Звіт про виконання

Інструкція по встановленню HDFS через Docker

Що вдалося зробити

Вдалося виконати усі кроки інструкції. Не виникало ніяких проблем. Приклади виконання кроків інструкції:

```
docker exec-it namenode /bin/bash

docker (com.docker.cli)

t1

docker (com.docker.cli)

**1

Last login: Sat Oct 22 01:26:07 on ttys001

**/Learning/BigData** git clone https://github.com/big-data-europe/docker-hadoop
Cloning into 'docker-hadoop'...
remote: Enumerating objects: 539, done.
remote: Counting objects: 100% (238/238), done.
remote: Compressing objects: 100% (33/33), done.
remote: Total 539 (delta 214), reused 205 (delta 205), pack-reused 301
Receiving objects: 100% (539/539), 108.02 KiB | 1.69 MiB/s, done.
Resolving deltas: 100% (251/251), done.

**/Learning/BigData** cd docker-hadoop

**/Le/B/docker-hadoop on master** sudo docker-compose up -d

**/Le/B/docker-hadoop on master** sud
```

Зі створенням контейнерів не виникало проблем. Все пройшло доволі швидко, адже docker-compose файл був правильно налаштований, на щастя.

```
~/Le/B/docker-hadoop > on master
                                       docker container ls
CONTAINER ID
                                                                               COMMAND
                                                                                                          CREATED
STATUS
                           PORTS
                                                                                 NAMES
6ff72626ee27
                bde2020/hadoop-resourcemanager:2.0.0-hadoop3.2.1-java8
                                                                               "/entrypoint.sh /run..."
                                                                                                          5 minutes ago
6ff/Z6Z6eez/ Bdcz250, 8088/tcp
Up 5 minutes (healthy) 8088/tcp
1f6e9a87ea77 bdc2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8
Up 5 minutes (healthy) 8042/tcp
                                                                                 resourcemanager
                                                                               "/entrypoint.sh /run..."
                                                                                                          5 minutes ago
                                                                                 nodemanager
                                                                               "/entrypoint.sh /run..."
                bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8
                                                                                                          5 minutes ago
                          0.0.0.0:9000->9000/tcp, 0.0.0.0:9870->9870/tcp
Up 5 minutes (healthy)
                                                                                 namenode
                                                                               "/entrypoint.sh /run..."
bb0c89a8e85c
                bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8
                                                                                                          5 minutes ago
Up 5 minutes (healthy)
                         9864/tcp
                                                                                 datanode
514b7ad457c7
                bde2020/hadoop-historyserver:2.0.0-hadoop3.2.1-java8
                                                                               "/entrypoint.sh /run..."
                                                                                                          5 minutes ago
Up 5 minutes (healthy) 8188/tcp
                                                                                 historyserver
d5040dd1a919
                zgp/parky_pandemonium:0.1
                                                                               "/bin/bash celery_en..."
                                                                                                          30 hours ago
Up 30 hours
                                                                                 parky_pandemonium_celery_1
de525452ac90
                                                                               "/bin/bash web_entry...
                zgp/parky_pandemonium:0.1
                                                                                                          30 hours ago
Up 30 hours
                            127.0.0.1:7000->8000/tcp
                                                                                 parky_pandemonium_web_1
d5292ec3211f
                                                                               "docker-entrypoint.s.."
                                                                                                          30 hours ago
                           33060/tcp, 127.0.0.1:8306->3306/tcp
Up 30 hours
                                                                                 parky_pandemonium_db_1
.
cc87f67d385d
                docker.elastic.co/elasticsearch/elasticsearch:6.5.4
                                                                               "/usr/local/bin/dock...
                                                                                                          30 hours ago
```

```
~/Le/B/docker-hadoop
                                                docker exec -it namenode /bin/bash
root@d6af7e9fe0c4:/# ls
KEYS dev
                            hadoop
                                                       mnt
                                                                          sbin
                                                                                  tmp
bin entrypoint.sh hadoop-data lib64 opt run
boot etc home media proc run.
                                                                         srv
                                             media proc run.sh sys
                                                                                 var
root@d6af7e9fe0c4:/# cd /tmp
root@d6af7e9fe0c4:/tmp# mkdir input
root@d6af7e9fe0c4:/tmp# echo "Hello World" >input/f1.txt
root@d6af7e9fe0c4:/tmp# echo "Hello Docker" >input/f2.txt
root@d6af7e9fe0c4:/tmp# hadoop fs -mkdir -p input
root@d6af7e9fe0c4:/tmp# hdfs dfs -put ./input/* input
2022-10-22 23:38:11,720 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false,
remoteHostTrusted = false
2022-10-22 23:38:11,855 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false,
remoteHostTrusted = false
root@d6af7e9fe0c4:/tmp# ls
hadoop-mapreduce-examples-2.7.1-sources.jar input
                                                            jetty-0.0.0.0-9870-hdfs-_-any-1155132808123346148.dir
hadoop-root-namenode.pid
hsperfdata_root
root@d6af7e9fe0c4:/tmp# hadoop jar hadoop-mapreduce-examples-2.7.1-sources.jar org.apache.hadoop.examples.WordCou
nt input output
2022-10-22 23:45:58,073 INFO client.RMProxy: Connecting to ResourceManager at resourcemanager/192.168.16.4:8032 2022-10-22 23:45:58,353 INFO client.AHSProxy: Connecting to Application History server at historyserver/192.168.1
```

Зі створенням файлів введення у контейнерів та копіювання jar файлу з логікою пройшло успішно, без будь яких помилок чи зауважень.

```
Peak Reduce Virtual memory (bytes)=8442822656
          Shuffle Errors
                   BAD ID=0
                   CONNECTION=0
                   IO_ERROR=0
                   WRONG_LENGTH=0
                   WRONG_MAP=0
                   WRONG_REDUCE=0
         File Input Format Counters
                   Bytes Read=25
         File Output Format Counters
Bytes Written=25
root@d6af7e9fe0c4:/tmp# hdfs dfs -cat output/part-r-00000
2022-10-22 23:46:52,192 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false,
remoteHostTrusted = false
Docker
Hello
World
```

Дані були успішно отримані. На мою думку, дві інструкції можуть бути об'єднані, адже в інструкції з другого посилання не вистачало пояснення команд та параметрів.