

Day2 AM Practical

“Hello world!” example

Open the Automatic QSUB Generator at <https://hydra-3.si.edu/tools/QSubGen/>

Now let's fill out our QSUB script:

1. Choose the compute time. Since our first job will be short, we can choose the short queue. Do this by selecting “short” from the drop down menu.
2. Don't worry about selecting a parallel environment because the 'serial' queue is already selected and we will be running a serial job.
3. Select the shell for the job. The default shell for all of hydra's biology users is bash. Please select it.
4. Skip the module section for now. We don't need anything special for this job.
5. Fill in the 'job specific' commands. For our first job, we will simply ask the program to output “Hello world!” to the stdout. You can do this with by entering `echo Hello world!`.
6. Now type in helloworld for your job name.
7. Now check 'Join stderr & stdout' and 'Change to cwd' (cwd stands for 'current working directory')
8. Ignore email for now.
9. Now proceed to the bottom of the script and click on `Check if OK`
10. If your job is validated you are good to go and can click on `Save`

Now we will upload our script:

1. Rename your file to helloworld.qsub and save it to your desktop.
2. Now upload it to hydra by opening your terminal window and entering `cd ~/Desktop` (on Mac) and issuing the command `scp helloworld.qsub frandsenp@hydra-login01.si.edu:/pool/cluster0/workshop/<username>`

Now we will run our job:

1. Login to hydra by typing in `ssh <username>@hydra-login01.si.edu`
2. Once in hydra change to your pool directory with the `cd /pool/cluster0/workshop/<username>` command.
3. Once in this file, you can submit your job using `qsub helloworld.qsub`.
4. If your file was submitted correctly you will receive confirmation that your job was submitted.
5. Wait a minute or two, then check if the logfile was written by using `ls`. There should be a file called `helloworld.job`.
6. Look at it using `less helloworld.job`. You should be able to see your Hello world output.

RAxML example

Open the Automatic QSUB Generator at <https://hydra-3.si.edu/tools/QSubGen/>

Now we will fill this out with specifics for RAxML

1. Choose the compute time. Since this job will also be short, we can choose the short queue. Do this by selecting “short” from the drop down menu.
2. We will be using a the “pthreads” version of RAxML so select multi-thread and enter 4 for the number of CPUs.
3. Select the shell for the job. The default shell for all of hydra's biology users is bash. Please select it.
4. Begin typing 'raxml' into the module window. A dropdown should appear with the `bioinformatics/raxml/8.1` module. Please click on this to select it.
5. Fill in the 'job specific' commands. Some sample raxml commands would be:
`raxmlHPC-PTHREADS-SSE3 -T $NSLOTS -s primates.dna.phy -n raxmltest -m GTRGAMMA -f a -N 100 -p 83487 -x`
The numbers after `-p` and `-x` are just random numbers to seed the parsimony starting tree search and the rapid bootstrap search.
6. Now type in raxmljob1 for your job name.
7. Now check 'Join stderr & stdout' and 'Change to cwd' (cwd stands for 'current working directory')

8. Check the 'Send email notifications' box and enter your email into the field.
9. Now proceed to the bottom of the script and click on `Check if OK`
10. If your job is validated you are good to go and can click on `Save`
11. Change your filename to `raxmltest.qsub`.

Now upload your file to hydra. Hint: you can use a similar method that you did with the 'Hello world!' example.

Now login to hydra and copy the `primates.dna.phy` file into your folder. You can do this by changing to your `/pool` directory and issuing the command `cp /pool/cluster0/workshop/Hydra_workshop_2015/Day2/data/primates.dna.phy .`

Now submit your job by issuing the command `qsub raxmltest.qsub`.

Once the job completes (you can check on it by issuing the `qstat` command), we will tar and download the data using the following commands:

1. `mkdir raxmlout`
2. `mv RAx* raxmlout`
3. `tar -cvzf raxmlout.tar.gz raxmlout`

open a new terminal window (command `t` (MAC), `control t` (Linux))

1. `cd ~/Desktop`
2. `scp <username>@hydra-login02.si.edu:/pool/cluster0/workshop/<username>/Day2/data/raxmlout.tar.gz .`

Once the download has finished issue the command: `tar -xvzf raxmlout.tar.gz`

Once you have confirmed that your data has been downloaded, go back to Terminal connected to hydra and clean up after yourself.

1. `rm raxmlout.tar.gz`
2. `rm -r raxmlout`