How to run jupyter notebook on a cluster

Date: 05/29/23

Step 1: Make sure jupyter-notebook is installed and in your system path so as to be recognized whenever we try running jupyter notebook on the server

```
1.1. Check it's installed:

    -> Eg., check that in your local file if it has been installed: /home/.local/bin
    -> If not installed then install it using 'pip install jupyter notebook'

1.2. Setting the Path

    -> Open the source file. Eg: vim .cshrc
    -> Set the path to the jupyter notebook program. Eg.: setenv PATH ${PATH}:/home/<username>/.
local/bin
    -> Run source file: source .cshrc
```

Step 2: Create and run batch file [bash_file].sh inside the same directory where you wish to run/open the jupyter notebook. The file should include the following:

```
#!/bin/bash
#BATCH -o jupyter-notebook-%j.log
#SBATCH -N 1
#SBATCH --ntasks-per-node=16
#SBATCH --mem=40G
#SBATCH --time=100:00:00
#SBATCH --job-name=jupyter-notebook
```

```
#SBATCH --mail-type=ALL
#SBATCH --mail-user=<username>@uh.edu

# get tunneling info
XDG_RUNTIME_DIR=""
node=$(hostname -s)
user=$(whoami)
cluster="<cluster> " #This can be carya,juniper,sabine,opuntia,sabine
port=<port_number> #This can be any 4 digit number (eg 8889). Note that this is important since the user can not reuse the same port number on multiple clusters
# print tunneling instructions jupyter-log
echo -e "

Command to create ssh tunnel:
ssh -N -f -L ${port}:${node}:${port} ${user}@${cluster}.rcdc.uh.edu

Use a Browser on your local machine to go to:
localhost:${port}  (prefix w/ https:// if using password)
token will show up shortly
"
# load modules or conda environments here if needed
module load python/3.9
# Run Jupyter
jupyter notebook --no-browser --port=${port} --ip=${node}
```

In the file above make sure to update:

- < username > #Your login username. Eg. klkusima
- < cluster > #Name of your cluster. Eg. carya
- < port_number > #The desired port number. Eg. 8889

Step 3: Submit the batch file sbatch .sh

- Note: The batch file should be save and ran from the folder/directory that contains your desired jupyter notebook file
- Once running that this will start the time (as specified in the wall clock) that the jupyter files can run. The jupyter notebook won't open if the job is pending and/or if it hasn't started yet.
- A slurm-<jobID>.out file will be created that has the link to the website and the corresponding token needed to access the cluster in your local computer

see example .out output file:

```
Command to create ssh tunnel:
ssh -N -f -L 8889:compute-5-10:8889 klkusima@carya.rcdc.uh.edu

Use a Browser on your local machine to go to:
localhost:8889  (prefix w/ https:// if using password)
token will show up shortly

/var/spool/slurm/slurmd/job1056332/slurm_script: line 28: module: command not found
[W 00:31:43.269 NotebookApp] Error loading server extension jupyter_lsp
    Traceback (most recent call last):
      File "/home/klkusima/.local/lib/python3.9/site-packages/notebook/notebookapp.py", line 2050, in init_server_extensions
        func(self)
      File "/home/klkusima/.local/lib/python3.9/site-packages/jupyter_lsp/serverextension.py", line 76, in load_jupyter_server_extension
        nbapp.io_loop.call_later(0, initialize, nbapp, virtual_documents_uri)
    AttributeError: 'NotebookApp' object has no attribute 'io_loop'
[W 2023-05-30 00:31:44.127 LabApp] 'port' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.
[W 2023-05-30 00:31:44.127 LabApp] 'ip' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.
[W 2023-05-30 00:31:44.127 LabApp] 'ip' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.
[W 2023-05-30 00:31:44.127 LabApp] 'ip' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.
[I 2023-05-30 00:31:44.130 LabApp] JupyterLab extension loaded from /home/klkusima/.local/lib/python3.9/site-packages/jupyterlab
[I 2023-05-30 00:31:44.130 LabApp] JupyterLab application directory is /home/klkusima/.local/share/jupyter/lab
```

```
[I 2023-05-30 00:31:44.131 LabApp] Extension Manager is 'pypi'.
[I 00:31:44.155 NotebookApp] Serving notebooks from local directory:
/project/grabow/klkusima/work/KMC/CO_Oxidation/Jupyter/KMC_output_varied_coverage
[I 00:31:44.155 NotebookApp] Jupyter Notebook 6.5.4 is running at:
[I 00:31:44.155 NotebookApp] http://compute-5-10:8889/?
token=c84a59cff8c5ca8bcef897b94fe2fb70a67747b498c88f98
[I 00:31:44.155 NotebookApp]  or http://127.0.0.1:8889/?
token=c84a59cff8c5ca8bcef897b94fe2fb70a67747b498c88f98
[I 00:31:44.155 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to
skip confirmation).
[C 00:31:44.172 NotebookApp]
..................
..................
..................
```

***** Steps below need to be repeated every time you need to open up the jupter job running on the server *****

Step 4: Create a tunnel on the local computer to contact the server's running job using the ssh prompt provided in the slurm output file. This establishes a bridge between the server and the local computer. The command to do this can be found in the beginning of the .out output file.

```
From the example above, run inside your local terminal:
    ssh -N -f -L 8889:compute-5-10:8889 klkusima@carya.rcdc.uh.edu

 * Note that this step may need to be repeated if your computer has restarted
```

Step 5: Copy the link (localhost:< port_number >) from the output file to a browser

-Side Note:-

- For the duration of the wall clock time, you will be able to access a jupyter notebook in the folder you specified. Once the wall clock time is over, you will need to restart from step 1.
- Displaying figure may not work in the cluster's version of jupyter notebook, below are the possible workarounds:

  i) apply **%matplotlib inline** at the very first line of your jupyter notebook to force matplolib to be used to plot the figure inline

  ii) save the figure as a png so you can open it externally. The code: **plt.savefig("plot1.png")**