HPC CHEAT SHEET

· Logging in:

Linux/MacOS - Terminal
Windows - Putty (Hostname: <username>@marvin.s.upf.edu)

ssh <username>@marvin.upf.edu

Connect with X11 forwarding:

ssh -X <username>@marvin.upf.edu

· Transferring files:

Upload file to cluster:
scp /PATH/TO/LOCAL <username>@marvin.upf.edu:/PATH/TO/REMOTE
Download file from cluster:
scp <username>@marvin.upf.edu:/PATH/TO/REMOTE /PATH/TO/LOCAL

· Storage:

Fileset	Description	Quota	Backup
homes	Non-replicable data (scripts,…)	50Gb ($ myquota)	weekly
scratch	Hot Data	No (Total: 297T)	No
projects	Cold Data	No (Total: 517T)	No

Hot/Cold data?

· Modules:

Modules are classified as apps (applications), libs (libraries), tools (compilers and others) and toolchains (a group of modules for work with)

Show available modules:
module av *OR* module av <name> (*E.g.: module av samtools*)
Load a module:
module load <class>/<name>/<version>
E.g.: module load SAMtools/1.6-foss-2016b
Unload a module:
module unload <class>/<name>/<version>
E.g.: module unload SAMtools/1.6-foss-2016b
List all loaded modules:
module list
Unload all loaded modules (and load the module structure):
module purge; module load modulepath/noarch

Useful SLURM commands

Send a batch job:
sbatch <parameters> <slurm_script>
* **Sbatch allows for a filename pattern to contain one or more replacement symbols:**
%j: **Jobid of the running job.** %a: **Job array ID (index) number.**
%u: **Username.** %x: **Job name.**

Start an interactive job:
interactive <parameters>
Start a task inside a job:
srun <program>
View state of pending / running jobs:
squeue -u <username>
View state of pending / running / finished jobs
sacct -j <job_id> // seff <job_id> // scontrol show job <job_id>
Cancel a job:
scancel <job_id>

· SLURM job submission examples:

Standard Serial job:

```
#!/bin/bash                                  #Bash directive
#SBATCH -J job_name                          # Job name identification
#SBATCH -e error_file.%j.err                 # Path to stderr file
#SBATCH -o output_file.%j.out                # Path to stdout file
#SBATCH --mail-user-=user@mail.com           # Send an email if the job
#SBATCH --mail-type=BEGIN,END,FAIL           #+ starts / ends / fails
#SBATCH -t 1-00:00                           # Time limit in dd-hh:mm
#SBATCH --mem-per-cpu=7G                     # memory per CPU
```

Standard Parallel MPI job (Generally for multi-node parallelization):

```
#!/bin/bash
#SBATCH -J job_name
#SBATCH -n 4                                 # MPI tasks
#SBATCH --mem-per-cpu=7G                     # memory per CPU
#SBATCH -t 1-00:00
<Load modules and run application>
```

Standard Parallel OMP job (Generally for intra-node parallelization):

```
#!/bin/bash
#SBATCH -J job_name
#SBATCH -n 1
#SBATCH -c 2                                 # OMP threads
#SBATCH -t 1-00:00
#SBATCH --mem-per-cpu=7G                     # memory per CPU
<Load modules>
export OMP_NUM_THREADS=${SLURM_CPUS_PER_TASK}
<command> + $OMP_NUM_THREADS
```

Parallel MPI+OMP job (4x2 cores):

```
#!/bin/bash
#SBATCH -J job_name
#SBATCH --ntask=4                            # MPI tasks
#SBATCH --cpus-per-task=2                    # OMP threads
#SBATCH -t 1-00:00
#SBATCH --mem-per-cpu=7G                     # memory per CPU
<Load modules>
export OMP_NUM_THREADS=${SLURM_CPUS_PER_TASK}
<run application>
```

GPU job:

```
#!/bin/bash
#SBATCH -J job_name
#SBATCH -gres=gpu:2                          # Request 2 GPGPU P100 per node
#SBATCH -N 1                                 # Number of nodes
```

· SLURM job environment variables:

SLURM sets some useful environment variables at job start:

SLURM_JOB_NAME	Name of the job.
SLURM_JOB ID	The ID of the job allocation.
SLURM_ARRAY_JOB_ID	Job array's master job ID number.
SLURM_ARRAY_TASK_ID	Job array's current ID (index) number.
SLURM_ARRAY_TASK_MIN	Job array's minimum ID (index) number.
SLURM_ARRAY_TASK_MAX	Job array's maximum ID (index) number.
SLURM_CPUS_PER_TASK	Number of cpus requested per task.
SLURM_NTASKS	Same as -n, --ntasks.
SLURM_CPUS_PER_TASK	Only set if the --cpus-per-task option is specified.
SLURM_SUBMIT_DIR	The directory from which sbatch was invoked.

· SLURM job environment variables:

SLURM does not accept spaces between # and SBATH word:

SBATCH -J job_name → INCORRECT

#SBATCH -J job_name → CORRECT

The following options can be omitted:

-c, --cpus-per-task	Default value = 1
--mem-per-cpu	Defalult value = 7G

For more useful information or technical support visit us at:
https://www.upf.edu/web/sct-sit