

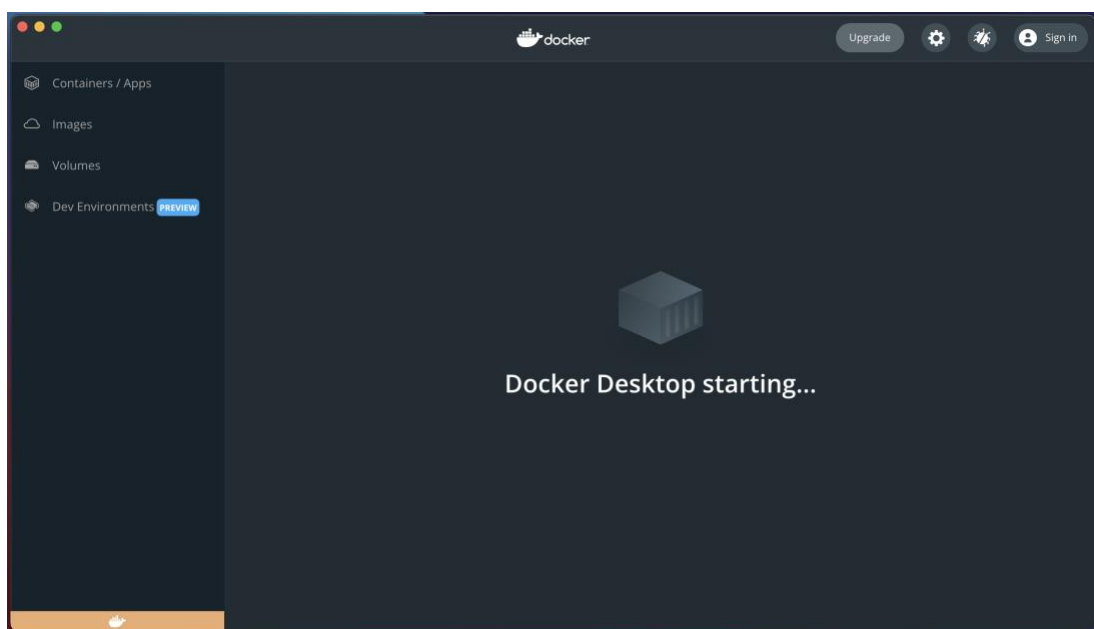
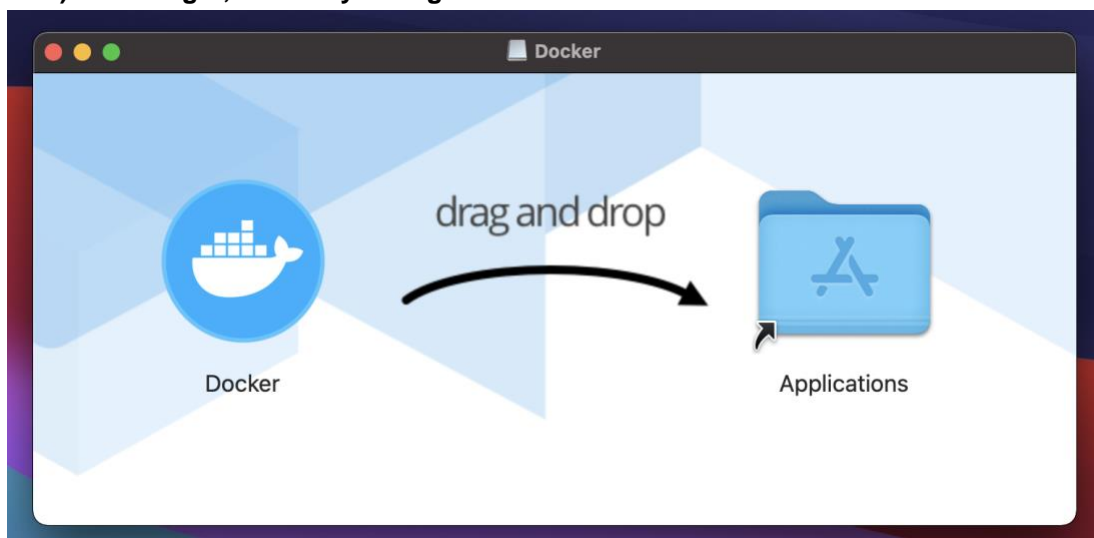
## DATOS BÁSICOS

Asignatura	Programación concurrente y distribuida
Curso Académico	21-22
Semestre	S1
Profesor	José Delgado Pérez
Alumnos	Rubén Servando Carrillo

## Actividad Individual AI3 Hadoop

### Enunciado Parte 1 → Instalación

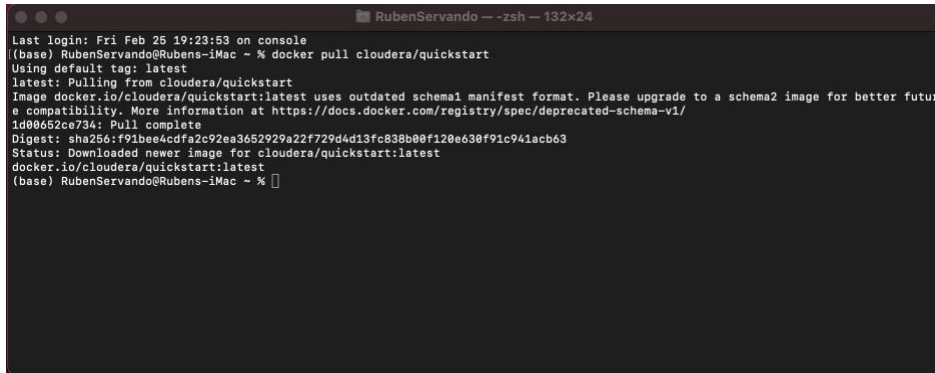
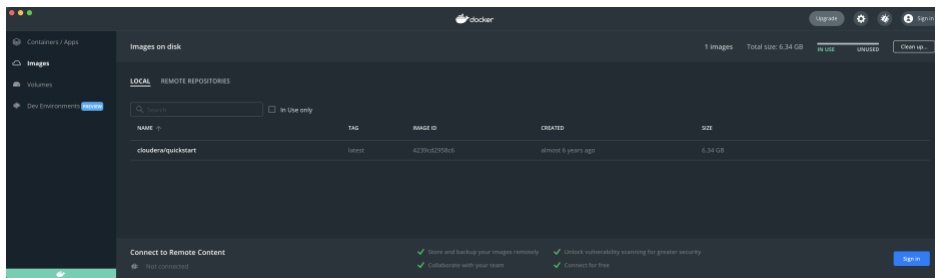
a) Descargar, instalar y configurar Docker.



Instalación de Docker Desktop en MacOS Big Sur

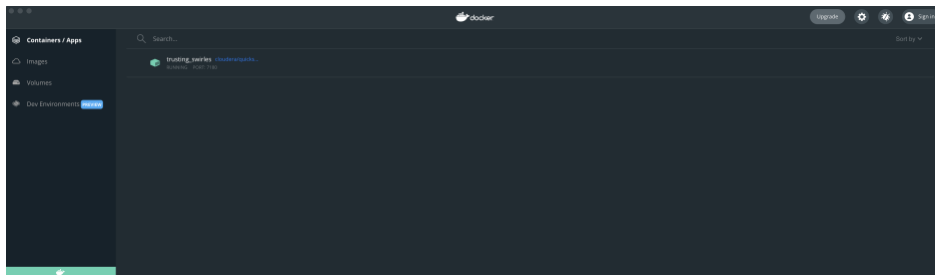


## b) Descargar imagen cloudera/Docker



Descarga de imagen Cloudera/Docker

## c) Crear un contenedor siguiendo las instrucciones del Hub de Docker Cloudera/Quickstart



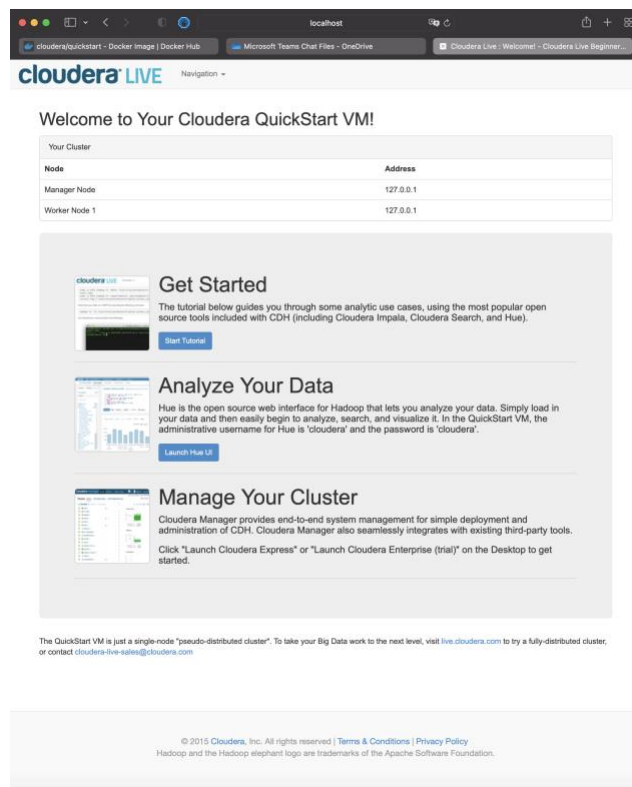


```
rubenservando@Rubens-Mac: ~$ docker ps
CONTAINER ID   IMAGE      COMMAND                  CREATED        STATUS        PORTS
179c9ee3b9f0e  cloudera/quickstart  "/usr/bin/docker-qui..."  16 minutes ago  Up 16 minutes  0.0.0.0:80->80/tcp, 0.0.0.0:7180->7180/tcp, 0.0.0.0:8888->8888/tcp
peaceful_mccarthy
(base) Rubenservando@Rubens-Mac: ~$
```

Descarga de imagen Cloudera/Docker

Comprobar que el contenedor está rodando:

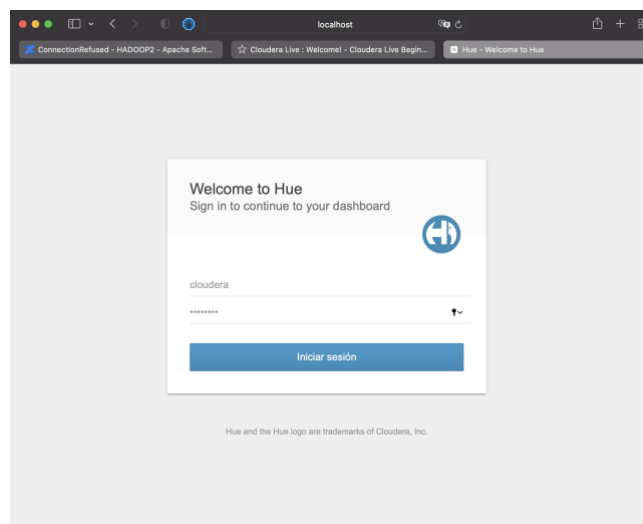
Abrir el navegador Safari e introducir → localhost/##/



Abrir otra pestaña de Safari e introducir → localhost:8888/about → iniciar sesión

user: cloudera

pass: cloudera



## MapReduce →



- b) Exportamos el proyecto en un fichero .jar para poder utilizarlo en el contenedor



## Conclusiones

Los pasos son:

1. Descargar Docker Desktop y arrancar → se detecta un problema de arranque que nos obliga a eliminar logs con el comando introducido en la terminal: `rm -rf Library/Containers/com.docker.*`  
Se encuentra la solución en le URL:  
<https://forums.docker.com/t/docker-is-starting-forever-solved-by-reinstall/13445>
2. Una vez arrancado el Docker Desktop, procedemos a descargar la imagen de cloudera, introducinendo en la terminal el comando siguiente: `docker pull cloudera/quickstart`  
que encontramos en la URL:  
<https://hub.docker.com/r/cloudera/quickstart>
3. Una vez descargada la imagen y comprobada la generación en el Docker Desktop, procedemos a levantar un contenedor introduciendo en la terminal el comando:  
`docker run --hostname=quickstart.cloudera --privileged=true -t -i -p 8888:8888 -p 80:80 -p 7180:7180 cloudera/quickstart /usr/bin/docker-quickstart`
4. Una vez comprobado en Docker Desktop que el contenedor ha levantado, debemos implementar el código y abrir el puerto, introduciendo en Safari, por un lado, `localhost/#/` → comprobar que el contenedor existe  
Por otro lado, abrir otra pestaña de Safari e introducir, `localhost:8888/about` → e iniciar sesión con  
*user: cloudera y password: cloudera*
5. para dejar programado el contador de palabras solicitado utilizando la IDE Eclipse.
6. Una vez exportado el fichero ejecutable (WordCounterDrive.jar) del programa de Eclipse, entonces introducimos en la terminal (A) la secuencia (línea a línea) de los comandos que se detallan a continuación para correrlo en la terminal:

```
sudo service hadoop-hdfs-namenode restart
```

```
sudo su hdfs
```

```
hadoop fs -mkdir /user/RubenServando
```

```
hadoop fs -chown cloudera /user/RubenServando
```

```
hadoop fs -mkdir /user/RubenServando/wordcounter /user/RubenServando/wordcounter/input
```

En otra pestaña de la terminal (B) que abriremos, vamos a cargar El Quijote del proyecto Gutenberg que hemos copiado dela siguiente URL , pegado en un archivo .txt y guardado en el escritorio:

<https://www.gutenberg.org/cache/epub/2000/pg2000.txt>

`docker ps` → el volumen en este caso se llama `peaceful_mccarthy` como vemos en la imagen

```

RubenServando - zsh - 193x24
...dera --privileged=true -t -i -p 8888:8888 -p 80:80 -p 7180:7180 cloudera/quickstart /usr/bin/docker-quickstart -- zsh
Last login: Mon Feb 28 16:19:23 on ttys000
(base) RubenServando@Rubens-iMac ~ % docker ps
CONTAINER ID   IMAGE          COMMAND                  CREATED    STATUS    PORTS                               NAMES
b70cee38f0a   cloudera/quickstart  "/usr/bin/docker-qui..."  16 minutes ago  Up 16 minutes    0.0.0.0:80->80/tcp, 0.0.0.0:7180->7180/tcp, 0.0.0.0:8888->8888/tcp    peaceful_mccarthy
(base) RubenServando@Rubens-iMac ~ % docker ps
CONTAINER ID   IMAGE          COMMAND                  CREATED    STATUS    PORTS                               NAMES
b70cee38f0a   cloudera/quickstart  "/usr/bin/docker-qui..."  2 hours ago  Up 2 hours    0.0.0.0:80->80/tcp, 0.0.0.0:7180->7180/tcp, 0.0.0.0:8888->8888/tcp    peaceful_mccarthy
(base) RubenServando@Rubens-iMac ~ % docker cp el_quijote.txt peaceful_mccarthy:/tmp/
```

```
cd Desktop
```

```
docker cp el_quijote.txt peaceful_mccarthy:/tmp/ → comando para meter el .txt en una carpeta /tmp
```



Volvemos a la pestaña **A** de la terminal anterior y comenzamos a introducir los comandos:

`hadoop fs -put /tmp/el_quijote.txt /user/RubenServando/wordcounter/input/`

Volvemos a la pestaña **B** de la terminal anterior para hacer o mismo con el fichero .jar de Eclipse:

`docker cp WordCounterDriver.jar peaceful_mccarthy:/tmp/` → comando para meter el .jar en una carpeta /tmp

```
Escritorio — zsh — 193x24
...dera --privileged=true -t -i -p 8888:8888 -p 80:80 -p 7180:7180 cloudera/quickstart /usr/bin/docker-quickstart
Last login: Mon Feb 28 16:19:23 on ttys000
(base) RubenServando@Rubens-iMac ~ % docker ps
CONTAINER ID   IMAGE          COMMAND                  CREATED        STATUS        PORTS                               NAMES
b70ceee38f0a   cloudera/quickstart   "/usr/bin/docker-qui-"   16 minutes ago   Up 16 minutes   0.0.0.0:80->80/tcp, 0.0.0.0:7180->7180/tcp, 0.0.0.0:8888->8888/tcp   peaceful_mccarthy
(base) RubenServando@Rubens-iMac ~ % docker ps
CONTAINER ID   IMAGE          COMMAND                  CREATED        STATUS        PORTS                               NAMES
b70ceee38f0a   cloudera/quickstart   "/usr/bin/docker-qui-"   2 hours ago     Up 2 hours     0.0.0.0:80->80/tcp, 0.0.0.0:7180->7180/tcp, 0.0.0.0:8888->8888/tcp   peaceful_mccarthy
(base) RubenServando@Rubens-iMac ~ % docker cp el_quijote.txt peaceful_mccarthy:/tmp/
ls -lat /Users/RubenServando/el_quijote.txt: no such file or directory
(base) RubenServando@Rubens-iMac Desktop ~ % docker cp el_quijote.txt peaceful_mccarthy:/tmp/
(base) RubenServando@Rubens-iMac Desktop ~ % docker cp WordCounterDriver.jar peaceful_mccarthy:/tmp/
(base) RubenServando@Rubens-iMac Desktop ~ %
```

Volvemos a la pestaña **A** de la terminal anterior y comenzamos a introducir los comandos:

`hadoop fs -put /tmp/ WordCounterDriver.jar /user/RubenServando/wordcounter/input/`

`exit`

`cp /tmp/WordCounterDriver.jar .`

`hadoop jar WordCounterDriver.jar /user/RubenServando/wordcounter/output`

```
RubenServando — @quickstart:/ — com.docker.cli - docker run --hostname=quickstart.cloudera --privileged=true -t -i -p 8888:8888 -p 80:80 -p 7180:7180 cloudera/quickstart /usr/bin/docker-quickstart
options are given and SOURCE and DEST are the same name for an existing, regular file.
Report cp bugs to bug-coreutils@gnu.org
GNU coreutils home page: <http://www.gnu.org/software/coreutils/>
General help using GNU software: <http://www.gnu.org/gethelp/>
For complete documentation, run: info coreutils 'cp invocation'
[root@quickstart /]# sudo su hdfs
bash-4.1$ hadoop fs -put /tmp/WordCounterDriver.jar /user/RubenServando/wordcounter/input/
bash-4.1$ hadoop jar WordCounterDriver.jar /user/RubenServando/wordcounter/input
Not a valid JAR: /WordCounterDriver.jar
bash-4.1$ exit
exit
[root@quickstart /]# hadoop jar WordCounterDriver.jar /user/RubenServando/wordcounter/input
Not a valid JAR: /WordCounterDriver.jar
[root@quickstart /]# cp /tmp/WordCounterDriver.jar .
[root@quickstart /]# hadoop jar WordCounterDriver.jar /user/RubenServando/wordcounter/input
Exception in thread "main" java.lang.ClassNotFoundException: /user/RubenServando/wordcounter/input
    at java.lang.Class.forName0(Native Method)
    at java.lang.Class.forName(Class.java:270)
    at org.apache.hadoop.util.RunJar.run(RunJar.java:214)
    at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
[root@quickstart /]# hadoop jar WordCounterDriver.jar /user/RubenServando/wordcounter/output
Exception in thread "main" java.lang.ClassNotFoundException: /user/RubenServando/wordcounter/output
    at java.lang.Class.forName0(Native Method)
    at java.lang.Class.forName(Class.java:270)
    at org.apache.hadoop.util.RunJar.run(RunJar.java:214)
    at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
[root@quickstart /]# cp /tmp/WordCounterDriver.jar .
cp: overwrite '/WordCounterDriver.jar'? y
[root@quickstart /]# hadoop jar WordCounterDriver.jar /user/RubenServando/wordcounter/output
Exception in thread "main" java.lang.ClassNotFoundException: /user/RubenServando/wordcounter/output
    at java.lang.Class.forName0(Native Method)
    at java.lang.Class.forName(Class.java:270)
    at org.apache.hadoop.util.RunJar.run(RunJar.java:214)
    at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
[root@quickstart /]#
```

LAMENTABLEMENTE EXISTE UN PROBLEMA DE COMPILACIÓN EN EL FICHERO .JAR, QUE NO CONSIGO RESOLVER POR FALTA DE CONOCIMIENTOS. EN EL VIDEO DE LA ACTIVIDAD QUEDA PATENTE EL CONFLICTO.