rorocloud Documentation

Release 1.1.3

rorodata

Jul 06, 2017

Contents

rorocloud is a serverless platform to run data science experiements and notebooks in the pre-configured compute environments in the cloud.

The simple interface allows data scientists to start running their experiements in the cloud with in minutes.

The primary interface to interact with the rorocloud service is a command-line tool called `rorocloud`.

CHAPTER 1

Installation

The rorocloud client can be installed using `pip`.

```
$ pip install -U rorocloud
```

You can check the version of the rorocloud client you can using:

```
$ rorocloud version
rorocloud, version 0.1.3
```

CHAPTER 2

Getting Started

Login

Login to rorocloud service to get started.

```
$ rorocloud login
email: anand@rorodata.com
password: ********
Login successful.
```

After login, your credentials will be saved permanently on your local computer. You can use the *rorocloud whoami* command to find the email address of logged in user.

```
$ rorocloud whoami
anand@rorodata.com
```

You need to have a valid login to use this service. Please write to us at `anand@rorodata.com` if you don't already have one.

Running Hello world

Once you are logged in, you can run commands in the cloud. Let us try with a simple one.

```
$ rorocloud run echo hello world
created new job ff4a0620
```

The rorocloud client submitted a new job to run the command `echo hello world` and the job id is `ff4a0620`. Let us look at the logs of the job.

```
$ rorocloud logs ff4a0620
starting the job
executing command
```

```
hello world
job finished with exit status 0
```

We can also run a command in foregrond by passing `--foreground` option.

```
$ rorocloud run --foreground echo helloworld
created new job cd5c7c7c
starting the job
executing command
helloworld
job finished with exit status 0
```

Running Jupyter Notebook

Jupyter notebooks are natively supported in rorodata. To start a jupyter notebook, run:

```
$ rorocloud run:notebook
created new job 60984179
starting the job
executing command
Jupyter notebook is available at:
https://60984179-nb.rorocloud.io/?token=rorocloud

The jupyter notebook server can be stopped using:
    rorocloud stop 60984179
```

That would start a jupyter notebook and the URL to access the notebook will be printed. The notebook server is protected using a token.

The jupyter notebook server will continue to run even after closing the browser window and it must be stoped using `rorocloud stop` command.

The notebooks will be stored in `/data/notebooks` directory.

Copying files

The *put* command copies a local file into the cloud.

For example, to copy a file *hello.py* from current directory to */data*:

```
$ rorocloud put hello.py /data/hello.py
```

Status of Jobs

The status of currently running jobs be seen using:

```
$ rorocloud status
JOBID     STATUS    WHEN            TIME     CMD
--------  --------  --------------  -------  ------------------------------
60984179  running   14 minutes ago  0:14:18  /opt/rorodata/jupyter-notebook
74ee24a1  running   24 minutes ago  0:24:47  python train.py
```

CHAPTER 3

Tips & Tricks

Downloading a dataset

To download a dataset or any other file to rorocloud, simply run:

```
$ rorocloud run wget http://example.com/your/file
```

Cloning a git repo

To clone a git repo:

```
$ rorocloud run git clone https://github.com/rorodata/rorocloud-examples.git
```

CHAPTER 4

Support

Please join our slack channel to discuss about rorocloud.

Found any issues? Please report on github.