Big Data Engineering with Hadoop & Spark

Assignment on Introduction to HDFS

Acadgild

Session 2: Assignment 2.1

This assignment is aimed at consolidating the concepts that was learned during the HDFS session of the course.

Task 1:

- Check whether /user/acadgild directory exists or not in the HDFS.
- If it doesn't exist, then create this.

```
[acadgild@localhost ~]$ hadoop fs -ls /user/
18/07/11 12:31:03 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes
Found 1 items
drwxr-xr-x   - acadgild supergroup          0 2018-07-07 10:54 /user/acadgild
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

- Create a directory /user/acadgild/hadoop.

```
[acadgild@localhost ~]$ hadoop fs -mkdir /user/acadgild/hadoop
18/07/11 12:38:54 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes
where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -ls /user/acadgild/
18/07/11 12:39:17 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes
where applicable
Found 3 items
drwxr-xr-x   - acadgild supergroup          0 2018-07-07 10:39 /user/acadgild/_sqoop
drwxr-xr-x   - acadgild supergroup          0 2018-07-11 12:38 /user/acadgild/hadoop
-rw-r--r--   1 acadgild supergroup         39 2018-07-07 10:54 /user/acadgild/test.txt
[acadgild@localhost ~]$
```

Task 2:

- Create a file in HDFS under directory /user/acadgild/hadoop, with name word-count.txt.

```
[acadgild@localhost ~]$ hadoop fs -touchz /user/acadgild/hadoop/word-count.txt
18/07/11 13:12:49 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes
where applicable
You have new mail in /var/spool/mail/acadgild
```

- Whatever we type on screen should get appended to the file.
- Try to type (on screen) few lines from any online article or textbook.

```
[acadgild@localhost ~]$ echo "This BigData Hadoop Session is very informative" | hdfs dfs -appendToFile - /user/acadgild/hadoop/word-count.txt
18/07/11 13:16:58 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes
where applicable
[acadgild@localhost ~]$ hadoop fs -cat /user/acadgild/hadoop/word-count.txt
18/07/11 13:17:36 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes
where applicable
This BigData Hadoop Session is very informative
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

Task 3:

- Create a file max-temp.txt in local FS.

```
[acadgild@localhost ~]$ nano max-temp.txt
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ cat max-temp.txt
10-01-1990,10
10-02-1991,12
10-03-1992,15
10-04-1993,16
10-05-1994,20
10-06-1995,23
10-07-1996,27
10-08-1997,24
10-09-1998,22
10-10-1999,21
10-11-2000,19
10-12-2001,13
10-01-2002,14
10-02-2003,18
10-03-2004,29
[acadgild@localhost ~]$
```

- Put some 10-15 records of date and temperature in "dd-mm-yyyy,temperature" format.

```
  GNU nano 2.0.9                          File: max-temp.txt                                        Modified

10-01-1990,10
10-02-1991,12
10-03-1992,15
10-04-1993,16
10-05-1994,20
10-06-1995,23
10-07-1996,27
10-08-1997,24
10-09-1998,22
10-10-1999,21
10-11-2000,19
10-12-2001,13
10-01-2002,14
10-02-2003,18
10-03-2004,29




^G Get Help      ^O WriteOut      ^R Read File      ^Y Prev Page      ^K Cut Text      ^C Cur Pos
^X Exit          ^J Justify       ^W Where Is       ^V Next Page      ^U UnCut Text    ^T To Spell
```

– Move this file to HDFS at /user/acadgild/hadoop.

```
[acadgild@localhost ~]$ hadoop fs -put max-temp.txt /user/acadgild/hadoop/max-temp.txt
18/07/11 13:29:30 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes
where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -ls /user/acadgild/hadoop/
18/07/11 13:30:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes
where applicable
Found 2 items
-rw-r--r--   1 acadgild supergroup        210 2018-07-11 13:29 /user/acadgild/hadoop/max-temp.txt
-rw-r--r--   1 acadgild supergroup         48 2018-07-11 13:17 /user/acadgild/hadoop/word-count.txt
[acadgild@localhost ~]$
```

Task 4:

– Change the permission of the file /user/acadgild/hadoop/max-temp.txt, such that only the owner and the group members have full control over the file.

– Others do not have any control over it.

```
[acadgild@localhost ~]$ hadoop fs -chmod -R 770 /user/acadgild/hadoop/max-temp.txt
18/07/11 13:40:59 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes
where applicable
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hadoop fs -ls /user/acadgild/hadoop/
18/07/11 13:41:20 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes
where applicable
Found 2 items
-rwxrwx---   1 acadgild supergroup        210 2018-07-11 13:29 /user/acadgild/hadoop/max-temp.txt
-rw-r--r--   1 acadgild supergroup         48 2018-07-11 13:17 /user/acadgild/hadoop/word-count.txt
[acadgild@localhost ~]$
```