**Hadoop Commands**

**Problem Statement Solution 01:**

* **Working with HDFS Commands**

1. # **Create the "dir01" folder in HDFS:**

hdfs dfs -mkdir /dir01

# **Move the input1.txt, input2.txt and input3.txt** **into the "dir01" folder:**

hdfs dfs -put input1.txt input2.txt input3.txt /dir01/

2. # **File names present in “/dir01”**

hdfs dfs -ls /dir01

3. # **Set replication factor to 5 for dir01**

hdfs dfs -setrep 5 /user/$USER/dir01

# **Get replication factor for dir01**

hdfs dfs -stat %r /dir01

4. # **Create directory scenario01**

hdfs dfs -mkdir /scenario01

# **Create level01 directory inside scenario01**

hdfs dfs -mkdir /scenario01/level01

# **Create level02 directory inside level01**

hdfs dfs -mkdir /scenario01/level01/level02

# **Copy input1.txt to scenario01**

hdfs dfs -put input1.txt /scenario01

# **Copy input2.txt to level01**

hdfs dfs -put input2.txt /scenario01/level01

# **Copy "input3.txt" to "level02"**

hdfs dfs -put input3.txt /scenario01/level01/level02

# **Recursively print only the file names present in scenario01**

hdfs dfs -ls -R /scenario01

**Problem Statement Solution 02:**

* **Working with YARN Commands Commands**

1. # **Capture the application id for the job**

yarn jar hadoop-mapreduce-jar.jar APPLICATION\_ID /user/dir01/input / /user/dir01/output/

1. # **Re-run map-reduce program**

yarn jar hadoop-mapreduce-jar.jar APPLICATION\_ID /user/dir01/input / /user/dir01/output/

# **Kill the application**

yarn application -kill <Application\_ID>

1. # **List of applications in running state**

yarn application -list -appStates RUNNING

1. # **View logs of jobs completed**

yarn logs -applicationId < Application\_ID >