

The IFB Core Cluster Infrastructure, etc ...

EB3I 2025

Gildas Le Corguillé, Julien Seiler
IFB Core Cluster taskforce

C'est quoi et pourquoi
un High Performance Computer /
Cluster de Calcul

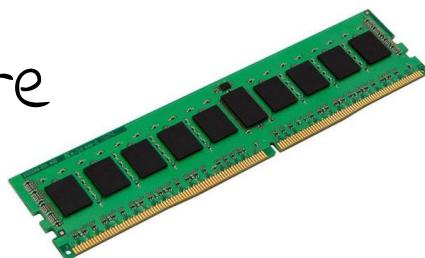
Mais avant, c'est quoi un ordinateur ?

Composants d'un ordinateur

Processeur
(4 coeurs)



RAM / Mémoire
(de travail)
(8 à 16 GB)



Disque dur
Stockage
(512 GB à 1 TB)



Carte ou puce
graphique



Réseau
filaire ou wifi



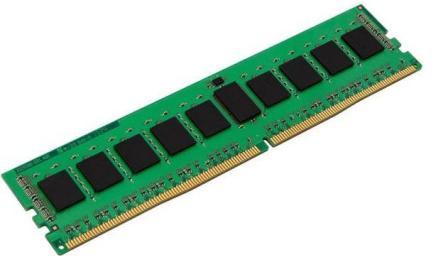
Alimentation



Dissipateur et
ventilateur



Processeur
(4 coeurs)



RAM / Mémoire
(de travail)
(8 à 16 GB)



Disque dur
Stockage
(512 GB à 1 TB)



Cuisinier·ères

Plan de travail pour
les ingrédients utiles à
la recette (gateau)

Zones de stockage :
cellier, réfrigérateur, garage, ...

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

8 à 16 Go de mémoire vive (RAM)

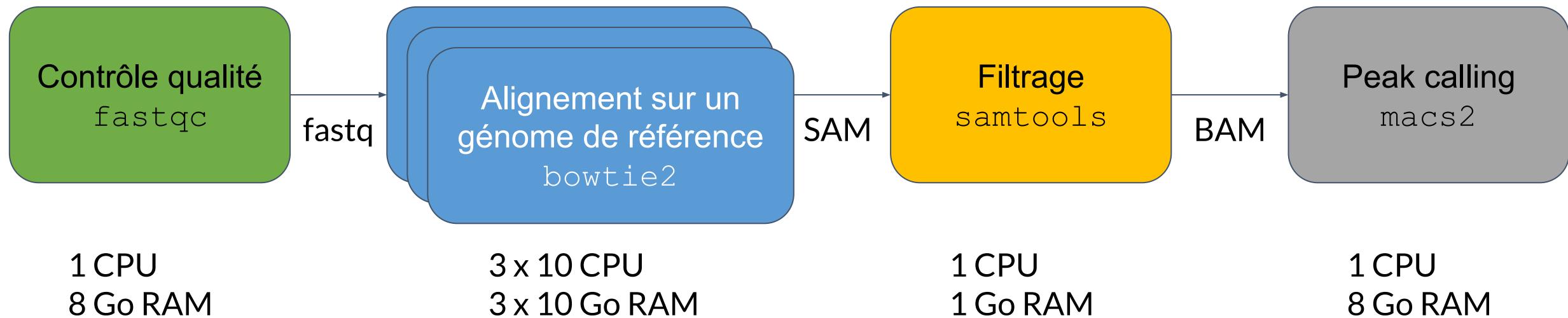
La mémoire vive est utilisée par le microprocesseur pour traiter les données

≈ 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 ?



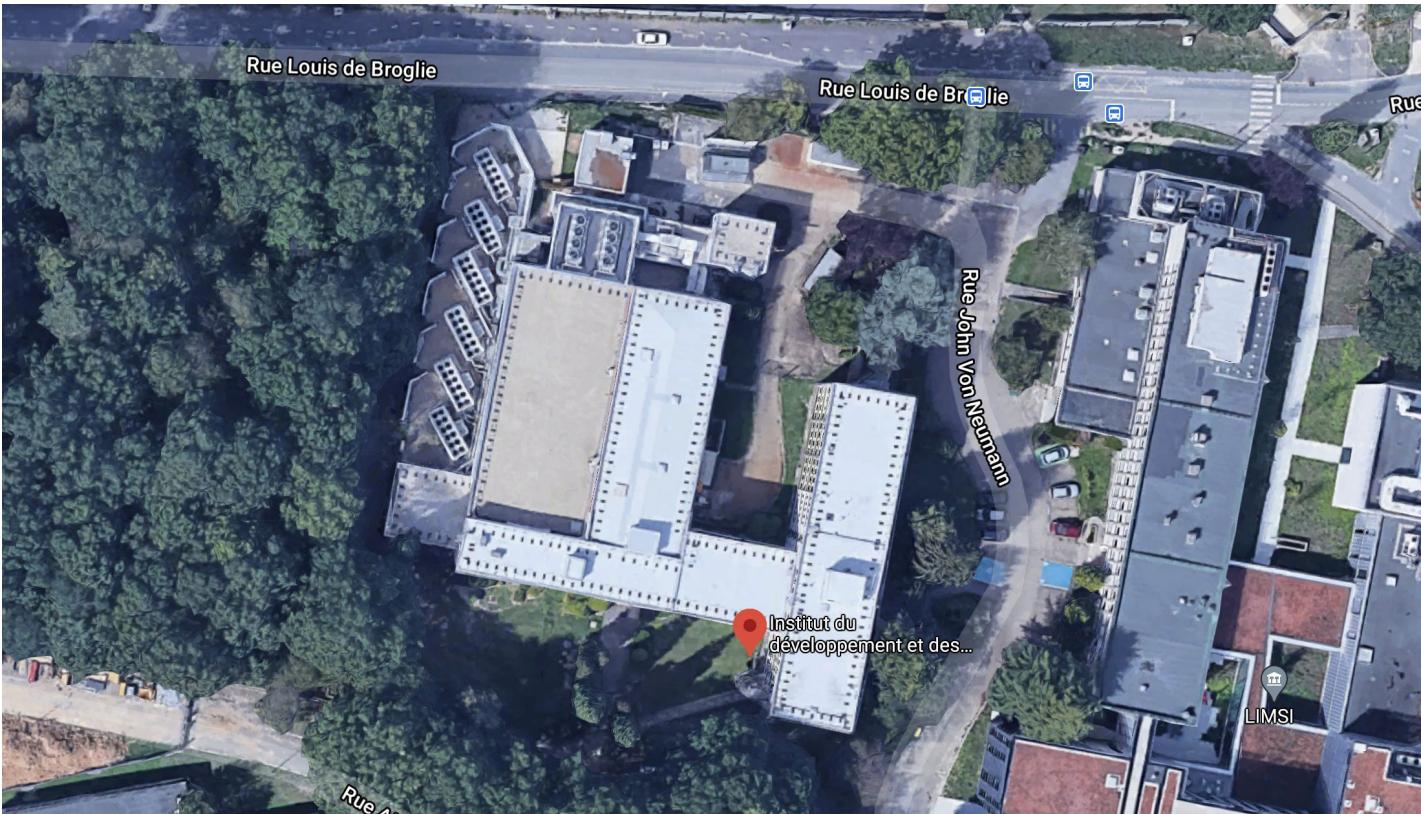
L'exécution de ce workflow nécessite au minimum toutes les ressources d'un ordinateur de bureau pendant plusieurs heures et ceci seulement pour **1 seul** fichier fastq.

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

Biiiggerrr



Du data center au cœur



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

Du data center au cœur

Groupes froid

Pour refroidir les équipements



Du data center au cœur

Groupe électrogène

Pour garantir l'alimentation électrique



Du data center au cœur



Les armoires de l'IFB

Chaque armoire peut contenir
85 super-ordinateurs

Du data center au cœur

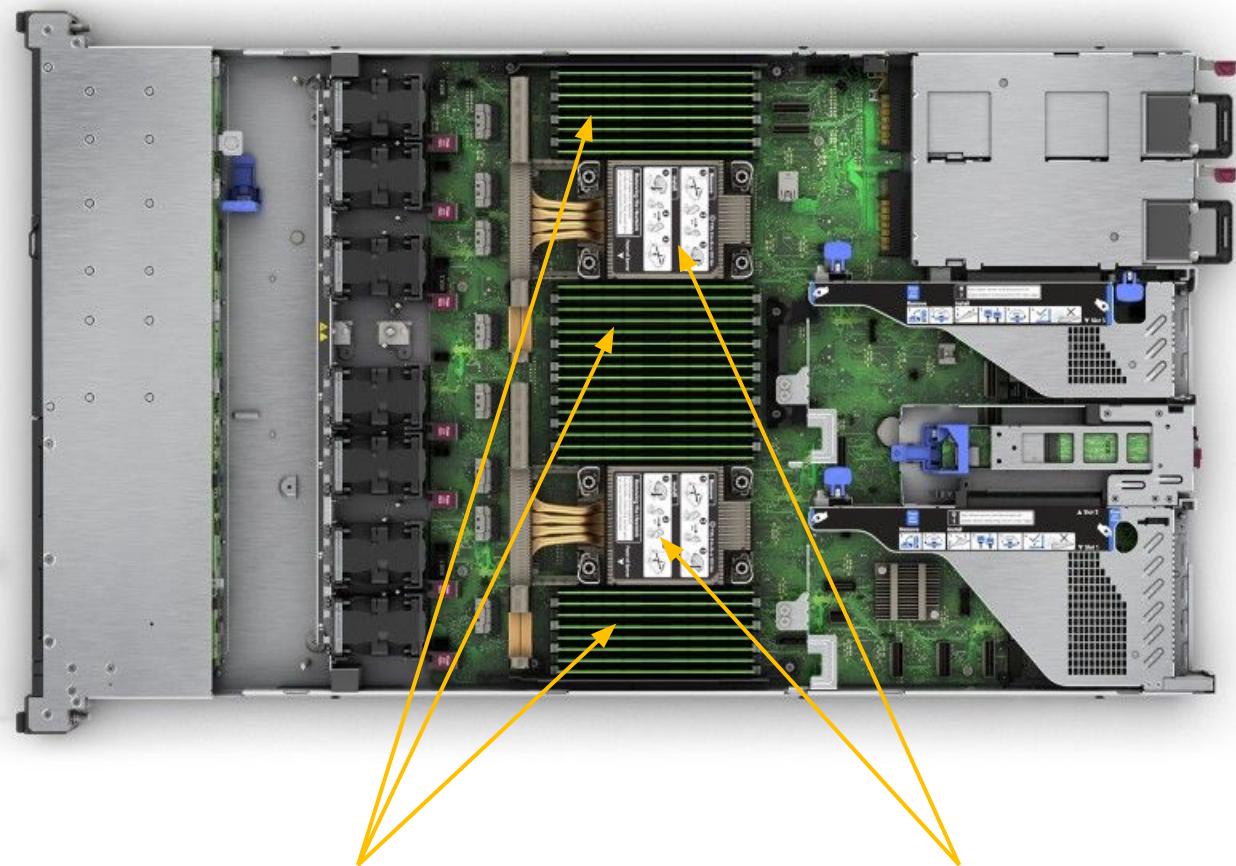


ordinateurs de calcul

Baies de stockage

Du data center au cœur

Un ordinateur ou **noeud** de calcul



Mémoires vives / RAM

Supports processeurs

Du data center au cœur

Un microprocesseur

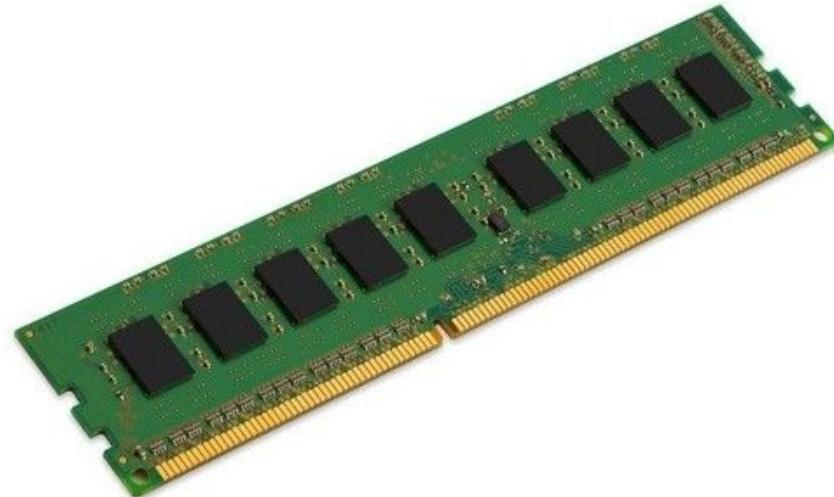


Un microprocesseur contient plusieurs **coeurs**
Chaque cœur se comporte comme un microprocesseur unique.

(nos nodes peuvent proposés jusqu'à 256 coeurs)

Du data center au cœur

La mémoire vive RAM



(nos nodes peuvent proposés jusqu'à 2 TB de RAM)

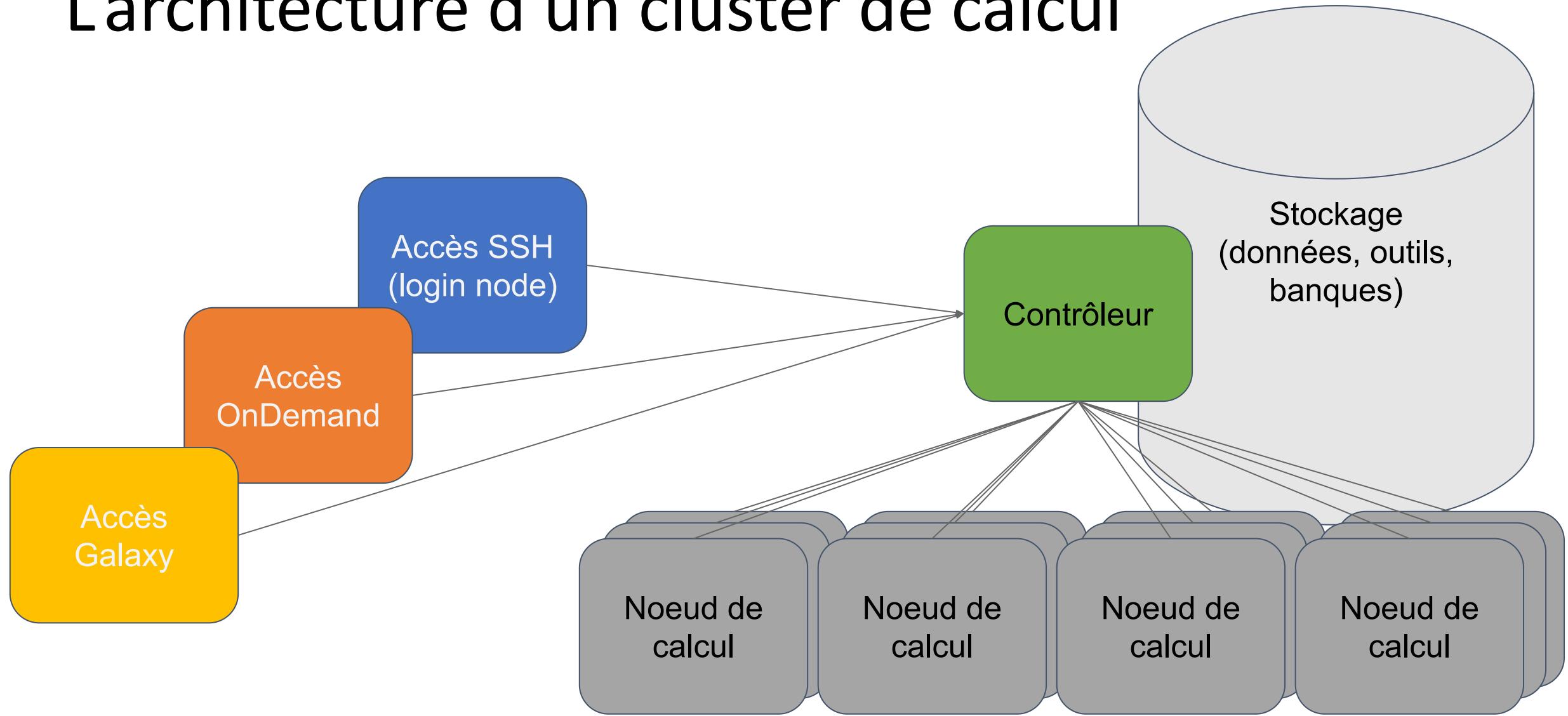
Du data center au cœur

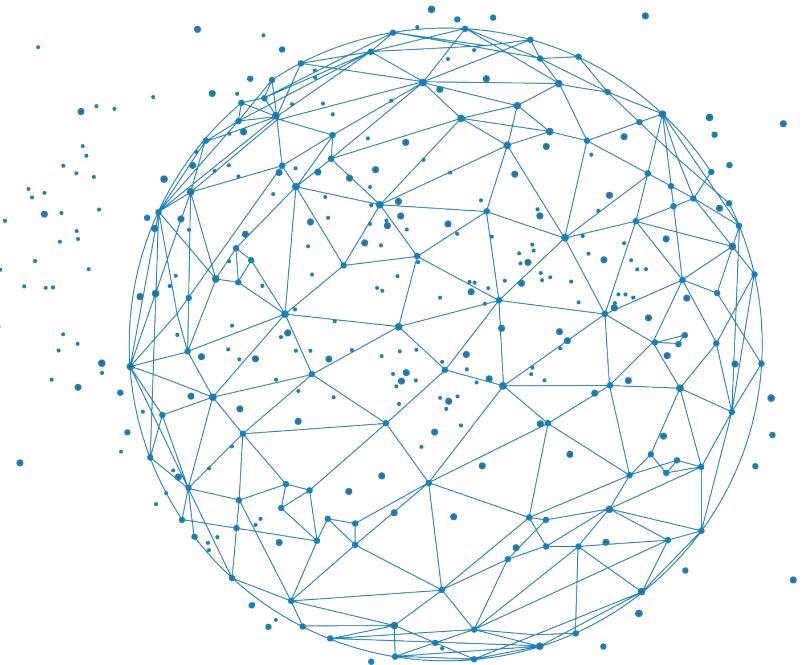
Carte graphique / GPU



(nos nodes peuvent proposés jusqu'à 2 cartes GPU A100 ou 4 L40s)

L'architecture d'un cluster de calcul





Institut Français de Bioinformatique (IFB) NNCR et IFB Core Cluster

Gildas Le Corguillé
Julien Seiler

Le National Network of Computing Resources

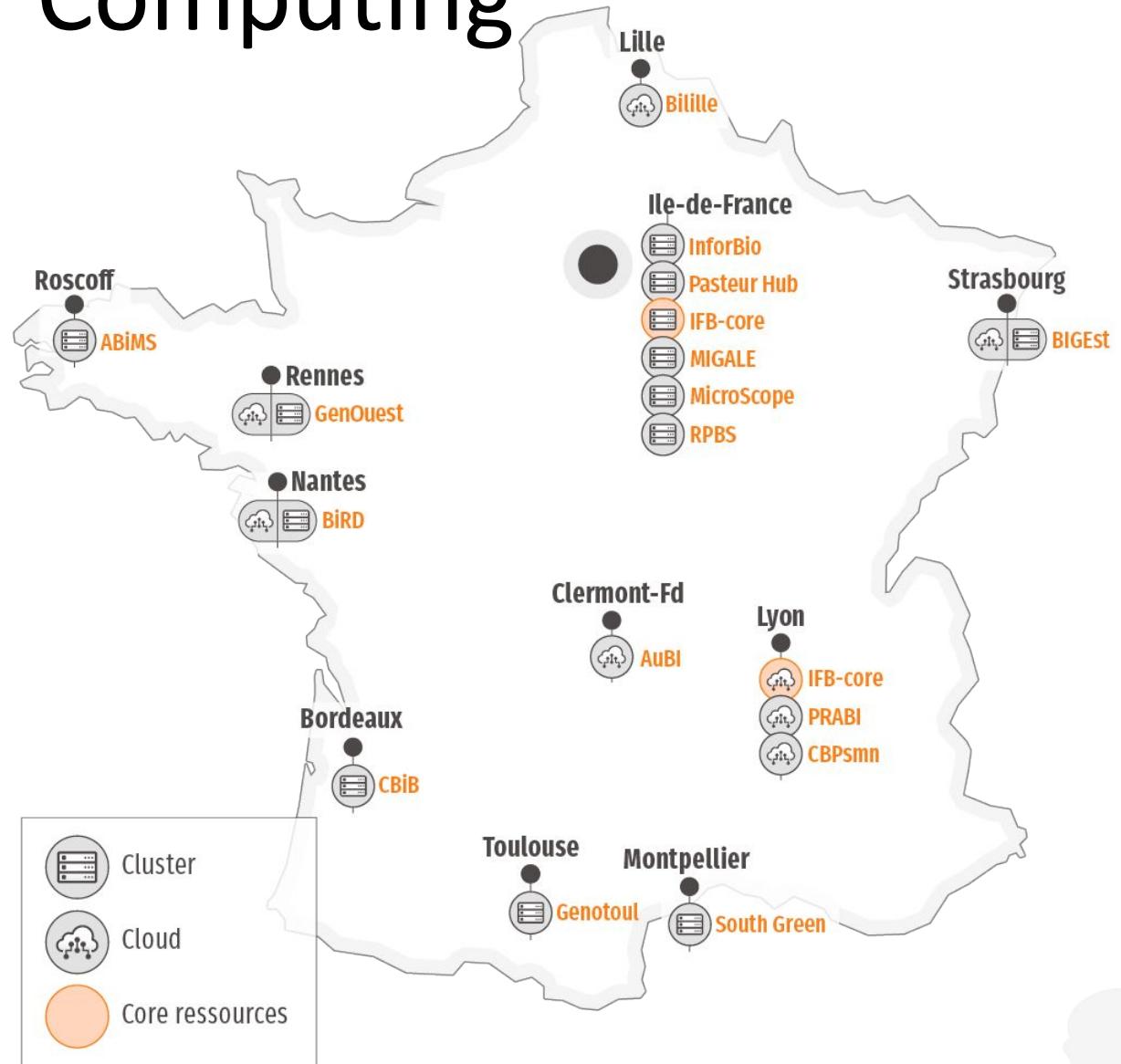
2 infrastructures complémentaires : Cloud et Cluster

Réseau national composé :

- D'équipements de plateformes régionales IFB
- De 2 infrastructures Core
(IFB core cluster + IFB core cloud)

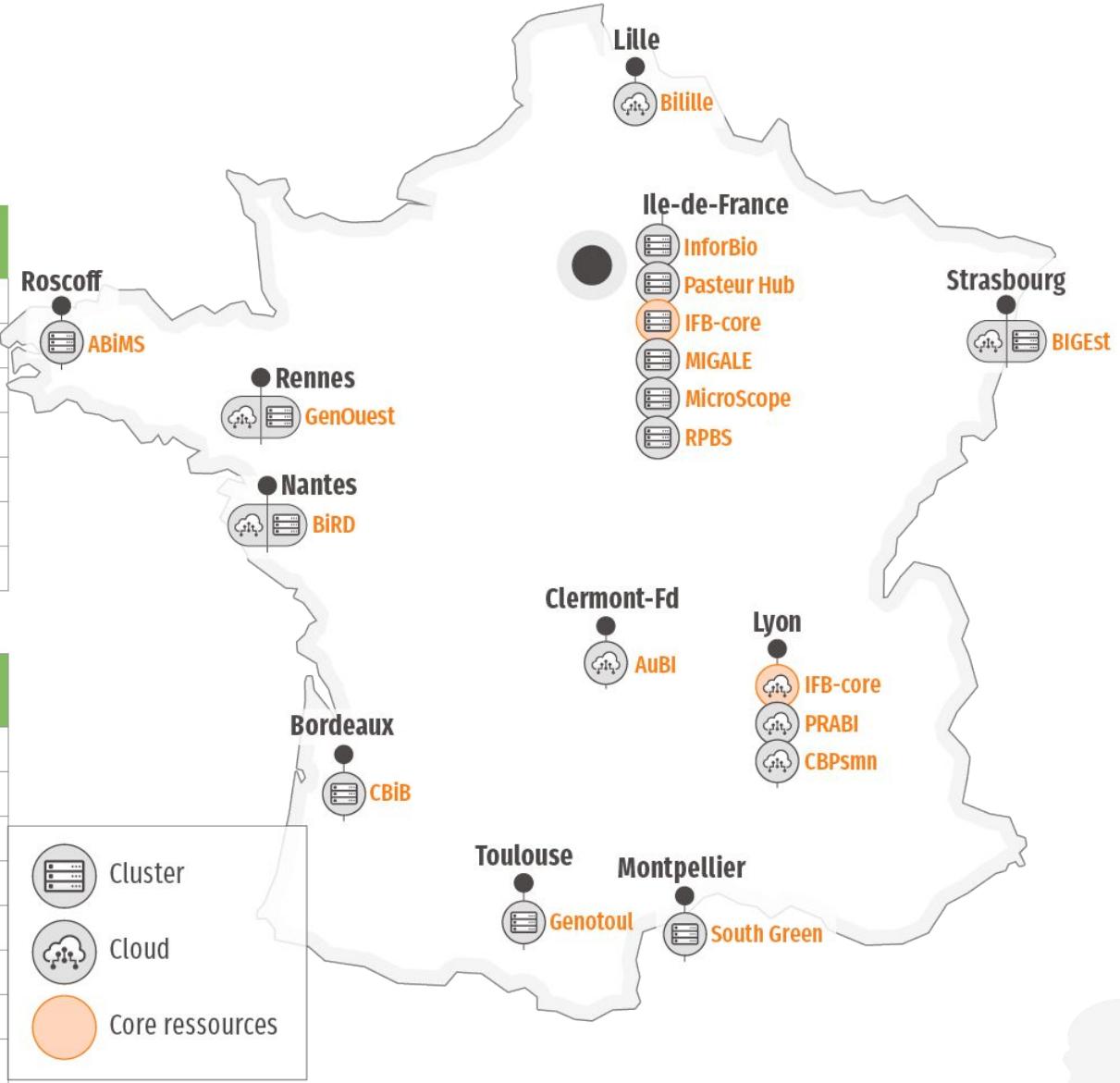
Administré par une “task force” mutualisée

- Les plateformes partage X% d'ETP d'ingénieurs
- Bénéfice mutuel des % d'ETP partagés
 - Forte émulation du personnel
 - Éviter la déperdition d'énergie
 - Partager l'expérience
 - Partager les développements
 - Traitement collaboratif des problèmes
 - Solutions robustes



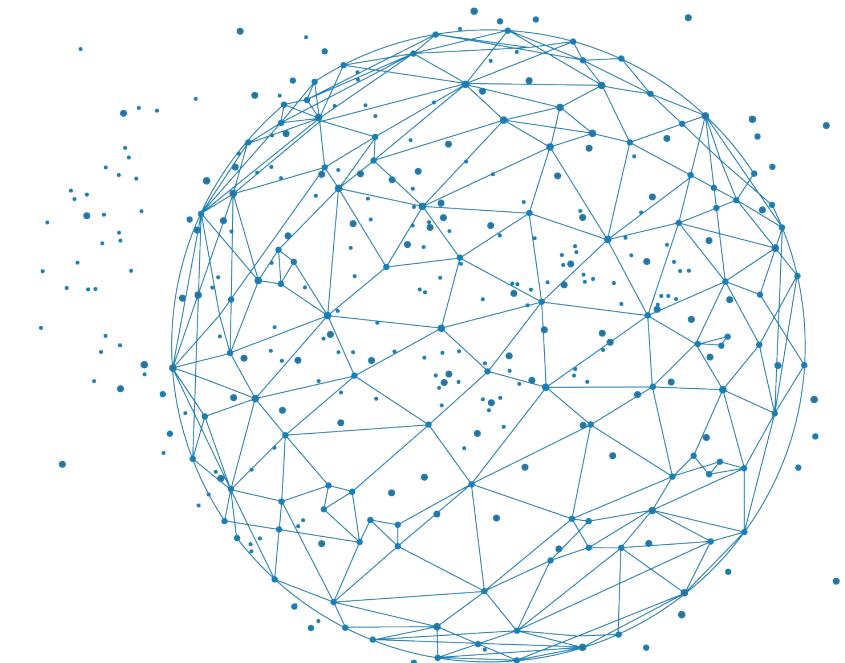
Le NNCR

| Platform | Location | Compute (#CPU-core HT) | Storage (#TB) | RAM (#GB) | GPU (#Cards) |
|---------------------------|---------------|---------------------------|------------------|----------------|-----------------|
| ifb-core-cluster | Orsay | 8 396 | 2 400 | 52 008 | 9 |
| ABiMS | Roscoff | 3 184 | 2 500 | 15 075 | 2 |
| GENOTOUL | Toulouse | 6 224 | 8 000 | 36 500 | 1 |
| GenOuest | Rennes | 1 866 | 2 300 | 11 616 | 7 |
| BiRD | Nantes | 902 | 600 | 6800 | - |
| MIGALE | Jouy en Josas | 1 016 | 350 | 7000 | - |
| Total IFB clusters | | 21 588 | 16 150 | 128 999 | 19 |





IFB Core Cluster



Le Core Cluster - Modalités d'accès

Utilisateurs / Equipes

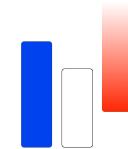
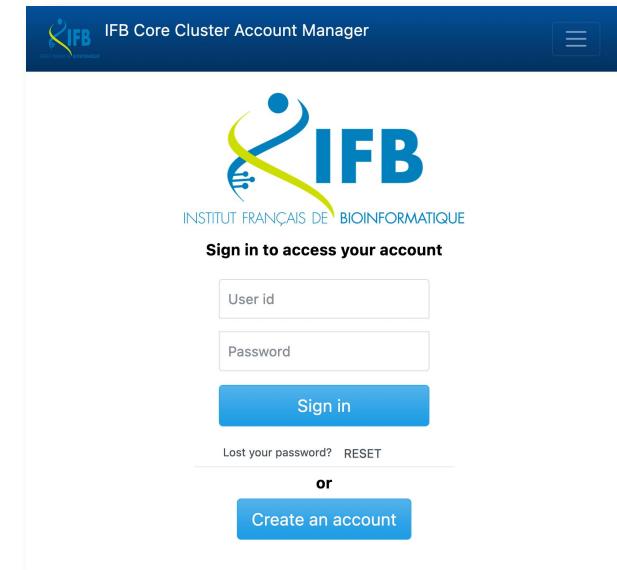
- Demander un compte nominatif gratuitement
- Demander un projet partagé gratuitement
 - 250 Go d'espace de stockage
 - Extensible à la demande

Formateurs

Nous pouvons héberger vos formations :

- Réservation de ressources de calcul
- Compte étudiants générique ou nominatifs
- Possibilité de demander l'installation d'outils

my.cluster.france-bioinformatique.fr



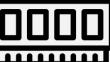
usegalaxy.fr
auto-inscription

Le Core Cluster - Le matériel

Actuellement

 85 nœuds

 4528 cores

 60 To RAM

 2 Po stockage

 6 cartes GPU
Nvidia A100



Le Core Cluster - Les services

The collage consists of five screenshots:

- Top Left:** A screenshot of the Galaxy France interface (<https://usegalaxy.fr>) showing a "Welcome to usegalaxy.fr" page with various bioinformatics tools listed under "Domain specific subdomains".
- Top Right:** A screenshot of an RStudio session window titled "R Studio OnDemand". It shows an R script running, producing a topographic map of Auckland's volcanic peaks, and a grid of scatter plots for Edgar Anderson's Iris Data.
- Middle Left:** A terminal window titled "glecorguille@clust-slurm-client:~" showing the output of module loading commands for fastQC, followed by a "fastQC" command and its version information.
- Middle Right:** A screenshot of a Jupyter Notebook titled "tutorial_slurm.ipynb". It shows a terminal session where a script for aligning data to the Arabidopsis thaliana genome is run, along with the resulting log output.
- Bottom Center:** Logos for several bioinformatics tools: AlphaFold (blue circular logo), TensorFlow (orange T logo), snakemake (green S logo), and nextflow (green and orange logo).



AlphaFold



TensorFlow



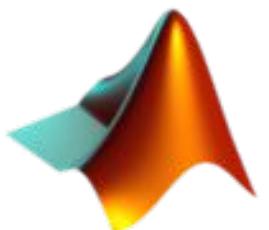
snakemake



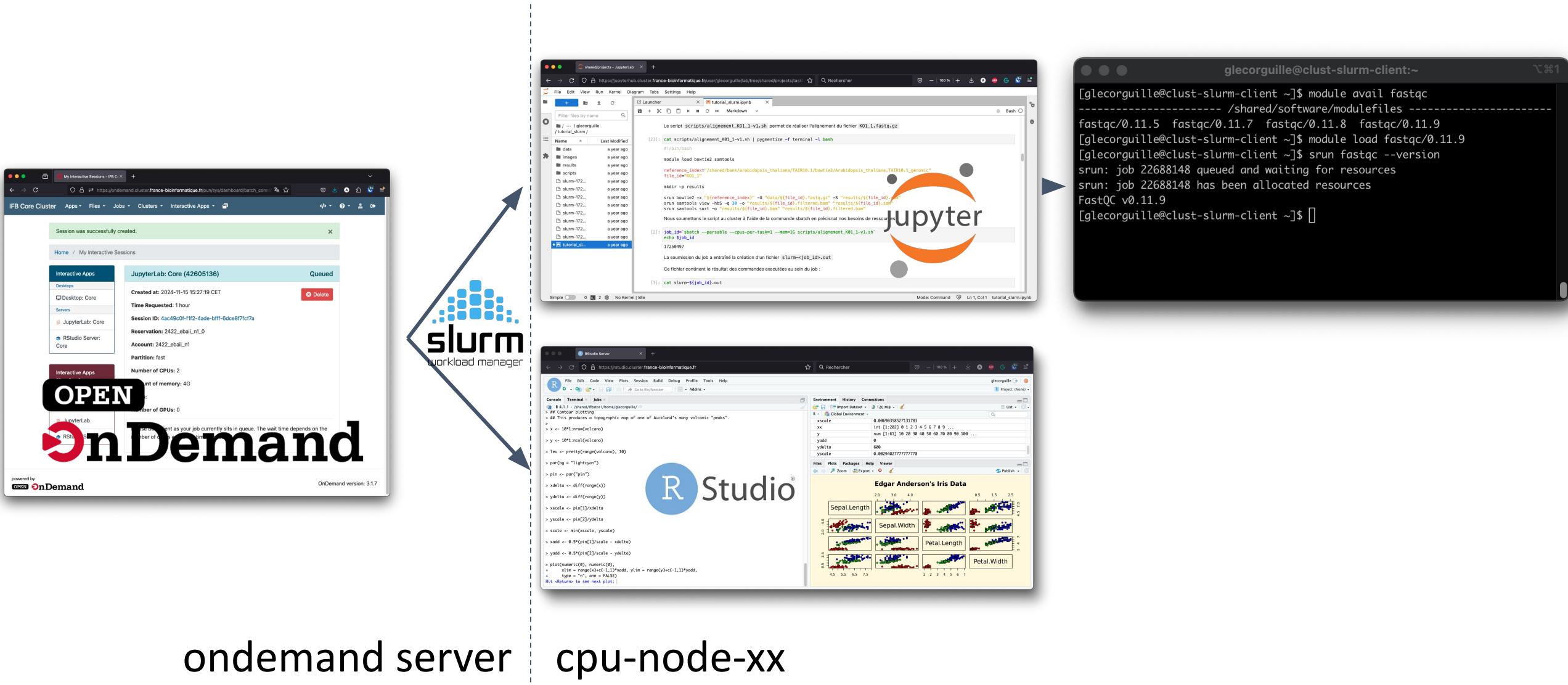
Open OnDemand

Un portail web pour l'utilisation des ressources de calcul

- Console de gestion SLURM en ligne
- Lancement de jobs et scripting assisté
- Accès web aux espaces de stockage
- Intégration de JupyterLab, RStudio, VSCode
- Intégration de RTrainer
- Accès environnement graphique (XFCE)



Open OnDemand



Open OnDemand

Utilisez votre compte IFB Core Cluster pour vous connecter

Lancer un serveur JupyterLab ou RStudio dans un job SLURM

Accédez aux mêmes données et outils que sur le cluster (ssh)



Amount of memory : 16G

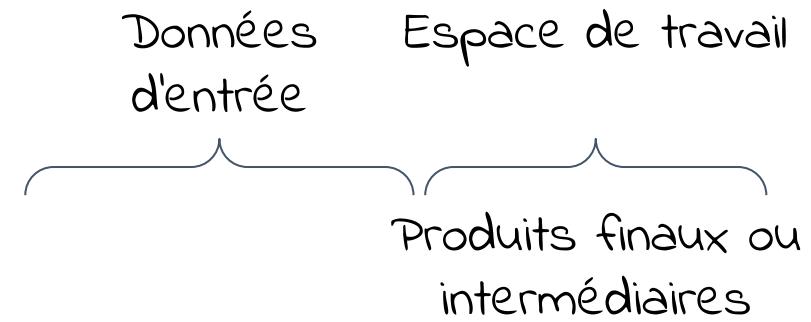


$16 = 16M$

The screenshot shows the 'Dashboard - IFB Core Cluster' page. At the top, there's a navigation bar with links for 'IFB Core Cluster', 'Apps', 'Files', 'Jobs', 'Clusters', 'Interactive Apps', and a user icon. Below the navigation, the 'OnDemand' logo is prominently displayed with the text 'OnDemand provides an integrated, single access point for all of your HPC resources.' A section titled 'Pinned Apps' features two cards: 'JupyterLab: Core System Installed App' and 'RStudio Server: Core System Installed App'. At the bottom of the dashboard, it says 'powered by OPEN OnDemand' and 'OnDemand version: 3.0.1'.

<https://ondemand.cluster.france-bioinformatique.fr>

Réservation de la RAM ?



Réservation de la RAM ?



Tools

Where is my tools?



Software environment



module load falco/1.2.5

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

744 tools
1237 versions
444 packages R
8 infrastructures

Besoin d'aide sur l'usage d'un outil scientifique ou sur un protocol ?

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

Rendez-vous sur :

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

The screenshot shows the homepage of the IFB Community Support forum. At the top, there is a navigation bar with the IFB logo, a search icon, and a user profile icon. Below the navigation bar, there are buttons for "all categories" and "all tags". A red "Categories" button is highlighted. Other buttons include "Latest", "New (5)", "Unread (22)", and "Top". There is also a "+ New Topic" button. The main content area is a table listing categories. The first category, "IFB Core Cluster", has 406 topics and includes posts about software requests and support. The second category, "Workflow4Metabolomics", has 46 topics and is described as being dedicated to questions about workflow4metabolomics.usegalaxy.fr. The third category, "Support scientifique et ...", has 30 topics and is described as providing help on scientific themes, workflows, software, or parameters.

| Category | Topics | Latest |
|------------------------------------|--------|--|
| IFB Core Cluster | 406 | Meilleure manière de... 1h Demande de logiciel ou de banque de données 19 unread Support IFB Core Cluster 7 unread 1 new |
| Workflow4Metabolomics | 46 | Troubleshooting - W... Sep '20 Troubleshooting - G... May '20 Erreur Python lors de l'e... 8h |
| Support scientifique et ... | 30 | Gatk spark beta • 11h Comment optimiser les r... 7d Utilisation gpu via sbatch 18d |