrateur système

François Gerbes
IFB, Roscoff
DevOps


Vous êtes ici
2017
2018
2019
IFBv2 submission
Task-Force Cluster
Développement d'un POC
Webinar 1
Pre-Production
Production
EBAI2018
DUBii2019
Opening step 1
SAB
Opening step 2

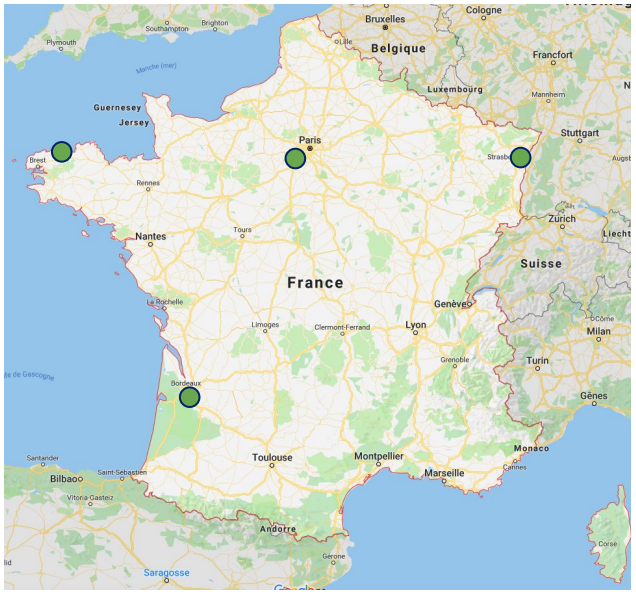
L’équipe Cluster

Ressources humaines

Environ 3.5 ETP dans les contributeurs actifs

- 5 permanents répartis sur 4 plateformes de bioinfo (1.5 ETP)
  - David Benaben CBIB
  - Olivier Inizan Migale
  - Gildas Le Corguillé ABiMS
  - Julien Seiler IGBMC
  - Guillaume Seith IGBMC
- 2 contractuels (2 ETP)
  - Nicole Charrière Hébergée à Migale
  - François Gerbes Hébergé à ABiMS

27 personnes inscrites sur le groupe de travail



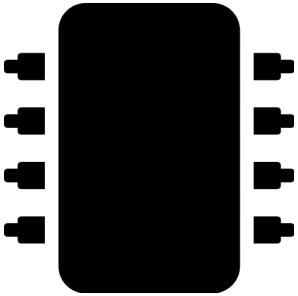
Cluster sans frontière

25714 messages depuis 14 mois (1800 / mois)

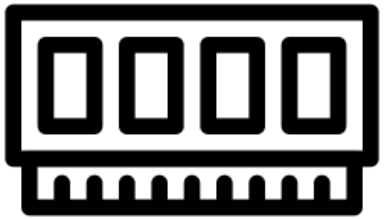
2 visio minimum par mois

Services proposés

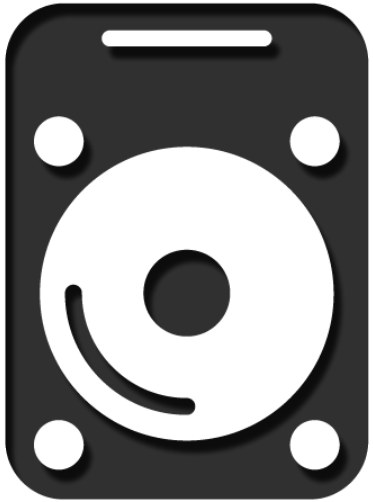
Infrastructure de calcul Unix



1968 coeurs



20 To RAM



400 To

Infrastructure de calcul Unix

68 **big** nodes

28 CPU + 256 Go RAM

1 **FAT** node

64 CPU + 3 To RAM

Infrastructure de calcul Unix

SSH

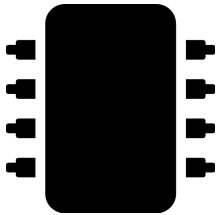
accès au cluster

SLURM

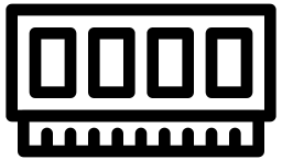
accès aux ressources / lancement des jobs

Infrastructure de calcul Unix

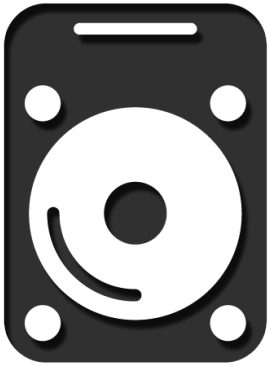
Infrastructure de calcul Unix



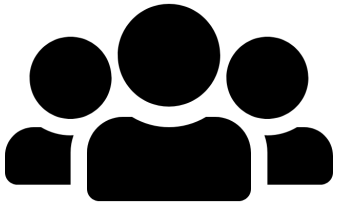
1968 coeurs



20 To RAM



400 To



Une communauté d’entraide

Infrastructure de calcul Unix

Catalogue de 112 outils

Les outils sont accessibles via `module`

ingularity

Infrastructure de calcul Unix

Il vous manque un outil ?

Vous maitrisez Git et Conda ou Singularity ?

Proposez le déploiement de l'outil via une *Merge Request* sur le dépôt de pilotage du déploiement des outils : https://gitlab.cluster.france-bioinformatique.fr

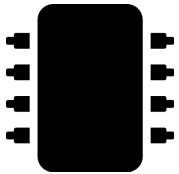
Vous savez expliquer ce dont vous avez besoin ?

Faite une demande sur la plate-forme de support communautaire : community.cluster.france-bioinformatique.fr

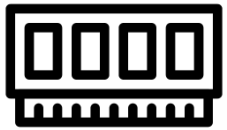
Vous êtes ultra pressé ?

Installez l'outil manuellement dans votre homedir ou votre espace projet

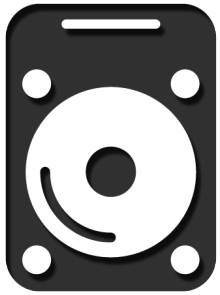
Infrastructure de calcul Unix



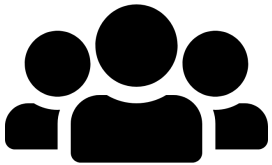
1968 coeurs



20 To RAM



400 To



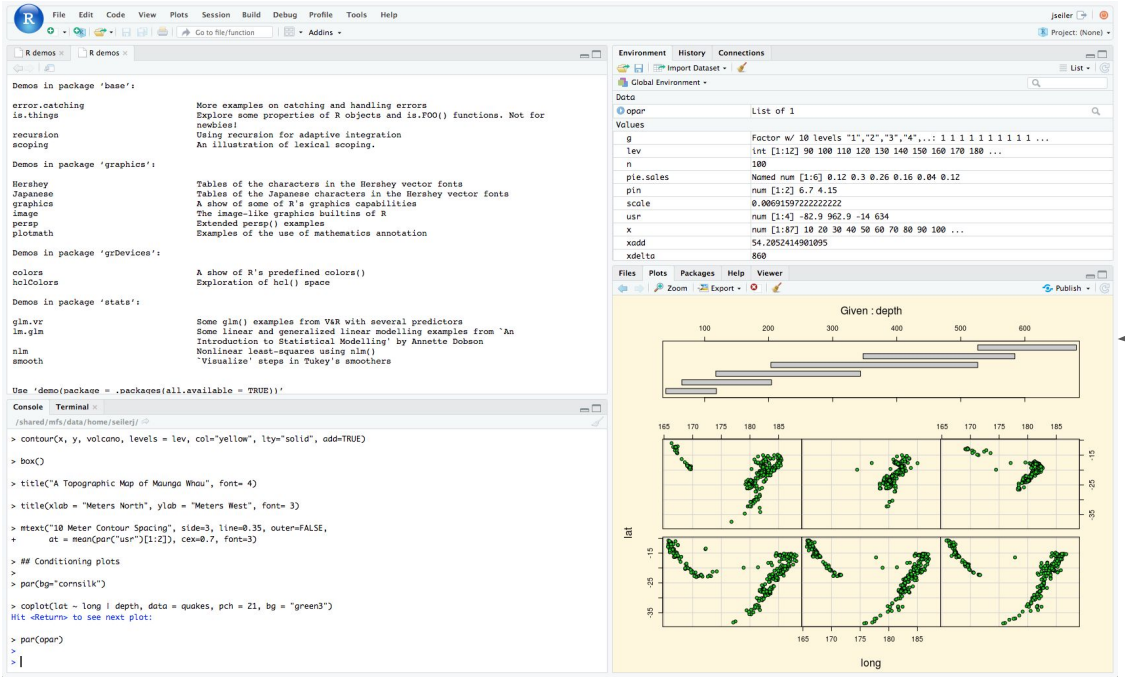
Une communauté d’entraide

112 outils clé en main

Serveur RStudio

https://rstudio.cluster.france-bioinformatique.fr

Un environnement de travail R dans votre navigateur



Serveur RStudio

https://rstudio.cluster.france-bioinformatique.fr

28 CPU + 256 Go RAM

- Environnement R 3.5.1 complet
- Installation de modules à la demande
- Terminal Unix en ligne
- Accès à l’ensemble du stockage cluster (homedir + projets)

Votre accès aux services cluster

1 compte Unix par utilisateur

SSH + RStudio / Homedir de 100Go

Votre accès aux services cluster

1 espace de travail par projet

1 To de stockage / 1 account SLURM

Accessible à **plusieurs** collaborateurs

Découpage possible en sous-projets

Votre accès aux services cluster

Demandez un compte utilisateur et/ou un espace projet ici :

https://www.france-bioinformatique.fr/fr/ifb-core-cluster-account-request

Votre accès aux services cluster

Et le tarif ?

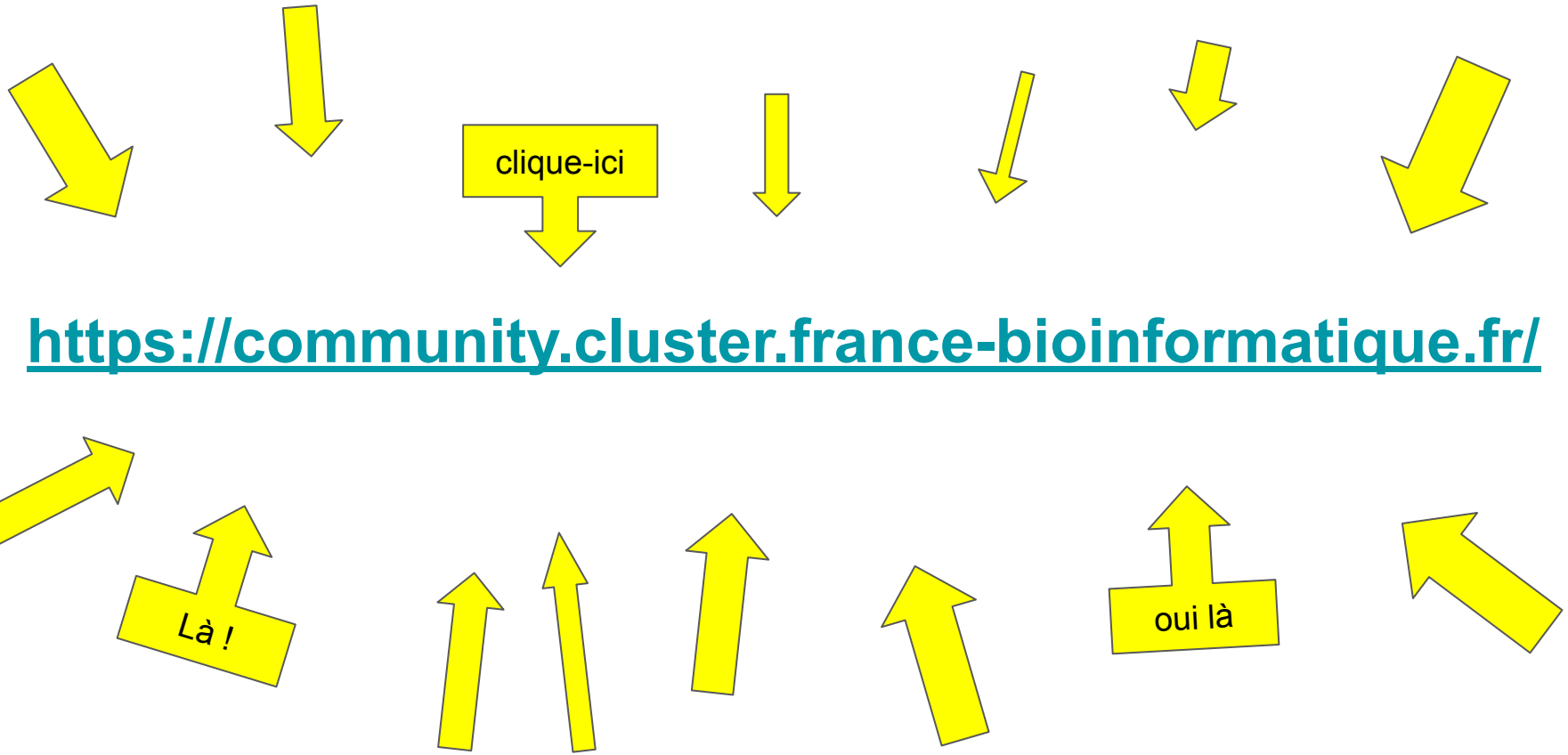
En cours de préparation...

L’action 1.4 de l’IFB étudie le meilleur modèle possible permettant de pérenniser l’infrastructure tout en assurant un accès simple aux services

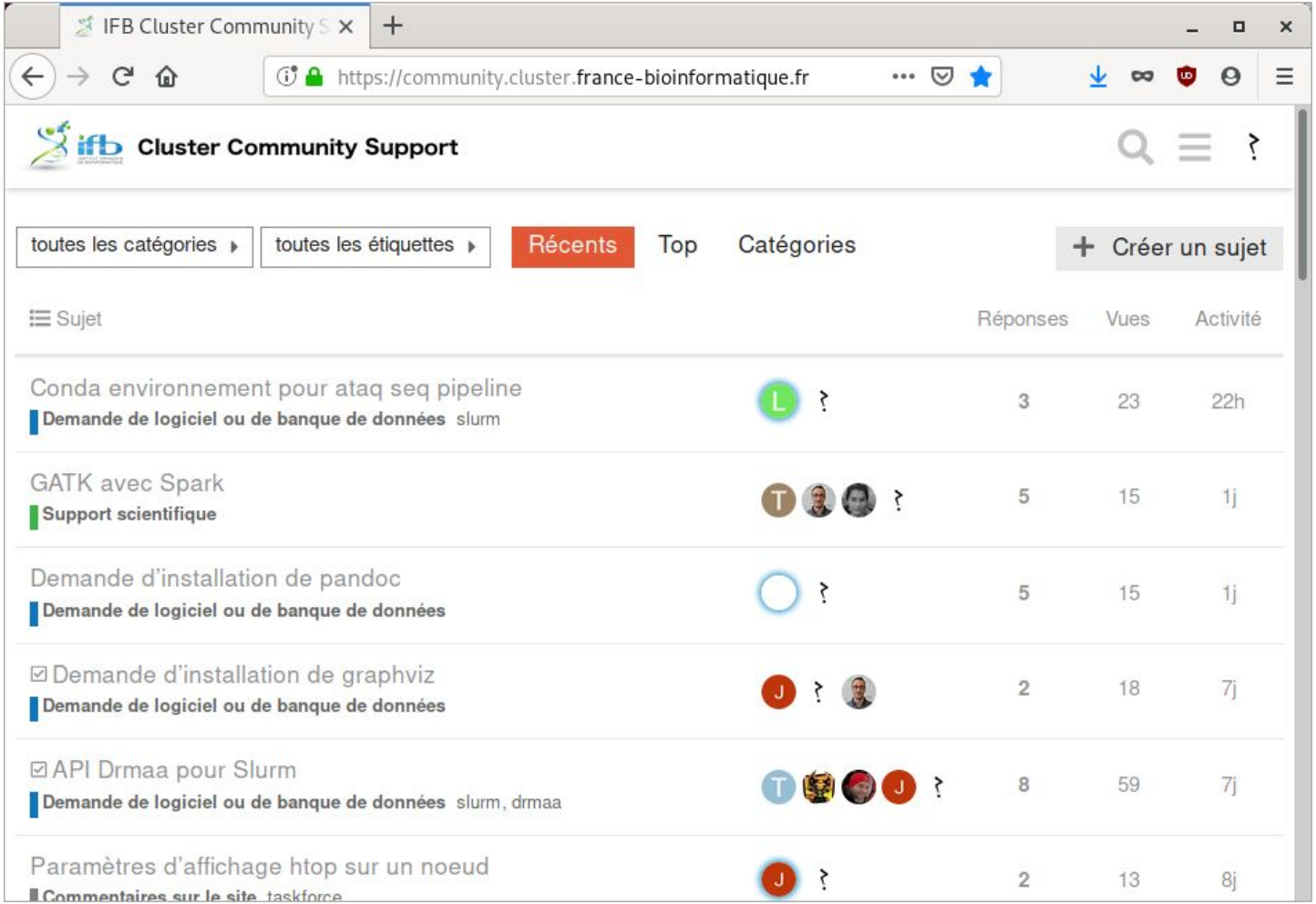
Services d’accompagnement

- Déploiement d’outils Unix ou modules R
- Déploiement d’applications pouvant s’appuyer sur les ressources de calcul
- Support à l’utilisation du cluster
- Organisation de formations (gestion des comptes des participants, mise à disposition des espaces partagés, création de bouquets logiciels personnalisés)

Community support



https://community.cluster.france-bioinformatique.fr/


IFB Cluster Community S
https://community.cluster.france-bioinformatique.fr
Cluster Community Support
toutes les catégories
toutes les étiquettes
Récents
Top
Catégories
Créer un sujet
Sujet
Réponses
Vues
Activité
Conda environnement pour ataq seq pipeline
Demande de logiciel ou de banque de données slurm
3
23
22h
GATK avec Spark
Support scientifique
5
15
1j
Demande d'installation de pandoc
Demande de logiciel ou de banque de données
5
15
1j
Demande d'installation de graphviz
Demande de logiciel ou de banque de données
2
18
7j
API Drmaa pour Slurm
Demande de logiciel ou de banque de données slurm, drmaa
8
59
7j
Paramètres d'affichage htop sur un noeud
Commentaires sur le site taskforce
2
13
8j

Plateforme communautaire IFB

- Plateforme basée sur Discourse (opensource GNU GPLv2)

- Ce “forum” se veut d’être un lieu d’**échange**, d’**entraide** et de **support**.
  - demander de l’aide autour du cluster (installation logiciel, questions autour du calcul)
  - discuter de bioinformatique

- Plusieurs catégories
  - Support IFB Core Cluster
  - Support scientifique
  - Non classé / Commentaires sur le site

- Demande public, anonyme ou privé (message direct)

Fonctionnement Contributif

Fonctionnement Contributif

Comment ça marche :

- Inscription sur https://gitlab.cluster.france-bioinformatique.fr/
- Demande d'accès en tant que développeur
- Exemple de contribution à partir d'un cas concret: https://community.cluster.france-bioinformatique.fr/t/demande-dinstallation-de-pandoc/170/4
- Création d'une branche sur le dépot: https://gitlab.cluster.france-bioinformatique.fr/taskforce/conda-env
- Edition du fichier sur l'interface de gitlab
- Création du Merge Request
- Demande de relecture

Fonctionnement Contributif

conda-env

2.3.4.3.yml 117 octets

```yaml
channels:
  - conda-forge
  - bioconda
  - default
dependencies:
  - bioconda::bowtie2=2.3.4.3
name: bowtie2-2.3.4.3
```

rstudio-packages

rstudio.yml 4,8 ko

```yaml
---



r_packages:
  - { name: BiocManager, type: cran, state: present } # For Bioconductor from

  - { name: ade4, type: cran, state: present }          # https://support.clus

  - { name: bioFAM/MOFA/MOFA, type: github, state: present } #  For DUBii 201

  - { name: caTools, type: cran, state: present }       # For DUBii 2019
  - { name: caret, type: cran, state: present }         # For DUBii 2019
  - { name: clues, type: cran, state: present }         # https://support.clus
  - { name: ComplexHeatmap, type: bioconductor, state: present } # https://sup
  - { name: corrplot, type: cran, state: present }      # For DUBii 2019
  - { name: cowplot, type: cran, state: present }       # For DUBii 2019
```

usegalaxy-fr-tools

metabolomics.yml 6,6 ko

```yaml
---



tools:
# label_preprocessing
## LCMS
  - name: msnbase_readmsdata
    owner: lecorguille
    tool_panel_section_id: metabolomics_preprocessing_lcms
    revisions:
      - 86a20118e743 # 2.8.2.1
  - name: xcms_xcmsset
    owner: lecorguille
    tool_panel_section_id: metabolomics_preprocessing_lcms
    revisions:
      - 3bd53dcfa438 # 3.4.4.1
```

Projets / Perspectives / Appel à contribution

Projets en cours

- Implémentation de *My*, un portal de création, et de gestion de compte (François Gerbes)
  Basé sur “My Accounts Manager” de la plateforme GenOuest
  https://github.com/genouest/genouestaccountmanager

- usegalaxy.fr
  - Sub-domaines pour entre autres Workflow4Metabolomics, ProteoRE, …
  - Possibilité de gestion communautaire via Git

Perspectives

- Harmonisation des pratiques entre les clusters de l’IFB
  - Bonnes pratiques
  - Diffusion de configuration serveur (rôles et playbook Ansible)
  - Partage d’outils et de banques
- Améliorer le support de Singularity
  - Sanctuarisation des environnement Conda
  - Transparence dans l'utilisation des images (Julien Seiler)
- Stockage : en 2019 extension de l’espace de stockage +1Po
- Hébergement de projet pilote IFB ou autres portails d’analyse
- Vos idées

Appel à contribution

- Participation au support aux utilisateurs sur le community forum
- Déploiement d’outils de bioinformatique
- Documentation utilisateur
- Déploiement d’instances thématique Galaxy
- Gestion des comptes utilisateurs
- Support de l’authentification EduGain
- Stockage des données

L’appel à contribution sur la plate-forme communautaire

Les services clusters de l’IFB

ssh://**core**.cluster.france.bioinformatique.fr	High Performance Computing
https://**rstudio**.cluster.france.bioinformatique.fr	R dans votre navigateur
https://**community**.cluster.france.bioinformatique.fr	Entraide et support communautaire
https://**gitlab**.cluster.france.bioinformatique.fr	Contribuer !