| **Name of Student:** Ajay Karthikesan | | | |
| --- | --- | --- | --- |
| **Roll Number:** 57 | | **Assignment Number:** 1 | |
| **Aim of Assignment:**  To study and execute 10-12 Hadoop commands | | | |
| **DOP:** 23.8.23 | | **DOS:** 4.9.23 | |
| **CO Mapped:**  - | **PO Mapped:**  - | **Faculty Signature:** | **Marks:** |

## 

## Practical No. 1

**Aim:** To study and execute 10-12 Hadoop commands

**Theory:**

HDFS is the primary or major component of the Hadoop ecosystem which is responsible for storing large data sets of structured or unstructured data across various nodes and thereby maintaining the metadata in the form of log files.

Following are the Hadoop commands,

1. hadoop fs: The File System (FS) shell includes various shell-like commands that directly interact with the Hadoop Distributed File System (HDFS) as well as other file systems that Hadoop supports, such as Local FS, WebHDFS, S3 FS, and others.

Syntax: hadoop fs < >

1. touchz: It creates an empty file. Syntax: bin/hdfs dfs -touchz <file\_path>
2. copyFromLocal (or) put: To copy files/folders from the local file system to hdfs store. This is the most important command. Local filesystem means the files present on the OS.

Syntax: bin/hdfs dfs -copyFromLocal <local file path> <dest(present on hdfs)>

1. test: This command is used to identify if the given destination is a directory or a file. It gives binary output. “demo ” is not a directory hence the output is 1, else the output is 0.

Syntax: hadoop fs -test -d <File OR Directory name>

1. mkdir : This command creates the directory in HDFS if it does not already exist. If the directory already exists in HDFS, then we will get an error message that file already exists.

Syntax: hadoop fs -mkdir <Directory name>

1. appendToFile: This command is used to append contents of multiple files and stores the result into a separate file.

Syntax: hadoop fs -appendToFile <fileName1 fileName2 fileName3(result)>

1. usage: The Hadoop fs shell command usage returns the help for an individual command.

Syntax: hadoop fs -usage <command>

1. Count: The Hadoop fs shell command count counts the number of files, directories, and bytes under the paths that matches the specified file pattern.

Options:

-q – shows quotas(quota is the hard limit on the number of names and amount of space used for individual directories)

-u – it limits output to show quotas and usage only

-h – shows sizes in a human-readable format

-v – shows header line

Syntax: hadoop fs -count [options] <path>

1. find : The Hadoop fs shell command finds all files that match the specified expression. If no path is specified, then it defaults to the present working directory. If an expression is not specified, then it defaults to -print.

Syntax: hadoop fs -find <path> … <expression>

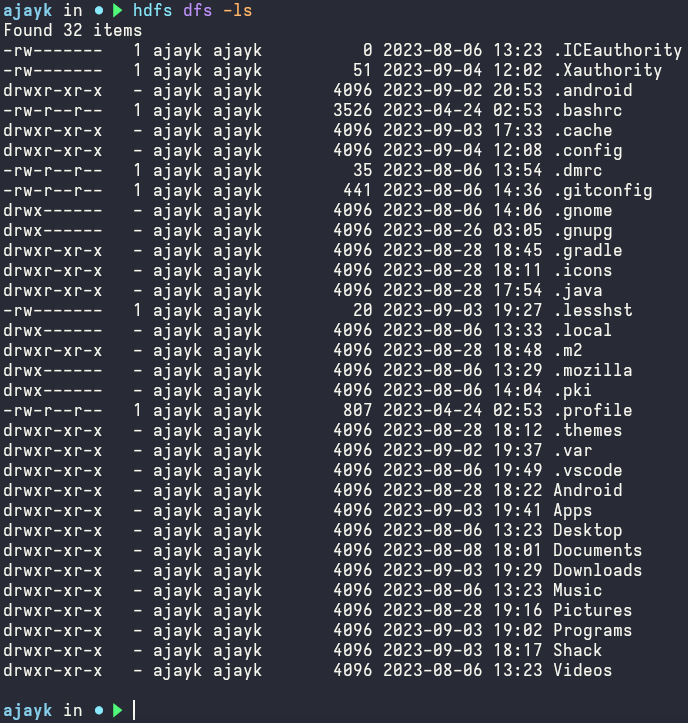
1. help : The Hadoop fs shell command help shows help for all the commands or the specified command.

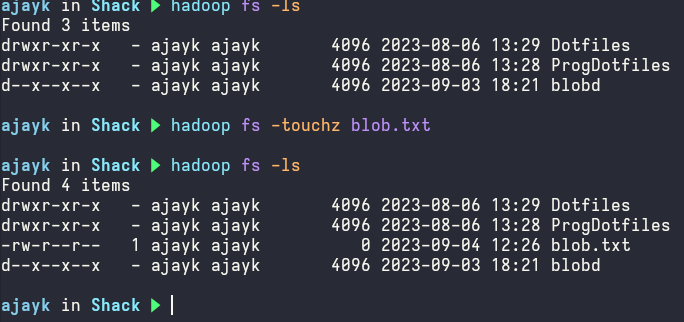
Syntax:hadoop fs -help [command]

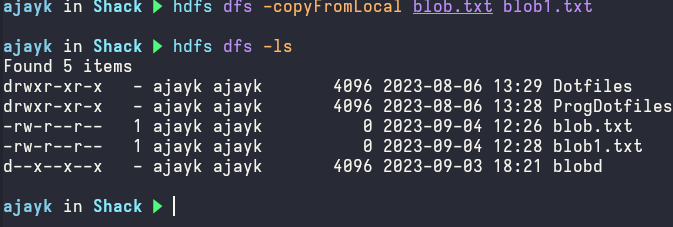
1. expunge : This command will delete all the files that are in trash without displaying any output.

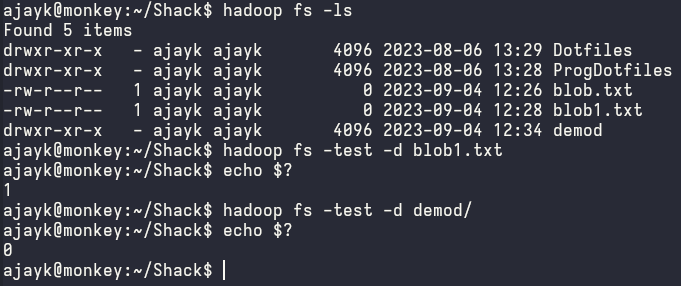
Syntax: hadoop fs -expunge.

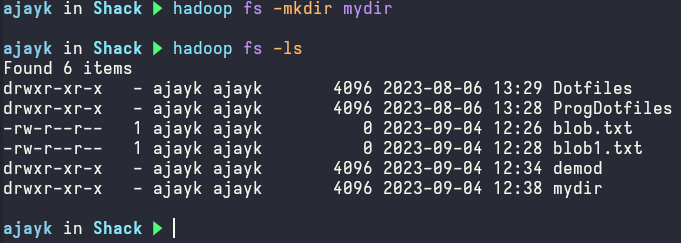
**Output:**

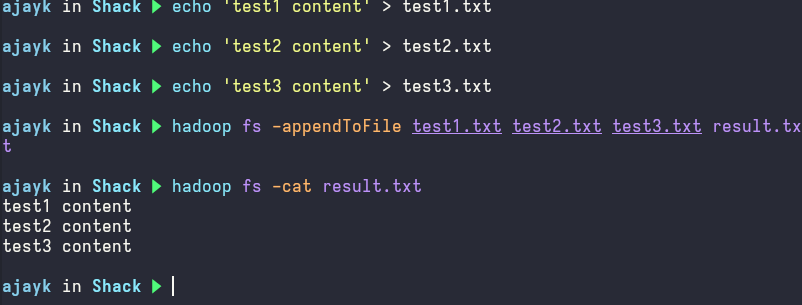


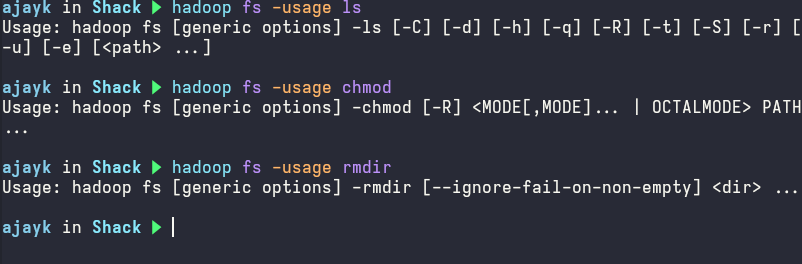


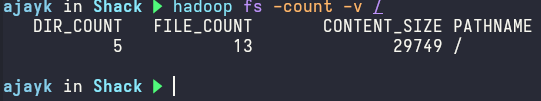




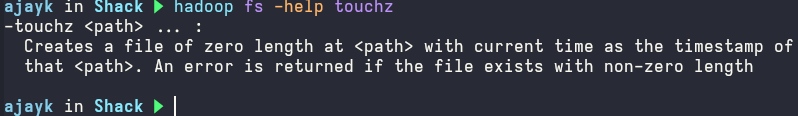


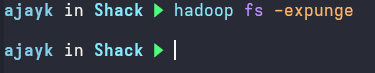












**Conclusion:**

I successfully studied and used the hadoop and hdfs commands.