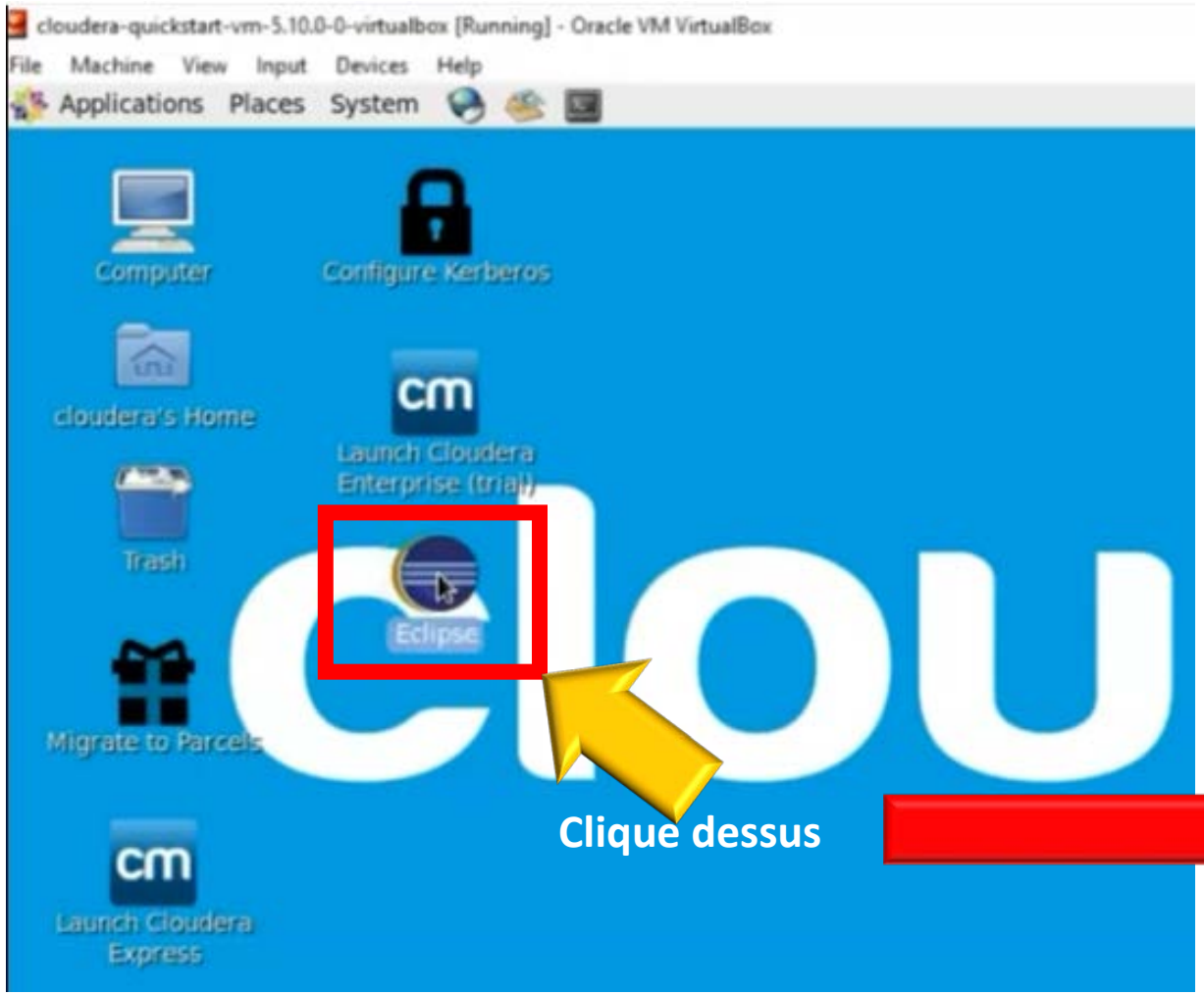
CLOUDERA

ensi ÉCOLE DES NOUVELLES SCIENCES & INGÉNIERIE

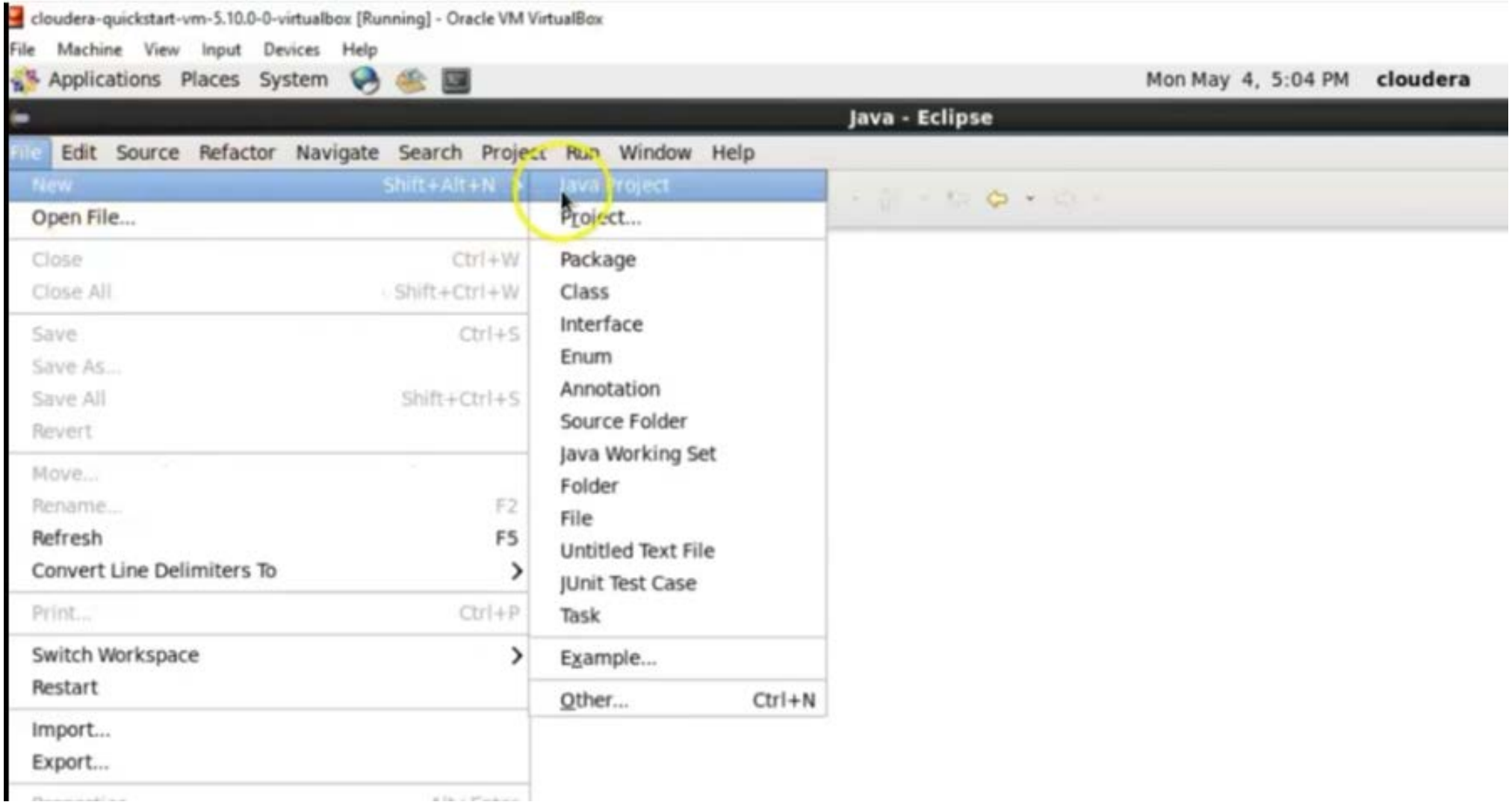
Atelier 1 Cloudera

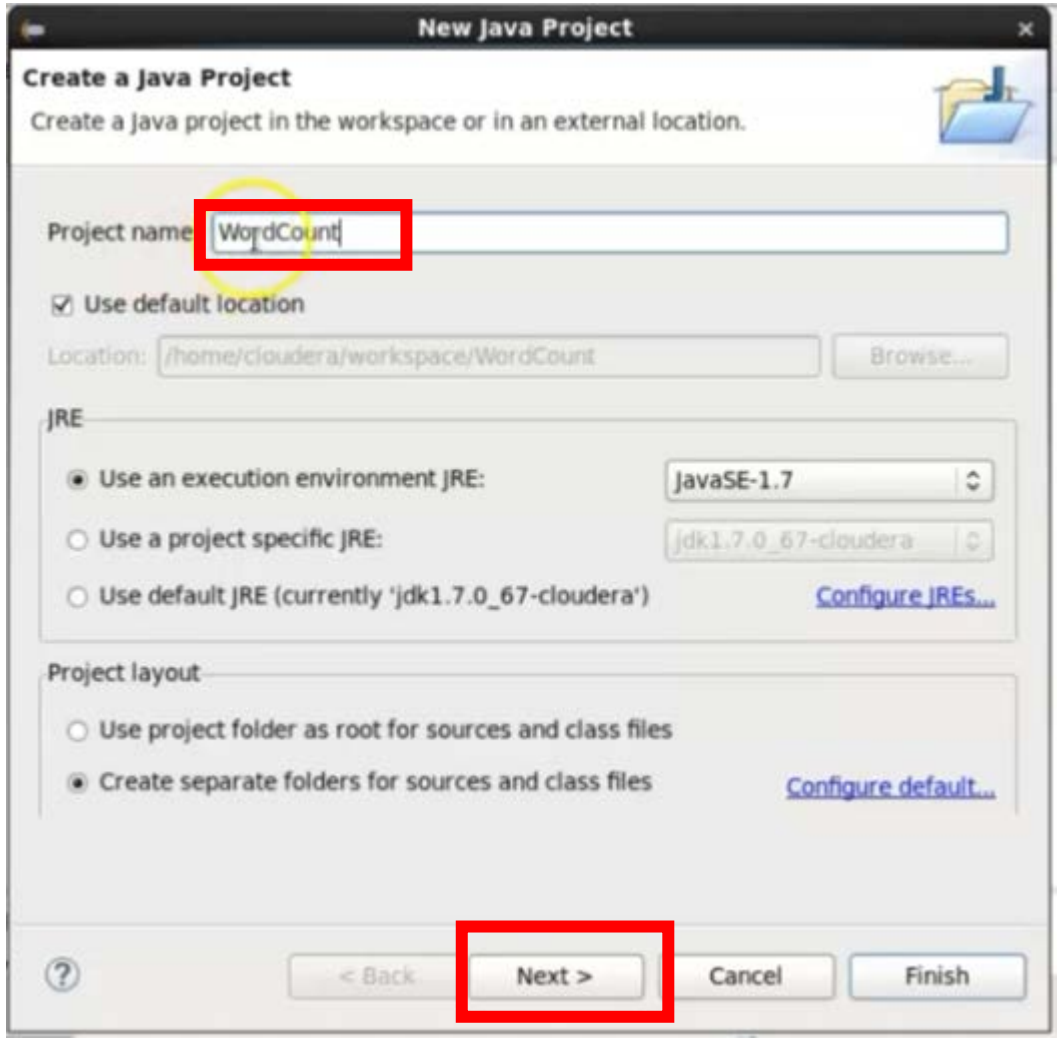
Partie 2
Map Reduce

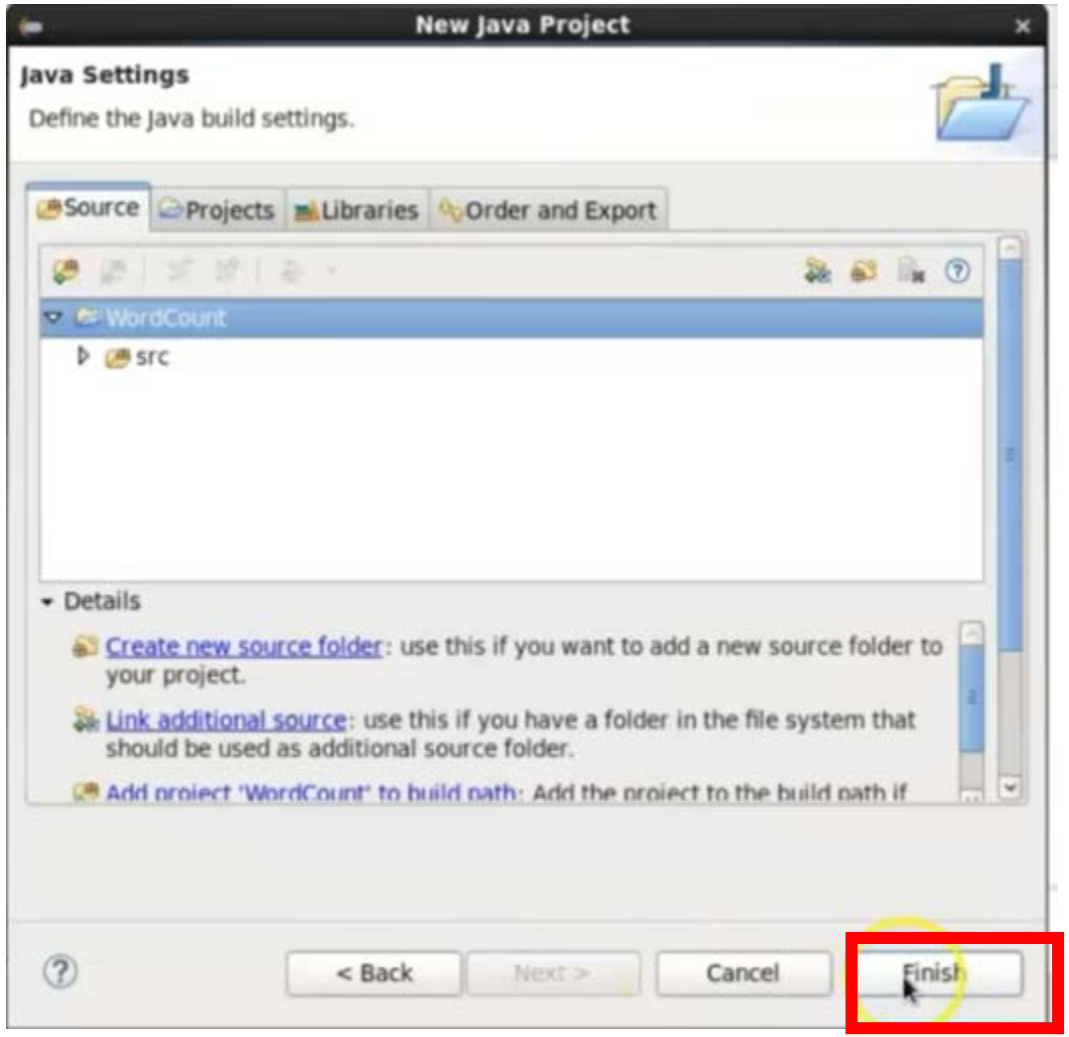
Med el Assad 21/22 5Ginfo Big Data

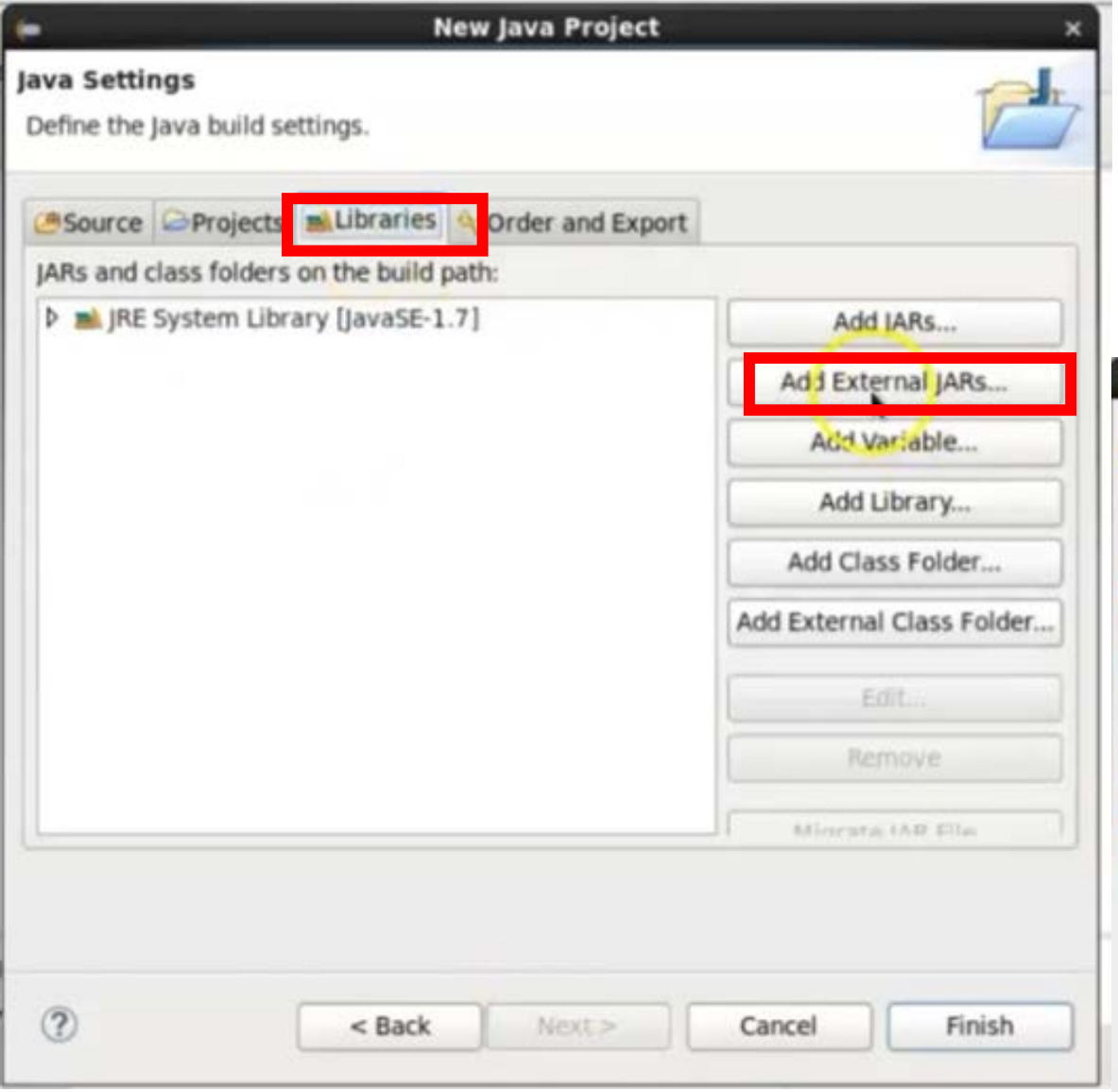

cloudera-quickstart-vm-5.10.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Computer
Configure Kerberos
cloudera's Home
Launch Cloudera Enterprise (trial)
Trash
Eclipse
Migrate to Parcels
Launch Cloudera Express
Clique dessus

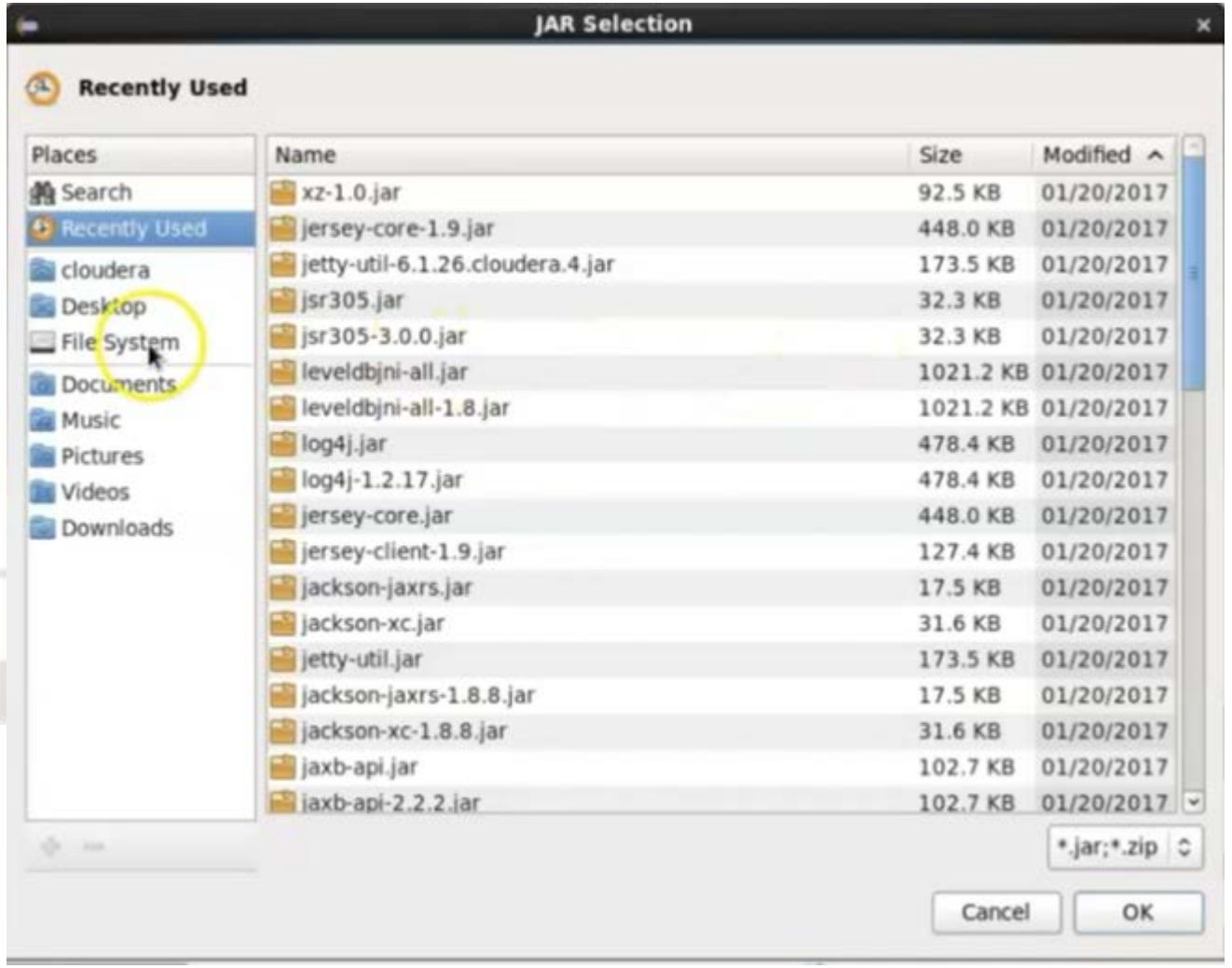
eclipse
LUNA

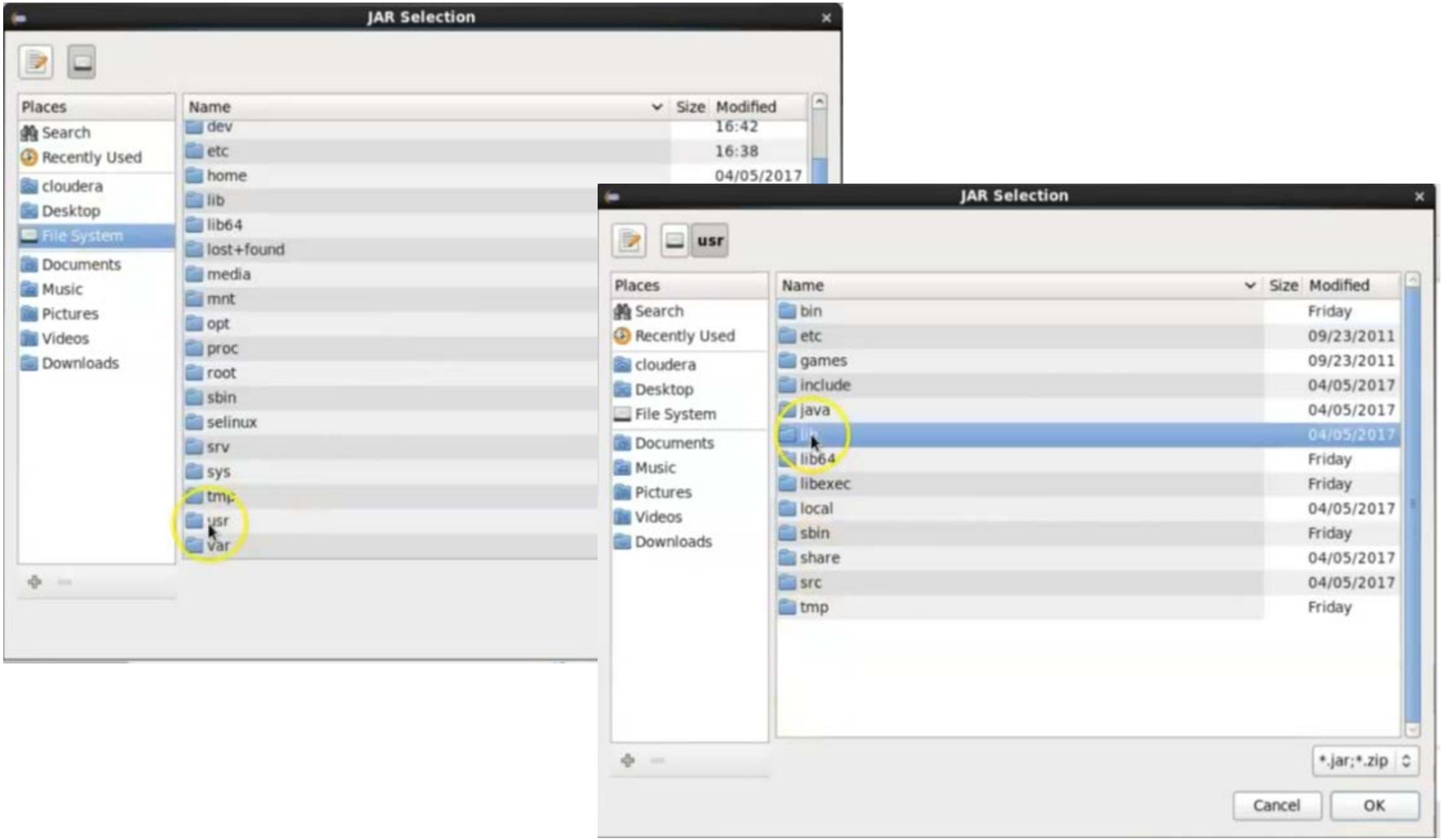

cloudera-quickstart-vm-5.10.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Mon May 4, 5:04 PM cloudera
Java - Eclipse
File Edit Source Refactor Navigate Search Project Run Window Help
New Shift+Alt+N
Open File...
Close Ctrl+W
Close All Shift+Ctrl+W
Save Ctrl+S
Save As...
Save All Shift+Ctrl+S
Revert
Move...
Rename... F2
Refresh F5
Convert Line Delimiters To
Print... Ctrl+P
Switch Workspace
Restart
Import...
Export...
Java Project
Project...
Package
Class
Interface
Enum
Annotation
Source Folder
Java Working Set
Folder
File
Untitled Text File
JUnit Test Case
Task
Example...
Other... Ctrl+N

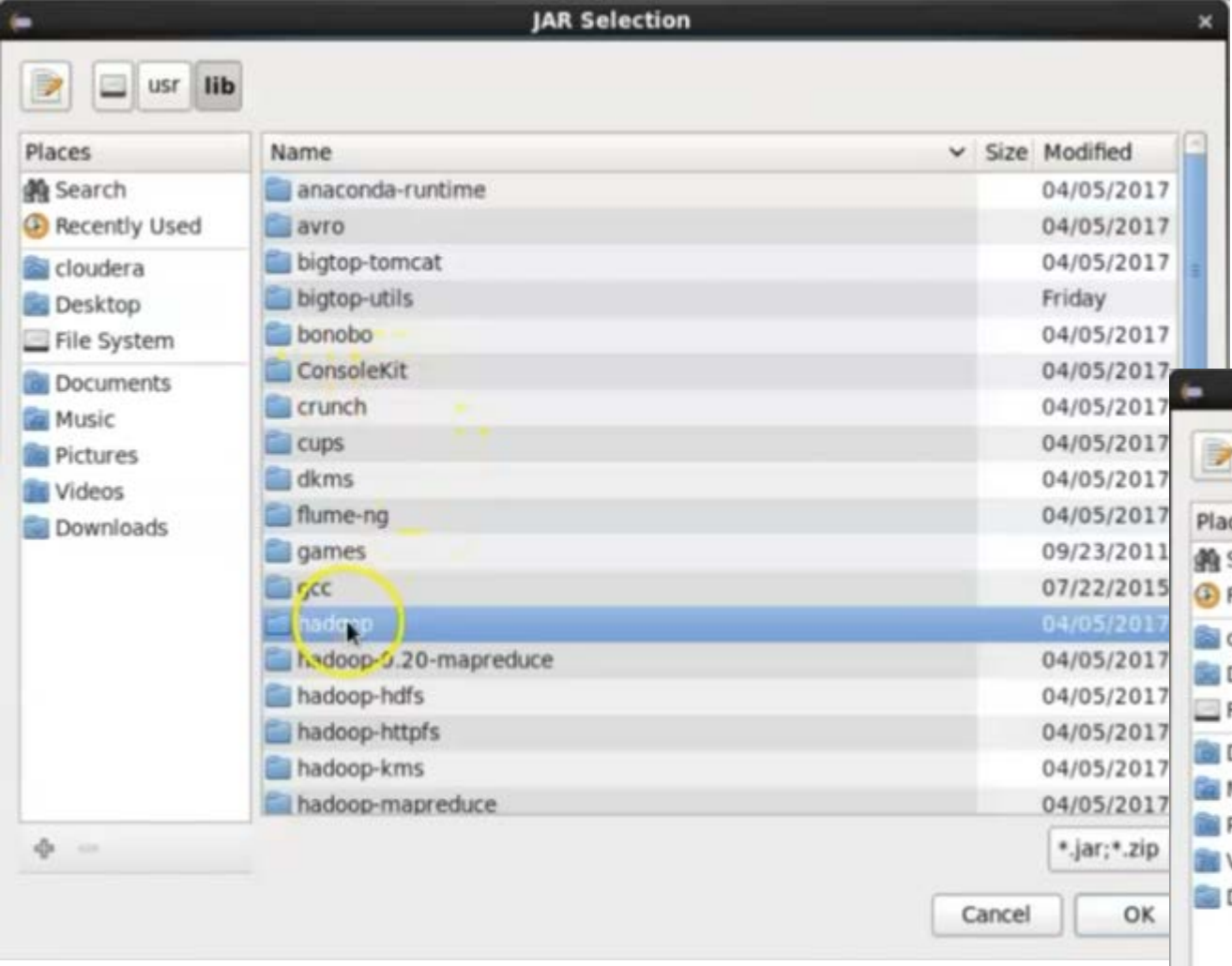

New Java Project
Create a Java Project
Create a Java project in the workspace or in an external location.
Project name: WordCount
Use default location
Location: /home/cloudera/workspace/WordCount
Browse...
JRE
Use an execution environment JRE: JavaSE-1.7
Use a project specific JRE: jdk1.7.0_67-cloudera
Use default JRE (currently 'jdk1.7.0_67-cloudera')
Configure JREs...
Project layout
Use project folder as root for sources and class files
Create separate folders for sources and class files
Configure default...
< Back
Next >
Cancel
Finish

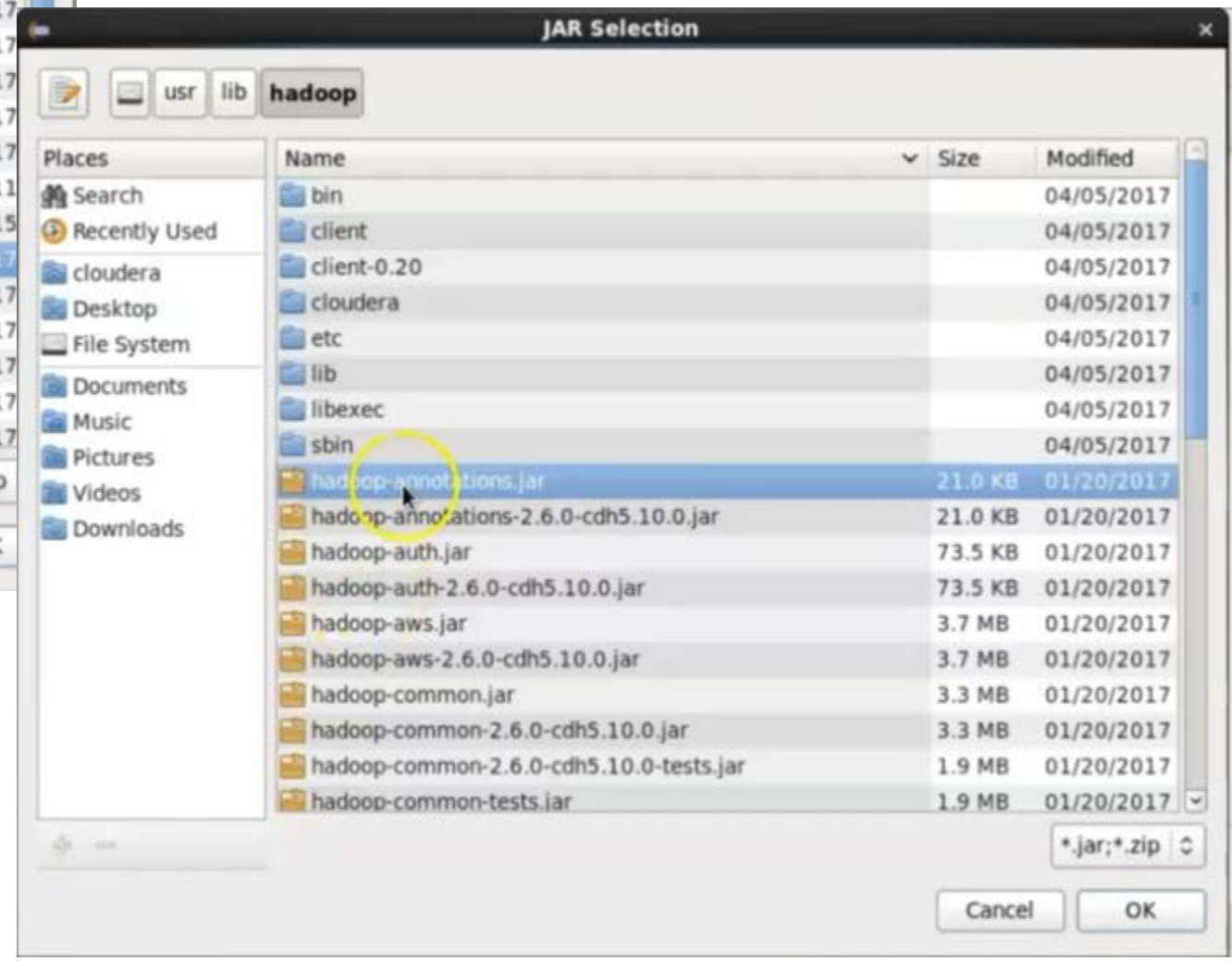

New Java Project
Java Settings
Define the Java build settings.
Source
Projects
Libraries
Order and Export
WordCount
src
Details
Create new source folder: use this if you want to add a new source folder to your project.
Link additional source: use this if you have a folder in the file system that should be used as additional source folder.
Add project 'WordCount' to build path: Add the project to the build path if
< Back
Next >
Cancel
Finish

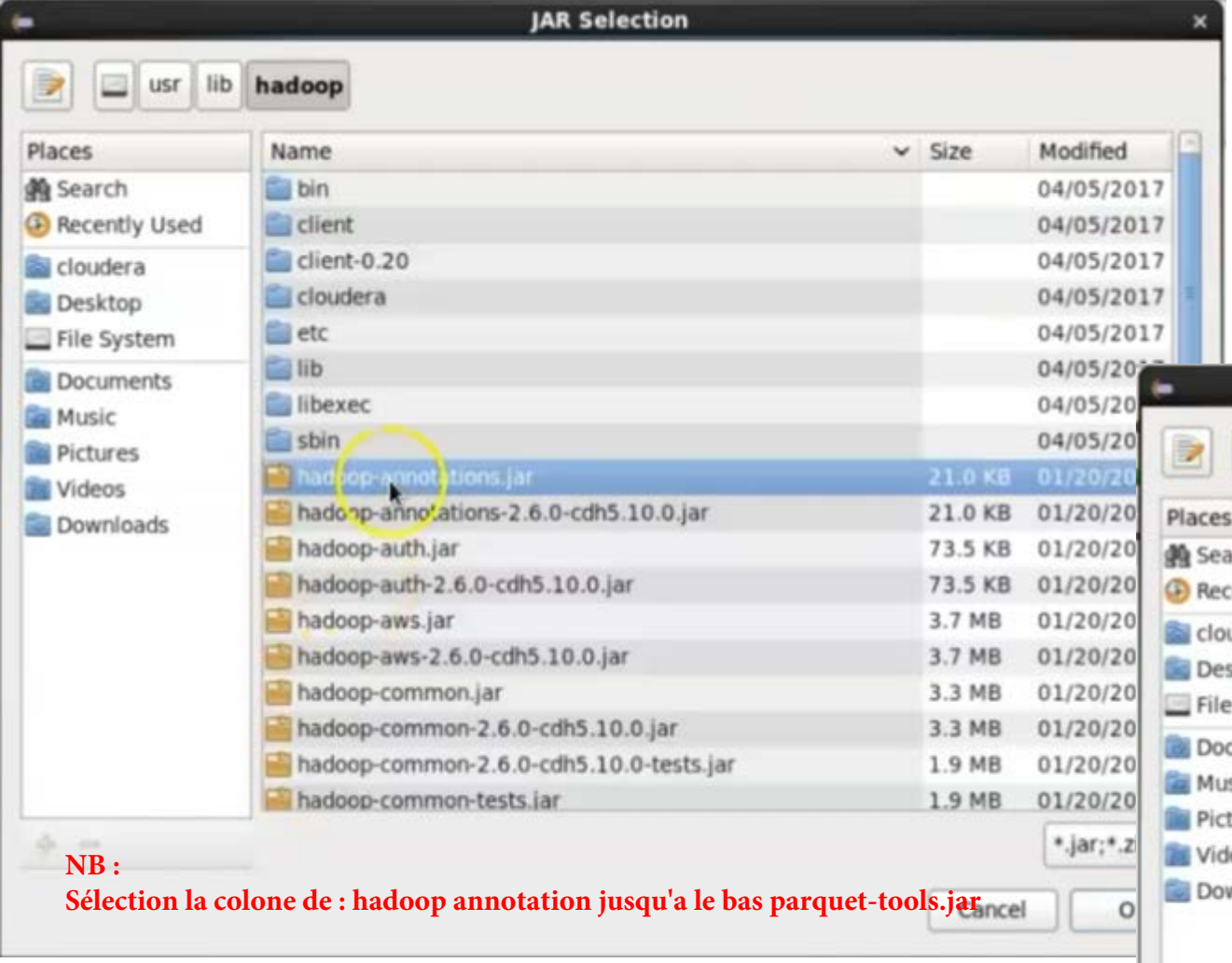

New Java Project
Java Settings
Define the Java build settings.
Source
Projects
Libraries
Order and Export
JARs and class folders on the build path:
JRE System Library [JavaSE-1.7]
Add JARs...
Add External JARs...
Add Variable...
Add Library...
Add Class Folder...
Add External Class Folder...
Edit...
Remove
< Back
Next >
Cancel
Finish

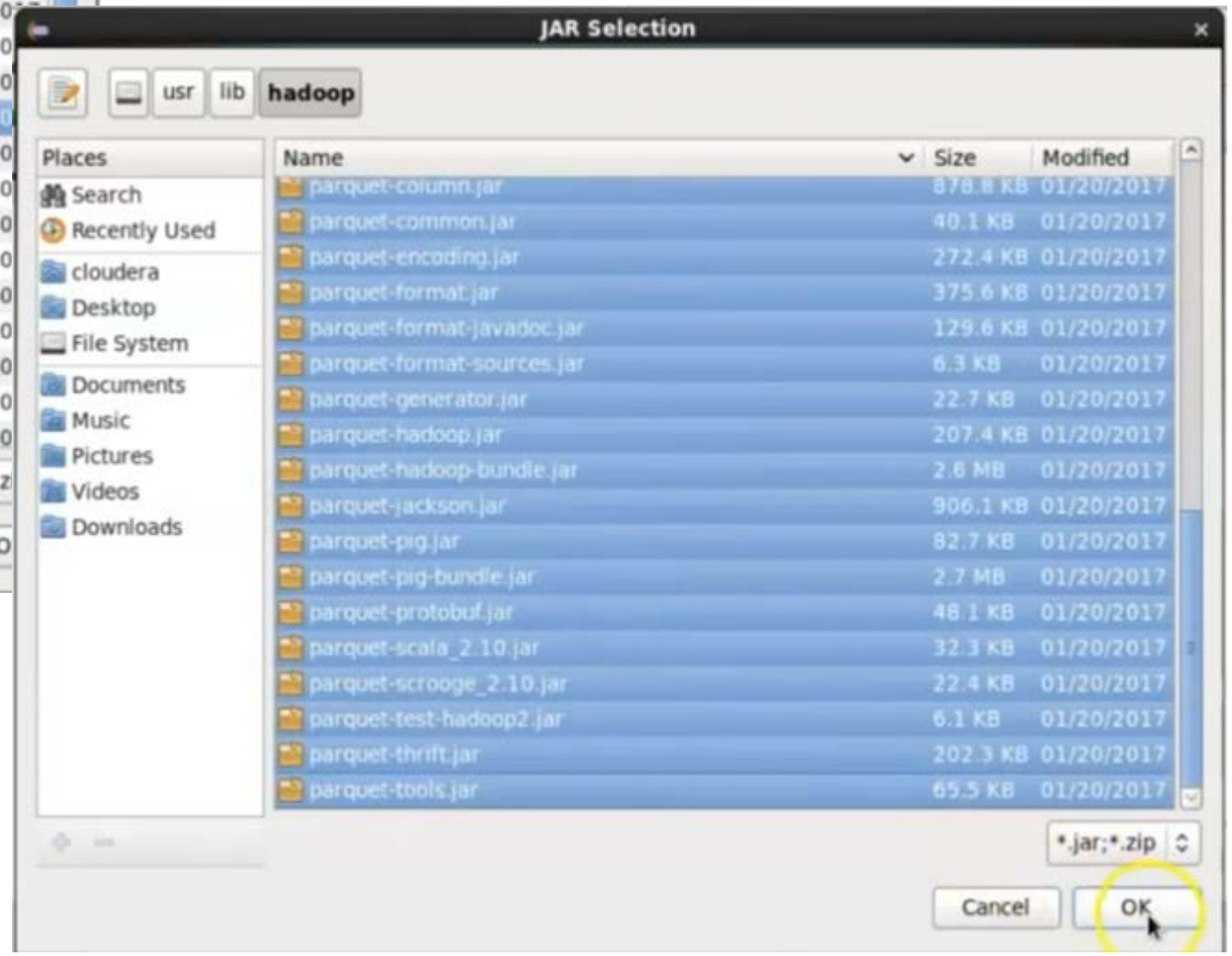

JAR Selection
Recently Used
Places
Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads
Name
Size
Modified
xz-1.0.jar 92.5 KB 01/20/2017
jersey-core-1.9.jar 448.0 KB 01/20/2017
jetty-util-6.1.26.cloudera.4.jar 173.5 KB 01/20/2017
jsr305.jar 32.3 KB 01/20/2017
jsr305-3.0.0.jar 32.3 KB 01/20/2017
leveldbjni-all.jar 1021.2 KB 01/20/2017
leveldbjni-all-1.8.jar 1021.2 KB 01/20/2017
log4j.jar 478.4 KB 01/20/2017
log4j-1.2.17.jar 478.4 KB 01/20/2017
jersey-core.jar 448.0 KB 01/20/2017
jersey-client-1.9.jar 127.4 KB 01/20/2017
jackson-jaxrs.jar 17.5 KB 01/20/2017
jackson-xc.jar 31.6 KB 01/20/2017
jetty-util.jar 173.5 KB 01/20/2017
jackson-jaxrs-1.8.8.jar 17.5 KB 01/20/2017
jackson-xc-1.8.8.jar 31.6 KB 01/20/2017
jaxb-api.jar 102.7 KB 01/20/2017
jaxb-api-2.2.2.jar 102.7 KB 01/20/2017
.jar;.zip
Cancel
OK


JAR Selection
Places
Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads
Name
Size
Modified
dev
16:42
etc
16:38
home
04/05/2017
lib
lib64
lost+found
media
mnt
opt
proc
root
sbin
selinux
srv
sys
tmp
usr
var
JAR Selection
usr
Places
Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads
Name
Size
Modified
bin
Friday
etc
09/23/2011
games
09/23/2011
include
04/05/2017
java
04/05/2017
lib
04/05/2017
lib64
Friday
libexec
Friday
local
04/05/2017
sbin
Friday
share
04/05/2017
src
04/05/2017
tmp
Friday
.jar;.zip
Cancel
OK


JAR Selection
usr
lib
Places
Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads
Name
Size
Modified
anaconda-runtime 04/05/2017
avro 04/05/2017
bigtop-tomcat 04/05/2017
bigtop-utils Friday
bonobo 04/05/2017
ConsoleKit 04/05/2017
crunch 04/05/2017
cups 04/05/2017
dkms 04/05/2017
flume-ng 04/05/2017
games 09/23/2011
gcc 07/22/2015
hadoop 04/05/2017
hadoop-0.20-mapreduce 04/05/2017
hadoop-hdfs 04/05/2017
hadoop-httpfs 04/05/2017
hadoop-kms 04/05/2017
hadoop-mapreduce 04/05/2017
.jar;.zip
Cancel
OK


JAR Selection
usr
lib
hadoop
Places
Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads
Name
Size
Modified
bin 04/05/2017
client 04/05/2017
client-0.20 04/05/2017
cloudera 04/05/2017
etc 04/05/2017
lib 04/05/2017
libexec 04/05/2017
sbin 04/05/2017
hadoop-annotations.jar 21.0 KB 01/20/2017
hadoop-annotations-2.6.0-cdh5.10.0.jar 21.0 KB 01/20/2017
hadoop-auth.jar 73.5 KB 01/20/2017
hadoop-auth-2.6.0-cdh5.10.0.jar 73.5 KB 01/20/2017
hadoop-aws.jar 3.7 MB 01/20/2017
hadoop-aws-2.6.0-cdh5.10.0.jar 3.7 MB 01/20/2017
hadoop-common.jar 3.3 MB 01/20/2017
hadoop-common-2.6.0-cdh5.10.0.jar 3.3 MB 01/20/2017
hadoop-common-2.6.0-cdh5.10.0-tests.jar 1.9 MB 01/20/2017
hadoop-common-tests.jar 1.9 MB 01/20/2017
.jar;.zip
Cancel
OK

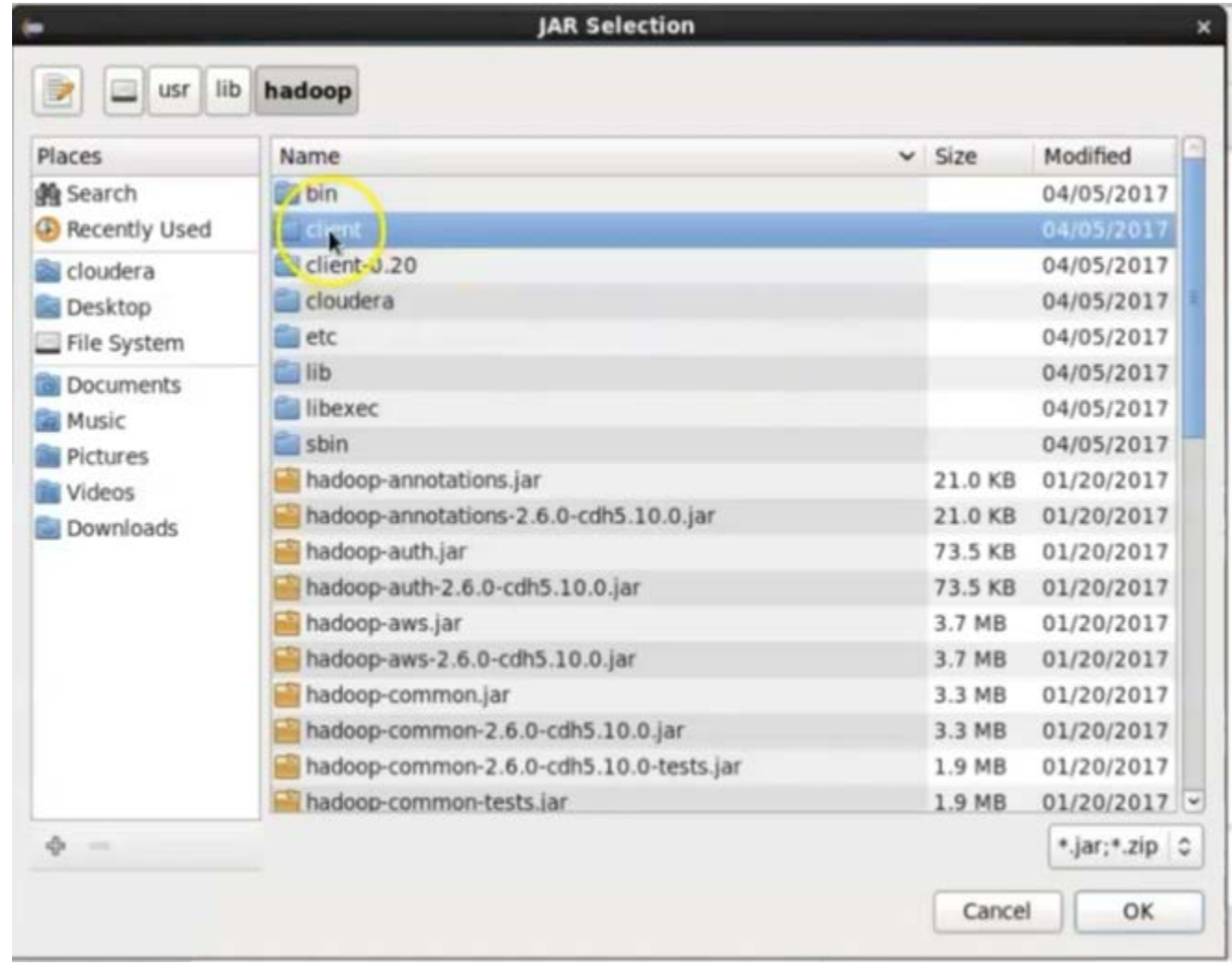


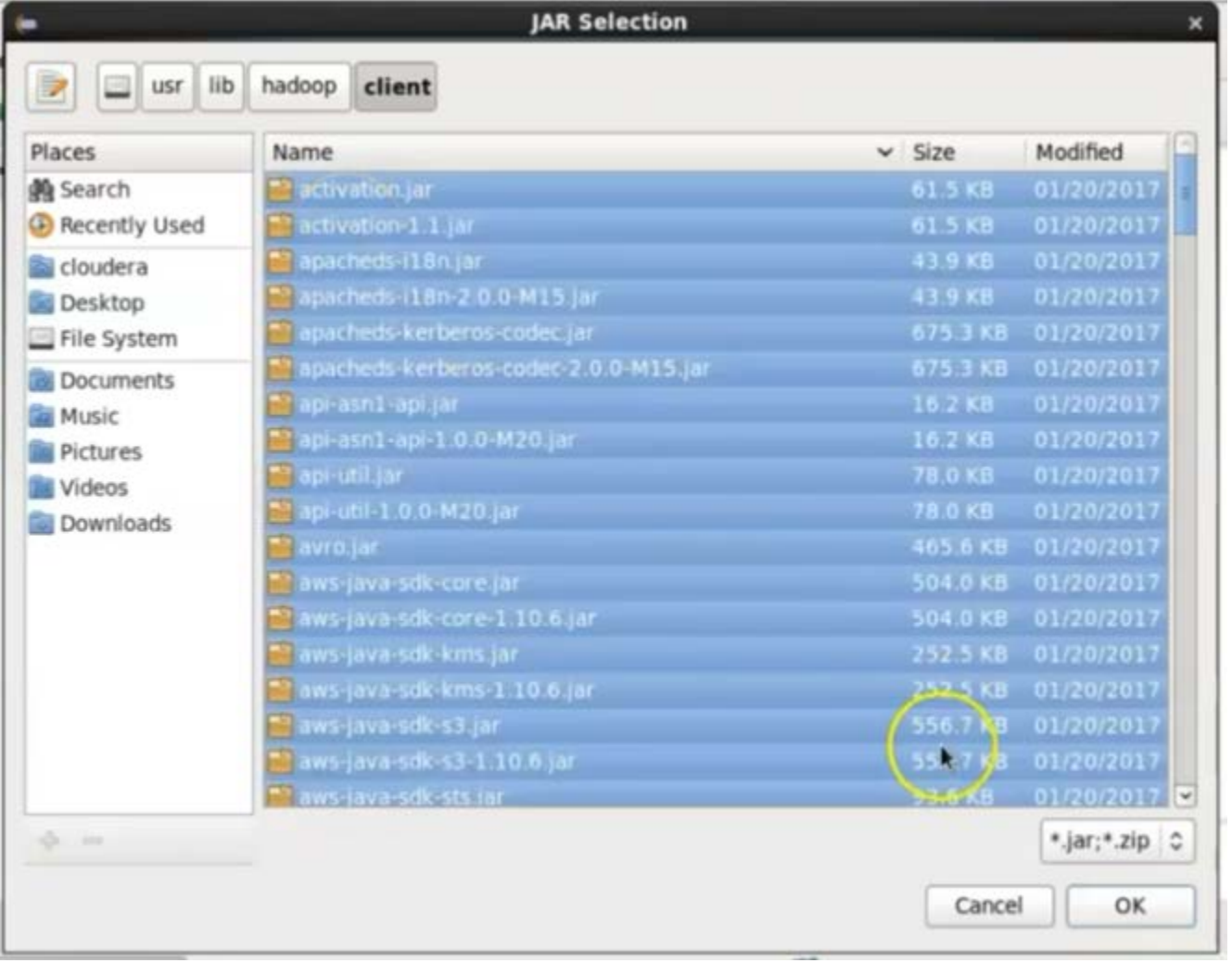


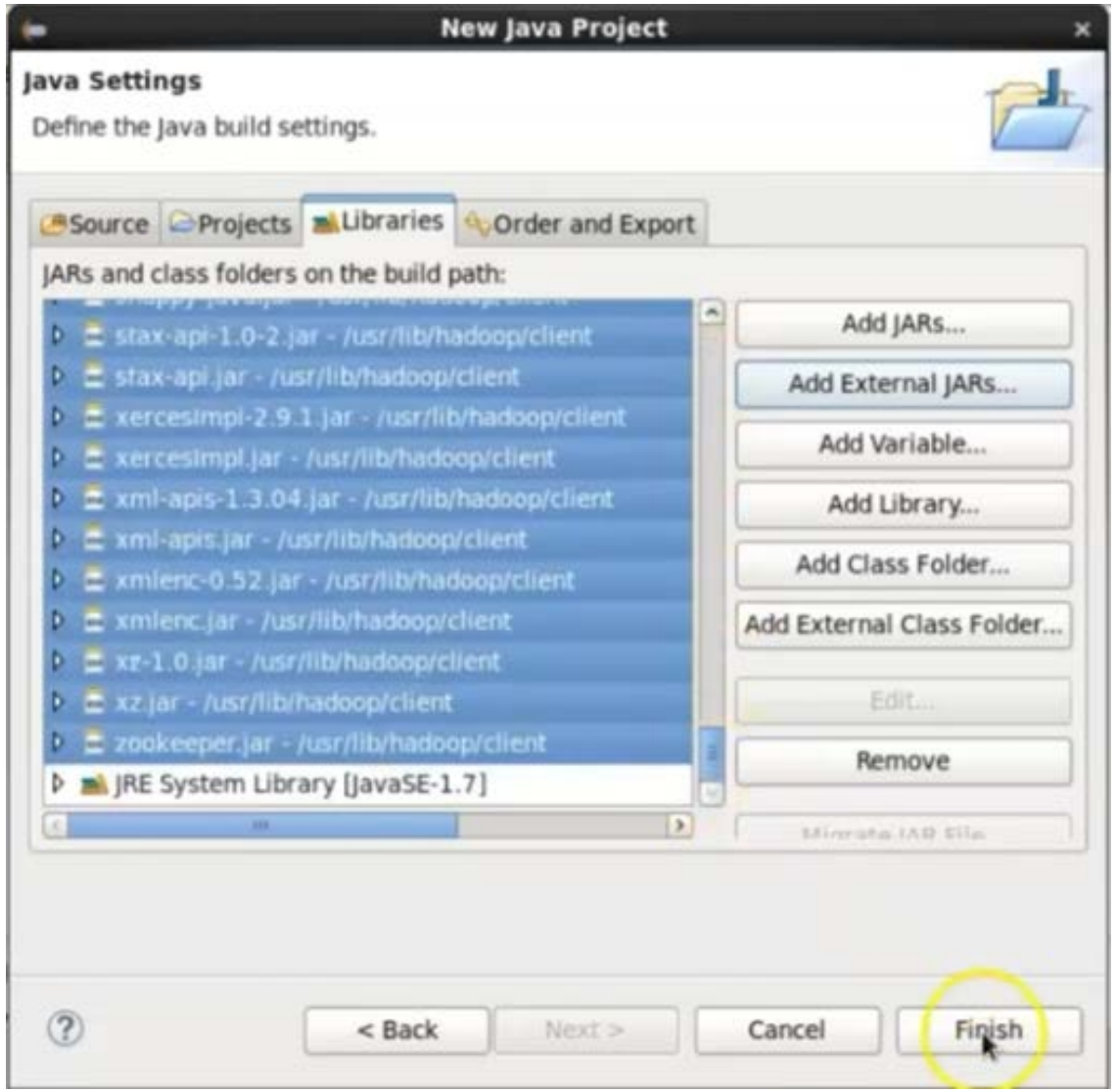
NB :

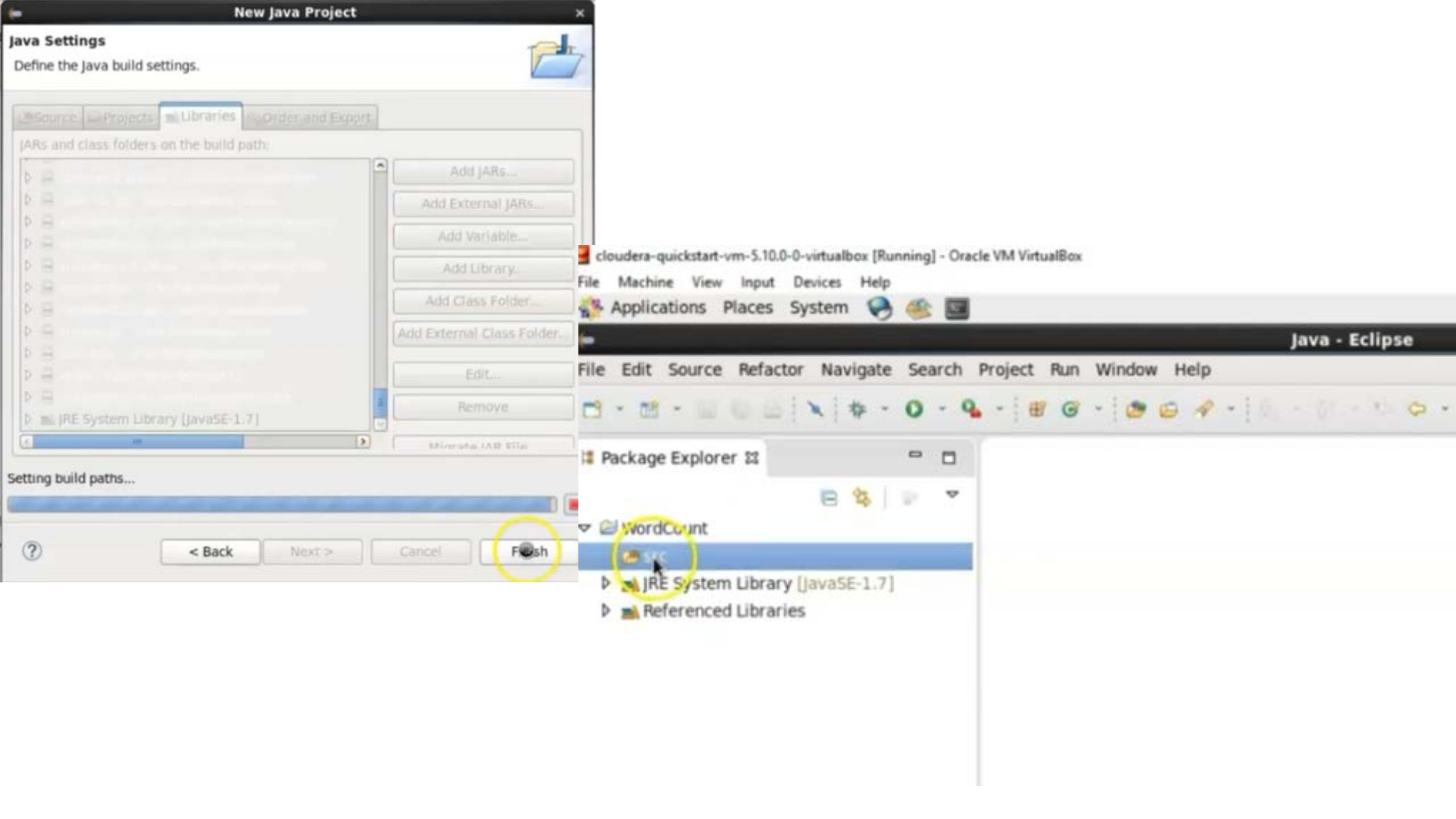
Sélection la colone de : hadoop annotation jusqu'a le bas parquet-tools.jar

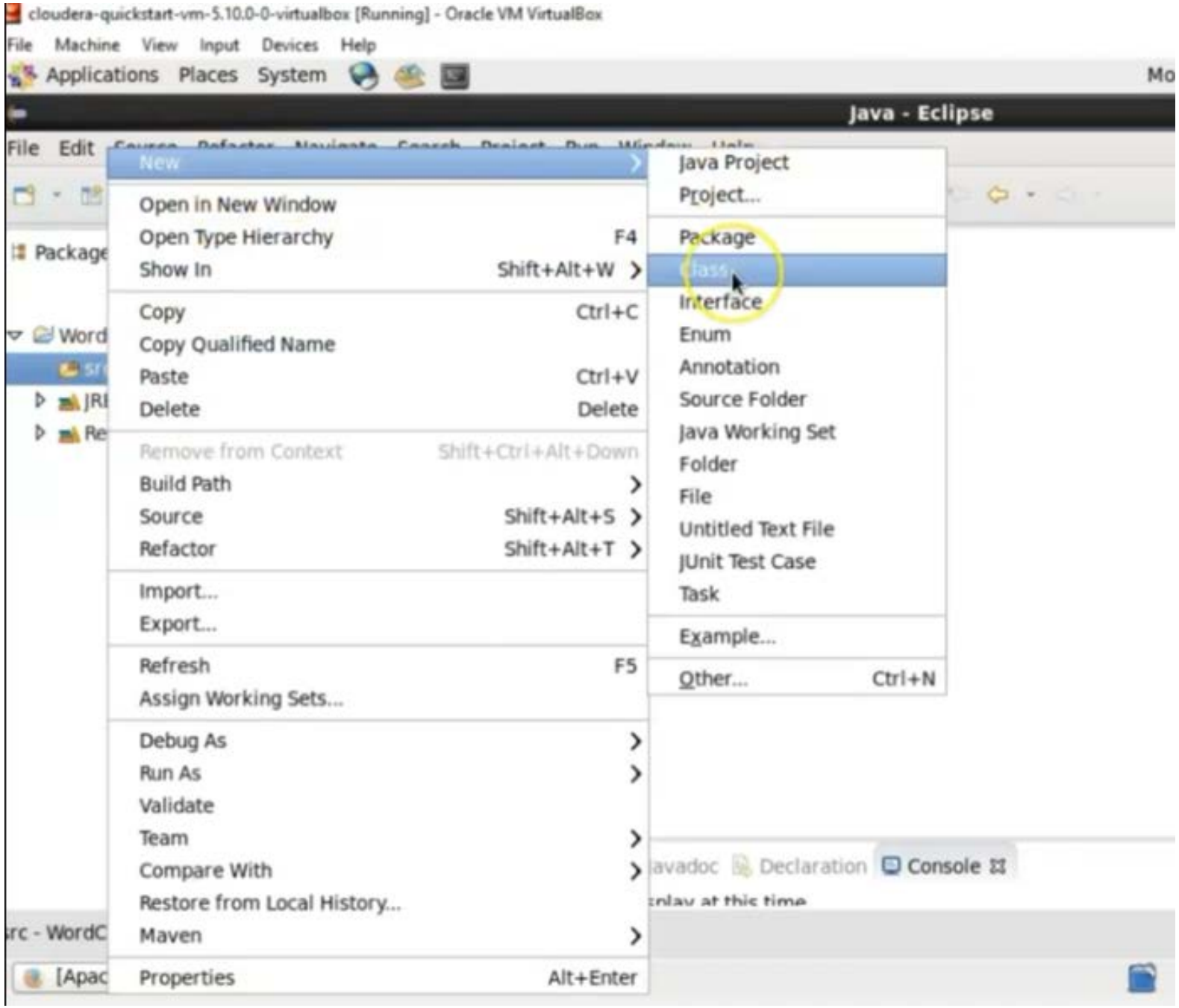
New Java Project
Java Settings
Define the Java build settings.
Source
Projects
Libraries
Order and Export
JARs and class folders on the build path:
hadoop-annotations-2.6.0-cdh5.10.0.jar - /usr/lib
hadoop-annotations.jar - /usr/lib/hadoop
hadoop-auth-2.6.0-cdh5.10.0.jar - /usr/lib/hadoop
hadoop-auth.jar - /usr/lib/hadoop
hadoop-aws-2.6.0-cdh5.10.0.jar - /usr/lib/hadoop
hadoop-aws.jar - /usr/lib/hadoop
hadoop-common-2.6.0-cdh5.10.0-tests.jar - /usr/
hadoop-common-2.6.0-cdh5.10.0.jar - /usr/lib/ha
hadoop-common-tests.jar - /usr/lib/hadoop
hadoop-common.jar - /usr/lib/hadoop
hadoop-nfs-2.6.0-cdh5.10.0.jar - /usr/lib/hadoop
hadoop-nfs.jar - /usr/lib/hadoop
Add JARs...
Add External JARs...
Add Variable...
Add Library...
Add Class Folder...
Add External Class Folder...
Edit...
Remove
< Back
Next >
Cancel
Finish

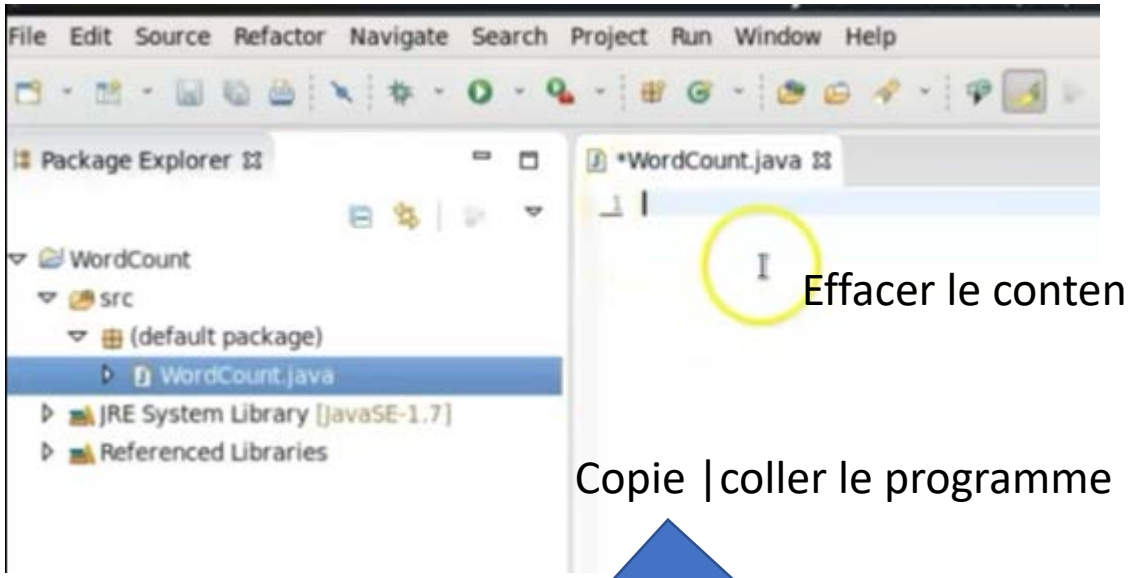

JAR Selection
usr
lib
hadoop
Places
Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads
Name
Size
Modified
bin 04/05/2017
client 04/05/2017
client-0.20 04/05/2017
cloudera 04/05/2017
etc 04/05/2017
lib 04/05/2017
libexec 04/05/2017
sbin 04/05/2017
hadoop-annotations.jar 21.0 KB 01/20/2017
hadoop-annotations-2.6.0-cdh5.10.0.jar 21.0 KB 01/20/2017
hadoop-auth.jar 73.5 KB 01/20/2017
hadoop-auth-2.6.0-cdh5.10.0.jar 73.5 KB 01/20/2017
hadoop-aws.jar 3.7 MB 01/20/2017
hadoop-aws-2.6.0-cdh5.10.0.jar 3.7 MB 01/20/2017
hadoop-common.jar 3.3 MB 01/20/2017
hadoop-common-2.6.0-cdh5.10.0.jar 3.3 MB 01/20/2017
hadoop-common-2.6.0-cdh5.10.0-tests.jar 1.9 MB 01/20/2017
hadoop-common-tests.jar 1.9 MB 01/20/2017
.jar;.zip
Cancel
OK


JAR Selection
usr
lib
hadoop
client
Places
Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads
Name
Size
Modified
activation.jar 61.5 KB 01/20/2017
activation-1.1.jar 61.5 KB 01/20/2017
apacheds-i18n.jar 43.9 KB 01/20/2017
apacheds-i18n-2.0.0-M15.jar 43.9 KB 01/20/2017
apacheds-kerberos-codec.jar 675.3 KB 01/20/2017
apacheds-kerberos-codec-2.0.0-M15.jar 675.3 KB 01/20/2017
api-asn1-api.jar 16.2 KB 01/20/2017
api-asn1-api-1.0.0-M20.jar 16.2 KB 01/20/2017
api-util.jar 78.0 KB 01/20/2017
api-util-1.0.0-M20.jar 78.0 KB 01/20/2017
avro.jar 465.6 KB 01/20/2017
aws-java-sdk-core.jar 504.0 KB 01/20/2017
aws-java-sdk-core-1.10.6.jar 504.0 KB 01/20/2017
aws-java-sdk-kms.jar 252.5 KB 01/20/2017
aws-java-sdk-kms-1.10.6.jar 01/20/2017
aws-java-sdk-s3.jar 556.7 KB 01/20/2017
aws-java-sdk-s3-1.10.6.jar 01/20/2017
aws-java-sdk-sts.jar 01/20/2017
.jar;.zip
Cancel
OK


New Java Project
Java Settings
Define the Java build settings.
Source
Projects
Libraries
Order and Export
JARs and class folders on the build path:
stax-api-1.0-2.jar - /usr/lib/hadoop/client
stax-api.jar - /usr/lib/hadoop/client
xercesImpl-2.9.1.jar - /usr/lib/hadoop/client
xercesImpl.jar - /usr/lib/hadoop/client
xml-apis-1.3.04.jar - /usr/lib/hadoop/client
xml-apis.jar - /usr/lib/hadoop/client
xmlenc-0.52.jar - /usr/lib/hadoop/client
xmlenc.jar - /usr/lib/hadoop/client
xz-1.0.jar - /usr/lib/hadoop/client
xz.jar - /usr/lib/hadoop/client
zookeeper.jar - /usr/lib/hadoop/client
JRE System Library [JavaSE-1.7]
Add JARs...
Add External JARs...
Add Variable...
Add Library...
Add Class Folder...
Add External Class Folder...
Edit...
Remove
< Back
Next >
Cancel
Finish

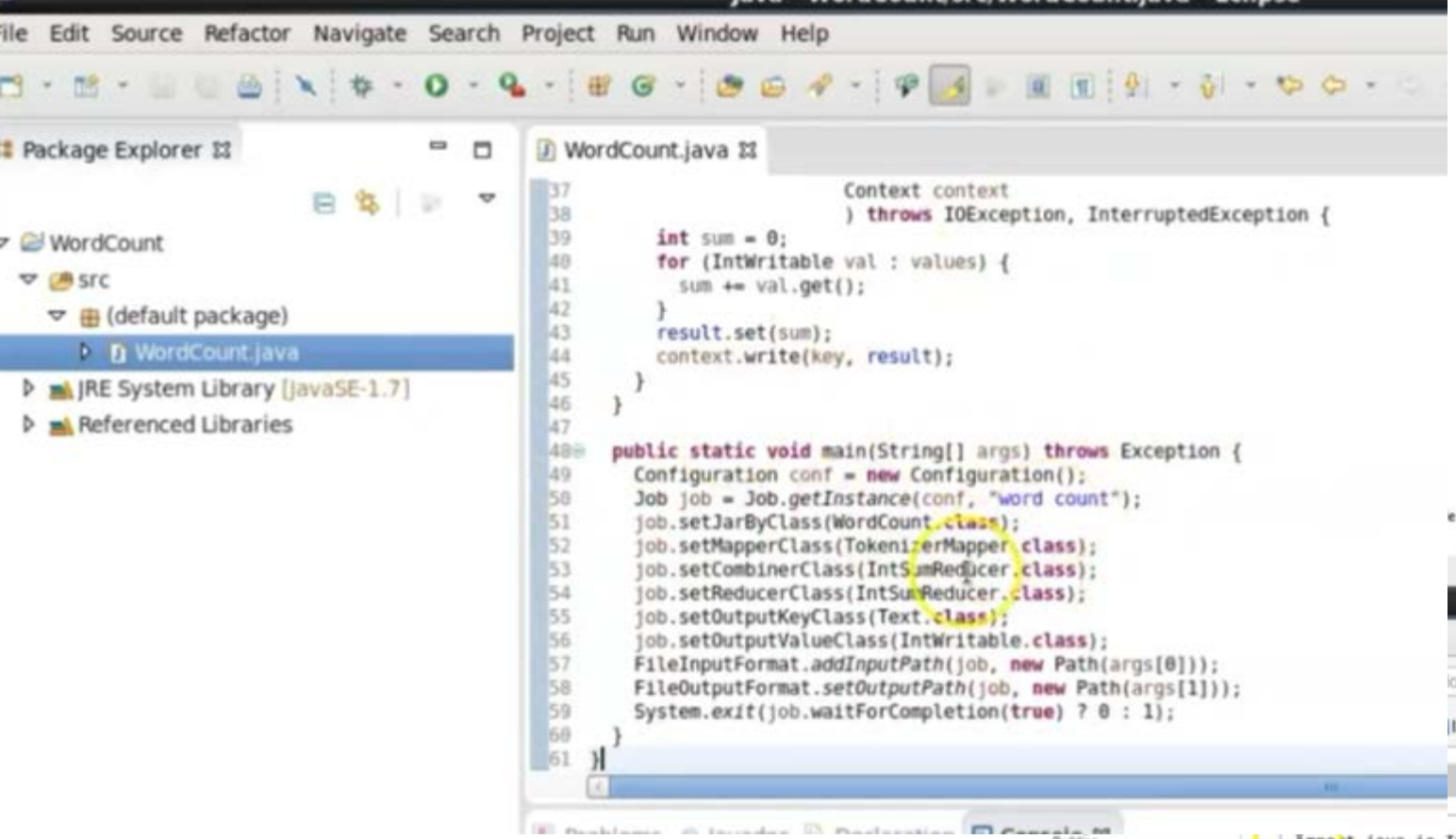

New Java Project
Java Settings
Define the Java build settings.
Source
Projects
Libraries
Order and Export
JARs and class folders on the build path:
JRE System Library [JavaSE-1.7]
Add JARs...
Add External JARs...
Add Variable...
Add Library...
Add Class Folder...
Add External Class Folder...
Edit...
Remove
Setting build paths...
< Back
Next >
Cancel
Finish
cloudera-quickstart-vm-5.10.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Java - Eclipse
File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer
WordCount
src
JRE System Library [JavaSE-1.7]
Referenced Libraries

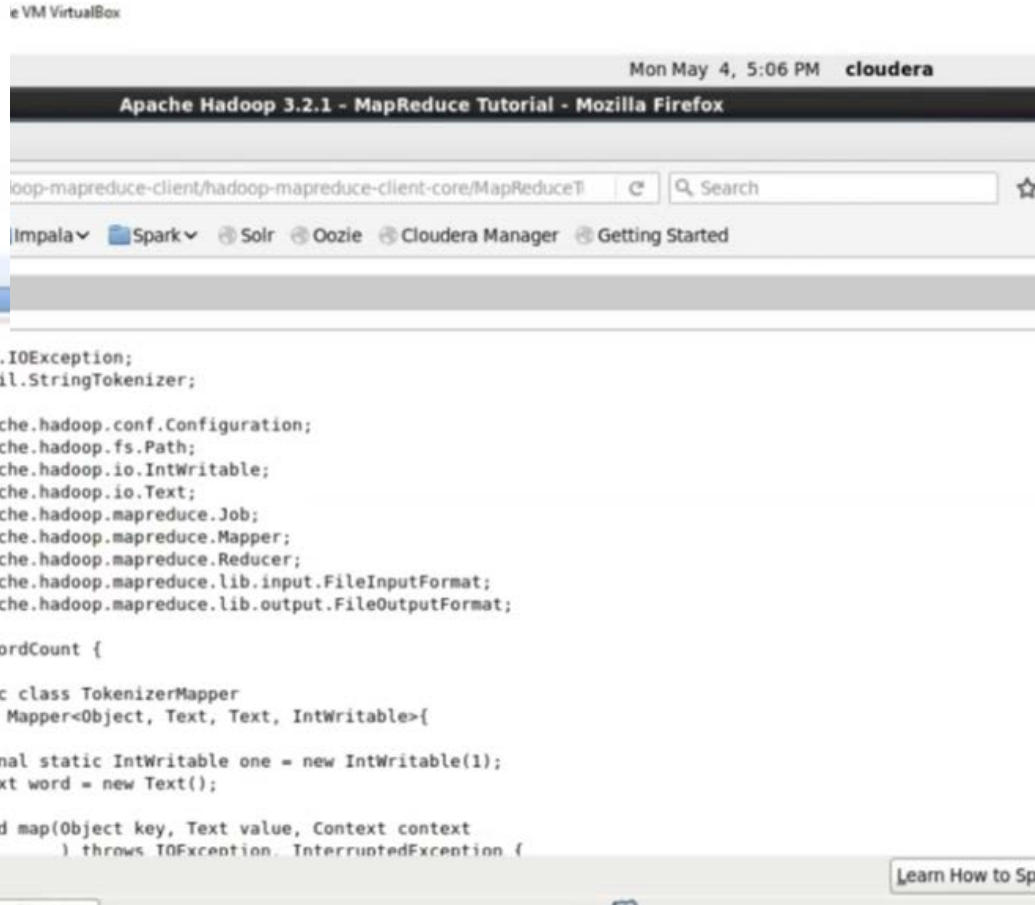

cloudera-quickstart-vm-5.10.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Java - Eclipse
File Edit
New
Open in New Window
Open Type Hierarchy F4
Show In Shift+Alt+W
Copy Ctrl+C
Copy Qualified Name
Paste Ctrl+V
Delete Delete
Remove from Context Shift+Ctrl+Alt+Down
Build Path
Source Shift+Alt+S
Refactor Shift+Alt+T
Import...
Export...
Refresh F5
Assign Working Sets...
Debug As
Run As
Validate
Team
Compare With
Restore from Local History...
Maven
Properties Alt+Enter
Java Project
Project...
Package
Class
Interface
Enum
Annotation
Source Folder
Java Working Set
Folder
File
Untitled Text File
JUnit Test Case
Task
Example...
Other... Ctrl+N
Declaration
Console

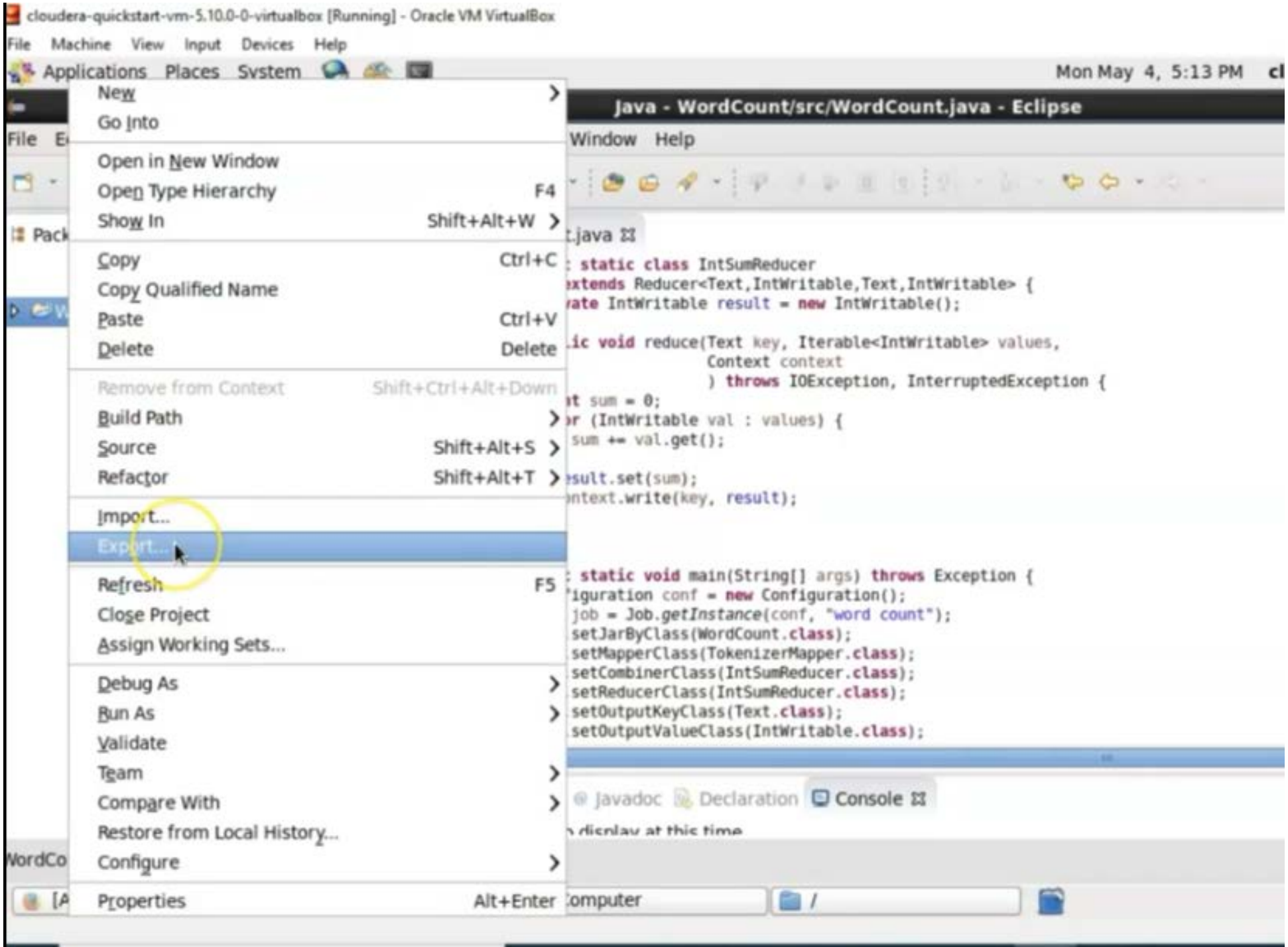


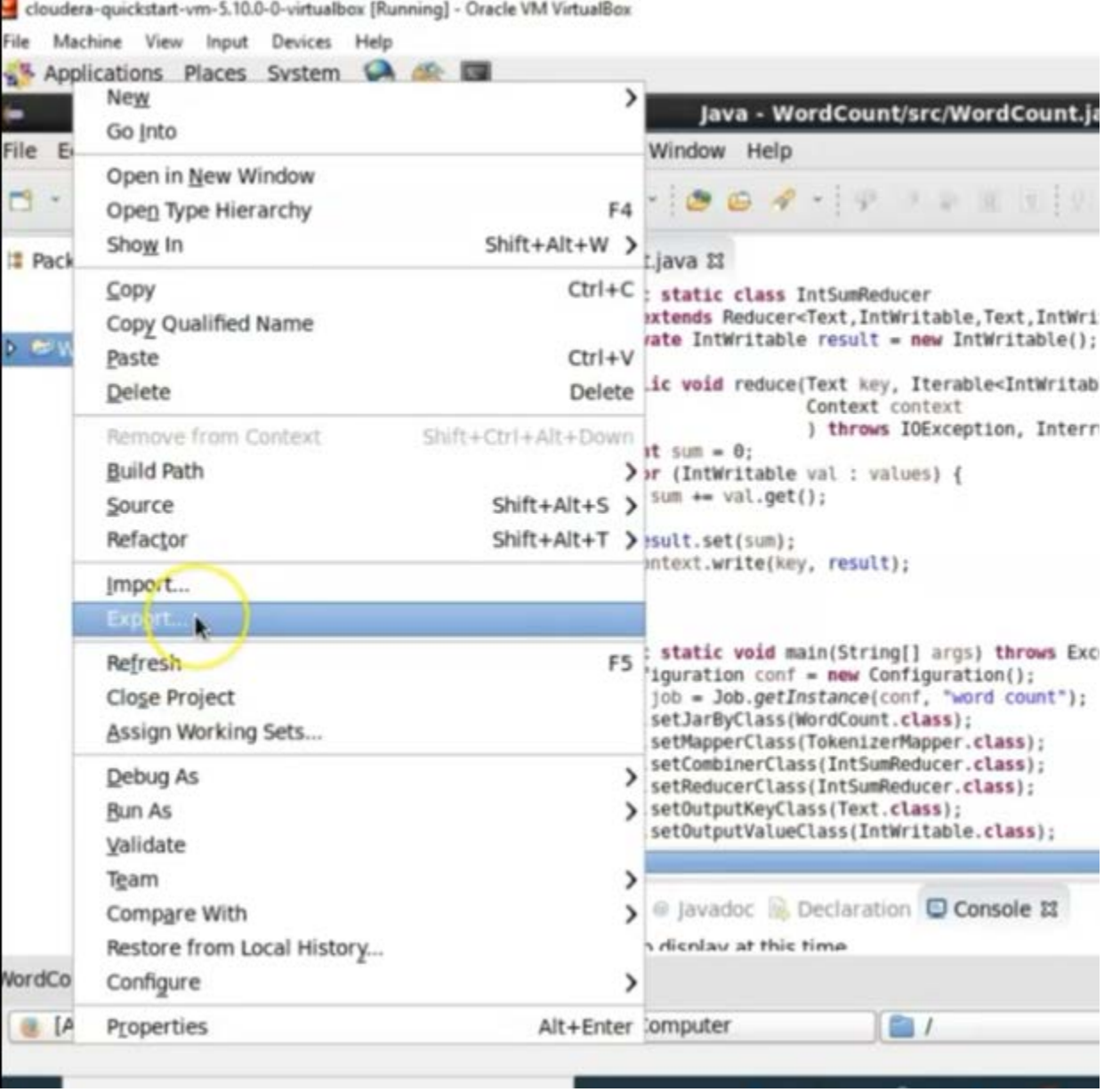
https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html

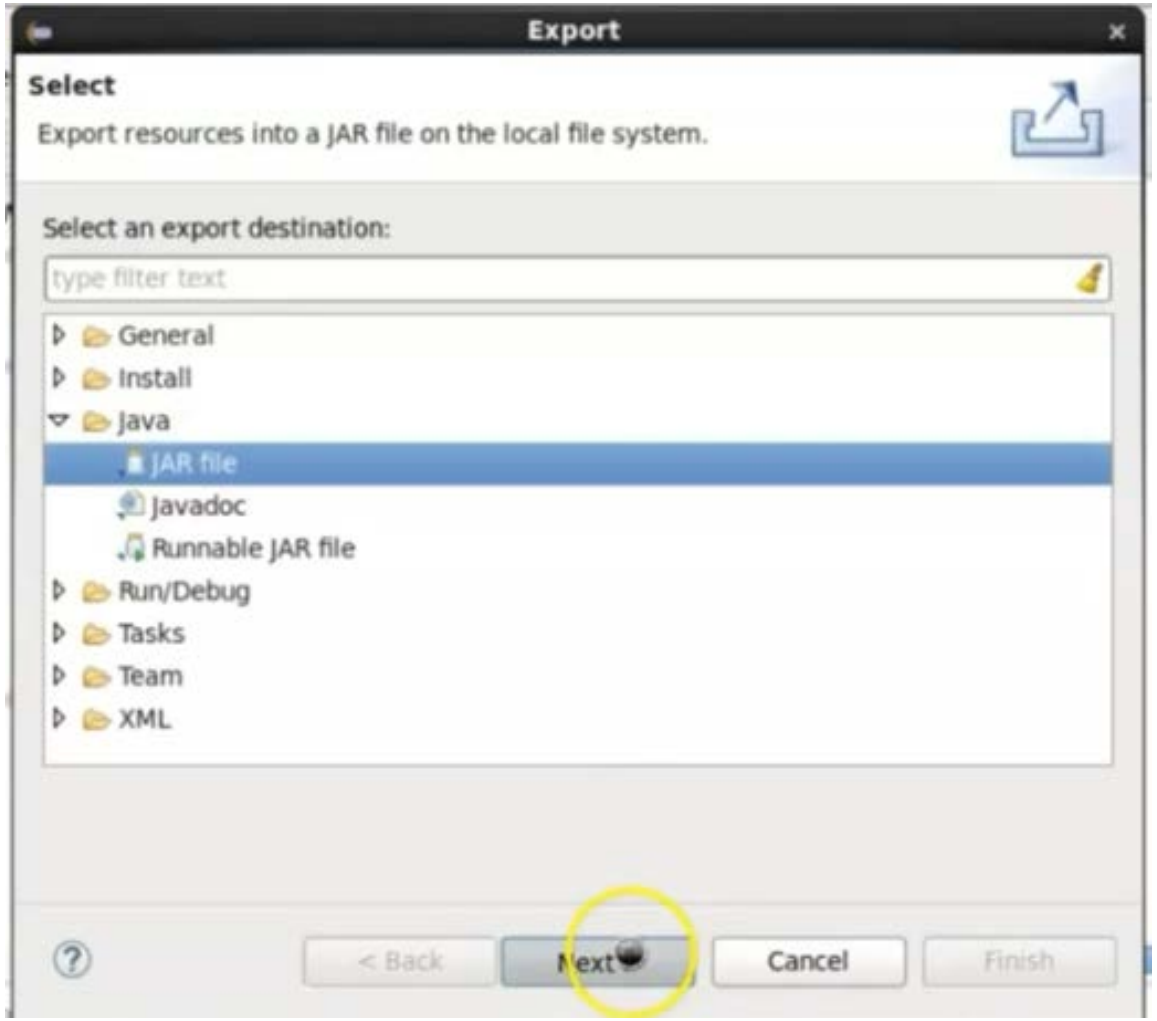
Voir référence

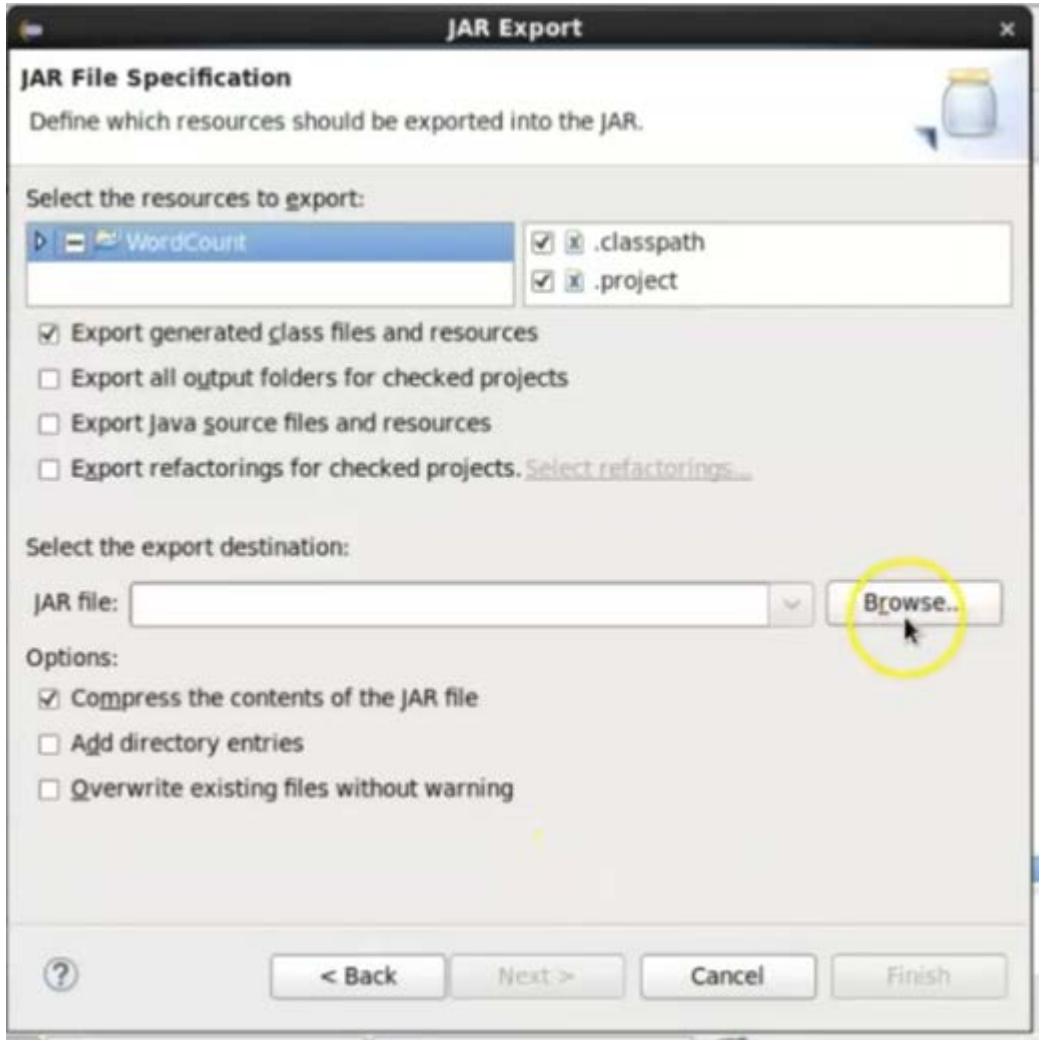

File Edit Source Refactor Navigate Search Project Run Window Help
Package Explorer
WordCount
src
(default package)
WordCount.java
JRE System Library [JavaSE-1.7]
Referenced Libraries
WordCount.java
Context context
) throws IOException, InterruptedException {
int sum = 0;
for (IntWritable val : values) {
sum += val.get();
}
result.set(sum);
context.write(key, result);
}
}
public static void main(String[] args) throws Exception {
Configuration conf = new Configuration();
Job job = Job.getInstance(conf, "word count");
job.setJarByClass(WordCount.class);
job.setMapperClass(TokenizerMapper.class);
job.setCombinerClass(IntSumReducer.class);
job.setReducerClass(IntSumReducer.class);
job.setOutputKeyClass(Text.class);
job.setOutputValueClass(IntWritable.class);
FileInputFormat.addInputPath(job, new Path(args[0]));
FileOutputFormat.setOutputPath(job, new Path(args[1]));
System.exit(job.waitForCompletion(true) ? 0 : 1);
}
}

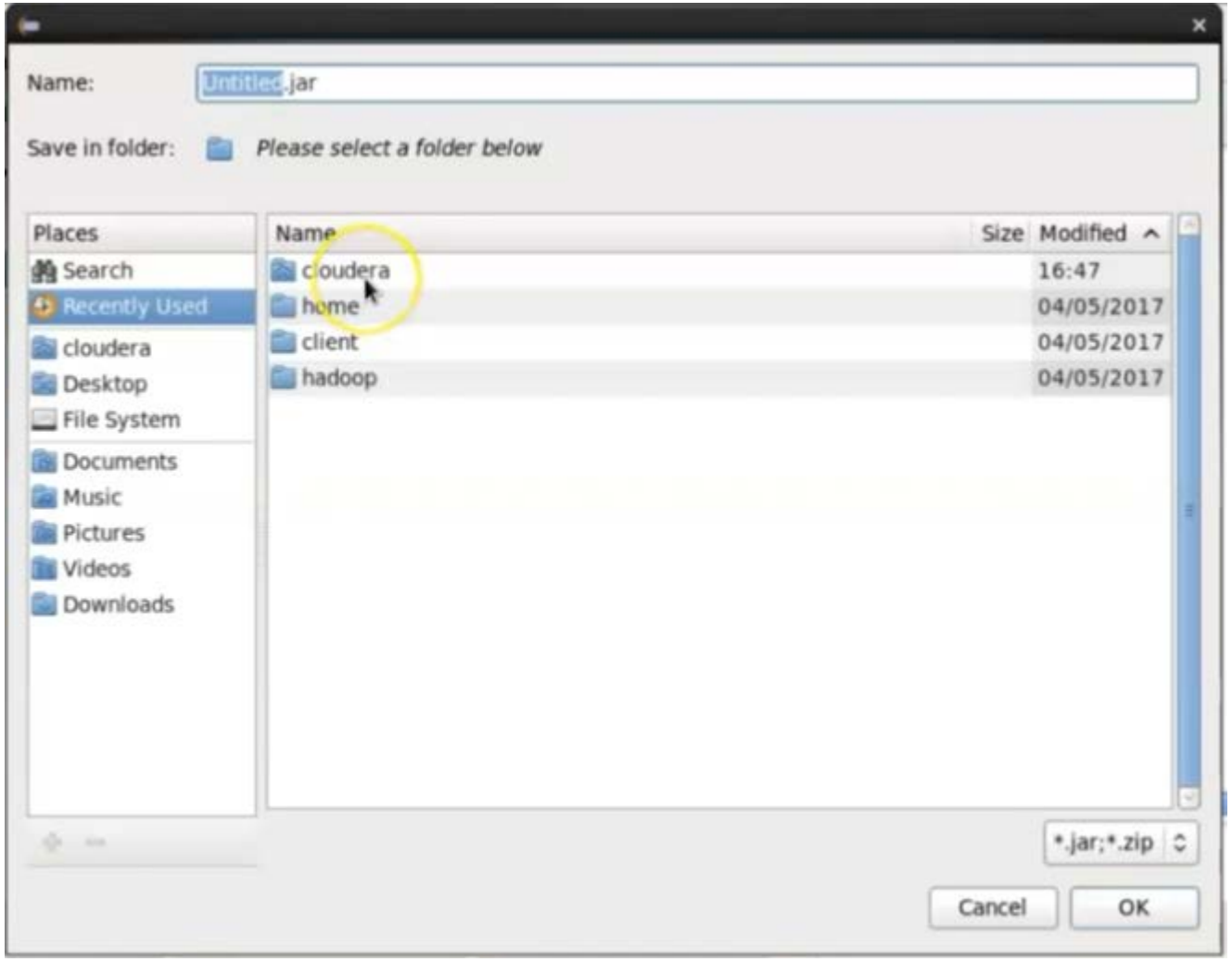

Mon May 4, 5:06 PM cloudera
Apache Hadoop 3.2.1 - MapReduce Tutorial - Mozilla Firefox
Impala
Spark
Solr
Oozie
Cloudera Manager
Getting Started
Building
Tools
Hadoop Streaming
Hadoop Archives
Hadoop Archive Logs
DistCp
GridMix
Rumen
Resource Estimator Service
Scheduler Load Simulator
Hadoop Benchmarking
Reference
Changelog and Release Notes
Java API docs
Unix Shell API
Metrics
Configuration
core-default.xml
hdfs-default.xml
hdfs-rbf-default.xml
mapred-default.xml
yarn-default.xml
Deprecated Properties
import java.io.IOException;
import java.util.StringTokenizer;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
public class WordCount {
public static class TokenizerMapper
extends Mapper<Object, Text, Text, IntWritable>{
private final static IntWritable one = new IntWritable(1);
private Text word = new Text();
public void map(Object key, Text value, Context context
Mozilla Firefox seems slow... to... start.
Learn How to Sp

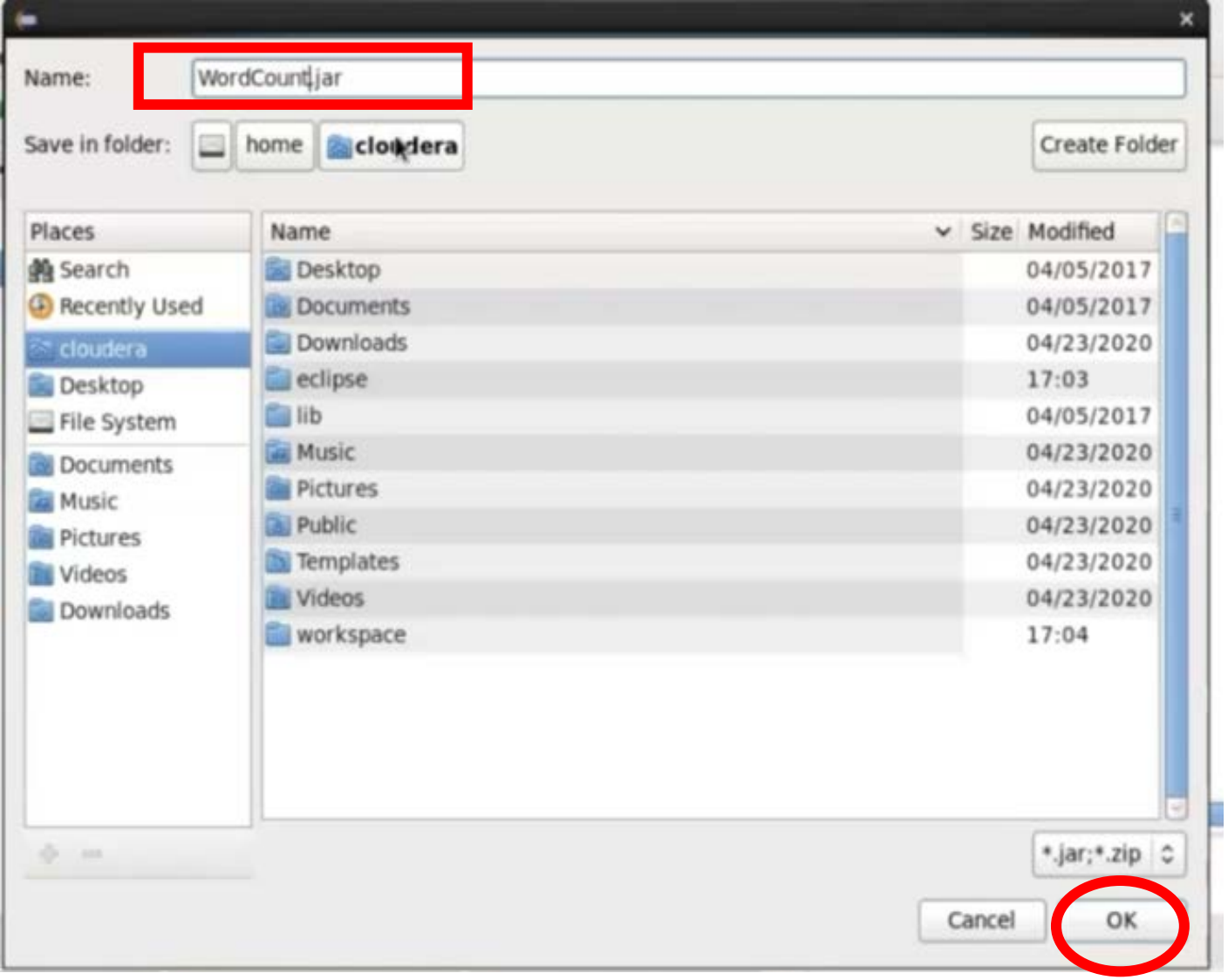

cloudera-quickstart-vm-5.10.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Mon May 4, 5:13 PM
Java - WordCount/src/WordCount.java - Eclipse
File
Window Help
New
Go Into
Open in New Window
Open Type Hierarchy F4
Show In Shift+Alt+W
Copy Ctrl+C
Copy Qualified Name
Paste Ctrl+V
Delete Delete
Remove from Context Shift+Ctrl+Alt+Down
Build Path
Source Shift+Alt+S
Refactor Shift+Alt+T
Import...
Export...
Refresh F5
Close Project
Assign Working Sets...
Debug As
Run As
Validate
Team
Compare With
Restore from Local History...
Configure
Properties Alt+Enter
static class IntSumReducer
extends Reducer<Text,IntWritable,Text,IntWritable> {
IntWritable result = new IntWritable();
void reduce(Text key, Iterable<IntWritable> values,
Context context
) throws IOException, InterruptedException {
sum = 0;
(IntWritable val : values) {
sum += val.get();
result.set(sum);
context.write(key, result);
static void main(String[] args) throws Exception {
Configuration conf = new Configuration();
job = Job.getInstance(conf, "word count");
setJarByClass(WordCount.class);
setMapperClass(TokenizerMapper.class);
setCombinerClass(IntSumReducer.class);
setReducerClass(IntSumReducer.class);
setOutputKeyClass(Text.class);
setOutputValueClass(IntWritable.class);
Javadoc Declaration Console
Computer
/

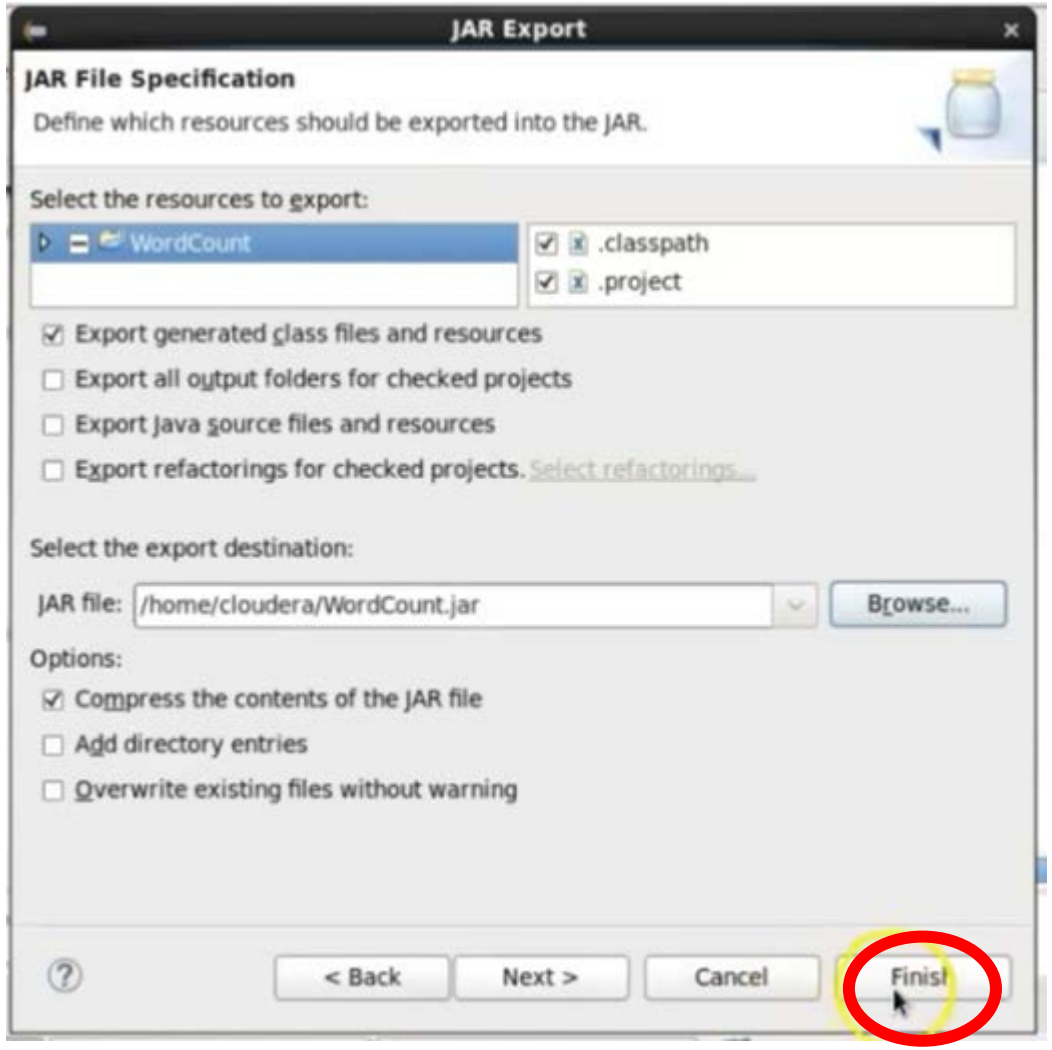

cloudera-quickstart-vm-5.10.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Java - WordCount/src/WordCount.ja
File E
Window Help
New
Go Into
Open in New Window
Open Type Hierarchy F4
Show In Shift+Alt+W
Copy Ctrl+C
Copy Qualified Name
Paste Ctrl+V
Delete Delete
Remove from Context Shift+Ctrl+Alt+Down
Build Path
Source Shift+Alt+S
Refactor Shift+Alt+T
Import...
Export...
Refresh F5
Close Project
Assign Working Sets...
Debug As
Run As
Validate
Team
Compare With
Restore from Local History...
Configure
Properties Alt+Enter
.java
static class IntSumReducer
extends Reducer<Text,IntWritable,Text,IntWri
rate IntWritable result = new IntWritable();
ic void reduce(Text key, Iterable<IntWritab
Context context
) throws IOException, Interr
t sum = 0;
r (IntWritable val : values) {
sum += val.get();
sult.set(sum);
ntext.write(key, result);
static void main(String[] args) throws Exc
iguration conf = new Configuration();
job = Job.getInstance(conf, "word count");
setJarByClass(WordCount.class);
setMapperClass(TokenizerMapper.class);
setCombinerClass(IntSumReducer.class);
setReducerClass(IntSumReducer.class);
setOutputKeyClass(Text.class);
setOutputValueClass(IntWritable.class);
Javadoc Declaration Console
display at this time
Computer
/

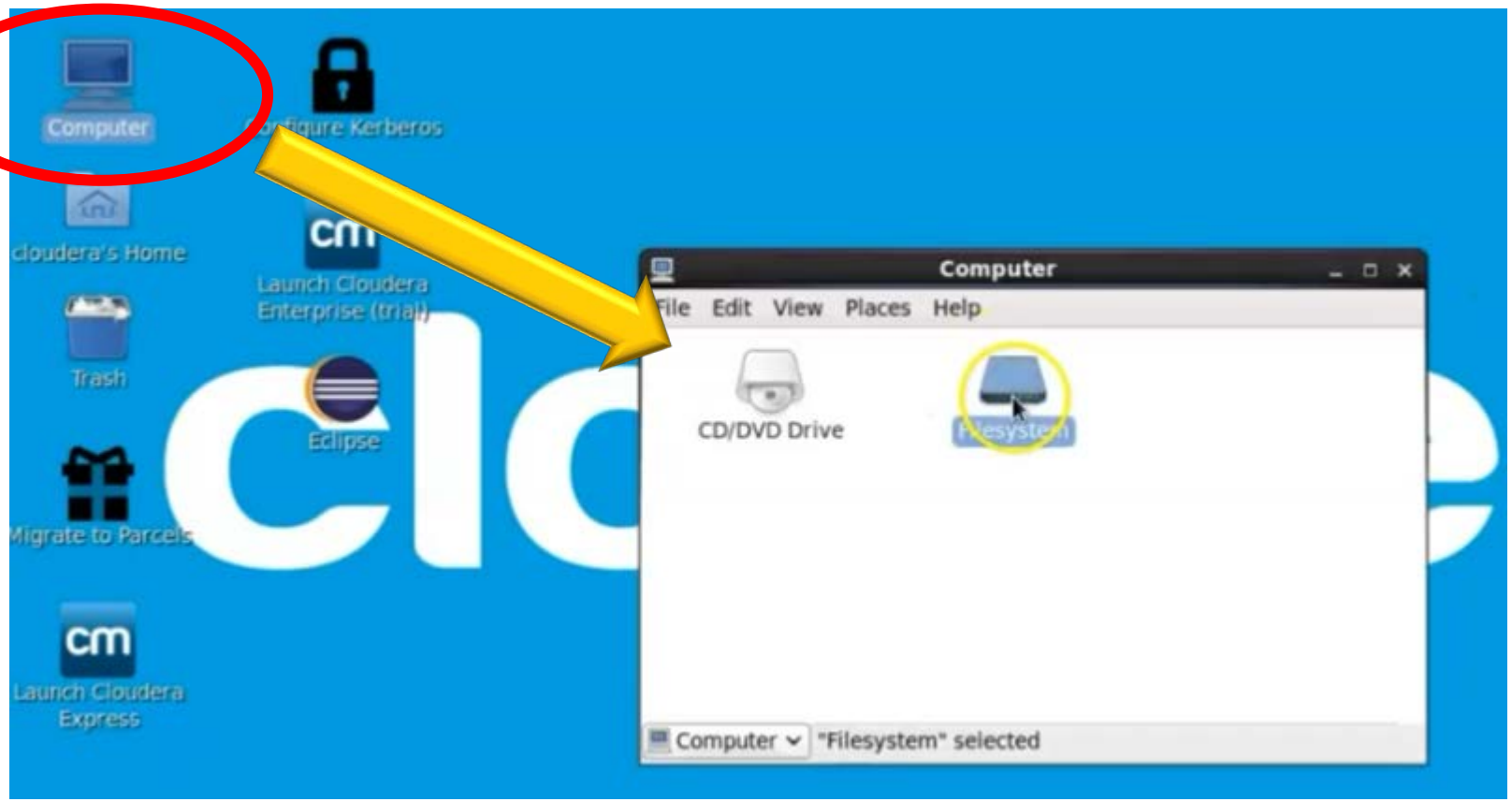

Export
Select
Export resources into a JAR file on the local file system.
Select an export destination:
type filter text
General
Install
Java
JAR file
Javadoc
Runnable JAR file
Run/Debug
Tasks
Team
XML
< Back
Next >
Cancel
Finish

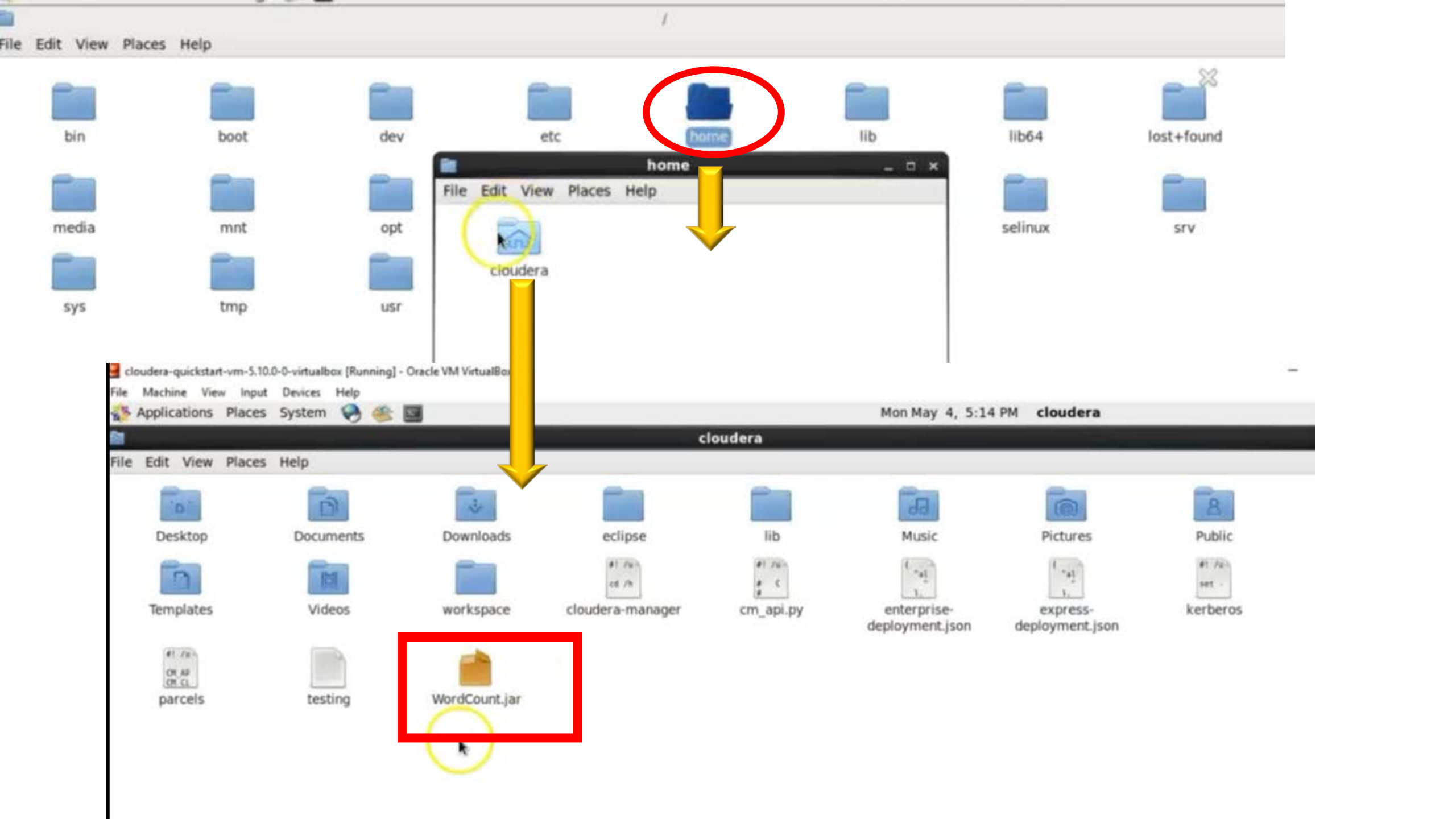

JAR Export
JAR File Specification
Define which resources should be exported into the JAR.
Select the resources to export:
WordCount
.classpath
.project
Export generated class files and resources
Export all output folders for checked projects
Export Java source files and resources
Export refactorings for checked projects. Select refactorings...
Select the export destination:
JAR file:
Browse...
Options:
Compress the contents of the JAR file
Add directory entries
Overwrite existing files without warning
< Back
Next >
Cancel
Finish


Name:
Untitled.jar
Save in folder:
Please select a folder below
Places
Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads
Name
Size
Modified
cloudera
16:47
home
04/05/2017
client
04/05/2017
hadoop
04/05/2017
.jar;.zip
Cancel
OK


Name:
WordCount.jar
Save in folder:
home
cloudera
Create Folder
Places
Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads
Name
Size
Modified
Desktop 04/05/2017
Documents 04/05/2017
Downloads 04/23/2020
eclipse 17:03
lib 04/05/2017
Music 04/23/2020
Pictures 04/23/2020
Public 04/23/2020
Templates 04/23/2020
Videos 04/23/2020
workspace 17:04
.jar;.zip
Cancel
OK


JAR Export
JAR File Specification
Define which resources should be exported into the JAR.
Select the resources to export:
WordCount
.classpath
.project
Export generated class files and resources
Export all output folders for checked projects
Export Java source files and resources
Export refactorings for checked projects. Select refactorings...
Select the export destination:
JAR file: /home/cloudera/WordCount.jar
Browse...
Options:
Compress the contents of the JAR file
Add directory entries
Overwrite existing files without warning
< Back
Next >
Cancel
Finish


Computer
Launch Cloudera Enterprise (trial)
Trash
Eclipse
Launch Cloudera Express
Computer
File Edit View Places Help
CD/DVD Drive
Computer
"Filesystem" selected


File Edit View Places Help
bin
boot
dev
etc
home
lib
lib64
lost+found
media
mnt
opt
selinux
srv
sys
tmp
usr
home
File Edit View Places Help
cloudera
cloudera-quickstart-vm-5.10.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Mon May 4, 5:14 PM cloudera
cloudera
File Edit View Places Help
Desktop
Documents
Downloads
eclipse
lib
Music
Pictures
Public
Templates
Videos
workspace
cloudera-manager
cm_api.py
enterprise-deployment.json
express-deployment.json
kerberos
parcels
testing
WordCount.jar

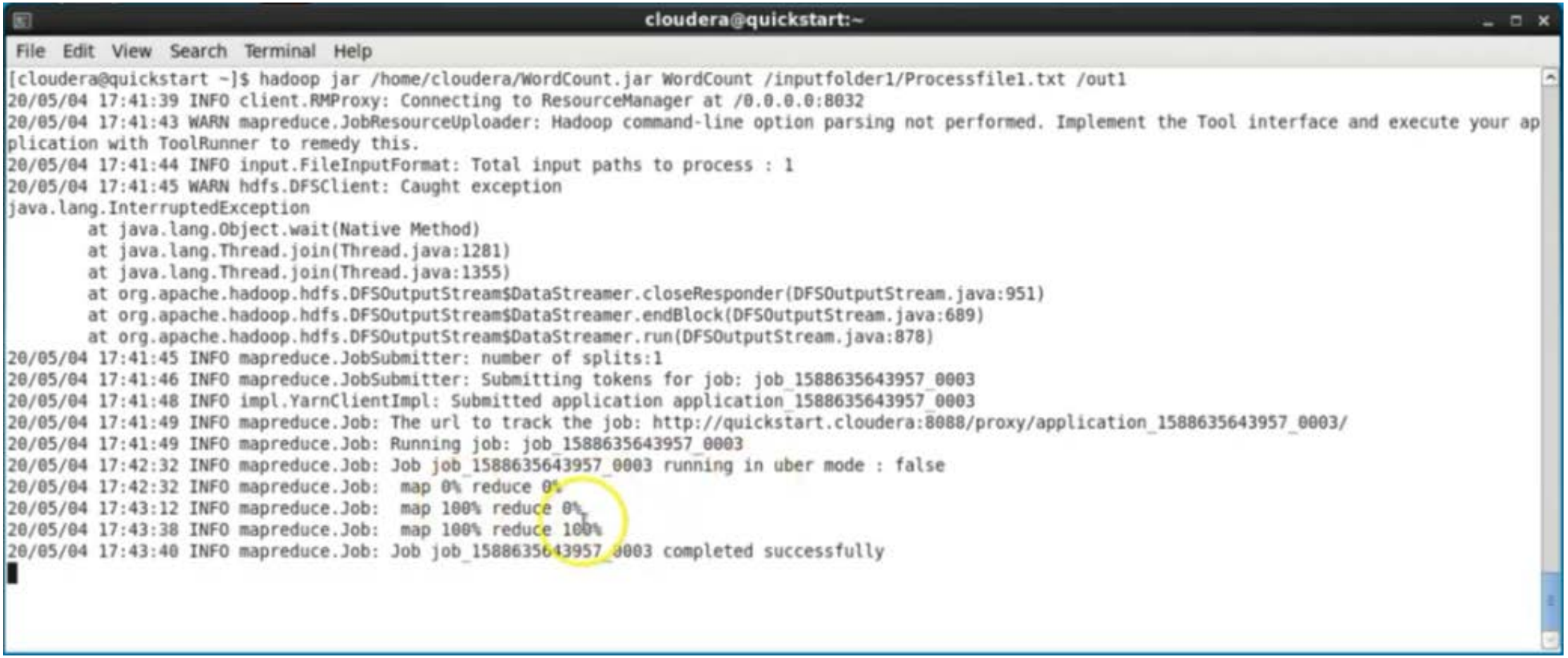
cloudera@quickstart:~

File Edit View Search Terminal Help

```
[cloudera@quickstart ~]$ ls
cloudera-manager  eclipse                      Music      testing
cm_api.py         enterprise-deployment.json   parcels    Videos
Desktop           express-deployment.json      Pictures   WordCount.jar
Documents         kerberos                     Public     workspace
Downloads         lib                          Templates
[cloudera@quickstart ~]$ pwd
/home/cloudera
[cloudera@quickstart ~]$ cat > /home/cloudera/Processfile1.txt
Riyaz Ahamed
Riyaz Ahamed
Riyaz Ahamed
Cloudera
Cloudera
Cloudera
Big Data
Big Data
Big Data
Cloudera
Riyaz Ahamed
Riyaz Ahamed
^Z
[1]+  Stopped                 cat > /home/cloudera/Processfile1.txt
[cloudera@quickstart ~]$
```

```
[cloudera@quickstart ~]$ hdfs dfs -ls /
Found 10 items
drwxrwxrwx   - hdfs     supergroup          0 2017-04-05 04:27 /benchmarks
drwxr-xr-x   - hbase    supergroup          0 2020-05-04 16:52 /hbase
drwxr-xr-x   - cloudera supergroup          0 2020-05-01 02:57 /inputfolder
drwxr-xr-x   - cloudera supergroup          0 2020-05-01 00:09 /inputnew
drwxr-xr-x   - cloudera supergroup          0 2020-05-01 03:03 /output1
drwxr-xr-x   - cloudera supergroup          0 2020-05-01 00:19 /output_new
drwxr-xr-x   - solr     solr                0 2017-04-05 04:29 /solr
drwxrwxrwt   - hdfs     supergroup          0 2020-04-23 22:28 /tmp
drwxr-xr-x   - hdfs     supergroup          0 2017-04-05 04:29 /user
drwxr-xr-x   - hdfs     supergroup          0 2017-04-05 04:29 /var
[cloudera@quickstart ~]$
```

```
[cloudera@quickstart ~]$ hdfs dfs -mkdir /inputfolder1
[cloudera@quickstart ~]$ hdfs dfs -put /home/cloudera/Processfile1.txt /inputfolder1/
[cloudera@quickstart ~]$ hdfs dfs -cat /inputfolder1/Processfile1.txt
Riyaz Ahamed
Riyaz Ahamed
Riyaz Ahamed
Cloudera
Cloudera
Cloudera
Big Data
Big Data
Big Data
Cloudera
Riyaz Ahamed
Riyaz Ahamed
[cloudera@quickstart ~]$
```


cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hadoop jar /home/cloudera/WordCount.jar WordCount /inputfolder1/Processfile1.txt /out1
20/05/04 17:41:39 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
20/05/04 17:41:43 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
20/05/04 17:41:44 INFO input.FileInputFormat: Total input paths to process : 1
20/05/04 17:41:45 WARN hdfs.DFSClient: Caught exception
java.lang.InterruptedException
at java.lang.Object.wait(Native Method)
at java.lang.Thread.join(Thread.java:1281)
at java.lang.Thread.join(Thread.java:1355)
at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.closeResponder(DFSOutputStream.java:951)
at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.endBlock(DFSOutputStream.java:689)
at org.apache.hadoop.hdfs.DFSOutputStream$DataStreamer.run(DFSOutputStream.java:878)
20/05/04 17:41:45 INFO mapreduce.JobSubmitter: number of splits:1
20/05/04 17:41:46 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1588635643957_0003
20/05/04 17:41:48 INFO impl.YarnClientImpl: Submitted application application_1588635643957_0003
20/05/04 17:41:49 INFO mapreduce.Job: The url to track the job: http://quickstart.cloudera:8088/proxy/application_1588635643957_0003/
20/05/04 17:41:49 INFO mapreduce.Job: Running job: job_1588635643957_0003
20/05/04 17:42:32 INFO mapreduce.Job: Job job_1588635643957_0003 running in uber mode : false
20/05/04 17:42:32 INFO mapreduce.Job: map 0% reduce 0%
20/05/04 17:43:12 INFO mapreduce.Job: map 100% reduce 0%
20/05/04 17:43:38 INFO mapreduce.Job: map 100% reduce 100%
20/05/04 17:43:40 INFO mapreduce.Job: Job job_1588635643957_0003 completed successfully

cloudera@quickstart:~

File Edit View Search Terminal Help

```
20/05/04 17:43:41 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=67
                FILE: Number of bytes written=242435
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=254
                HDFS: Number of bytes written=41
                HDFS: Number of read operations=6
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
        Job Counters
                Launched map tasks=1
                Launched reduce tasks=1
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=36101
                Total time spent by all reduces in occupied slots (ms)=23876
                Total time spent by all map tasks (ms)=36101
                Total time spent by all reduce tasks (ms)=23876
                Total vcore-seconds taken by all map tasks=36101
                Total vcore-seconds taken by all reduce tasks=23876
                Total megabyte-seconds taken by all map tasks=36967424
                Total megabyte-seconds taken by all reduce tasks=24449024
        Map-Reduce Framework
                Map input records=12
                Map output records=20
```

cloudera@quickstart:~

File Edit View Search Terminal Help

```
                Total vcore-seconds taken by all reduce tasks=23876
                Total megabyte-seconds taken by all map tasks=36967424
                Total megabyte-seconds taken by all reduce tasks=24449024
        Map-Reduce Framework
                Map input records=12
                Map output records=20
                Map output bytes=208
                Map output materialized bytes=67
                Input split bytes=126
                Combine input records=20
                Combine output records=5
                Reduce input groups=5
                Reduce shuffle bytes=67
                Reduce input records=5
                Reduce output records=5
                Spilled Records=10
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=945
                CPU time spent (ms)=4550
                Physical memory (bytes) snapshot=337055744
                Virtual memory (bytes) snapshot=3008081920
                Total committed heap usage (bytes)=226365440
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
```

```
[cloudera@quickstart ~]$ hdfs dfs -ls /out1
Found 2 items
-rw-r--r--   1 cloudera supergroup          0 2020-05-04 17:43 /out1/_SUCCESS
-rw-r--r--   1 cloudera supergroup         41 2020-05-04 17:43 /out1/part-r-00000
[cloudera@quickstart ~]$
```

```
[cloudera@quickstart ~]$ hdfs dfs -cat /out1/part-r-00000
Ahamed  5
Big     3
Cloudera        4
Data    3
Riyaz   5
[cloudera@quickstart ~]$
```

```
[cloudera@quickstart ~]$ cat /home/cloudera/Processfile1.txt
Riyaz Ahamed
Riyaz Ahamed
Riyaz Ahamed
Cloudera
Cloudera
Cloudera
Big Data
Big Data
Big Data
Cloudera
Riyaz Ahamed
Riyaz Ahamed
[cloudera@quickstart ~]$
```

Référence

https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html