

Cluster Quick Guide

Computers you can log into

karst.uits.iu.edu

256 node, 16 cores per node, 16GB ram per core, for general jobs

mason.indiana.edu

18 nodes, 32 cores per node, 512GB ram per node, for high memory genomics jobs.

How much space do I have?

You can check it with:

`quota`

Modules

Modules make certain programs available to you in your path without having to find/install them.

View module list:

`module av #or module avail`

Load a module:

`module load <module name>`

Unload a module:

`module unload <module name>`

Swap a module (useful when you want to switch versions of a module):

`module swap <module name you want to use>`

List the modules you have loaded:

`module list`

Determine if a module needs another module loaded first (dependency) and to get more information on the module in general:

module display <module name>

Job Queues

You can run jobs without having to be signed in or wait on them. They are submitted to the "queue". Any job that will take more than 20 minutes must be submitted to the queue (there is one for Mason and one for Karst, just depends on where you are logged in).

Job Submission

In order to submit to the queue, you need to make a job file. They work kind of like an executable.bash would - they run the code in the file. However, additional information is needed in the header. It looks like this (all #comments are added):

```
#PBS -k oe #keep output and error
#PBS -m a b e #mail me when job : a - abort, b - begins, e - ends
#PBS -M <your email>
#PBS -N <name of job>
#PBS -l nodes=1:ppn=1,vmem=16gb,walltime=2:00:00 #See note below
```

#The -l flag requires you to set how long and what resources you are requesting. Please see <https://kb.iu.edu/d/bdkd> for information on this setting.

```
#Load modules required
module load java/1.7.0_40
module load fastqc
```

```
#you must enter the directory where your data is, with an absolute path
cd /N/dc2/scratch/ss93/Analysis
```

```
#call the program as you would normally
fastqc NewData.fq
```

Managing your jobs

What is the status of my job (for example if you are me/ss93)?

qstat -u ss93

NOTE: q - waiting in the queue, r - running, c - complete

Delete a job (the number is listed when you use qstat

qdel 60123

What's going on with my job?

Your job will create a file, entitled whatever you used in the -N flag above (i.e. RunTrimmerMagna1) followed by a .o and the job number (i.e. RunTrimmerMagna1.o23453). You will also see a .e file of the same sort (i.e. RunTrimmerMagna1.e23453). These are your STDOUT and STDERR outputs, respectively. They will give you the output on the job

Tips

- Try running commands before putting them in a job file. This ensure they will run correctly. You can always stop them with ctrl c.
- Create a blank job file and just copy it to another file name when you are creating a new job submission
- Have some sort of naming paradigm for jobs. I call all mine RunWhatever. Some people give them a .job extension. It doesn't matter what you do, but it helps when you are looking for them in your folders.
- You can run multiple things in parallel if you put an '&' at the end of each command and then 'wait' at the end of the file. This will run all commands in the background at the same time, and wait until all background jobs are done before terminating. If you forget 'wait' the job will launch, push everything to the background, run out of code to run, and terminate. If you use &, you must use 'wait'!

For more in depth information, try the KnowledgeBase:

<https://kb.iu.edu/index.html>