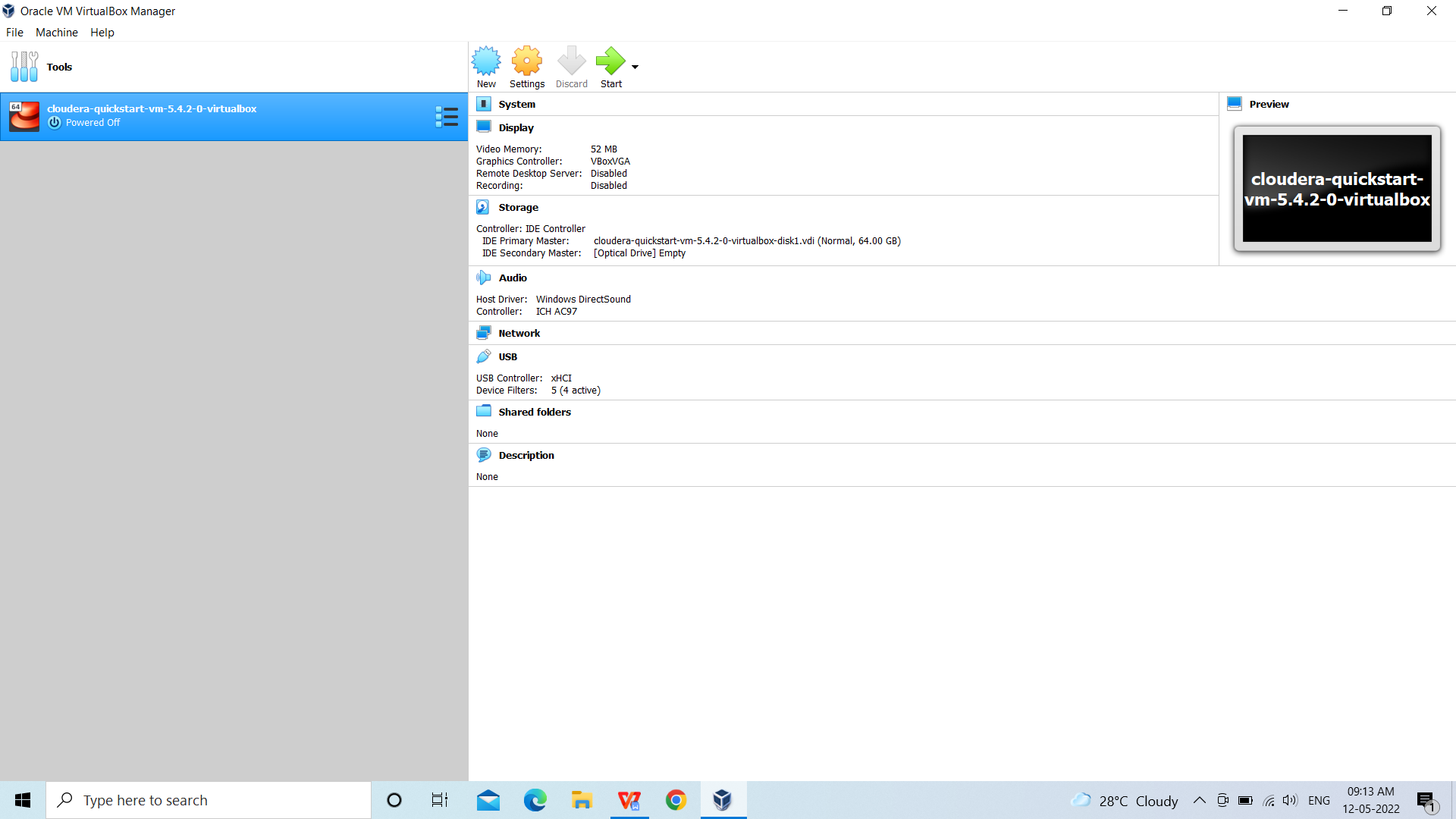
**Steps to execute Map-Reduce Program on Cloudera**

1. On Ubuntu OS open terminal type **virtualbox**
2. Start cloudera by clicking on green arrow



1. Upload input file from cloudera to HDFS file system using following command . Input file must be in cloudera home.

File name you want on HDFS

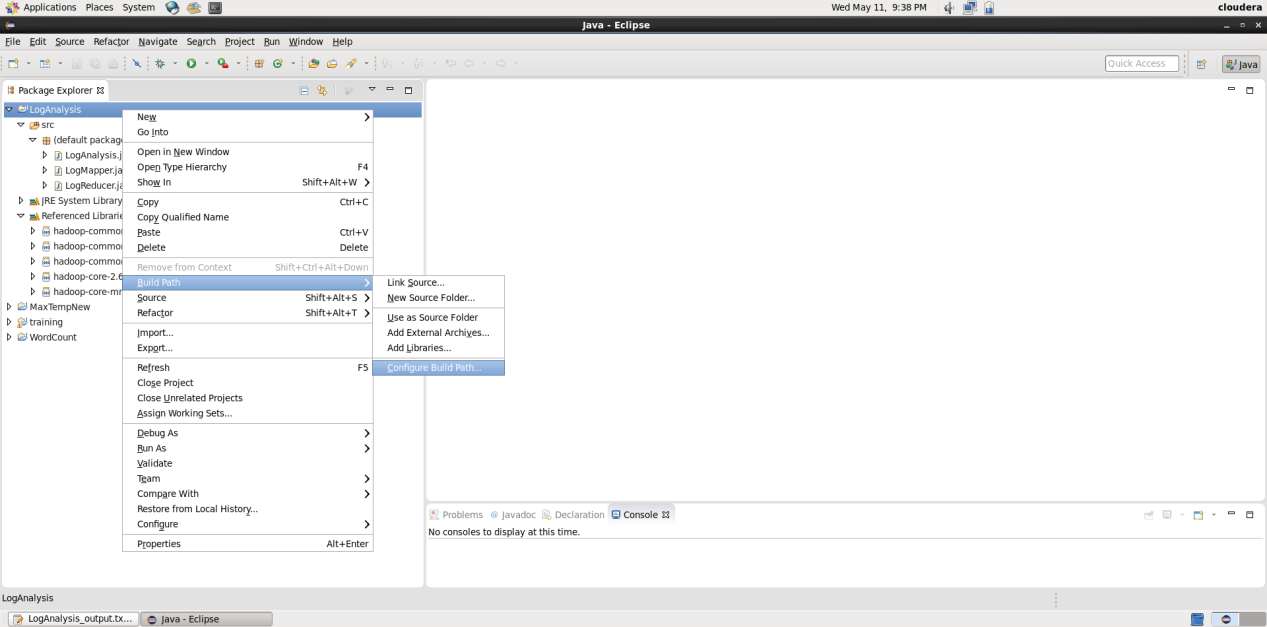
[cloudera@quickstart ~]$ hadoop fs -put f1.txt sample.txt

Input file name

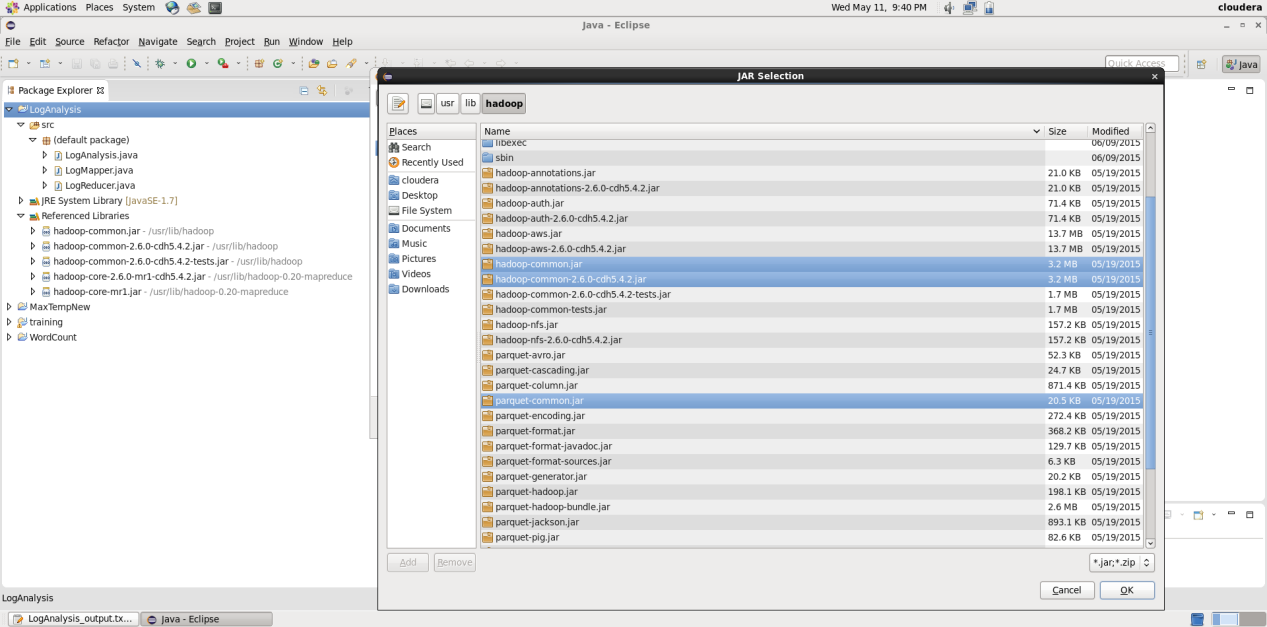
1. Open eclipse and create new Java Project. (e.g. LogAnalysis)
2. Add three separate java source files (1. mapper 2. reducer 3. Driver)

Right click to Java Project you created in step 4-----> select new class---> Give name(LogAnalysis,LogMapper and LogReducer)

1. Import following jar files using Right Click to Java Project---> Build Path---> Configure Build Path---> Add External Jar Files ---

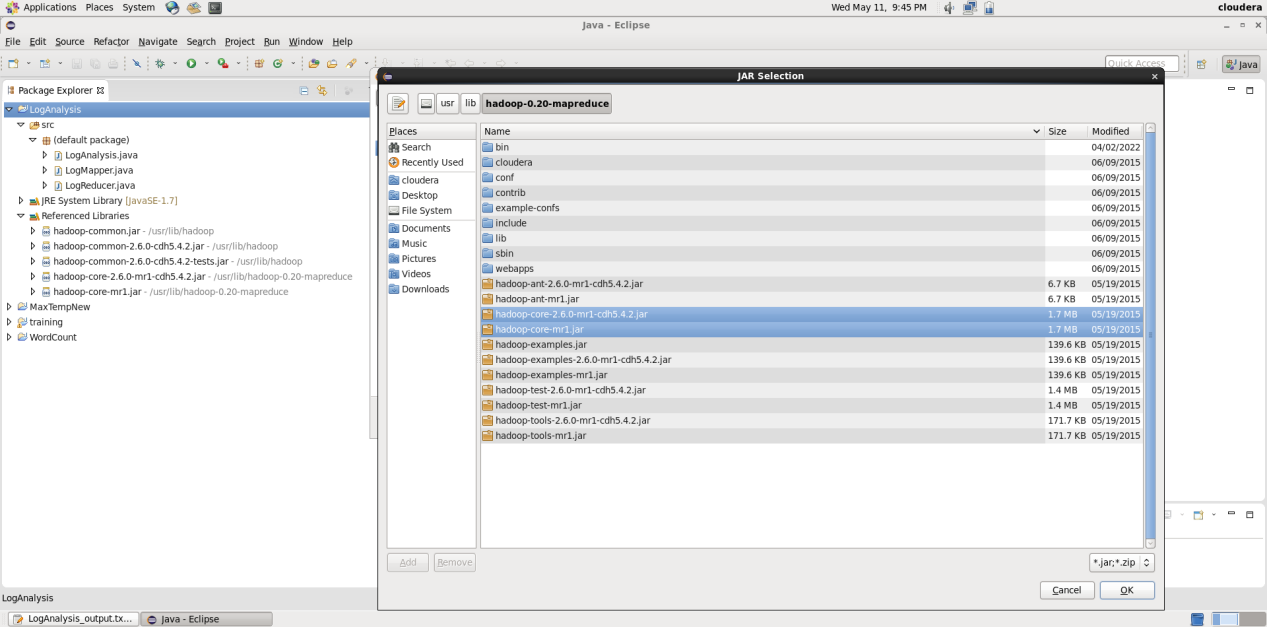


Add following jar files from /usr/lib/hadoop folder----hadoop-common.jar, hadoop-common-2.4,



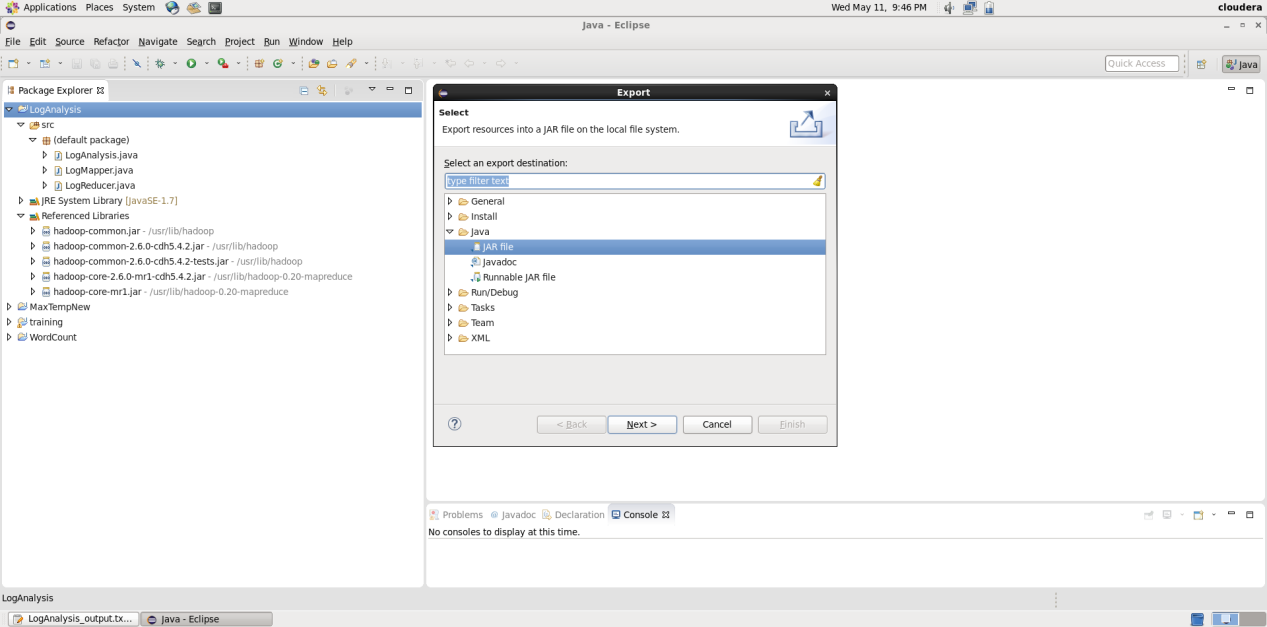
From /usr/lib/hadoop-0.20-mapreduce folder

Add mr.jar files



1. Make your java project Error free
2. Create Jar file of your java project.

Right click to Project ---->click Export----> select jar file--->click Next----> give Jar file name ( same as project name)



1. Create output directory using command

Hadoop fs -mkdir logop

10 Execute map reduce program using following command

[cloudera@quickstart ~]$ hadoop jar LogAnalysis.jar LogAnalysis access\_log.txt logop

1. After executing step 10 in logop two files will be created as below

[cloudera@quickstart ~]$ hadoop fs -ls logop

Found 2 items

-rw-r--r-- 1 cloudera cloudera 0 2022-03-21 02:24 logop/\_SUCCESS

-rw-r--r-- 1 cloudera cloudera 4551 2022-03-21 02:24 logop/part-00000

1. Display output using following command

[cloudera@quickstart ~]$ hadoop fs -cat logop/part-00000