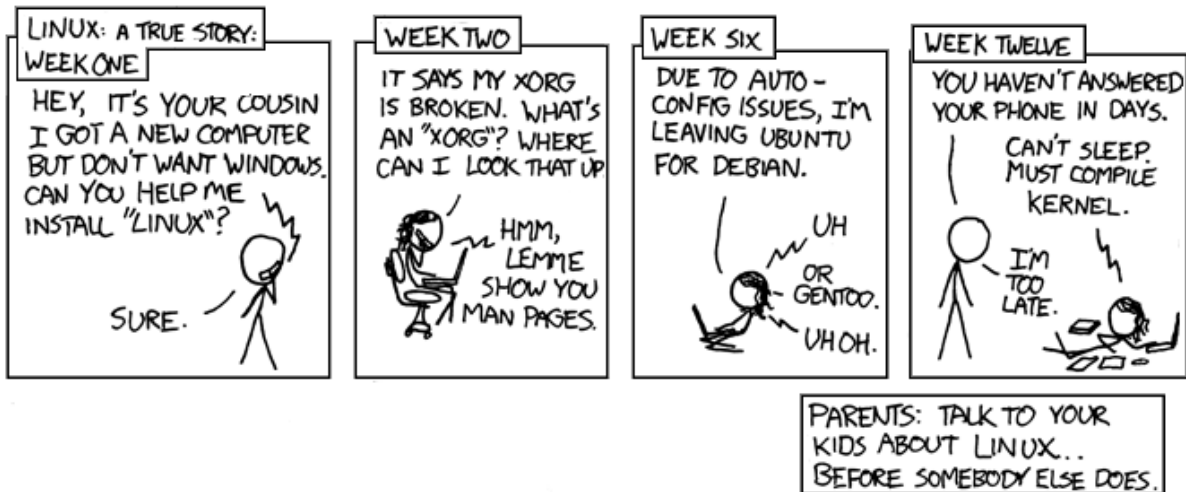# Usage of Cirrus

Sebastian Schubert and Anselm Arndt

admin-klimageo@lists.hu-berlin.de

30th May 2018

Contents

1 Access to Cirrus and file transfer

First, the administrator of cirrus has to create a user account.

SSH login with password is only allowed from some computers of the Geographical Department. From outside (and optionally also from the aforementioned computers), users can login with an SSH key. If only access with password is required, skip the respective steps in the following.

Note that after 5 wrong attempts to login, your IP address will be banned for one day. Please contact the administrator to remove this ban.

1.1 From Linux

1.1.1 Creation of an SSH key (for access from outside of Geography department, otherwise optional)

Create a SSH key (optionally with a password):

```
$ ssh-keygen -t rsa -b 4096 -f ~/.ssh/id_rsa
```

If you are currently situated at computer within the Geography Department's net, enable the key with

```
$ ssh-copy-id USERy@cirrus.geo.hu-berlin.de
```

Alternatively, send the administrator of cirrus the created public key `~/.ssh/id_rsa.pub`. The content of the public key will be added on the cluster to `~/.ssh/authorized_keys`.

1.1.2 SSH login

Use

```
$ ssh -X USER@cirrus.geo.hu-berlin.de
```

to login (substitute USER with your username on the cluster).

1.1.3 Access to web services

The web page of cirrus, https://cirrus.geo.hu-berlin.de/ is only available from within the Geography Department. In order to access the page, an SSH key (section 1.1.1) is required. After executing

```
$ ssh -N -L 1234:localhost:443 USER@cirrus.geo.hu-berlin.de
```

in a terminal and keeping that terminal open, open https://localhost:1234/ in a browser. The SSL certificate error can be ignored.

1.1.4 File transfer

Files can be transfer with the `scp` programme:

```
$ scp /source/file/on/local/machine USER@cirrus.geo.hu-berlin.de:/target/file/on/cirrus
```

or

```
$ scp USER@cirrus.geo.hu-berlin.de:/source/file/on/cirrus /target/file/on/local/machine
```

1.2 From Windows

Download PuTTY. In addition to the main program, you will need PuTTYgen for the generation of an SSH key. If you require to run graphical applications on Cirrus (aka X11 applications), download VcXsrv and install it. For file transfer from and to Cirrus, WinSCP is required. If you only want to transfer data, you need only PuTTYgen to generate an SSH key (section 1.2.1) and WinSCP (section 1.2.3).

1.2.1 Creation of an SSH key (for access from outside of Geography department, otherwise optional)

For access with an SSH key, start PuTTYgen and create an RSA key with 4096 (bits):
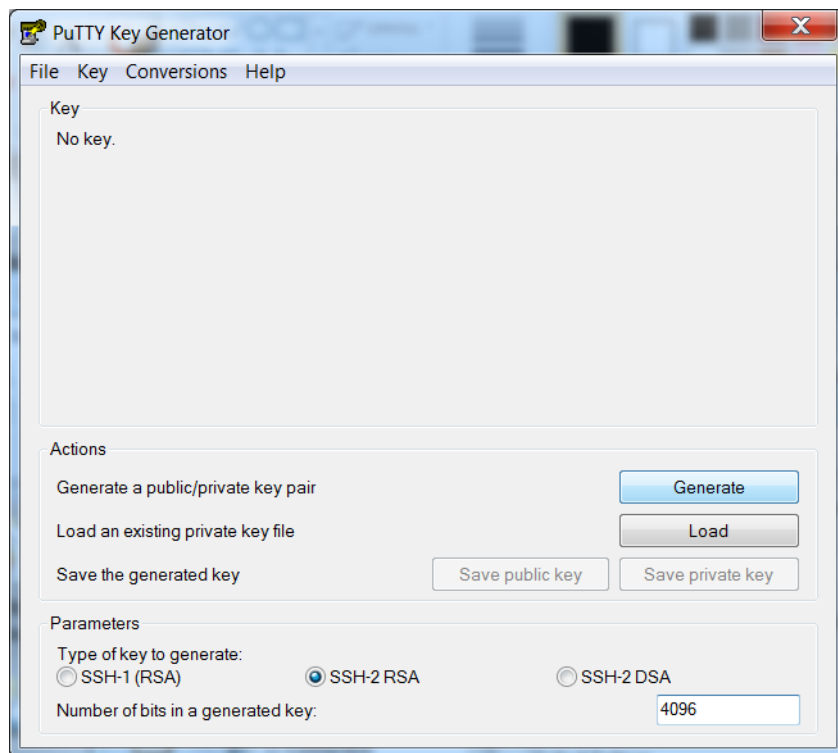

The usage of a password is optional. The results look like
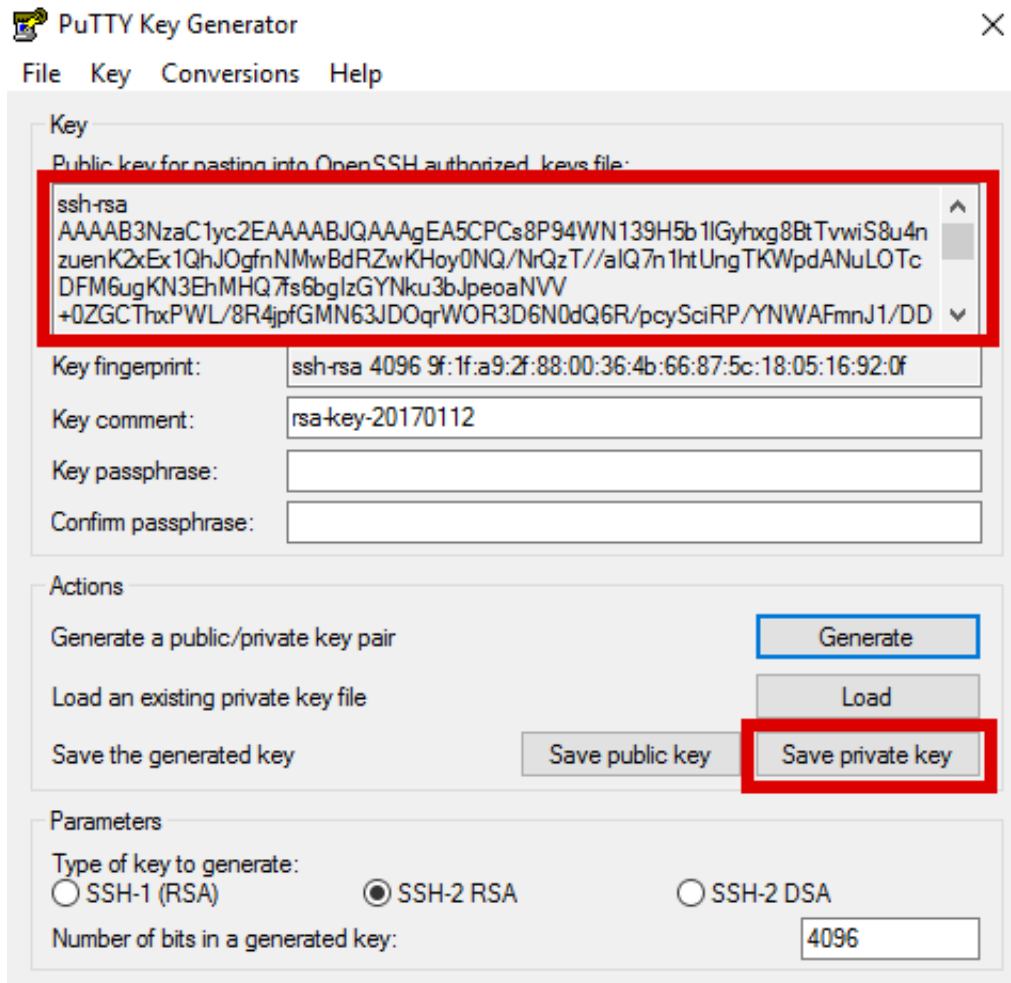
Please save the private key inaccessible to other users on you system. Furthermore, please send the administrator of Cirrus by email the public key, i.e. the content of the window marked above. The administrator will add the public key to `~/.ssh/authorized_keys` in your home directory.

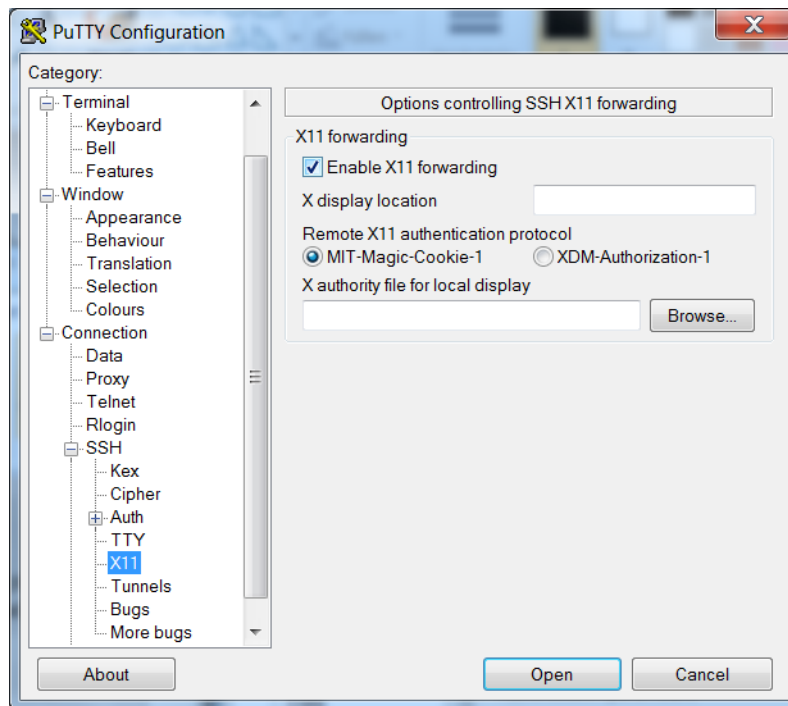### 1.2.2 SSH login and access to web services

Start VcXsrv (optional for graphical applications) and PuTTY. Enter `USER@cirrus.geo.hu-`
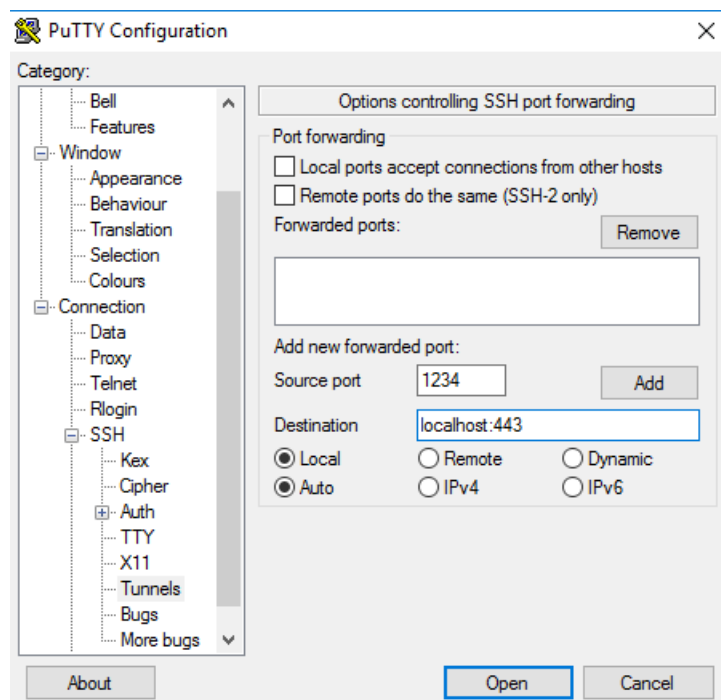`↪ berlin.de` as the Host Name, where `USER` is your username on the cluster:
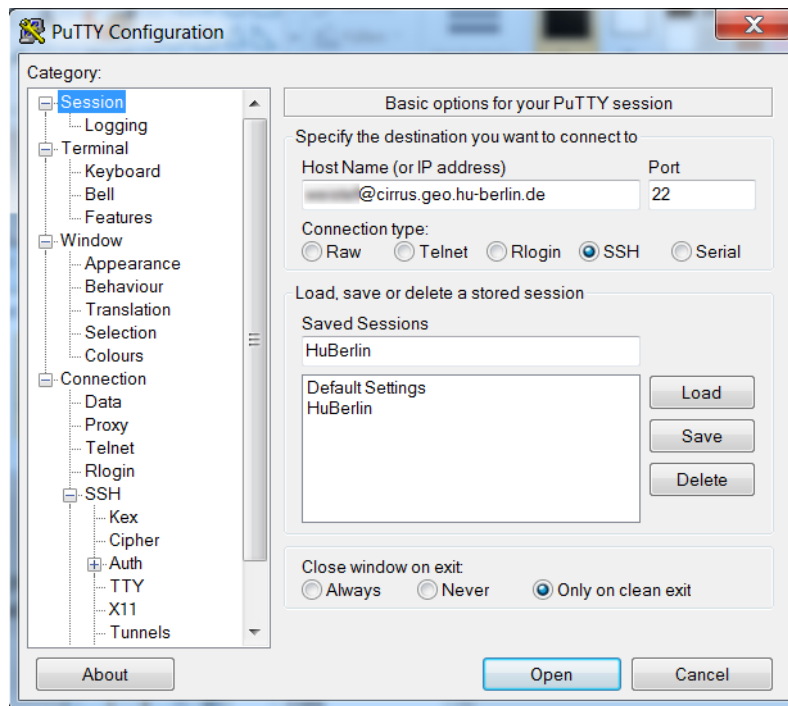
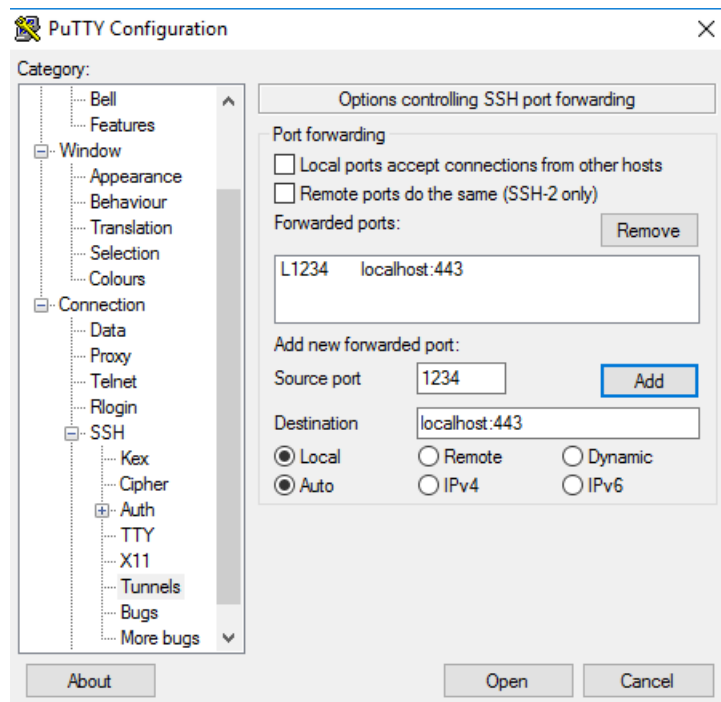Enter the location of the private SSH key:
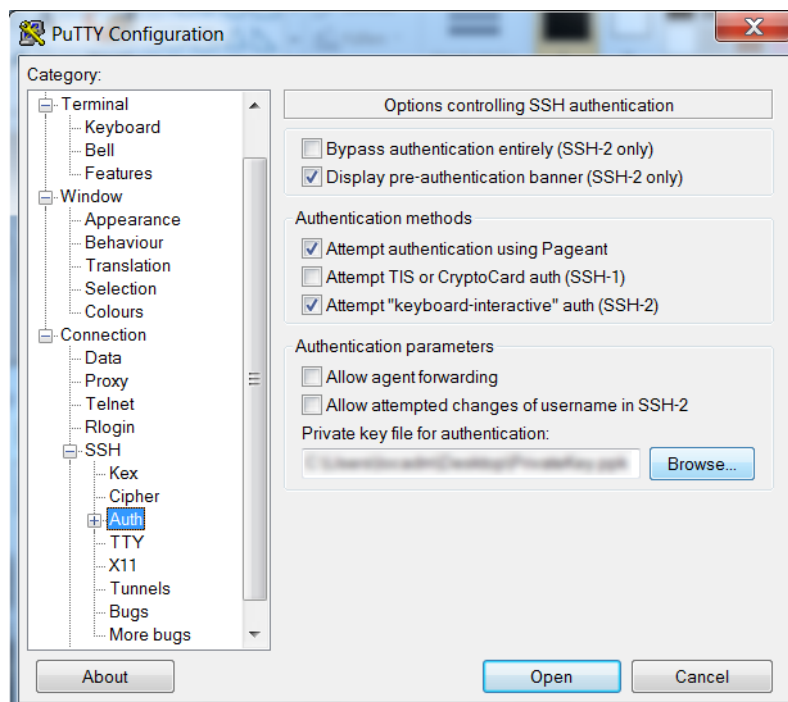
Optionally, enable X11 forwarding for graphical applications:

Furthermore, if you require access to Cirrus' web page and you are not within the net of Geography Department, enter the following in the Tunnel section:

Click "Add" to produce

After all settings are done, save the Session as indicated on the first screenshot of the PuTTY Configuration. Click Open to start the connection.

1.2.3 File transfer

Start WinSCP. Select File protocol "SCP", enter Host name `cirrus.geo.hu-berlin.de`, your User name and passwort:
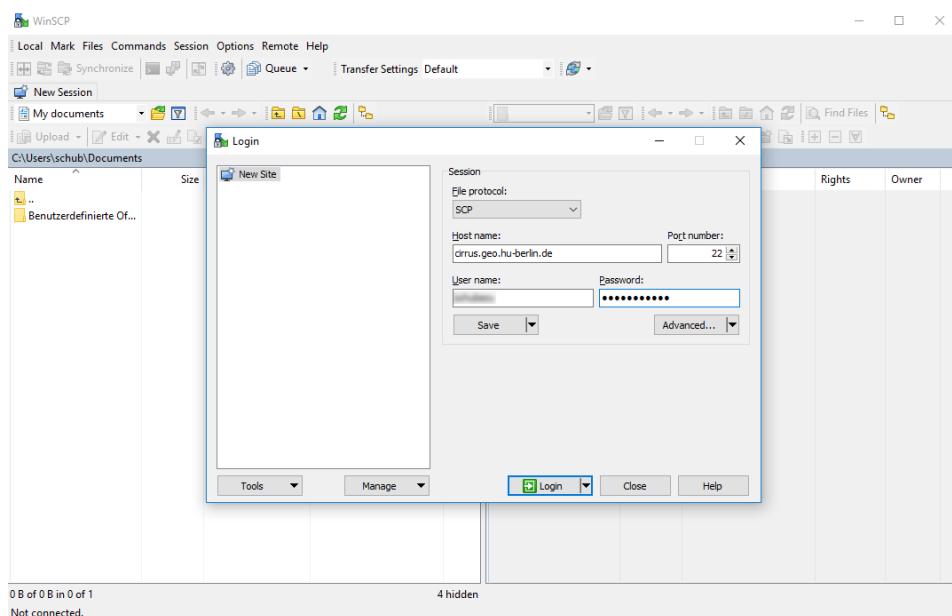
If you require the SSH private key for access from outside of the Geography department, click on "Advanced..." and select the "Private key file" in SSH → Authentication.
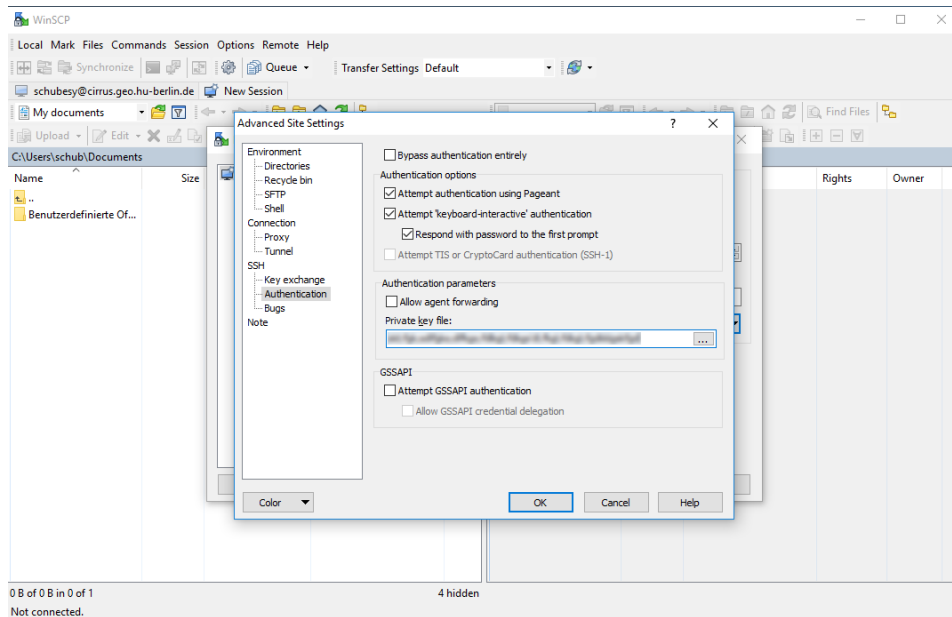
In the main window, press "Login". When you connect for the first time, you will be asked to confirm the SSH fingerprint. Press yes here.
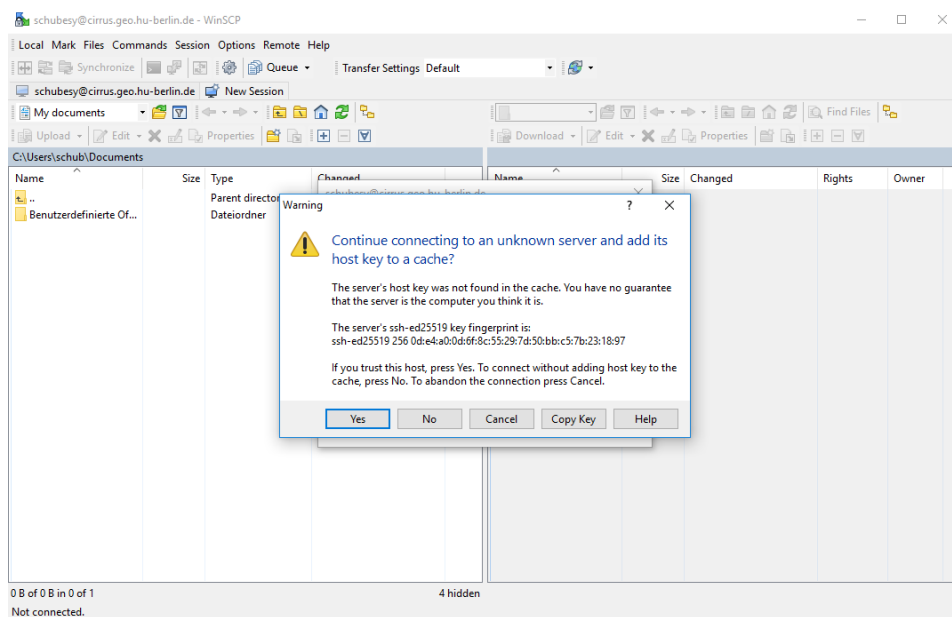
Once successfully connected, you will you computer on the left-hand side of the WinSCP window and Cirrus on the right-hand side. You can copy files with drag-and-drop from one side to the other.
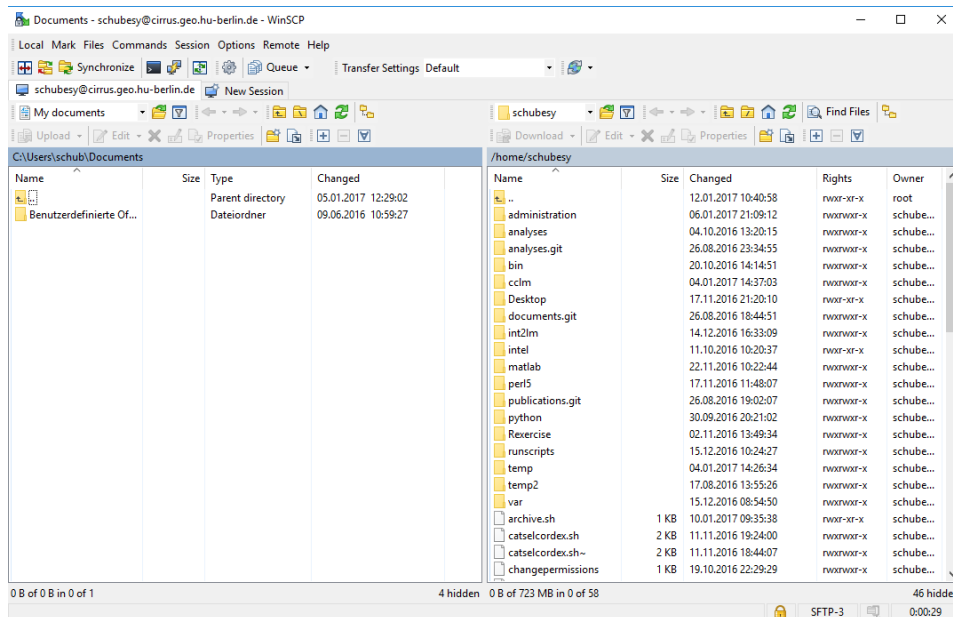
By default, the shown folder on Cirrus is your home folder (see next section). You can navigate to the folders under `/data/` by double-clicking on `..` twice. Here, double-click on `data` and then either on `scratch` or on `projects`.

2 Folder structure

On a Linux system, all folders are integrated in one large folder tree; there are no drives. Cirrus is set-up in such a way that for users three folders are important:

/home/USER home folder of each user; not suited for data intensive storage; daily backup

/data/scratch/USER personal workspace for data intensive tasks; no backup

/data/projects group workspace for data intensive tasks, no backup

The current usage and quota in each of the folders can be queried with

```
$ xfs_quota -c 'quota -h'
```

3 Submitting jobs to the batch queuing system

On Cirrus, the batch queuing system SLURM is installed. It is used to distribute computation jobs to the compute nodes to which users do not have direct access.

3.1 The batch job file

In order submit a job to the queuing system, a batch job of the form

```
#!/bin/bash

#SBATCH --job-name="SOME STRINGS"
#SBATCH --qos=short
```

```
#SBATCH --account=XYZ
#SBATCH --nodes=5
#SBATCH --ntasks-per-node=20
##SBATCH --ntasks=100
#SBATCH --time=24:00:00
#SBATCH --workdir=/path/to/workdir
#SBATCH --output=/path/to/%j.out
#SBATCH --error=/path/to/%j.out
#SBATCH --mail-type=ALL


...
some_commands
...
```

The lines starting with `#SBATCH` are commands for SLURM: The most import parameter is the number of compute cores you require. You can either give the total number of cores with `ntasks`, or the number of `nodes` times `ntasks-per-node`. Note, that every node has 20 cores.

The usage of cores and wall time is limited depending on the chosen `qos` (quality of service). At the time of writing, there are three classes available:

short up to 120 cores for up to 1 day

medium up to 20 cores for up to 7 days

long up to 5 cores for up to 30 days

The current set-up can be queried with

```
$ sacctmgr show qos format=name,priority,maxwall,mintres,maxtres,grptres
```

The maximum `time` depends on the `qos`. Entering a lower `time` than the maximum allowed wall time in the `qos` can improve the queuing position.

Furthermore, every user is associated with one or different accounts. There are accounts for projects, teaching and a misc account. The account will only be used for cluster reports. You can find your accounts with

```
$ sacctmgr show assoc where user=...
```

and you default account with

```
$ sacctmgr show user
```

The default `workdir` as well as file names for `output` and `error` streams can be set-up. Note that no folders are created are by SLURM.

The `job-name` is a string for easy recognition of the job. With the `mail-type` option you enable automatic email notification.

After defining all required SLURM parameters in the header of the batch script, an arbitrary number of bash commands can be given. Note that programs that require MPI communication have to be started with `srun -n $SLURM_NTASKS`. The `-n $SLURM_NTASKS` argument tells `srun` how many cores are acquired by this job:

```
srun -n $SLURM_NTASKS mpi_programm
```

3.2 Submitting a job and the queue

A batch script is submitted to the SLURM queue with

```
$ sbatch batchscript
```

With

```
$ squeue
```

all current running and queue jobs are listed. The general load of the cluster is printed with

```
$ sinfo
```

Alternatively, the GUI program

```
$ sview
```

is available.