

slurm

cheat sheet

Daemons

slurmctld	Executes on cluster's "head" node to manage workload.
slurmd	Executes on each compute node to locally manage resources.
slurmdbd	Manages database of resources limits, licenses, and archives accounting records.
slurmrestd	Interface to Slurm via REST API.

Job Submission

sbatch - Submit a batch script for later execution.

srun - Obtain a job allocation (as needed) and execute an application.

salloc - Obtain a job allocation.

--array=<indexes> (e.g. "--array=1-10")	Job array specification.
--account=<name>	Account to be charged for resources used.
--begin=<time>(e.g. "--begin=18:00:00")	Initiate job after specified time.
--clusters=<name>	Cluster(s) to run the job. (sbatch command only)
--constraint=<features>	Required node features.
--cpus-per-task=<count>	Number of CPUs required per task.
--dependency=<state:jobid>	Defer job until specified job(s) reach specified state.
--error=<filename>	File in which to store job error messages.
--exclusive[=user]	Allocated nodes can not be shared with other jobs/users.
--export=<name[=value]>	Export identified environment variables.

--gres=<name[:count]>	Generic resources required per node.
--job-name=<name>	Job name.
--label	Prepend task ID to output. (srun command only)
--licenses=<name[:count]>	License resources required for entire job.
--mem=<MB>	Memory required per node.
--mem-per-cpu=<MB>	Memory required per allocated CPU.
-N<minnodes[-maxnodes]>	Node count required for the job.
-n<count>	Number of tasks to be launched.
--nodelist=<names>	Specific host names to include in job allocation.
--output=<name>	File in which to store job output.
--partition=<names>	Partition/queue in which to run the job.
--qos=<name>	Quality Of Service.
--signal=[B:]<num>[@time]	Signal job when approaching time limit.
--time=<time>	Wall clock time limit.
--wrap=<command_string>	Wrap specified command in a simple "sh" shell. (sbatch command only)

Accounting

sacct - Display accounting data.

--allusers	Displays all users' jobs.
--accounts=<name>	Displays jobs with specified accounts.
--endtime=<time>	End of reporting period.
--format=<spec>	Specify the fields shown in the output.
--partition=<names>	Comma separated list of partitions to select jobs and job steps from.
--state=<state_list>	Display jobs with specified states.
--starttime=<time>	Start of reporting period.
--long	Provides output with more fields listed. .

sacctmgr - View and modify account information.

Options:

--immediate	Commit changes immediately.
--parsable	Output delimited by ' '

Commands:

add <ENTITY> <SPECS> create <ENTITY> <SPECS>	Add an entity. Identical to the create command.
delete <ENTITY> where <SPECS>	Delete the specified entities.
show <ENTITY> [<SPECS>]	Display information about the specific entity.
modify <ENTITY> where <SPECS> set <SPECS>	Modify an entity.

Entities:

account	Account used to group users.
association	Entity on cluster, consisting of cluster, account, user and (optionally) partition.
cluster	ClusterName parameter in the slurm.conf.
qos	Quality of Service.

sreport - Report job usage and cluster utilization.

--allusers	Displays all users' jobs.
--accounts=<name>	Displays jobs with specified accounts.

Report Types:

cluster	AccountUtilizationByUser, UserUtilizationByAccount, UserUtilizationByWckey, Utilization, WCKeyUtilizationByUser
job	SizesByAccount, SizesByAccountAndWckey, SizesByWckey
reservation	Utilization
user	TopUsage

Job Management

squeue - View information about jobs.

--account=<name>	View only jobs with specified accounts.
--clusters=<name>	View jobs on specified clusters.
--format=<spec> (e.g. "--format=%i,%j")	Output format to display. Specify fields, size, order, etc.
--jobs<job_id_list>	Comma separated list of job IDs to display.
--name=<name>	View only jobs with specified names.
--partition=<names>	View only jobs in specified partitions.
--priority	Sort jobs by priority.
--qos=<name>	View only jobs with specified Qualities Of Service.
--start	Report the expected start time and resources to be allocated for pending jobs in order of increasing start time.
--state=<names>	View only jobs with specified states.
--users=<names>	View only jobs for specified users.

sinfo - View all information about nodes and partitions.

--all	Display information about all partitions.
--dead	If set, only report state information for non-responding (dead) nodes.
--format=<spec>	Output format to display.
--iterate=<seconds>	Print the state at specified interval.
--long	Print more detailed information.
--Node	Print information in a node-oriented format.
--partition=<names>	View only specified partitions.
--reservation	Display information about advanced reservations.
-R	Display reasons nodes are in the down, drained, fail or failing state.
--state=<names>	View only nodes in specified states.

scancel - Signal jobs, job arrays, and/or job steps.

--account=<name>	Operate only on jobs running in the specified account.
--name=<name>	Operate only on jobs with specified name.
--partition=<names>	Operate only on jobs in the specified partition/queue.
--qos=<name>	Operate only on jobs using the specified quality of service.
--reservation=<name>	Operate only on jobs using the specified reservation.
--state=<names>	Operate only on jobs in the specified state.
--user=<name>	Operate only on jobs from the specified user.
--nodelist=<names>	Operate only on jobs using the specified compute nodes.

scontrol - Used view and modify configuration and state. Also see the **sview** graphical user interface version.

--all	When <i>show</i> command is used, then display all partitions, their jobs, and job steps.
--details	Causes <i>show</i> command to provide more details when available.
--future	Report nodes in FUTURE state.
--oneliner	Print information on one line.

Commands:

create <i>SPECIFICATION</i>	Create a new partition or reservation.
delete <i>SPECIFICATION</i>	Delete the entry with the specified <i>SPECIFICATION</i> .
reconfigure	All Slurm daemons will re-read the configuration file.
requeue <i>JOB_LIST</i>	Requeue a running, suspended or completed batch job.
setdebugflags	Add or remove DebugFlags of the slurmctld daemon.
show <i>ENTITY ID</i>	Display the state of the specified entity with the specified identification.
update <i>SPECIFICATION</i>	Update job, step, node, partition, or reservation configuration per the supplied specification.

Environment Variables

SLURM_ARRAY_JOB_ID	Set to the job ID if part of a job array.
SLURM_ARRAY_TASK_ID	Set to the task ID if part of a job array.
SLURM_CLUSTER_NAME	Name of the cluster executing the job.
SLURM_CPUS_PER_TASK	Number of CPUs requested per task.
SLURM_DEBUG_FLAGS	Specify debug flags for sbatch to use.
SLURM_JOB_ACCOUNT	Account name.
SLURM_JOB_ID	Job ID.
SLURM_JOB_NAME	Job Name.
SLURM_JOB_NODELIST	Names of nodes allocated to job.
SLURM_JOB_NUM_NODE	Number of nodes allocated to job.
SLURM_JOB_PARTITION	Partition/queue running the job.
SLURM_JOB_UID	User ID of the job's owner.
SLURM_JOB_USER	User name of the job's owner.
SLURM_RESTART_COUNT	Number of times job has restarted.
SLURM_PROCID	Task ID (MPI rank).
SLURM_STEP_ID	Job step ID.
SLURM_STEP_NUM_TASKS	Task count (number of MPI ranks).