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, nuerical simulation, statistical modeling, machine learning and more.

Show-and-tell: An example notebook
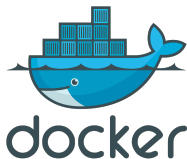
`https://github.com/losc-tutorial/quickview`

Setting up your Jupyter notebook environment
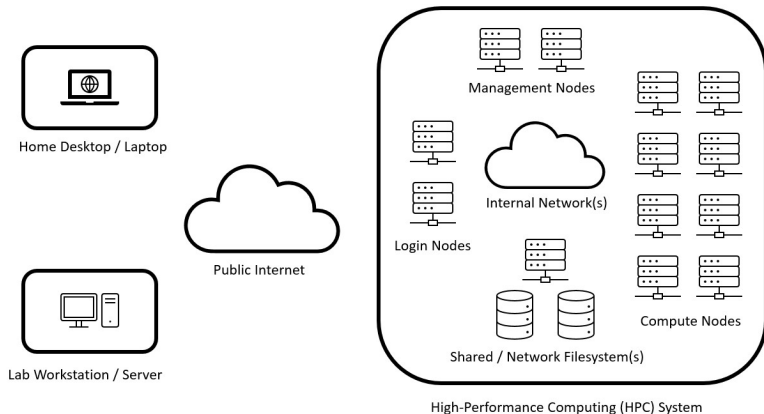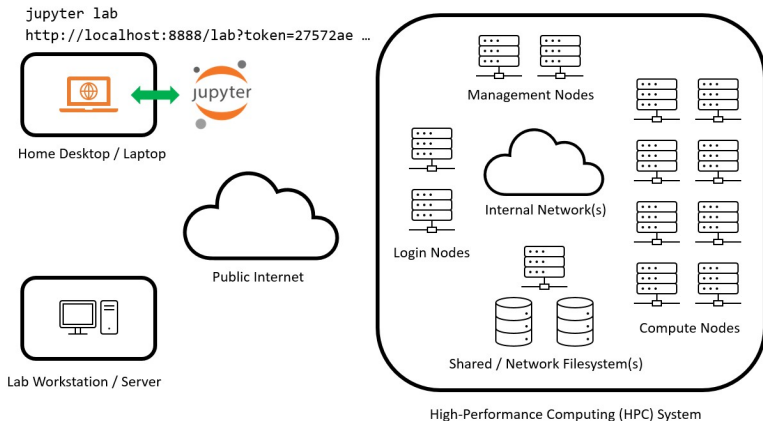
How-to run Jupyter notebooks on Expanse

Running your Jupyter notebooks, locally (HTTP)

Running your Jupyter notebooks, locally (HTTP)

DO NOT RUN ON LOGIN NODES!

Home Desktop / Laptop

Lab Workstation / Server

Public Internet

Management Nodes

Jupyter

Login Nodes

Internal Network(s)

Shared / Network Filesystem(s)

Compute Nodes

High-Performance Computing (HPC) System
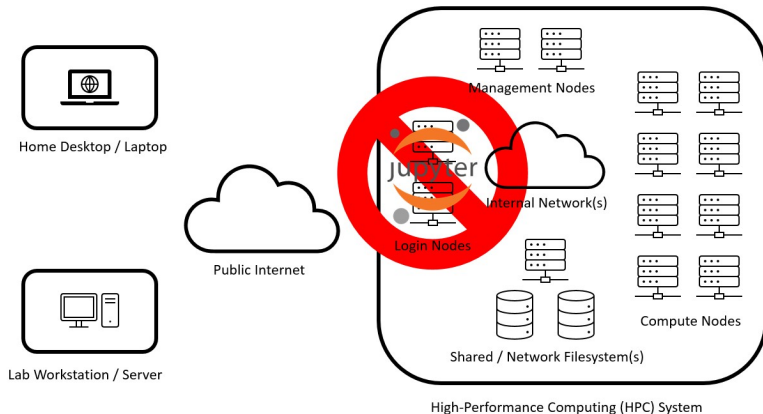
Running your Jupyter notebooks, remotely (HTTPS)

https://provoke-unlatch-dust.expanse-user-content.sdsc.edu?token=27572ae …

Home Desktop / Laptop

HTTPS

Public Internet

Lab Workstation / Server

Management Nodes

Internal Network(s)

Login Nodes

Shared / Network Filesystem(s)

Compute Nodes

High-Performance Computing (HPC) System
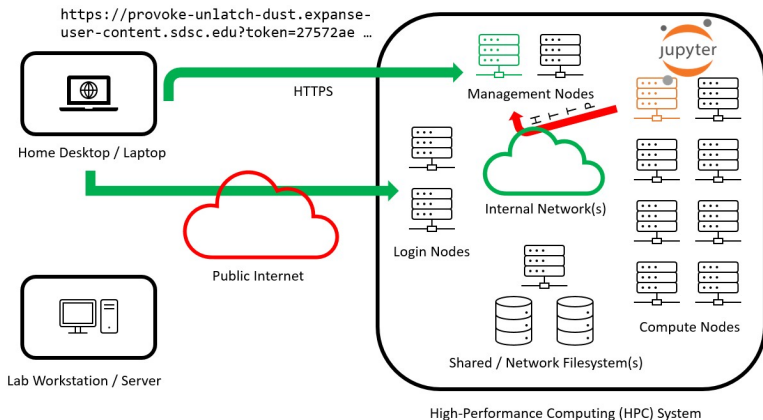
What is galyleo?

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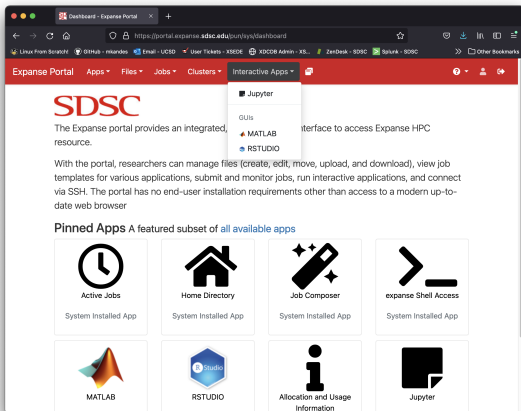   ```
   export PATH="/cm/shared/apps/sdsc/galyleo:${PATH}"
   ```

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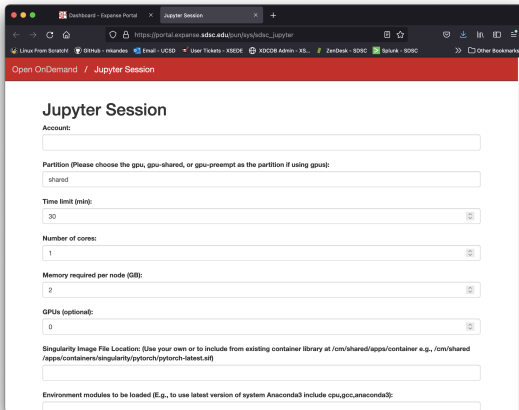
3. Copy and paste the HTTPS URL into your web browser.

# Expanse User Portal - Interactive Apps - Jupyter



portal.expanse.sdsc.edu

# Expanse User Portal - Interactive Apps - Jupyter



portal.expanse.sdsc.edu

# Questions?