Table of Contents

Introduction

This document shows you how to set up RStudio on a virtual machine (VM). By the end, you will have a computer that can be scaled to your R-programming needs.

The target audience:

- Researchers
- Data scientists
- Lab technicians

Programming in RStudio is beyond the scope of this guide and will not be covered.

Primer concepts (optional)

The following are useful concepts and terms to understand this document.

Virtual machine

A physical computer (usually in a remote location) that *virtually* divides and allocates its computing resources (i.e., CPU, RAM, memory, etc.). Each allocation is a VM that can run its own operating systems and applications. Users can access the VM through the internet and easily increase/decrease its computing resources.

Linux

A collection of operating systems. Many VM providers offer Linux as a popular option. Different "versions" of Linux are known as distributions. We will use the Ubuntu distribution.

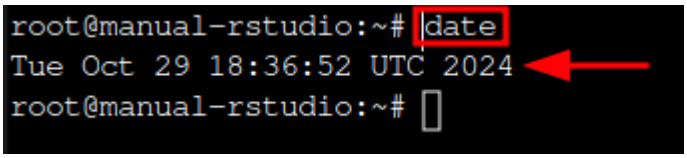
Command-line interface (CLI)

A program used to enter text-only commands to a computer. Commands are typed out and the computer outputs the results.

The following is a typical Linux command-line interface display:

1. The user who is logged into the computer - in this example `root`.
2. The name of the computer - in this example `manual-rstudio`.
3. The current directory - in this example `~` (which denotes the user's home directory).
4. The area to type commands.

The following shows a typical Linux command input and its resulting output:



1. The command `date` is typed and entered.
2. The output is `Tue Oct 29 18:36:52 UTC 2024`.

Requirements

- A computer with internet access
- A DigitalOcean account

We will use DigitalOcean as our VM provider. They offer a free trial and have pre-configured VMs ready to use. DigitalOcean calls their VMs "Droplets".

> Note: Double curly brackets `{{}}` denote placeholders for *your* values.

Quick Setup

Steps to set up a pre-configured Droplet that includes RStudio.

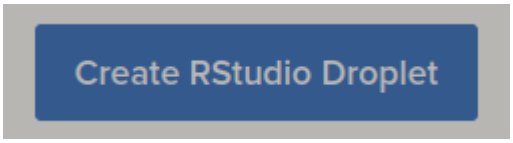
1. Sign in

Log in to your DigitalOcean account.

2. Create the Droplet

Select the pre-configured Droplet: RStudio by Simply Statistics.

Click `Create RStudio Droplet`.



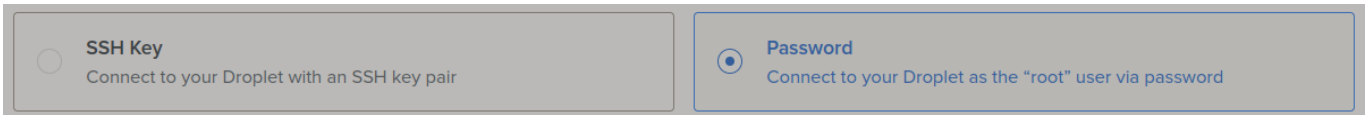
3. Configure the settings

- **Choose Region:** Select the location closest to you.
- **Datacenter:** Leave the defaults.
- **Choose an image:** Select Rstudio by Simply Statistics.

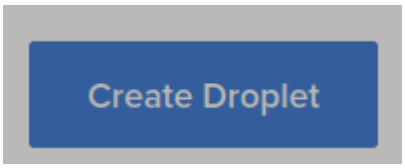
- **Choose Size:** Select the specifications you need. Click here for more specifications details.

> Note: Free trial accounts may need to request access to the Dedicated CPU/Premium CPUs.

- **Backups:** Select if needed.
- **Choose Authentication Method:** Select the Password method for simplicity. Create a password for your Droplet.



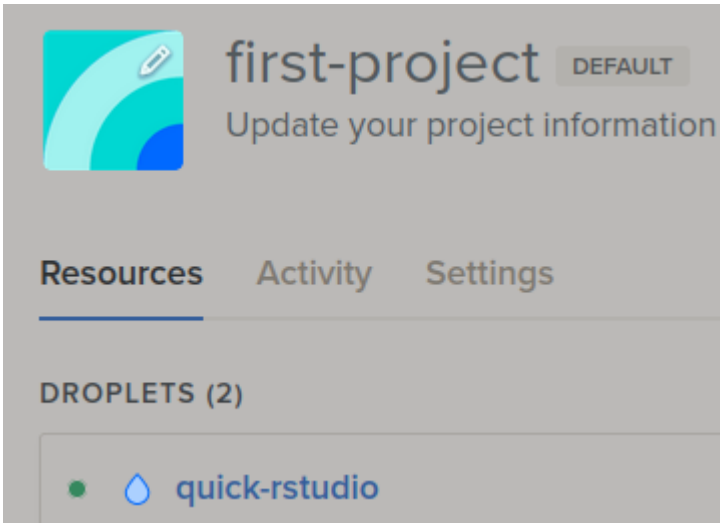
- **Finalize Details:** Change the Hostname to help identify your Droplet. Click `Create Droplet`.



Click here for more settings details.

4. Set up a new user

Navigate to your Projects dashboard. Your newly-created Droplet is shown. Click your Droplet.



Make note of your **ipv4 address** in the top menu bar. To the right, click `Console`.

A CLI window pops up. Look to the bottom for the line `root@{{your-hostname}}:~#`.

```
root@quick-rstudio:~#
```

Enter the following command into the CLI:

```
adduser {{username}}
```

Follow the output instructions:

1. Enter a new password.

   > Note: The password is not displayed while typing.

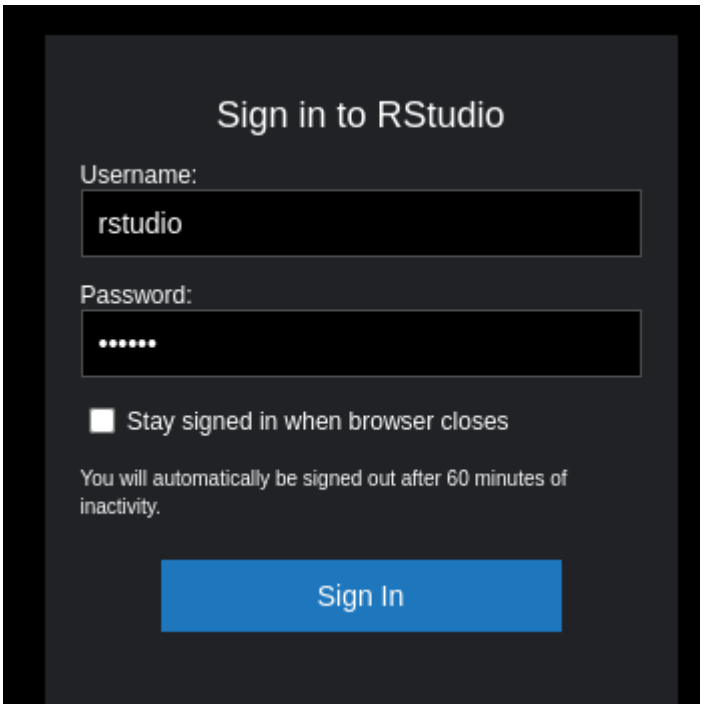
2. Enter the basic information or leave as-is to skip.
3. Enter `y` to confirm the information.

5. Access RStudio

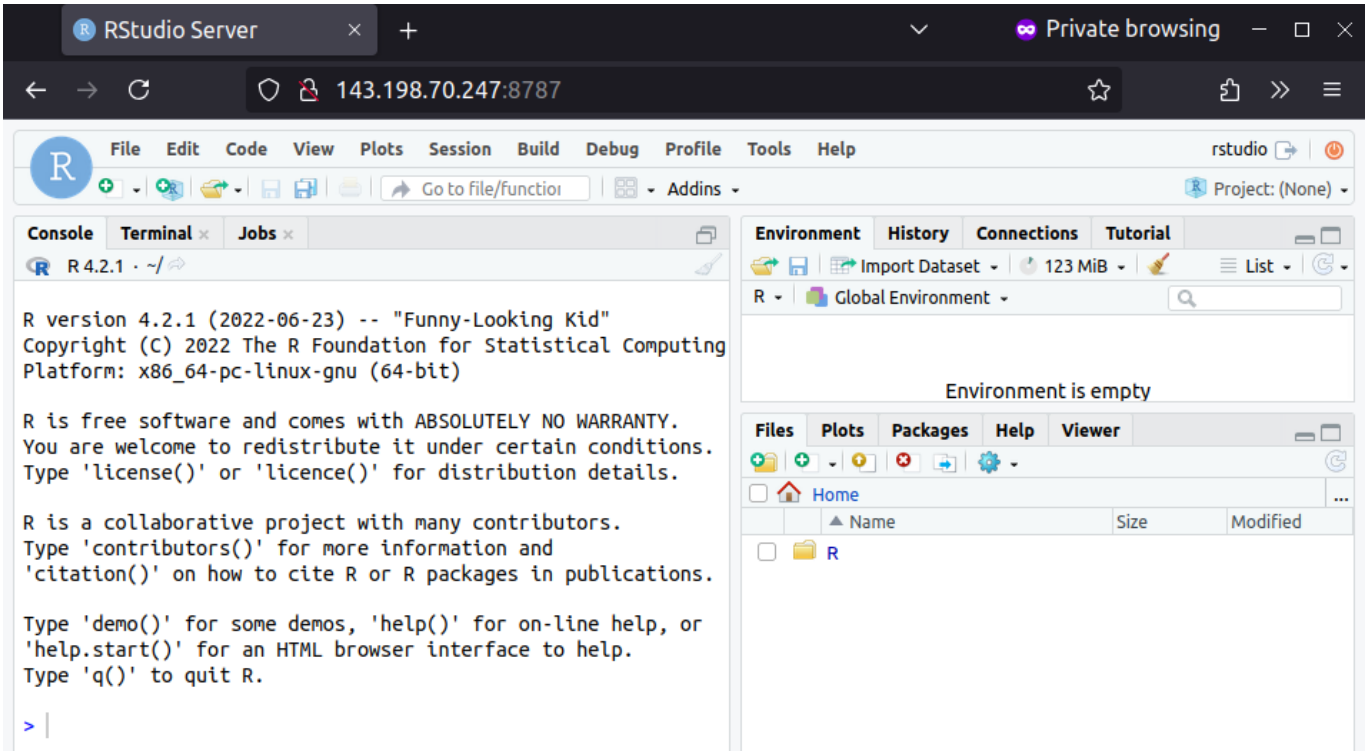
Open a new browser on your local computer. Enter your **ipv4 address** and `:8787` into the URL address bar. It should look like:

```
{{your.ipv4.address}}:8787
```

Enter your new user credentials into the RStudio sign in page.



Congrats! You now have access to RStudio.



Manual Setup

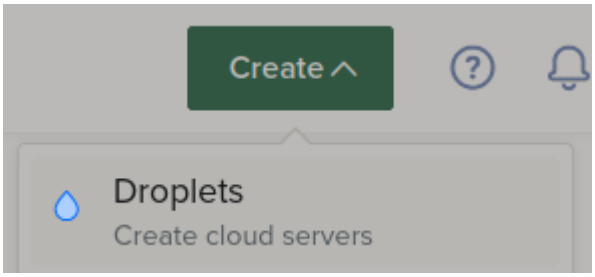
Set up a Droplet with RStudio.

1. Sign in

Log in to your DigitalOcean account.

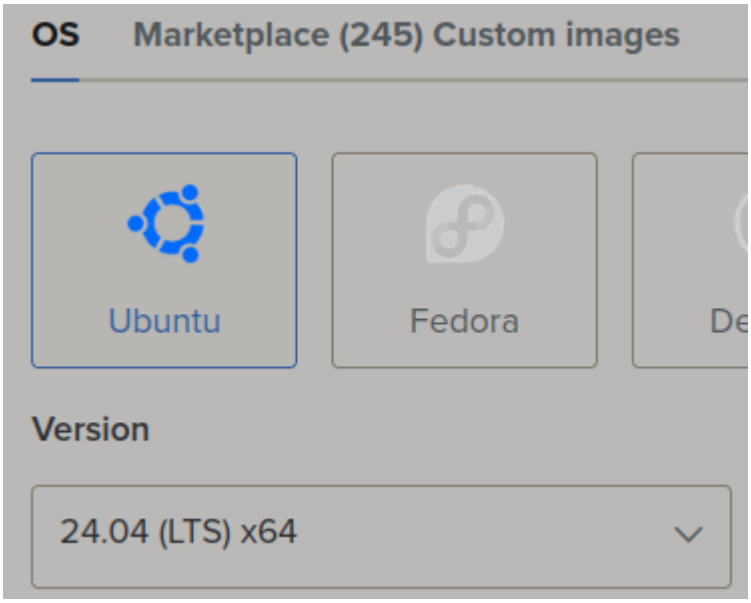
2. Create the Droplet

Navigate to your Projects dashboard. Click `Create` and then `Droplets`.



3. Configure the settings

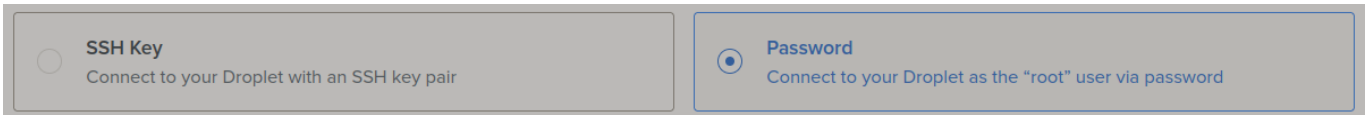
- **Choose Region:** Select the location closest to you.
- **Datacenter:** Leave the defaults.
- **Choose an image:** Leave the defaults Ubuntu and its version.



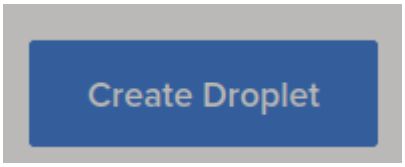
- **Choose Size:** Select the specifications you need. Click here for more specifications details.

> Note: Free trial accounts may need to request access to the Dedicated CPU/Premium CPUs.

- **Backups:** Select if needed.
- **Choose Authentication Method:** Select the Password method for simplicity. Create a password for your Droplet.



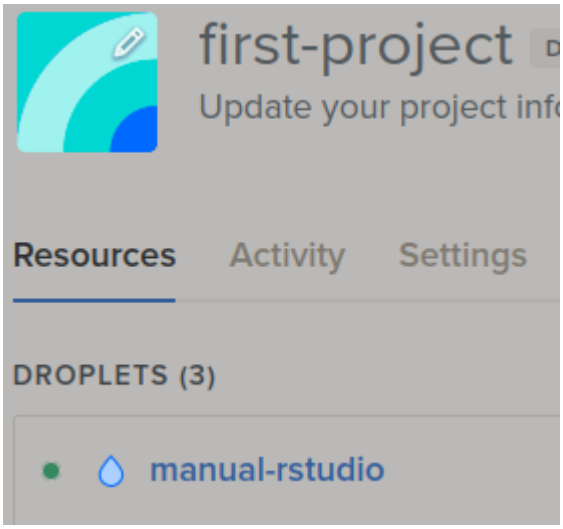
- **Finalize Details:** Change the Hostname to help identify your Droplet. Click `Create Droplet`.



Click here for more settings details.

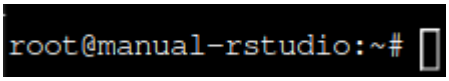
4. Install R and RStudio

Navigate to your Projects dashboard. Your newly-created Droplet is shown. Click your Droplet.



In the top menu bar, note your **ipv4 address**. To the right, click the `Console`.

A CLI window pops up. Look to the bottom for the line `root@{{your-hostname}}:~#`.



> Note: We used Ubuntu version 24.04 at the time of writing. Be sure to check the official install instructions if you are on a different image/version.

Enter the following commands into the CLI:

1. `sudo apt update`
2. `sudo apt upgrade -y`
3. `sudo apt install r-base -y`
4. `sudo apt install gdebi-core -y`
5. `wget https://download2.rstudio.org/server/jammy/amd64/rstudio-server-2024.09.0-375-amd64.deb`
6. `sudo gdebi rstudio-server-2024.09.0-375-amd64.deb`
    - Enter `y` to continue.

> Note: A pop-up window may appear during installation asking for configuration settings. Leave the default choice and press enter to continue.

RStudio Server is now active. You can always check the status with the command:

```
systemctl status rstudio-server
```

```
● rstudio-server.service - RStudio Server
     Loaded: loaded (/usr/lib/systemd/system/rstudio-server.service; enabled; preset: enabled)
     Active: active (running) since Sat 2024-10-26 00:18:36 UTC; 1s ago
    Process: 21513 ExecStart=/usr/lib/rstudio-server/bin/rserver (code=exited, status=0/SUCCESS)
   Main PID: 21514 (rserver)
      Tasks: 1 (limit: 9489)
     Memory: 1.9M (peak: 40.9M)
        CPU: 1.010s
     CGroup: /system.slice/rstudio-server.service
             └─21514 /usr/lib/rstudio-server/bin/rserver

Oct 26 00:18:36 manual-rstudio systemd[1]: Starting rstudio-server.service - RStudio Server...
Oct 26 00:18:36 manual-rstudio systemd[1]: Started rstudio-server.service - RStudio Server.
```

5. Set up a new user

Enter the following command into the CLI:

```
adduser {{username}}
```

Follow the output instructions:

1. Enter a new password.

   > Note: The password is not displayed while typing.

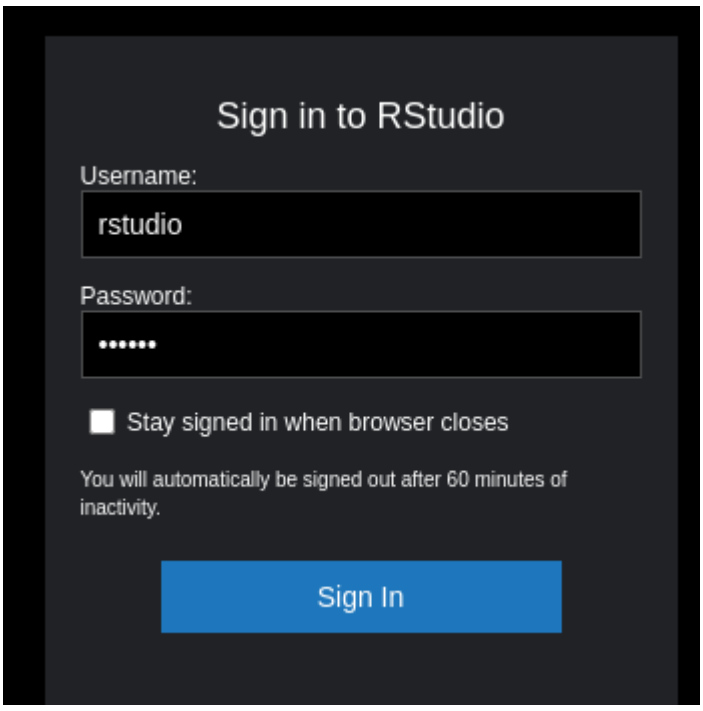
2. Enter the basic information or leave as-is to skip.
3. Enter `y` to confirm the information.

5. Access RStudio

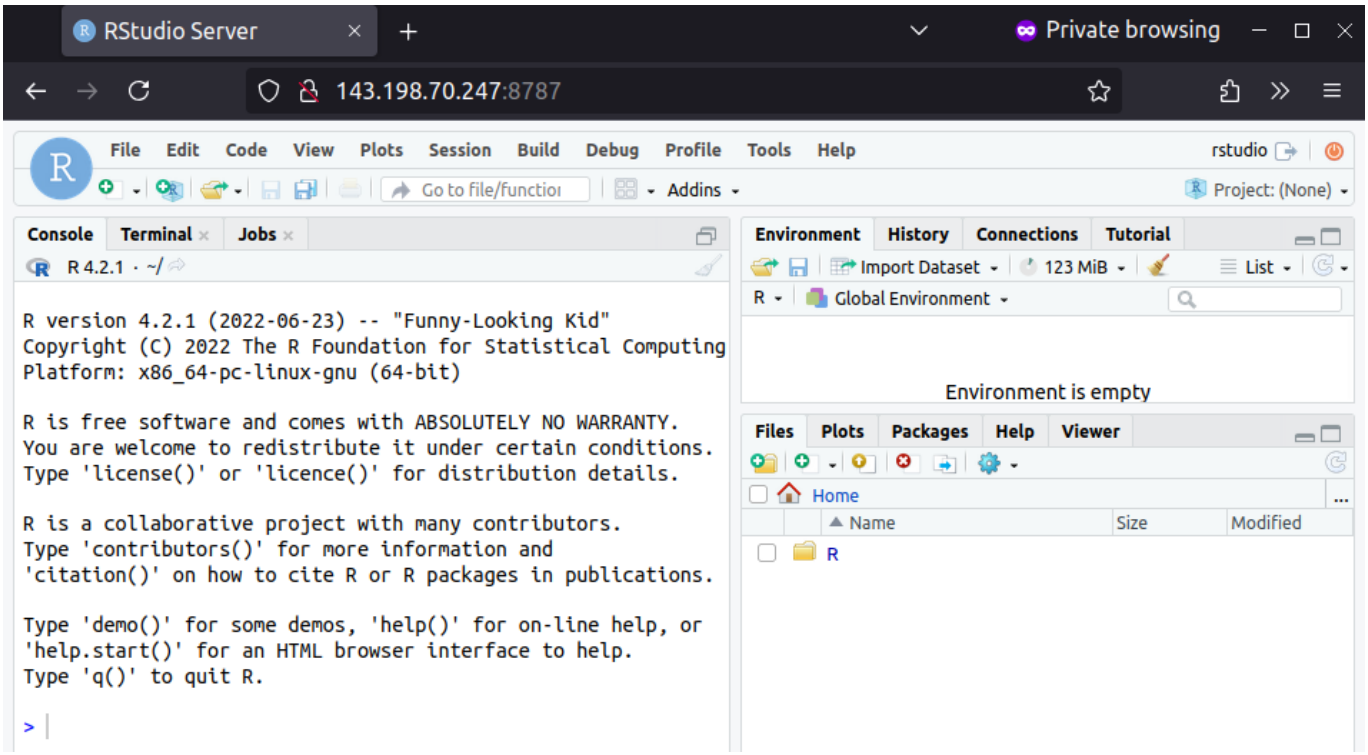
Open a new browser on your local computer. Enter your **ipv4 address** and `:8787` into the URL address bar. It should look like:

```
{{your.ipv4.address}}:8787
```

Enter your new user credentials into the RStudio sign in page.



Congrats! You now have access to RStudio.



Final Considerations

The following options and tools may be useful.

Elevate Privileges

You can elevate your new user's privileges with `sudo`. In the Droplet CLI, and as `root` user, enter the command:

```
usermod -aG sudo {{username}}
```

Now the new user can temporarily elevate its privileges by typing `sudo` in front of their commands (e.g., `sudo apt update`).

Transfer Files

Below are a few options to transfer files from your local computer to your Droplet. Your local computer's operating system will dictate the available options.

Windows

WinSCP

WinSCP is a separate software with a graphical user interface. It provides a "click-and-drag" method to transfer files.

1. Download WinSCP
2. Connect to your Droplet
3. Transfer files

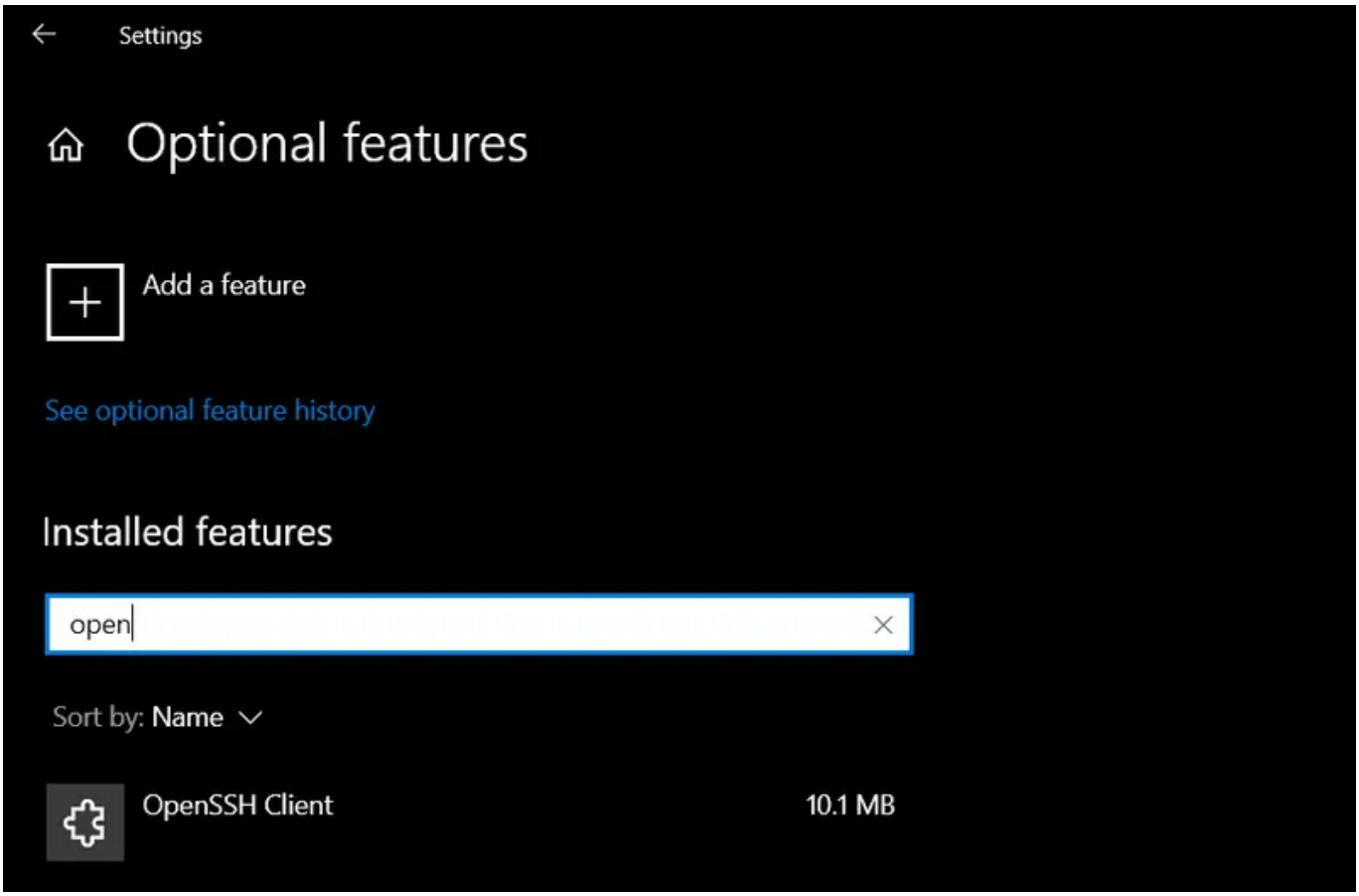
OpenSSH

OpenSSH is a CLI tool included in certain Windows versions (see the note below), but may be disabled by default.

> Note: Requires Windows 10 (build 1809 or later) and PowerShell (5.1 or later)

Enable OpenSSH with PowerShell or:

1. From the Start menu, search and open "optional features".
2. Click "Add a feature".
3. Search "OpenSSH Client" and click install.



Enter the following command in your local CLI to transfer files:

```
scp {{C:\path\to\local\file}} {{username}}@{{vm.ip.address}}:
{{/path/to/remote/directory}}
```

Click here for more `scp` options.

> Note: Windows uses backslashes `\` while Mac/Linux uses forward slashes `/` for directories.

Mac/Linux

rsync

`rsync` is a CLI tool included in most Mac/Linux versions.

Enter the following command in your local CLI to transfer files:

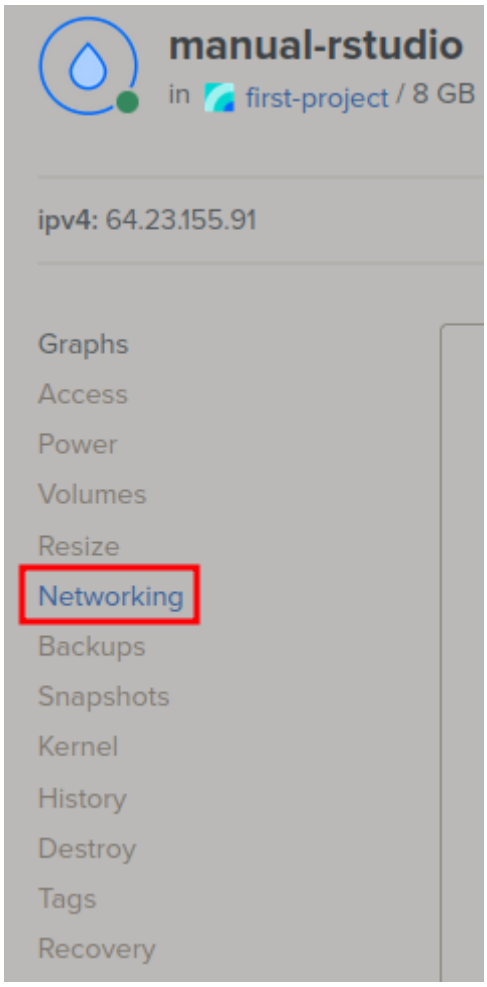
```
rsync -avzP -e ssh {{/path/to/local/file}} {{username}}@{{vm.ip.address}}:
{{/path/to/remote/directory/}}
```

Click here for more `rsync` options.

Firewall

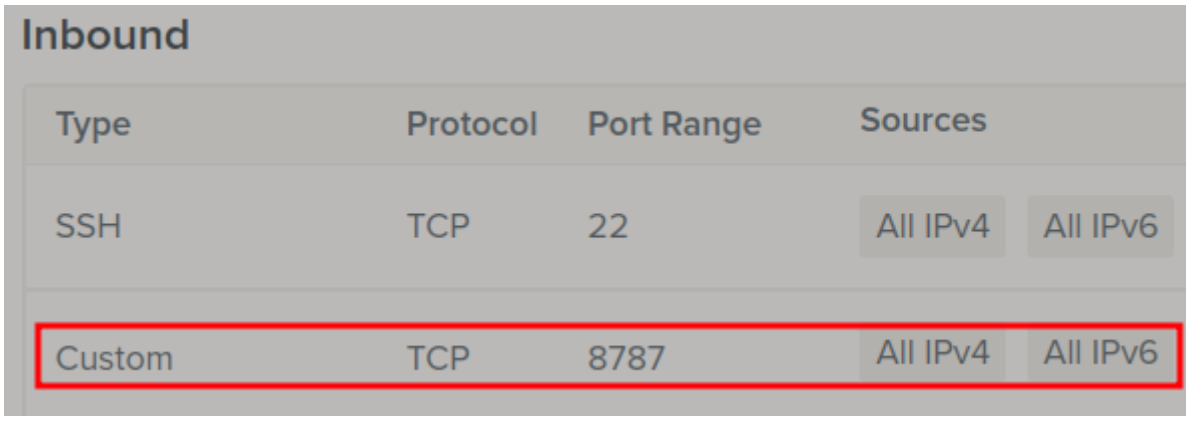
A firewall can help secure your VM against unauthorized access.

Navigate to your Droplet:



1. In the left-panel menu, click `Networking`.
2. Scroll down to Firewalls, click `Edit`.
3. Click `Create Firewall`.

- **Name:** Create a name for the firewall.
- **Inbound Rules:** Select the `New rule` drop-down and click `Custom`. Change the Port Range to `8787` and then save it.



- **Outbound Rules:** Leave the defaults.
- **Apply to Droplets:** Select your Droplet and then click `Create Firewall`.