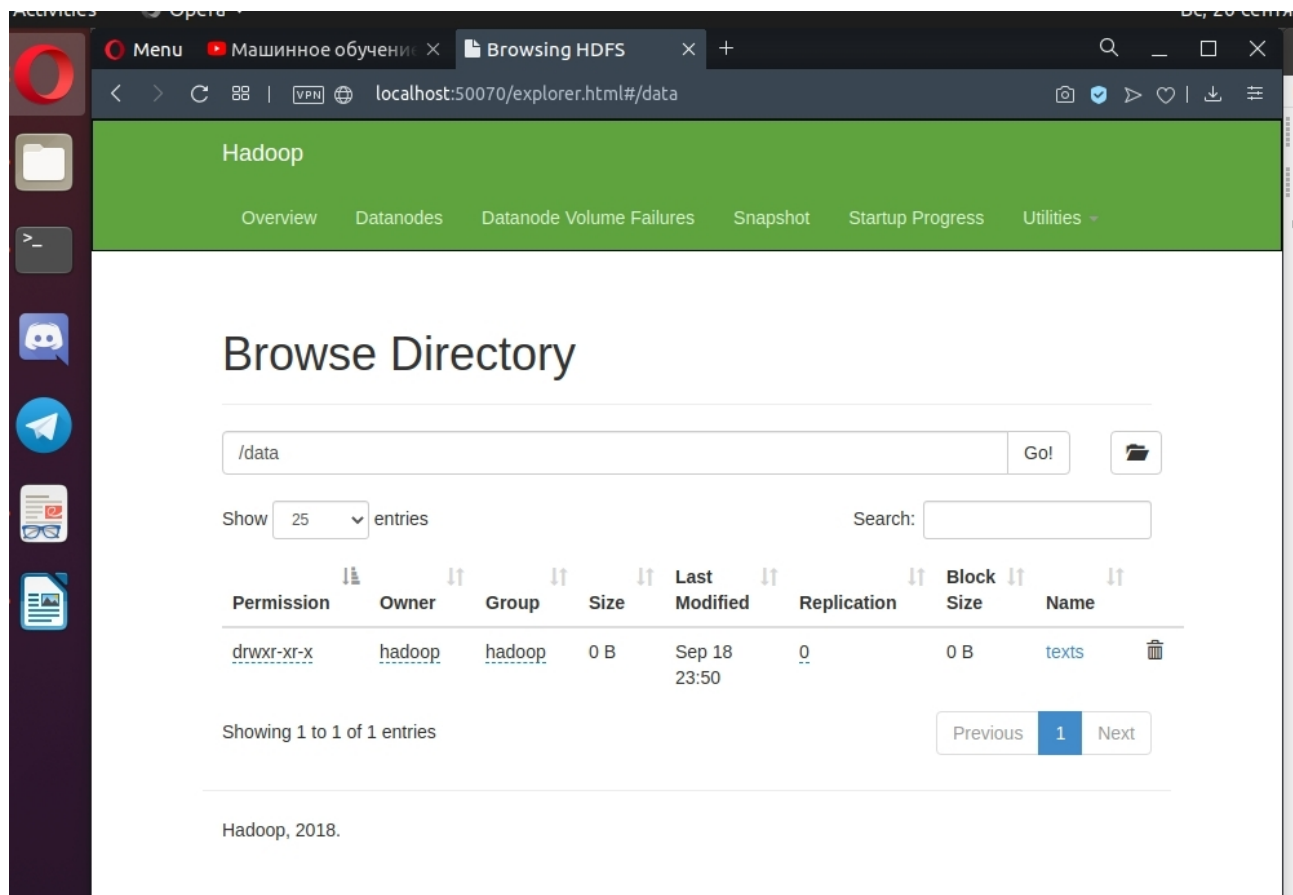


Beginner:

1. `ssh -i emr.pem -N -L 50070:ec2-3-249-21-2.eu-west-1.compute.amazonaws.com:50070 hadoop@ec2-3-249-21-2.eu-west-1.compute.amazonaws.com`

2. Количество подпапок: 1 (папка /text)



Intermediate -ls:

1. `hdfs dfs -ls /data/texts` или `hdfs dfs -ls -C /data/texts`

2. `hdfs dfs -ls -h /data/texts`

3. *actual*

Intermediate -du:

1. `hdfs dfs -du -s /data/texts`

Intermediate -mkdir u -touchz:

1. `hdfs dfs -mkdir odegov_ilya`

2. `hdfs dfs -mkdir odegov_ilya/sub_folder`

3. *Trash* — аналог корзины в *Windows*, эта папка хранит файлы удаленные из *hdfs*, что позволяет избежать случайного удаления файлов

`hdfs dfs -rm -skipTrash ./odegov_ilya/subfolder/henry.txt`

4. `hdfs dfs -touchz ./odegov_ilya/sub_folder/test.txt`

5. `hdfs dfs -rm ./odegov_ilya/sub_folder/test.txt`

6. `hdfs dfs -rm -r ./odegov_ilya`

Intermediate -put, -cat, -tail, -distcp:

1. `hdfs dfs -mkdir odegov_ilya`, затем

`hadoop distcp s3://texts-bucket/henry.txt ./odegov_ilya/henry.txt`

2. `hdfs dfs -cat ./odegov_ilya/henry.txt`

3. `hdfs dfs -tail ./odegov_ilya/henry.txt`

4. `hdfs dfs -cat ./odegov_ilya/henry.txt | head`

5. `hdfs dfs -cp ./odegov_ilya/henry.txt ./odegov_ilya/subfolder/henry.txt`

Advanced -setrep -w

1. `hdfs dfs -setrep -w 3 ./odegov_ilya/henry.txt`

Увеличение фактора репликации заняло 10...15 секунд.

Уменьшение фактора репликации заняло чуть большее время чем увеличение.

2. `hdfs fsck ./odegov_ilya/henry.txt -files -blocks -locations`

3. `hdfs fsck -blockId blk_1073743978`