



2018 HPC Course

Slurm on Fram

Saerda Halifu





Slurm On Fram

- Traditional cluster
- No extra layer like on hexagon
- You have to apply for CPU hours to run job
- #SBATCH -A option is mandatory





Type of Slurm Jobs

- **normal:** The default type. Most jobs are normal jobs. They run on between 4 and 32 standard nodes, but different projects can apply for more nodes if there is a need.
- **bigmem:** Jobs that require more memory (RAM). bigmem jobs run on special nodes with more RAM.(512G~6T)





Type of Slurm Jobs

- **preproc:** Short, 1-node jobs for things like preprocessing and postprocessing, that don't need many cpus and has maximum walltime of 1 day
- **devel:** Short jobs, meant for quick development and testing. They get higher priority to start sooner, but can only use up to 10 nodes in total. devel jobs also have a maximum walltime of 1 hour





Type of Slurm Jobs

- **optimist:** Low priority jobs that can start whenever there are free resources, but will be requeued when other jobs need the resources. Meant for jobs that use checkpointing.





Specify Different Job

- normal jobs must specify account (--account), walltime limit (--time) and number of nodes (--nodes)
- bigmen jobs must specify --partition=bigmem. In addition, they must specify wall time limit, number of CPUs/nodes and memory per CPU.





Specify Different Jobs

- preproc jobs must specify `--qos=preproc`.
A preproc job is allocated 1 node, exclusively, and has a maximum walltime of 1 day.
- devel jobs must specify `--qos=devel`. A devel job is like a normal job, except that it can use between 1 and 10 nodes, devel jobs also have a maximum walltime of 1 hour.





Specify Different Job

- optimist jobs must specify `--qos=optimist`, and should specify the number of nodes needed (between 4 and 32), but they should not specify wall time limit. They run on the same nodes as normal jobs.





Priority Principle

The principle is that a job's priority increases 1 point per minute the job is waiting in the queue, and once the priority reaches 20,000, the job gets a reservation in the future. It can start before that, and before it gets a reservation, but only if it does not delay any jobs with reservations.





Fram Job Priorities

- devel jobs start with 19,990, so get a reservation in 10 min
- normal, bigmem and preproc jobs start with 19,940, so get a reservation in 60 min
- optimist jobs start with 1 and end at 10,080, so they never get a reservation





Fram Job Priority

- Check what QOS is available
`sacctmgr show qos format=name,priority`





Interactive slurm Job

```
salloc --account=<my_account> --time <HH:MM> -N <no_of_nodes>
```

When you landed on the node you start your code with :

```
srun ./my_programm
```





support@metacenter.no

Thank You !



UNIVERSITETET I BERGEN