

Práctica con Comandos HDFS

Cada actividad incluye un comando específico para interactuar con HDFS y una breve explicación de su propósito. Realice las capturas de pantalla necesarias que justifiquen la realización de esta práctica.

Cada comando de HDFS deberá tener la siguiente nomenclatura:

- `hdfs dfs -comando`
- `hadoop fs -comando`

1. Subir un Archivo al HDFS

`hadoop fs -put /ruta/local/archivo.txt /ruta/hdfs/`

Propósito: Familiarizarse con el comando de carga de archivos.

Actividad: Sube un archivo de texto pequeño desde tu sistema local a HDFS.

```
[cloudera@quickstart Desktop]$ nano tarea.txt
[cloudera@quickstart Desktop]$ cat tarea.txt
hola
[cloudera@quickstart Desktop]$ pwd
/home/cloudera/Desktop
[cloudera@quickstart Desktop]$ # Crear la ruta de destino si no existe
[cloudera@quickstart Desktop]$ hadoop fs -mkdir -p /user/usuario/ventas/
[cloudera@quickstart Desktop]$
[cloudera@quickstart Desktop]$ # Copiar el archivo al destino en HDFS
[cloudera@quickstart Desktop]$ hadoop fs -put /home/usuario/datos/ventas_2024.txt /user/usuario/ventas/
put: `/home/usuario/datos/ventas_2024.txt': No such file or directory
[cloudera@quickstart Desktop]$ hadoop fs -put /home/cloudera/Desktop/tarea.txt /user/usuario/ventas/
```

2. Listar el Contenido de un Directorio en HDFS

`hadoop fs -ls /ruta/hdfs/`

Propósito: Verifica que el archivo subido esté en la ubicación correcta.

Actividad: Lista el contenido del directorio HDFS para verificar la carga.

```
[cloudera@quickstart Desktop]$ hadoop fs -ls /user/usuario/ventas/
Found 1 items
-rw-r--r-- 1 cloudera supergroup 5 2024-12-15 04:54 /user/usuario/ventas/tarea.txt
[cloudera@quickstart Desktop]$
```

3. Copiar un Archivo desde HDFS al Sistema Local

`hadoop fs -get /ruta/hdfs/archivo.txt /ruta/local/`

Propósito: Asegurarse de que los archivos pueden recuperarse desde HDFS.

Actividad: Descarga el archivo de HDFS y compáralo con el original para verificar que el contenido coincida

```
[cloudera@quickstart Desktop]$ hadoop fs -get /user/usuario/ventas/tarea.txt ./
[cloudera@quickstart Desktop]$ ls
Eclipse.desktop Enterprise.desktop Express.desktop Kerberos.desktop Parcels.desktop tarea.txt
[cloudera@quickstart Desktop]$ cat tarea.txt
hola
```

4. Crear un Directorio en HDFS

hadoop fs -mkdir /ruta/hdfs/nuevo_directorio

Propósito: Practicar la creación de estructuras en HDFS.

Actividad: Crea un directorio en HDFS y lista su contenido para verificar su creación.

```
[cloudera@quickstart ~]$ hadoop fs -mkdir /user/usuario/compras
[cloudera@quickstart ~]$ hadoop fs -ls /user/usuario/
Found 2 items
drwxr-xr-x  - cloudera supergroup          0 2024-12-15 06:10 /user/usuario/com
pras
drwxr-xr-x  - cloudera supergroup          0 2024-12-15 04:54 /user/usuario/ven
tas
```

5. Cambiar Permisos en HDFS

hadoop fs -chmod 755 /ruta/hdfs/archivo_o_directorio

Propósito: Controlar el acceso en HDFS.

Actividad: Cambia los permisos de un archivo o directorio y observa cómo esto afecta a otros usuarios.

```
[cloudera@quickstart ~]$ hadoop fs -chmod 755 /user/usuario/ventas/tarea.txt

[cloudera@quickstart ~]$
[cloudera@quickstart ~]$ hadoop fs -ls -l /user/usuario/ventas/tarea.txt
-ls: Illegal option -l
Usage: hadoop fs [generic options] -ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [<path> ...]
[cloudera@quickstart ~]$ hadoop fs -ls /user/usuario/ventas/tarea.txt
-rwxr-xr-x  1 cloudera supergroup          5 2024-12-15 04:54 /user/usuario/ventas/tarea.txt
```

6. Eliminar un Archivo o Directorio en HDFS

hadoop fs -rm /ruta/hdfs/archivo.txt

```
[cloudera@quickstart ~]$ hadoop fs -rm /user/usuario/ventas/tarea.txt
Deleted /user/usuario/ventas/tarea.txt
[cloudera@quickstart ~]$ hadoop fs -ls /user/usuario/ventas/tarea.txt
ls: `/user/usuario/ventas/tarea.txt': No such file or directory
```

hadoop fs -rm -r /ruta/hdfs/nuevo_directorio

```
[cloudera@quickstart ~]$ hadoop fs -rm -r /user/usuario/ventas/
Deleted /user/usuario/ventas
[cloudera@quickstart ~]$ hadoop fs -ls /user/usuario/ventas/
ls: `/user/usuario/ventas/': No such file or directory
```

Propósito: Practicar la eliminación de archivos y directorios en HDFS.

Actividad: Elimina los archivos y directorios creados y confirma su eliminación.

7. Mostrar el Contenido de un Archivo en HDFS

hadoop fs -cat /ruta/hdfs/archivo.txt

Propósito: Verificar el contenido de un archivo en HDFS sin necesidad de descargarlo.

Actividad: Visualiza el contenido de un archivo en HDFS.

```
[cloudera@quickstart Desktop]$ hadoop fs -cat /user/usuario/practica/tarea2.txt
esto es un ejemplo de contenido para tarea2
```

8. Mover un Archivo dentro de HDFS

```
hadoop fs -mv /ruta/hdfs/archivo.txt /nueva/ruta/hdfs/
```

Propósito: Practicar la reorganización de archivos en HDFS.

Actividad: Mueve un archivo a otra ubicación en HDFS.

```
[cloudera@quickstart Desktop]$ hadoop fs -mv /user/usuario/practica/tarea2.txt /user/usuario/compras/
[cloudera@quickstart Desktop]$ hadoop fs -ls /user/usuario/compras/
Found 1 items
-rw-r--r--  1 cloudera supergroup          44 2024-12-15 06:40 /user/usuario/compras/tarea2.txt
```

9. Copiar un Archivo dentro de HDFS

```
hadoop fs -cp /ruta/hdfs/archivo.txt /otra/ruta/hdfs/
```

Propósito: Practicar la duplicación de archivos en HDFS.

Actividad: Copia un archivo a otra ubicación dentro de HDFS.

```
[cloudera@quickstart Desktop]$ hadoop fs -cp /user/usuario/compras/tarea2.txt /user/usuario/practica/
[cloudera@quickstart Desktop]$ hadoop fs -ls /user/usuario/practica/
Found 1 items
-rw-r--r--  1 cloudera supergroup          44 2024-12-15 07:09 /user/usuario/practica/tarea2.txt
```

10. Buscar un Archivo en HDFS

```
hadoop fs -find /ruta/hdfs/ -name archivo*.txt
```

Propósito: Practicar la búsqueda de archivos en HDFS.

Actividad: Encuentra archivos en HDFS usando un patrón de búsqueda.

```
[cloudera@quickstart Desktop]$ hadoop fs -find / -name tarea2.txt
/user/usuario/compras/tarea2.txt
/user/usuario/practica/tarea2.txt
```

11. Comprensión de Tolerancia a Fallos en HDFS. Verificar Replicación y Salud de Archivos en HDFS

```
hadoop fsck /ruta/hdfs/
```

Propósito: Entender cómo HDFS mantiene la tolerancia a fallos mediante réplicas de datos.

Actividad: Ejecuta el comando `fsck` en un directorio para ver la replicación de archivos y el estado de salud de los bloques en HDFS. Observa si HDFS muestra advertencias o errores en la replicación.

```
[cloudera@quickstart Desktop]$ hadoop fsck /  
DEPRECATED: Use of this script to execute hdfs command is deprecated.  
Instead use the hdfs command for it.  
  
Connecting to namenode via http://quickstart.cloudera:50070/fsck?ugi=FSCK started by cloudera (auth:SIMPLE) from /127.0.0.1 for path / at :  
.....  
/hbase/data/hbase/meta/1588230740/info/ad0c1bf764ff4ca0a884ef2b12d950  
8775264580 block blk_1073742828  
  
/hbase/data/hbase/meta/1588230740/info/ad0c1bf764ff4ca0a884ef2b12d950  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....Status: CORRUPT  
Total size:      864007289 B  
Total dirs:     133  
Total files:    952  
Total symlinks: 0  
Total blocks (validated): 947 (avg. block size 912362 B)
```