Python HPC School
Setting up a Jupyter notebook on an HPC node

Oscar J. Castro-Lopez and THE UL-HPC Team

December, 2023

University of Luxembourg

Introduction

Introduction to the Workshop

Welcome to the workshop! We will explore techniques to significantly improve the performance of Python code using parallel programming.

Introduction to the Workshop

Welcome to the workshop! We will explore techniques to significantly improve the performance of Python code using parallel programming.

Python is known for its ease of use and versatility, but it may not always be the fastest choice for computationally intensive tasks.

Introduction to the Workshop

Welcome to the workshop! We will explore techniques to significantly improve the performance of Python code using parallel programming.

Python is known for its ease of use and versatility, but it may not always be the fastest choice for computationally intensive tasks.

We'll focus on practical techniques and tools for harnessing the power of parallel computing to boost Python performance.

Key Topics We'll Cover:

1. Setup a Jupyter notebook with the HPC cluster
2. Taking time and profiling python code
3. Optimizing with Numpy broadcasting
4. Using multiple cores with Python packages

Key Topics We'll Cover:

1. Setup a Jupyter notebook with the HPC cluster
2. Taking time and profiling python code
3. Optimizing with Numpy broadcasting
4. Using multiple cores with Python packages

By the end of this workshop, you'll have the knowledge and tools to make your Python applications faster and more efficient using these parallel computing techniques.

Let's get started!

Material

- To set up Jupyter, please refer to the accompanying presentation slides.
- Subsequent topics will be covered using Jupyter Notebooks.

Setup a Jupyter notebook with the HPC cluster

Jypyter notebook

Jupyter is a flexible, popular literate-computing web application for creating notebooks containing code, equations, visualization, and text. Notebooks are documents that contain both computer code and rich text elements (paragraphs, equations, figures, widgets, links).

This tutorial is based on the tutorial available in the ULHPC technical documentation:

`https://hpc-docs.uni.lu/services/jupyter/`

Connecting to the cluster

On a terminal type the following:

```
ssh aion-cluster
```

Then it would show something like the following prompt:

```
Enter passphrase for key '/home/my_user/.ssh/id_rsa':
```

Input your password and you should be connected to the aion cluster and see the welcome message.

Install Jupyter

We are going to:

1. Start an interactive session.
2. Create a virtual environment and use it.[1]
3. Update the pip package manager.
4. Install Jupyter.

[1] for more details on venv in `https://docs.python.org/3/library/venv.html`

Install Jupyter

We are going to:

1. Start an interactive session.
2. Create a virtual environment and use it.[1]
3. Update the pip package manager.
4. Install Jupyter.

To do that we are going to use the following commands:

```
si -t70 --reservation=pyschool-interactive
python -m venv jupyter_env
source jupyter_env/bin/activate
python -m pip install --upgrade pip
python -m pip install jupyter ipykernel
```

[1] for more details on venv in `https://docs.python.org/3/library/venv.html`

Configure Jupyter

1. After installing Jupyter and IPython, you can proceed to configure your installation by setting environment variables to match your specific requirements.
2. However, for the purpose of this tutorial, we will utilize the default settings.
3. Our primary configuration task will be to set up a password for your Jupyter environment.

Configure Jupyter

1. After installing Jupyter and IPython, you can proceed to configure your installation by setting environment variables to match your specific requirements.
2. However, for the purpose of this tutorial, we will utilize the default settings.
3. Our primary configuration task will be to set up a password for your Jupyter environment.

In a terminal, run the following command and setup a password:

```
jupyter notebook password
```

Use a password you are going to remember or store it somewhere.

To customize your installation, please take a look at:
`https://jupyter-notebook.readthedocs.io/en/5.7.5/index.html`

Register the kernel

Jupyter Notebook makes sure that the IPython kernel is available, but you have to manually add a kernel with a different version of Python or a virtual environment.

Register the kernel using the following command:

```
python -m ipykernel install --sys-prefix --name jupyter_env
```

Jupyter and the virtual environment are now installed and ready.

Starting a Jupyter Notebook

You can initiate Jupyter as a SLURM batch job by using the following script:

```
#!/bin/bash -l
#SBATCH -J Jupyter_aion_oc       # Name of the job
#SBATCH -N 1
#SBATCH --ntasks-per-node=1
#SBATCH -c 128                  # Cores assigned to each tasks
#SBATCH --time=0-10:00:00
#SBATCH -p batch
#SBATCH --reservation=pyschool-day1          # Change it the second day to pyschool-day2
#SBATCH --mail-user your.email@uni.lu        # Optional if you want to be notified
#SBATCH --mail-type BEGIN,END,FAIL           # Optional if you want to be notified

source jupyter_env/bin/activate

jupyter notebook --ip $(hostname -i) --no-browser  &
pid=$!
sleep 5s
jupyter notebook list
jupyter --paths
jupyter kernelspec list
echo "Enter this command on your laptop: ssh -p 8022 -NL 8888:$(hostname -i):8888 ${USER}@access-${
     SLURM_CLUSTER_NAME}.uni.lu " > notebook.log
wait $pid
```

Starting a Jupyter Notebook (cont.)

For future use, you have the flexibility to tailor the script according to your requirements. However, for the sake of this tutorial, it's recommended to leave most settings as they are. Please only modify the script by changing the job name or email notification preferences.

Starting a Jupyter Notebook (cont.)

For future use, you have the flexibility to tailor the script according to your requirements. However, for the sake of this tutorial, it's recommended to leave most settings as they are. Please only modify the script by changing the job name or email notification preferences.

Now, in a terminal connected to the HPC cluster let's launch the script:

```
sbatch jupyter_launch_aion.sh
```

The name of the file is: `'jupyter_launch_aion.sh'` (change it to the name you used for your file).

If your email was configured you will receive an email when the job begins. Otherwise you can check your jobs with the command `sq`.

Connect to a Jupyter Notebook (forwarding with linux)

Once your job is running, we are going to use `ssh` forwarding to connect to the notebook from your laptop.

Open a new terminal (outside the cluster front-end) and copy-paste the contents of the file notebook.log.

```
cat notebook.log
```

This is going to print something like the following:

```
Enter this command on your laptop:  ssh -p 8022 -NL 8888:172.21.13.1:8888
yourusername@access-aion.uni.lu
```

Copy and paste the text from “ssh...“ to the end in a terminal.

Connect to a Jupyter Notebook (forwarding with Windows/mobaxterm)

Once your job is running, we are going to use `ssh` forwarding to connect to the notebook from your laptop with MobaXterm.

We are going to use the same information from the output of notebook.log. Let's suppose the output is:

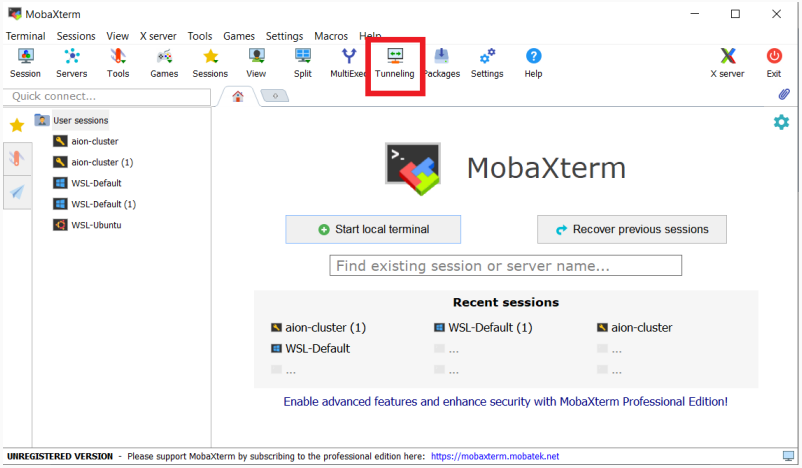
```
Enter this command on your laptop:  ssh -p 8022 -NL 8888:172.21.13.1:8888
yourusername@access-aion.uni.lu
```

This means that:

- Forwarder port is "8888".
- Remote server IP "172.21.13.1".
- Remote port "8888".
- SSH Server is "access-iris.uni.lu".
- SSH Login is "yourusername".
- SSH Port is "8022".

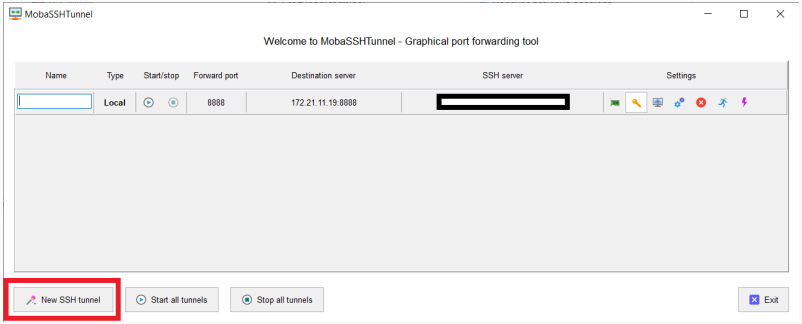
Forwarding with MobaXterm pt 1

1. On MobaXterm windows Open MobaSSHTunnel (click on Tunneling):



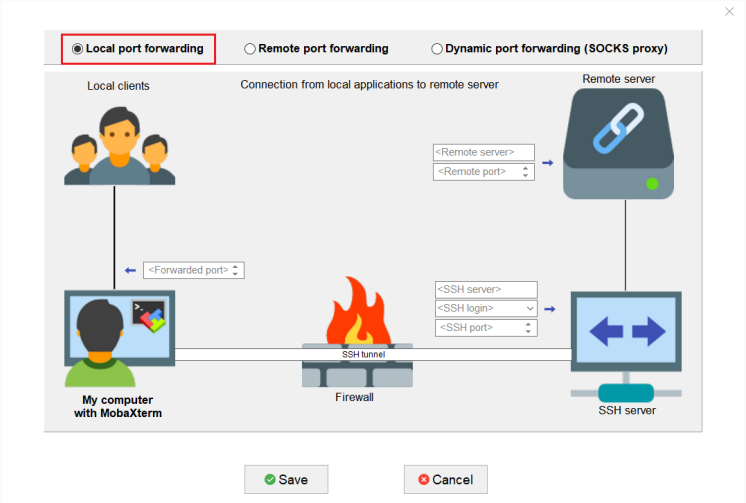
Forwarding with MobaXterm pt 2

2. Create New SSH tunnel:



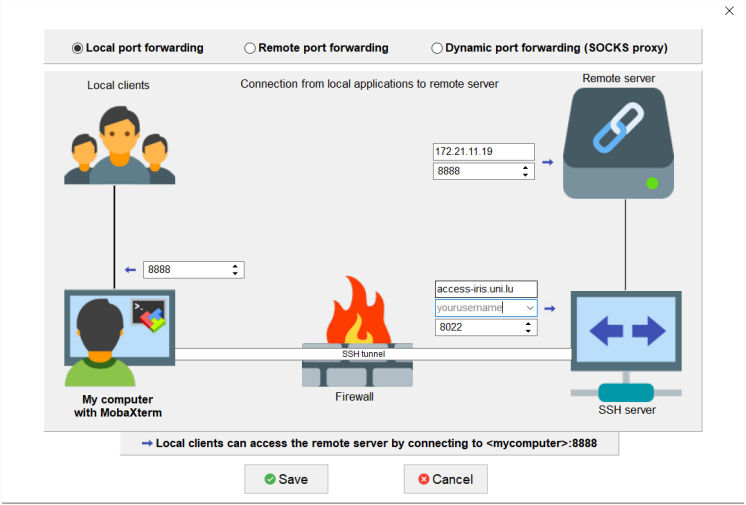
Forwarding with MobaXterm pt 3

3. The following window is going to open:



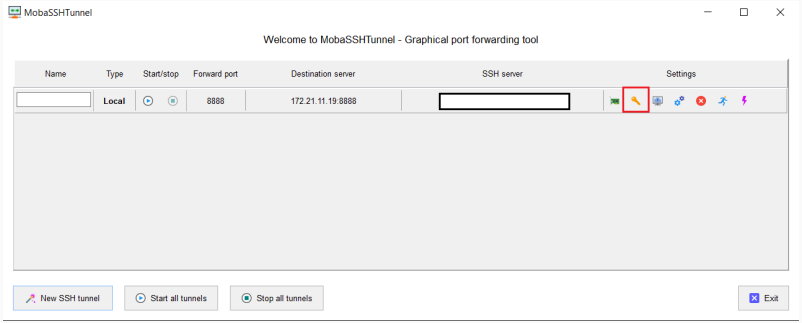
Forwarding with MobaXterm pt 4

4. Type the values taken from the output of notebook.log:



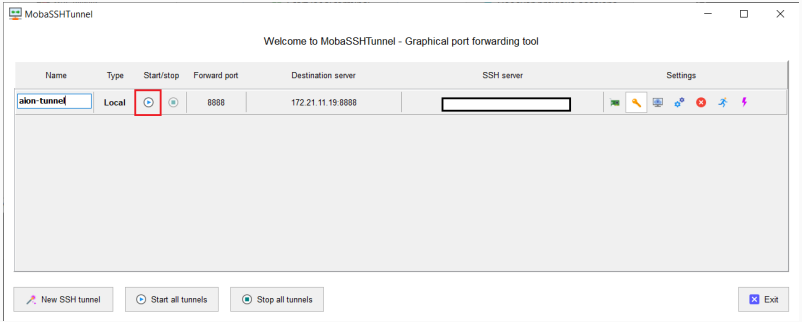
Forwarding with MobaXterm pt 5

5. Add the SSH key:



Forwarding with MobaXterm pt 6

6. Click on the start button (optionally set a name for your forwarding):



Connect to a Jupyter Notebook (cont.)

To access the Jupyter Notebook interface, follow these steps:

1. Open your web browser.
2. Enter the following URL in the address bar: `http://127.0.0.1:8888/`
3. Press Enter to navigate to the provided URL.
4. Jupyter Notebook should request a password. Enter your password in the provided field.
5. If the password is correct, you will be granted access to the Jupyter Notebook interface, where you can view your files and begin working.

Download the notebooks from Github

To download the notebooks for the session, clone the repository from Github:

```
git clone https://github.com/ULHPC/python-school.git
```

Optionally you can direct download them from the URL
`https://github.com/ULHPC/python-school`

Extra

Useful commands

Exit a virtual environment:

```
source deactivate
```

Delete a virtual environment (folder name equal to the name of the venv):

```
rm -r folder_name
```

Terminate a job in the cluster

```
scancel job_id
```

End

Thank you!

oscar.castro@uni.lu