How to run mapreduce program in terminal

WordCount Execution in Hadoop-2.6.5

Step1:Copy your WordCount.java file from local drive to hduser

sudo mv /home/bhaskar/Desktop/WordCount.java /home/hduser

Step2: Compile WordCount.java

Use the sun.tools.javac.Main: to pass it an array of Strings equivalent to the command line parameters. You might do this to bulk-compile large numbers of files without reloading the javac compiler for each file.

hadoop com.sun.tools.javac.Main WordCount.java

Step3: create a jar

Users can bundle their Map Reduce code in a jar file and execute it using this command.

jar cf wc.jar WordCount*.class

Step4: Create Input Directory

MapReduce expects the Input and Output paths to be the directories in HDFS and not local unless the Cluster is configured in Local mode. Also the Input directory must exist and the Output should not.

This directory must be created with all the input files in it. Use HDFS commands to create them.

hdfs dfs -mkdir -p input_dir

Step5:Move input file to Input directory

hdfs dfs -put /home/bhaskar/Desktop /test.txt input_dir

Step6:hadoop to list your file in input directory

hadoop fs -ls input_dir

Step7: Run the application

hadoop jar /home/hduser/wc.jar WordCount input_dir output

Step8: HDFS Name Node Web Interface

The name node web UI shows you a cluster summary including information about total/remaining capacity, live and dead nodes. Additionally, it allows you to browse the HDFS namespace and view the contents of its files in the web browser.

http://localhost:50070/.

Note:

- Summary of the HDFS file system is listed on the overview page

- Name Node status and storage information will also be displayed at the bottom.

- Data Nodes information will be displayed on Datanodes tab
- We can browse file system from Utilities menu on screen
- We can browse through each directory by clicking on each directory name as shown below for /user directory
- Browse Log files generated from various events like running map reduce jobs, running HDFS or YARN daemons. We can select these logs from Utilities menu on screen.