Heinrich Heine
Universität Düsseldorf

Using HPC of HHU

Create an account

- You need to create an account on https://www.zim.hhu.de/forschung/high-performance-computing/antrag. Fill in the following information:
  - Institut: “Institut für Informatik”
  - Institutsleiter: Prof. Dr. Martin Lercher
  - Verantwortlicher des Projekts: Alexander Kroll
  - Zustand des Projekts: ”Neues Projekt”
  - Projektkürzel: “DL4MoleculesCourse”
  - Kurze Projektbeschreibung: “This account will be used for small programming tasks that are part of the lecture ‘Applications of Transformer Networks in Bio- and Cheminformatics’ by Alexander Kroll.”

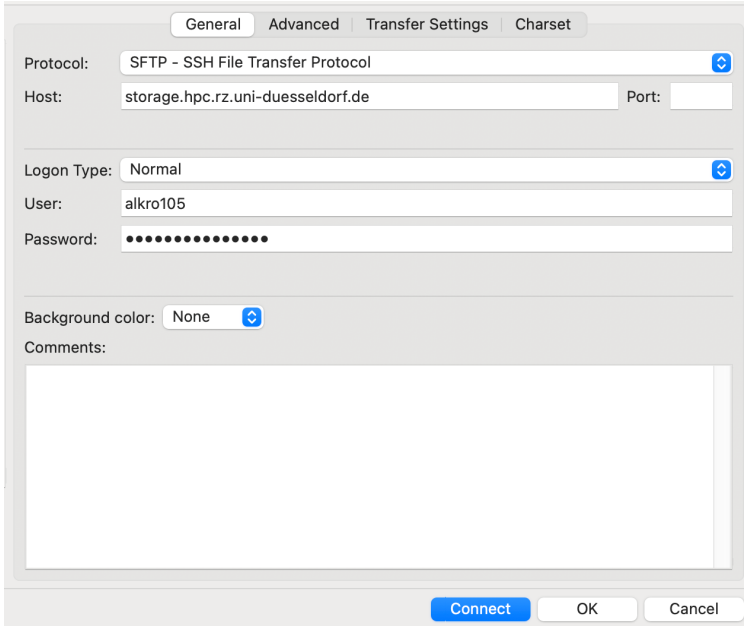
Requirements to log in to HPC

Heinrich Heine
Universität Düsseldorf

- You need to be in the campus network to log in to HPC. Options:
  - Connect to university wifi
  - Connected to campus network via VPN:
    - Follow instructions on https://www.zim.hhu.de/en/servicekatalog/networks/vpn

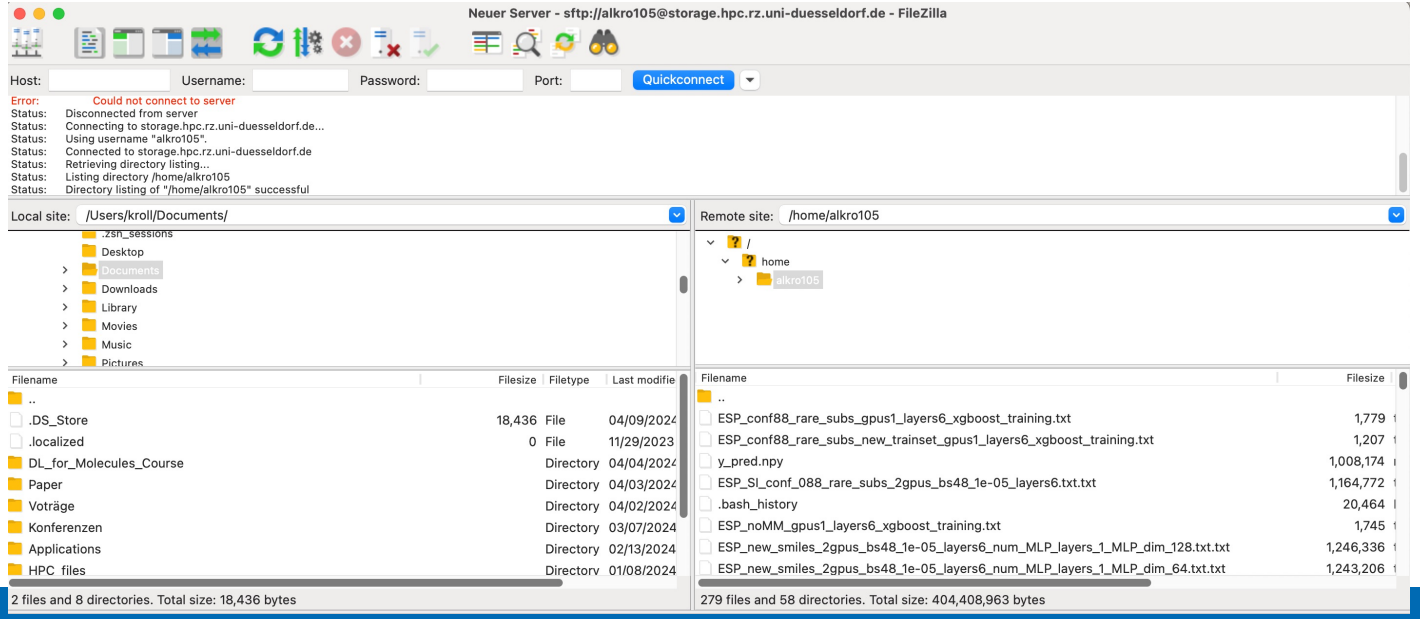
Managing files with FileZilla

- You need to move files between your PC and the HPC
- You can do this using FileZilla: https://filezilla-project.org/ (available for all platforms)
  - Download the FileZilla Client
- Log in to HPC using FileZilla:
  - Connect to new server (ctrl + s / cmd + s) by entering the information on the right
  - User is your university username



Managing files with FileZilla (2)

- After you are connected you can see
  - Left side: The files on your PC
  - Right side: Files in your HPC folder “/home/user”
- You can move files be right clicking on them



Jobs are submitted via Bash scripts

Heinrich Heine Universität Düsseldorf

- Example for bash script that installs the python package torch

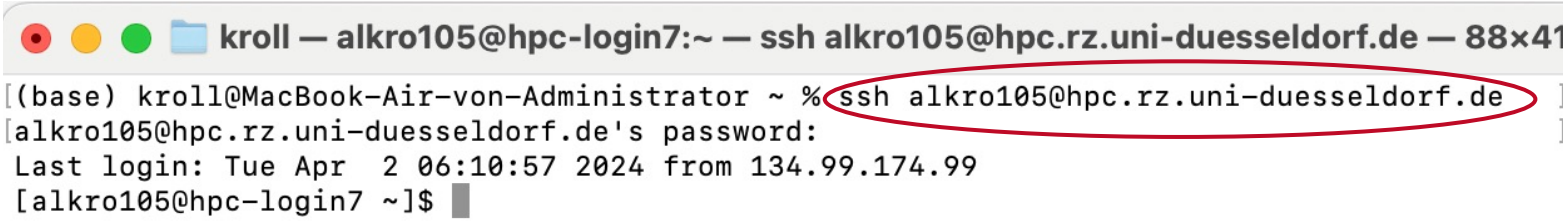
```
1   #!/bin/bash
2   #PBS -l select=1:ncpus=1:mem=6GB
3   #PBS -l walltime=05:00:00
4   #PBS -A kcat-prediction
5   #PBS -N package_install
6   #PBS -j oe
7   #PBS -o "package_install.log"
8
9   cd /home/alkro105/
10
11  module load Python/3.8.3
12  module load CUDA/11.7.1
13
14  pip install --user -i http://pypi.repo.test.hhu.de/simple/ --trusted-host pypi.repo.test.hhu.de torch==2.0.0+cu117 --upgrade
```

- Line 2: Required resources: 1 CPU and 6GB RAM
- Line 3: Required runtime: After 5 hours the job will quit if not already finsihed
- Line 4: Project name that you entered when registering for the HPC
- Line 5: Job name; choose a suitable name
- Line 7: After job is finished you will find the logs in a file with this name
- Line 9: Move to your home folder
- Line 11: Load Python version for which you want to install the Python package
- Line 12: Load CUDA to install torch with CUDA enabled
- Line 14: Command to install Python package “torch” version 2.0.0 with CUDA enabled

- You need to upload this file to HPC via FileZilla if you want to execute it

Submitting jobs via terminal

- You need to connect to HPC via your terminal:



- Once logged-in, you will be in your home folder "/home/user"
  - Operating system is Linux:
    - You can see all the files in your current folder using the command ls
  - You can submit a job using the command $qsub\ name\_of\_bash\_script.sh$
  - You can see the status of your jobs
    - $qstat\ -u\ username$
      - Status "Q" stands for queued; "R" stands for running; After job is executed it will vanish from this list
    - You can monitor your jobs on myjam3.hhu.de

Running a Python Script

Heinrich Heine
Universität Düsseldorf

- Implement Bash script to execute Python script:

```
1   #!/bin/bash
2   #PBS -l walltime=01:30:00
3   #PBS -l select=1:ncpus=1:mem=8GB
4   #PBS -A kcat-prediction
5   #PBS -N run_python_script
6   #PBS -j oe
7   #PBS -o "run_python_script.log"
8
9   cd /home/alkro105
10  module load Python/3.8.3
11  python3 /home/alkro105/python_script.py
```

Chang to project name that you entered when registering for the HPC

- Upload your Python script and your bash script to your HPC folder via FileZilla
- Submit Bash script via terminal