

The IFB Core Cluster Infrastructure

EBAII 2023

Gildas Le Corguillé, Thomas Denecker & Julien Seiler
IFB Core Cluster taskforce

Besoin d'aide ?

Il manque un outil !

J'ai plus de place !

HEEELP !

Je ne trouve pas un index...

Rejoignez la communauté IFB

Rendez-vous sur :

<https://community.france-bioinformatique.fr>

The screenshot shows the IFB Community Support forum interface. At the top, there is a search bar and a user profile icon. Below that, there are filters for 'all categories' and 'all tags'. The main content area is divided into sections for different categories. The first section is 'IFB Core Cluster' with 406 topics. It includes a list of topics such as 'Demande de logiciel ou de banque de données', 'Support IFB Core Cluster', and 'Meilleure manière de...'. The second section is 'Workflow4Metabolomics' with 46 topics, including 'Troubleshooting - W...', 'Troubleshooting - G...', and 'Erreur Python lors de l'e...'. The third section is 'Support scientifique et ...' with 30 topics, including 'Galk spark beta', 'Comment optimiser les r...', and 'Utilisation gpu via sbatch'.

High Performance Computer

Votre ordinateur peut-il faire de la bioinformatique ?



Un ou deux microprocesseurs

Un microprocesseur est chargé de l'exécution des instructions élémentaires demandées par le logiciel

4 à 8 Go de mémoire vive (RAM)

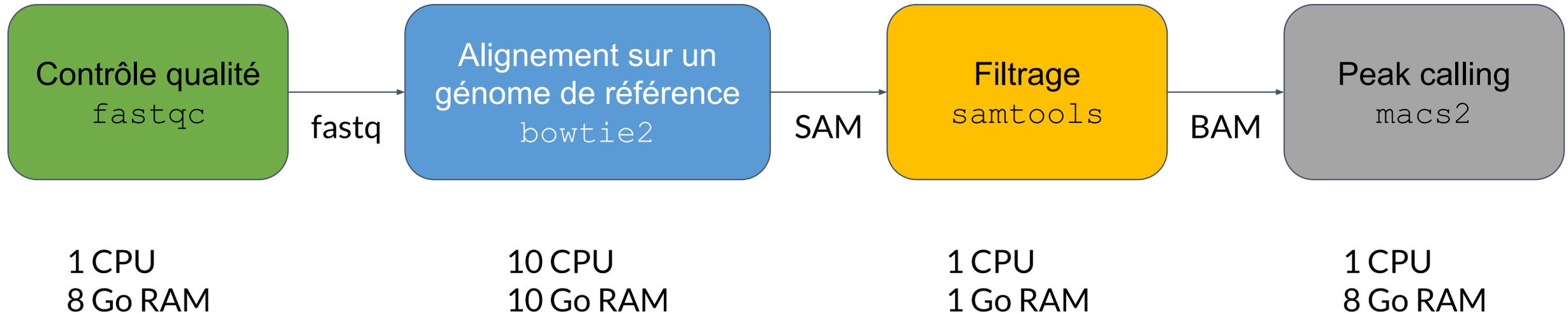
La mémoire vive est utilisée par le microprocesseur pour traiter les données

≈ 1 To d'espace de stockage

L'espace de stockage est utilisé pour conserver de grandes quantités de données de manière plus permanente



Votre ordinateur peut-il faire de la bioinformatique ?



L'exécution de ce workflow nécessite au minimum toutes les ressources d'un ordinateur de bureau pendant plusieurs heures et ceci seulement pour 1 seul fichier fastq.

Pour faire ce type d'analyse nous avons besoin d'ordinateurs plus puissants !

Du data center au coeur



Le Data Center de l'IDRIS
Un bâtiment conçu pour accueillir des infrastructures informatiques

Du data center au coeur

Groupes froid
Pour refroidir les
équipements



Du data center au coeur

Groupe électrogène
Pour garantir l'alimentation
électrique



Du data center au coeur



Les armoires de l'IFB

Chaque armoire peut contenir
80 super-ordinateurs

Du data center au coeur

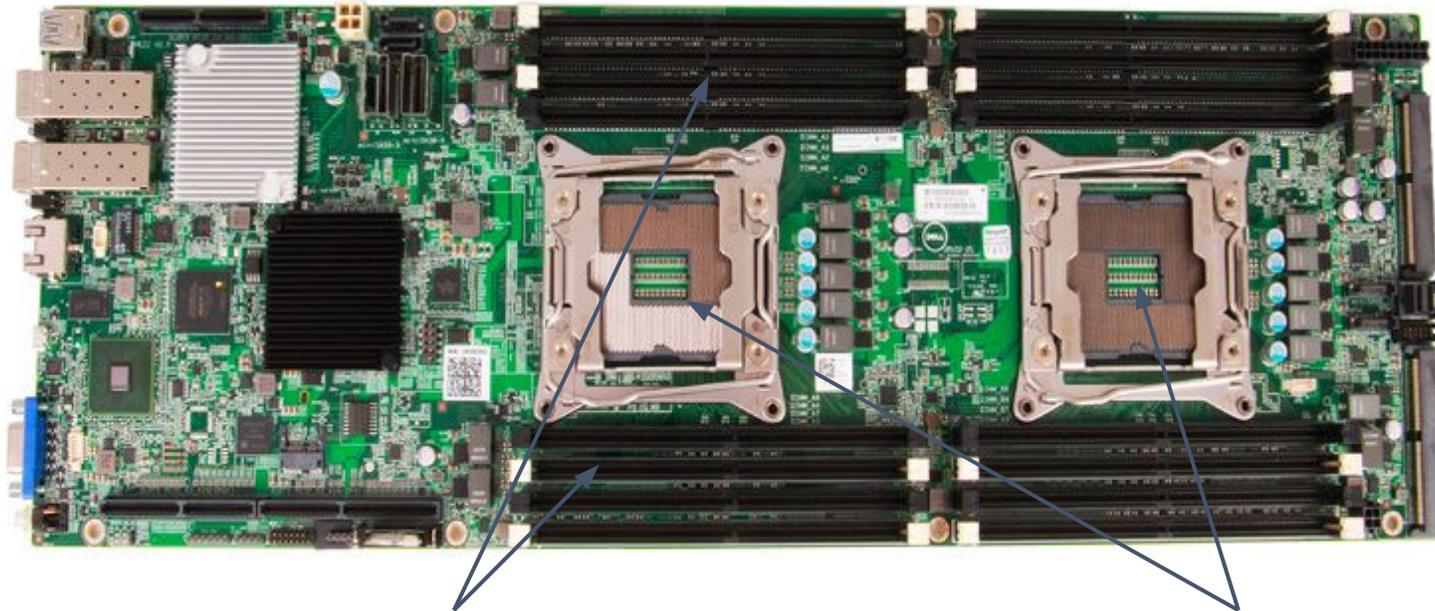


ordinateurs de calcul

Baies de stockage

Du data center au coeur

Un ordinateur ou **noeud** de calcul



Mémoire vive

Supports processeurs

Du data center au coeur

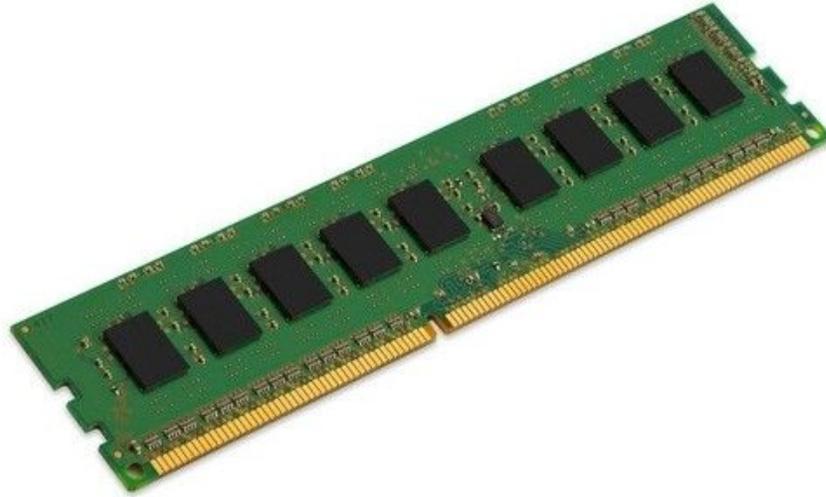
Un microprocesseur



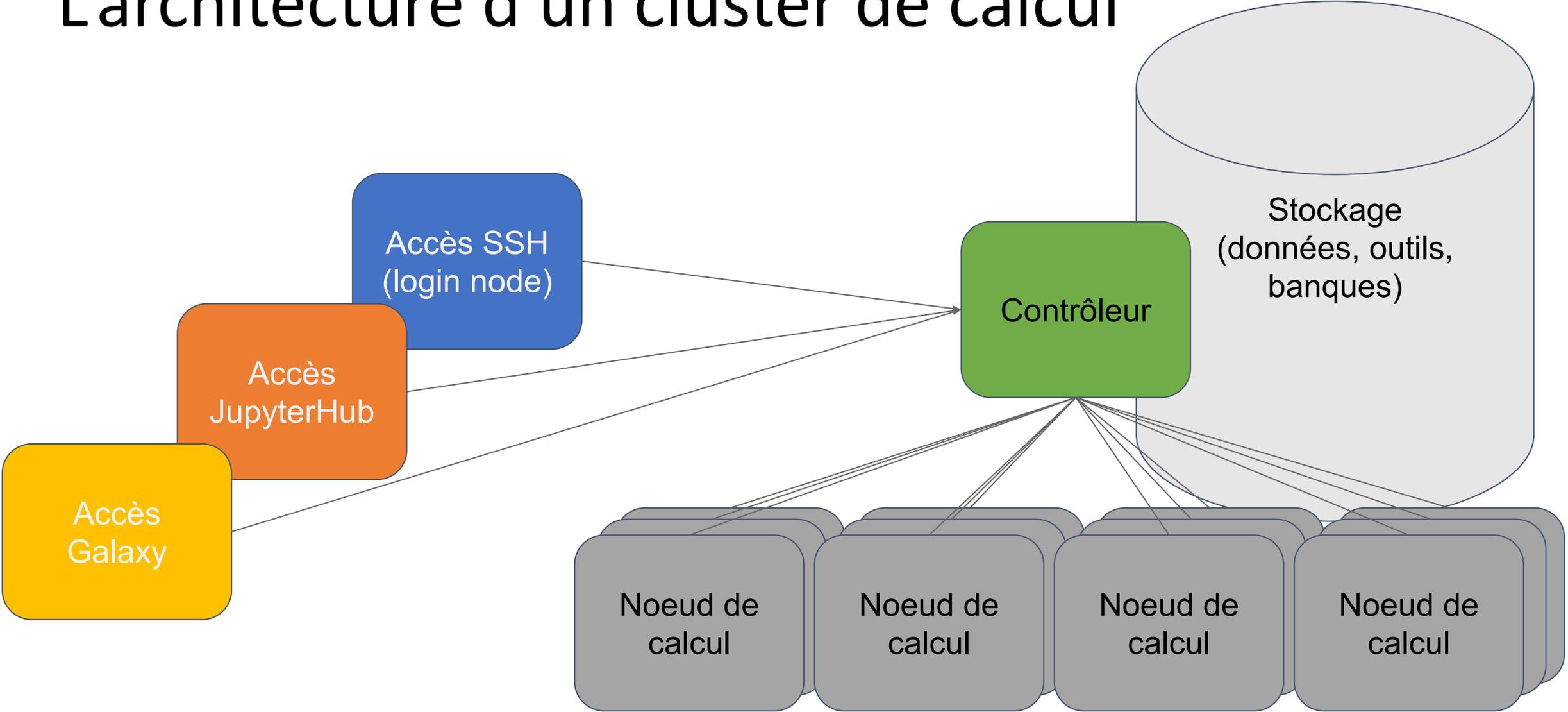
Un microprocesseur contient plusieurs **coeurs**
Chaque coeur se comporte comme un microprocesseur unique.

Du data center au coeur

La mémoire vive RAM



L'architecture d'un cluster de calcul



Software environment



module load fastqc/0.11.9



BIOCONDA[®]

par défaut



si une licence doit être acceptée
si ce n'est pas dans Bioconda et difficile à intégrer ou urgent
si un conteneur Docker existe

Pour toute demande d'installation : <https://community.france-bioinformatique.fr/>

Besoin d'un outil ?

Demande d'installation de FusionInspector

IFB Core Cluster | Demande de logiciel ou de banque de données

This is the first time Camille has posted — let's welcome them to our community! 20d

C Camille

Bonjour,

Je me permets de vous contacter car dans le cadre de mon projet de master 2, j'aurais besoin de l'outil FusionInspector pour lequel mon ordinateur personnel n'est pas assez puissant. C'est un outil permettant d'analyser les fusions en RNAseq. Ne sachant pas comment procéder, serait-il possible de vous l'installer ? Voici l'adresse <https://github.com/FusionInspector/FusionInspector/wiki/installing-FusionInspector>.

Je vous serais très reconnaissante de votre aide.

Cordialement,

Camille Baron encadrée par Audrey Gros à l'université de Bordeaux

✓ Solved by [gildaslecorguille](#) in [post #4](#)
Hop: module load fusion-inspector/2.2.1

created 20d last reply 19d 4 replies 25 views 2 users 2 links

Installation (nonpareil) @team.software

IFB Core Cluster | Demande de logiciel ou de banque de données

E echase 3 Feb

Bonjour @team.software

Serait-il possible d'installer nonpareil sur le cluster, SVP ?

[GitHub](#)

[Imrodriguez/nonpareil](#) 1
Estimate metagenomic coverage and sequence diversity - Imrodriguez/nonpareil

Merci pour votre temps,

Emily

✓ Solved by [julien](#) in [post #3](#)
nonpareil 3.3.4 est à présent disponible sur le cluster : module load nonpareil/3.3.4 Bonne journée, Julien

created 3 Feb last reply 4 Feb 2 replies 28 views 2 users 2 likes 1 link

Installation pyslim @team.software

IFB Core Cluster | Demande de logiciel ou de banque de données

G Guillaume_Lan-Fong 24d

Bonjour,

Faisant suite à mon post pour l'installation de SLiM, j'aimerais savoir s'il était possible d'installer le package python "pyslim" (<https://pyslim.readthedocs.io/en/latest/installation.html>) permettant de travailler sous python sur les .trees générés par SLiM.

L'installation devrait être possible via un simple : `python3 -m pip install pyslim`

En vous remerciant par avance,

Lan-Fong Guillaume

✓ Solved by [gildaslecorguille](#) in [post #7](#)
Hop: module load pyslim/0.501

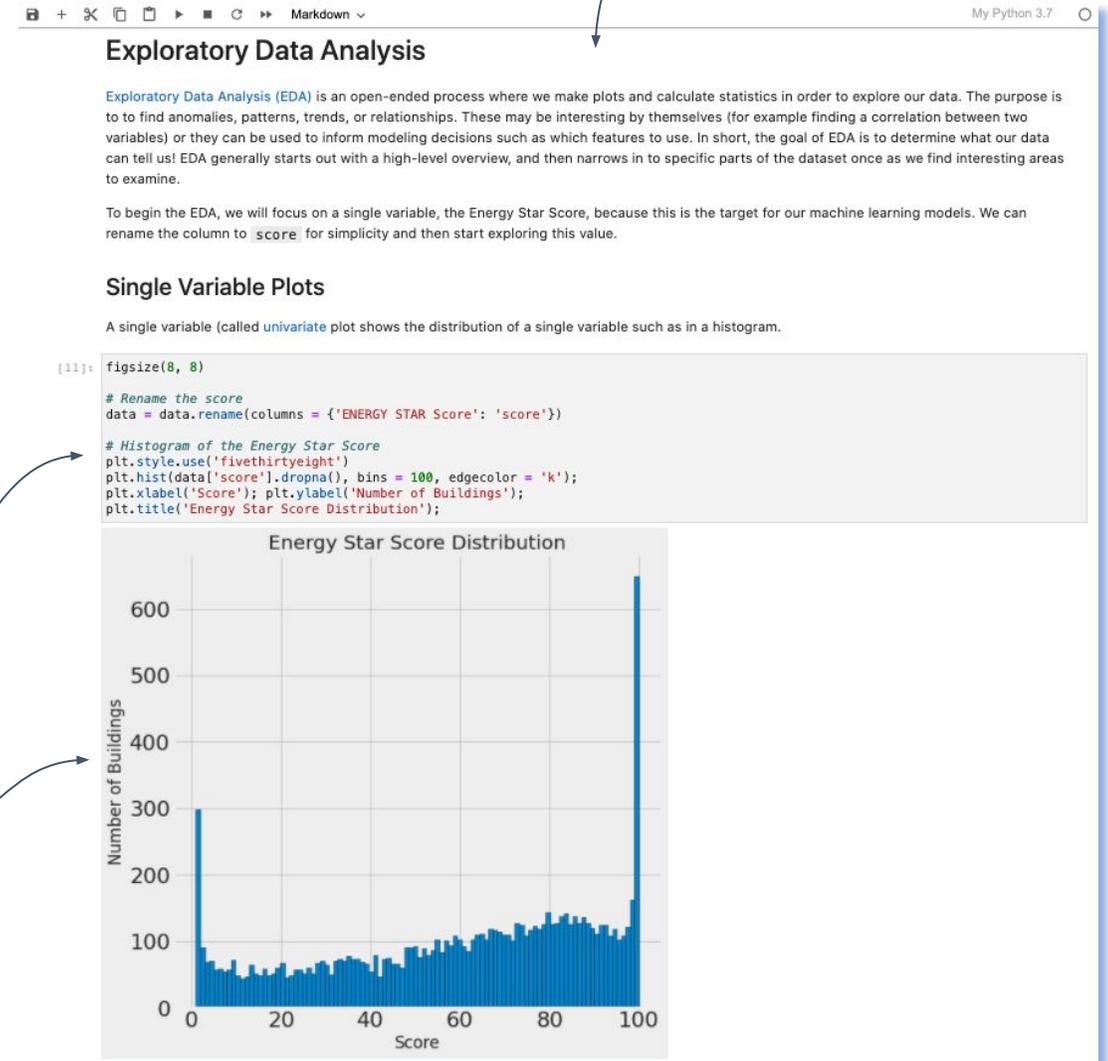
created 24d last reply 11d 7 replies 38 views 3 users 1 like 1 link

Jupyter

Qu'est-ce qu'un Notebook Jupyter ?

- Un fichier spécial avec l'extension `.ipynb`
- Combinaison de **Markdown** et de **code**
- Le code peut être exécuté à l'intérieur du notebook
- La sortie du code est intégrée directement dans le notebook

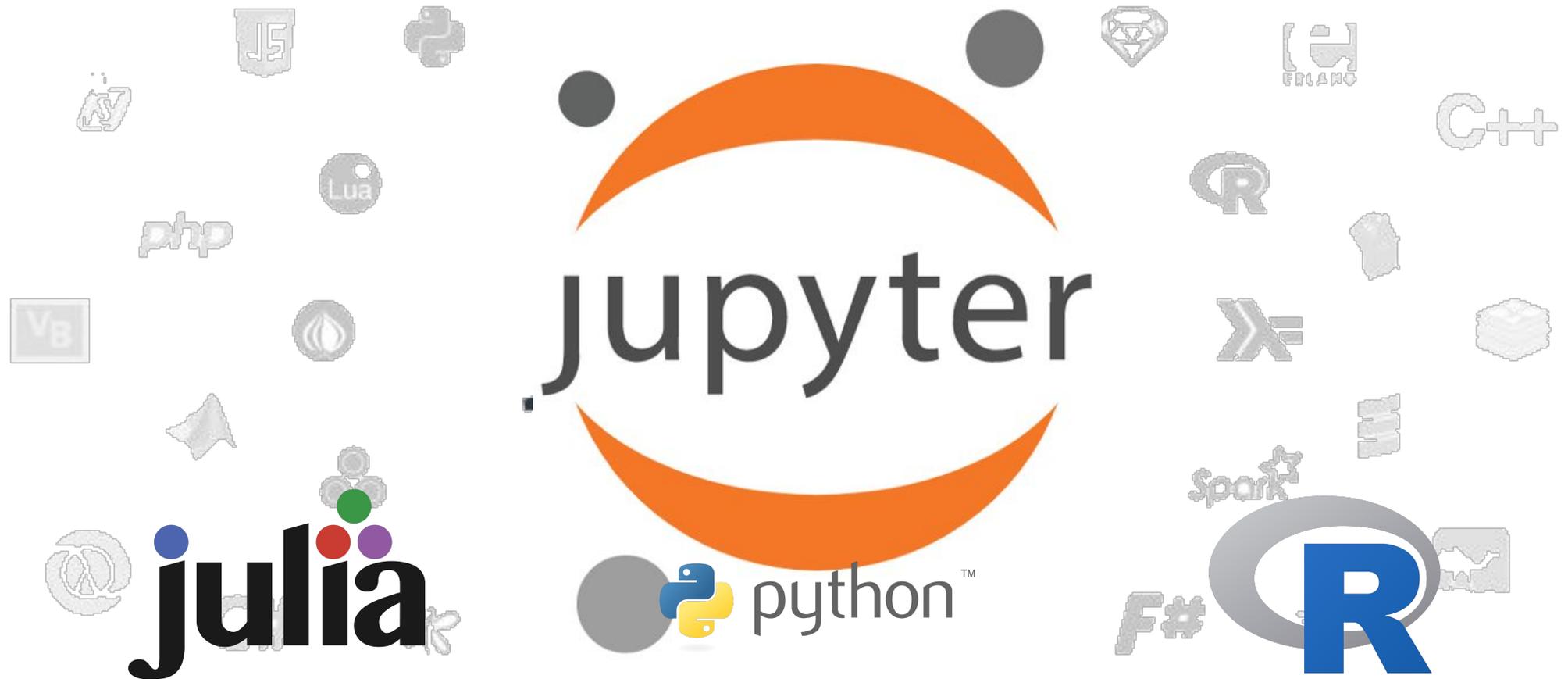
Cellule
Markdown



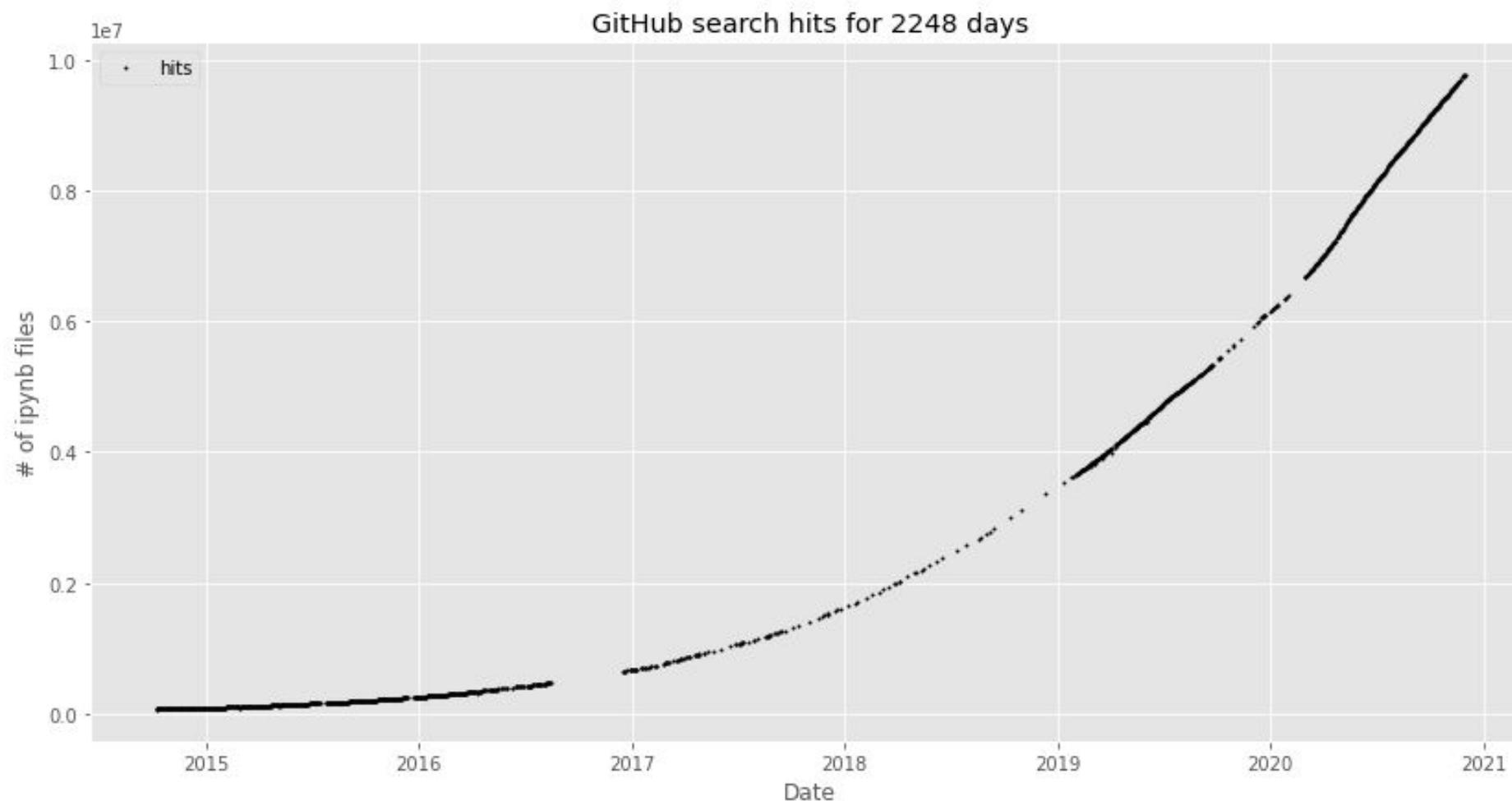
Cellule
de code

Sortie du
code

Qu'est-ce qu'un Notebook Jupyter ?



Les notebooks sont populaires



Pourquoi s'intéresser aux notebooks ?

- Les notebooks vous permettent d'analyser des données et de rédiger des rapports en un seul endroit
- Ils prennent en charge la visualisation des données en temps réel
- Vous pouvez facilement inclure une section interactive dans un notebook
- En accord avec le principe de la science reproductible

**Les Notebooks Jupyter sont
le cahier de laboratoire pour la science des données.**

Qu'est-ce que JupyterLab

JupyterLab est le logiciel permettant de manipuler les fichiers notebook.

C'est une "application web"

On y accède au travers d'un navigateur web

Pour utiliser JupyterLab, il faut un "serveur" JupyterLab

jupyterlab

Qu'est-ce que JupyterHub

JupyterHub est une application web qui vous permet de créer des serveurs JupyterLab sur un cluster ou une infrastructure en nuage.



authentication
spawner
proxy

IFB core cluster

En résumé

Notebook Jupyter :

le format de fichier mêlant Markdown (documentation) et Code (Python, R, Bash, etc.)

JupyterLab :

l'application permettant de manipuler des fichiers notebook

JupyterHub :

le service permettant de lancer des serveurs JupyterLab sur le cluster

SLURM

Présentation et démonstration ->

https://gitlab.com/ifb-elixirfr/cluster/tutoriel_slurm

IFB, une offre nationale

L'IFB Core Cluster n'est pas la seule infrastructure

NNCR - National Network
Computing Resources

Cluster	Localisation du Data center	Coeurs	RAM (Go)	Stockage (To)
IFB Core	IDRIS - Orsay	8 434	55 770	2 000
Genotoul	Toulouse	6 128	34 304	3 000
ABiMS	Roscoff	3 200	15 589	2 500
GenOuest	Rennes	1 824	7 500	2 300
Migale	Jouy en Josas	1 084	7 000	350
BiRD	Nantes	560	4 000	500

Besoin d'aide ?

Il manque un outil !

J'ai plus de place !

HEEEELP !

Je ne trouve pas un index...

Rejoignez la communauté IFB
(on a des cookies...)

Rendez-vous sur :

<https://community.france-bioinformatique.fr>

The screenshot shows the IFB Community Support forum interface. At the top, there is a search bar and a user profile icon. Below that, there are filters for 'all categories' and 'all tags'. The main content area is divided into sections for different categories. The first section is 'IFB Core Cluster' with 406 topics. It includes sub-sections for 'Demande de logiciel ou de banque de données', 'Support IFB Core Cluster', and 'Meilleure manière de...'. The second section is 'Workflow4Metabolomics' with 46 topics, including 'Troubleshooting - W...', 'Troubleshooting - G...', and 'Erreur Python lors de l'e...'. The third section is 'Support scientifique et ...' with 30 topics, including 'Galk spark beta', 'Comment optimiser les r...', and 'Utilisation gpu via sbatch'.