

Big Data Aplicada, Tarea 2

1. Crea una carpeta en HDFS llamada /user/tuusuario/tarea1.

Browse Directory

The screenshot shows the HDFS Browse Directory interface. The address bar displays the path `/user/carlos/`. Below the address bar, there is a search bar and a table of entries. The table has columns for Permission, Owner, Group, Size, Last Modified, Replication, Block Size, and Name. One entry is listed: `drwxr-xr-x` for Permission, `Administrador` for Owner, `supergroup` for Group, `0 B` for Size, `Oct 15 21:24` for Last Modified, `0` for Replication, `0 B` for Block Size, and `tarea1` for Name. Below the table, it says "Showing 1 to 1 of 1 entries".

2. Crea dos archivos de texto en tu ordenador, por ejemplo archivoA.txt y archivoB.txt, con frases diferentes.

The screenshot shows a Windows File Explorer window. The address bar displays the path `Documentos > carlos > Big_Data_Aplicada > Tarea-2`. The main area shows a table of files with columns for Nombre, Fecha de modificación, Tipo, and Tamaño. Two files are listed: `archivoA` and `archivoB`, both with a size of 1 KB and a modification date of 15/10/2025 21:29.

3. Sube ambos archivos a HDFS dentro de la carpeta creada.

The screenshot shows the HDFS Browse Directory interface. The address bar displays the path `/user/carlos/tarea1`. Below the address bar, there is a search bar and a table of entries. The table has columns for Permission, Owner, Group, Size, Last Modified, Replication, Block Size, and Name. Two entries are listed: `-rw-r--r--` for Permission, `Administrador` for Owner, `supergroup` for Group, `43 B` for Size, `Oct 15 21:37` for Last Modified, `1` for Replication, `128 MB` for Block Size, and `archivoA.txt` for Name. The second entry is `-rw-r--r--` for Permission, `Administrador` for Owner, `supergroup` for Group, `42 B` for Size, `Oct 15 21:37` for Last Modified, `1` for Replication, `128 MB` for Block Size, and `archivoB.txt` for Name. Below the table, it says "Showing 1 to 2 of 2 entries".

4. Lista el contenido de la carpeta en HDFS y comprueba que están los dos archivos.

The screenshot shows a Windows Command Prompt window. The command `hdfs dfs -ls /user/carlos/tarea1` has been entered. The output shows two items found in the directory:

```
C:\Users\Administrador\Documents\carlos>hdfs dfs -ls /user/carlos/tarea1
Found 2 items
-rw-r--r-- 1 Administrador supergroup 43 2025-10-15 21:37 /user/carlos/tarea1/archivoA.txt
-rw-r--r-- 1 Administrador supergroup 42 2025-10-15 21:37 /user/carlos/tarea1/archivoB.txt
```

Big Data Aplicada, Tarea 2

5. Copia uno de los archivos dentro de la misma carpeta, poniéndole otro nombre (por ejemplo, copia archivoA.txt como archivoA_copia.txt).

Directory listing for /user/carlos/tarea1:

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	Administrador	supergroup	43 B	Oct 15 21:37	1	128 MB	archivoA.txt
-rw-r--r--	Administrador	supergroup	43 B	Oct 15 21:47	1	128 MB	archivoA_copia.txt
-rw-r--r--	Administrador	supergroup	42 B	Oct 15 21:37	1	128 MB	archivoB.txt

```
C:\Users\Administrador\Documents\carlos>hdfs dfs -cp /user/carlos/tarea1/archivoA.txt /user/carlos/tarea1/archivoA_copia.txt
```

6. Muestra el contenido de cada archivo usando el comando adecuado.

File contents for archivoA.txt, archivoA_copia.txt, and archivoB.txt:

- archivoA.txt: Primer archivo de texto de la tarea de HDFS
- archivoA_copia.txt: Primer archivo de texto de la tarea de HDFS
- archivoB.txt: Segundo y último archivo de la tarea HDFS

```
C:\Users\Administrador\Documents\carlos>hdfs dfs -cat /user/carlos/tarea1/archivoA_copia.txt
```

7. Borra uno de los archivos (el que quieras) de la carpeta en HDFS.

Browse Directory

Directory listing for /user/carlos/tarea1:

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	Administrador	supergroup	43 B	Oct 15 21:37	1	128 MB	archivoA.txt
-rw-r--r--	Administrador	supergroup	43 B	Oct 15 21:47	1	128 MB	archivoA_copia.txt

```
C:\Users\Administrador\Documents\carlos>hdfs dfs -rm /user/carlos/tarea1/archivoB.txt
```

8. Entrega: los comandos usados y una breve explicación de qué hace cada uno.

- `hdfs dfs -mkdir {direccion/nombre}`: crea una carpeta en la dirección indicada con el nombre que se le añade.
- `hdfs dfs -put {direccion_archivo} {direccion_hadoop}`: añade a la carpeta indicada de hadoop el archivo indicado en la primera direccion
- `hdfs dfs -ls {direccion}`: devuelve una lista de los archivos de la dirección indicada.
- `hdfs dfs -cp {direccion_original} {direccion_copia/nombre}`: recoge el archivo indicado en la dirección original y lo añade en la dirección a la que se quiere pegar la copia con el nombre indicado.
- `hdfs dfs -cat {direccion}`: lee el contenido del archivo indicado.
- `hdfs dfs -rm {direccion}`: elimina el archivo/carpeta de la dirección indicada.