Slurm Script

Basic Slurm script commands

Slurm script command	Description
#!/bin/bash	Sets the shell that the job will be executed on the compute node
#SBATCH --ntasks=1 #SBATCH --n1	Requests for 1 processors on task, usually 1 cpu as 1 cpu per task is default.
#SBATCH --time=0-05:00 #SBATCH -t 0-05:00	Sets the maximum runtime of 5 hours for your job
#SBATCH --mail-user= <email>	Sets the email address for sending notifications about your job state.
#SBATCH --mail-type=BEGIN #SBATCH --mail-type=END #SBATCH --mail-type=FAIL #SBATCH --mail-type=REQUEUE #SBATCH --mail-type=ALL	Sets the scedualing system to send you email when the job enters the follwoing states: BEGIN,END,FAIL,REQUEUE,ALL
#SBATCH --job-name=my-named-job	Sets the Jobs name

Slurm script commands

Slurm script command	Description
#SBATCH −ntasks=X	Requests for X tasks. When cpus-per-task=1 (and this is the default) this requests X cores. When not otherwise constraint these CPUs may be running on any node
#SBATCH --nodes=X	Request that a minimum of X nodes be allocated to this job
#SBATCH --nodes=X-Y	Request that a minimum of X nodes and a maximum of Y nodes be allocated to this job
#SBATCH --cpus-per-task=X	Request that a minimum of X CPUs per task be allocated to this job
#SBATCH --tasks-per-node=X	Requests minimum of X task be allocated per node

Slurm script commands

Slurm script commands	Description of effects
#SBATCH --ntasks=1 #SBATCH --cpus-per-task=1	Requests 1 CPU (Serial) cpus-per-task is set to 1 by default and may be omitted.
#SBATCH --cpus-per-task=X #SBATCH --ntasks=1 #SBATCH --nodes=1	Requests for X CPUs in 1 task on 1 node (OpenMP) Both ntasks and nodes are set to 1 by default and may be omitted
#SBATCH --ntasks=X #SBATCH --tasks-per-node=X #SBATCH --cpus-per-task=1	Requests for X CPUs and tasks on 1 node cpus-per-task is set to 1 by default and may be omitted.
#SBATCH --ntasks=X #SBATCH --nodes=1 #SBATCH --cpus-per-task=1	Requests for X CPUs and tasks on 1 node cpus-per-task is set to 1 by default and may be omitted.

Slurm script commands

Slurm script commands	Description of effects
#SBATCH --ntasks=X #SBATCH --cpus-per-task=1	Requests X CPUs and tasks (MPI) cpus-per-task is set to 1 by default and may be omitted.
#SBATCH --ntasks=X #SBATCH --ntasks-per-node=Y #SBATCH --cpus-per-task=1	Requests for X CPUs and tasks with Y CPUs and tasks per node cpus-per-task is set to 1 by default and may be omitted.
#SBATCH --ntasks=X #SBATCH --nodes=1 #SBATCH --cpus-per-task=1	Requests for X CPUs and tasks on the same node, cpus-per-task is set to 1 by default and may be omitted.
#SBATCH --ntasks=X #SBATCH --nodes=1 #SBATCH --cpus-per-task=1	Requests for X CPUs and tasks on the 1 node cpus-per-task is set to 1 by default and may be omitted.

Slurm script commands

Slurm script command	Description
#SBATCH --ntasks=1 #SBATCH --cpus-per-task=1	Requests 1 cpu in 1 task. (Serial) cpus-per-task is set to 1 by default and may be omitted.
#SBATCH --cpus-per-task=N #SBATCH --ntasks=1 #SBATCH --nodes=1	Requests for X processors on the same node (OpenMP) Both ntasks and nodes are set to 1 by default and may be omitted
#SBATCH --ntasks=X	Requests for X processors which may be running on any node (MPI).
#SBATCH --nodes=X #SBATCH --ntasks=Y	Requests minimum of X nodes for the Y tasks. (MPI job)
#SBATCH --array=0-4	Requests Job array of 5 jobs with indexes 0-4
#SBATCH --array=1,3,5,7,9	Requests Job array of 5 jobs with indexes 1,3,5,7,9
#SBATCH --array=0-X%Y ex: #SBATCH --array=0-4%2	Requests Requests Job array of X jobs with only a maximum of Y jobs running at the same time

Slurm script commands

PBS script command	Description
#SBATCH --mem=4000	Requests 4000 MB of memory in total
#SBATCH --mem-per-cpu=4000	Requests 4000 MB of memory per cpu
#SBATCH --licenses=sas:2	Requests 2 SAS licenses
#SBATCH --gres=gpu:1	Requests that your job get 1 GPU allocated per node
#SBATCH --exclusive	Requests that your job run only on nodes with no other running jobs
#SBATCH --dependency=after:job_id1	Requests that the the job start after job (jobid1) has **started**
#SBATCH --dependency=afterany:job_id1, job_i2	Requests that the the job start after ether job (jobid1) or job (jobud2) has **finished**
#SBATCH --dependency=afterok:job_id1	Requests that the the job start after job (jobid1) has **finished successfully**

Slurm script commands

PBS script command	Description
#SBATCH --account=acc_name	To submit a job to a specific accounting group such as RAC/RAS allocation or different role
#SBATCH --tmp=200G	Asks for 200Gb of temporary disk space
#SBATCH --constraint=blue	To ask for a node feature or constraint set by cluster admin. Here we are looking for “blue” nodes.
#SBATCH --partition=partition_name	To ask for the job to run in a specific partition or queue by name, (unlike Moab there can be more than 1 partition per Job)
--prolog=<executable>	Run by srun only, runs the executable before the step
--epilog=<executable>	Run by srun only, runs the executable after the step finishes

SLURM Environment Variables

Environment Variable	Description
SLURM_JOB_NAME	User specified job name
SLURM_JOB_ID	Unique slurm job id
SLURM_NNODES	Number of nodes allocated to the job
SLURM_NTASKS	Number of tasks allocated to the job
SLURM_ARRAY_TASK_ID	Array index for this job
SLURM_ARRAY_TASK_MAX	Total number of array indexes for this job
SLURM_MEM_PER_CPU	Memory allocated per CPU
SLURM_JOB_NODELIST	List of nodes on which resources are allocated to Job
SLURM_JOB_CPUS_PER_NODE	Number of CPUs allocated per Node
SLURM_JOB_PARTITION	List of Partition(s) that the job is in.
SLURM_JOB_ACCOUNT	Account under which this job is run.

Getting information on your Job

Command	What its used for
squeue -u <username>	List all current jobs for a user
squeue -u <username> -t PENDING	List all pending jobs for a user
squeue -u <username> -t RUNNING	List all running jobs for a user
squeue -p <partitionname>	List all the jobs in a partition
scontrol show job <jobid>	List information on Job
scontrol show jobid -dd <jobid>	List detailed information on Job
sstat --format=AveCPU,MaxRSS,MaxVMSize,JobID -j <jobid>	List info resource used by your completed job : average cpu time, Max memory, Max virtual memory, JobId
sacct −u <username> --format=JobID,JobName,AveCPU,MaxRSS,MaxVMSize,JobID,Elapsed	List resources used by all jobs of a user
sprio	List job priority information

Controlling jobs

Command	What its used for
scancel <jobid>	Cancel job
scancel -u <username>	Cancel all the jobs for a user
scancel -t PENDING -u <username>	Cancel all the pending jobs for a user:
Scancel -name JobName	Cancel one or more jobs by name
scontrol hold <jobid>	Hold a job, prevent it form starting
scontrol resume <jobid>	Release a job hold, allowing the job to try to start
scontrol requeue <jobid>	Requeue a running, suspended or finished job into pending state
scontrol requeuehold<jobid>	First requeue the job than put a hold on it.
squeue -u <username> -ho %A -t R	List running jobs by user
squeue --start	Show expected start time of jobs. (This can change)

Getting information on you and your group

Command	What its used for
sacctmgr list Users USERS=<username>	List user and their default account (accounting group)
sacctmgr show user <username> withassoc	List user and their default account (accounting group) and shows more extensive information
sshare	Shows usage info for user.

Getting information on your Cluster

Command	What its used for
sinfo --states=idle	Show idle node on cluster
sinfo -R	Show down, drained and draining nodes and their reason
sinfo --Node --long	Show detailed node info.
scontrol show reservation	Shows reservations on the cluster
scontrol create reservation user=root starttime=now duration=infinite flags=maint nodes=<nodeid>	

Administrating your Cluster

Command	What its used for
scontrol create reservation user=root starttime=now duration=infinite flags=maint nodes=<nodeid>	Create a maintance reservation on node nodeid