ACE3P CW10 Supercomputing Cheat Sheet, Sep 22 2010

Franklin is a supercomputer at NERSC (National Energy Research Center) in Berkeley, CA. It has a scheduler to run computational jobs based on a queue. The user has to use a job submission script to run a job. Once a job is submitted, the scheduler will decide when and where the job will be executed.

- You first need to get a NERSC Franklin account
  (send your information to ace3p@slac.stanford.edu :
  First and Last names, nationality, phone, address, and email)

- General overview at http://www.nersc.gov/nusers/systems/franklin/
- Or go directly to
  http://www.nersc.gov/nusers/systems/franklin/running_jobs/

- Transfer mesh and input file(s) from your desktop to your directory on Franklin's high-performance file system:

  scp –pr <problem_dir>
  `franklin.nersc.gov:/scratch/scratchdirs/<user>/`

  (Note, that since we want access to the high-performance file system, we need to run our ACE3P simulations from within the /scratch directory)

- Login on Franklin:
  `ssh franklin.nersc.gov`

- Go to the above directory:
  `cd /scratch/scratchdirs/<user>/<problem_dir>/`

- Note that you can find the AC3P executables for Franklin in the directory
  `~candel/.community/franklin/`

How to run a Job on Franklin:
http://www.nersc.gov/nusers/systems/franklin/running_jobs/

- The following is a sample job submission script. You can find it at `~candel/.community/franklin/run.pbs`.

```
#PBS -N JOBNAME
#PBS -q regular              <- which queue to use (regular is fine)
#PBS -l mppwidth=128         <- requested number of CPUs (multiple of 4)
#PBS -l walltime=00:30:00    <- requested time (don't need to use it all)
#PBS -j eo                   <- combine standard output + error
cd $PBS_O_WORKDIR            <- go to the right place on compute nodes

### Omega3P (with growing amount of memory per process)
#aprun -n 128 -N 4 ~candel/.community/franklin/omega3p input.o3p
#aprun -n  64 -N 2 ~candel/.community/franklin/omega3p input.o3p
#aprun -n  32 -N 1 ~candel/.community/franklin/omega3p input.o3p

### T3P with VolumeMonitor -> Mode conversion (also parallel)
aprun -n 128 -N 4 ~candel/.community/franklin/t3p input.t3p
aprun -n  64 -N 2 ~candel/.community/franklin/acdtool postprocess
volmontomode in.t3p
```

Notes on the parameters of aprun, the "parallel job starter"

-n specifies how many processes overall (distributed computing with MPI)
-N specifies how many processes per compute node are started

Franklin has 4 CPUs and 8 GB of memory per compute node.

`#PBS -l mppwidth=<NCPUS>` ***says how many processors (+memory!) your reserve***

Depending on problem size, you can fine-tune the available memory per process:
`aprun -n <NCPUS>  -N 4` ***starts 4 processes per node, each has around 2 GB***
`aprun -n <NCPUS>/2 -N 2` ***starts 2 processes per node, each has around 4 GB***
`aprun -n <NCPUS>/4 -N 1` ***starts 1 process per node, it has all the 8 GB for itself!***

For Omega3P/S3P and large problems, it may be better to start <NCPUS>/4 processes with 8 GB each than to use <NCPUS> processes with 2 GB each, as memory scalability is hard and the O/S and support libraries also occupy memory. For all other codes, it is usually OK to use all CPUs (-N 4).

Remember to use `Solver: CG` for T3P, to reduce memory requirements.

- Submit the job
  ```
  qsub run.pbs
  ```

- Look at the status of the batch jobs, and get the Job ID
  ```
  qstat -u <username>
  ```

- Watch the job's printout as it runs
  ```
  tail –f <Job ID>.nid0000X.ER
       (e.g, tail –f 6733359.nid0003.ER)
  ```

- Cancel the job
  ```
  qdel <Job ID>   (e.g., qdel 6733359)
  ```

- Change the request time (before the job starts)
  ```
  qalter –l walltime=HH:MM:SS <Job ID>
  ```

- Copy the results back to your computer (from your computer)
  ```
  scp -pr franklin.nersc.gov:/<directory> .
  ```

- Clean up on Franklin:
  1) Move the results back into your project directory
  mv <directory> /project/projectdirs/gc2/<user>/

  or, even better
  2) upload it into HPSS (high performance storage system) for archival

  please consult http://www.nersc.gov/nusers/systems/hpss/