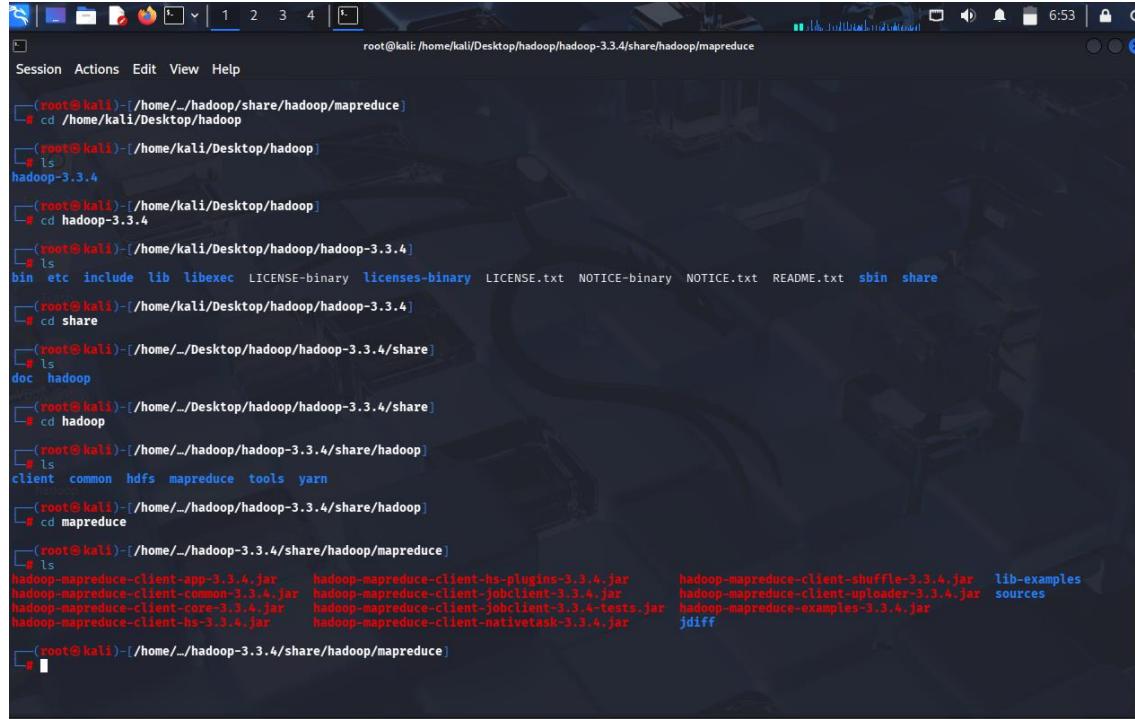


Practica Hadoop

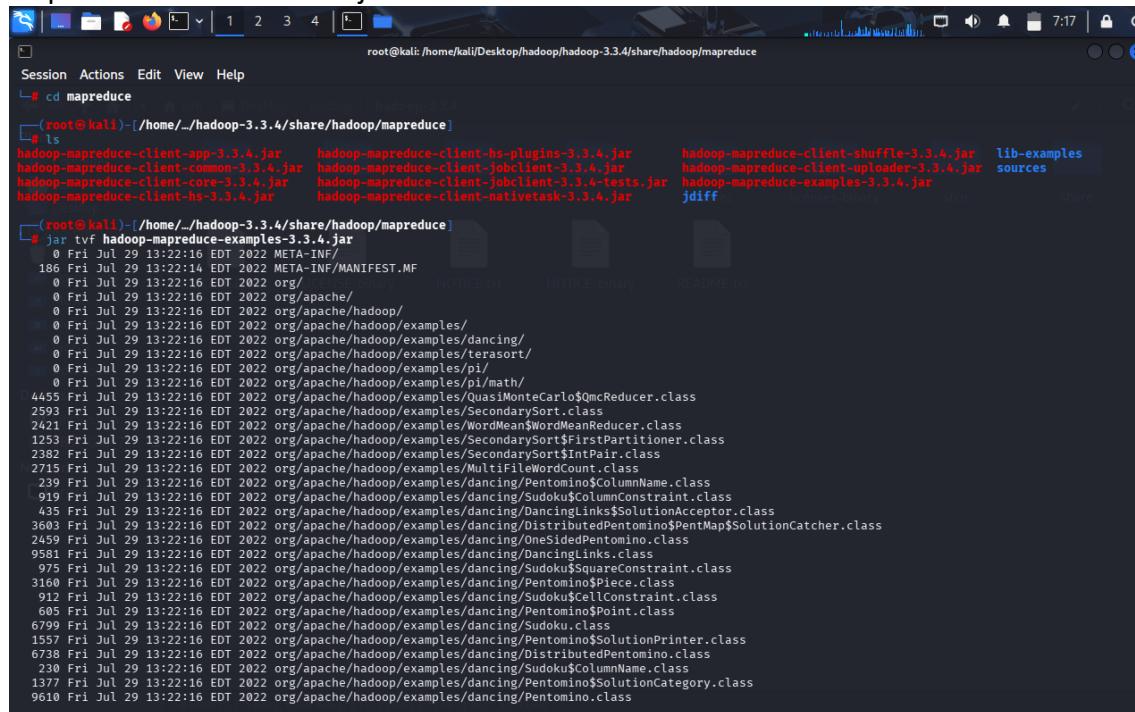
1. Vamos al directorio de hadoop



```
root@kali: /home/_/hadoop/share/hadoop/mapreduce
# cd /home/kali/Desktop/hadoop
# ls
hadoop-3.3.4
# cd hadoop-3.3.4
# ls
bin etc include lib libexec LICENSE-binary licenses-binary LICENSE.txt NOTICE-binary NOTICE.txt README.txt sbin share
# cd share
# ls
doc hadoop
# cd hadoop
# ls
client common hdfs mapreduce tools yarn
# cd mapreduce
# ls
hadoop-mapreduce-client-app-3.3.4.jar hadoop-mapreduce-client-hs-plugins-3.3.4.jar hadoop-mapreduce-client-shuffle-3.3.4.jar lib-examples
hadoop-mapreduce-client-common-3.3.4.jar hadoop-mapreduce-client-jobclient-3.3.4.jar hadoop-mapreduce-client-uploader-3.3.4.jar sources
hadoop-mapreduce-client-core-3.3.4.jar hadoop-mapreduce-client-jobclient-3.3.4-tests.jar hadoop-mapreduce-examples-3.3.4.jar
hadoop-mapreduce-client-hs-3.3.4.jar hadoop-mapreduce-client-nativetask-3.3.4.jar jdiff

#
```

Inspeccionamos el archivo jar.



```
root@kali: /home/_/hadoop/share/hadoop/mapreduce
# cd mapreduce
# ls
hadoop-mapreduce-client-app-3.3.4.jar hadoop-mapreduce-client-hs-plugins-3.3.4.jar hadoop-mapreduce-client-shuffle-3.3.4.jar lib-examples
hadoop-mapreduce-client-common-3.3.4.jar hadoop-mapreduce-client-jobclient-3.3.4.jar hadoop-mapreduce-client-uploader-3.3.4.jar sources
hadoop-mapreduce-client-core-3.3.4.jar hadoop-mapreduce-client-jobclient-3.3.4-tests.jar hadoop-mapreduce-examples-3.3.4.jar
hadoop-mapreduce-client-hs-3.3.4.jar hadoop-mapreduce-client-nativetask-3.3.4.jar jdiff

# jar tf hadoop-mapreduce-examples-3.3.4.jar
0 Fri Jul 29 13:22:16 EDT 2022 META-INF/
186 Fri Jul 29 13:22:14 EDT 2022 META-INF/MANIFEST.MF
0 Fri Jul 29 13:22:16 EDT 2022 org/
0 Fri Jul 29 13:22:16 EDT 2022 org/apache/
0 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/
0 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/
0 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/
0 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/terasort/
0 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/pi/
0 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/pi/math/
4455 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/QuasiMonteCarlo$QmcReducer.class
2593 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/SecondarySort.class
2421 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/WordMean$WordMeanReducer.class
1253 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/SecondarySort$FirstPartitioner.class
2382 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/SecondarySort$IntPair.class
2715 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/MultifileWordCount.class
239 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Pentomino$ColumnName.class
919 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Sudoku$ColumnConstraint.class
435 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/DancingLinks$SolutionAcceptor.class
3603 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/DistributedPentomino$PentMap$SolutionCatcher.class
2459 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/OnesidedPentomino.class
9581 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/DancingLinks.class
975 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Sudoku$SquareConstraint.class
3160 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Pentomino$Piece.class
912 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Sudoku$CellConstraint.class
605 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Pentomino$Point.class
6799 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Sudoku.class
1557 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Pentomino$SolutionPrinter.class
6738 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/DistributedPentomino.class
230 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Sudoku$ColumnName.class
1377 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Pentomino$SolutionCategory.class
9610 Fri Jul 29 13:22:16 EDT 2022 org/apache/hadoop/examples/dancing/Pentomino.class
```

2. Preparamos el directorio de entrada local y primera ejecución

```
root@kali: /home/kali/Desktop/hadoop/hadoop-3.3.4/share/hadoop/mapreduce
Session Actions Edit View Help
[root@kali] ~]
# cd home/kali/Desktop/hadoop/hadoop-3.3.4/
cd: no such file or directory: home/kali/Desktop/hadoop/hadoop-3.3.4

[root@kali] ~]
# cd home/kali/Desktop/hadoop/hadoop-3.3.4
cd: no such file or directory: home/kali/Desktop/hadoop/hadoop-3.3.4

[root@kali] ~]
# cd /home/kali/Desktop/hadoop/hadoop-3.3.4/
[root@kali] /home/kali/Desktop/hadoop/hadoop-3.3.4]
# ls
bin etc include lib libexec LICENSE-binary licenses-binary LICENSE.txt NOTICE-binary NOTICE.txt README.txt sbin share
[root@kali] /home/kali/Desktop/hadoop/hadoop-3.3.4]
# ls
bin etc include lib libexec LICENSE-binary licenses-binary LICENSE.txt NOTICE-binary NOTICE.txt README.txt sbin share

[root@kali] /home/kali/Desktop/hadoop/hadoop-3.3.4]
# cd /home/kali/Desktop/hadoop/hadoop-3.3.4/share/hadoop/mapreduce
[root@kali] /home/.../hadoop-3.3.4/share/hadoop/mapreduce]
# mkdir /tmp/entrada
mkdir: cannot create directory '/tmp/entrada': File exists

[root@kali] /home/.../hadoop-3.3.4/share/hadoop/mapreduce]
# cp /home/kali/Desktop/hadoop/etc/hadoop/*.xml /tmp/entrada
cp: cannot stat '/home/kali/Desktop/hadoop/etc/hadoop/*.xml': No such file or directory

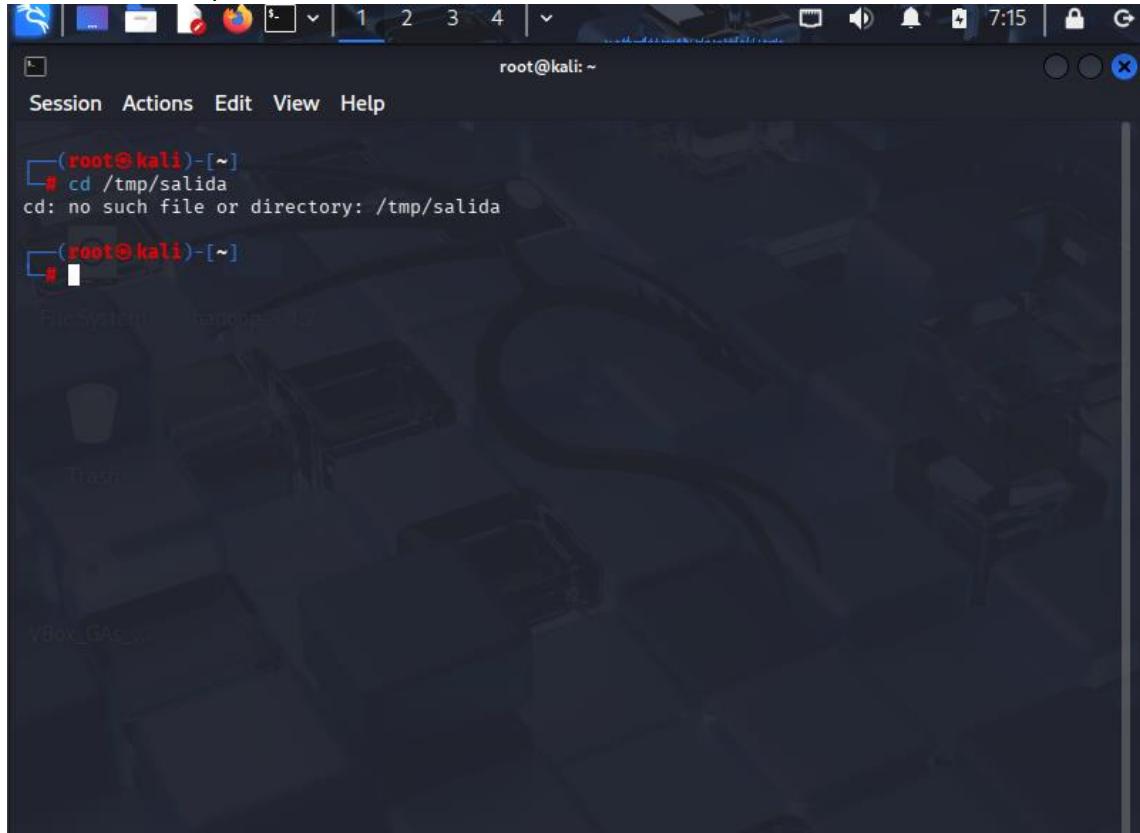
[root@kali] /home/.../hadoop-3.3.4/share/hadoop/mapreduce]
# cp /home/kali/Desktop/hadoop/hadoop-3.3.4/etc/hadoop/*.xml /tmp/entrada
[root@kali] /home/.../hadoop-3.3.4/share/hadoop/mapreduce]
# ls /tmp/entrada
capacity-scheduler.xml hadoop-policy.xml hdfs-site.xml kms-acls.xml mapred-site.xml
core-site.xml hdfs-rbf-site.xml httpfs-site.xml kms-site.xml yarn-site.xml
yarn-env.cmd yarn-env.sh yarnservice.log4j.properties
yarn-site.xml

[root@kali] /home/.../hadoop-3.3.4/share/hadoop/mapreduce]
#
```

```
root@kali: ~
Session Actions Edit View Help
[root@kali] ~]
# hadoop jar hadoop-mapreduce-examples-3.4.2.jar grep /tmp/entrada /tmp/salida 'dfs[a-z.]+'
JAR does not exist or is not a normal file: /root/hadoop-mapreduce-examples-3.4.2.jar

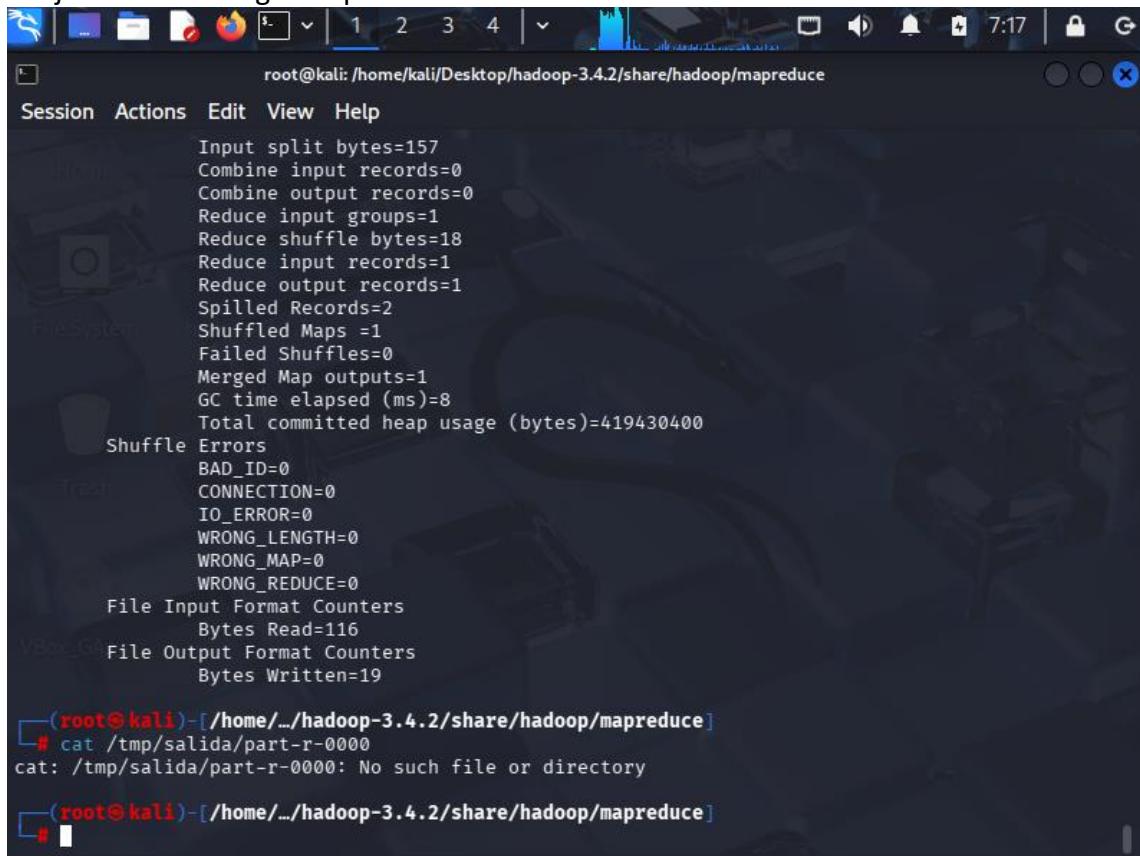
[root@kali] ~]
#
```

3. Miramos /tmp/salida



```
root@kali: ~
Session Actions Edit View Help
└─(root@kali)-[~]
# cd /tmp/salida
cd: no such file or directory: /tmp/salida.
```

4. Ejecutamos la segunda prueba



```
root@kali: /home/kali/Desktop/hadoop-3.4.2/share/hadoop/mapreduce
Session Actions Edit View Help
Input split bytes=157
Combine input records=0
Combine output records=0
Reduce input groups=1
Reduce shuffle bytes=18
Reduce input records=1
Reduce output records=1
Spilled Records=2
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=8
Total committed heap usage (bytes)=419430400
Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=116
File Output Format Counters
  Bytes Written=19

└─(root@kali)-[/home/.../hadoop-3.4.2/share/hadoop/mapreduce]
# cat /tmp/salida/part-r-0000
cat: /tmp/salida/part-r-0000: No such file or directory

└─(root@kali)-[/home/.../hadoop-3.4.2/share/hadoop/mapreduce]
#
```