

MANUAL INSTALACIÓN DE PYTHON Y DJANGO

Nasly Yisedt Bermúdez Tarazona

Ernesto Alonso Cortés Prada

Facultad de Ciencias Naturales e Ingenierías

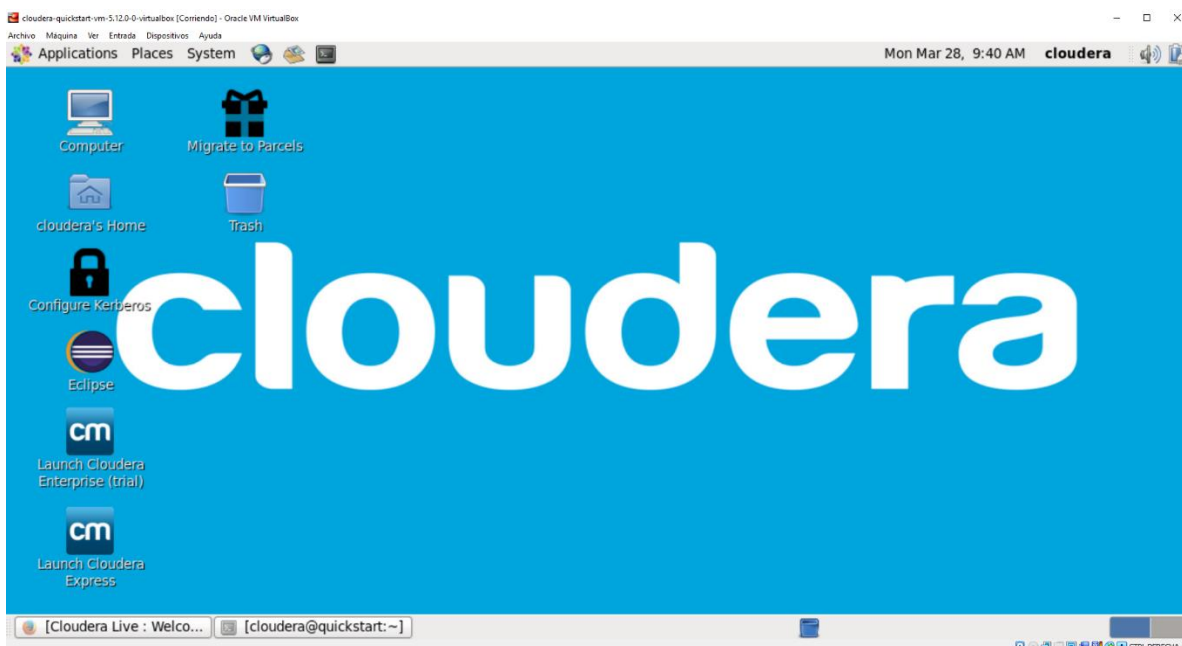
E198 Big Data

Ing. Eliecer Montero Ojeda

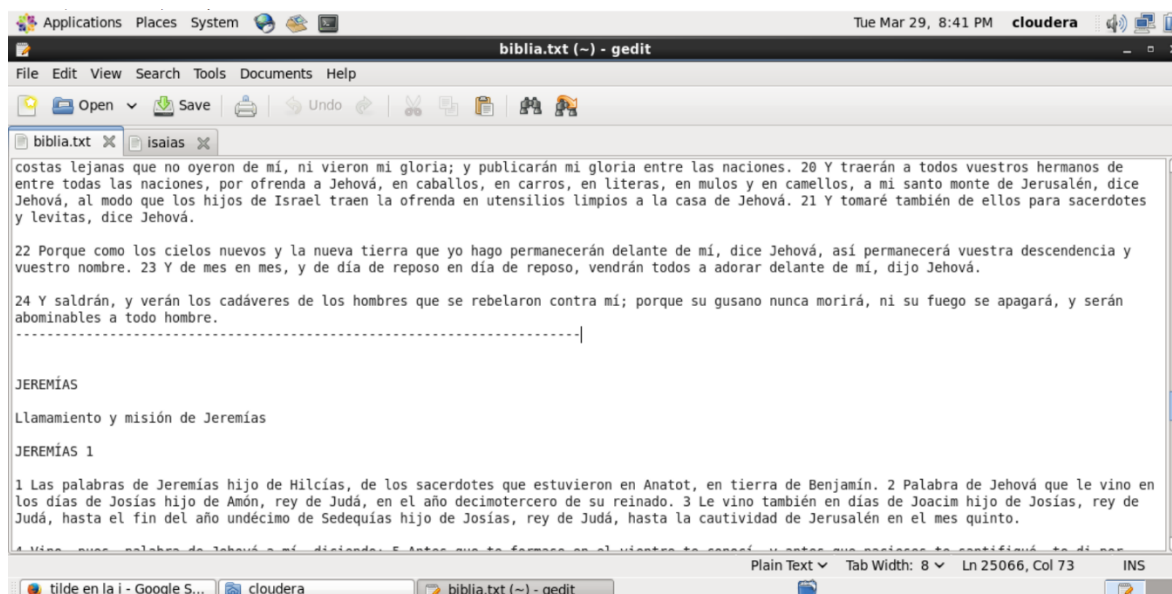
29 de Marzo de 2022

Pasos a seguir para desarrollar Map-reduce con Hadoop en el sistema Linux Cloudera

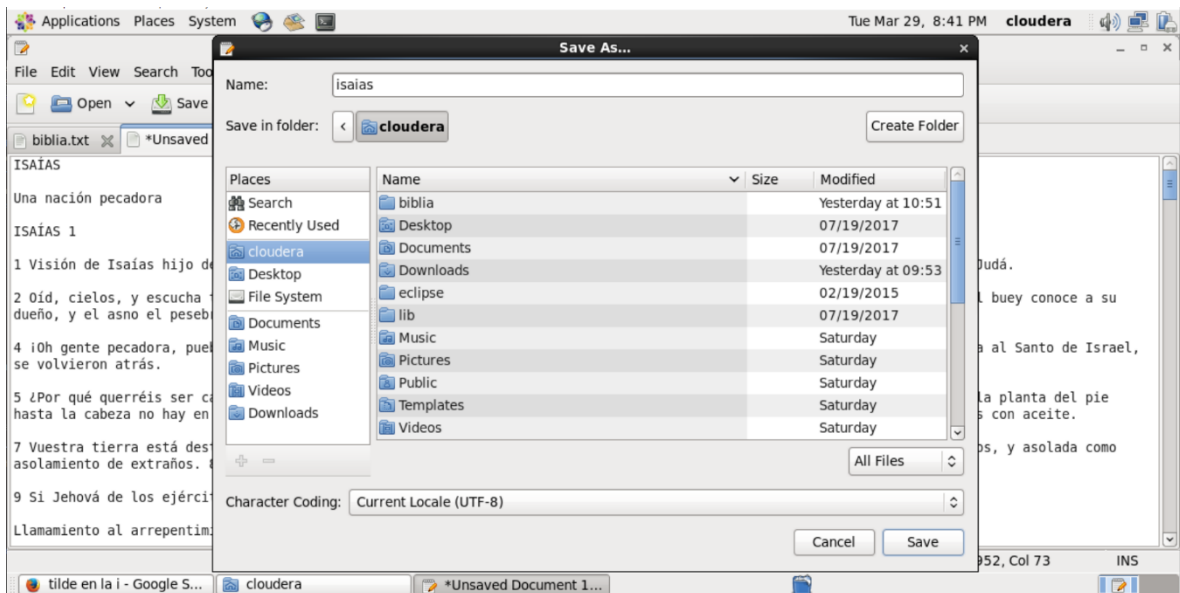
1. Se procede a descargar la imagen ISO del sistema operativo CLOUERA, el cual ofrece un mejor manejo en el ambiente Hadoop y el análisis de datos. Se importa e instala la imagen ISO en el programa virtual box para simular el sistema.



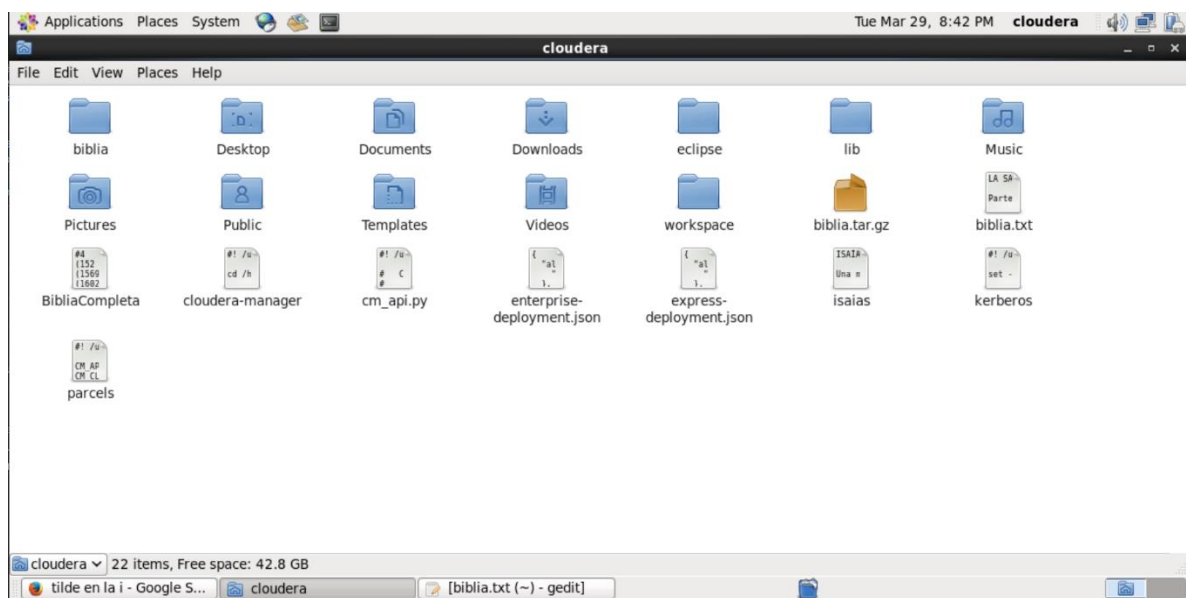
2. Se descarga la biblia en txt.



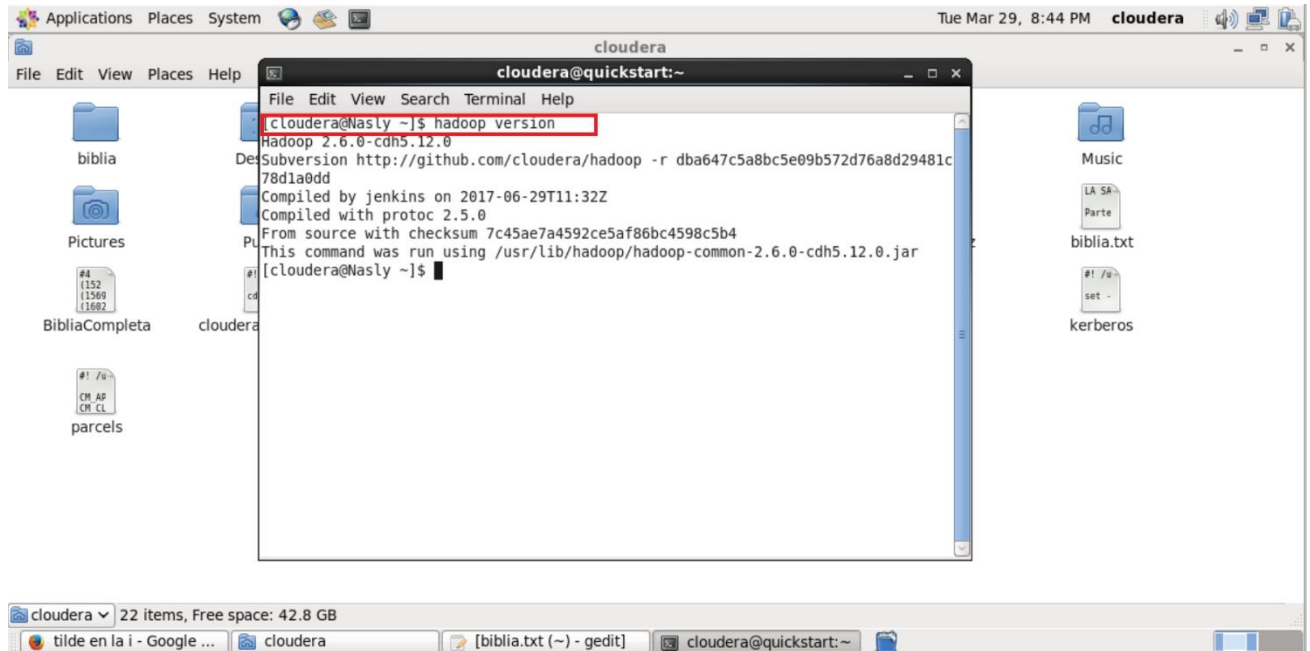
3. Se guarda el libro de la biblia solicitado Isaías en la carpeta cloudera



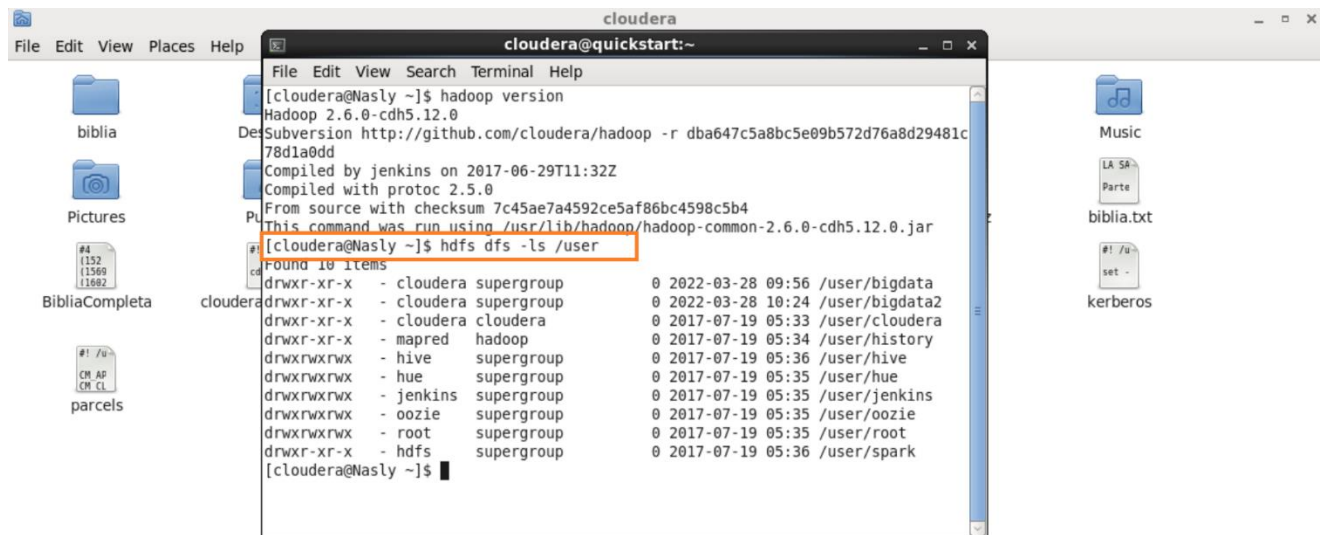
4. Se verifica el documento guardado



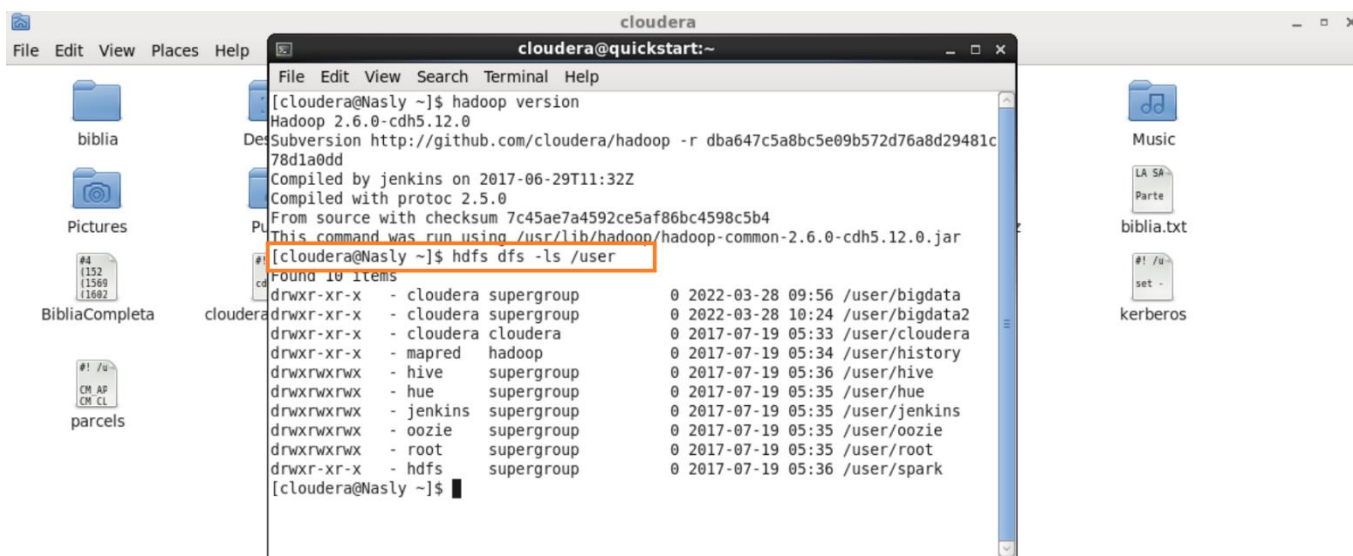
5. En la terminal de Cloudera se verifica



6. Se lista user en donde se va a trabajar Hadoop



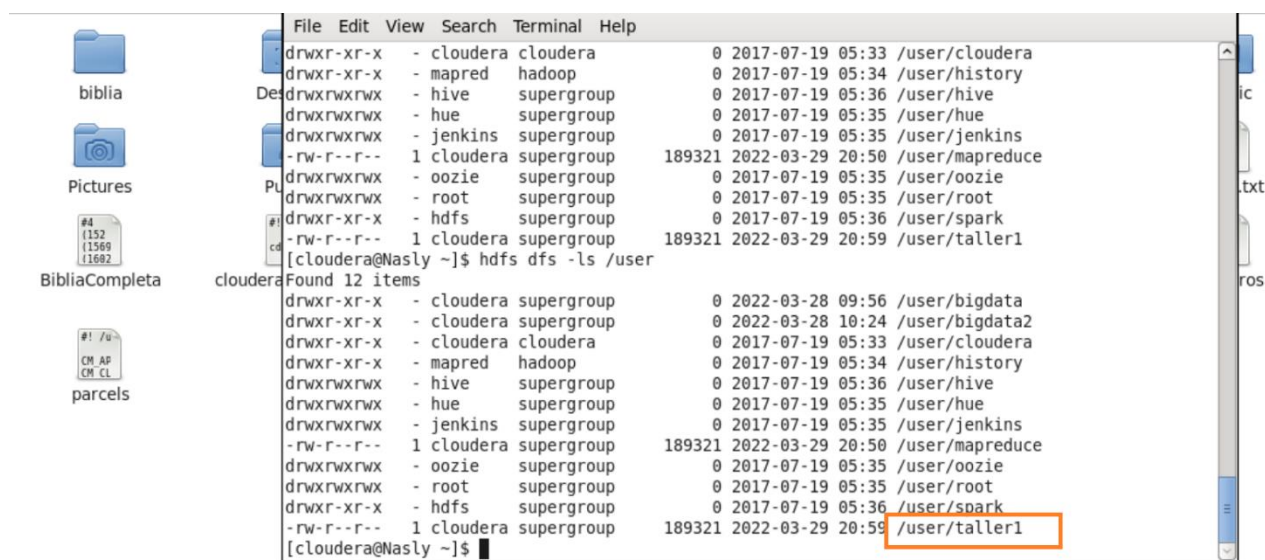
7. Se procede a crear la carpeta taller 1 donde se va a guardar el archivo que se trabajar con map reduce.



The screenshot shows a terminal window titled 'cloudera@quickstart:~'. The user has run 'hadoop version' and 'hdfs dfs -ls /user'. The output of 'hdfs dfs -ls /user' is as follows:

```
[cloudera@Nasly ~]$ hdfs dfs -ls /user
Found 10 items
drwxr-xr-x - cloudera supergroup          0 2022-03-28 09:56 /user/bigdata
drwxr-xr-x - cloudera supergroup          0 2022-03-28 10:24 /user/bigdata2
drwxr-xr-x - cloudera cloudera            0 2017-07-19 05:33 /user/cloudera
drwxr-xr-x - mapred hadoop                 0 2017-07-19 05:34 /user/history
drwxrwxrwx - hive supergroup              0 2017-07-19 05:36 /user/hive
drwxrwxrwx - hue supergroup               0 2017-07-19 05:35 /user/hue
drwxrwxrwx - jenkins supergroup           0 2017-07-19 05:35 /user/jenkins
drwxrwxrwx - oozie supergroup             0 2017-07-19 05:35 /user/oozie
drwxrwxrwx - root supergroup              0 2017-07-19 05:35 /user/root
drwxr-xr-x - hdfs supergroup              0 2017-07-19 05:36 /user/spark
```

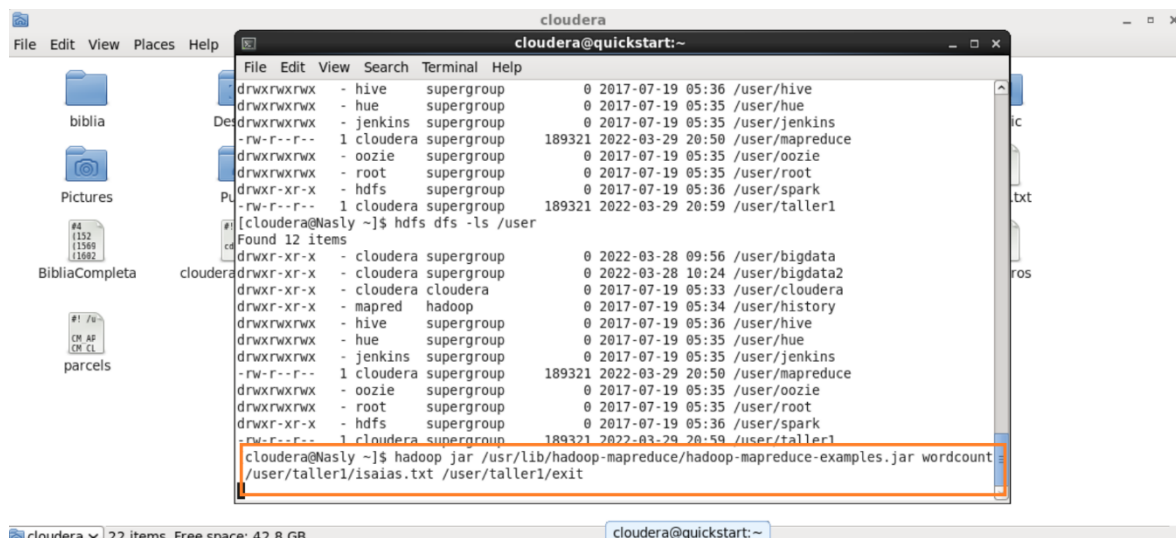
8. Se verifica en la carpeta User que se creara taller 1



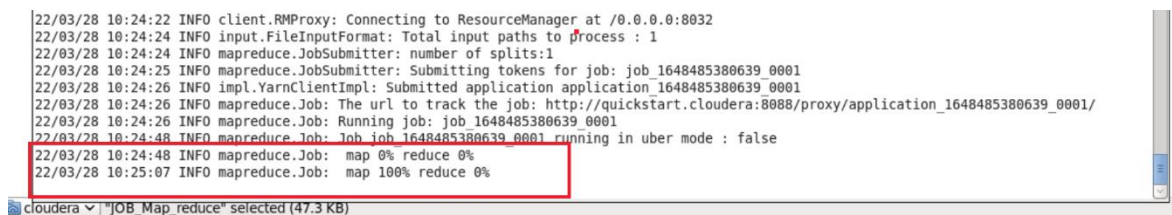
The screenshot shows a terminal window titled 'cloudera@quickstart:~'. The user has run 'hdfs dfs -ls /user'. The output of 'hdfs dfs -ls /user' is as follows:

```
[cloudera@Nasly ~]$ hdfs dfs -ls /user
Found 12 items
drwxr-xr-x - cloudera cloudera            0 2017-07-19 05:33 /user/cloudera
drwxr-xr-x - mapred hadoop                 0 2017-07-19 05:34 /user/history
drwxrwxrwx - hive supergroup              0 2017-07-19 05:36 /user/hive
drwxrwxrwx - hue supergroup               0 2017-07-19 05:35 /user/hue
drwxrwxrwx - jenkins supergroup           0 2017-07-19 05:35 /user/jenkins
-rw-r--r-- 1 cloudera supergroup          189321 2022-03-29 20:50 /user/mapreduce
drwxrwxrwx - oozie supergroup             0 2017-07-19 05:35 /user/oozie
drwxrwxrwx - root supergroup              0 2017-07-19 05:35 /user/root
drwxr-xr-x - hdfs supergroup              0 2017-07-19 05:36 /user/spark
-rw-r--r-- 1 cloudera supergroup          189321 2022-03-29 20:59 /user/taller1
```

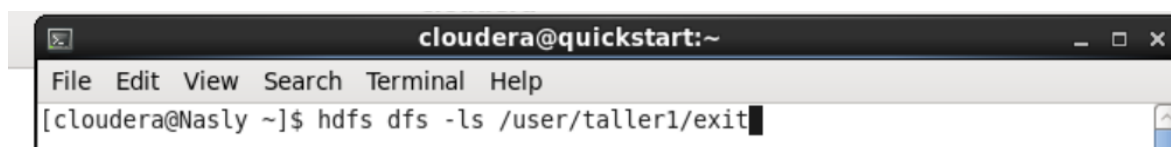
9. Se procede a realizar el map reduce con la librería de Hadoop y se crea una carpeta de salida exit donde se albergará nuestro resultante



10. Se procede a verificar el proceso



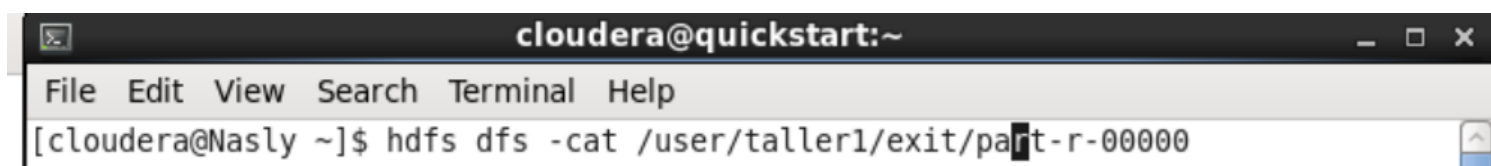
11. Se procede a listar taller 1 para visualizar el archivo resultante



12. Se verifica que se creó un archivo txt llamado part-r-00000

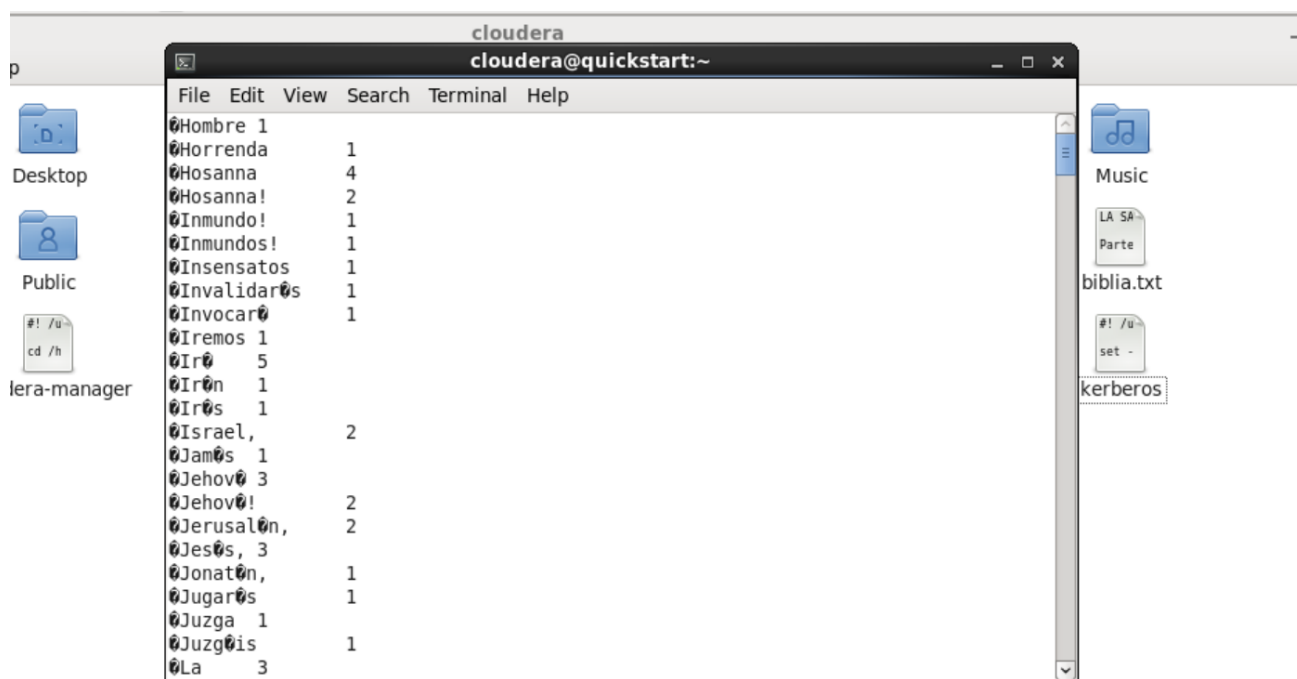
```
Found 2 items
-rw-r--r-- 1 cloudera supergroup
ida/_SUCCESS
-rw-r--r-- 1 cloudera supergroup
ida/part-r-00000
```

13. Se procede a abrir el archivo resultante part-r-00000



```
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@Nasly ~]$ hdfs dfs -cat /user/taller1/exi/part-r-00000
```

14. Se verifica que se le realizó map reduce al libro Isaías con Word Count de Hadoop.



| Word | Count |
|------------|-------|
| Hombre | 1 |
| Horrenda | 1 |
| Hosanna | 4 |
| Hosanna! | 2 |
| Inmundo! | 1 |
| Inmundos! | 1 |
| Insensatos | 1 |
| Invalidar | 1 |
| Invocar | 1 |
| Iremos | 1 |
| Ir | 5 |
| Iron | 1 |
| Iros | 1 |
| Israel, | 2 |
| Jam | 1 |
| Jehov | 3 |
| Jehov! | 2 |
| Jerusal | 2 |
| Jes | 3 |
| Jonat | 1 |
| Jugar | 1 |
| Juzga | 1 |
| Juzg | 1 |
| La | 3 |

15. Se verifica todo el documento

```

)
    todos 1
    t 1
    us 1
    vendisteis 1
    vendreis 1
    vive 1
    vivir 2
    viviron 1
    volver 2
    vuestros 1
    vuolvete 2
nager
    xito 2
    xito. 2
    xtasis 1
    xtasis. 1
    xtasis; 1
    y 64
    yeme 1
    yeme. 2
    yeme: 1
    yeme; 1
    ste 1
[cloudera@Nasly ~]$
```



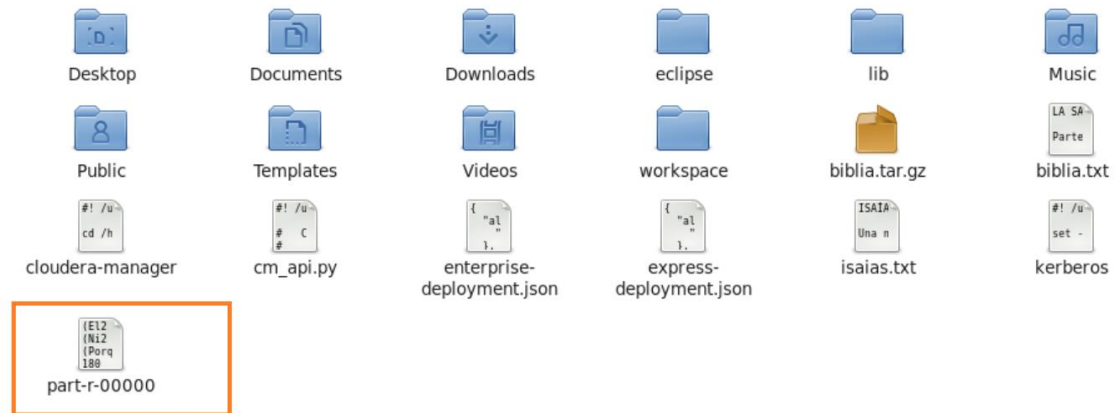
16. Se procede a copiar el archivo resultante a la carpeta de Cloudera home

```

xtasis. 1
xtasis; 1
y 64
yeme 1
yeme. 2
yeme: 1
yeme; 1
ste 1
[cloudera@Nasly ~]$ hdfs dfs -get /user/taller1/exit/part-r-00000
```



17. Se verifica la carpeta principal de Cloudera



18. También se puede verificar el proceso completo en el quickstart de Cloudera del explorador principal

| | | | | | | | |
|--|----------|------------|-----------|--------------------------------|---|--------|---------------------------|
| quickstart.cloudera:50070/explorer.html#/user | | | | | | | |
| Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started | | | | | | | |
| drwxr-xr-x | mapred | hadoop | 0 B | Wed Jul 19 05:34:54 -0700 2017 | 0 | 0 B | history |
| drwxrwxrwx | hive | supergroup | 0 B | Wed Jul 19 05:36:42 -0700 2017 | 0 | 0 B | hive |
| drwxrwxrwx | hue | supergroup | 0 B | Wed Jul 19 05:35:24 -0700 2017 | 0 | 0 B | hue |
| drwxrwxrwx | jenkins | supergroup | 0 B | Wed Jul 19 05:35:00 -0700 2017 | 0 | 0 B | jenkins |
| -rw-r--r-- | cloudera | supergroup | 184.88 KB | Tue Mar 29 20:50:53 -0700 2022 | 1 | 128 MB | mapreduce |
| drwxrwxrwx | oozie | supergroup | 0 B | Wed Jul 19 05:35:35 -0700 2017 | 0 | 0 B | oozie |
| drwxrwxrwx | root | supergroup | 0 B | Wed Jul 19 05:35:16 -0700 2017 | 0 | 0 B | root |
| drwxr-xr-x | hdfs | supergroup | 0 B | Wed Jul 19 05:36:36 -0700 2017 | 0 | 0 B | spark |
| -rw-r--r-- | cloudera | supergroup | 184.88 KB | Tue Mar 29 20:59:05 -0700 2022 | 1 | 128 MB | taller1 |