

Mox_gnu_parallel

There are many sources of information about GNU parallel on the web. They show the wide variety of ways to use GNU parallel. However, below is the simplest usage of GNU parallel.

In below commands, the file mywork contains a sequence of commands with one command per line. Note that all the commands should be independent of each other since GNU parallel may run the commands in any order. Each command should only use one core. The -j 28 option means that 28 processes will be started on the node to process the commands in the mywork file. Usually, the number of processes should be chosen to be equal to the number of cores on the node. However, if memory is limited on the node then you may choose this number to be smaller than the number of cores on the node. Ikt nodes have 16 cores per node. Mox nodes bought before August 2018 have 28 cores per node. Mox nodes bought after August 2018 have 32 cores per node. Mox nodes in 2019 have 40 cores.

Single Node

Use the options "--nodes=1 --ntasks-per-node=28" with your srun command or sbatch script to ensure that GNU parallel can use one node with 28 cores on the node.

```
module load parallel-20170722

cpu_count=28

cat mywork | parallel -j $cpu_count
```

Multiple Nodes

Use the options "--nodes=2 --ntasks-per-node=28" with your srun command or sbatch script to ensure that GNU parallel can use 2 nodes with 28 cores on each node.

```
module load parallel-20170722

cpu_count=28

scontrol show hostnames > list_of_nodes

cat mywork | parallel --sshloginfile list_of_nodes -j $cpu_count
```

ckpt queue

If you are using the ckpt queue then below GNU parallel options are useful. The "--joblog mylogfile" option keeps track of which tasks have been completed and the "--resume" option runs those tasks which have not been completed. Hence if your job gets interrupted and put back on the ckpt queue then when it starts again, GNU parallel will only run those tasks which were not completed earlier. (The specific task which was interrupted will be run again since it did not get completed earlier.)

```
cat mywork | parallel --joblog mylogfile --resume -j 28
```

```
===== ignore below =====
```

```
cpu_count=$(grep -c proc /proc/cpuinfo)
```