Connecting to our servers

We offer two servers called **CN45** and **CN106**, both these servers are available for intensive computational work. Both of these servers are connected to our **CEPH** partition, which is able to store a large amount of data. All server addresses in this document are constructed as follows:

- **<SERVER_NAME>.science.ru.nl**

CN45 and CN106 can only be reached from the RU network.

This means you'll either:
- Need to be in the office
- Connect to the network via VPN
- SSH tunnel through lilo, a server provided by CNCZ

They can be reached by either an SSH client or an FTP client. Both servers connect to the same **CEPH** structure but have a separate scratch directory. All servers require your science account.

Requesting access

To request access you will need to email slrinzema@science.ru.nl your science username and the team you are a part of. Please note that a username is required, this is the same as your login for your science email and is usually a combination of your first name's first letter plus your last name.

Accessing the RU-network

Besides being at the office you can use one of the following two VPN's:
- eduVPN ([info here](#))
- Cncz's VPN ([info here](#))

Or you can tunnel through the Linux login server provided by CNCZ called lilo. Its address is:

- **lilo.science.ru.nl**

When accessing a server with mobaXterm (specified in the next chapter) you can set **lilo** as an **SSH gateway**. This setting can be found under **network settings** (for mobaXterm)**.**
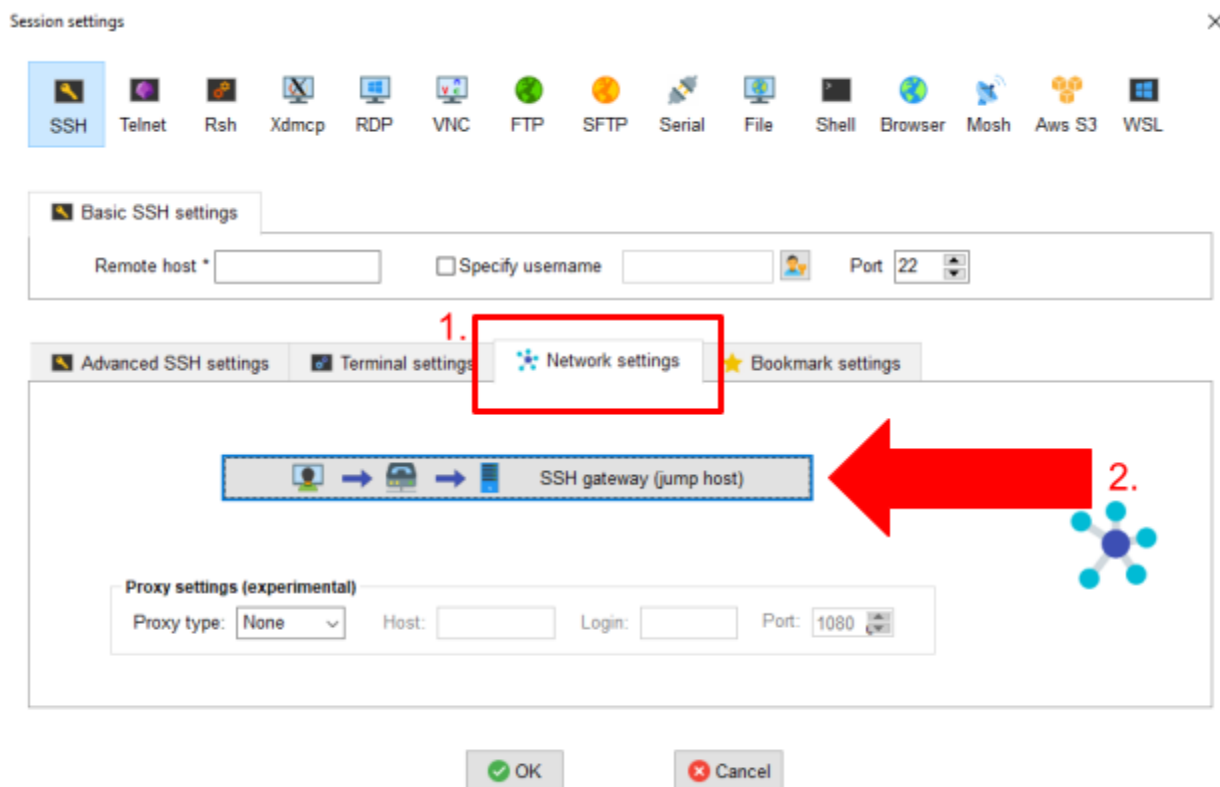

Fig 1. Mobaxterm SSH gateway option

When clicking the button enter **lilo.science.ru.nl** as the gateway host, your science account username, and 22 for the port.

When using a native ssh client on Linux or macOS you can use the following command:

```
ssh <USERNAME>@lilo.science.ru.nl
```

Accessing a server

Our server addresses are:

- **cn45.science.ru.nl**
- **cn106.science.ru.nl**

When using mobaXterm you can click the **session** button in the top left and select the **SSH** option. In the Basic SSH settings section, you can enter one of the addresses mentioned above as the **remote host**, select **specify username** and enter your **science account username**.
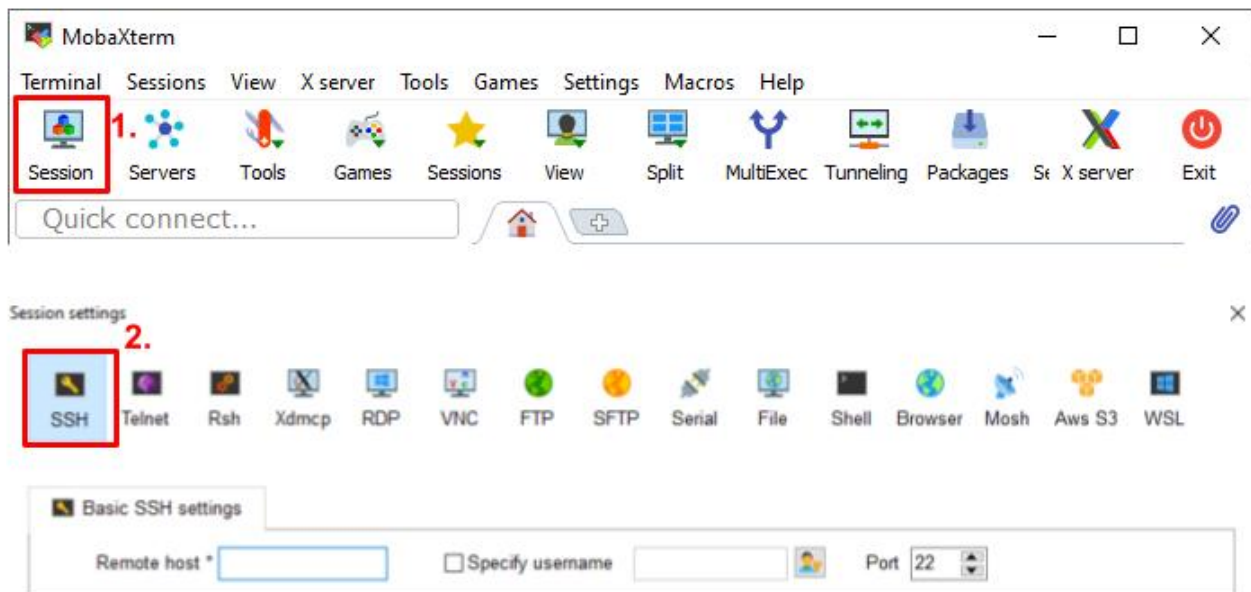

Fig 2. MobaXterm ssh session options.

When using a native ssh client the command is as follows (using cn45 as an example)

```
ssh <USERNAME>@cn45.science.ru.nl
```
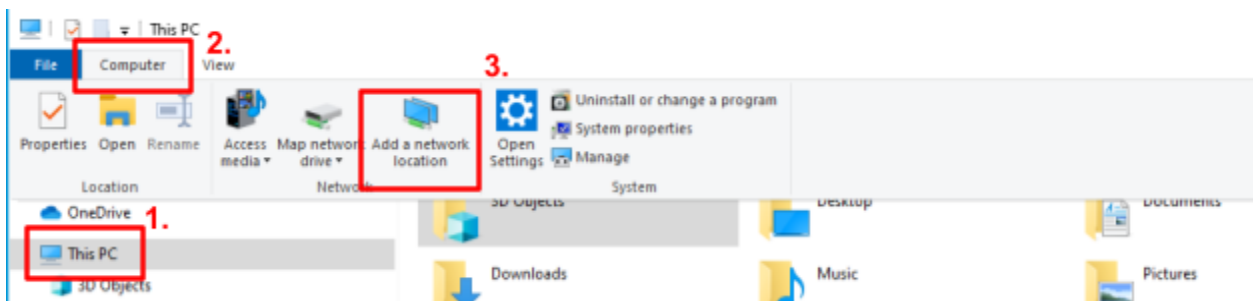
Mapping a network drive

There are a lot of external programs (WinSCP or FileZilla for example) to connect to an ssh server and move data around. While this is a viable option, we also offer the ability to map a **department** directory stored on our **CEPH** partition directly to your pc. Essentially making it just another directory on your pc.

Mapping a network drive requires the address. Our three departments, **cellbio**, **molbio,** and **moldevbio**, all have their own addresses:

- **\\molbio-srv.science.ru.nl\cellbio**
- **\\molbio-srv.science.ru.nl\molbio**
- **\\molbio-srv.science.ru.nl\moldevbio**

<div style="background-color:red; color:white; text-align:center">

DO NOT STORE DATA DIRECTLY IN THE DEPARTMENT DIRECTORY.
PLEASE CHECK
THE CHAPTER <u>DATA STORAGE</u> FOR INFORMATION ON HOW TO STORE YOUR DATA

</div>

Mapping a network drive in **Windows 10** is done by opening file explorer. Selecting "This PC", click the "computer" tab on top and select "Add a network location".

Fig 3. The first steps to mapping a network location in Windows 10.

After these steps, you can continue until you are required to specify an address. Here you enter the address you want to map and use your **science account** to get access.

For **Mac OS** the steps are as follows:
From the mac os finder hit Command+K to bring up the "connect to server" window.
Enter the path to the network drive you want to map, please note that mac requires you to prepend "smb:" to the address and replace "\" with "/":

- **smb://molbio-srv.science.ru.nl/<DEPARTMENT>**

Remote port forwarding

Some programs, such as jupyter notebook, run in a browser on the server. You can access this browser when using X forwarding, but with considerable delays. Alternatively, you forward a port from cn45 to your PC and remove the delay. To do so run the following command (with your username) in a local terminal. This can be a "local session" in mobaXterm:

```
# SSH, with port forwarding, over a gateway server
ssh -fN -l username -L 2222:cn45:22 lilo.science.ru.nl && \
ssh -fN -l username -p 2222 -L 8888:localhost:8888 localhost;
```

Next, run jupyter notebook as usual on the server (without browser to speed it up) jupyter notebook --no-browser, and access it in your browser at http://localhost:8888/.

- Note 1: in this example, port 8888 on cn45 (:8888 in the command) is forwarded to port 8888 on your PC (8888: in the command).
- Note 2a: you connect to cn45 over port 2222. If this port is refusing connection you can try any other.
- Note 2b: you tunnel tp cn45 through lilo over port 22. If this port is refusing connection you can also try port 80 or 443.
- Note 3: if you have multiple notebooks/programs running that normally output to port 8888, they will output to a different port (like 8889).
- Note 4: the port forwarding will keep running until you kill it (or exit mobaXterm or restart your machine).

Data storage and directory structure

Our server is divided into a few different locations. Most of these are shared between servers, except the scratch. Only CEPH is important if you're not planning on working on our servers and only want to store large data.
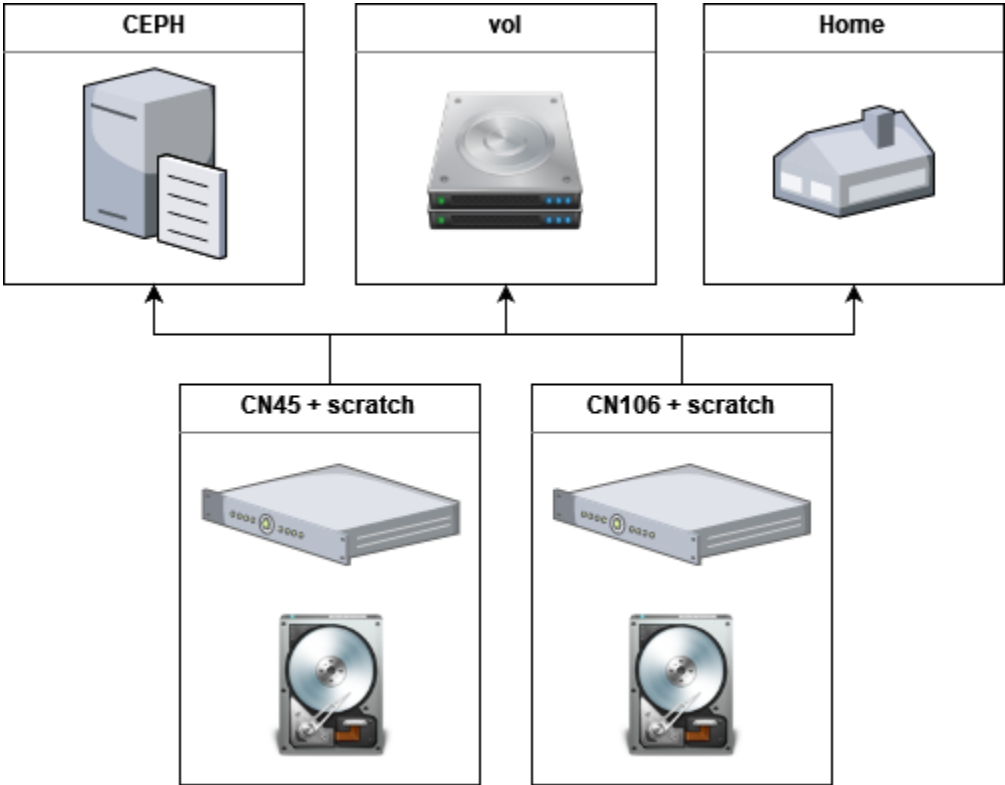
Fig 4. An overview of the server structure.

Data storage

We provide a large data storage called **CEPH**. its base path is **/ceph/rimlsfnwi**, it is further divided into three subdirectories:

data

This location is used for the storage of large data to work on or as a **backup**. Here it is protected against disk failure, removed or changed data is not recoverable. To protect against user error you could store your data here and work on a copy elsewhere.

The data directory is subdivided into our departments: **cellbio**, **molbio,** and **moldevbio**. These are in turn subdivided by **research teams** which are the surnames of the team leaders.

Please find the department/team directory that you belong to and make your personal directory, a good practice is to name it after your science account. An example is:

- **/ceph/rimlsfnwi/data/moldevbio/heeringen/slrinzema**

Or a more generalized path is:

- **/ceph/rimlsfni/data/<DEPARTMENT>/<TEAM>/<SCIENCE ACOUNT>**

Mapping a department directory as a network drive is the same as entering the path from:

- **/ceph/rimlsfni/data/<DEPARTMENT>**

raw_data

Here is where data originating from our machines and generated fastq files are stored.

web_share

This location is web-accessible and can be used to host and share data or make a trackhub. It is divided into two subdirectories **mbdata** and **trackhub.** In mbdata you can make a personal directory and trackhub is used to host trackhubs, but is currently writable by a limited user group. The URLs for these locations are:

- **https://mbdata.science.ru.nl**
- **https://trackhub.science.ru.nl/<SUB DIRECTORY>**

They can be mapped as a **network drive** as follows:

- **\\molbio-srv.science.ru.nl\mbdata**
- **\\molbio-srv.science.ru.nl\mbtrackhub**

Workin on data

There are two primary locations (besides your local computer) you can use to work on data. When you log in through an ssh connection you start in your **home directory**. The other location is **scratch** which is unique for each server.

Home

This location is provided by CNCZ and more information is provided on their [wiki](#). It has limited storage space and automated emails will be sent if you are nearing capacity. As it is used by a lot of basic Unix processes and environment variables it is suggested you do not store large data here. Limit it to small files and small scripts.

Its path is:

/home/<USERNAME>

Scratch

The scratch disk is used to store large data while you work on it. Copying data here while working on it protects it from overwriting or accidental deletion. However, it is a shared space so please be considerate of others and don't leave your data on here forever. The base path is:

/scratch/

Conda

We provide a shared disk for your conda environments at:

/vol/mbconda

These environments can get out of control in terms of storage so please be considerate. If you have never used conda please check out [the following guide](#). **Please refrain from installing conda yourself,** instead, use our provided script. Use the following command to install:

```
bash /vol/mbconda/install_conda.sh
```

Data backup

CEPH

Ceph has no 'snapshot' backup ability, as it is generally too large for this feature. This means you can't restore your data to how it was yesterday. If you require 'snapshot' backup or version control please read about **surfdrive**. It can be used to store data that isn't being worked on right away.

- A great option to store data you're not working on
- **Is not** protected against user error (deletion or changes)

Surfdrive

Surf provides a cloud backup service of up to 500Gb and provides version control.
Surfdrive can be accessed here: https://surfdrive.surf.nl/files

- Has version control, meaning you can reset your files to a previous version

Mbconda

Our conda disk is being backed up by cncz. DO NOT STORE DATA HERE