Instructions for Max-Temperature Analysis

CS555: Big Data Computing Lab

16th August 2021

1. Install Maven
    a. Open the terminal and type:

       sudo apt-get install maven

2. create project directory : (**Temperature**)
    a. Browse to home directory

       cd /home/iitp

    b. Create project directory

       mkdir Temperature

3. Create source files:
    a. Browse into Temperature directory

       cd Temperature

    b. Create new folder **src** inside Temperature directory

       mkdir src

    c. Browse into src directory

       cd src

    d. Create new folder **main** inside src directory

       mkdir main

    e. Browse into **main** directory

       cd main

    f. Create new folder **java** inside main directory

       mkdir java

    g. Browse into java directory

cd java

h. Create source files within java directory
   i. **nano TotalDriver.java**
   ii. **nano TotalMapper1.java**
   iii. **nano TotalMapper2.java**
   iv. **nano TotalReducer.java**

   paste the lines from the source code provided

   Note:- **To save file: Press- CTRL + o followed by Enter button**

   To Exit Press:- CTRL + x from nano editor

i. Browse into Temperature directory

   cd /home/iitp/Temperature

j. Create pom.xml file inside Temperature directory

   nano pom.xml

   paste the lines from the xml file provided and follow the steps in note of step (h)

k. Compile the java source files

   mvn clean && mvn compile && maven package

4. Create input directory (inputdata) for input files

   cd /home/iitp/Temperature

   mkdir inputdata

5. Copy the input file (input.txt) into inputdata folder

   cd /home/iitp/Temperature/inputdata

   nano file1.txt

   nano file2.txt

   paste the lines from the provided input file

6. Start all hadoop services
   a. Browse to hadoop installation sbin sub-directory

      cd /home/iitp/hadoop-2.6.0/sbin

   b. start all services

./start-all.sh

Note:- Enter password when prompted

7. Create input directory on HDFS
    a. browse to hadoop installation bin folder

        cd /home/iitp/hadoop-2.6.0/bin

    b. create directory (**Temperature**)

        ./hadoop fs -mkdir /Temperature

    c. create sub-directory (**inputdata**) under Temperature directory

        ./hadoop fs -mkdir /Temperature/inputdata

8. Copy the input text file from local directory to HDFS
    a. browse to hadoop installation bin folder

        cd /home/iitp/hadoop-2.6.0/bin

    b. Copy both input files (file1.txt and file2.txt) from Local

        ./hadoop dfs -put /home/iitp/Temperature/inputdata/file1.txt /Temperature/inputdata/

        and

        ./hadoop dfs -put /home/iitp/Temperature/inputdata/file2.txt /Temperature/inputdata/

9. Running the program
    a. browse to the bin directory of hadoop installation

        cd /home/iitp/hadoop-2.6.0/bin

    b. Running in terminal

        ./hadoop jar /home/iitp/Temperature/target/MaxTemp-1.0.0.jar TotalDriver 1 /Temperature/inputdata/file1.txt /Temperature/inputdata/file2.txt /Temperature/outputdata/

    c. Finding outputs

        ./hadoop dfs -cat /Temperature/outputdata/part-r-00000