

Hadoop Lab

CSE 328

Assignment-3

Gyanendra Kr. Shukla
CSE 1
191112040

Q1. Check whether your Hadoop installation is correct or not using start-all and stop-all commands.

I exported some environment variables to the `start-all.sh` and `stop-all.sh` files.

```
1 export HDFS_NAMENODE_USER="sok"
2 export HDFS_DATANODE_USER="sok"
3 export HDFS_SECONDARYNAMENODE_USER="sok"
4 export YARN_RESOURCEMANAGER_USER="sok"
5 export YARN_NODEMANAGER_USER="sok"
```

Then, I ran the `start-all.sh` and `stop-all.sh` files in the `sbin` directory of the Hadoop installation.

```
sok@languor: ~/hadoop-3.3.1 x sok@languor: ~
--(sok@languor)-[~/hadoop-3.3.1]
$ bin/hadoop
Usage: hadoop [OPTIONS] SUBCOMMAND [SUBCOMMAND OPTIONS]
or hadoop [OPTIONS] CLASSNAME [CLASSNAME OPTIONS]
where CLASSNAME is a user-provided Java class

OPTIONS is none or any of:

buildpaths          attempt to add class files from build tree
--config dir        Hadoop config directory
--debug             turn on shell script debug mode
--help             usage information
hostnames list[,of,host,names] hosts to use in slave mode
hosts filename      list of hosts to use in slave mode
loglevel level      set the log4j level for this command
workers            turn on worker mode

SUBCOMMAND is one of:

Admin Commands:

daemonlog          get/set the log level for each daemon

Client Commands:

archive            create a Hadoop archive
checknative        check native Hadoop and compression libraries availability
classpath          prints the class path needed to get the Hadoop jar and the required libraries
conftest          validate configuration XML files

--(sok@languor)-[~/hadoop-3.3.1/sbin]
$ sudo ./start-all.sh
Starting namenodes on [languor]
Starting datanodes
Starting secondary namenodes [languor]
Starting resourcemanager
Starting nodemanagers
--(sok@languor)-[~/hadoop-3.3.1/sbin]
--(sok@languor)-[~/hadoop-3.3.1/sbin]
$ sudo ./stop-all.sh
Stopping namenodes on [languor]
Stopping datanodes
Stopping secondary namenodes [languor]
Stopping nodemanagers
Stopping resourcemanager
--(sok@languor)-[~/hadoop-3.3.1/sbin]
$ |
```

Q2. Create a new directory in your Hadoop installation.

I created a new directory named `hadoopdata` in my home directory using the command

```
1 | bin/hadoop fs -mkdir /home/sok/hadoopdata
```

```
sok@languor: ~/hadoop-3.3.1 x sok@languor: ~
[sok@languor]~/hadoop-3.3.1
$ bin/hadoop fs -mkdir /home/sok/hadoopdata
[sok@languor]~/hadoop-3.3.1
$ ls ../
assignments Documents hadoop-3.3.1.tar.gz Music psqllinstall.sh tftesting Videos
cabal-userguide Downloads hadoopdata nohup.out Public thinclient_drives workshop
data.txt dpvis helix os reactugh toyoS tryredox
Desktop hadoop-3.3.1 Klox.jar Pictures Templates
[sok@languor]~/hadoop-3.3.1
$
```

Q3. Create file data.txt on Desktop and the copy it to the directory created. Display the content of HDFS folder.

I created a file name `data.txt` by simplying echoing the contents in a file. I then copied the file using `-copyFromLocal` command. Then, I displayed the contents of the folder using the command `fs -ls` command.

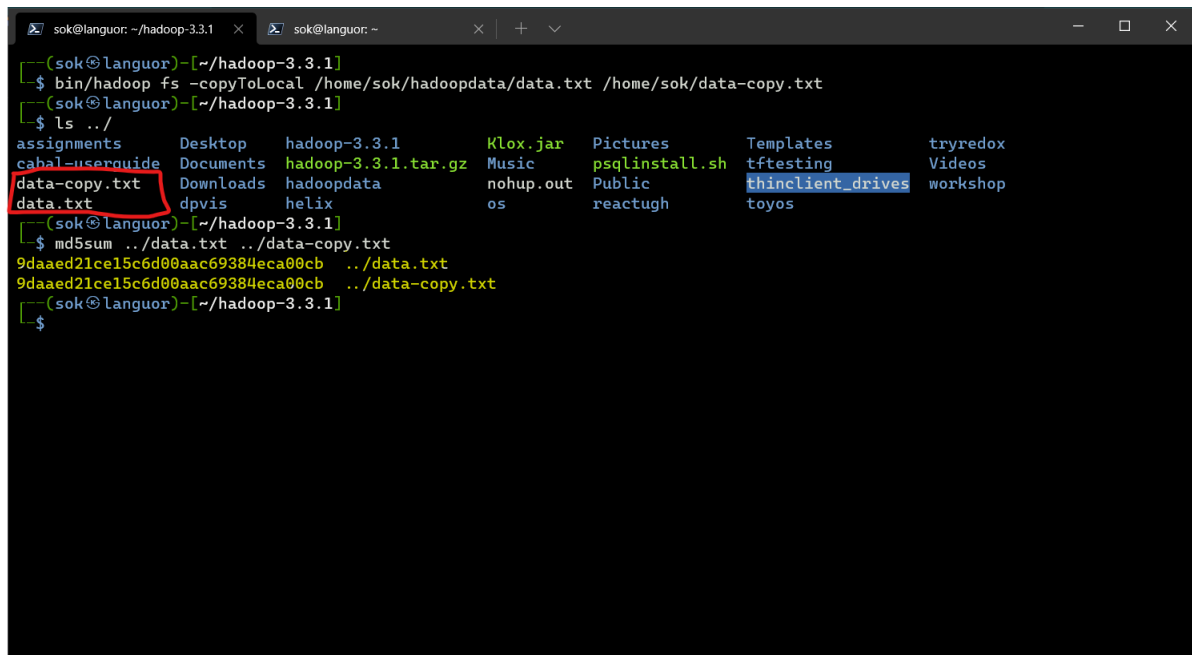
```
1 # creating a file
2 echo "this is some data in a data.txt file on my desktop" > ../data.txt
3
4 # copying the file to the hadoop directory
5 bin/hadoop fs -copyFromLocal home/sok/data.txt /home/sok/hadoopdata
6
7 # Displaying the contents of hdfs folder
8 bin/hadoop fs -ls /home/sok/hadoopdata
```

```
sok@languor: ~/hadoop-3.3.1 x sok@languor: ~
[sok@languor]~/hadoop-3.3.1
$ echo "this is some data in a data.txt file on my desktop" > ../data.txt
[sok@languor]~/hadoop-3.3.1
$ ls ../
assignments Documents hadoop-3.3.1.tar.gz Music psqllinstall.sh tftesting Videos
cabal-userguide Downloads hadoopdata nohup.out Public thinclient_drives workshop
data.txt dpvis helix os reactugh toyoS tryredox
Desktop hadoop-3.3.1 Klox.jar Pictures Templates
[sok@languor]~/hadoop-3.3.1
$ cat ../data.txt
this is some data in a data.txt file on my desktop
[sok@languor]~/hadoop-3.3.1
$ bin/hadoop fs -copyFromLocal /home/sok/data.txt /home/sok/hadoopdata
[sok@languor]~/hadoop-3.3.1
$ bin/hadoop fs -ls /home/sok/hadoopdata/
Found 1 items
-rw-r--r-- 1 sok sok 51 2022-02-04 09:49 /home/sok/hadoopdata/data.txt
[sok@languor]~/hadoop-3.3.1
$
```

Q4. Create copy of data.txt present in hadoop on Desktop with name "data-copy.txt" then compare them using md5 command.

I copied the `data.txt` file by using `-copyToLocal` command. Then, I used the command `md5sum` to check the md5sum of the file.

```
1 # copying the file from hadoop to the desktop
2 bin/hadoop fs -copyToLocal /home/sok/hadoopdata/data.txt /home/sok/data-
  copy.txt
3
4 # checking the md5sum of the file
5 md5sum ../data.txt ../data-copy.txt
```



```
sok@languor: ~/hadoop-3.3.1
[sok@languor]~/hadoop-3.3.1$ bin/hadoop fs -copyToLocal /home/sok/hadoopdata/data.txt /home/sok/data-copy.txt
[sok@languor]~/hadoop-3.3.1$ ls ../
assignments Desktop hadoop-3.3.1 Klox.jar Pictures Templates tryredox
cabal-userguide Documents hadoop-3.3.1.tar.gz Music psqlinstall.sh tftesting Videos
data-copy.txt Downloads hadoopdata nohup.out Public thinclient_drives workshop
data.txt dpvis helix os reactugh toyos
[sok@languor]~/hadoop-3.3.1$ md5sum ../data.txt ../data-copy.txt
9daaed21ce15c6d00aac69384eca00cb ../data.txt
9daaed21ce15c6d00aac69384eca00cb ../data-copy.txt
[sok@languor]~/hadoop-3.3.1$
```