

# Setup your VM on surf research cloud

R. Grouls  
Version 0.3

September 2, 2024

# Contents

<b>1</b>	<b>Create a VM</b>	<b>2</b>
1.1	Introduction . . . . .	2
1.2	Creating a VM . . . . .	2
1.2.1	Requesting a wallet . . . . .	2
1.2.2	Workspace access setup . . . . .	4
1.2.3	SSH key . . . . .	4
1.2.4	New workspace . . . . .	4
<b>2</b>	<b>Connect to your VM</b>	<b>8</b>
2.1	Pausing and Resuming your VM . . . . .	8
2.2	Finding your IP address . . . . .	8
2.3	checking authorized keys . . . . .	10
<b>3</b>	<b>Setup your VM</b>	<b>11</b>
3.1	Clone and install the course . . . . .	11
3.2	Optional: eyecandy . . . . .	11
3.2.1	starship . . . . .	11
<b>4</b>	<b>TLDR</b>	<b>12</b>

# Chapter 1

## Create a VM

### 1.1 Introduction

For this semester, we will be using the SURF Research Cloud to run virtual machines. This has a lot of advantages:

1. Uniformity of hardware and software, so we can help you better.
2. Lesson material is tested on the VMs, so you can be sure that everything will work.
3. Instructions are optimized for the VMs, so you can follow them step by step.
4. We will have GPUs available on the VMs, so you can run deep learning models with considerably less time.

This manual will guide you through the process of setting up your VM on the SURF Research Cloud. The first requirement is that you should receive an invite to the SURF Research Cloud. If you have not received an invite, please contact us. We will need your @student.han.nl email address to send you an invite.

### 1.2 Creating a VM

Click on the invite link in your email. Log in with your @han email account. Note: make sure to drop the student part of your email address. So if your email is name@student.han.nl, you should use name@han.nl to log in.

#### 1.2.1 Requesting a wallet

The first screen you encounter is an overview where you can enter the SURF Research Cloud, see Figure 1.1. Click on "Open".

After this, you should see a screen similar to Figure 1.2. Here you can request a wallet.

#### **The wallet**

The wallet is a way to pay for the VMs. You will get a budget of 5000 credits, and it depends on which VM you choose how many credits you will use. The VM we will use costs around 50 credits per 24 hours for a 2 core CPU with 16GB of RAM, and a GPU will cost you about 400 credits per 24 hours. Put as a subject "SRC account for MADS-HAN" and the SURF helpdesk will approve your request, typically during workhours on the same day.

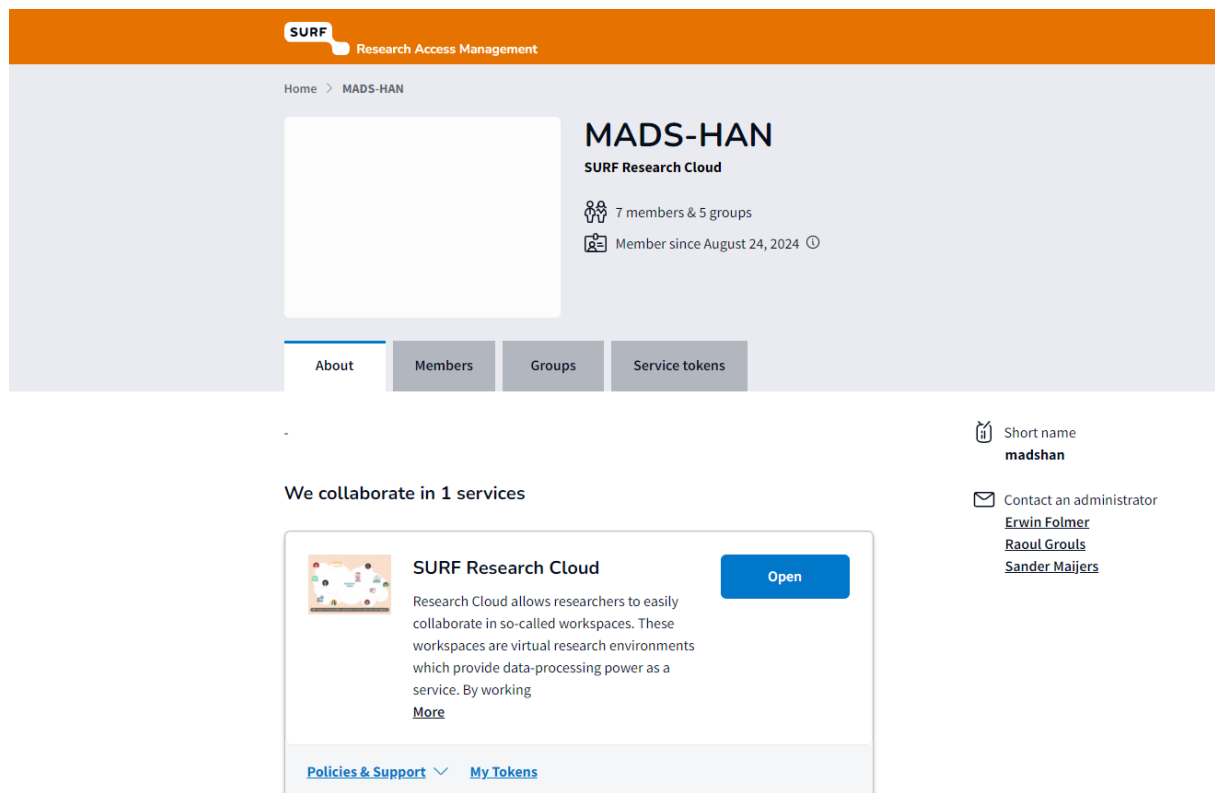


Figure 1.1: Open the SURF Research Cloud

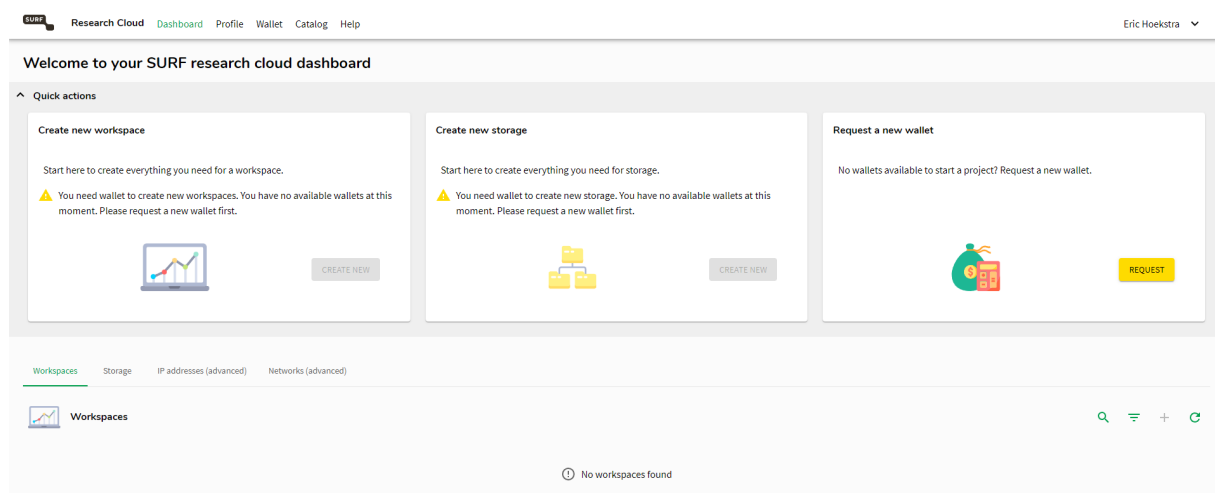


Figure 1.2: Request a wallet

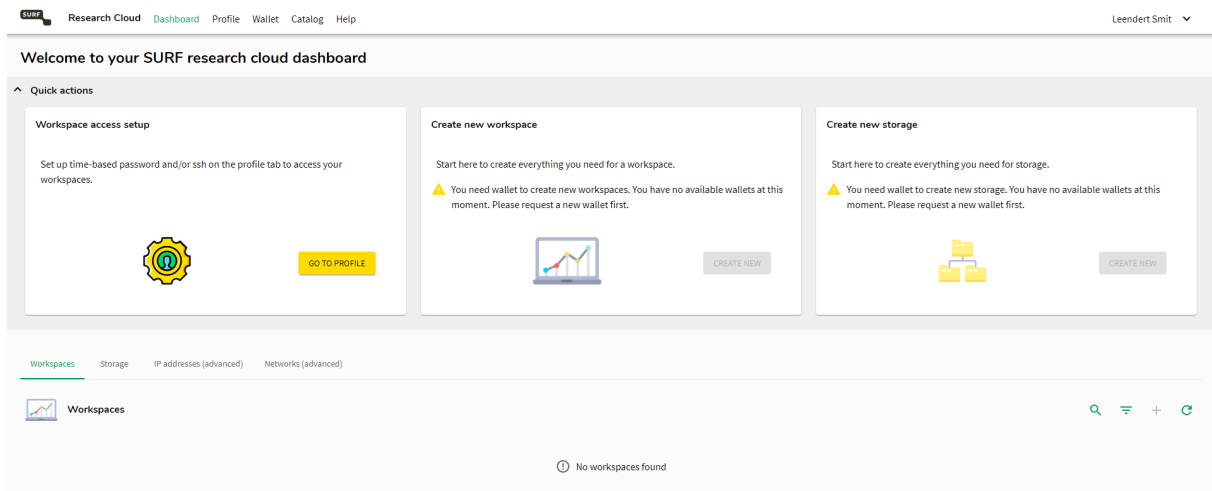


Figure 1.3: Choose 'workspace access setup'.

## 1.2.2 Workspace access setup

After you get your wallet approved, you should see something as shown in Figure 1.3. Choose “workspace access setup”. This will bring you to a screen where you need to set up your SSH key.

### 1.2.3 SSH key

You will need a SSH key. If you don’t already have one and need a refresher on how to create one, see the following link: <https://docs.github.com/en/authentication/connecting-to-github-with-ssh/generating-a-new-ssh-key-and-adding-it-to-the-ssh-agent>

You should add your public key to the workspace setup.

#### The SSH keys

1. Please, pay attention **where** you store your keys on your system. If necessary, write down the location. On linux and mac it will typically be in `~/.ssh`, on windows it can be anywhere depending on your setup. Your public key is a string of characters inside a text file; if you open the `.pub` file with a text editor, you should see something like `ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQgQDQ...`. This is what you need to copy-paste into the SURF Research Cloud.
2. Never share your private key, only the public key (ending in `.pub`).

## 1.2.4 New workspace

After setting up SSH, you should be able to create a workspace like in Figure 1.4.

You will run to some choices. Note that probably step 1 and 2 will be automatically selected and thus skipped. Choose the following:

1. Choose a collaborative organisation: MADS-HAN
2. Choose a wallet: SRC account for MADS-HAN
3. Catalog item: search for 2204 and choose Ubuntu 2204 - SUDO enabled, see Figure 1.5.
4. Choose cloud provider: SURF HPC Cloud

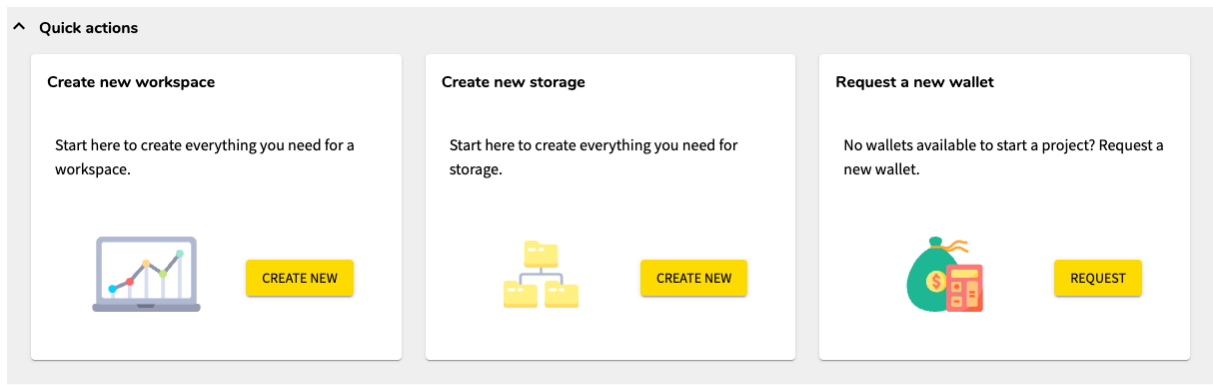


Figure 1.4: Choose 'Create new' under workspace.

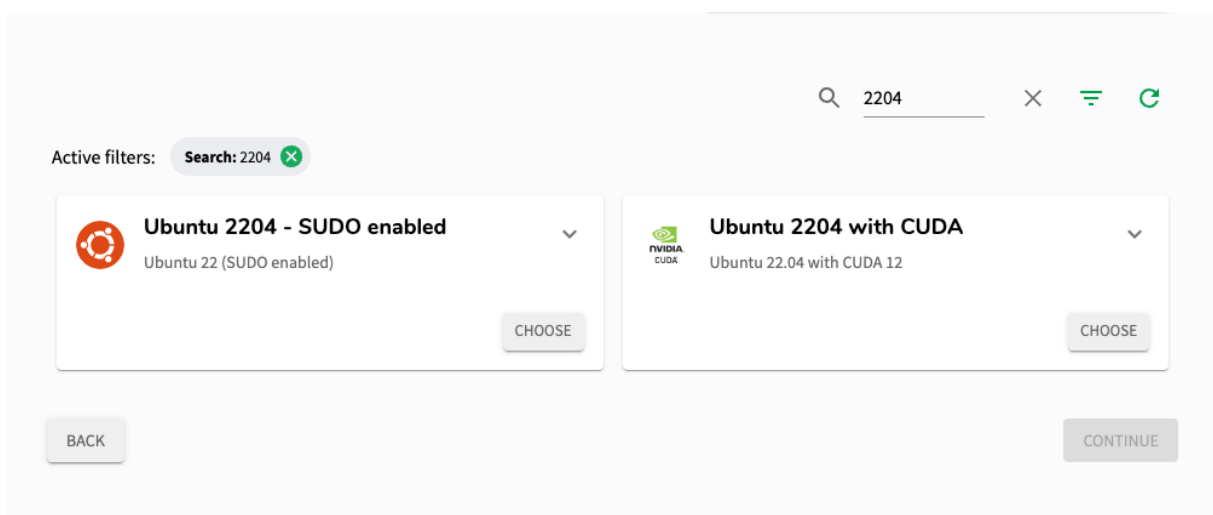


Figure 1.5: Choose Ubuntu 2204















Overview of your choices		
	<b>Collaborative organisation</b> MADS-HAN	
	<b>Wallet</b> SRC account for MADS-HAN	
	<b>Catalog item</b> Ubuntu 2204 - SUDO enabled	
	<b>Dataset(s)</b> No datasets selected	
	<b>Cloud provider</b> Cloud provider: SURF HPC Cloud Flavours: Ubuntu 22.04, 2 Core- 16 GB RAM 	
	<b>Options</b> Options: No storage selected Options: No IP address selected Options: No network selected	
	<b>Name</b> End date: 01-08-2025 (02:21) Workspace name: ubuntu 2204 - sudo enabled Hostname: ubuntu2204sudo Description: not set	

Figure 1.6: Your overview should look like this

5. Choose size: 2 Core - 16GB RAM
6. Choose the expiration date of the machine: 01-08-2025
7. Workspace name: use your LastnameInitial-semestername. For example, mine would be groulsr-UOS1 or groulsr-UOS3.
8. Workspace Hostname: use your LastnameInitial. For example, mine would be groulsr.

After you went through these steps, you should get an overview like in Figure 1.6. We will be experimenting with GPUs later on in the course; but because you have a limited budget of compute, start with the CPU first. For GPU, you should

- a choose 'Ubuntu 2204 with CUDA' in step 3 (see Figure 1.5) and
- b for size choose 'RTX2080 - 1 GPU' in step 5 (there are not enough A10s for everyone available, the RTX should be available for everyone).
- c To discern them, add 'GPU' to the workspace name and hostname in step 7. For example, mine would be 'groulsr-UOS3-GPU' and 'groulsrgpu'.

After clicking "submit" you should see your VM being created. This can take a while.

#### Naming conventions

1. UOS1 (first semester class 2024) and UOS3 (third semester class 2023) will be creating VMs in the same workspace, so it is **really important** to stick to the naming conventions! We will **delete** VMs that do not adhere to the naming conventions!
2. use your LastnameInitial-semester for the workspace name and LastnameInitial for the hostname. Add -GPU for a GPU VM, eg. groulsr-UOS1 or groulsr-UOS3-GPU etc.



# Chapter 2

## Connect to your VM

### 2.1 Pausing and Resuming your VM

You will get 5000 credits, which equals about 200 euros. Running a small VM (2 cores, 16GB) will cost you about 50 credits per day, which means you would be able to keep it running for about 100 days. Please note that even if you don't actually use the VM, it will still cost you credits: it is still "on", regardless if you are logged in. So, when you are not using the VM, please pause it. You can do this in two ways:

1. Log in to the SURF Research Cloud, navigate to your Dashboard, and click "Resume" or "Pause" on your workspace. Note that it might very well be possible we will have more than 60 VMs here, which is the reason why we are strict on the naming conventions, see 1.2.4 for more details.
2. Use `surfcontroller`, a package we developed in Python. You can install it by running `pip install surf-controller`. You can then run `surfcontroller` from your terminal and select your VMs to pause/resume, see Figure 2.1.

#### Autoshutdown @ 21:00

1. We will **automatically pause all VMs every day at 21:00**. Experience learns that there will always be people who forget to shutdown their VM, and if you forget your GPU for a week you will spend most of your credits.
2. If you want to do long computations (eg for hypertuning your model) you can request a lift from the autoshutdown with your teacher.

### 2.2 Finding your IP address

After the VM is created, you can use the dropdown menu to find the URL with the IP address of your VM. You can use this IP address to connect to your VM, see Figure 2.4. Make sure your VM is running before you try to connect. You can open VScode and when you have installed the "Remote Development" extension from Microsoft, you will have a small blue icon in the bottom left corner, see Figure 2.2. Click on it and choose "Remote-SSH: Connect to Host...", see Figure 2.3. You can now enter `user@IP-address` of your VM and connect to it. You will find this under 'URL' in the dropdown menu. For example, I would enter `rgrouls@145.38.192.86` to connect.

The first time you will get a question saying this is an unknown host and if you want to add its fingerprint. You can accept this and continue. After this, you should have entered your VM. You can now start working on your VM.

```
[ ] rgrouls-MADS(paused)
[ ] rgrouls-MADS-GPU(paused)

== Username : rgrouls == surfcontroller version 0.3.1 ==
Press
'j' to move down,
'k' to move up,
'Enter' to select,
'a' to select all,
'f' to toggle filter,
'n' to rename user,
'p' to pause,
'r' to resume,
'u' to update status,
's' for ssh access,
'l' to toggle logs,
'q' to quit
```

Figure 2.1: Surfcontroller screenshot

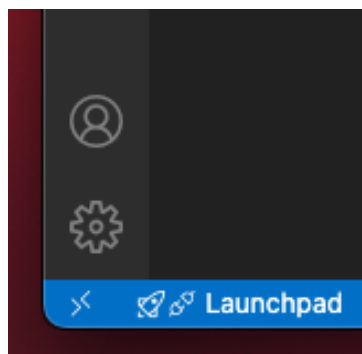


Figure 2.2: Click the small blue icon in the bottom left corner

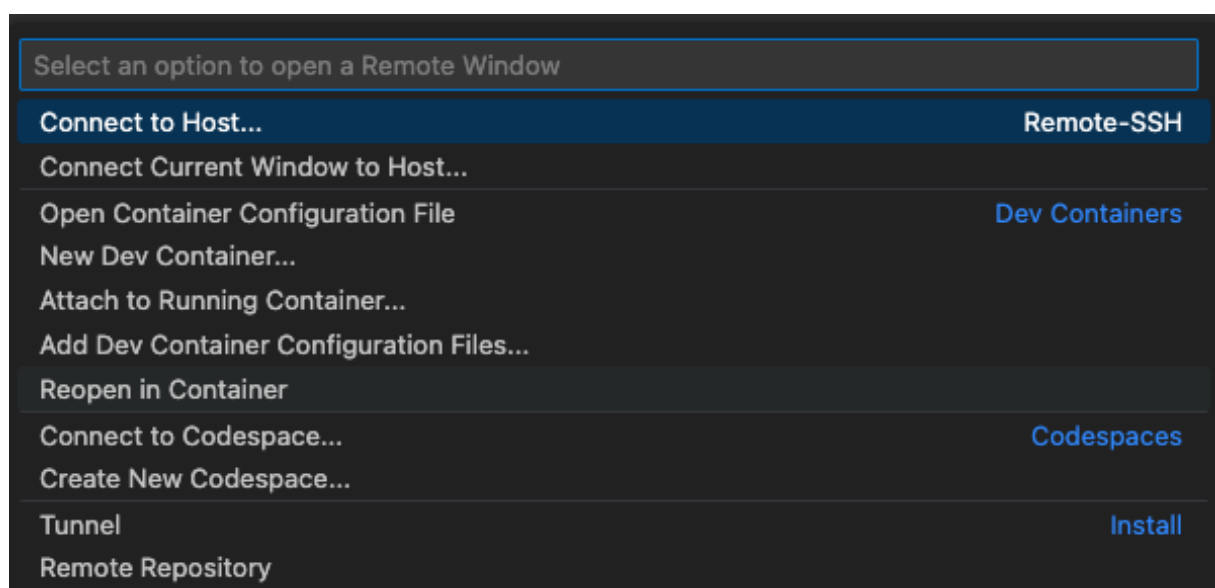


Figure 2.3: connect to host

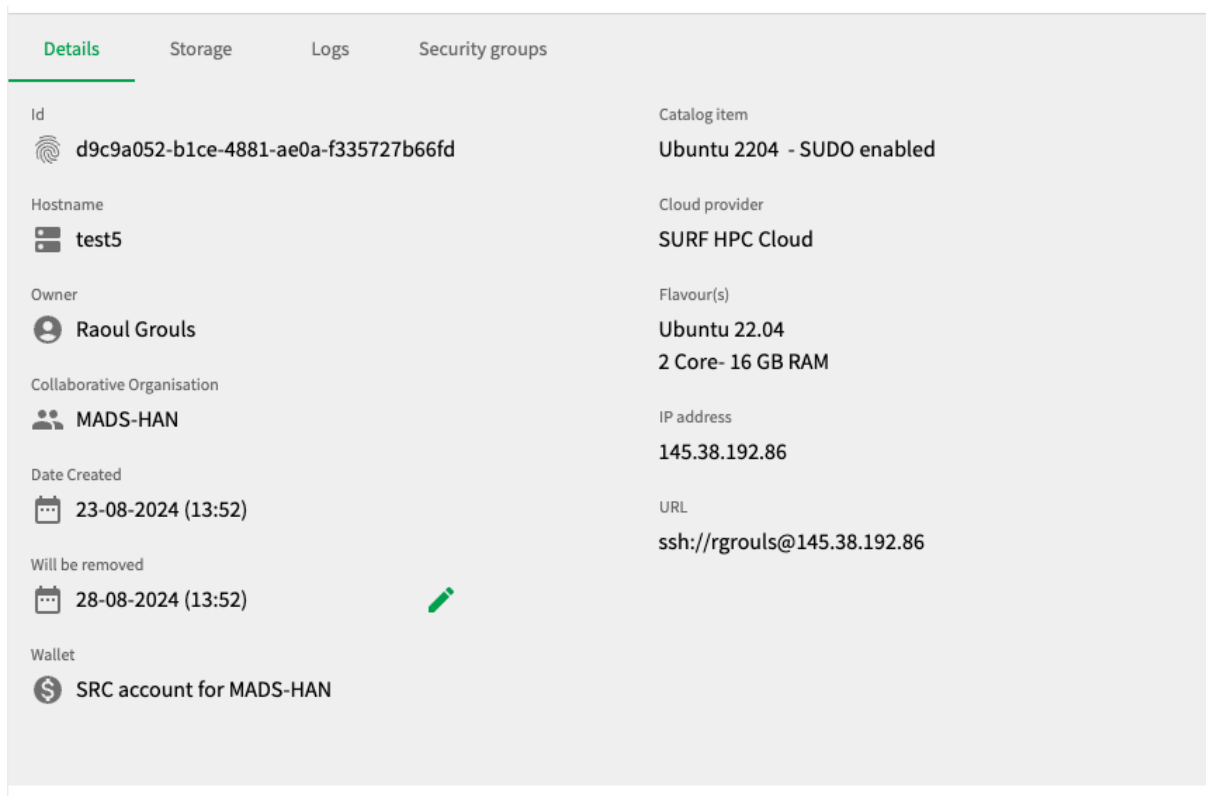


Figure 2.4: Find you IP address and url

## 2.3 checking authorized keys

If you still need to enter a password when connecting, you might need to check your ssh key. The ssh-key will serve as a way to prove you are you, and you can connect to the VM without a password. To do this, you need to add your public key to the authorized keys on the VM.

**On the VM:**

1. navigate with cd to '`~/ssh`'
2. check if there is a file called '`authorized_keys`'
3. if not, create one with '`touch authorized_keys`'
4. open the file with '`nano authorized_keys`'
5. copy-paste your public key in this file <sup>1</sup> and save the file.

<sup>1</sup>see <https://docs.github.com/en/authentication/connecting-to-github-with-ssh/generating-a-new-ssh-key-and-adding-it-to-the-ssh-agent> for details

# Chapter 3

## Setup your VM

### 3.1 Clone and install the course

In the README of the course <https://github.com/raoulg/MADS-MachineLearning-course>, you can find the instruction to 1. install python and rye 2. clone the repo 3. setup git and install the requirements.

### 3.2 Optional: eyecandy

I typically like to install some packages for eyecandy.

#### 3.2.1 starship

1. run `curl -fsSL https://starship.rs/install.sh | bash` in your terminal.
2. open your `~/.bashrc` file with `nano ~/.bashrc`
3. add `eval "$(starship init bash)"` to the file

For more info, or screenshots, got to <https://starship.rs/>

## Chapter 4

# TLDR

### TLDR

please read all the boxes similar to these and you're good.