

ModuLair: Streamlining Python venv Management on HPC

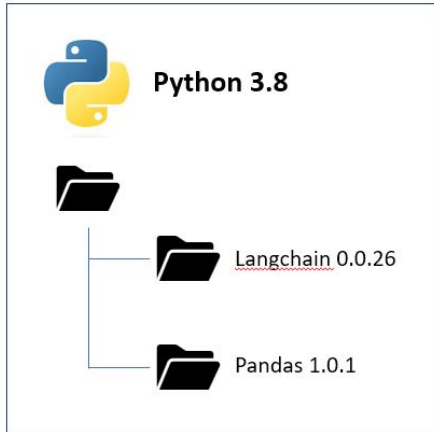
Gabriel Floreslovo*, **Surada Suwansathit***, Marinus Pennings, Honggao Liu
Texas A&M University

*both authors contributed equally to this research.

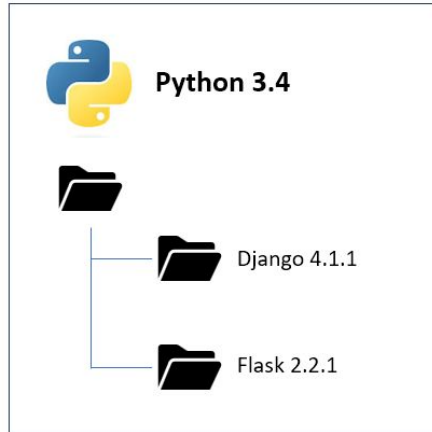
Python Virtual Environment (venv)

- User-specific, portable, isolated software setup
- Install software without modifying on system-wide package
- Recreate the same computational environment across different systems

Virtual Environment 1



Virtual Environment 2



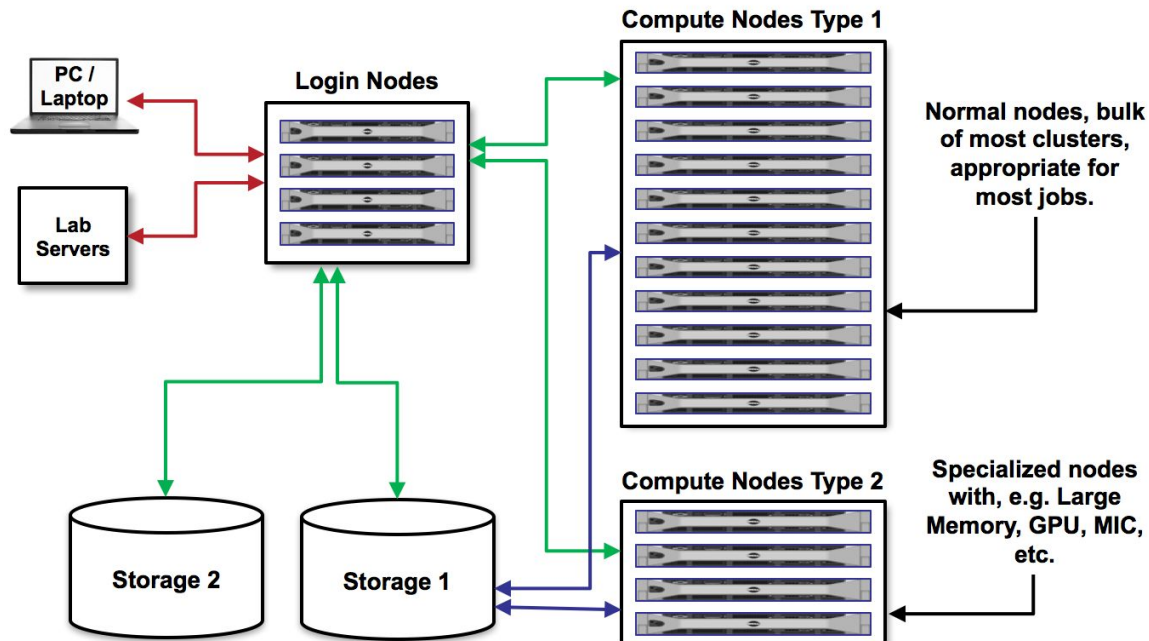
Research Data Management

Python environments are part of data provenance and workflow reproducibility

- venvs capture the environment used to process, analyze, and produce results

Why venvs Are Hard on HPC

- HPCs have complex environments
 - Modules
 - Quotas
 - Steep learning curve



What is Modulair?

- Python Virtual Environment Framework for HPC
- Works with EasyBuild and non-EasyBuild clusters
- Tracks dependencies for reproducibility
- Offers both CLI and GUI tools



<https://github.com/tamu-edu/dor-hprc-venv-manager>

Core Features

- **create_venv**: auto-detects toolchain or dependencies
- **activate_venv**: loads modules + sources environment
- **list_venvs**: shows metadata and environments
- **delete_venv**: removes environment and updates registry

```
[gabr13lenstation@faster1 ~]$ list_venvs
These are your virtual environments currently available through 'activate_venv':

Name: hi Description: stuff Python Version: Python/3.10.4 GCC version: GCCcore/11.3.0
Name: test Description: test Python Version: Python/3.10.4 GCC version: GCCcore/11.3.0
Name: space Description: Python Version: Python/3.11.5 GCC version: GCCcore/13.2.0
Name: testerr Description: Python Version: Python/3.11.5 GCC version: GCCcore/13.2.0

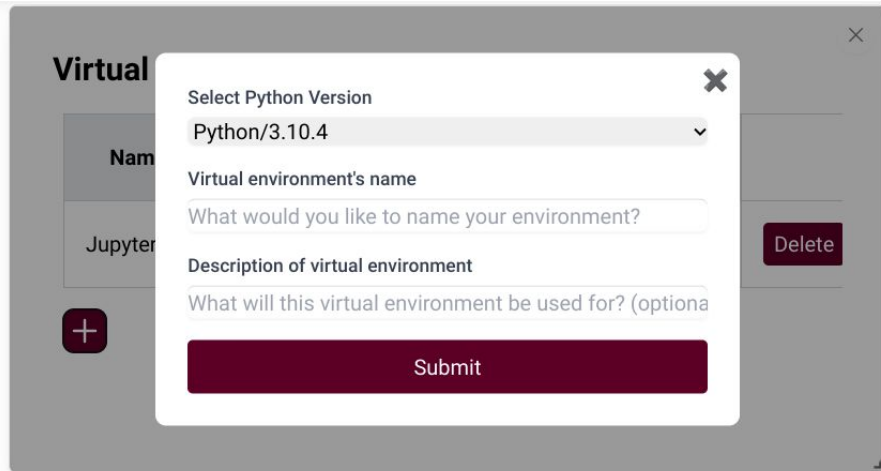
For example, to activate 'hi' run this command:
$ source activate_venv hi

[gabr13lenstation@faster1 ~]$ source activate_venv space
When deactivating this environment, remember to run deactivate first:
$ deactivate

And then purge modules AFTERWARDS (if desired)
$ ml purge

Otherwise, there may be undefined behavior
(space) [gabr13lenstation@faster1 ~]$ python --version
Python 3.11.5
```

GUI #1 TAMU HPRC User Support Dashboard



A screenshot of a web application window titled "Virtual" with a close button (X) in the top right corner. The window contains a form for creating a virtual environment. The form has a "Select Python Version" dropdown menu with "Python/3.10.4" selected. Below this is a text input field for "Virtual environment's name" with the placeholder text "What would you like to name your environment?". There is also a text input field for "Description of virtual environment" with the placeholder text "What will this virtual environment be used for? (optional)". A "Submit" button is at the bottom of the form. To the right of the form is a "Delete" button. A "+" button is visible in the bottom left corner of the window.

Virtual

Select Python Version

Python/3.10.4

Virtual environment's name

What would you like to name your environment?

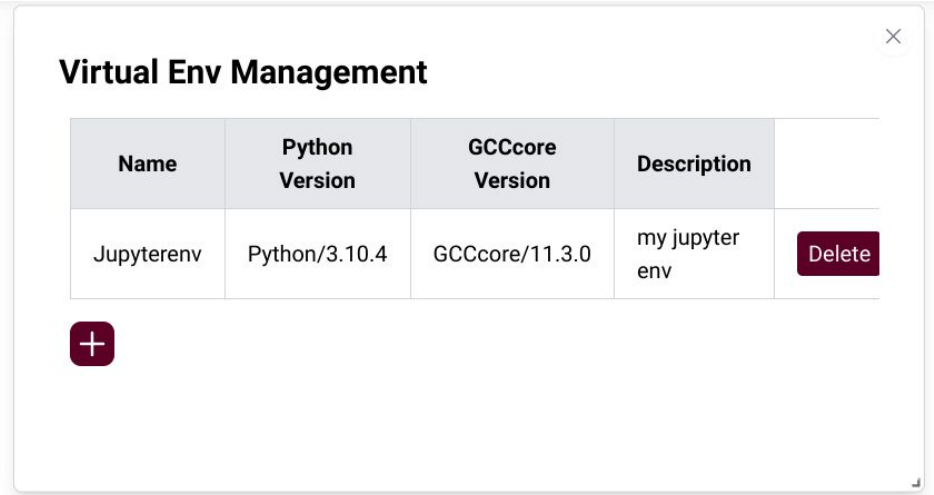
Description of virtual environment

What will this virtual environment be used for? (optional)

Submit

Delete

Creation Form



A screenshot of a web application window titled "Virtual Env Management" with a close button (X) in the top right corner. The window displays a table with four columns: "Name", "Python Version", "GCCcore Version", and "Description". There is a "Delete" button next to each row. A "+" button is visible in the bottom left corner of the window.

Name	Python Version	GCCcore Version	Description	
Jupyterenv	Python/3.10.4	GCCcore/11.3.0	my jupyter env	Delete

Listing Interface

GUI #2 Drona Composer

Job Name

Job1

Location

Change

/scratch/user/surada/drona_composer/runs/

Environments

Python

▼

+

Python Environment ?

Private virtual env

▼

Main Python script ?

local

cluster

My environments ?

|-- Choose an option --

▼

malwa

trnsf

jyp

Python Environment ?

Create new private virtual env

▼

Name of virtual env ?

my_venv

Python Version ?

|-- Choose an option --

▼

Python 3.11.5

Python 3.10.4

Python 3.9.6

Python 3.8.6

Python 3.7.0

Main Python script ?

local

cluster

Upload files/data ?

Select an option ▼

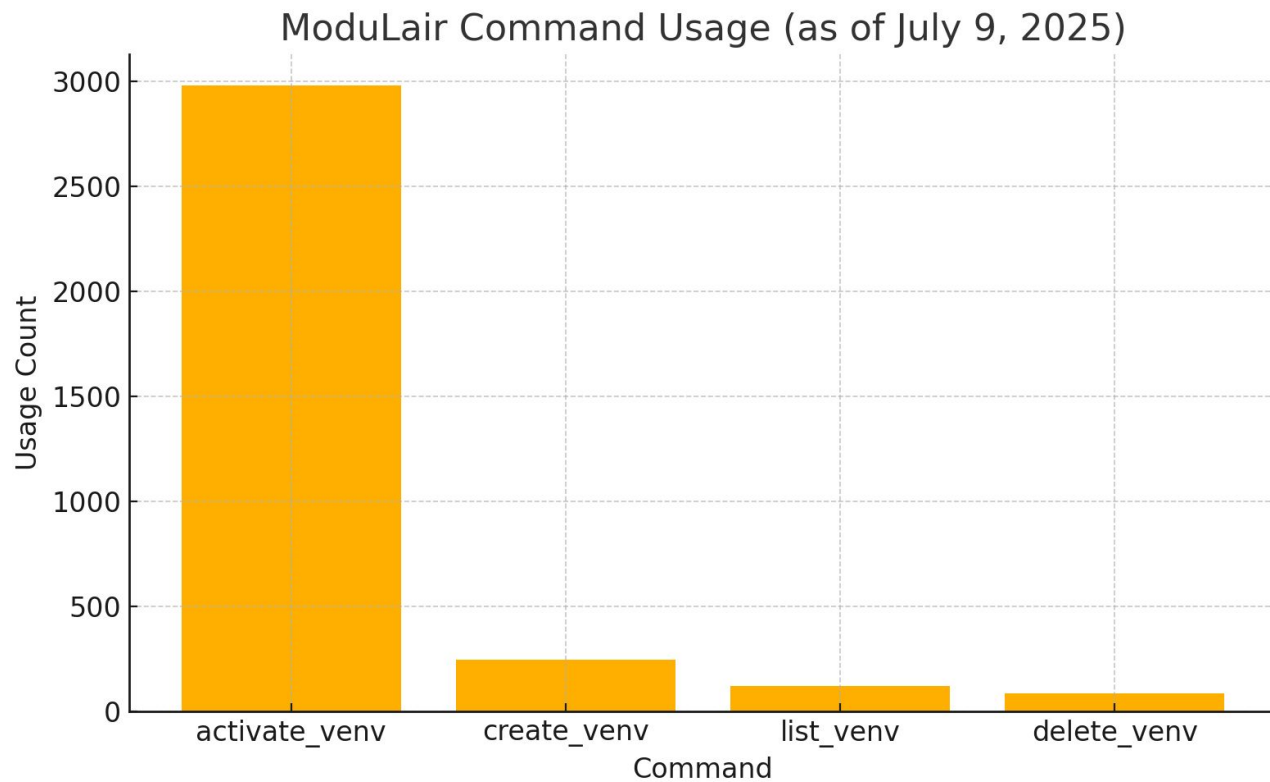
Add

Set Job requirements ?

☐

Generate

Usage Stats



Contacts and Acknowledgements

Surada Suwansathit

help@hprc.tamu.edu

Kenton Romero (former HPRC student technician) wrote the original bash script that inspired this paper. We gratefully acknowledge the National Science Foundation for support of this work under award number 2112356 (ACES: Accelerated Computing for Emerging Sciences), and to Elizabeth Leake (HPRC Project Manager of Advanced Cyberinfrastructure) and TAMU Writing Center for proofreading.

