

Working with the Charité high performance cluster (HPC)



- 1) Obtain access to the **cluster**: write an email to sc-hpc-helpdesk@charite.de requesting access to the cluster from your charité email mentioning your Charité username (typically 5-7 characters, alphanumerical)
 - 2) Obtain access to the **cluster documentation**: visit <https://git.bihealth.org/charite-sc-public/sc-wiki/> , when first trying to access this, you are prompted to apply for access, which you should do (one sentence description of what you want to do with the cluster is sufficient) and which will hopefully be granted quickly
- Access the cluster via command line at
 - `ssh <username>@s-sc-frontend1.charite.de`

Working with the Charité high performance cluster (HPC)



```
libuchauer — [REDACTED]@s-sc-frontend2:~ — ssh [REDACTED]@s-sc-frontend2.cha...

Last login: Mon Jun  2 10:42:43 on ttys000
[libuchauer@MBP-DJ52QWFF76 ~ % ssh [REDACTED]@s-sc-frontend2.charite.de
[REDACTED]@s-sc-frontend2.charite.de's password:
Welcome: Scientific Computing Team @Charite
This is frontend node s-sc-frontend2.charite.de of the Charite
Compute Cluster.

You can find the User Documentation, various HOWTOs and contact
information in our Gitlab Wiki:

https://git.bihealth.org/charite-sc-public/sc-wiki
Have a lot of fun!: Scientific Computing Team @Charite
Last login: Mon Jun  2 10:42:51 2025 from 10.43.72.188
[[REDACTED]@s-sc-frontend2 ~]$ ls
PBMC_lens  ards  covid_lav  data  examples  projects  scratch.txt  software
[[REDACTED]@s-sc-frontend2 ~]$
```

Working with the Charité high performance cluster (HPC)



- 1) Obtain access to the cluster: write an email to sc-hpc-helpdesk@charite.de requesting access to the cluster from your charité email mentioning your Charité username (typically 5-7 characters, alphanumerical)
 - 2) Obtain access to the cluster documentation: visit <https://git.bihealth.org/charite-sc-public/sc-wiki/> , when first trying to access this, you are prompted to apply for access, which you should do and which will hopefully be granted quickly
- Access the cluster via command line at
 - `ssh <username>@s-sc-frontend1.charite.de`
 - **Access the jupyter hub (recommended for beginners and interactive analyses)** at
 - <https://s-sc-hub.charite.de/hub/>

Working with the Charité high performance cluster (HPC)



A screenshot of a web browser showing the JupyterHub login page. The browser's address bar displays the URL `s-sc-hub.charite.de/hub/login?next=%2Fhub%2F...`. The page features the JupyterHub logo at the top left. In the center, there is a login form with an orange header containing the text "CHARITÉ" and "Sign in". Below this, the form asks for a "Username:" and a "Password:", each with a corresponding text input field. At the bottom of the form is an orange "Sign in" button.

A screenshot of the "Server Options" page. The title "Server Options" is centered at the top. Below it, the text "Select a job profile:" is followed by a dropdown menu that currently shows "Standard Compute". Below the dropdown is a large orange button labeled "Start".

A screenshot of the server status page. At the top, it says "Your server is starting up." followed by "You will be redirected automatically when it's ready for you." Below this text is a long, light gray progress bar. Further down, the text "Cluster job running... waiting to connect" is displayed. At the bottom left, there is a link labeled "Event log".

Working with the Charité high performance cluster (HPC)



JupyterLab (auto-G) interface showing the Launcher view. The left sidebar shows a file browser with folders like 'ards', 'covid_lav', 'data', 'examples', 'PBMC_lens', 'projects', 'software', 'jupyterhub...', and 'scratch.txt'. The main area displays the Launcher with options for R, Terminal, Text File, Markdown File, Python File, R File, and Show Contextual Help. The bottom status bar shows 'Simple' mode and 'Launcher' tab.

JupyterLab (auto-G) interface showing the D4_metacells.ipynb notebook. The left sidebar shows a file browser with folders like 'metaout', 'A-r_Hughe...', 'archive_D...', 'B-r_Hughe...', 'C-p_plotti...', 'circ_meta...', 'circadian_...', 'CircMed_...', 'D1_merge...', 'D2_make_...', 'D3_circadi...', 'D4_metac...', 'Droin_adat...', 'Droin_circ...', 'Droin_mt_...', 'GSE11923...', 'GSE15913...', 'Hughes_2...', 'Hughes_2...', 'Hughes_2...', and 'Hughes_2...'. The main area displays the notebook content with a title 'Metacells object from circadian gene space' and code cells. The bottom status bar shows 'Simple' mode and 'D4_metacells.ipynb' tab.

```
[40]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import scanpy as sc
import anndata as ad
sns.set_context('paper')
import metacells as mc

import utils
import plots
from importlib import reload
reload(utils)
reload(plots)
from utils import sample_colors, sum_counts_by_
from plots import plot_scatter_gene_expression_

from scipy.stats import circmean

[60]: # read single cell data
adata_circadian = ad.read_h5ad("Droin_circadian
adata_circadian.obs['time'] = adata_circadian.c
#adata_circadian.var_names_make_unique()

[61]: adata_circadian.X = adata_circadian.layers["col

[62]: adata_circadian.X.sum(axis = 1).mean()

[62]: 21.241701

[63]: adata_all = ad.read_h5ad("Droin_mt_ribo_filtere
```

Working with the Charité high performance cluster (HPC)



- 1) Obtain access to the cluster: write an email to sc-hpc-helpdesk@charite.de requesting access to the cluster from your charité email mentioning your Charité username (typically 5-7 characters, alphanumerical)
 - 2) Obtain access to the cluster documentation: visit <https://git.bihealth.org/charite-sc-public/sc-wiki/> , when first trying to access this, you are prompted to apply for access, which you should do and which will hopefully be granted quickly
- Access the cluster via command line at
 - `ssh <username>@s-sc-frontend1.charite.de`
 - Access the jupyter hub (recommended for beginners) at
 - <https://s-sc-hub.charite.de/hub/>
 - **The cluster can be accessed only from within the Charité network**
 - **If you want to work remotely, you need VPN access with extra privileges** (“Zusatzantrag O” filled according to the instructions here <https://git.bihealth.org/charite-sc-public/sc-wiki/-/wikis/Resources/User%20Documentation/User%20Guide:%20HPC%20@Charite#connection-to-the-cluster-from-outside-the-charite-network-over-vpn> and screenshot on the next slide)

Working with the Charité high performance cluster (HPC)



Connection to the cluster from outside the Charite network over VPN

If you need to connect to the cluster through VPN, you have to apply for [VPN permissions](#) (Antrag O). The current list of the relevant hosts and ports is here:

- s-sc-frontend1.charite.de, 22/ssh und 443/https
- s-sc-frontend2.charite.de, 22/ssh und 443/https
- s-sc-frontend3.charite.de, 22/ssh und 443/https
- s-sc-hub.charite.de, 443/https
- s-sc-metrics.charite.de, 443/https

To simplify the VPN application process, the VPN-Team of the Charite IT (vpn@charite.de) has defined the group rule `charite-hpc-user-access`. You can simply enter the name of this group in your VPN application (Erweiterte Berechtigung »Spezifische Dienste« -> Zielsystem: Gruppenregel charite-hpc-user-access)

Note on the VPN application VPN permissions are neither necessary nor sufficient for cluster access. A granted VPN application does not imply login access to the frontend nodes. To be granted access to the cluster, please send a request to the SC Helpdesk as explained above.

Working with the Charité HPC – live demo of jupyter hub, documentation and dashborad



- **Access to jupyter hub** <https://s-sc-hub.charite.de/hub/>
- **General user guide** (includes info on how to use prebuilt conda environments to notebook kernels): <https://git.bihealth.org/charite-sc-public/sc-wiki/-/wikis/Resources/User-Documentation/User-Guide:-HPC-@Charite>
- How to submit jobs to SLURM (large and/or repetitive jobs, like alignment jobs, that do not require interaction) <https://git.bihealth.org/charite-sc-public/sc-wiki/-/wikis/Resources/User-Documentation/User-Guide:-HPC-@Charite>
- **Dashboard** of the cluster, shows how busy it is currently <https://s-sc-metrics.charite.de/>
- Applying for larger data storage folders works via email, as outlined here (PI or data owner involvement required) <https://git.bihealth.org/charite-sc-public/sc-wiki/-/wikis/Resources/User-Documentation/User-Guide:-HPC-@Charite#data-storage>
- Connecting **Rstudio** to the cluster is possible, but requires some fiddling <https://git.bihealth.org/charite-sc-public/sc-wiki/-/wikis/Resources/HOWTOs/Rstudio>

Of note - Charité HPC is not the same as BIH HPC



- In addition to the system described above, BIH also provides a Compute Cluster which can be used by Charité and MDC users
- The BIH HPC is extremely well documented, see here <https://hpc-docs.cubi.bihealth.org/>
- Obtaining access works via the group leader who first needs to apply for a user group to be installed
- Similar to Charité HPC, it also includes an interactive portal, the OnDemand portal <https://hpc-portal.cubi.bihealth.org/pun/sys/dashboard>
- This allows interactive jupyter notebook sessions as well as Rstudio sessions