Introduction to Linux for HPC

Lecturer: Maria Chiara Debernardi

Language

English

Course description and objectives

The main objective of this short, hands-on course is to provide participants with the essential Linux and High Performance Computing (HPC) skills needed to effectively run Python-based data science and AI workloads on Bocconi's shared HPC facility. Through guided exercises on the actual cluster, students will learn how to connect securely, navigate the Linux filesystem, manage files, submit and monitor batch jobs with SLURM, and work efficiently on the command line when working with large datasets and computationally intensive models.

Before the start of the course, students will complete a short self-assessment test on Blackboard to evaluate their initial familiarity with Linux and the command line. The same test will be proposed at the end of the course so that each student can measure their own progress.
*Please note that these tests are intended **only** for self-evaluation and that the results will not be used in any curricular course.*

At the end of the course, participants will be able to:

- access Bocconi's shared HPC facility via SSH and understand the difference between login nodes and compute nodes
- navigate the Linux filesystem on the cluster and manage files and folders (create, move, copy, remove) safely from the command line
- transfer data between their local computer and the HPC system
- submit, monitor, and cancel batch jobs with SLURM, and interpret the resulting log and output files
- use software modules and basic Python environments to run their own code reproducibly on the cluster
- search for files and filter text efficiently using Linux tools and basic regular expressions

Audience

The course is **reserved for Master of Science students** enrolled in the Data Science and Business Analytics (**DSBA**) or Artificial Intelligence (**AI**) programs who want to learn the Linux operating system to access and use Bocconi's HPC resources.

Prerequisites

No prior Linux knowledge is required, but basic coding experience in Python is assumed. It is advisable to be familiar with basic computer operations, such as file management.

Guidelines

Registration:
You can sign up for the course only through the yoU@B student Diary, in the "**Sign-up for various activities**" box (please note that the box appears only when registrations open. Before then it will not be visible).
You can only cancel your registration by Diary **no later** than the registration deadline for the course itself. No other ways of cancellation are allowed.
Registration will be confirmed a few days before the start of the course through a message posted in the yoU@B student Diary.

Attendance:

- Attendance of **75% or more** of the lessons: obtainment of the **Open Badge**
- Attendance of **less than 25%** of the lessons: placement on **Exclusion List**

Duration

8 academic hours

Teaching mode

This course will be taught only in person. Neither online participation nor video recordings will be provided.

Calendar

Lecture	Date	Time	Room
1	Thu 12/02/2026	18.15 - 19.45	4 (Sarfatti)
2	Thu 19/02/2026	18.15 - 19.45	4 (Sarfatti)
3	Thu 26/02/2026	18.15 - 19.45	4 (Sarfatti)
4	Thu 05/03/2026	18.15 - 19.45	4 (Sarfatti)

Note: lessons will be held in a traditional classroom, and **all the students must bring their own device**.

Syllabus of the course

Lecture	Topics

1	**Managing files and folders**
	- The HPC resources overview
	- Login node *vs*. compute nodes
	- SSH login
	- From Windows/Mac file manager to the Linux command line: parallels and analogies
	- Linux Command Line Interface (CLI)
	  o Up and down arrows
	  o Tab key for completion
	  o Interrupting commands with Ctrl+C
	  o `nano` (to write small scripts or edit them)
	- Help in Linux: `man` command and `--help` option
	- The `tree` command
	- Linux commands:
	  o `whoami, hostname`
	  o `ls (-l, -a), pwd`
	  o `mkdir, cd`
	  o Special directory shortcuts: `~, ., .., /, -` (`$OLDPWD`)
	  o `mv, cp (-r)`
	  o `file, touch, cat, less, head, tail`
	  o `rm (-r, -f), rmdir` (dangerous, not undoable!)
	  o `clear`
	  o `exit`
	- Absolute *vs*. relative path (`realpath`)
	- Copying files between HPC and local PC (`scp`)
	Exercises

Lecture	Topics

2	**The batch system**
	- The batch system on an HPC cluster
	- SLURM basics
	  o Submitting jobs: `sbatch`
	  o Monitoring the queue and resources: `squeue`, `sinfo`
	  o Inspecting job details: `scontrol`
	  o Cancelling jobs: `scancel`
	- Writing a simple SLURM job script
	  o Structure of a batch script
	  o Requesting resources (time, memory, CPUs)
	  o Redirecting output and error to log files
	- Software environments
	  o Using modules: `module avail/load/list/unload`
	  o Building a minimal Python environment with *conda*
	  o Activating the environment in a job script using Python
	- Practice: submitting jobs to the scheduler
	  o Example: simple Python scripts that read/write small files (e.g., with *open* and *pandas*)
	  o Checking logs, understanding common errors (missing module, wrong path, exceeded time limit, etc.)
	- Storage and `quota`
	  o Default storage areas on the HPC (e.g., home, project space)
	  o How quotas and storage limits affect job runs
	- Linux commands:
	  o `echo`
	  o `df -h, du -sh, quota`
	Exercises

Lecture	Topics
3	**Searching files and contents on HPC** - Searching for files and directories - Searching text inside files - Pipes and simple text processing - Linux commands: o `find` (`-name`, `-type`, `-mtime`) o `grep` (`-i`, `-c`, `-n`, `-v`, `-r`, `-E`) o Chaining commands with `|` o `sort` o `uniq` - Introduction to regular expressions (RegEx) o Matching characters: `.`, `[…]` o Quantifiers: *, +, ? o Anchors: ^, $ o Alternation and grouping: `|`, (…) *Exercises*
4	**Recap and insights** - Copying your public SSH key to the cluster - Additional useful commands (optional): o `curl` (`-o`, `-s`) o `chmod` (`u +-rwx`) o `diff` o `history` o `wc` - Best practices for efficient job scheduling and resource utilization - Recap exercise - Final self-assessment *Exercises*

Software used

Bocconi HPC account with Linux OS and Python

Enrolled students will receive an email from Bocconi's Technology Office to activate their own HPC account before the first lesson. Instructions will be provided in Blackboard.

The course focuses on command-line and batch-job usage; interactive tools such as Jupyter or remote IDEs will not be covered.

Suggested bibliography

Materials will be provided by the instructor during the course and will be accessible on Blackboard.
Recommended online resources:

- https://learnbyexample.github.io/cli-computing/
- https://www.freecodecamp.org/news/the-linux-commands-handbook
- https://ryanstutorials.net/linuxtutorial/
- https://learning.lpi.org/en/learning-materials/010-160/

Available seats

This activity is limited to **110** participants. Registrations cannot be conducted once this number has been reached or after the registration period ends.
Please remember that you may unsubscribe from ITEC courses only through the yoU@B Diary and **only before the registration deadline**.