



ZEWAIL CITY OF SCIENCE AND TECHNOLOGY

CIE 547

CLOUD COMPUTING TECHNOLOGY

Twitter Influencers

Fatma Moanes NourEl-Din
Mai Al-Yamani

201700346
201701207

Contents

1	Requirements	1
2	Program Results	1
2.1	Influencers	1
3	Running the code on different Cloud Service Providers	2
3.1	CSPs specifications	2
3.1.1	AWS	2
3.1.2	Azure	3
3.2	CSPs Comparison	4

1 Requirements

The project aims to create a program that finds the top 10 influencers that have the highest number of followers on Twitter and run this program on at least 2 cloud service providers (CSP). In this project, we ran it on Amazon Web Service (AWS) and Microsoft Azure. The report flow goes from introducing the created program, showing the results of the program, and presenting a comparison between the results on each CSP.

2 Program Results

2.1 Influencers

Table 1. below shows the program results of the top 10 influencers with highest number of followers on Twitter.

Account ID	No. of followers
115485051	3383
40981798	3216
43003845	2735
813286	2647
22462180	2471
34428380	2462
15913	2133
7861312	2074
3359851	1905
11348282	1707

Table 1: Top 10 influencers on Twitter

3 Running the code on different Cloud Service Providers

3.1 CSPs specifications

3.1.1 AWS

```

ubuntu@ip-172-31-92-91: ~/Project
Setting up cpp (4:9.3.0-lubuntu2) ...
Setting up gcc-9 (9.4.0-lubuntu1~20.04.1) ...
Setting up libpython3-dev:amd64 (3.8.2-0ubuntu2) ...
Setting up libstdc++-9-dev:amd64 (9.4.0-lubuntu1~20.04.1) ...
Setting up gcc (4:9.3.0-lubuntu2) ...
Setting up g++-9 (9.4.0-lubuntu1~20.04.1) ...
Setting up python3.8-dev (3.8.10-0ubuntu1~20.04.4) ...
Setting up g++ (4:9.3.0-lubuntu2) ...
update-alternatives: using /usr/bin/g++ to provide /usr/bin/c++ (c++) in auto mo
de
Setting up build-essential (12.8ubuntu1.1) ...
Setting up python3-dev (3.8.2-0ubuntu2) ...
Processing triggers for man-db (2.9.1-1) ...
Processing triggers for libc-bin (2.31-0ubuntu9.7) ...
ubuntu@ip-172-31-92-91:~$ pip install pandas
Collecting pandas
  Downloading pandas-1.4.2-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_64.
  whl (11.7 MB)
    |████████████████████████████████████████| 11.7 MB 171 kB/s
Collecting pytz>=2020.1
  Downloading pytz-2022.1-py2.py3-none-any.whl (503 kB)
    |████████████████████████████████████████| 503 kB 32.5 MB/s
Collecting python-dateutil>=2.8.1
  Downloading python_dateutil-2.8.2-py2.py3-none-any.whl (247 kB)
    |████████████████████████████████████████| 247 kB 50.6 MB/s
Collecting numpy>=1.18.5; platform_machine != "aarch64" and platform_machine !=
"arm64" and python_version < "3.10"
  Downloading numpy-1.22.4-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_64.
  whl (16.9 MB)
    |████████████████████████████████████████| 16.9 MB 91 kB/s
Requirement already satisfied: six>=1.5 in /usr/lib/python3/dist-packages (from
python-dateutil>=2.8.1->pandas) (1.14.0)
Installing collected packages: pytz, python-dateutil, numpy, pandas
  WARNING: The scripts f2py, f2py3 and f2py3.8 are installed in '/home/ubuntu/.l
ocal/bin' which is not on PATH.
  Consider adding this directory to PATH or, if you prefer to suppress this warn
ing, use --no-warn-script-location.
Successfully installed numpy-1.22.4 pandas-1.4.2 python-dateutil-2.8.2 pytz-2022
.1
ubuntu@ip-172-31-92-91:~$ ls
Project
ubuntu@ip-172-31-92-91:~$ cd Project/
ubuntu@ip-172-31-92-91:~/Project$ ls
main.py  twitter.csv

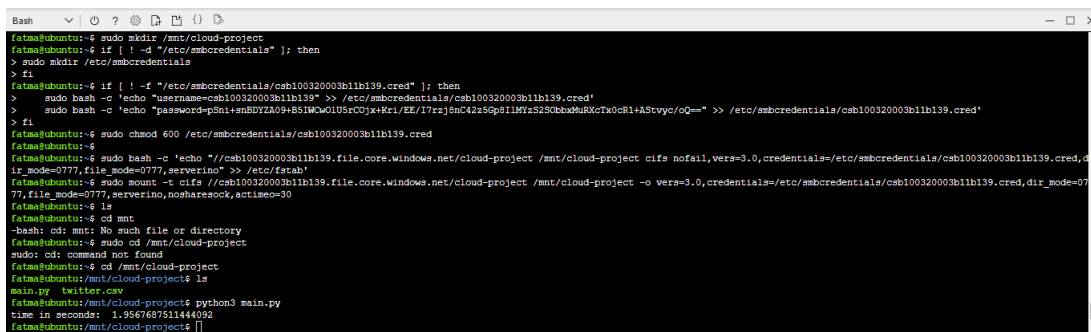
```

Figure 1: Using putty to run the program on the AWS VM

Amazon EC2 instance:

- Instance type: t2.micro
- Platform: Ubuntu Linux/UNIX
- Volume size (GiB): 8
- Number of vCPUs: 1

3.1.2 Azure



```
Bash
fatma@ubuntu:~$ sudo mkdir /mnt/cloud-project
fatma@ubuntu:~$ if [ ! -d "/etc/smbcredentials" ]; then
> sudo mkdir /etc/smbcredentials
> fi
fatma@ubuntu:~$ if [ ! -f "/etc/smbcredentials/csb100320003b1b139.cred" ]; then
> sudo bash -c 'echo "username=csb100320003b1b139" >> /etc/smbcredentials/csb100320003b1b139.cred'
> sudo bash -c 'echo "password=psni+snB072A09+851WCwG1U5rCOjx+Krl/EZ/17rzj6uc42z5Gp811M1z5250btaMuRxcRl+ASTvyc/oQ==" >> /etc/smbcredentials/csb100320003b1b139.cred'
> fi
fatma@ubuntu:~$ sudo chmod 600 /etc/smbcredentials/csb100320003b1b139.cred
fatma@ubuntu:~$
fatma@ubuntu:~$ sudo bash -c 'echo "//csb100320003b1b139.file.core.windows.net/cloud-project /mnt/cloud-project cifs nofail,vers=3.0,credentials=/etc/smbcredentials/csb100320003b1b139.cred,dir_mode=0777,file_mode=0777,serverino" >> /etc/fstab'
fatma@ubuntu:~$ sudo mount -t cifs //csb100320003b1b139.file.core.windows.net/cloud-project /mnt/cloud-project -o vers=3.0,credentials=/etc/smbcredentials/csb100320003b1b139.cred,dir_mode=0777,file_mode=0777,serverino,nosharesock,actimeo=30
fatma@ubuntu:~$ ls
fatma@ubuntu:~$ cd /mnt
-bash: cd: /mnt: No such file or directory
fatma@ubuntu:~$ sudo cd /mnt/cloud-project
sudo: cd: command not found
fatma@ubuntu:~$ cd /mnt/cloud-project
fatma@ubuntu:/mnt/cloud-project$ ls
main.py  twitter.csv
fatma@ubuntu:/mnt/cloud-project$ python3 main.py
time in seconds: 1.9567697511444092
fatma@ubuntu:/mnt/cloud-project$
```

Figure 2: Azure cloud shell

For some reason, while provisioning the first VM ever on Azure, choosing a 1vCPU was not an option. It displayed a message saying "Insufficient quota - family limit". Therefore, we choose a VM having 2 vCPUs, and we will compare the results accordingly.

Azure Virtual Machine:

- Instance type: Standard D2as v4 by Microsoft
- Platform: Ubuntu Server 20.04 LTS - Gen2
- Volume size (GiB): 8
- Number of vCPUs: 2

3.2 CSPs Comparison

Table 2 shows the comparison between AWS and Microsoft Azure regarding the execution time of the program, the starting time of the VM, and the price of the VM.

	AWS	Microsoft Azure
Execution time	1.0754 sec	1.9568 sec
Starting VM time	25 sec	1 min 22 sec
VM price	\$0.007 per hour	\$0.1070 per hour

Table 2: CSPs Comparison

Surprisingly, the results show that AWS VM (with 1 vCPU only), was approximately 1.8 times faster than Azure VM (2 vCPU). Besides, AWS interface was way more friendly and intuitive. The time needed to start the VM was much faster in AWS also.