MANUAL INSTALACIÓN DOCKER Y DOCKER COMPUSE

1. Instalar Docker:

<https://docs.docker.com/desktop/setup/install/windows-install/>

1. Verificar los requisitos de tu sistema operativo para elegir la descarga apropiada.

Por ejemplo, si tuviese:

1. un sistema operativo Windows 11 Home Single Language con
2. Tipo de Sistema: Sistema operativo de 64 bits, descargaríamos:

**Docker Desktop for Windows-x86\_64 on the Microsoft Store.**

**(“Nota: en caso de no encontrar la versión X86, contemple que la opción de AMD es compatible, para los Intel”)**

1. Al descargar muy probablemente te solicite alguna actualización de Windows (Microsoft.WSL\_2.5.9.0\_x64\_ARM64), si es así, descargar e instalar.
2. Verificar que no cuentes con alguna restricción para poder trabajar, algún FireWire, VPN o Proxy. (muchas veces los antivirus traen un Firewall por default, el cual deberá deshabilitar sobre todo para que no te genere problemas al momento de instalar el/los conteiner’s).

**Nota:**

Usted podría encaso de no deshabilitar los Firewalls, experimentar un error similar al siguiente:

failed to copy: httpReadSeeker: failed open: unexpected status code https://registry-1.docker.io/v2/bde2020/hadoop-namenode/blobs/sha256:3192219afd04f93d90f0af7f89cb527d1af2a16975ea391ea8517c602ad6ddb6: 503 Service Unavailable

1. Al tener todo instalado, se comenzará con la actividad.

Practica 1

1. Crea tus archivos en block de notas
   1. docker-compose.yml

El contenido de dicho archivo es:

**services:**

**namenode:**

**image: bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8**

**container\_name: namenode**

**hostname: namenode**

**restart: always**

**ports:**

**- "9870:9870"**

**- "9000:9000"**

**volumes:**

**- hadoop\_namenode:/hadoop/dfs/name**

**environment:**

**- CLUSTER\_NAME=test**

**env\_file:**

**- ./hadoop.env**

**datanode1:**

**image: bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8**

**container\_name: datanode1**

**restart: always**

**volumes:**

**- hadoop\_datanode1:/hadoop/dfs/data**

**environment:**

**SERVICE\_PRECONDITION: "namenode:9870"**

**env\_file:**

**- ./hadoop.env**

**volumes:**

**hadoop\_namenode:**

**hadoop\_datanode1:**

* 1. hadoop.env

El contenido de dicho archivo es:

**CORE\_CONF\_fs\_defaultFS=hdfs://namenode:9000**

* 1. Colocarlos en raíz de C en un directorio con el nombre hadoop-cluster

1. Ingresa a CMD (Símbolo del Sistema)
2. Posiciónese en la ruta: **C:\>cd hadoop-cluster**

Ejecutar la siguiente instrucción, para iniciar el clúster.

Esto descargará las nuevas imágenes y creará los contenedores. Puede tardar un poco la primera vez

1. **C:\hadoop-cluster>docker-compose up -d**

**Texto

El contenido generado por IA puede ser incorrecto.**

1. **Accede al Shell:** Para ejecutar comandos HDFS, necesitas entrar al contenedor del NameNode. Usa este comando:
   1. docker exec -it namenode bash

**Windows PowerShell**

1. PS C: \Users\ana\_m> echo HOLA MUNDO

HOLA MUNDO

1. C: \Users\ana\_m> echo “HOLA HDFS DESDE MI MAQUINA LOCAL”

HOLA HDFS DESDE MI MAQUINA LOCAL

1. C: \Users\ana\_m> echo “HOLA HDFS DESDE MI MAQUINA LOCAL”>> ana.txt
2. C: \Users\ana\_m> dir

Desplegar el contenido del directorio donde debe aparecer el archivo generado en este caso: -a---- 30/06/2025 09:16 p. m. 68 ana.txt

Texto

El contenido generado por IA puede ser incorrecto.

1. PS C:\Users\ana\_m> more ana.txt

HOLA HDFS DESDE MI MAQINA LOCAL

Una vez dentro, tu terminal estará en el NameNode y podrás empezar a usar los

comandos de la siguiente sección.

1. Comandos Esenciales de HDFS para Practicar Los comandos de HDFS son muy parecidos a los de un terminal de Linux, pero generalmente van precedidos por hdfs dfs.

Nota: dentro de este contenedor, también puedes usar hadoop fs como alias de hdfs dfs.

1. Listar y Navegar
   1. ● Listar archivos y directorios:

hdfs dfs -ls /

1. Crear Directorios
   1. Crear un directorio:

hdfs dfs -mkdir /midirectorio

1. Crear una estructura de directorios anidada:

hdfs dfs -mkdir -p /proyectos/bigdata/dataset1

1. **Subir y Bajar Archivos (put)**
   1. Subir un archivo local a HDFS (put): Primero, crea un archivo de prueba en tu máquina local (fuera de docker).

# En tu máquina local (fuera de docker)

**echo "Hola HDFS desde mi máquina local" > archivo\_prueba.txt**

# Ahora sube el archivo al contenedor

**docker cp archivo\_prueba.txt namenode:/**

# Vuelve a entrar al contenedor

**docker exec -it namenode bash**

# Y ahora sí, súbelo a HDFS desde la raíz del contenedor

**hdfs dfs -put /archivo\_prueba.txt /midirectorio/**

1. **Bajar un archivo de HDFS a local (get):**

# Dentro del contenedor, baja el archivo a la raíz del contenedor

**hdfs dfs -get /midirectorio/archivo\_prueba.txt /archivo\_bajado.txt**

# Luego sal del contenedor (`exit`) y cópialo a tu máquina real

**docker cp namenode:/archivo\_bajado.txt .**

1. **Ver Contenido de Archivos**
   1. **Mostrar todo el contenido de un archivo (cat):**

**hdfs dfs -cat /midirectorio/archivo\_prueba.txt**

1. **Eliminar Archivos y Directorios** 
   1. **Eliminar un archivo:**

**hdfs dfs -rm /midirectorio/archivo\_prueba.txt**

1. **Eliminar un directorio y todo su contenido (con cuidado):**

**hdfs dfs -rm -R /proyectos Obtener Información**

1. **Ver el espacio usado por archivos/directorios: }**

**hdfs dfs -du -h /**