

Scheduler Policy

Table of Contents

- [Definitions for clarification:](#)
- [Policy Review and Revisions](#)

“*Aeolus is managed following an enhanced community condominium model.*” This idea, taken from the [Aeolus Investment Policy](#), while it sounds inviting, it carries potential implications. The *community* part of *community condominium model* is how the overall costs for running Aeolus HPC can be kept low. By design, Aeolus has the capability of catering to highly variable and specific computational needs. The goal of the Scheduler Policies is to help guide distribution of demand (computational load) in a fair and predictable manner.

1. Access for those that need it will be granted to the cluster investors within two business days.
2. Resources will be allocated to run a lab's jobs based on resources funded by that lab.
3. A lab can have access to compute resources beyond those the lab may claim through investment, when such resources are not needed by those who funded them, during low compute demand periods.
4. All users have mutual access to others' compute resources (nodes) when they are not being utilized.
5. Jobs submitted by some funding lab and in conformity with directions in the [Job Submission Docs](#) will have priority access to that lab's funded resources.
6. When a lab submits a job that exceeds that lab's funded resources, the job will be queued and scheduled for execution when compute demand allows.

Definitions for clarification:

Term	Definition
Scheduler	Maui is the scheduler on Aeolus. Maui applies (enforces) policies describing how system resources (nodes, processors, memory) should be allocated to run jobs submitted by users via the qsub command, for all defined queues. Scheduler is blind to resource competition from foreground and background jobs that might be run without using the qsub command; thus compute jobs should be submitted through qsub.
Resource	A resource is 1) any descriptor used in specifying a job's requirements to run successfully, or 2) a characteristic of a queue which can be used to match a suitable job to that queue to allow it to run. Examples are: nodes, ppn (processors), ram (memory), wall-time, a node-series-identifier, or a CPU-architecture-identifier.
Resource Manager	Torque is the resource manager on Aeolus, and it interacts with the scheduler, Maui.
Queue	1) A FIFO (first-in first-out) list of jobs for sequential initiation and that may be based on each lab's resources, or 2) a class defined for a subset of available compute resources as might be required for a particular lab's job.
Reservation	A scheduling policy written to enforce availability of resources to a queue and/or lab parameters.
Preemption	Event of a job being halted in deference to another job with superior claim on queue or resources. Preemption can occur if a running job has a relevant flag (preemption) set. This generally requires that the code is designed to be interrupted and restarted later without significant loss of progress or computation.

Policy Review and Revisions

The VCEA HPC Committee will review and revise these policies as necessary, and at least annually. This is the April 2017 version.