

# Laboratorio

Módulo 01

# Para poder realizar el Laboratorio se recomienda:

- Leer los contenidos previos.



# Cluster Hadoop

Se pueden elegir entre distintas opciones para realizar la práctica:

1) Instalación de Docker en tu sistema operativo:

<https://hub.docker.com/>

Si el sistema operativo usado es Linux:

<https://docs.docker.com/engine/install/>

# Cluster Hadoop

2) La segunda opción tiene que ver con que si el sistema operativo usado no es Linux, creemos una máquina virtual utilizando VirtualBox:

<https://www.virtualbox.org/>

Luego, seguir los pasos del punto 1, con Linux instalado.

Nota: En en el material de descarga hay un VDI con un Ubuntu Server ya instalado, el cual es posible utilizar.

Usuario: ubuntu

Contraseña: ubuntu

# Cluster Hadoop

Una vez que tenemos Docker instalado, habiendo escogido cualquiera de las opciones ya podemos crear el cluster Hadoop siguiendo las siguientes instrucciones:

1-Probemos la instalación de Docker:  
"sudo docker run hello-world"

```
Hello from Docker!  
This message shows that your installation appears to be working correctly.  
  
To generate this message, Docker took the following steps:  
1. The Docker client contacted the Docker daemon.  
2. The Docker daemon pulled the "hello-world" image from the Docker Hub.  
   (amd64)  
3. The Docker daemon created a new container from that image which runs the  
   executable that produces the output you are currently reading.  
4. The Docker daemon streamed that output to the Docker client, which sent it  
   to your terminal.  
  
To try something more ambitious, you can run an Ubuntu container with:  
$ docker run -it ubuntu bash  
  
Share images, automate workflows, and more with a free Docker ID:  
https://hub.docker.com/  
  
For more examples and ideas, visit:  
https://docs.docker.com/get-started/
```

# Cluster Hadoop

Descarguemos desde GitHub lo necesario para ejecutar el cluster:

```
"git clone https://github.com/dariomiguellopez/curso\_big\_data\_modulo1"  
"cd curso_big_data_modulo1"
```

Revisemos el archivo "start-container.sh", los archivos dentro de la carpeta "config" y la URL "localhost:8088".

Finalmente sigamos las instrucciones propuestas para ver en funcionamiento HDFS y MapReduce sobre Hadoop.

# ¡Muchas gracias!

¡Sigamos trabajando!