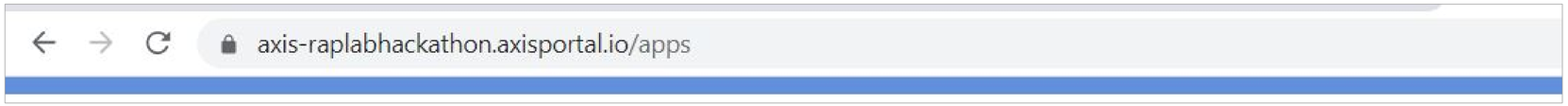
Cluster Login Instructions

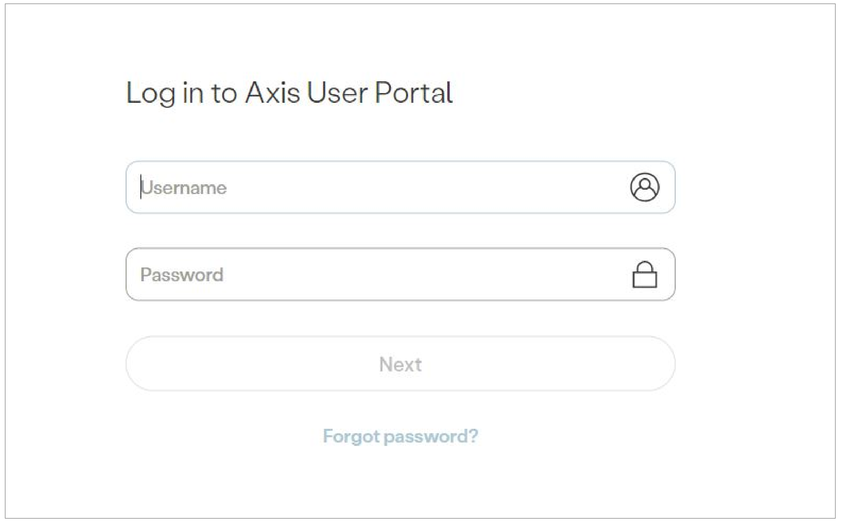
Login Instructions

Login Credentials

- Step 1:
  - Open the browser and enter the login URL: **https://axis-raplabhackathon.axisportal.io/apps**



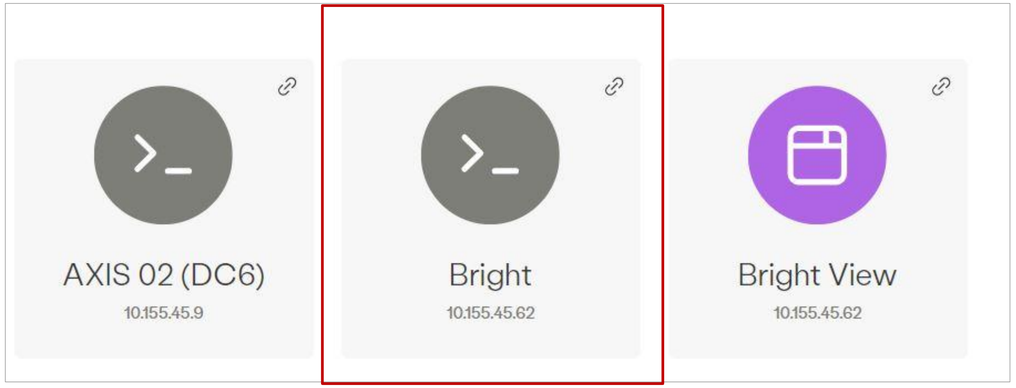
- Step 2:
  - Enter your cluster login credentials



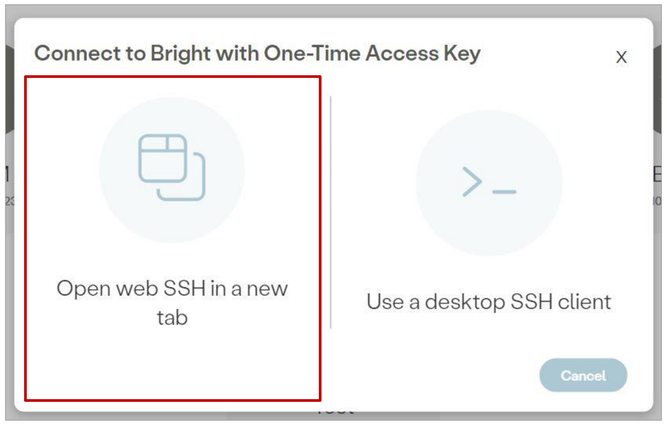
Login Instructions

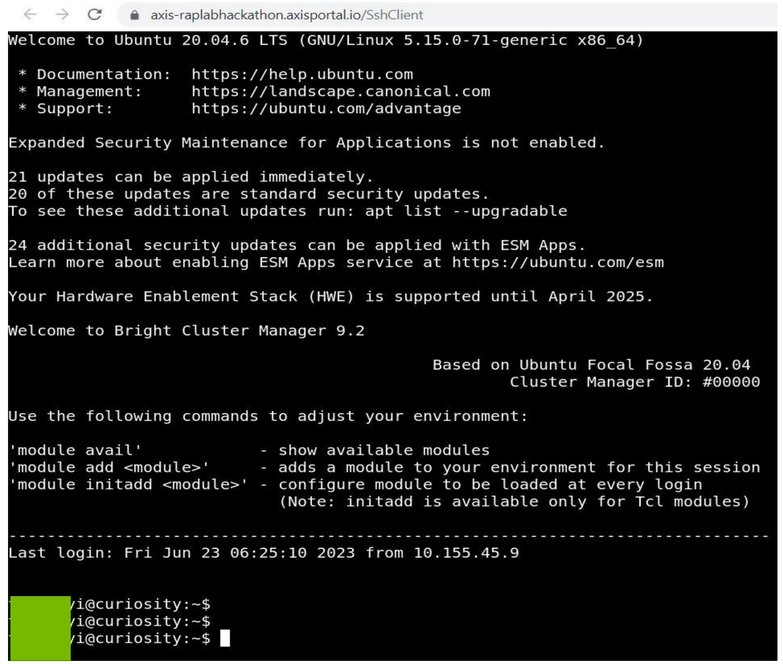
Select Bright option

- Step 3:
  - Click on **Bright**



- Step 4:
  - Click on **Open web SSH in a new tab**

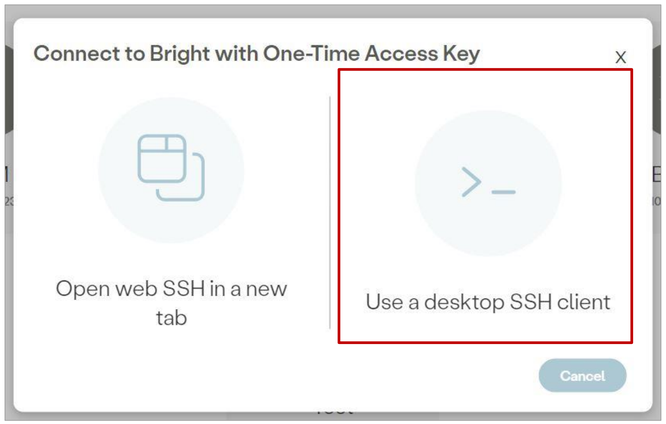


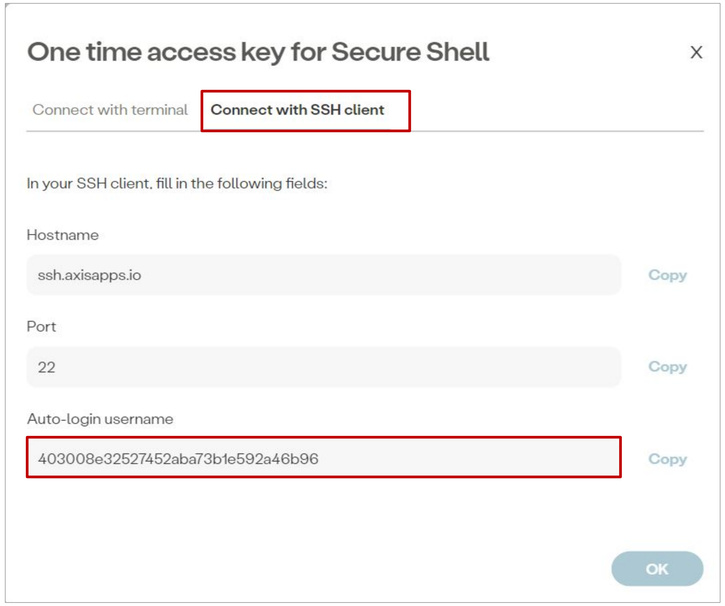


Login Instructions

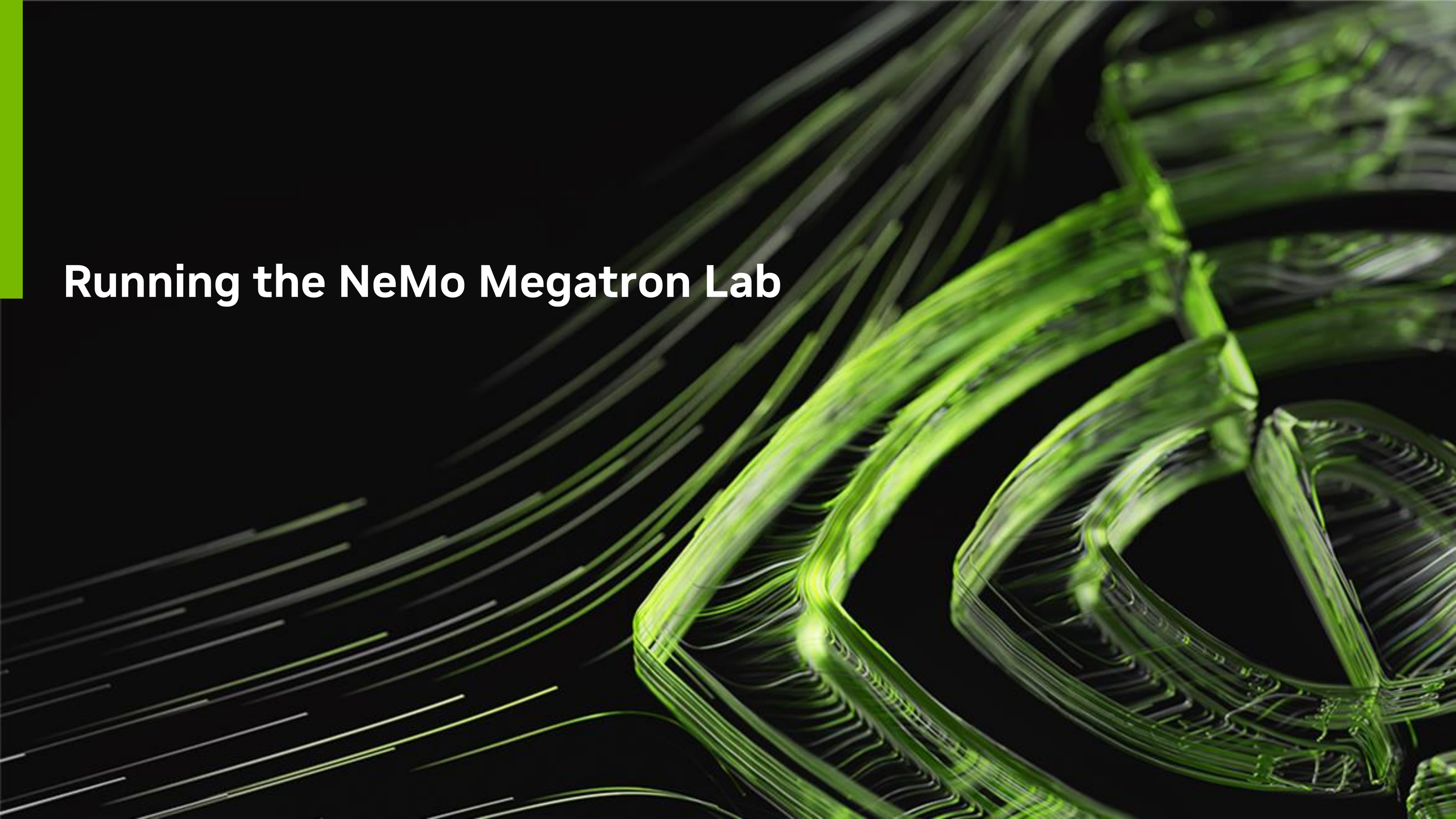
Copy the Hashcode

- Step 5:
  - Click on the option **Use a desktop SSH client.** Click on **Connect with SHH client** and copy the hashcode





Running the NeMo Megatron Lab



NeMo Megatron-GPT Lab

Running the Lab Script and Port-forwarding (1/2)

- The Nemo Megatron lab contains the following hands-on notebooks:
  1) Nemo Fundamentals
  2) Question Answering
  3) Prompt-Tuning/P-Tuning
  4) The Megatron-GPT Language Model Inferencing
  5) Challenge

Steps:

- From the browser terminal, run the sbatch script and press on the Enter key to submit job:

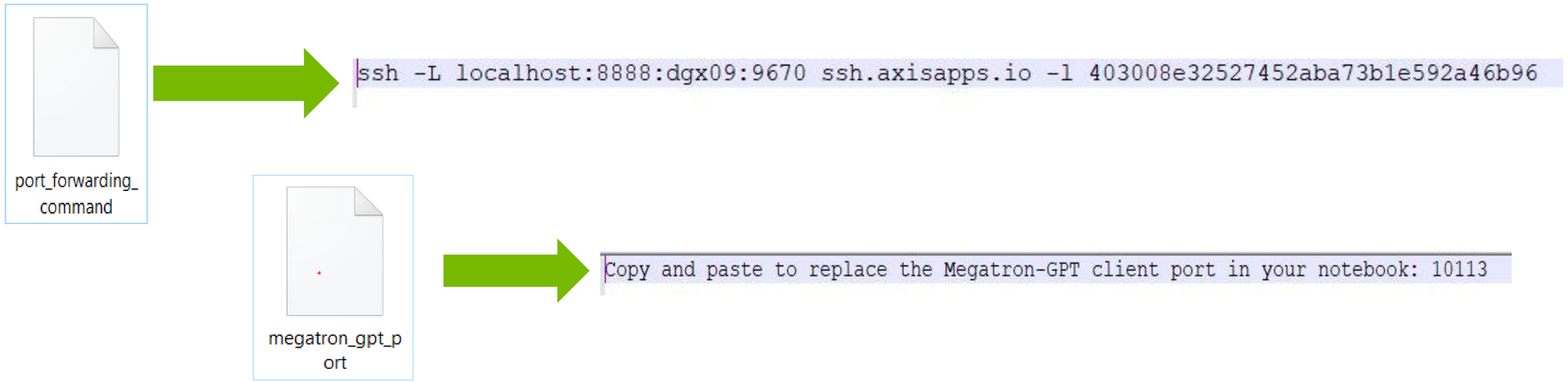
sbatch /bootcamp_scripts/NLP/**sbatch-llm** <hashcode>

```
yi@curiosity:~$ sbatch /bootcamp_scripts/NLP/sbatch-llm cad857f39ec14668b098892bef85cc3
```

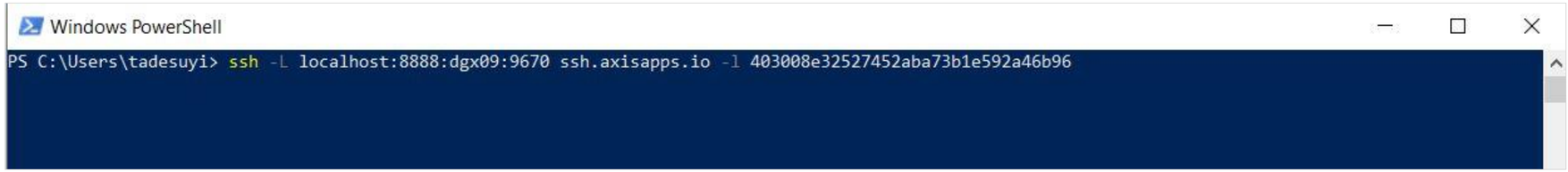
NeMo Megatron-GPT Lab

Running the Lab Script and Port-forwarding (2/2)

- In approximately **8mins** a port forwarding (port_forwarding_command) and a megatron-gpt inferencing port (megatron_gpt_port) files will be created in your home directory.
  - To see the files created, type the command this into the terminal: ls
  - To display the **port_forwarding_command** file content, type the command: cat port_forwarding_command
  - To display the content of **megatron_gpt_port** file, type the command: cat megatron_gpt_port



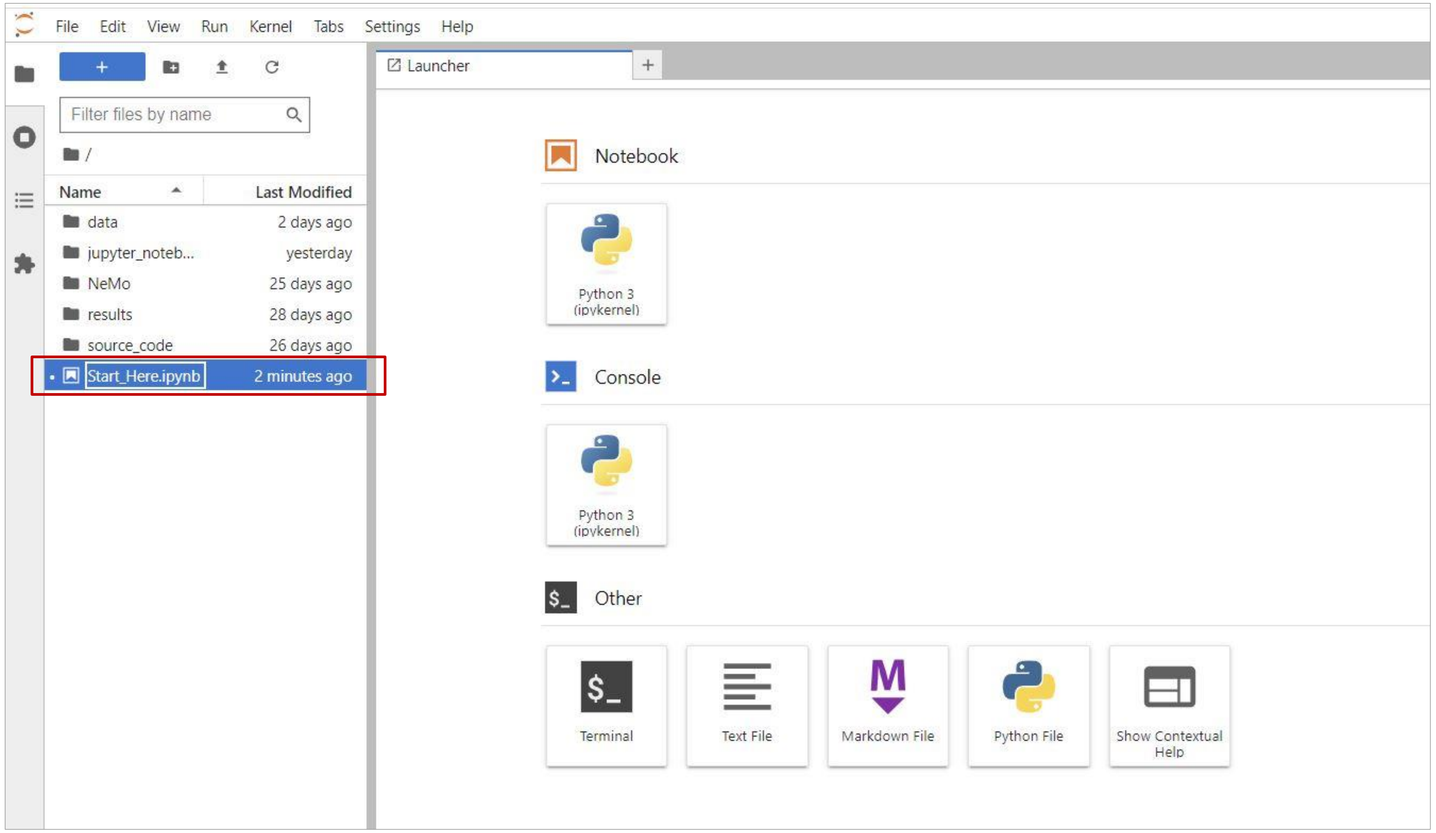
- Copy and paste the content of the port-forwarding file into your local machine or laptop terminal (e.g. CMD prompt, PowerShell, Terminal, etc.). Press on the Enter key.



NeMo Megatron Lab

Running the Lab

- Run the notebooks via the browser :
  - Open a new tab on your browser at http://localhost:8888
  - Click on the Start_Here.ipynb notebook



NeMo Megatron Lab

Start with Nemo Fundamentals

- Under the “NeMo Megatron,” Click on Nemo Fundamentals.

Table of Content

The following contents will be covered:

Troubleshooting

- Everyone can only run 1 job at a time.
- The jobs will run for 12 hours so you do not need to re-submit the sbatch script. All you need is the content of the "port_forwarding_command" to reconnect to the cluster.
- Each person will get access to 1 GPU only.
- The Hash key is valid for 24 hours.
- Do not use VPN or eduroam
- The access will be revoked by the end of the day.

If you get the following error in the final Step :

```
bind [127.0.0.1]:8888: Address already in use
channel_setup_fwd_listener_tcpip: cannot listen to port: 8888
Could not request local forwarding.
```

You will need to change the port 8888 in your ssh command to some other port, examples would be 8890, 8900 etc.

For example: `ssh -L localhost:8890:dgx05:9538 ssh.axisapps.io -l 7aab57129c95450cb80c4df5c13836bd`

Troubleshooting

- Useful commands:
  - Check the number of jobs submitted via squeue --me
  - Cancel the submitted job via scancel {job id}
  - Check the slurm output via cat slurm-{job id}.out
- To check if the new “`port_forwarding_command`” has been created, run `ls port_forwarding_command`
- Type the commands instead of copy-pasting to avoid errors.
- Please direct the questions to the “**cluster-support**” channel on the Slack.