

Durée / Programme : 3 journées

Présentation des plateformes et des équipes

Initiation à l'interface Galaxy

Administration et wrappers Galaxy

Présentation du cluster de calculs intensifs Genotoul

Sarah Maman
Déc. 2016

Présentation de la capacité de calcul du cluster

Connexion au serveur Genotoul

Utilisation de quelques commandes basiques

File system

Les outils de base (find, sort, cut, grep ...)

Transférer et compresser des fichiers

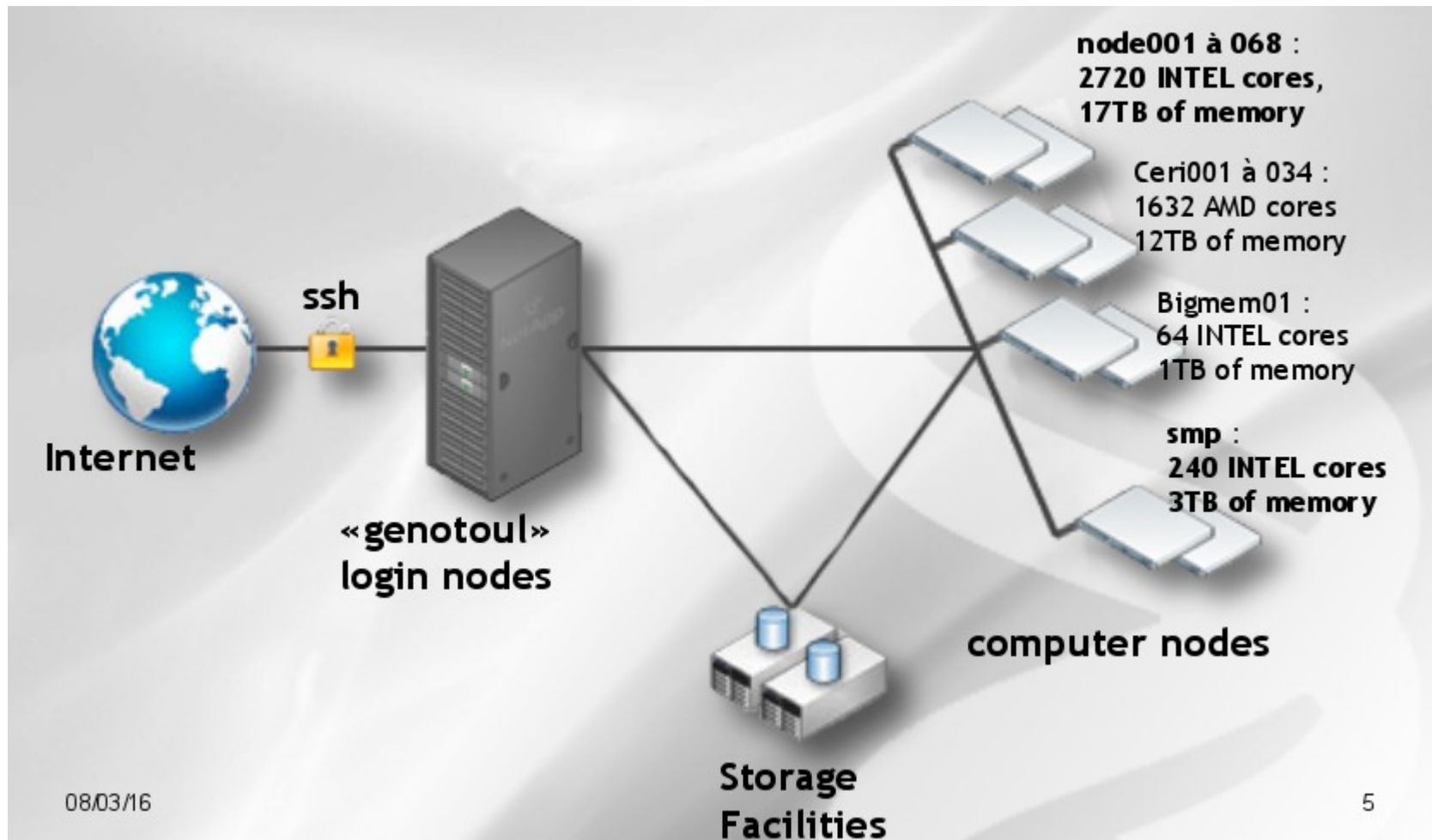
Soumettre un job et le gérer



Linux

Vous vous connecter en ssh sur les noeuds de Genotoul (login fleur).
 Si besoin, demande de compte : <http://bioinfo.genotoul.fr/index.php?id=81>

Vous avez ainsi accès à des ressources de stockage et de calculs.



From Windows

- **Xming** (Windows graphic)
- **Putty** (Connection)



From Linux / Mac

- **ssh username@genotoul.toulouse.inra.fr**
(command line)

```

Sarah@localhost:~
Fichier Édition Affichage Rechercher Terminal Aide
[Sarah@localhost ~]$ ssh -YX smaman@genotoul.toulouse.inra.fr
smaman@genotoul.toulouse.inra.fr's password:
Last login: Fri Oct 21 09:50:34 2016 from lan31-4-82-227-130-176.fbx.proxad.net
    
```

=====
 Bienvenue sur les serveurs d'accès de la plate-forme bioinformatique
 =====

Les serveurs "genotoul" ne sont PAS destinés à lancer des traitements de données mais uniquement à soumettre des jobs sur le cluster de calcul.
 Pour cela, utiliser les commandes "qsub" (batch), "qsh" (interactif) ou "qlogin" (redirection graphique).
 Tout traitement lancé sur les serveurs genotoul sera immédiatement interrompu par les administrateurs systèmes.

=====
 Information concernant les espaces disques
 =====

La **partition /home** est réservée aux fichiers de configuration (quota utilisateur 100Mo)
 La **partition /save** est réservée pour les données à sauvegarder (quota utilisateur 250Go)
 La **partition /work** est réservée pour les fichiers temporaires de calcul (quota utilisateur 1To)
 => ATTENTION CE VOLUME N'EST PAS SAUVEGARDE ET PEUT ÊTRE PURGÉ SI BESOIN
 => IL EST DE VOTRE **RESPONSABILITÉ DE GÉRER VOS DONNÉES (organisation, volumétrie, pertinence, ancienneté)**
 Pour connaître votre consommation d'espace disque, utilisez la commande suivante :

du -csh /DIR_NAME/USER_NAME/*

=====
 Informations concernant l'utilisation de la mémoire
 =====

Il existe une limitation de 8Go de RAM par processus utilisateur (sur le cluster).
 Pour obtenir plus de mémoire, veuillez consulter la FAQ de notre site web (cf + bas)

=====
 Informations concernant le quota de temps de calcul
 =====

Il existe un **quota de temps de calcul annuel** de 100.000 pour les académiques (500H pour les entreprises privées).
 Au-delà, il faut renseigner le formulaire de demande de ressources exceptionnelles.
 Vous pouvez vérifier votre quota de calcul avec la commande: qquota_cpu <login>

=====
 Support
 =====

Pour plus d'informations, consulter le site web :
<http://bioinfo.genotoul.fr/>
 Pour toute demande de support, adressez-vous à :
support.genopole@toulouse.inra.fr

Pas d'accès direct en ssh sur les noeuds de calculs

CAR file d'attente : gestion des priorités/ressources par GRID SCHEDULER.



DONC

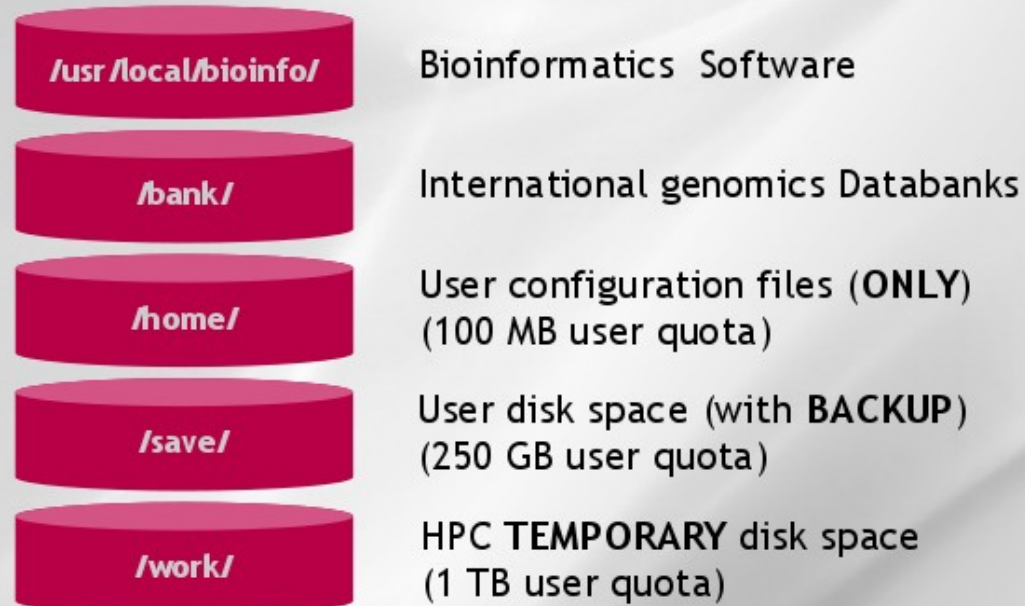
ssh pour un accès au frontal

PUIS

QLOGIN pour lancer des calculs **UNIQUEMENT SUR VOTRE /WORK**

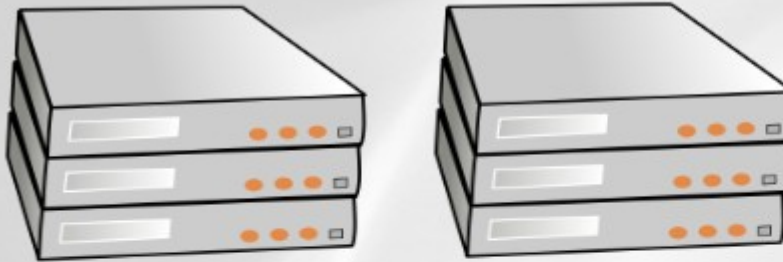
exit pour sortir du cluster

Disk spaces

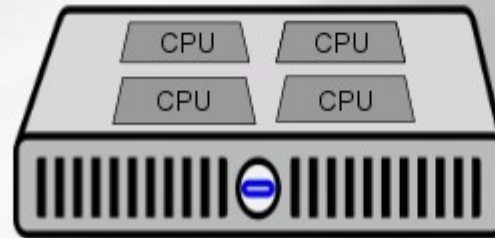




- Cluster : set of nodes

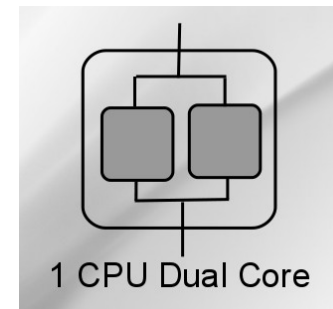


- Node : Huge computer (with several CPUs)



08/03/16

1 cluster = une série de noeuds
1 noeud = un énorme ordinateur de plusieurs CPU
1 CPU = Central Processing Unit = 2 coeurs



Plusieurs queues de traitement sont disponibles selon les ressources demandées

Queue	Access	Priority	Max time	Max slots
Workq (default)	everyone	300	96H	4120
unlimitq	everyone	100	unlimited	680
smpq	on demand	0	unlimited	240
hypermemq	On demand	0	unlimited	96
Interq (qlogin)	everyone	100	48H	40

Les ressources sont calculées et limitées selon le groupe linux auquel appartient l'utilisateur : INRA, Région, contributeur, autre.

Defaults parameters

- Workq
- 1 core
- 8 GB memory maximum
- Write only /work directory (temporary disk space)
- 1 TB quota disk per user (on /work directory)
- 120 days files without access automatic purged
- 100 000H annually computing time (more on demand)

1- ssh sur le serveur frontal genotoul

2- Se positionner dans /work/user/

3- Ecrire un script : myscript.sh
#!/bin/bash <--- shebang bash
ls -ltrah <--- Ligne de commande

4 - qlogin

5 - cd /work/user/ → Oui, oui, de nouveau après le qlogin ;)

6 - Lancer votre script

```
sh myscript.sh OU qsub toto.sh
```

```
[smaman@node001:/work/smaman]$ qsub toto.sh
```

```
Your job 5117398 ("toto.sh") has been submitted
```

```
[smaman@node001:/work/smaman]$ qstat -u user
```

```
[smaman@node001:/work/smaman]$ls -ltrah
```

```
-rw-r--r--  1 smaman      BIOINFO  23 7 nov. 15:43 toto.sh
```

```
-rw-r--r--  1 smaman      BIOINFO   0 7 nov. 15:44 toto.sh.e5117398
```

```
drwxrwxrwx 64 smaman      BIOINFO 16K 7 nov. 15:44 .
```

```
-rw-r--r--  1 smaman      BIOINFO 16K 7 nov. 15:44 toto.sh.o5117398
```

```
[smaman@node001:/work/smaman]$ exit
```

```
logout
```

```
/SGE/ogs/inra/tools/qlogin_wrapper.sh exited with exit code 0
```

Exercice : lançons un premier job avec plus de paramètres du qsub

- N job_name : to give a name to the job
- q queue_name : to specify the batch queue
- o output_file_name : to redirect output standard
- e error_file_name : to redirect error file
- l mem=8G: to ask for 8GB of memory (minimum reservation)
- l h_vmem=10G : to fix the maximum consumption of memory

```
$qsub -l mem=32G -l h_vmem=36G myscript.sh  
Your job 15661  
("mon_script.sh") has been submitted
```



qstat

```
job-ID prior name user state submit/start queue slots ja-task-ID
```

Job-ID : Job identifier
prior : priority of job
name : job name
user : user name
state : actualy state of job (see follow)
submit/start at : submit/start date
Queue : batch queue name
slots : number of slots asked for the job
ja-task-ID : job array task identifier (see follow)

State: d(letion), e(rror), h(old), w(aiting), t(ransferring), r(unning).

Plus de détails à propos du qstat : man qstat

cd : change directory

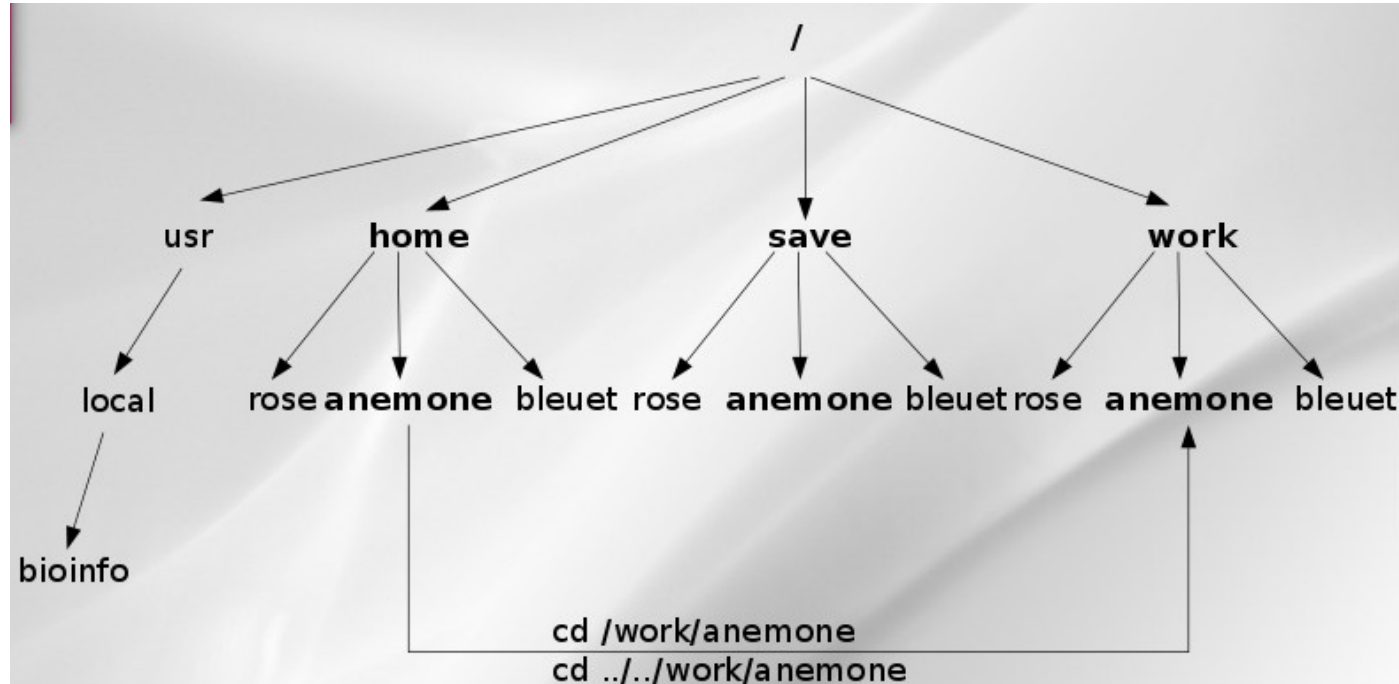
pwd : print working directory

ls [nom_répertoire] : list directory contents

who : show who is logged on the server

passwd : update user's authentication token

history : display the commands history



Comprendre le résultat d'un ls -ltrah

```

-rw-r--r-- 1 cnoirot BIOINFO 3685591 Jun  9 11:56 Diapo_F10a.odp
drwxr-xr-x 3 cnoirot BIOINFO  4096 Jul  8 14:56 igv
-rwxr-xr-x 1 cnoirot BIOINFO    20 Apr 16 11:21 monscript.sh
-rw-r--r-- 1 cnoirot BIOINFO 954415 Oct  3 2009 Presentation_pyrocleaner.odp
lrwxrwxrwx 1 cnoirot BIOINFO    13 Mar 15 2009 save -> /save/cnoirot
lrwxrwxrwx 1 cnoirot BIOINFO    13 Mar 18 2009 work -> /work/cnoirot

```

Changer les permissions avec chmod

chmod [options] filename

modifies the permissions of a file

- **u** : user, **g** : group, **o** : other, **a** : all
- **r** : read, **w** : write, **x** : execute

Créer un lien symbolique avec ln -s file-source file-destination



`mkdir directory_name`

`touch new_file_name`

`cp file_source file_destination`

`cp -r dir_source dir_destination`

`mv file dir/.`

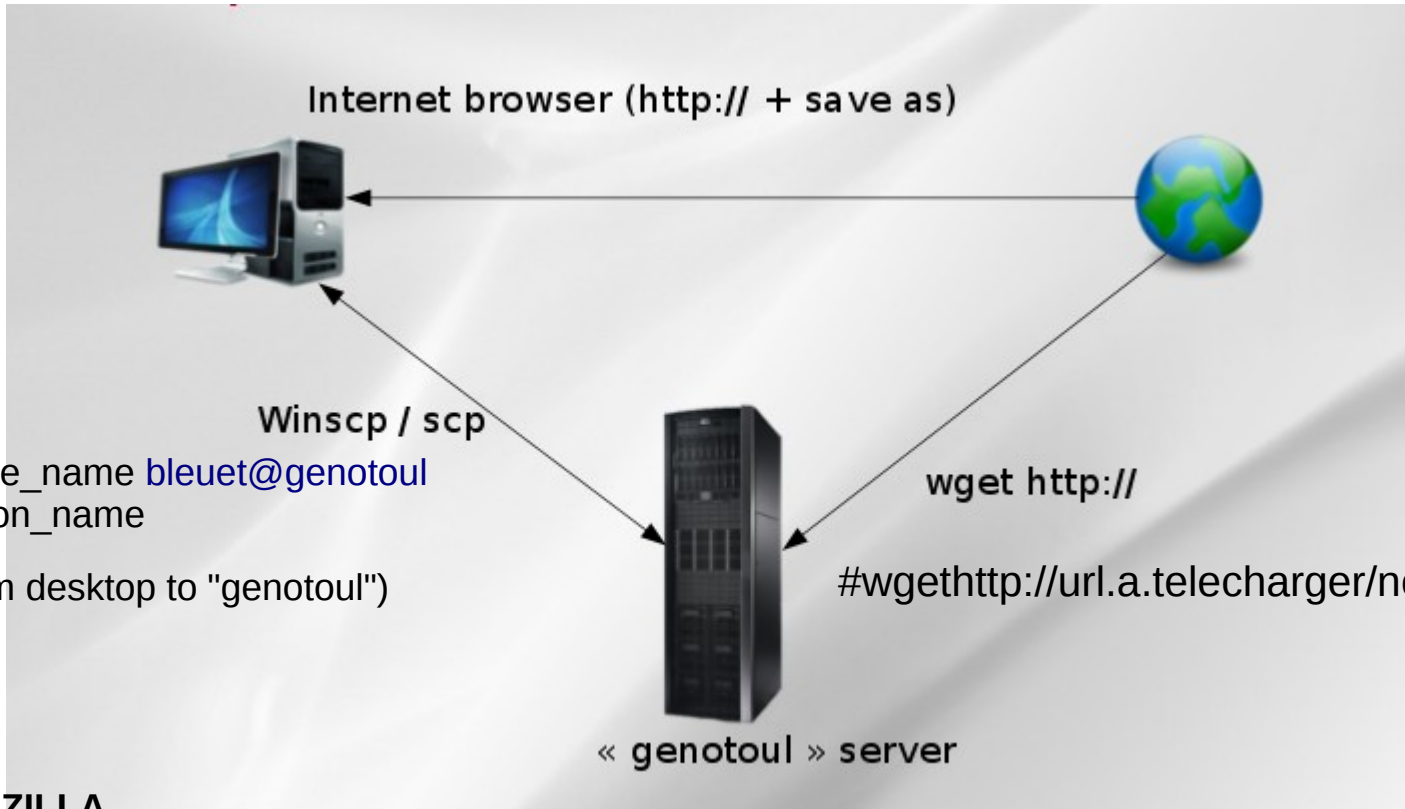
`mv old_file_name new_file_name`

`ls *.txt`

`find . -type f *.txt`

`cat file - more file - less file`

`geany - gedit - vi`



```
scp source_name bleuet@genotoul  
:destination_name
```

(copy from desktop to "genotoul")

```
#wgethttp://url.a.telecharger/nom_fichier
```

Ou FILEZILLA

Merci pour votre écoute

Remerciements



Fonds Européen
de Développement Régional

