

Звіт до роботи №4

Робота з HDFS.

Запустіть Hadoop. Виконайте наступні команди (В якості відповіді надайте Screen Shots):

```
~/Le/B/docker-hadoop > on master docker ps
```

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS
6ff72626ee27	bde2020/hadoop-resource-manager:2.0.0-hadoop3.2.1-java8	"/entrypoint.sh /run..."	5 weeks ago	Up 31 seconds (healthy)	8088/tcp
1f6e9a87ea77	bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8	"/entrypoint.sh /run..."	5 weeks ago	Up About a minute (healthy)	8042/tcp
d6af7e9fe0c4	bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8	"/entrypoint.sh /run..."	5 weeks ago	Up About a minute (healthy)	0.0.0.0:9000->9000/tcp, 0.0.0.0:9870->9870/tcp
bb0c89a8e85c	bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8	"/entrypoint.sh /run..."	5 weeks ago	Up About a minute (healthy)	9864/tcp
514b7ad457c7	bde2020/hadoop-historyserver:2.0.0-hadoop3.2.1-java8	"/entrypoint.sh /run..."	5 weeks ago	Up About a minute (healthy)	8188/tcp

```
historyserver
```

listing directories and files

hdfs dfs -ls /

```
~/Learning/BigData/docker-hadoop > on master docker exec -it namenode /bin/bash
root@d6af7e9fe0c4:/# hdfs dfs -ls /
Found 4 items
drwxrwxrwt - root root 0 2022-10-22 23:46 /app-logs
drwxr-xr-x - root supergroup 0 2022-10-22 23:24 /rmstate
drwx----- - root supergroup 0 2022-10-22 23:45 /tmp
drwxr-xr-x - root supergroup 0 2022-10-22 23:37 /user
```

create a directory in hdfs

hdfs dfs -mkdir -p /MyDirectory

```
root@d6af7e9fe0c4:/# hdfs dfs -mkdir -p /MyDirectory
root@d6af7e9fe0c4:/#
```

upload a file at the root

hdfs dfs -copyFromLocal /DataJek/helloWorld.txt /MyDirectory

Because docker running need copy file from local to the container first

```
UW PICO 5.09 docker (com.docker.cli) File: ./helloWorld.txt nano (nano) Modified
Test Hello World

~/Learning/BigData cat ./helloWorld.txt
Test Hello World

~/Learning/BigData docker cp ./helloWorld.txt d6af7e9fe0c4:/tmp
root@d6af7e9fe0c4:/# ls ./tmp/
hadoop-mapreduce-examples-2.7.1-sources.jar
hadoop-root-namenode.pid
helloWorld.txt
hsperfdata_root
root@d6af7e9fe0c4:/# hdfs dfs -copyFromLocal ./tmp/helloWorld.txt /MyDirectory
2022-11-29 14:09:17,775 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
```

verify helloWorld.txt is uploaded to root

hdfs dfs -ls /MyDirectory

```
2022-11-29 14:09:17,775 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
root@d6af7e9fe0c4:/# hdfs dfs -ls /MyDirectory
Found 1 items
-rw-r--r-- 3 root supergroup 17 2022-11-29 14:09 /MyDirectory/helloWorld.txt
```

check the size of the directory

hdfs dfs -du -s -h -v /MyDirectory

```
2022-11-29 14:13:31,746 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
root@d6af7e9fe0c4:/# hdfs dfs -du -s -h -v /MyDirectory
SIZE DISK_SPACE_CONSUMED_WITH_ALL_REPLICAS FULL_PATH_NAME
17 51 /MyDirectory
```

view the contents of the file

hdfs dfs -cat /MyDirectory/helloWorld.txt

```
2022-11-29 14:13:31,746 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
root@d6af7e9fe0c4:/# hdfs dfs -cat /MyDirectory/helloWorld.txt
Test Hello World
```

or

hdfs dfs -text /MyDirectory/helloWorld.txt

```
2022-11-29 14:13:54,861 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
root@d6af7e9fe0c4:/# hdfs dfs -text /MyDirectory/helloWorld.txt
Test Hello World
```

download the previously uploaded file

hdfs dfs -copyToLocal /MyDirectory/helloWorld.txt /Downloads/

First of all need to create such folder `Downloads`.

```
root@d6af7e9fe0c4:/# mkdir Downloads
root@d6af7e9fe0c4:/# ls
Downloads KEYS bin boot dev entrypoint.sh etc hadoop hadoop-data home lib lib64 media mnt opt proc root run run.sh/sbin srv sys tmp usr var
root@d6af7e9fe0c4:/# hdfs dfs -copyToLocal /MyDirectory/helloWorld.txt /Downloads/
2022-11-29 14:15:17,703 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
```

find operation

hdfs dfs -find / -iname "hello*"

```
root@d6af7e9fe0c4:/# hdfs dfs -find / -iname "hello*"
/MyDirectory/helloWorld.txt
```

Перевірте налаштування dfs.namenode.http-address за замовчанням порт 9870.

Перейдіть за адресою <http://localhost:9870/dfshealth.html>

Ви маєте побачити стан.

Overview 'namenode:9000' (active)

Started:	Tue Nov 29 15:49:18 +0200 2022
Version:	3.2.1, rb3cbbb467e22ea829b3808f4b7b01d07e0bf3842
Compiled:	Tue Sep 10 18:56:00 +0300 2019 by rohitsharmaks from branch-3.2.1
Cluster ID:	CID-71698876-7499-43be-b2a2-748b21e3f7fe
Block Pool ID:	BP-743849815-192.168.16.6-1666481079137

Summary

Security is off.

Safemode is off.

49 files and directories, 23 blocks (23 replicated blocks, 0 erasure coded block groups) = 72 total filesystem object(s).

Heap Memory used 53.36 MB of 257.5 MB Heap Memory. Max Heap Memory is 1.71 GB.

Non Heap Memory used 53.61 MB of 54.88 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	58.42 GB
Configured Remote Capacity:	0 B
DFS Used:	616.03 KB (0%)
Non DFS Used:	37.32 GB
DFS Remaining:	18.1 GB (30.98%)
Block Pool Used:	616.03 KB (0%)
DataNodes usages% (Min/Median/Max/stdDev):	0.00% / 0.00% / 0.00% / 0.00%

Виконайте наступні команди:

Retrieve the topology of the cluster

`hdfs dfsadmin -printTopology`

```
root@d6af7e9fe0c4:/# hdfs dfsadmin -printTopology
Rack: /default-rack
192.168.16.3:9866 (datanode.docker-hadoop_default)
```

Get an admin report

`hdfs dfsadmin -report`

```

root@d6af7e9fe0c4:/# hdfs dfsadmin -report
Configured Capacity: 62725623808 (58.42 GB)
Present Capacity: 19433353244 (18.10 GB)
DFS Remaining: 19432722432 (18.10 GB)
DFS Used: 630812 (616.03 KB)
DFS Used%: 0.00%
Replicated Blocks:
  Under replicated blocks: 23
  Blocks with corrupt replicas: 0
  Missing blocks: 0
  Missing blocks (with replication factor 1): 0
  Block groups with corrupt internal blocks: 0
  Missing block groups: 0
  Low redundancy blocks with highest priority to recover: 0
  Pending deletion blocks: 0

-----
Decommission Status : Normal
Configured Capacity: 62725623808 (58.42 GB)
DFS Used: 630812 (616.03 KB)
Non DFS Used: 40075554788 (37.32 GB)
DFS Remaining: 19432722432 (18.10 GB)
DFS Used%: 0.00%
DFS Remaining%: 30.98%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Last contact: Tue Nov 29 14:17:47 UTC 2022
Last Block Report: Tue Nov 29 13:49:23 UTC 2022
Num of Blocks: 23

```

Get the status of a Datanode

hdfs dfsadmin -getDatanodeInfo localhost:9867

```

root@d6af7e9fe0c4:/# hdfs dfsadmin -getDatanodeInfo localhost:9867
Uptime: 5212, Software version: 3.2.1, Config version: core-3.0.0,hdfs-1

```

Retrieve the FS image

hdfs dfsadmin -fetchImage /Downloads

```

root@d6af7e9fe0c4:/# hdfs dfsadmin -fetchImage /Downloads
2022-11-29 14:20:01,441 INFO namenode.TransferFsImage: Opening connection to http://namenode:9870/imagetransfer?getImage=1&txid=latest
2022-11-29 14:20:01,594 INFO common.Util: Combined time for file download and fsync to all disks took 0.00s. The file download took 0.00s at 3000.00 KB/s. Synchronous (fsync) write to disk
of /Downloads/fsimage_00000000000000000713 took 0.00s.

```

Attempt to start the Datanode

hdfs datanode

[illegible]

In operation

Show entries

Search:

Node	Http Address	Last contact	Last Block Report	Capacity	Blocks	Block pool used	Version
✓ bb0c89a8e85c:9866 (192.168.16.3:9866)	http://bb0c89a8e85c:9864	1s	46m	58.42 GB <div><div></div></div>	23	624 KB (0%)	3.2.1
✓ d6af7e9fe0c4:9866 (192.168.16.4:9866)	http://d6af7e9fe0c4:9864	0s	9m	58.42 GB <div><div></div></div>	23	624 KB (0%)	3.2.1

Showing 1 to 2 of 2 entries

Previous **1** Next

Перейдіть за адресою <http://localhost:9864/datanode.html>