

Présentation clusters

Franck Dary

JTT

18 mars 2021

Clusters disponibles :

- LIS : ~30GPU, ~200CPU — pas de limite de temps.
- Mesocentre : Milliers de CPU — 250.000 heures pour l'équipe.
- Jean-Zay : Milliers de GPU — Environ 10.000 heures par projet.

Se connecter avec ses identifiants lis-lab, en passant par saphir, exemple :

```
ssh -J prénom.nom@saphir2.lis-lab.fr prénom.nom@sms-ext.lis-lab.fr
```

Si pas de compte demander à support@lis-lab.fr

Par ssh : `ssh user@login.mesocentre.univ-amu.fr`

Où user devrait être première lettre du prénom + nom.

La demande se fait sur le site <https://www.edari.fr/> :

- Créer un compte.
- Déclaration de compte calcul, renseigner :
 - ▶ Status
 - ▶ Laboratoire
 - ▶ IP : 139.124.22.4
 - ▶ FQDN : saphir2.lidil.univ-mrs.fr
- Création de dossier type "Accès Dynamique" :
 - ▶ Description du projet : pas besoin de plus de 5 lignes.
 - ▶ Nombre d'heures GPU : 10.000 sans soucis.
 - ▶ Stockage : 1To sans soucis.

Sur Jean-Zay :

- Maintenance (inaccessible, pas de jobs) tous les mardis matin de 8h à 12h.
- Aucun espace (~3Go) dans \$HOME. Une solution, tout mettre dans \$WORK sauf .bash_profile et .bashrc.

Apprendre à utiliser les clusters

Ressources utiles :

- Wiki LIS : <https://gitlab.lis-lab.fr/cluster/wiki/-/wikis/home>
- Wiki Idris : <http://www.idris.fr/jean-zay/> (très complet).

Même ordonnanceur Slurm sur les 3 clusters donc très similaires.

Exemple job Jean-Zay

```
#!/usr/bin/env bash
#SBATCH --array=0-1000%25
#SBATCH --job-name=monSuperProgramme
#SBATCH --output=%A_%a.out
#SBATCH --error=%A_%a.err
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=10
#SBATCH --gres=gpu:1
#SBATCH --hint=nomultithread
#SBATCH --qos=qos_gpu-t4
#SBATCH --time=48:00:00
module purge
module load gcc/9.1.0
module load python/3.7.5
./programme1 $SLURM_ARRAY_TASK_ID
./programme2 $SLURM_ARRAY_TASK_ID
```


Exemple job Mesocentre

```
#!/usr/bin/env bash

#SBATCH --array=0-55%200
#SBATCH --job-name=trainOld
#SBATCH --output=%A_%a.out
#SBATCH --error=%A_%a.err
#SBATCH --ntasks=1
#SBATCH --cpus-per-task=4
#SBATCH --partition=skylake
#SBATCH -A b250
#SBATCH --time=48:00:00

module purge
module load userspace/all
module load gcc/10.2.0
module load python3/3.8.6

./programme $SLURM_ARRAY_TASK_ID
```

Astuce array commandes différentes

```
#SBATCH --array=0-55%200
#SBATCH --job-name=trainOld
#SBATCH --output=%A_%a.out

names=('nom_job_1' 'nom_job_2' 'nom_job_3')
commands=('macaon train --machine "{
Name : test
Size : 200
}" --silent' 'bert decode german' 'cowsay ')

oldOut=${SLURM_ARRAY_JOB_ID}_${SLURM_ARRAY_TASK_ID}.out"
newOut=${names[$SLURM_ARRAY_TASK_ID]}.stdout"
mv $oldOut $newOut

eval "${commands[$SLURM_ARRAY_TASK_ID]}"
```