

HDFS Task 1

Answer .

What is the Namenode's URI and which file is it configured in?

Ans: Name node is the master and management node of Hadoop system.in core-site.xml file

2. Where on a local file system will Namenode store its image and which file is it configured in?

Ans: FSImage is where Namenode stores its Image and it is stored in XML.config

3. Where on a local file system will Datanode store its blocks and which file is it configured in?

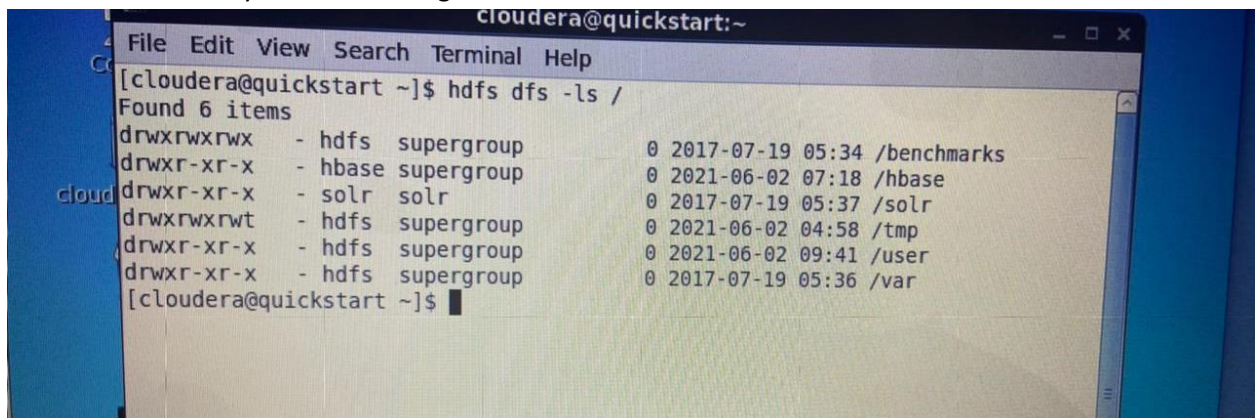
Ans:In Hadoop home etc folder the hdfs-site.xml file is there in that file data node property is configured

4. What is the block replication and which file is it configured in?

Ans: Its a copy of the data of node of rack, each data should be placed in different nodes of different rack.It is configured in hdfs-site.xml

Perform:

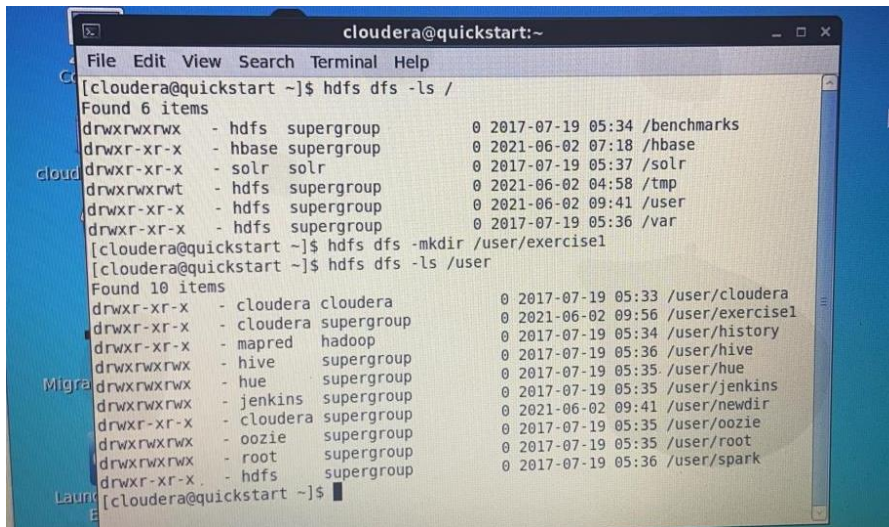
1. Start HDFS and verify that it's running ?



The screenshot shows a terminal window titled "cloudera@quickstart:~". The user has entered the command `hdfs dfs -ls /`. The output shows "Found 6 items" followed by a list of directories with their permissions, owner, group, size, modification time, and path.

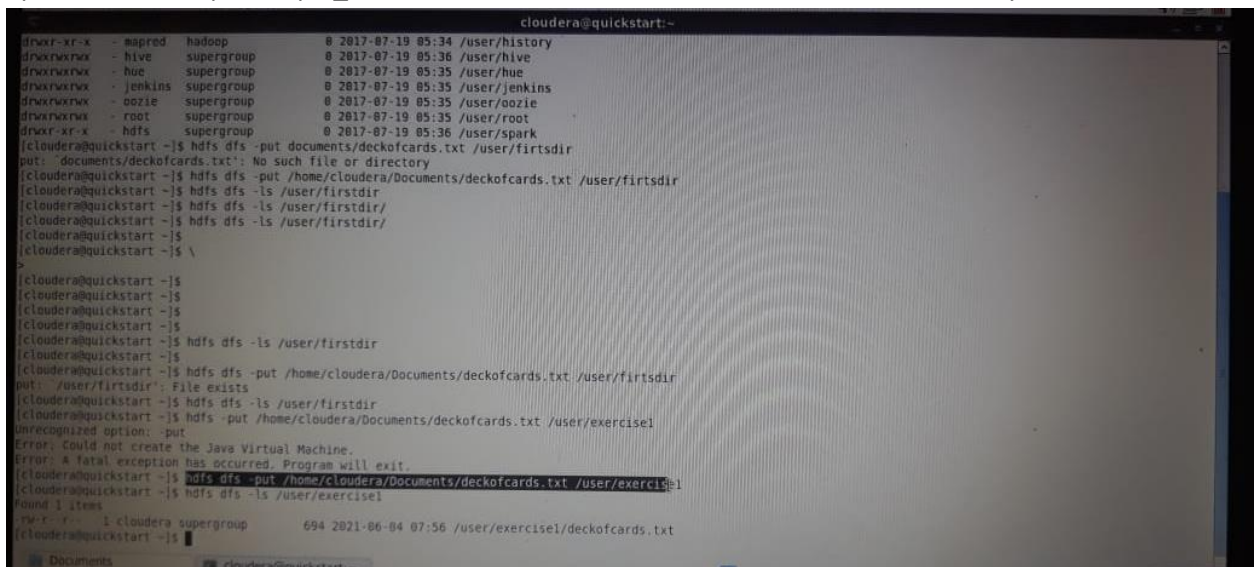
Permissions	Owner	Group	Size	Modification Time	Path
drwxrwxrwx	-	hdfs	super	group	0 2017-07-19 05:34 /benchmarks
drwxr-xr-x	-	hbase	super	group	0 2021-06-02 07:18 /hbase
drwxr-xr-x	-	solr	solr		0 2017-07-19 05:37 /solr
drwxrwxrwt	-	hdfs	super	group	0 2021-06-02 04:58 /tmp
drwxr-xr-x	-	hdfs	super	group	0 2021-06-02 09:41 /user
drwxr-xr-x	-	hdfs	super	group	0 2017-07-19 05:36 /var

2. Create a new directory /exercise1 on HDFS ?



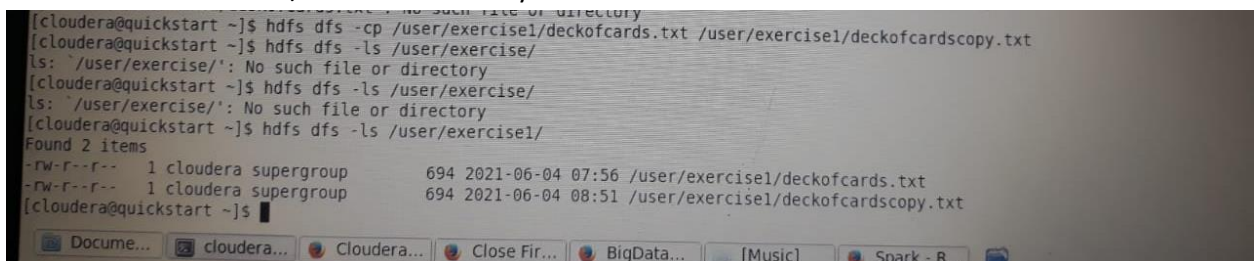
```
cloudera@quickstart:~$ hdfs dfs -ls /
Found 6 items
drwxrwxrwx - hdfs supergroup 0 2017-07-19 05:34 /benchmarks
drwxr-xr-x - hbase supergroup 0 2021-06-02 07:18 /hbase
drwxr-xr-x - solr solr 0 2017-07-19 05:37 /solr
drwxrwxrwt - hdfs supergroup 0 2021-06-02 04:58 /tmp
drwxr-xr-x - hdfs supergroup 0 2021-06-02 09:41 /user
drwxr-xr-x - hdfs supergroup 0 2017-07-19 05:36 /var
cloudera@quickstart:~$ hdfs dfs -mkdir /user/exercise1
cloudera@quickstart:~$ hdfs dfs -ls /user
Found 10 items
drwxr-xr-x - cloudera cloudera 0 2017-07-19 05:33 /user/cloudera
drwxr-xr-x - cloudera supergroup 0 2021-06-02 09:56 /user/exercise1
drwxr-xr-x - mapred hadoop 0 2017-07-19 05:34 /user/history
drwxrwxrwx - hive supergroup 0 2017-07-19 05:36 /user/hive
drwxrwxrwx - hue supergroup 0 2017-07-19 05:35 /user/hue
drwxrwxrwx - jenkins supergroup 0 2017-07-19 05:35 /user/jenkins
drwxr-xr-x - cloudera supergroup 0 2021-06-02 09:41 /user/newdir
drwxr-xr-x - oozie supergroup 0 2017-07-19 05:35 /user/oozie
drwxrwxrwx - root supergroup 0 2017-07-19 05:35 /user/root
drwxr-xr-x - hdfs supergroup 0 2017-07-19 05:36 /user/spark
cloudera@quickstart:~$
```

3. Upload GitHub repo sample_data/deckofcards.txt to HDFS under /exercise1 directory



```
cloudera@quickstart:~$ hdfs dfs -put documents/deckofcards.txt /user/exercise1
put: 'documents/deckofcards.txt': No such file or directory
cloudera@quickstart:~$ hdfs dfs -put /home/cloudera/Documents/deckofcards.txt /user/exercise1
put: /home/cloudera/Documents/deckofcards.txt: File exists
cloudera@quickstart:~$ hdfs dfs -ls /user/exercise1
Found 1 items
-rw-r--r-- 1 cloudera supergroup 694 2021-06-04 07:56 /user/exercise1/deckofcards.txt
cloudera@quickstart:~$
```

4. View the content of the /exercise1 directory



```
cloudera@quickstart:~$ hdfs dfs -cp /user/exercise1/deckofcards.txt /user/exercise1/deckofcardscopy.txt
cloudera@quickstart:~$ hdfs dfs -ls /user/exercise1/
ls: '/user/exercise1/': No such file or directory
cloudera@quickstart:~$ hdfs dfs -ls /user/exercise1/
ls: '/user/exercise1/': No such file or directory
cloudera@quickstart:~$ hdfs dfs -ls /user/exercise1/
Found 2 items
-rw-r--r-- 1 cloudera supergroup 694 2021-06-04 07:56 /user/exercise1/deckofcards.txt
-rw-r--r-- 1 cloudera supergroup 694 2021-06-04 08:51 /user/exercise1/deckofcardscopy.txt
cloudera@quickstart:~$
```

5. Determine the size of the hamlet.txt file in KB that resides on HDFS (not local directory)

```
cloudera@quickstart:~$ hdfs dfs -get /user/exercisel/deckofcards.txt ../Desktop/text/
get: ../Desktop/text/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /Desktop/text/
get: /Desktop/text/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /Desktop/text/
get: /Desktop/text/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /home/cloudera/Desktop
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /home/cloudera/Music
get: /home/cloudera/Music/deckofcards.txt: File exists
[cloudera@quickstart ~]$ hdfs dfs -cp /user/exercisel/deckofcards.txt /user/exercisel/deckofcardscopy.txt
cp: /user/exercisel/deckofcardscopy.txt: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -cp /user/exercisel/deckofcards.txt /user/exercisel/deckofcardscopy.txt
[cloudera@quickstart ~]$ hdfs dfs -ls /user/exercise/
ls: /user/exercise/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -ls /user/exercise/
ls: /user/exercise/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -ls /user/exercisel/
Found 2 items
-rw-r--r-- 1 cloudera supergroup 694 2021-06-04 07:56 /user/exercisel/deckofcards.txt
-rw-r--r-- 1 cloudera supergroup 694 2021-06-04 08:51 /user/exercisel/deckofcardscopy.txt
[cloudera@quickstart ~]$ hdfs dfs rm -rf /user/exercisel/deckofcards.txt
rm: Unknown command
Did you mean -rm? This command begins with a dash.
[cloudera@quickstart ~]$ hdfs dfs -rf /user/exercisel/deckofcards.txt
-rf: Unknown command
[cloudera@quickstart ~]$ hdfs dfs rm -rf /user/exercisel/deckofcards.txt
rm: Unknown command
Did you mean -rm? This command begins with a dash.
[cloudera@quickstart ~]$ hdfs dfs -rm -rf /user/exercisel/deckofcards.txt
-rm: Illegal option -rf
Usage: hadoop fs [generic options] -rm [-f] [-r|-R] [-skipTrash] <src> ...
[cloudera@quickstart ~]$ hdfs dfs -rm /user/exercisel/deckofcards.txt
Deleted /user/exercisel/deckofcards.txt
[cloudera@quickstart ~]$
```

6. Print the first 25 lines to the screen from deckofcards.txt on HDFS

```
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /home/cloudera/Documents
get: /home/cloudera/Documents/deckofcards.txt: File exists
[cloudera@quickstart ~]$ hdfs dfs -cat /user/exercisel/deckofcards.txt | head -n 25
BLACK|SPADE|2
BLACK|SPADE|3
BLACK|SPADE|4
BLACK|SPADE|5
BLACK|SPADE|6
BLACK|SPADE|7
BLACK|SPADE|8
BLACK|SPADE|9
BLACK|SPADE|10
BLACK|SPADE|J
BLACK|SPADE|Q
BLACK|SPADE|K
BLACK|SPADE|A
BLACK|CLUB|2
BLACK|CLUB|3
BLACK|CLUB|4
BLACK|CLUB|5
BLACK|CLUB|6
BLACK|CLUB|7
BLACK|CLUB|8
BLACK|CLUB|9
BLACK|CLUB|10
BLACK|CLUB|J
BLACK|CLUB|Q
BLACK|CLUB|K
[cloudera@quickstart ~]$
```

7. Copy deckofcards.txt to deckofcardsCopy.txt

```
cloudera@quickstart:~$ hdfs dfs -get /user/exercisel/deckofcards.txt ../Desktop/text/
get: ../Desktop/text/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /Desktop/text/
get: /Desktop/text/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /Desktop/text/
get: /Desktop/text/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /home/cloudera/Desktop
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /home/cloudera/Music
get: /home/cloudera/Music/deckofcards.txt: File exists
[cloudera@quickstart ~]$ hdfs dfs -cp /user/exercisel/deckofcards.txt /user/exercisel/deckofcardscopy.txt
cp: /user/exercisel/deckofcardscopy.txt: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -cp /user/exercisel/deckofcards.txt /user/exercisel/deckofcardscopy.txt
[cloudera@quickstart ~]$ hdfs dfs -ls /user/exercise/
ls: /user/exercise/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -ls /user/exercise/
ls: /user/exercise/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -ls /user/exercisel/
Found 2 items
-rw-r--r-- 1 cloudera supergroup 694 2021-06-04 07:56 /user/exercisel/deckofcards.txt
-rw-r--r-- 1 cloudera supergroup 694 2021-06-04 08:51 /user/exercisel/deckofcardscopy.txt
```


8. Copy deckofcards.txt back to local file system and name it deckofcards.copy.txt

```
BLACK|CLUB|
BLACK|CLUB|9
BLACK|CLUB|10
BLACK|CLUB|J
BLACK|CLUB|Q
BLACK|CLUB|K
[cloudera@quickstart ~]$ hdfs dfs -cat /user/exercisel/deckofcards.txt |head -n
head: option requires an argument -- 'n'
Try 'head --help' for more information.
cat: Unable to write to output stream.
[cloudera@quickstart ~]$ hdfs dfs -cat /user/exercisel/deckofcards.txt |wc -l
53
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /home/cloudera/Documents
get: /home/cloudera/Documents/deckofcards.txt: File exists
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /home/cloudera/Documents
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /home/cloudera/Music
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /home/cloudera/Music
get: /user/firstdir/deckofcards.txt: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt ../Desktop/
get: ../Desktop/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt ../Desktop/text/
get: ../Desktop/text/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /Desktop/text/
get: /Desktop/text/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /Desktop/text/
get: /Desktop/text/: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -get /user/exercisel/deckofcards.txt /home/cloudera/Desktop
[cloudera@quickstart ~]$
```

9. Check the entire filesystem for inconsistencies/problems

```
[cloudera@quickstart ~]$ hdfs fsck /user/exercisel/deckofcards.txt
Connecting to namenode via http://quickstart1.cloudera:50070/fuse?ugi=cloudera%2Fuser%2Fexercisel%2Fdeckofcards.txt
FSCK started by cloudera (auth:SIMPLE) from /10.0.2.15 for path /user/exercisel/deckofcards.txt at Thu Jun 03 09:24:50 PDT 2021
1
. Status: HEALTHY
Total size: 695 B
Total dirs: 0
Total files: 1
Total symlinks: 0
Total blocks (validated): 1 (avg. block size 695 B)
Minimally replicated blocks: 1 (100.0 %)
Over-replicated blocks: 0 (0.0 %)
Under-replicated blocks: 0 (0.0 %)
Mis-replicated blocks: 0 (0.0 %)
Default replication factor: 1
Average block replication: 1.0
Corrupt blocks: 0
Missing replicas: 0 (0.0 %)
Number of data-nodes: 1
Number of racks: 1
FSCK ended at Thu Jun 03 09:24:50 PDT 2021 in 10 milliseconds

The filesystem under path '/user/exercisel/deckofcards.txt' is HEALTHY
[cloudera@quickstart ~]$
```

10. Delete deckofcards.txt from HDFS

```
Did you mean -rm? This command begins with a dash.
[cloudera@quickstart ~]$ hdfs dfs -rm -rf /user/exercisel/deckofcards.txt
-rm: Illegal option -rf
Usage: hadoop fs [generic options] -rm [-f] [-r|-R] [-skipTrash] <src> ...
[cloudera@quickstart ~]$ hdfs dfs -rm /user/exercisel/deckofcards.txt
Deleted /user/exercisel/deckofcards.txt
[cloudera@quickstart ~]$
```

11. Delete the /exercisel directory from HDFS

```
-rw-r--r-- 1 cloudera supergroup 694 2021-06-04 08:51 /user/exercisel/deckofcards.copy.txt
[cloudera@quickstart ~]$ hdfs dfs rm -rf /user/exercisel/deckofcards.txt
rm: Unknown command
Did you mean -rm? This command begins with a dash.
[cloudera@quickstart ~]$ hdfs dfs -rf /user/exercisel/deckofcards.txt
-rf: Unknown command
[cloudera@quickstart ~]$ hdfs dfs rm -rf /user/exercisel/deckofcards.txt
rm: Unknown command
Did you mean -rm? This command begins with a dash.
[cloudera@quickstart ~]$ hdfs dfs -rm -rf /user/exercisel/deckofcards.txt
-rm: Illegal option -rf
Usage: hadoop fs [generic options] -rm [-f] [-r|-R] [-skipTrash] <src> ...
[cloudera@quickstart ~]$ hdfs dfs -rm /user/exercisel/deckofcards.txt
Deleted /user/exercisel/deckofcards.txt
[cloudera@quickstart ~]$
```

