

Exploitation bio info et calcul au Mésocentre Clermont Auvergne

Nadia Goué
David Grimbichler

Le Mésocentre, techniquement ...

- **Calcul**

- 43 nœuds de 128~256 Go RAM + 16 ~ 48 processeurs
- 1 SMP de 12 to RAM + 384 processeurs
- 4 nœuds avec GPU (CUDA), au total 13 GPU (2 P40, 8 V100, 3 P100)

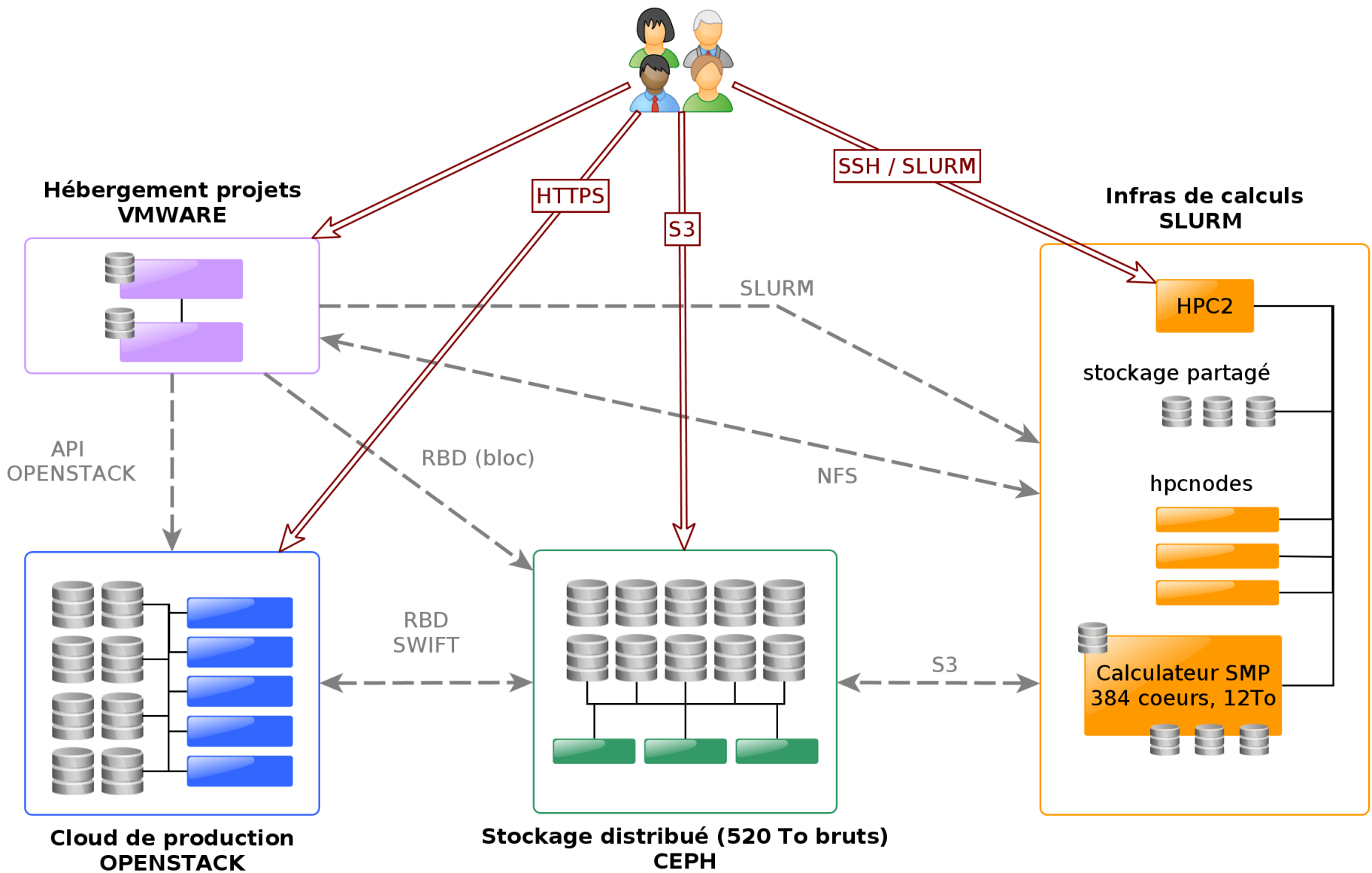
- **Stockage**

- 20 To pour les comptes cluster
- 500 To + 800 To en stockage capacitif
- Stockages spécifiques labos (20 To + 8 To + 16 To)

- **Cloud IaaS**

- 48 nœuds de 40 processeurs + 128~256 Go RAM
- 10 nœuds contrôleurs, réseau 10G
- Pour tout ce qui n'est pas adapté au calcul sur cluster

+ des gens et des compétences ☺



Actions du Mésocentre (1)

Leitmotiv : facilitateur pour les labos et les chercheurs

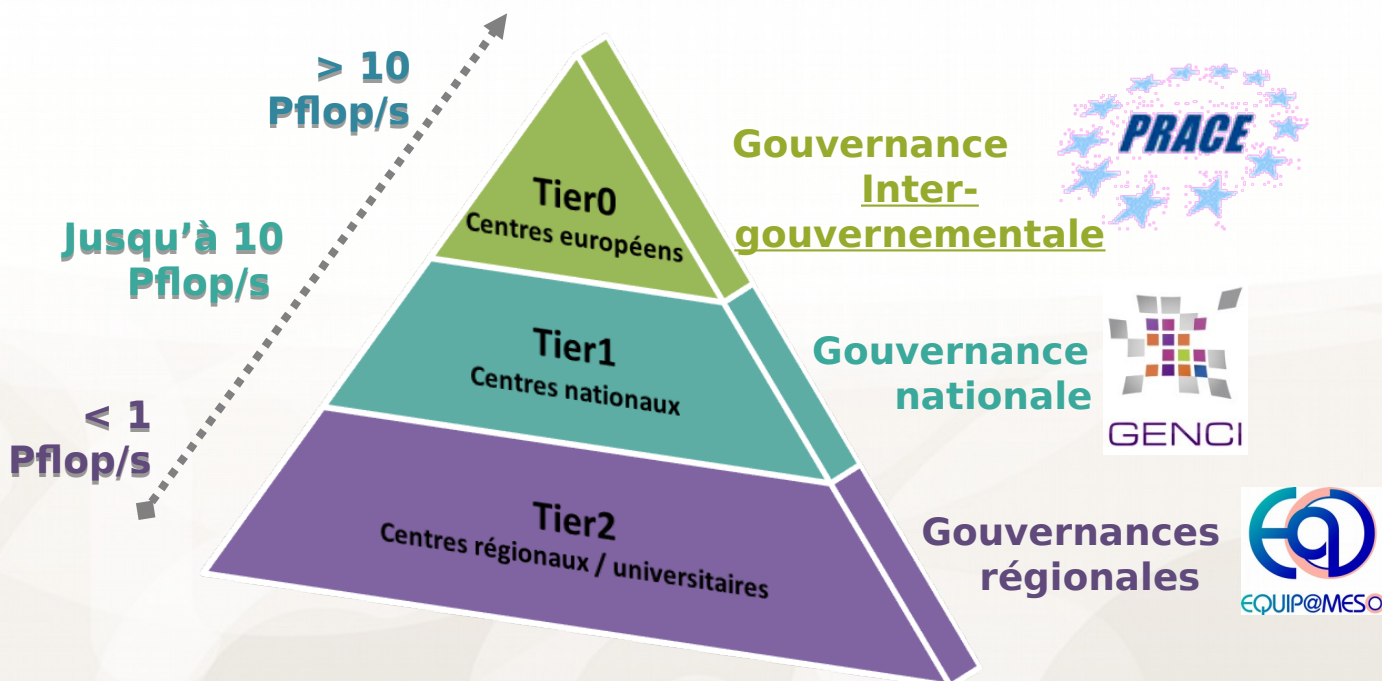
- **Développements :**
 - Plugin Matlab / S3 (Institut Pascal / Sigma)
 - Scripts et exemples pour l'utilisation des ressources
- **Compilations spécifiques, patches d'outils et de codes :**
 - getfem, scotch, metis, mumps, qhull (LMV)
 - MEEP (IP), Beast/Beagle (LMGE), ...
- **Solutions spécifiques :**
 - R Shiny démo (Cerdi), phrogs (LMGE), DefVolc (LMV), ...
 - Ceba (Cap2025) / ConnecSens (CPER Audace)
- **Utilisation d'outils propriétaires :**
 - Comsol (LMV), Matlab (tout le monde), Casino (IP), CUDA (CPER Audace), Photoscan (OPGC)...

Actions du Mésocentre (2)

Leitmotiv : facilitateur pour les labos et les chercheurs

Campagnes sur les ressources de calcul des centres nationaux :

- Casino (IP), Gromacs (PIAF)



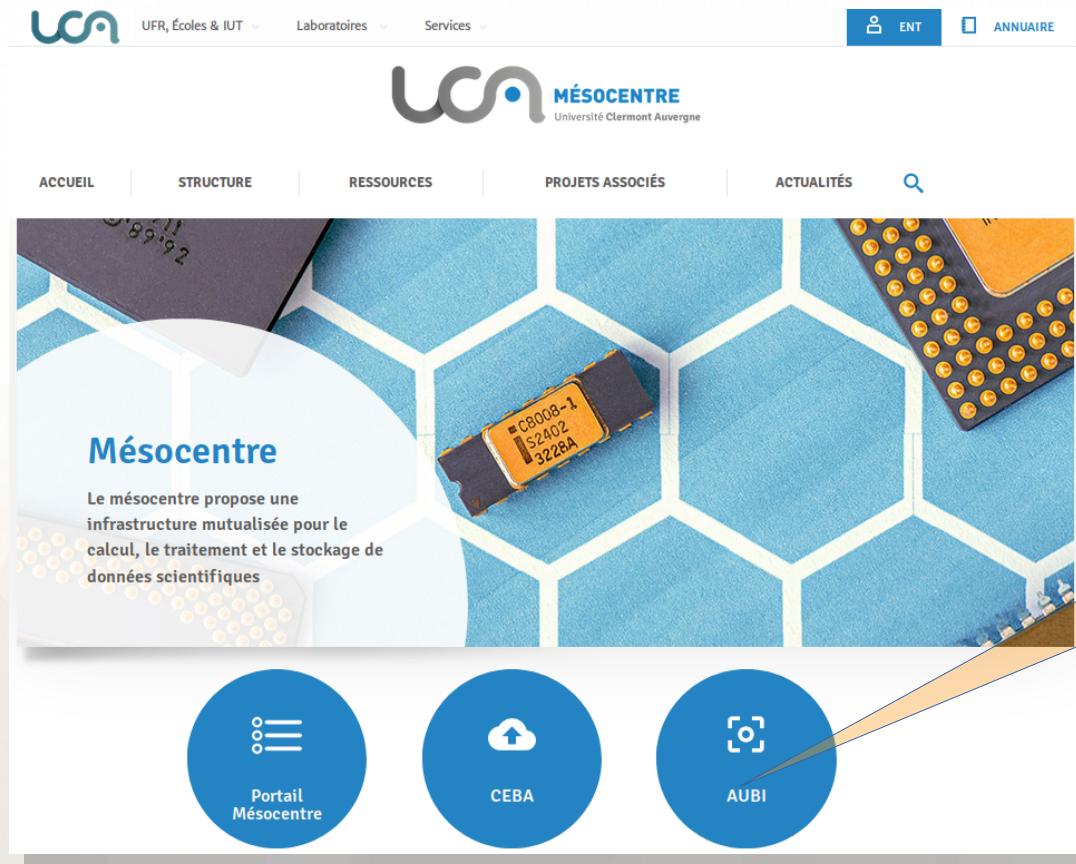
Actions du Mésocentre (3)

Leitmotiv : facilitateur pour les labos et les chercheurs

- **Accompagnement, analyse pour**
 - Utilisation optimale des GPU
 - Meilleure utilisation des ressources calcul :
 - Surveillance de certains jobs
 - Conseil aux utilisateurs
 - Profiling si nécessaire
- **Formations sur l'utilisation des ressources**
- **Cloud :**
 - Inculquer l'esprit des VM éphémères
 - Proposer des images pour les demandes spécifiques :
 - Smrtlink (LMGE), Triannot (GDEC), R (Infosol / INRA), ...
 - Accompagner les utilisateurs
 - Augmenter les capacités de calcul

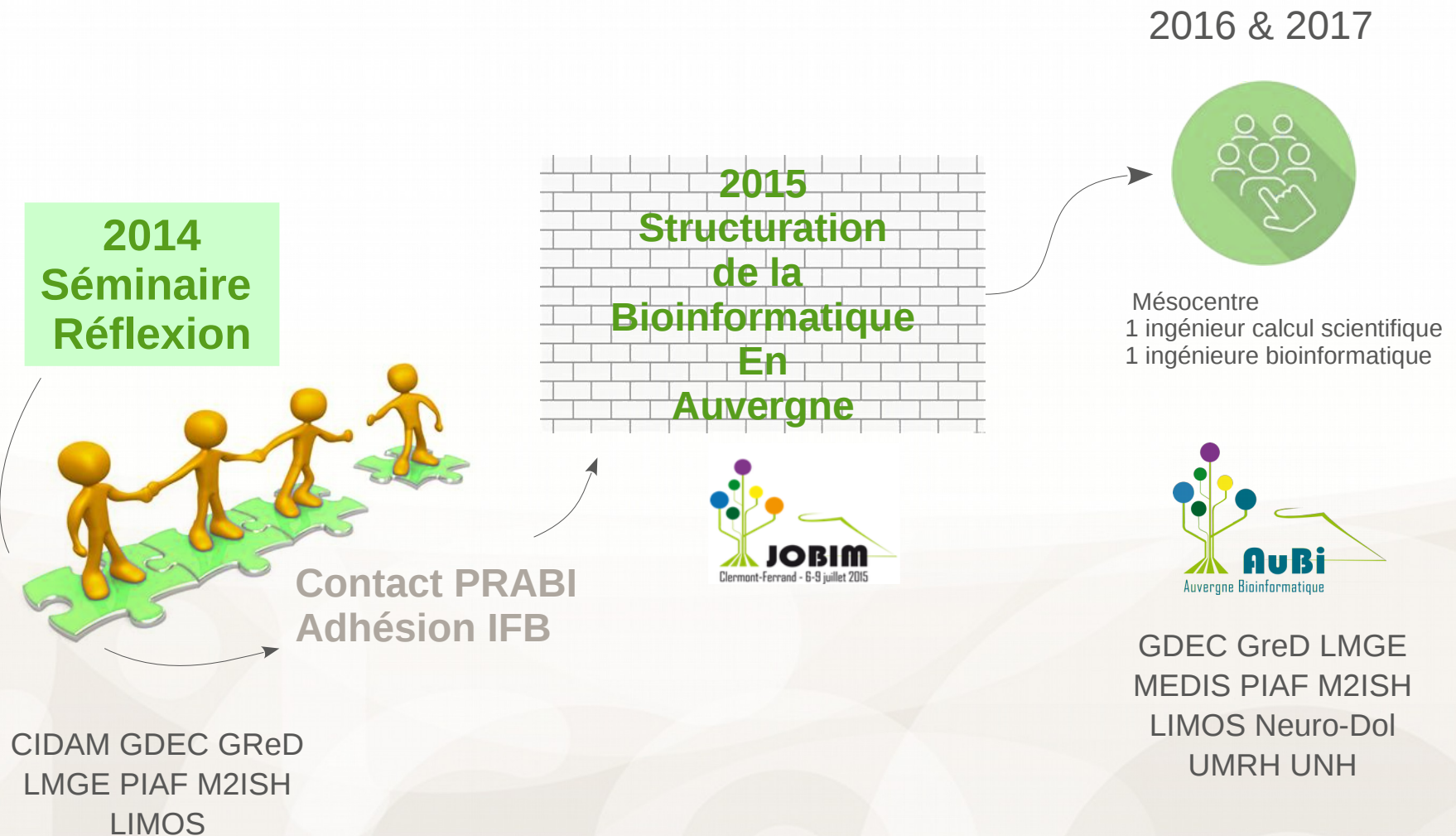
Service mutualisé de l'UCA & Associés

mesocentre.uca.fr

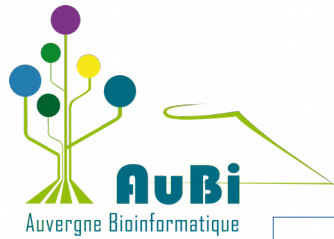


news.aubi@listes.uca.fr

Genèse de la plateforme AuBi



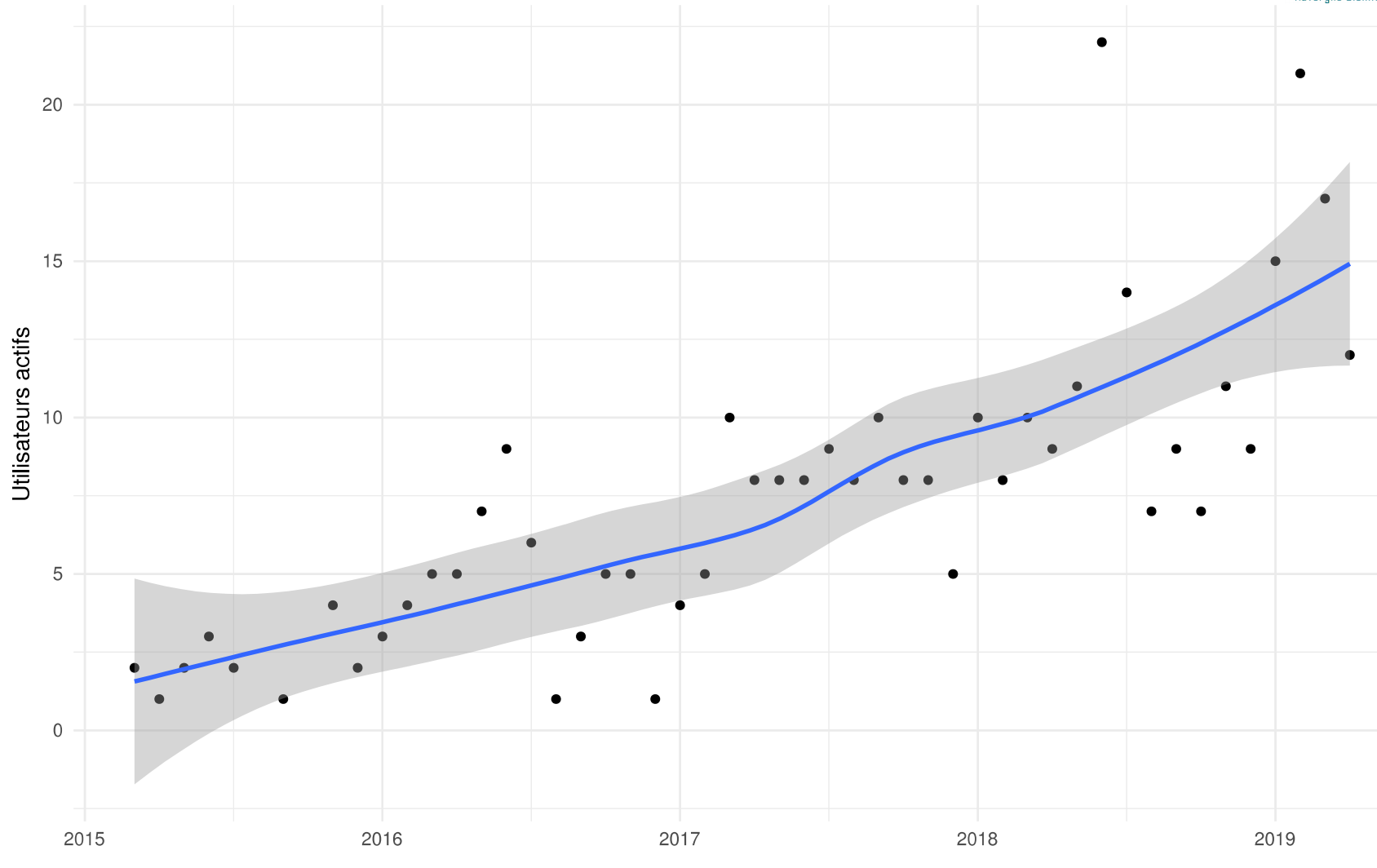
AuBi dans son réseau bioinfo



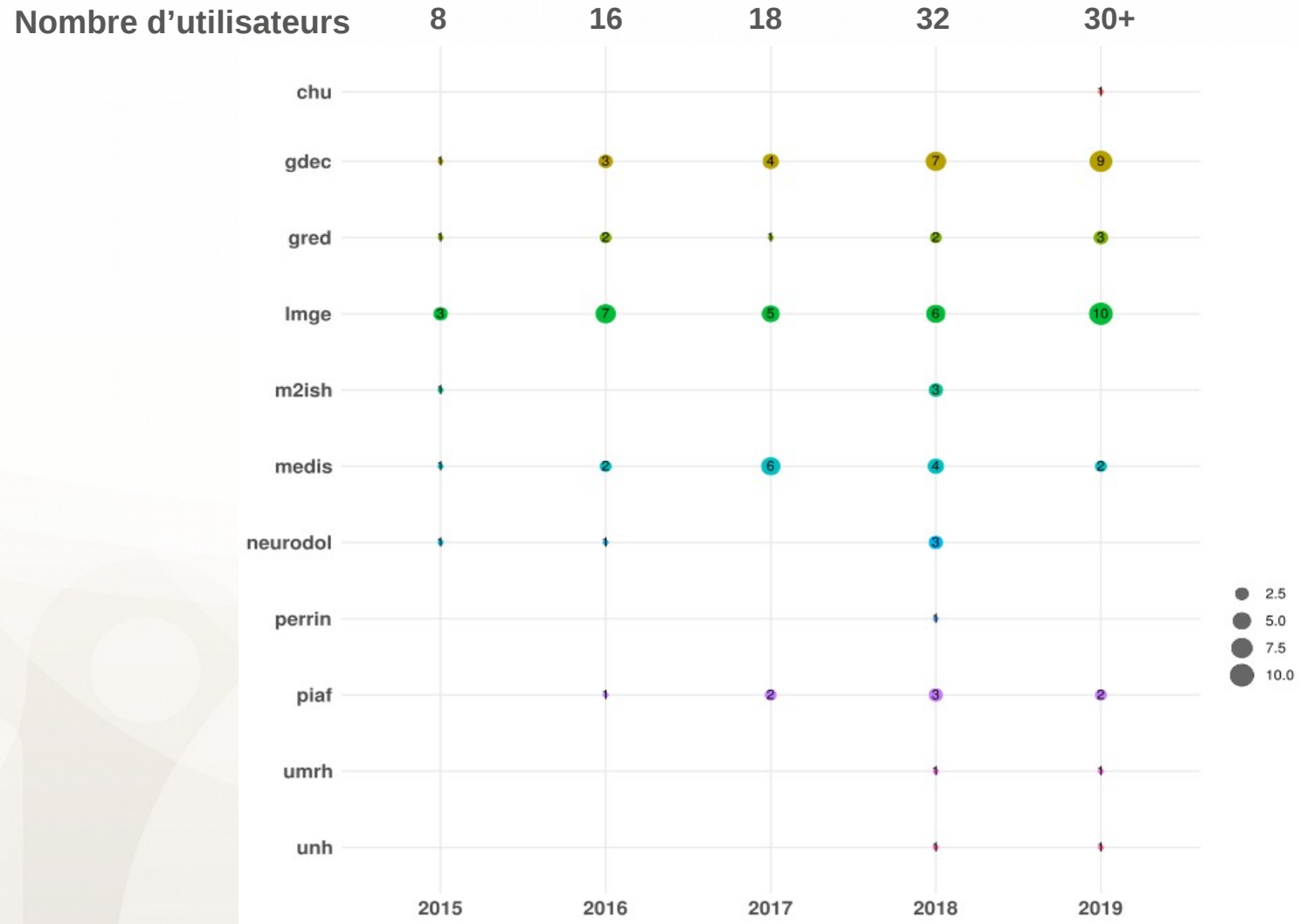
Quels sont les besoins en bioinformatique ?

- **Calcul**
 - CPU
 - GPU
 - Stockage
 - réseau
- **Accès aux collections de données publiques**
 - NCBI
 - UCSC
 -
- **Formation et accompagnement des utilisateurs**
 - Unix
 - Cluster & NGS
- **Mise à disposition de VM**
 - Galaxy
 - VM à la demande

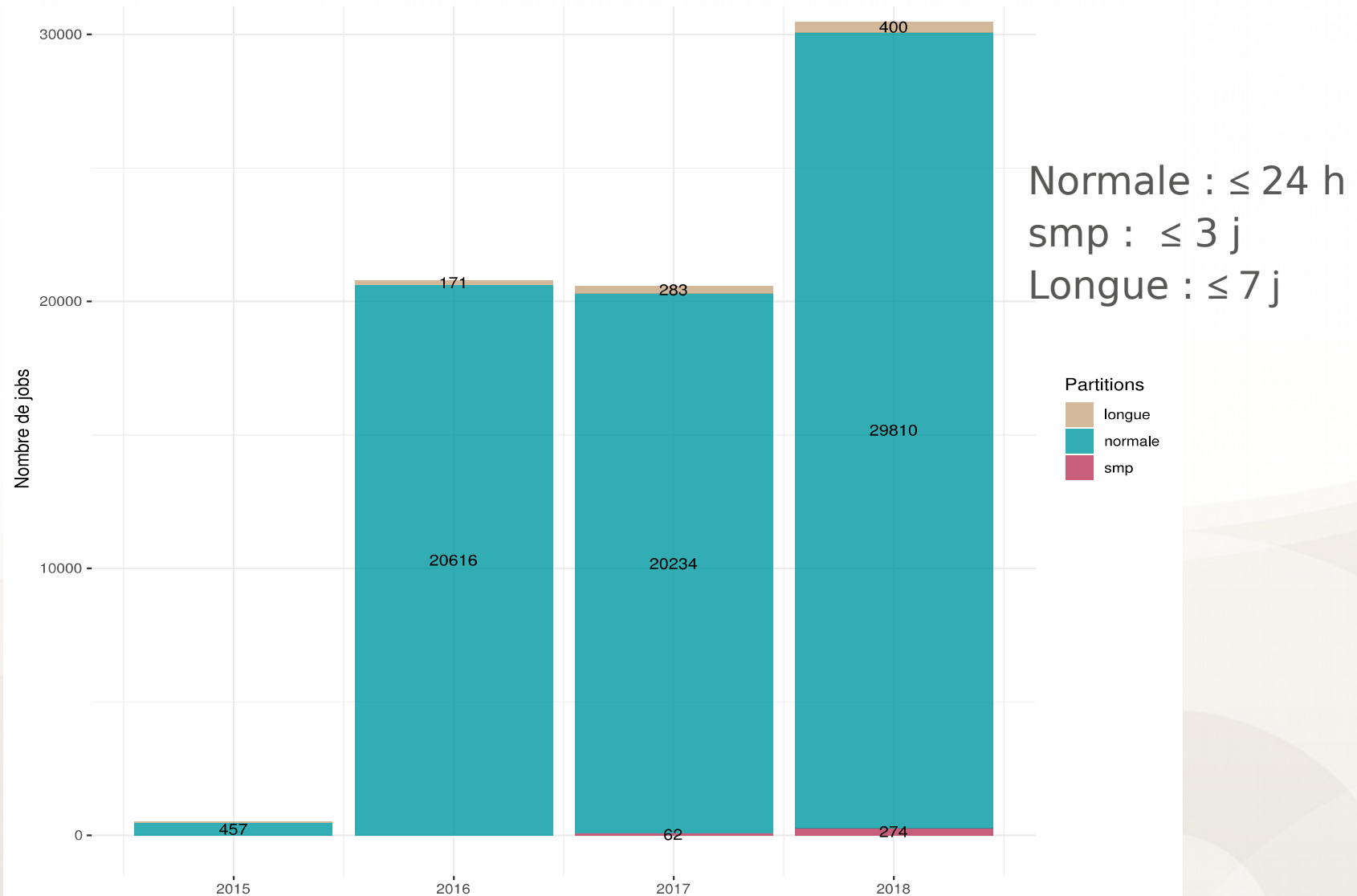
86 comptes ouverts



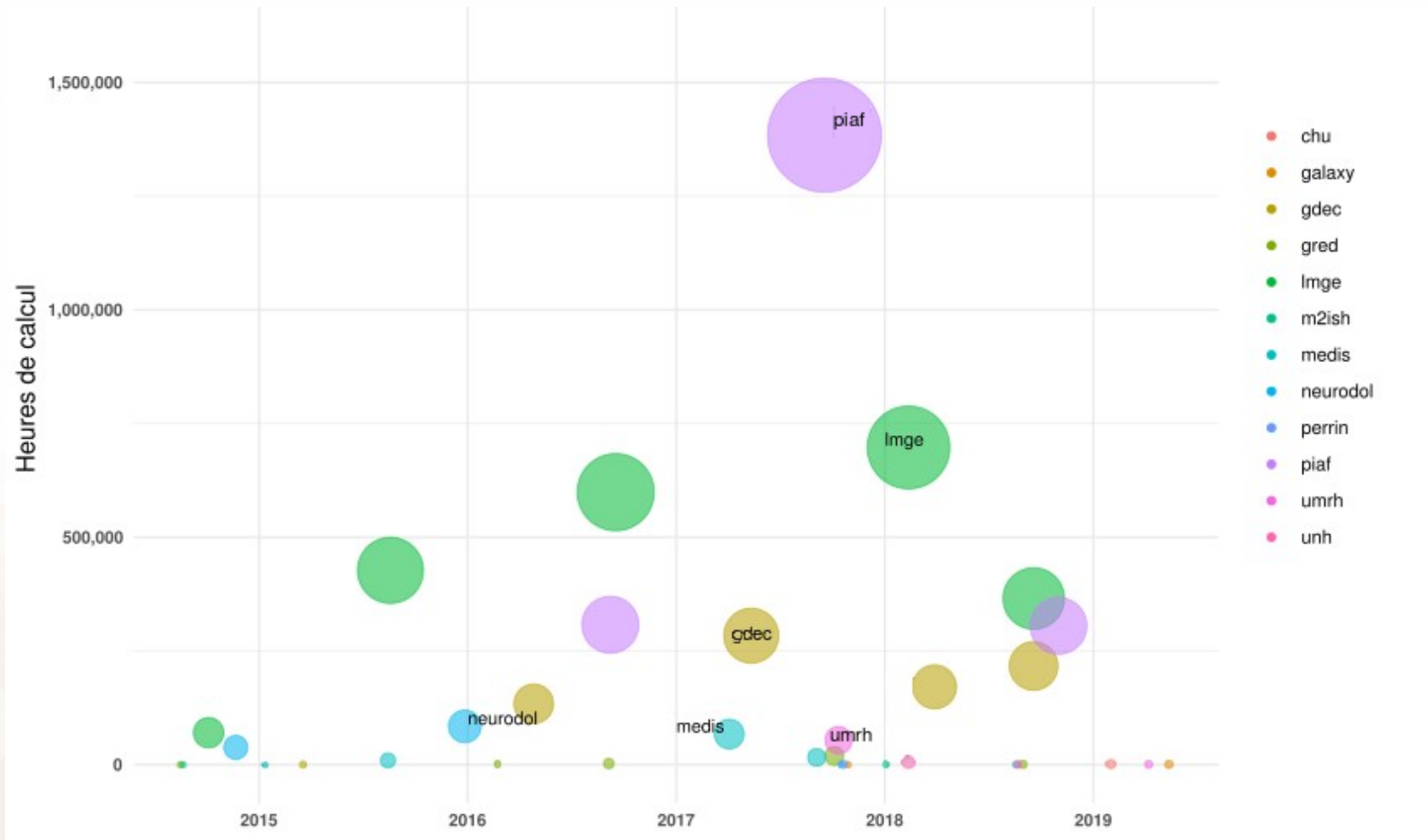
Qui sont les utilisateurs communauté AuBi ?



Evolution des demandes en calcul



Evolution des demandes en calcul



Projet Structure-fonction aquaporines du peuplier

Biologie in silico : Comment les aquaporines régulent-elles le passage de l'eau au travers des membranes biologiques ?



Besoins

Champ de forces tous atomes

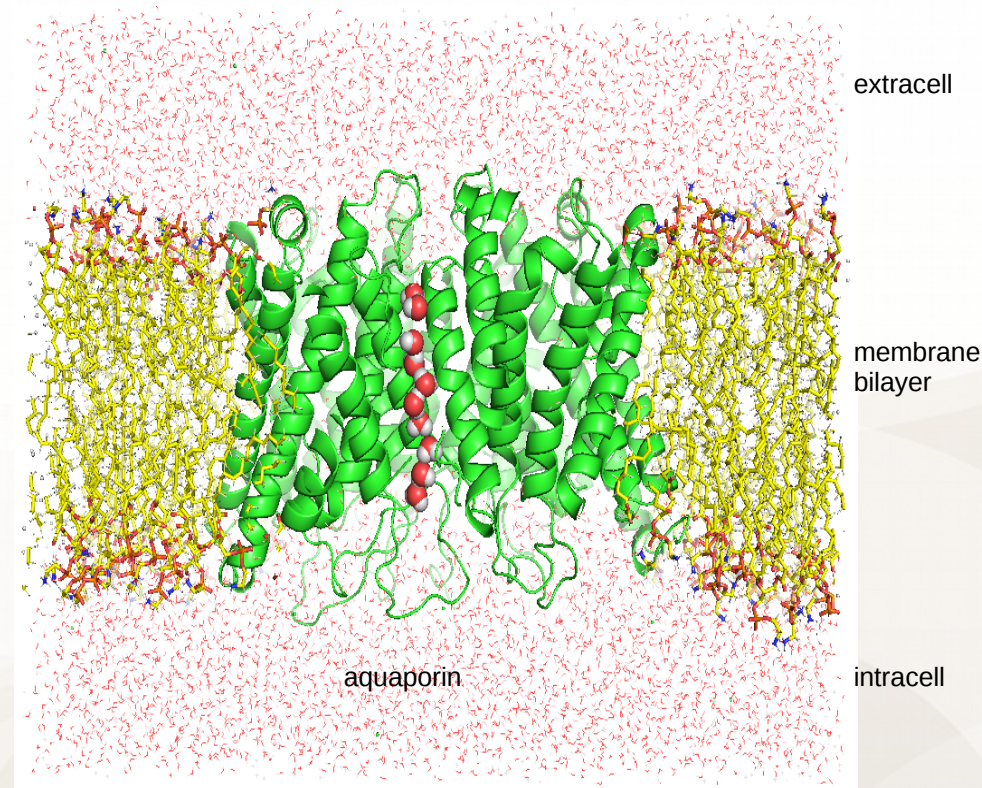
(~250 000 atomes)² interactions

50 000 000 itérations pour 100 ns

Moyens

2x GPU Tesla P100 + 32 CPU

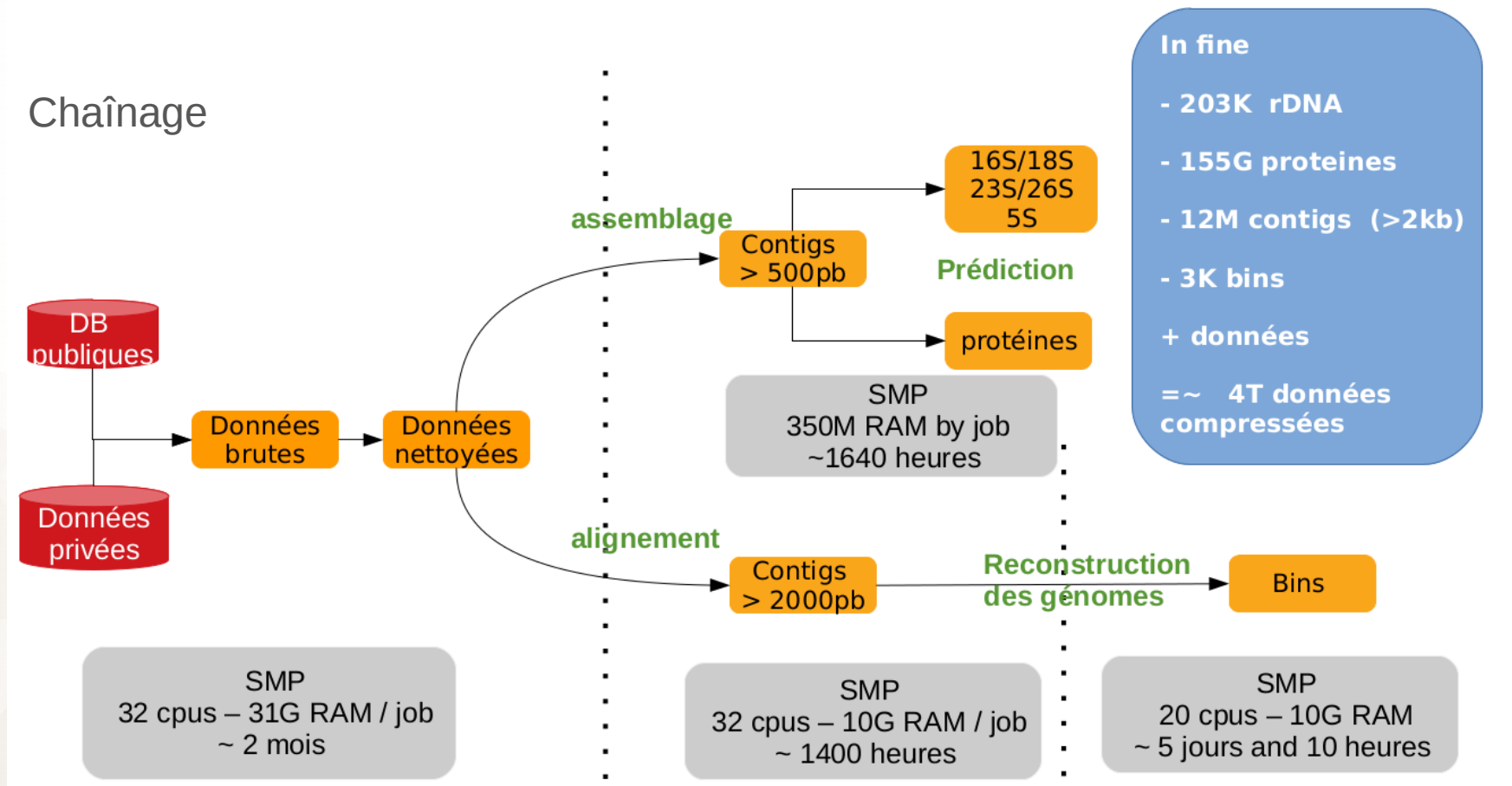
2 jours.calcul⁻¹ *wall clock time*



Projet MetaFreshWater

Méta-analyse de 496 métagénomiques d'eaux douces

Chaînage



Stockage CEPH

58 To de données
Triplicats (174 To d'espace consommé)
38 comptes utilisateurs Mésocentre

mode PRIVÉ

3 GDEC
2 PIAF
1 LMGE



mode PROJET

Aquaporine (3 To, PIAF)
Eureka (5 To, LMGE)
GDEC (2 To)
GReD
Microstore (3 To, LMGE)
Neuro-Dol (15 To)
NoyauHD (5 To, GReD)

Exemple d'utilisation des ressources en stockage



Espace de stockage centralisé

634 Gb utilisé

Accroître la volumétrie
utilisable sur le cluster
stockage temporaire

Sécurisation des données



Partager des données
multi-utilisateurs

Limiter la duplication
de fichiers volumineux

Meilleure reproductibilité

Service web Galaxy

galaxy.mesocentre.uca.fr

Galaxy / AuBi

Analyze Data Workflow Visualize Shared Data Admin Help User

Using 21%

Tools

search tools

DATA MANAGER

Get Data

Send Data

Collection Operations

BASIC TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

Extract Features

Fetch Sequences

Fetch Alignments

NGS COMMON TOOLSETS

NGS: SAM Tools

NGS: BED Tools

NGS: Picard Tools

NGS: GATK Tools

NGS: DeepTools

NGS ANALYSIS

NGS: QC and manipulation

NGS: Mapping

NGS: RNA Analysis

NGS: small RNA analysis

NGS: Peak Calling

NGS: Assembling

ANNOTATION, ONTOLOGIES

Motif Tools

METAGENOMIC ANALYSIS

FROGS

MOTHUR

UCA
MÉSOCENTRE
UNIVERSITÉ
Clermont Auvergne

Hello, AuBi Galaxy is running on Mesocentre Clermont Auvergne!
Please send an email to support.dsi@uca.fr for any feedback and questions.

Take an interactive tour: [Galaxy UI](#) [History](#) [Scratchbook](#)

Warning!

All jobs running on AuBi Galaxy are sent to UCA Mesocentre cluster. Consequently, Mesocentre platform reserves the right to purge all files not accessed since 120 days. Data and results produced in this Galaxy instance are not saved and are under user responsibility. The use of softwares or banks availables on Galaxy must be done in respect for intellectual property.

For help downloading big size datasets: [tutorial SFTP](#)

Galaxy is an open platform for supporting data intensive research. Galaxy is developed by [The Galaxy Team](#) with the support of [many contributors](#).

The [Galaxy Project](#) is supported in part by [NHGRI](#), [NSF](#), [The Huck Institutes of the Life Sciences](#), [The Institute for CyberScience at Penn State](#), and [Johns Hopkins University](#).

History

search datasets

Unnamed history
52 shown, 66 deleted
268.4 MB

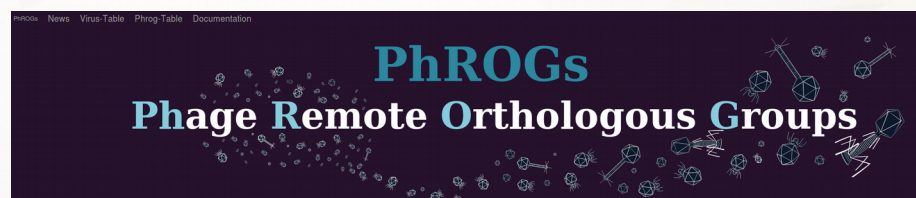
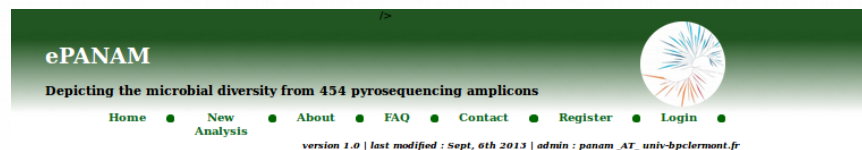
- 118: HISAT2 on data 12: aligned reads (BAM)
- 115: Bowtie2 on data 15 and data 16: mapping stats
- 114: Bowtie2 on data 15 and data 16: aligned reads (BAM)
- 113: Bowtie2 on data 15 and data 16: aligned reads (L)
- 112: Bowtie2 on data 15 and data 16: unaligned reads (L)
- 111: FROGS Affiliation OTU: report.html
- 110: FROGS Affiliation OTU: affiliation.biom
- 99: FROGS Affiliation postprocess: aggregation_composition.txt
- 98: FROGS Affiliation postprocess: sequences.fasta
- 97: FROGS Affiliation postprocess: affiliation.biom
- 93: FROGS Affiliation OTU: report.html



2 GDEC
1 GReD
2 LMGE
2 MEDIS
1 PIAF

Mise à disposition de VM

Outils d'analyses développés par les laboratoires



Merci de votre attention !

