CCIN2P3

cnrs

Centre de Calcul
de l'Institut National de Physique Nucléaire
et de Physique des Particules

GPUs availability via the Jupyter notebooks platform service

Bernard CHAMBON, April 30th 2021

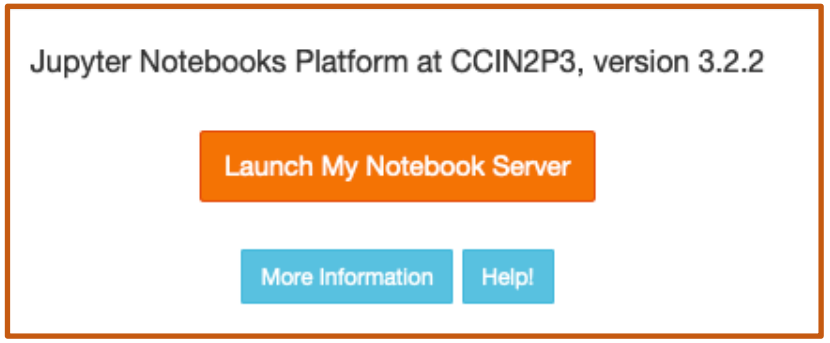
The Jupyter notebooks platform > Reminder

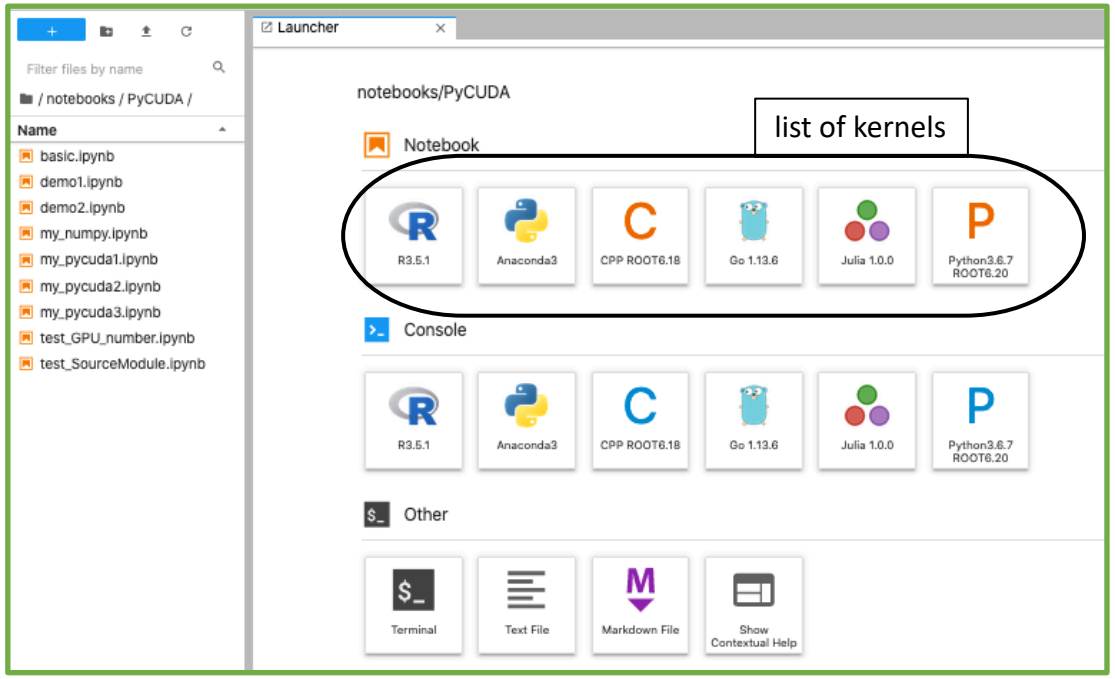
- Objective of this service
  - The Jupyter notebooks platform provides notebooks with JupyterLab interface
  - It allows you to develop & run code interactively by using several programming languages; It provides a UNIX terminal

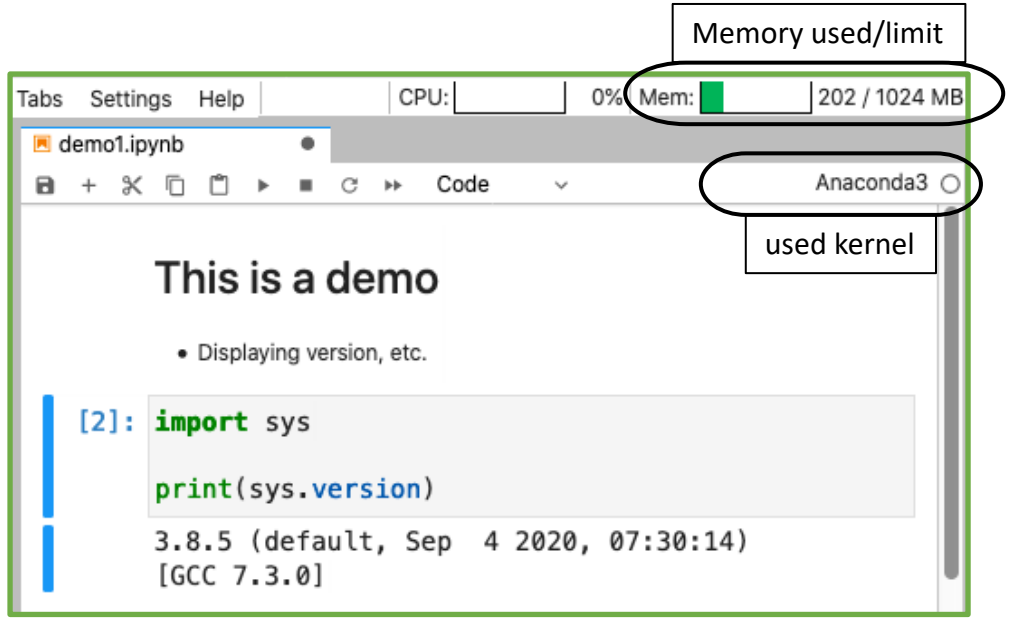
- Access to this service
  - Accessible for all CC-IN2P3 users
  - Authenticating via Keycloak SSO, by using 'computing account' credentials (same as for ssh cca.in2p3.fr)
  - Url https://notebook.cc.in2p3.fr

- Runtime environment
  - You will run JupyterLab with your primary and all secondary groups
  - You will access the following storage areas : /pbs/home/u/username, /pbs/throng, /pbs/software, /sps/xyz, /cvmfs/xyz
  - Available programming language : requires to setup a kernel
    - Anaconda 3 : kernel provided with Docker image, nothing to do.
    - Kernels ready to be used from /pbs/software : ROOT via Python, ROOT via C++ Cling interpreter, R, Go, Julia. See doc Adding kernels
  - Resources control
    - Memory limit set per group or per user. Default is 2GB
    - Idle notebook timeout of 72h

The Jupyter notebooks platform > Reminder







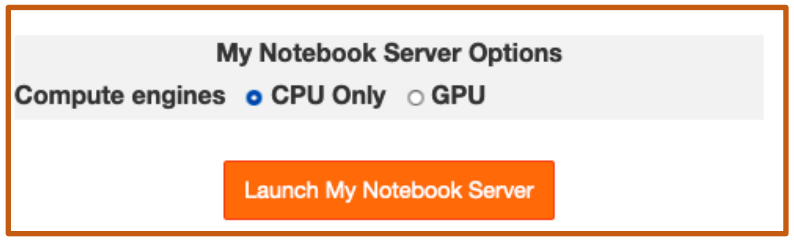
The Jupyter notebooks platform > GPUs

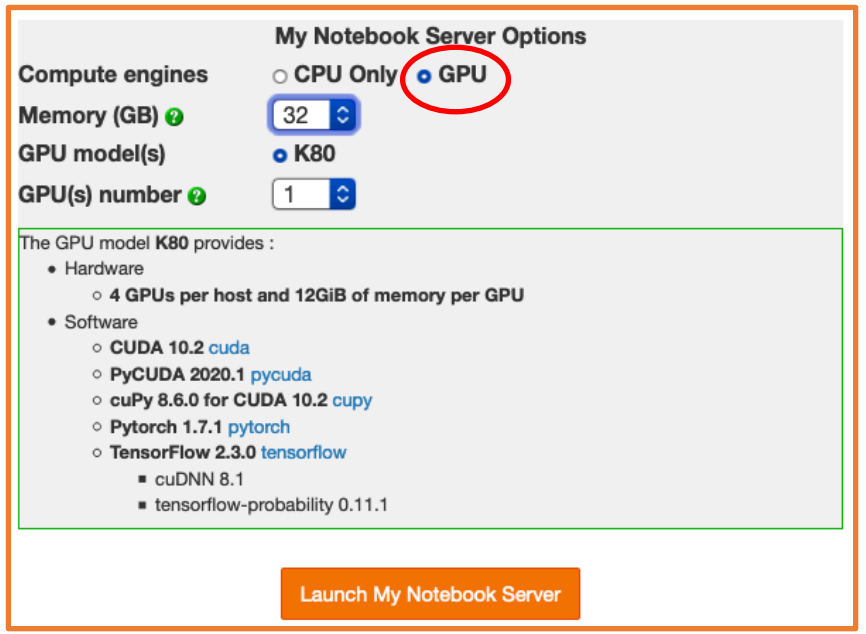
- Objective of this feature
  - Provide notebook running on host with GPUs devices
  - Each user will have it’s own set of GPUs (= NOT shared with other users)

- Access to this feature
  - For granted users only
  - Granted users will get an options form, allowing to select
    - The model and the numbers on GPUs, according to the instant availability
    - The memory limit for the notebooks server

- Runtime environment
  - GPU notebooks server will run on host with GPUs model K80
  - GPU software provided : CUDA, PyCUDA, cuPy, Pytorch, TensorFlow + cuDNN (details of versions from the options form)
  - Same storage areas as seen previously

The Jupyter notebooks platform > GPUs





Memory (32GB = selected limit)

GPUs availability via the Jupyter notebooks platform service

The Jupyter notebooks platform > GPUs

- Status (2021/04/30)
  - One bare metal host : model K80, 4 GPUs per host, RAM 135 GB (125 GiB), Ethernet 10Gbs
  - Docker image with installs for CUDA 10.2, PyCUDA 2020.1, Pytorch 1.7.1, TensorFlow 2.3 + cuDNN 8.1
  - **GPUs availability currently under test, for “beta testers”**

- Planned
  - Another K80, currently being installed
  - Add K80 hosts (from batch production farm), to provide 10 hosts by the end of the year

- How to test
  - Ask UserSupport to get a granted access
  - But for now, there is only one GPU host
  - Log in https://notebook.cc.in2p3.fr Log out : File > Log Out (mandatory to release GPU slot(s))
  - Documentation and Ticketing URLs are available from login page