JupyterHub Lunchbox session

Vlaanderen
is supercomputing

VLAAMS
SUPERCOMPUTER
CENTRUM

Innovative Computing
for A Smarter Flanders

vscentrum.be

Poll

https://pollev.com/wouterlampaert916

I intend to access JupyterHub ...

to develop CPU/GPU software

to pre- or post-process data interactively

to create visualizations

for educational purposes

I am ...

already using Jupyter Notebooks for my work

planning to use Jupyter Notebooks in the future

How frequently do you use Jupyter Notebooks?

Rarely

Few times a week

Every day

General – Best practices

- Easy use of Jupyter Notebooks on the HPC facilities
  - Interactive session

- Best used for:
  - Data exploration
  - Visualization
  - Education
    - According to standard rules for HPC courses
  - Prototyping/development

- **Don’t use with:**
  - Version control
  - Complicated dataflows

Getting started

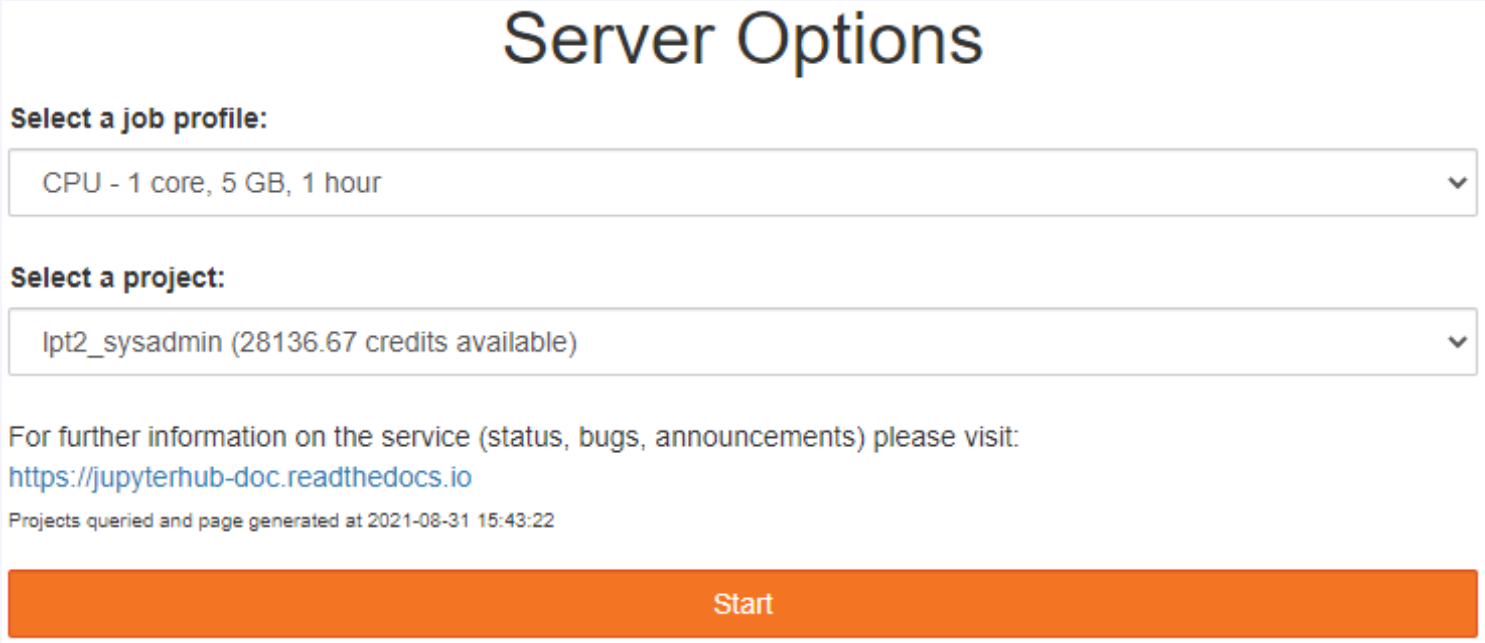
- Open your browser and go to https://jupyterhub.hpc.kuleuven.be/hub

- Select the required resources
  - 1, 4 or 12h
  - CPU or GPU
  - Best lowest possible resources, increase if necessary

- Select desired project

- Standard location is $VSC_DATA

- If server doesn't start in 5 mins, process will be killed → try again



Kernels

- Two kernels available
  - Python 3.6.8
  - R 3.6.3
    - RStudio also available in NX, but kernel is better for intense calculations

- Limited packages available, recommended to create your own kernel

Creating your own kernel

- R or Python kernel

- Use conda:
  - Install Miniconda in $VSC_DATA:

    Python: https://docs.vscentrum.be/en/latest/software/python_package_management.html#install-miniconda

    R: https://docs.vscentrum.be/en/latest/software/r_package_management.html#installing-miniconda

  - Create new conda environment or use existing environment
    - Install necessary packages
      $ conda install <package_name>

Creating your own kernel: Python

- Activate your environment:
  $ source activate <envname>

- Include the IPython kernel, as well as ipython_genutils
  $ conda install ipykernel
  $ conda install ipython_genutils

- Install environment in *$VSC_HOME/.local* with ipykernel:
  $ python -m ipykernel install --prefix=${VSC_HOME}/.local/ --name '<envname>' --display-name '<name>`

- New packages can be added from within JupyterHub.
  $ conda install <package_name>

  - only one conda command per cell
  - Restarting the kernel is necessary

Creating your own kernel: R

- Activate your environment:
  $ source activate <envname>
- Install IRKernel and jupyter_client:
  $ conda install jupyter_client r-irkernel –c conda-forge
- Install the environment to $VSC_HOME/.local with IRkernel:
  Rscript -e 'IRkernel::installspec(prefix="${VSC_HOME}/.local/", name=" <envname>", displayname=" <display_name>")'
- Conda packages cannot be installed from within Jupyter notebook

Removing a kernel

- Remove kernel in $VSC_HOME/.local
  $ rm $VSC_HOME/.local/share/jupyter/kernels/<kernel_name> -r

- Be careful: removing a conda env and not removing the kernel as above will result in a non-working kernel. No error/warning is given in the Jupyter Notebook

Limitations

- Not possible to load any extra modules.
  - Have a look for a conda/pip alternative if possible.
  - In case you have a specific case where you can only use a specific VSC module, please contact the servicedesk

- Other executables (e.g. Tensorboard) cannot be used in Jupyter at the moment

New kernel installation and notebook examples

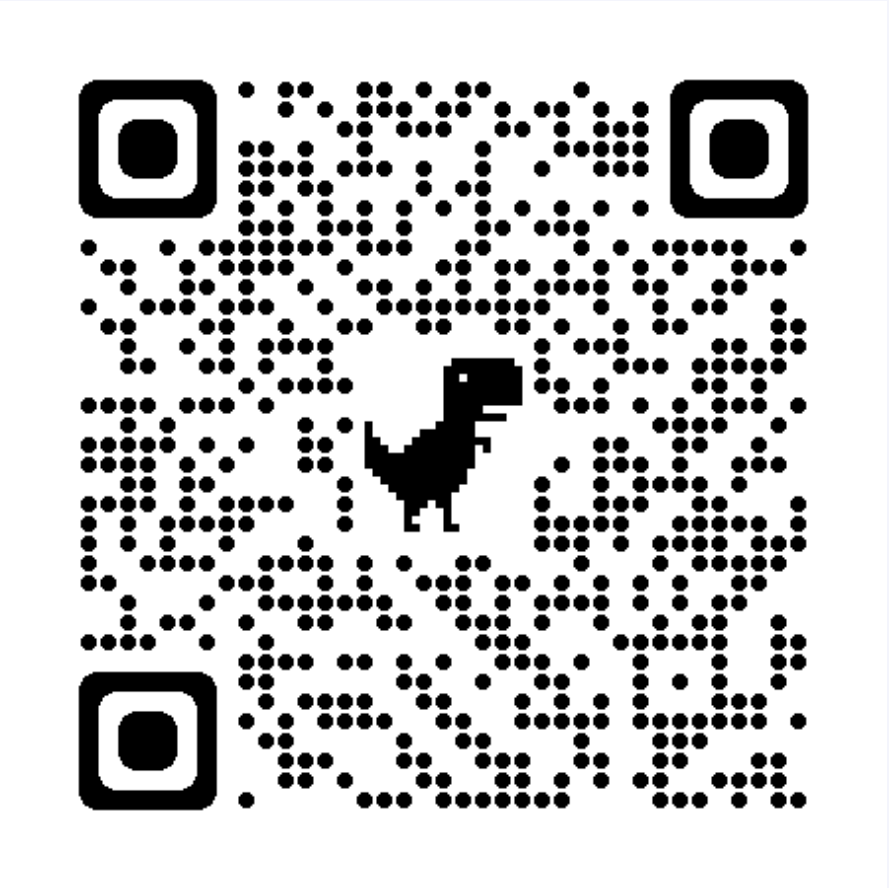
Python

- MNIST on GPU

R

- K-means clustering

Poll



https://pollev.com/wouterlampaert916

I intend to use ...

resources with CPUs only

resources with GPUs

Desired number of cores when using CPU resources:

Not applicable

1

2

4

More

Useful duration of the JupyterHub session

1 hour

4 hours

12 hours

Which kernel do you use?

Python

R

Bash