Queue Usage and Policies

- Queues
  - Investor queues
  - The University of Iowa (UI) queue
  - The cluster wide queue
- Guidelines for selecting a queue

This document describes the queue structure of the HPC cluster system at a high level. It is intended to provide information to people so that decisions about job submissions can be made to best utilize the resources. The types of queues will be described and the relationships between them will be discussed. That will hopefully give people an idea of how the system is structured and how to make the best use of it.

The campus HPC cluster system at the University of Iowa use Sun Grid Engine (SGE). SGE is a type of software known as a resource manager and queuing system. In simple terms, the job of SGE is to take resource requests that a user makes for a computations and dispatch jobs to nodes on the cluster that provide the needed resources. Our cluster systems are built around the condominium model which means that people (investors) buy resources and allow others to use them when the investor does not need them. In addition to that, there is also some portion of the HPC systems that is funded by the University of Iowa and available to everyone. In order to make this model work, the purchased nodes are assigned to queues which will be named after the investor group that purchased them. Thus, one of the most important resources that one can specify is the queue that a job will run in.

Queues

The HPC cluster system consist of many queues that jobs can be submitted to. The purpose of having multiple queues is to provide different access levels to people depending on whether they are members of groups that have invested in the cluster by purchasing nodes or not. The queues break down on the following lines:

1. Investor queues. These queues are allocated to specific investor groups. There are physical machines assigned to the queues that represent the type of compute node that was purchased. Only users who are members of an investor group have access to these queues.
2. University general queue. This queue is made up of compute nodes that were purchased centrally by the University of Iowa. Everyone who has an account on the HPC cluster system can access this queue, but there is a limit to the number of running jobs per user.
3. General purpose queue. This queue covers the entire cluster and is available to everyone who has an account. There is no job limit but it is a lower priority queue than the other two types. Since this queue resides on the same machines as the investor queues, jobs in this general queue will get evicted when an investor needs their compute server to run a job.

Investor queues

Investor queues are generally named according to how the investor wants them named. This generally translates into the research group, a department, or the PI's name.

You will probably know if you are in a group or not but if you are not sure send an email to research-computing@uiowa.edu.

To see which Investor group you are associated with (if any) use the following command:

```
whichq
```

The above information is useful when you are trying to determine where to submit your job. Each investor queue is only available to members of the respective group. If you are a member of an investor group and your job(s) will fit in the number of slots allocated for that queue then those resources are guaranteed to be available for use to that group. If your job(s) will not fit in the allotment then you have the option of using the University or general queue. It is anticipated that members of an investment group will have their own system for deciding who runs what on their dedicated resources. As an example, if you are a member of the CGGER investment group and want to determine how many slots are currently available, the following command can be used.
The above indicates that 464 slots are available out of the 560 total slots allocated to that queue.

The University of Iowa (UI) queue

A significant portion of the HPC cluster systems at UI were funded centrally. These nodes are put into a queue named, 'UI', or prefixed with 'UI-' for special resources, such as high memory nodes, which have a queue called 'UI-HM'. These queues are available to everyone who has an account on an HPC system. Since that is a fairly large user base there are limits placed on these shared queues. The link below show the available UI queues and the restrictions in place.

Argon UI queues

Note that the number of slots available in the UI queue can vary depending on whether anyone has purchased a reservation of nodes. The UI queue is the default queue and will be used if no queue is specified. This queue is available to everyone who has an account on a UI HPC cluster system. In addition to the UI queue there is also the UI-DEVELOP queue on Argon. This queue has dedicated nodes with a wall clock limit of 24 hours.

Please use the development queue for testing new jobs at a smaller scale before committing many nodes to your job.

In addition to the above, the HPC system has some nodes that are not part of any investor queue. These are in the all.q queue and are used for node rentals and future purchases. The number of nodes for this purpose varies.

The cluster wide queue

Finally, there is cluster wide queue called the all.q queue. This queue encompasses all of the nodes and contains all of the available job slots. It is available to everyone with an account and there are no running job limits. However, it is a low priority queue instance on the same nodes as the higher priority investor and UI queue instances. The all.q queue is subordinate to these other queues and jobs running in it will give up the nodes they are running on when jobs in the high priority queues need them. The term we use for this is "job eviction". Jobs running in the all.q queue are the only ones subject to this.

Guidelines for selecting a queue

It may not always be obvious, particularly if you are a member of an investor group, which is the best queue to submit a job to. As a guideline, if you are in an investor group and there are enough free slots in your queue for your job(s) then you should use the investor queue. If you are not in an investor group, or there are not enough free slots in your investor queue, you should submit parallel jobs to the UI queue. If not submitting to an investor queue, and if your jobs are serial jobs, they should generally be submitted to the all.q queue. Unless you have a small number of jobs, and/or can not risk them getting evicted, then use the UI queue.

References to the UI queue also includes the UI-HM queue.
While not indicated in the above, a parallel job can be submitted to the all.q queue. Since a parallel job likely runs on more than one node, the likelihood of a job getting evicted is increased. Thus, it is recommended that parallel jobs be submitted to the UI queue in preference to the all.q queue.