Getting Started with the National Systems Video Series

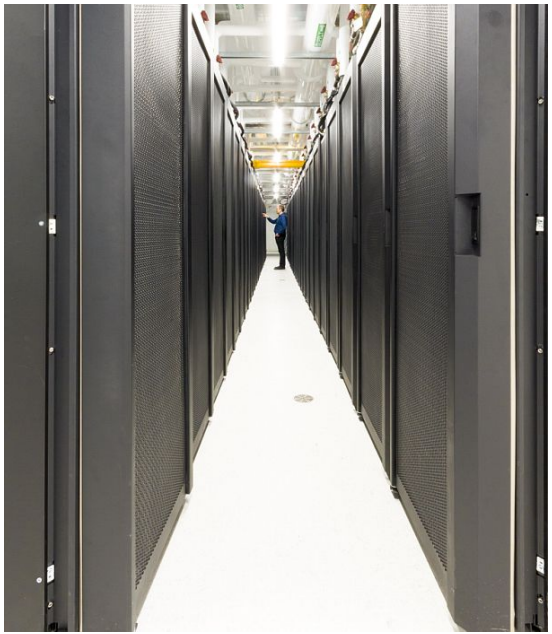
compute canada

Managing jobs

by Sergey Mashchenko, SHARCNET, McMaster University

Basic facts

- GPx clusters (Cedar and Graham) are shared HPC facilities used by thousands of researchers.
- In most cases, one cannot expect to login to the cluster and run a code instantly.
- Instead, users have to submit their computations to the scheduler as batch jobs.
- The scheduler used on GPx systems is **SLURM**.
- The general rule for any scheduler is the more resources (cpu cores, memory, GPUs, job runtime) you ask for, the longer the job has to wait in queue.



compute canada | calcul canada

SLURM commands

- sbatch: to submit a job
- squeue: to list your current jobs
- sacct: to list completed jobs
- scancel: to kill / cancel a job

How to minimize the wait time

- Specify the job runtime only slightly (~10-20%) larger than the estimated value.
- Only ask for the memory your code will actually need (with a bit of a cushion).
- Usually the more constraints you provide (e.g. whole nodes only, specific nodes, or specific subset of nodes), the longer the wait time becomes. Try to minimize the number of node constraints.
- Do not package what is essentially a bunch of serial jobs into a parallel (MPI / threaded) job - it is much faster to schedule many independent serial jobs than a single parallel job using the same number of cpu cores.

References

For more details about job scheduling on Graham and Cedar, check the following links:

- https://docs.computecanada.ca/wiki/Running_jobs
- https://docs.computecanada.ca/wiki/Job_scheduling

Official SLURM documentation (tutorials, man pages, FAQs etc):

- https://slurm.schedmd.com/tutorials.html

compute canada | calcul canada