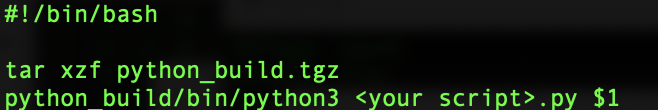
**Open Science Grid Setup Tutorial**

Last updated: 11/15/2022

To start: ask David Zurek for an OSG account. Other questions related to the OSG can be directed to Sajesh Singh ([ssingh@amnh.org](mailto:ssingh@amnh.org)), who is the IT person who knows everything about OSG.

Prepare to submit jobs:

1. Create a folder with all the following files (assuming you are running python): [fib.sub](https://github.com/lyx12311/osg_tutorial/blob/main/fib.sub), [wrapper.sh](https://github.com/lyx12311/osg_tutorial/blob/main/wrapper.sh), and [python\_build.tgz](https://zenodo.org/record/7324844/files/python_build.tgz?download=1).
2. wrapper.sh is the shell scrip to run your code, which typically contains the following:
   1. “tar xzf python\_build.tgz” to unzip the python folder
   2. “python\_build/bin/python3 <your script>.py $1” to run the python script on $1, which $1 will be obtained from the fib.sub file.



1. fib.sub is the submission file, the main things to change are:
   1. This part is typically at the end of the fib.sub file. It indicates the arguments to pass down into the shell script.
      1. “queue fib from <filename>” this will execute wrapper.sh on every node with each line in <filename> as argument.

E.g., if <filename> includes 3 lines that are:

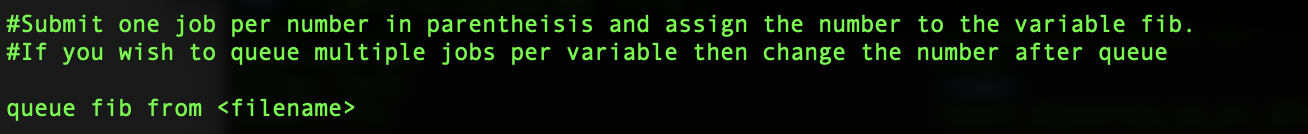
aaa

bbb

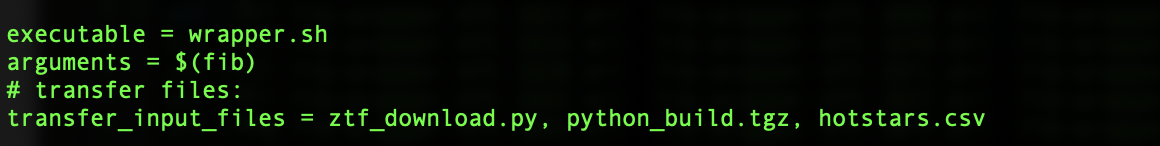
ccc

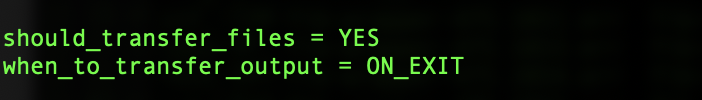
It will execute ./wrapper.sh aaa, ./wrapper.sh bbb, ./wrapper.sh ccc, on different nodes.

\* This is very good for repeating jobs, e.g., each line is a light curve file you want to process.



* 1. This part is typically at the beginning of the file.
     1. “executable = wrapper.sh” is the the executable shell script.
     2. “arguments = $(fib)” is the argument that is passed to the shell script. $1 in the shell script in this case, and fib is typically taken from a text file. It will run.
     3. The last line represents the files you want to transfer to the node, it will typically include all the files you need to run the code besides the wrapper.sh and fib.sub file.



* 1. This part is typically in the middle of the file, it will indicate whether you want to transfer the output data to your home where you submit the script. 

Submit jobs:

condor\_submit fib.sub

\*Note: make sure your file has less than 5000 lines as you can only submit 5000 jobs at once.

Cheat sheet:

* To check your jobs:
  + condor\_q <username>
* If your job is on hold and you want to know the reason:
  + condor\_q -better-analyze <job\_ID>
* If you fixed the problem that holds your job and you want to release them:
  + condor\_release <username>
* To install new packages in your python:
  + Extract the tgz file for python\_build
    - tar xzvf python\_build.tgz
  + Run the following command:
    - python\_build/bin/pip3 install <PACKAGE NAME>
  + Recreate the tgz file
    - tar czvf python\_build.tgz python\_build

To cite OSG:

This research was done using services provided by the OSG Consortium \citep{OSG1,OSG2}, which is supported by the National Science Foundation awards \#2030508 and \#1836650.

@inproceedings{OSG1,

title = {The open science grid},

author = {

Pordes, Ruth

and Petravick, Don

and Kramer, Bill

and Olson, Doug

and Livny, Miron

and Roy, Alain

and Avery, Paul

and Blackburn, Kent

and Wenaus, Torre

and W{\"u}rthwein, Frank

and Foster, Ian

and Gardner, Rob

and Wilde, Mike

and Blatecky, Alan

and McGee, John

and Quick, Rob

},

doi = {10.1088/1742-6596/78/1/012057},

booktitle = {J. Phys. Conf. Ser.},

volume = {78},

series = {78},

pages = {012057},

year = {2007},

}

@inproceedings{OSG2,

title = {The pilot way to grid resources using glideinWMS},

author = {

Sfiligoi, Igor

and Bradley, Daniel C

and Holzman, Burt

and Mhashilkar, Parag

and Padhi, Sanjay

and Wurthwein, Frank

},

doi = {10.1109/CSIE.2009.950},

booktitle = {2009 WRI World Congress on Computer Science and Information Engineering},

volume = {2},

series = {2},

pages = {428--432},

year = {2009},

}