BDAT 1002 – Data Systems Architecture

In-class Lab 1 – HDFS and Linux Commands Instructions:

Namood-e-Sehar

St ID: 200474523

Date: May 30th, 2023

Answer the following questions. For any questions that require you to execute a Linux or HDFS command, take a screenshot(s) of the command.

1. Create a folder in a directory in your Linux operating system and name it with your first name.

A screenshot of a computer program

Description automatically generated with medium confidence

1. Create a text file called yourname.txt in the folder you created in Step 1 and place some text in it using any text editor tool.

**Answer:**

A screen shot of a computer code

Description automatically generated with low confidence

1. Create a folder in HDFS under the root directory called BigDataFirstName where “FirstName” is your name.

**Answer:**

A screen shot of a computer

Description automatically generated with medium confidence

1. Copy the text file you created in step 2 from the Linux local file system into the HDFS folder BigData that you created in step 3.

**Answer:**



A picture containing text, screenshot, font

Description automatically generated

5. In the class exercises, we have most of the HDFS commands but not all of them. Research a new HDFS command and use it in an example. Provide a link for any external sources you use to answer this question (in other words, cite any sources you use).

**Answer:**

The command is text and its usage is Hadoop fs -text [filename]

It simply takes the file and simply output its content.

https://www.hdfstutorial.com/hdfs-commands/

A screenshot of a computer

Description automatically generated with medium confidence

6. The class exercises use “hadoop fs” as the start to HDFS commands but we can also use “hdfs dfs”. Research on the web what is the difference between these two commands. Provide a link for any external sources you use to answer this question (in other words, cite any sources you use).

**Answer**:

Hadoop fs is good when you are working with different file systems, your local file system can be one. And it is the broader approach.

But on the other hand hdfs dfs are useful when you are working with Hadoop distributed file system so in short Hadoop fs is flexible command but hdfs dfs is only limited to HDFS.

https://www.bigdatainrealworld.com/what-is-the-difference-between-hadoop-fs-hadoop-dfs-and-hdfs-dfs-commands/

7. We went over the most important Linux commands in class but there are many more. Try to find a Linux command that we did not cover in class but you find it useful. Describe the command and use it in an example and take a screenshot showing the usage.

**Answer**:

The command I find interesting is echo and it simply print the thing you type next to it.

A picture containing text, screenshot, font

Description automatically generated

https://www.digitalocean.com/community/tutorials/linux-commands