Beagle2 Scheduling Policy

Beagle2 Queue & Scheduling Policies

This document describes the scheduling policies in place for Beagle2.

If you have questions about the terms used or about how to interact with the scheduler, please refer to the Scheduling FAQ.

If you have questions about how to submit or interact with jobs, please refer to the Job Management FAQ.

- Job Queues
- Job Priorities
- Fairshare Policy
- Job limits
- Node allocation
- Reservations
  - Standing Reservations
  - Advanced Reservations

Job Queues

There are two queues (defined in Torque) on Beagle2. Each queue was created for a specific phase of the code development cycle:

development:
- to test parallelism of your code on a small scale
- use this queue after you've gotten your code to compile and run using the interactive queue.
- up to 3 nodes
- max walltime: 00:30:00
- Monday through Friday between 8:00AM and 5:00PM 3 nodes are reserved for jobs that request less than 1 hour walltime.
- No particular queue or PBS switch needs to be specified.

batch:
- default queue where you'll run your production code at scale on Beagle2
- there is 728 standard compute nodes
- you can request up to 500 nodes per job

<table>
<thead>
<tr>
<th>Queue Name</th>
<th>Max Walltime</th>
<th>Max Nodes</th>
<th>Default Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development</td>
<td>30 minutes</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Batch</td>
<td>2 days</td>
<td>500</td>
<td>1</td>
</tr>
<tr>
<td>Advanced</td>
<td>No max walltime, no max number of nodes. Only certain people are authorized to use it because it has no restrictions.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Use `qstat -q` to check properties of a queue.

There is no need to request a specific queue. The scheduler will automatically put your job in the appropriate queue based on your node and walltime parameters.
If you don't specify job walltime or node count, your job will be assigned a default of 1 CPU for 10 minutes and placed in the development queue.

If you do specify a specific queue but don't specify a walltime, you will still be assigned a default walltime of 10 minutes but will have starting priority of the queue requested.

If for e.g., the development queue is full while the batch queue appears to have enough available nodes to run the job, it is possible to change queue by simply typing qmove <new_queue> <job id>, e.g., qmove batch <job id>.

Job Priorities

Beagle2 is essentially a general purpose resource and so, in general, no special preference is given to any user, job, or group, though if there are research deadlines or other short term needs, we may temporarily increase priorities for specific users or groups.

The normal policy is every user's job increases in priority by one for every minute that it sits in the queue. Please read the general scheduling FAQ for more specific information about how the scheduler uses priorities.

Fairshare Policy

Beagle2 does enforce a fairshare policy. Beagle2 uses 12 hour windows and tracks usage over the last 14 windows - the last week's usage. The window decay factor is 80%. Usage is tracked based on processor equivalent seconds per job. The fairshare target for a user is 10% and for a project a ceiling of 25%.

To show fairshare information type:

```bash
mdiag -f
```

Job limits

Other than walltime or node limits for the testing queues, there are no limits currently placed on the size or duration of a job.

Node allocation

Each job "slot" correlates to a node, so Beagle2 can accommodate up to 732 concurrently jobs.

User can request up to 500 nodes per one job.

Nodes cannot be shared by other users or even jobs of the same user.

It is possible to start multiple process on a node from one job. Please refer to the Beagle FAQ for more information on how to do this. Please refer to Job Management FAQ for more information on general job submission parameters.

Reservations

Standing Reservations

There are 3 standing reservations that correlate to the testing queues that reserve dedicated nodes for those queues. These reservations are active 24 hours a day, 7 days a week.

Advanced Reservations

We do allow advanced reservations on Beagle2 if you know you will need a certain number of nodes for a specific time frame. This can be useful if you have a demo for your program manager, have researchers onsite for a specific time, or need dedicated resources to meet a deadline (but we expect this situation to be uncommon, as most deadlines can be planned for in advance).
You can request an advanced reservation by emailing beagle-support@lists.uchicago.edu and specifying:

- how many nodes you need
- what time frames you need them for
- what users should have access to the reserved nodes.

To ensure that the required nodes will be available when you need them, you should request the reservation as early as possible, at least 4 days before you actually need them. So that nodes do not sit idle unnecessarily, please let us know if you no longer need the reservation or are done with it early so we can release the nodes back for other's use.