

Présentation des équipements HPC@LR

Comment soumettre son premier job

Objectifs

- ▶ Renouveler le cluster octantio
- ▶ Plateforme test pour le futur centre Meso@LR

Le cluster thau

- ▶ 1 nœud d'administration Dell R620
- ▶ 15 nœuds de calcul Dell R630
 - Bi processeur Intel Xeon E5-2680 v3 2,5 Ghz (2 x 12c)
 - 64 Go DDR4 2133 Mhz RDIMM

360 cœurs

14,4 Tflop puissance crête

Le cluster thau

- ▶ Espace disque unique entre thau et octantio mais comptes différents.
- ▶ Si vous avez un compte sur octantio
=> nouveau projet
- ▶ Si vous n'avez pas de compte
=> procédure classique.

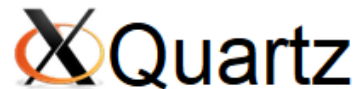
Connexion ssh

- ▶ Depuis Linux

- [chapuis@local ~]\$ ssh -X chapuisatb@thau-login.hpc-lr.univ-montp2.fr

- ▶ Depuis MacOS X

Pas d'interface X11 par défaut

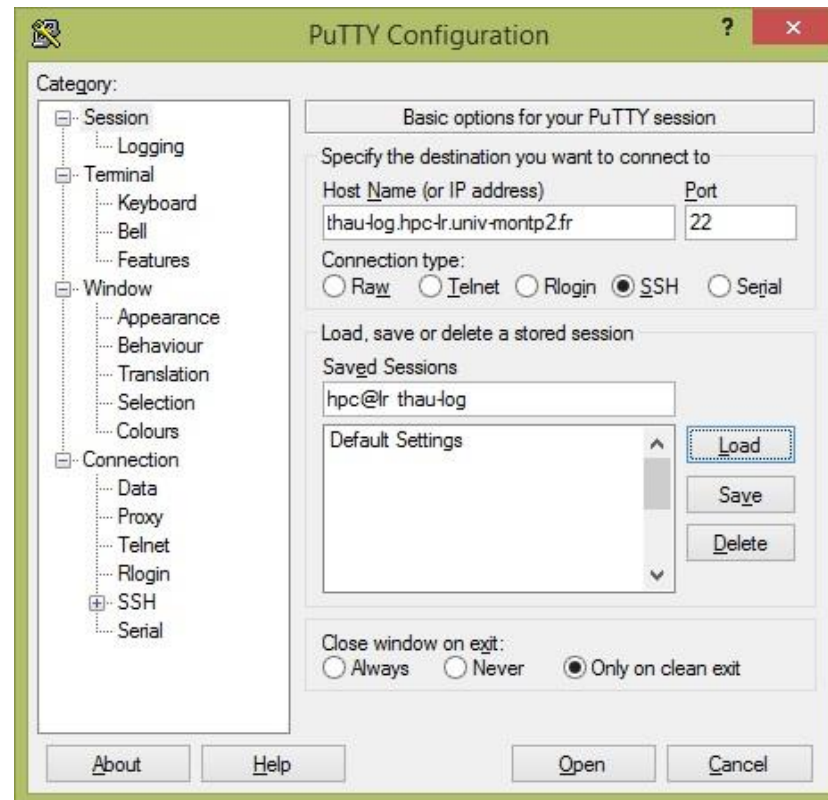


<http://xquartz.macosforge.org/landing/>

Connexion ssh depuis Windows

PuTTY

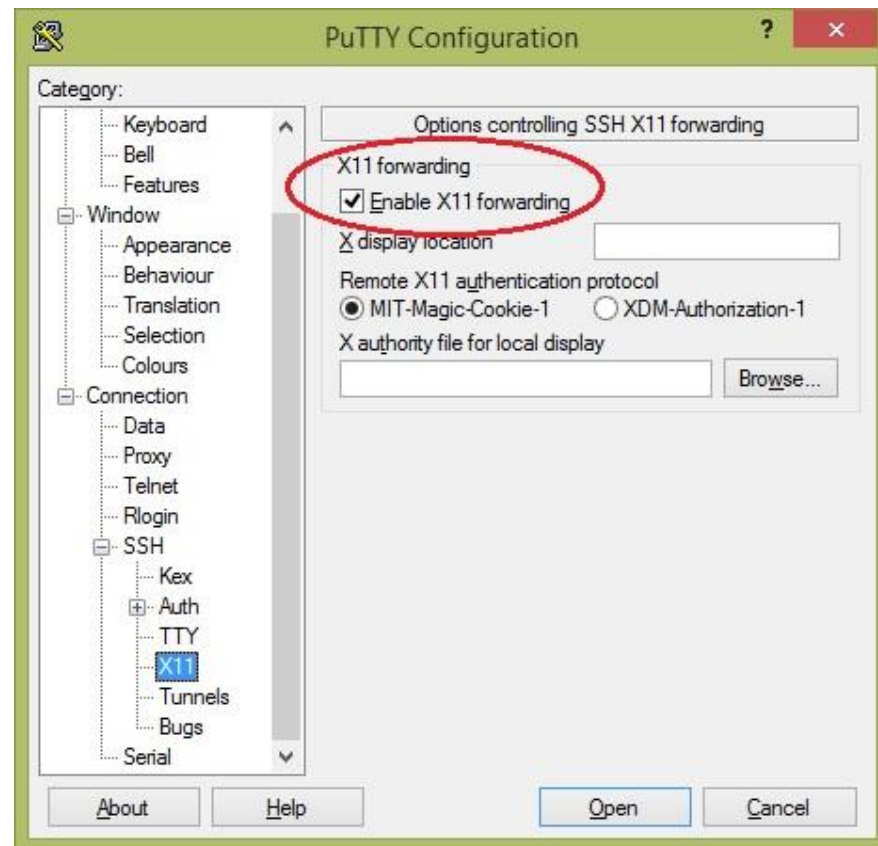
<http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html>



Connexion ssh depuis Windows



<http://www.straightrunning.com/XmingNotes/>



Environment modules

Permet de renseigner les variables d'environnement selon vos besoins (PATH, LD_LIBRARY_PATH...)

```
chapuis@thau-log:~  
[chapuis@thau-log ~]$ module avail  
----- /cm/shared/modulefiles -----  
blacs/openmpi/gcc/64/1.1patch03      hwloc/1.8.1                          iozone/3_420  
blas/gcc/64/1                        intel/compiler/64/15.0/2015.3.187    lapack/gcc/64/3.5.0  
bonnie++/1.97.1                      intel/ipp/64/8.2/2015.3.187          mpich/ge/gcc/64/3.1  
cmgui/7.0                             intel/mkl/64/11.2/2015.3.187        mpiexec/0.84_432  
default-environment                  intel/sourcechecker/64/15.0/2015.3.187 mvapich/gcc/64/1.2rc1  
fftw2/openmpi/gcc/64/double/2.1.5    intel/tbb/32/4.3/2015.3.187        mvapich2/gcc/64/2.0b  
fftw2/openmpi/gcc/64/float/2.1.5     intel/tbb/64/4.3/2015.3.187        mvapich2/gcc/64/2.1rc1  
fftw3/openmpi/gcc/64/3.3.3           intel-cluster-checker/2.2.1         netcdf/gcc/64/4.3.1.1  
gcc/4.8.2                             intel-cluster-runtime/ia32/3.7      netperf/2.6.0  
globalarrays/openmpi/gcc/64/5.3      intel-cluster-runtime/intel64/3.7   openblas/dynamic/0.2.8  
hdf5/1.6.10                           intel-cluster-runtime/mic/3.7       openmpi/gcc/1.8.4  
hdf5_18/1.8.12                       intel-tbb-oss/ia32/43_20141023oss   scalapack/gcc/64/1.8.0  
hpl/2.1                                intel-tbb-oss/intel64/43_20141023oss slurm/14.03.0  
[chapuis@thau-log ~]$
```

```
[chapuis@thau-log ~]$ echo $MODULEPATH  
/cm/shared/modulefiles
```


Environment modules

- ▶ Format de la commande:
 - module <paramètre> <*nom_du_module*>
 - avail : liste des modules disponibles
 - show : affiche les informations d'un module
 - add / load : charge un module
 - list : liste des modules chargés
 - rm / unload : décharge un module
 - purge : décharge tous les modules

Environment modules

- `initadd` : ajoute un module dans le fichier d'initialisation du shell
- `initrm` : enlève un module du fichier d'initialisation du shell

<http://modules.sourceforge.net/man/module.html>

Le gestionnaire SLURM

Simple Linux Utility for Resource Management

- ▶ gestionnaire de ressources
- ▶ Les +
 - Open source
 - Très répandu
 - Intègre des outils de monitoring



Le gestionnaire SLURM

SLURM	LoadLeveler	PBS / Torque
sbatch	llsubmit	qsub
squeue	llq	qstat
scancel	llcancel	qdel

Le gestionnaire SLURM

- ▶ sbatch, srun
- ▶ squeue
- ▶ scancel <num_job>

- ▶ sinfo (donne des informations sur les nœuds et sur les partitions)
- ▶ scontrol show jobid <num_job>
- ▶ sacct, sreport

Format long	Format court	Description
--job-name	-J	Nom du job
--account		Nom du projet
--nodes	-N	Nombre de nœuds
--ntasks	-n	Nombre de tâches au total
--ntasks-per-node		Nombre de tâches par nœud (doit correspondre au nombre total de tâches divisé par le nombre de nœuds)
--ntasks-per-core		Nombre de tâches par cœurs
--partition	-p	Groupe de machines sur lequel le job va tourner

Format long	Format court	Description
--mem		Mémoire réservée par nœud en MegaBytes? (Sans cette option la totalité de la mémoire du nœud est allouée)
--time	-t	Durée maximum d'exécution du job (Passé ce délai, le job est automatiquement arrêté)
--output	-o	Nom du fichier de sorties. Par défaut slurm-%j.out ou '%j' est le numéro du job
--error	-e	Nom du fichier d'erreur. Par défaut slurm-%j.err ou '%j' est le numero du job
--mail-user		Adresse mail où envoyer les notifications

- ▶ [chapis@thau-log thau]\$ srun -N 2 -n 2 -p defq hostname
- ▶ [chapis@thau-log thau]\$ sbatch ./job-thau.sh

```
#!/bin/sh
#SBATCH --job-name=test
#SBATCH -N 2
#SBATCH -n 6
#SBATCH --account=projet
#SBATCH --ntasks-per-node=3
#SBATCH --ntasks-per-core=1
#SBATCH --partition=defq
```

```
echo "Running on: $SLURM_NODELIST"
mpirun ./prog_f90.exe
```


Le gestionnaire SLURM

- ▶ sbatch, srun
- ▶ squeue
- ▶ scancel <num_job>

- ▶ sinfo (donne des informations sur les nœuds et sur les partitions)
- ▶ scontrol
- ▶ sacct, sreport

- ▶ [chapis@thau-log thau]\$ scontrol show jobid <num_job>
- ▶ [chapis@thau-log thau]\$sreport cluster AccountUtilizationByUser
user=\$LOGNAME -t hour start=2015-01-01T00:00:00
- ▶ [chapis@thau-log ~]\$ sacct -u \$LOGNAME