

A quickstart guide to Jupyter notebooks

Marty Kandes, Ph.D.

Computational & Data Science Research Specialist
High-Performance Computing User Services Group
San Diego Supercomputer Center
University of California, San Diego

CIML Summer Institute
Wednesday, June 22nd, 2022
09:50 AM - 10:10 AM PT

Today's Session

- ▶ What is a Jupyter notebook?
- ▶ Setting up your Jupyter notebook environment
- ▶ How-to run Jupyter notebooks on Expanse



<https://jupyter.org>

What is a Jupyter notebook?

A Jupyter notebook is a web-based application that allows you to create and shared documents that can run live code, visualize data, and display explanatory text, including equations. Use cases include: data cleaning and transformation, numerical simulation, statistical modeling, machine learning and more.

Show-and-tell: An example notebook

`https://github.com/losc-tutorial/quickview`

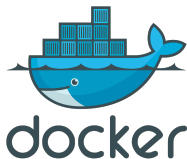
Setting up your Jupyter notebook environment

--sif

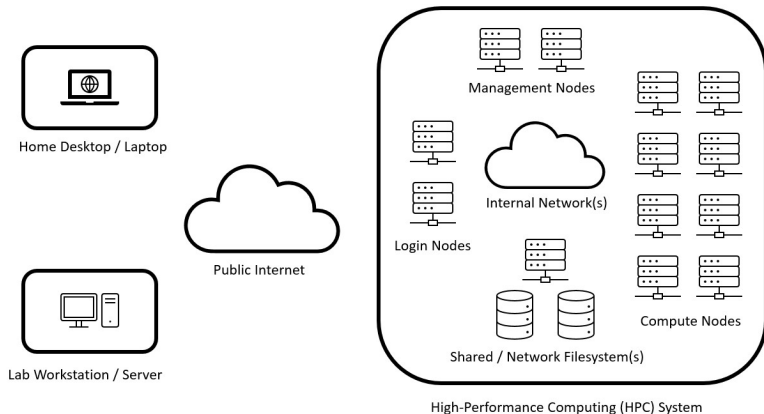
--env-modules



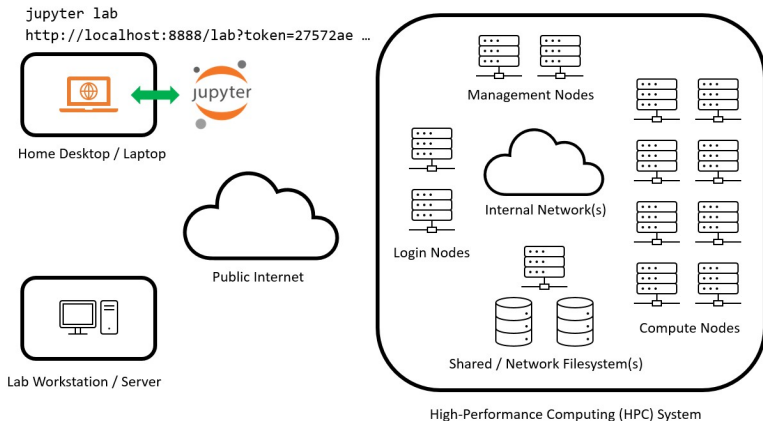
--conda-env



How-to run Jupyter notebooks on Expanse



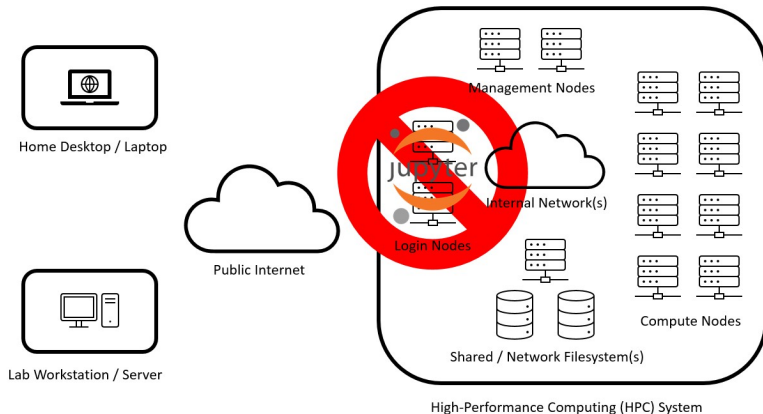
Running your Jupyter notebooks, locally (HTTP)



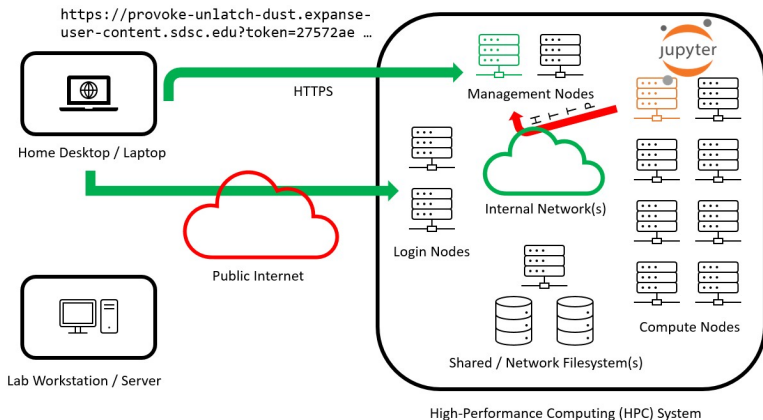
Running your Jupyter notebooks, locally (HTTP)

```
mkandes — jupyter-lab — 80x24
mkandes@castleyankee ~ % jupyter lab
[I 2021-12-13 12:29:20.163 ServerApp] jupyterlab | extension was successfully linked.
[I 2021-12-13 12:29:20.532 ServerApp] nbclassic | extension was successfully linked.
[I 2021-12-13 12:29:20.601 ServerApp] nbclassic | extension was successfully loaded.
[I 2021-12-13 12:29:20.603 LabApp] JupyterLab extension loaded from /usr/local/Cellar/jupyterlab/3.2.5/libexec/lib/python3.9/site-packages/jupyterlab
[I 2021-12-13 12:29:20.603 LabApp] JupyterLab application directory is /usr/local/Cellar/jupyterlab/3.2.5/libexec/share/jupyter/lab
[I 2021-12-13 12:29:20.607 ServerApp] jupyterlab | extension was successfully loaded.
[I 2021-12-13 12:29:20.608 ServerApp] Serving notebooks from local directory: /Users/mkandes
[I 2021-12-13 12:29:20.608 ServerApp] Jupyter Server 1.13.1 is running at:
[I 2021-12-13 12:29:20.608 ServerApp] http://localhost:8888/lab?token=27572ae29daf84aa8b69bc79884233f56f9ca94ea754fca6
[I 2021-12-13 12:29:20.608 ServerApp] or http://127.0.0.1:8888/lab?token=27572ae29daf84aa8b69bc79884233f56f9ca94ea754fca6
[I 2021-12-13 12:29:20.608 ServerApp] Use Control-C to stop this server and shutdown all kernels (twice to skip confirmation).
[C 2021-12-13 12:29:20.622 ServerApp]
```


DO NOT RUN ON LOGIN NODES!



Running your Jupyter notebooks, remotely (HTTPS)



What is galyleo?

galyleo is a shell utility to help you launch Jupyter notebooks on high-performance computing (HPC) systems – like Expanse – in a simple, secure way. It works with the SDSC's Satellite Reverse Proxy Service and a batch job scheduler like SLURM to provide each Jupyter notebook server you start with its own one-time, token-authenticated HTTPS connection between the compute resources of the HPC system the notebook server is running on and your web browser.

`https://github.com/mkandes/galyleo`

Getting Started with galyleo in 3 Easy Steps

1. Prepend its installation location to your PATH.

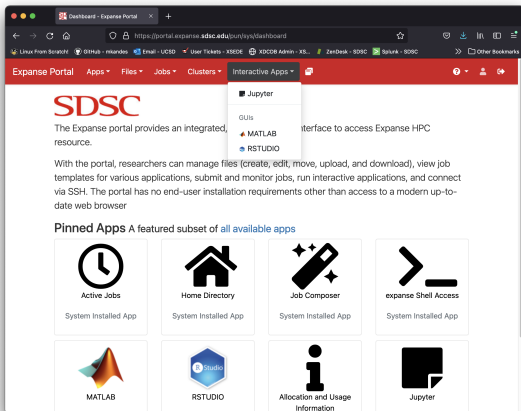
```
export PATH="/cm/shared/apps/sdsc/galyleo:${PATH}"
```

2. launch your Jupyter notebook session.

```
galyleo launch --account abc123 --partition shared --cpus 1 \
  --memory 2 --time-limit 00:30:00 --env-modules cpu,gcc,anaconda
```

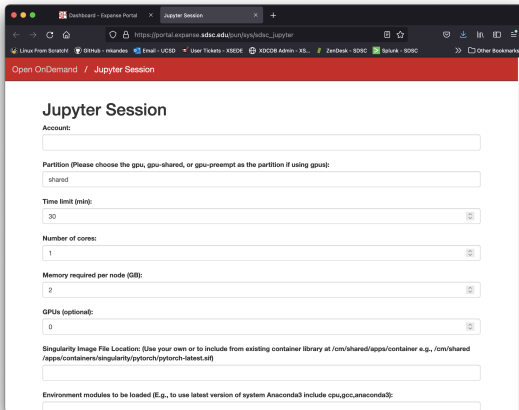
3. Copy and paste the HTTPS URL into your web browser.

Expanse User Portal - Interactive Apps - Jupyter



`portal.expense.sdsc.edu`

Expanse User Portal - Interactive Apps - Jupyter



The screenshot shows a web browser window with the URL `https://portal.expanse.sdsc.edu/#!/apps/sdsc_jupyter`. The page has a red header bar with the text "Open OnDemand / Jupyter Session". Below the header, the main content area is titled "Jupyter Session". It contains several form fields for configuring a Jupyter session:

- Account:** A text input field.
- Partition (Please choose the gpu, gpu-shared, or gpu-preempt as the partition if using gpus):** A dropdown menu with "shared" selected.
- Time limit (min):** A text input field with "30" and a spinner icon.
- Number of cores:** A text input field with "1" and a spinner icon.
- Memory required per node (GiB):** A text input field with "2" and a spinner icon.
- GPUs (optional):** A text input field with "0" and a spinner icon.
- Singularity Image File Location:** A text input field with the placeholder text: "(Use your own or to include from existing container library at /cm/shared/apps/container e.g., /cm/shared/apps/containers/singularity/pytorch/pytorch-latest.sif)".
- Environment modules to be loaded (E.g., to use latest version of system Anaconda3 include cpu,gcc,anaconda3):** A text input field.

`portal.expanse.sdsc.edu`

Questions?

