Course: MMDS

*Tutorial 02*

**HDFS**

# **Basic Operation**

## **Starting HDFS**

1. Format the configured HDFS filesystem

bin/hdfs namenode -format

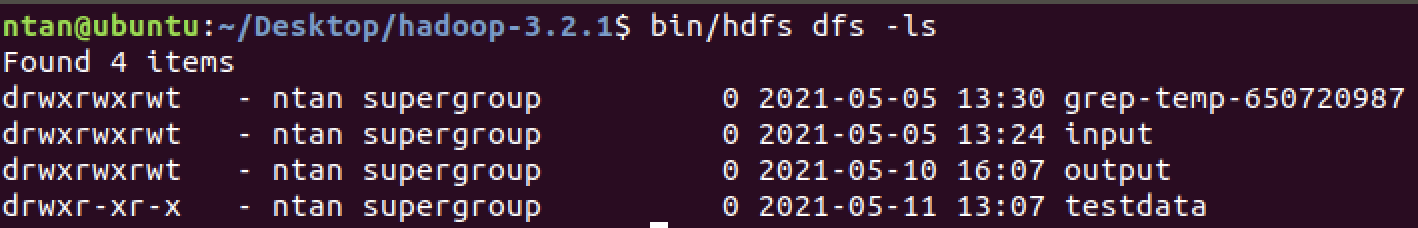
1. Start namenode, datanode, secondary namenode

sbin/start-dfs.sh

## **Listing Files in HDFS**

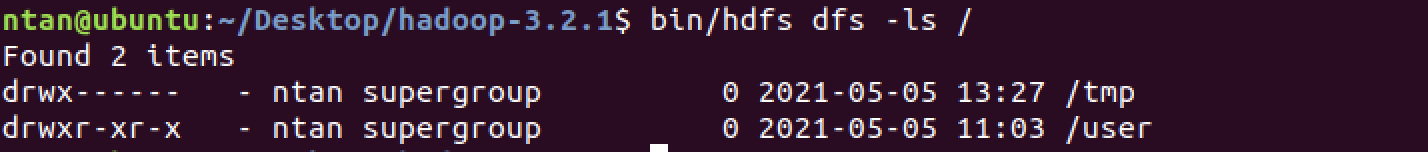
1. Find the list of files in a directory, status of file

bin/hdfs dfs -ls <path/directory>



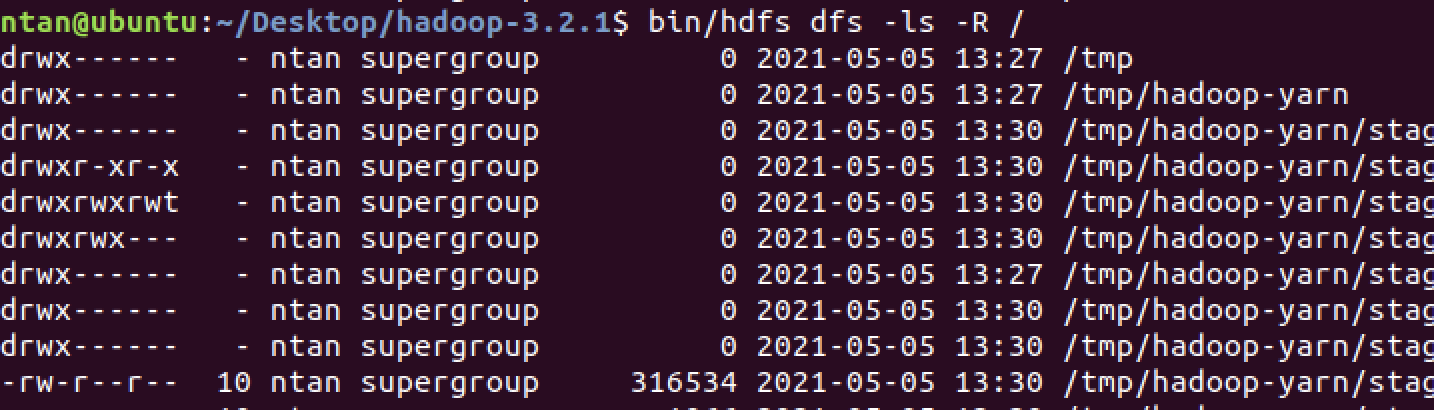
1. Note: root directory is /

bin/hdfs dfs -ls /



1. Recursively displays entries in all subdirectories of path

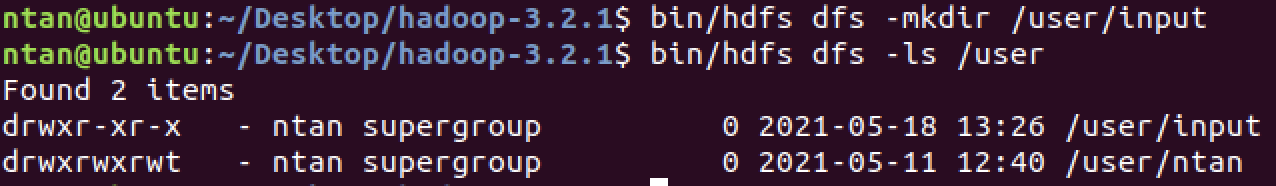
bin/hdfs dfs -ls -R <path/directory>



## **Inserting Data into HDFS**

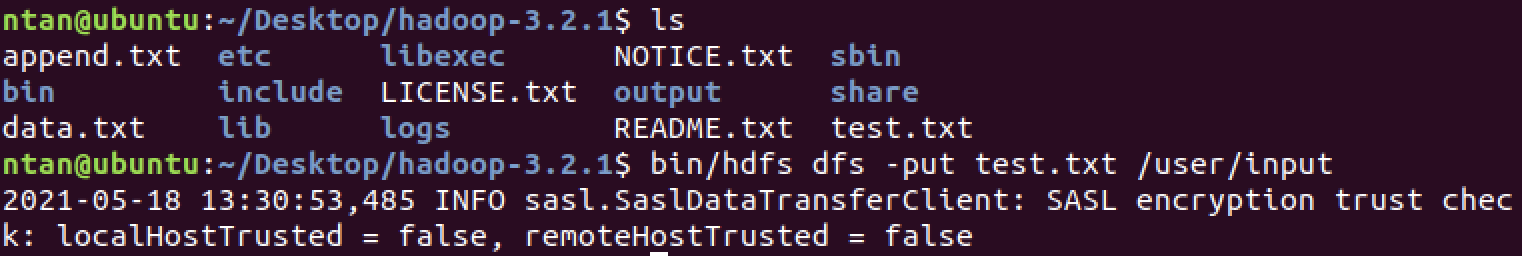
1. Create a directory

bin/hdfs dfs -mkdir <path>



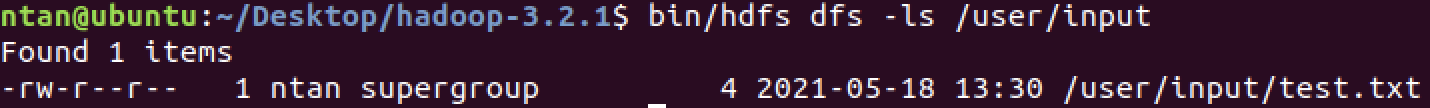
1. Transfer and store data file from local systems to HDFS

bin/hdfs dfs -put <local file/folder> <HDFS directory>



1. Verify the file

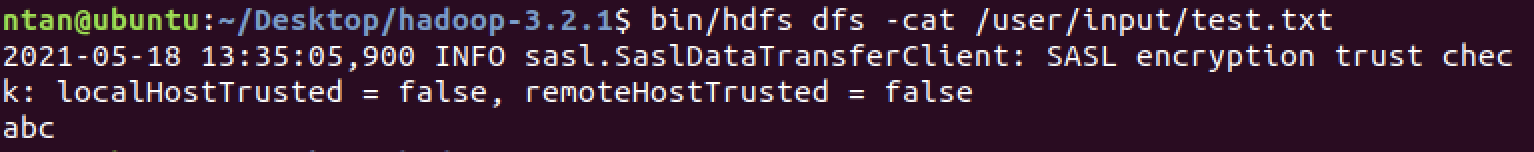
bin/hdfs dfs -ls /user/input



## **Retrieving Data from HDFS**

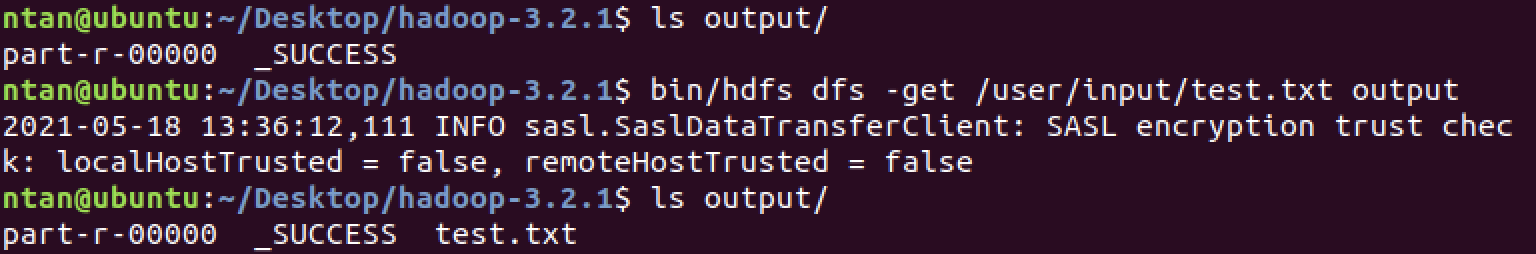
1. View the data from HDFS using cat

bin/hdfs dfs -cat /user/input/test.txt



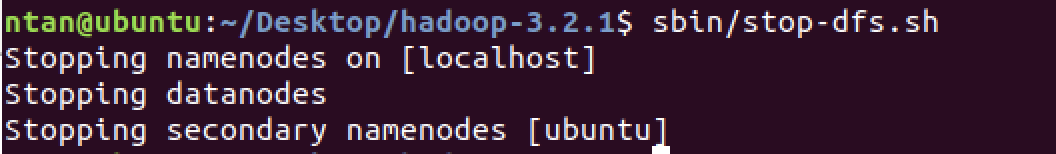
1. Get the file from HDFS

bin/hdfs dfs -get /user/input/test.txt output



1. Shut down the HDFS

sbin/stop-dfs.sh



# **HDFS Command Syntax**





# **Notice for examinations**

The commands below commonly appear in your examination questions.

* jps
* copyFrom/ToLocal
* get
* put
* mv
* mkdir

# **References**

* <https://www.tutorialspoint.com/hadoop/hadoop_hdfs_operations.htm>
* <https://images.linoxide.com/hadoop-hdfs-commands-cheatsheet.pdf>