

Advanced Data Analytics — Lecture 1b

Crash-course on the Computational Infrastructure

Dimitrios Karyampas

Department of Economics, University of Lausanne, Switzerland

Slides by Simon Scheidegger (2023–2025), adapted with permission

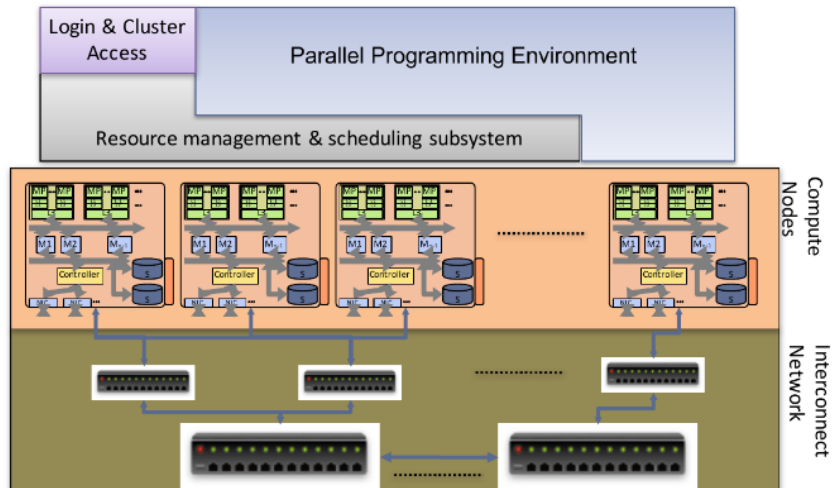
September 14th, 2026 | 12:30 - 14:00: Internef 125 | 16:30 - 18:00: Anthropole 3185

- ▶ Crash-course on the Computational Infrastructure
 - ▶ Make first steps on a Linux Cluster (Login, short overview of basic unix commands like cd, pwd, cp, scp,...)
 - ▶ Get lecture notes/codes for the lecture
- ▶ Hello World in Python

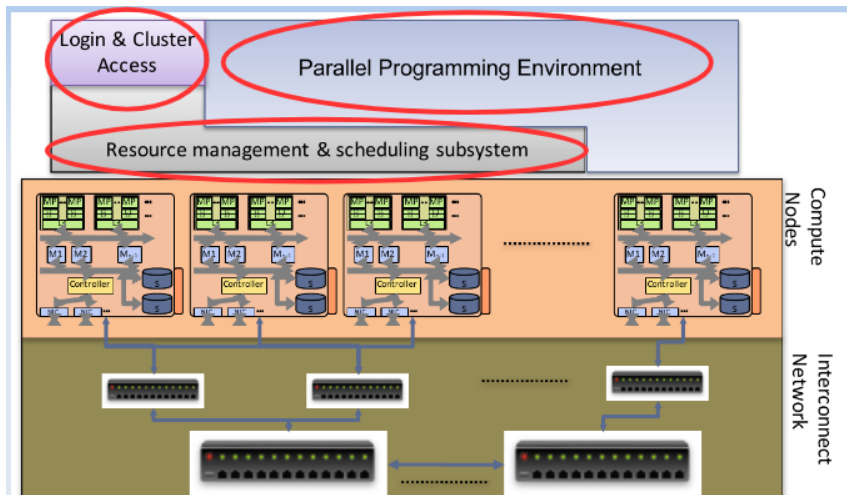
**MORE DETAILS IN THE TA
SESSION**

- ▶ For this course, we use Nuolos.cloud: JupyterLab in the browser, every package installed, nothing on your laptop.
- ▶ Its setup is very similar to any other system
- ▶ Log in at <https://app.nuolos.cloud> with the account you enrolled with. Enrolment is closed; if you cannot get in, email the teaching assistant.
- ▶ Step-by-step guide:
<https://unil-ada.github.io/nuolos.html>

An abstract compute cluster



An abstract compute cluster



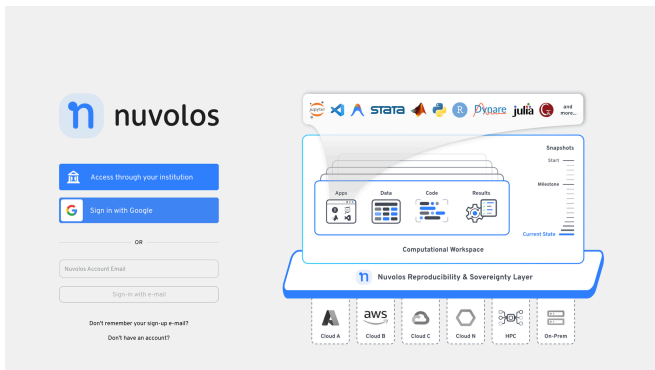
A real compute cluster



Credit: Blue Waters research cluster at NCSA, Illinois, USA (active 2007-2021).

Access to Nuvolos Services

LOG IN: app.nuvolos.cloud
with the account you enrolled with. Enrolment is closed.



After login: your own instance

HEC Lausanne → Advanced Data Analytics 2026 → an instance named after your email.
Nobody else sees it. The Overview shows the workspace guide.

The screenshot displays the Nuvolos web interface. The top navigation bar shows the breadcrumb: ORGANIZATION (HEC Lausanne) > SPACE (Advanced Data Analytics 2026) > INSTANCE (anna.smirnova@unil.ch) > SNAPSHOT (Current State). The left sidebar contains a menu with options: SPACE (Advanced Data Analy), Instances, INSTANCE (anna.smirnova@unil.ch), Snapshots, Overview (selected), Files, Tables, Applications, Staging, and Assignments. Below the menu, a 'SNAPSHOT' section shows 'Current State' with a storage usage bar at 0.02% (1.9 MB of 10.0 GB used) and the user profile 'Anna Smirnova'. The main content area is titled 'Overview' and features a card for 'Advanced Data Analytics – your workspace'. This card explains that the workspace is persistent and contains a table with the following data:

Folder	What it is
refresher/	Twelve short Python notebooks, from calculator to pandas. Start here if you have not written Python for a while.
week-NN/	The runnable material for week NN: demo notebooks, scripts, small data files. A new folder appears after each Monday's session.

Below the table, it states: 'Slides and PDFs are not here – they are on the course site, where you read them: <https://unil-ada.github.io> (schedule, weekly materials, the project, the refresher as readable pages).' A section titled 'The one rule' explains that 'week-NN/' folders are overwritten when material is redistributed and advises users to keep their work outside these folders. The interface also includes a 'Resources' sidebar on the right with links for Files, Tables, Applications, and Snapshots.

Your files

Files: README.md and refresher/ are there from day one; a week-NN/ folder appears after each Monday. Week folders are overwritten on updates – keep your own work outside them.

The screenshot shows the Nuvolos interface. The left sidebar contains a navigation menu with the following items: SPACE (Advanced Data Analy), INSTANCES, INSTANCE (anna.smirnova@un...), SNAPSHOTS, OVERVIEW, FILES (selected), TABLES, APPLICATIONS, STAGING, and ASSIGNMENTS. The main area is titled 'Files' and shows a list of files in the 'WORKSPACE'.

File	Last modified	Size	Actions
refresher	2026-09-13 Time: 15:06 GMT+2	1.9 MB	...
README.md	2026-09-13 Time: 15:06 GMT+2	2.0 kB	...

Below the file list, the 'README.md' file is previewed. It contains the following text:

Advanced Data Analytics – your workspace

This is your Nuvolos workspace for the course. Everything in this folder is yours and persists between sessions.

What is here

Folder	What it is
refresher/	Twelve short Python notebooks, from calculator to pandas. Start here if you have not written Python for a while.
week-NN/	The runnable material for week NN: demo notebooks, scripts, small data files. A new folder appears after each Monday's session.

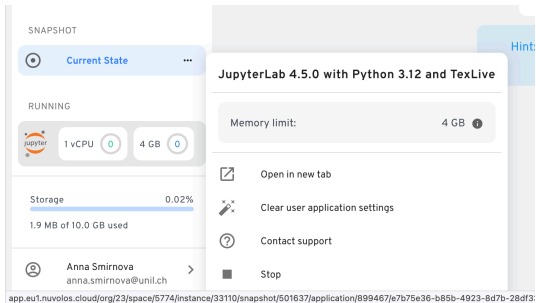
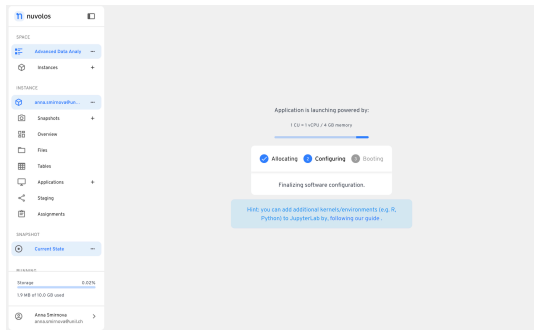
Applications: start JupyterLab

Applications → the power button next to *JupyterLab 4.5.0 with Python 3.12 and TeX Live*.
Every package used in the course is already installed in it.

The screenshot shows the NuvoLabs web interface. The left sidebar has a navigation menu with sections: SPACE (Advanced Data Analy), INSTANCES (Instances), INSTANCE (anna.smirnova@unil.ch), and SNAPSHOT (Current State). The main content area is titled 'Applications' and shows a table of applications. The table has columns: Type, Name, Size, Credit/hour, Status, Description, Space usage, and Action. One application is listed: 'JupyterLab 4.5.0 with Python 3.12 and TeXLive' with a size of 1 CU and a status of 'Running' (indicated by a power button icon). The bottom of the sidebar shows storage usage (1.9 MB of 10.0 GB used) and the user profile (Anna Smirnova).

Type	Name	Size	Credit/hour	Status	Description	Space usage	Action
	JupyterLab 4.5.0 with Python 3.12 and TeXLive	1 CU	-	Running	JupyterLab 4.5.0 + Python 3.12 + TeXLive (no collaboration)	0 B	...

Starting takes a minute; stop it when you are done



Left: the first start takes about a minute. Right: the running app in the sidebar – *Open in new tab* to get back to it, *Stop* when you finish. Your files are saved continuously; stopping loses nothing.

JupyterLab

File browser on the left, notebook on the right. Open the folder first, then the notebook; run cells with Shift+Enter. Start with `refresher/jupyter_intro.ipynb`.

The screenshot displays the JupyterLab interface. On the left is the 'nuvotos' sidebar with sections for 'Instances', 'Snapshots', 'Overview', 'Files', 'Tables', 'Applications', 'Staging', and 'Assignments'. The 'Files' section is active, showing a file browser view of the '/files/refresher/' directory. A table lists files with their names and modification times:

Name	Modified
jupyter_intro.ipynb	22h ago
python_basics_1...	22h ago
python_basics_2...	22h ago
python_basics_3...	22h ago
python_basics_4...	22h ago
python_basics_5...	22h ago
python_basics_6...	22h ago
python_basics_7...	22h ago
python_basics_8...	22h ago
python_basics_9...	22h ago
python_basics_10...	22h ago
python_basics_11...	22h ago

The main area shows the 'jupyter_intro.ipynb' notebook. The top cell contains Python code:

```
[3]: import IPython
import numpy as np
from scipy.io import wavfile
from matplotlib import pyplot as plt
```

The notebook content includes the Jupyter logo and the title 'Introduction to Jupyter'. Below the title is a bulleted list:

- **Jupyter** (previously known as IPython Notebook) is an application that facilitates the **development, execution, and presentation of Python code**, all in one interface.
- Jupyter also supports text, *LaTeX*, hyperlinks, images, sound files, and more.

The 'Cells' section follows, explaining that a Jupyter notebook consists of separate sections called cells. It lists two main types:

- **Markdown cells** (like this cell), containing for instance **text**, *LaTeX*, or **images**.
- **Code cells**, containing Python code that can be executed within the notebook.

Additional details about code cells are provided:

- Each code cell can be run independent of other cells, as if the code from that cell were being pasted into IPython and executed line by line.
- When a cell is run, its output will be displayed directly after it.
- Run the following code cell by clicking on it, holding **Ctrl** or **Shift**, and pressing **Enter**.
- You can also click the run button (the **play** button) on the tool bar at the top of the page.

The 'Menu bar' section states: 'The top level menus expose actions along with their keyboard shortcuts.' It lists several action categories:

- **File:** actions related to files and directory
- **Edit:** actions related to editing documents
- **View:** actions to alter the appearance of JupyterLab
- **Run:** actions for running code in notebooks and consoles
- **Kernel:** actions for managing kernels (processes for running code)
- **Help:** List of links for help on JupyterLab and kernels

The bottom status bar shows 'Simple', '1', 'Python 3 (ipykernel) | Idle', 'Mode: Command', 'Ln 1, Col 1', 'jupyter_intro.ipynb', and a bell icon.

Basic Linux commands

Command	Description
<code>pwd</code>	Print name of current/working directory
<code>cd [Directory]</code>	Change directory (no directory → change to home)
<code>ls [Directory]</code>	List directory contents (no directory → list current)
<code>cat FILE</code>	Concatenate files and print on the standard output
<code>mkdir DIRECTORY</code>	Make directories
<code>mkdir -p DIRECTORY</code>	Make directories, make parent directories as needed
<code>cp SOURCE. . . DIRECTORY</code>	Copy files and directories
<code>cp -r SOURCE. . . DIRECTORY</code>	Copy files and directories, copy directories recursively
<code>mv SOURCE. . . DIRECTORY</code>	Move (rename) files
<code>man COMMAND</code>	An interface to the on-line reference manuals

Basic Linux commands

Command	Description
<code>ssh -X foo@host.com</code>	OpenSSH SSH client (remote login program), access to host.com with user foo
<code>scp foo@host.com:/home/bar ./</code>	Secure copy (remote file copy program), copy file bar from /home on host.com to directory
<code>scp bar foo@host.com:/home/</code>	Secure copy (remote file copy program), copy file bar from the local host to /home on host.com
<code>git clone git@github.com:whatever folder-name</code>	The stupid content tracker, Clone a repository (whatever) into a new directory (folder-name)
<code>git checkout</code>	Checkout a branch or paths to the working tree.