

# Full DCC Job Scheduling and SLURM User Guide

This document describes the process for submitting and running jobs to the DCC under the [Slurm Workload Manager](#).

### What is a cluster, and why do I need SLURM?

Our [HPC cluster](#) is made up of a number of compute nodes, each with a varying level of scratch space, processors, memory, and sometimes GPUs. Users submit jobs to **SLURM** that specify the instructions (code) and software they want to run along with a description of the computing resources needed. SLURM, a combined scheduler and resource manager, then runs those jobs on the DCC.

## Cluster Basics and Terminology

**account:** Your DCC group is your SLURM account. A SLURM account grants users rights and manages limits in relation to SLURM partitions. You can see your DCC group(s) in [Research Toolkits](#).

**batch:** Instructions run in a file without user interaction, typically referred to as 'run in batch' or 'batch processing'. This is the most powerful way to use the DCC.

**batch script:** A text file containing a series of resource specifications, such as the number of nodes, amount of memory, and other requirements needed to run, and commands to be run. This script is submitted to the Slurm workload manager, which schedules and executes the job on available compute nodes within a high-performance computing (HPC) cluster.

**core:** A processing unit within a computer chip (CPU). The CPU in a node performs computations.

**GPU:** A graphics processing unit (GPU) is a specialized processor which can generate computer graphics, but in HPC, designed to accelerate computation-intensive tasks by performing parallel operations, making it ideal for tasks such as simulations, deep learning, and scientific computations.

**interactive session or job:** SLURM allows users to request a real time interactive session on a compute node through [OnDemand](#) or via SSH using `srun` so you can use a computing node real-time to code, debug, and/or watch your request run in real-time.

**job:** A task or set of tasks, submitted to the cluster scheduler to be executed.

**node:** A physical machine in a cluster, including login, compute, and transfer nodes

Type	Description
Login nodes	The login nodes are a place where users can login, edit files, view job results and submit new jobs. Login nodes are a shared resource and should not be used to run application workloads. There are limits on the login nodes.
Data Transfer node / Globus	The data transfer nodes are available for moving data to be accessible to and from the cluster. On the DCC, is done primarily via Globus.

Type	Description
Compute node	The compute nodes are the computers where jobs are run. To run jobs on the compute nodes, users need to access a head node and schedule their program to be run on the compute nodes once the requested resources are available.

**partition:** A partition, also called a job queue, is a subset of compute nodes on the cluster. Each partition is configured with a set of limits (including allowed users) that specify the requirements for jobs run in that partition. The DCC has a set of partitions for all users as well as partitions for specific use cases or lab groups.

**SLURM:** Is the open source, fault tolerant, scalable cluster management and job scheduling system we run on the Duke Compute Cluster.

- **When:** It finds and allocates the computing resources that fulfill the job's request at the soonest available time. The time to start the job will be faster if the batch script is written well for what the job will need (i.e. not asking for just the right amount of resources).
- **How:** When a job is scheduled to run, the scheduler instructs the resource manager to launch the application(s) across the job's allocated resources. This is also known as "running the job".
- **Results:** When a job has completed, you can check the job logs for informational updates or errors and the resources used via sacct or sreport. For long term usage, see our [usage reports](#).