

Fédération des plateformes Jupyter

Violaine Louvet, CNRS

Mathrice, 28 mars 2018



- ▶ Des **initiatives** autour de la mise en place de serveurs Jupyter dans de plus en plus de labos
- ▶ Des **technologies** très évolutives, de plus en plus intéressantes
- ▶ Des **usages** variés : enseignement, recherche, valorisation



Qu'est-ce que c'est ?



Notebooks Jupyter

Notebooks : programmes contenant à la fois du texte en markdown et du code en Julia, Python, R...

Jupyter : Application web utilisée pour écrire des notebooks

A quoi ça sert ?

- ▶ Très pratique pour présenter des résultats pas à pas, expliquer des contextes.
- ▶ Tester des bouts de code.
- ▶ Présenter, échanger du code.
- ▶ Assurer une reproductibilité des traitements.
- ▶ Utiliser un environnement de programmation clé en main.
- ▶ Permettre l'accès à des outils, codes sans avoir à les installer.
- ▶ Déporter les calculs.

-> voir présentation Matthieu

Qui est intéressé ?



- ▶ **Enseignement, formation** : pour disposer d'un outil utilisable par tous sans installation particulière. Permettant le partage et la mise en place de TP.
- ▶ **Recherche** : outil pour le test, le partage, la mise en place de pipeline, la simplification d'accès aux outils numériques.
- ▶ **Valorisation** : donner l'accès à des codes numériques de laboratoires pour les entreprises.

Le Groupe Calcul

Pour faciliter la pratique dans le cadre des formations organisées



- ▶ Développer un **portail national** qui va s'appuyer sur des instances BinderHub locales
- ▶ **Déploiement sur openstack** localement sur chaque site
- ▶ Préparation d'un **environnement** permettant de déployer facilement
- ▶ S'appuyer sur une collaboration avec les développeurs de Binderhub

-> voir présentation David

L'exemple de cocalc <https://cocalc.com/>

Un peu ce qu'on veut faire mais sous forme commerciale.

- ▶ Les plus : facilité de mise en oeuvre, ergonomie, très grand nombre de kernels disponibles, outils collaboratifs embarqués, etc.
- ▶ Les moins : la ressource coûte vite très cher.



- ▶ **Mathrice**
 - ▶ Orsay, **LAL**
 - ▶ Palaiseau, **Polytechnique** (Initiative HPC@Math - financement - avec de la main d'œuvre du Mésocentre PHYMATH)
 - ▶ Grenoble, **GRICAD**
 - ▶ Reims, **ROMEO** (mésocentre)
 - ▶ Pau, **LMA**
 - ▶ Lyon, **ICJ**
 - ▶ Strasbourg, **IPHC (et France Grilles) et IRMA**
-
- ▶ Mise en place d'une liste de discussions
 - ▶ Première réunion en février.
 - ▶ Réunion de travail en mars avec les développeurs de binderHub



- ▶ **Demande croissante** : la question de la possibilité de traiter de grosses masses de données, et d'accéder à travers le notebook à des **moyens de calcul importants**
- ▶ Pas encore quelque chose qui émerge réellement dans le **monde du HPC** : aucune session à SC17 par exemple. Stratégiquement, on est un peu en avance.
- ▶ Des outils qui deviennent **incontournables** pour certains usages : exemple des challenges liés aux données (Center for Data Science de Saclay ou Data Institute de Grenoble par exemple)
- ▶ Des frémissements dans les **projets** à différents niveaux:
 - ▶ projets européens où la notion de notebook commencent à apparaître (y compris hors opendreamkit)
 - ▶ projet régional avec aspect valorisation / services aux entreprises important
- ▶ Un grand intérêt de **différentes communautés** : santé, matériaux, SHS ...