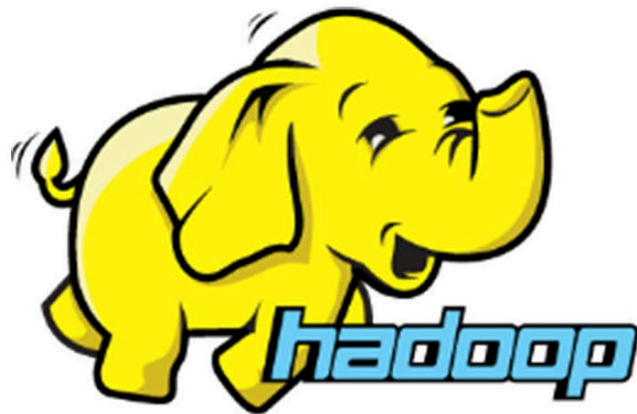


Práctica con Comandos HDFS



Adrián Yared Armas de la Nuez



Contenido

1. Enunciado.....	3
2. Subir un Archivo al HDFS.....	3
2.1 Propósito.....	3
2.2 Comando.....	3
2.3 Resultado.....	3
3. Listar el Contenido de un Directorio en HDFS.....	4
3.1 Propósito.....	4
3.2 Comando.....	4
3.3 Resultado.....	4
4. Copiar un Archivo desde HDFS al Sistema Local.....	4
4.1 Propósito.....	4
4.2 Comando.....	4
4.3 Resultado.....	4
5. Crear un Directorio en HDFS.....	4
5.1 Propósito.....	5
5.2 Comando.....	5
5.3 Resultado.....	5
6. Cambiar Permisos en HDFS.....	5
6.1 Propósito.....	5
6.2 Comando.....	5
6.3 Resultado.....	5
7. Eliminar un Archivo o Directorio en HDFS.....	5
7.1 Propósito.....	5
7.2 Comando.....	5
7.3 Resultado.....	6
8. Mostrar el Contenido de un Archivo en HDFS.....	6
8.1 Propósito.....	6
8.2 Comando.....	6
8.3 Resultado.....	6
9. Mover un Archivo dentro de HDFS.....	6
9.1 Propósito.....	6
9.2 Comando.....	6
9.3 Resultado.....	6
10. Copiar un Archivo dentro de HDFS.....	7
10.1 Propósito.....	7
10.2 Comando.....	7



Práctica con Comandos HDFS

10.3 Resultado.....	7
11. Mover un Archivo dentro de HDFS.....	7
11.1 Propósito.....	7
11.2 Comando.....	7
11.3 Resultado.....	8
12. Comprensión de Tolerancia a Fallos en HDFS. Verificar Replicación y Salud de Archivos en HDFS.....	8
12.1 Propósito.....	8
12.2 Comando.....	8
12.3 Resultado.....	8
12.4 Interpretación del resultado.....	9

1. Enunciado

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

2. Subir un Archivo al HDFS

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

2.1 Propósito

Familiarizarse con el comando de carga de archivos.

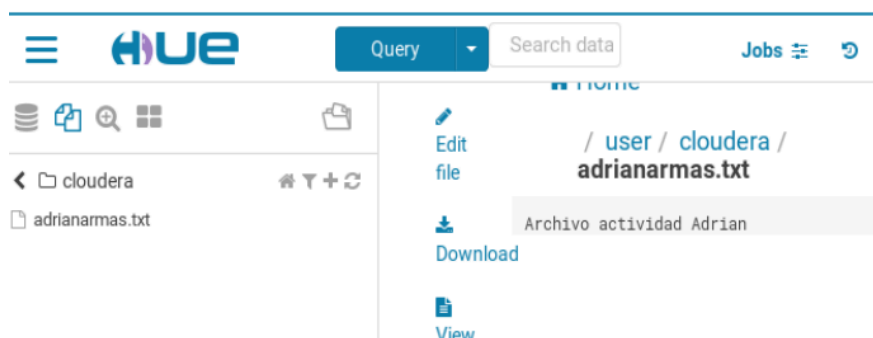
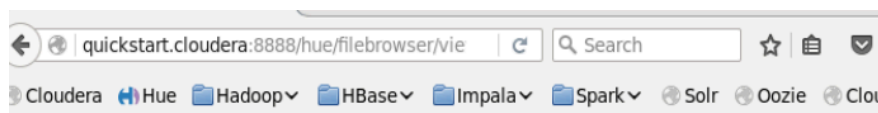
2.2 Comando

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

2.3 Resultado

```
[cloudera@quickstart ~]$ nano adrianarmas.txt
[cloudera@quickstart ~]$ ls
adrianarmas.txt  Downloads  lib  Templates
cloudera-manager eclipse  Music  Videos
cm_api.py        enterprise-deployment.json parcels workspace
Desktop          express-deployment.json  Pictures
Documents        kerberos  Public
```

```
[cloudera@quickstart ~]$ hadoop fs -put adrianarmas.txt
[cloudera@quickstart ~]$
```





3. Listar el Contenido de un Directorio en HDFS

Lista el contenido del directorio HDFS para verificar la carga.

3.1 Propósito

Verifica que el archivo subido esté en la ubicación correcta

3.2 Comando

```
hadoop fs -ls
```

3.3 Resultado

```
[cloudera@quickstart ~]$ hadoop fs -ls
Found 1 items
-rw-r--r--  1 cloudera cloudera      25 2024-11-13 10:15 adrianarmas.txt
[cloudera@quickstart ~]$ █
```

4. Copiar un Archivo desde HDFS al Sistema Local

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

4.1 Propósito

Asegurarse de que los archivos pueden recuperarse desde HDFS

4.2 Comando

```
hadoop fs -get adrianarmas.txt
```

4.3 Resultado

```
[cloudera@quickstart ~]$ hadoop fs -get adrianarmas.txt
get: `adrianarmas.txt': File exists
[cloudera@quickstart ~]$ █
```

5. Crear un Directorio en HDFS

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

5.1 Propósito

Practicar la creación de estructuras en HDFS.



5.2 Comando

`hadoop fs -mkdir /ruta/hdfs/CarpetaAdrianArmas`

5.3 Resultado

```
[cloudera@quickstart ~]$ hadoop fs -mkdir CarpetaAdrianArmas  
[cloudera@quickstart ~]$
```

6. Cambiar Permisos en HDFS

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

6.1 Propósito

Controlar el acceso en HDFS

6.2 Comando

`hadoop fs -chmod 755 /ruta/hdfs/CarpetaAdrianArmas`

6.3 Resultado

Doy todos los permisos a mi carpeta:

```
[cloudera@quickstart ~]$ hadoop fs -chmod 755 CarpetaAdrianArmas  
[cloudera@quickstart ~]$
```

7. Eliminar un Archivo o Directorio en HDFS

Elimina los archivos y directorios creados y confirma su eliminación

7.1 Propósito

Practicar la eliminación de archivos y directorios en HDFS

7.2 Comando

`hadoop fs -rm /ruta/hdfs/adrianarmas.txt`
`hadoop fs -rm -r /ruta/hdfs/CarpetaAdrianArmas`

7.3 Resultado

Elimino el archivo txt:



Práctica con Comandos HDFS

```
[cloudera@quickstart ~]$ hadoop fs -rm adrianarmas.txt  
Deleted adrianarmas.txt
```

Elimino el directorio:

```
[cloudera@quickstart ~]$ hadoop fs -rm -r CarpetaAdrianArmas  
Deleted CarpetaAdrianArmas
```

8. Mostrar el Contenido de un Archivo en HDFS

Visualiza el contenido de un archivo en HDFS.

8.1 Propósito

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

8.2 Comando

```
hadoop fs -cat archivo.txt
```

8.3 Resultado

Creo de nuevo el archivo:

```
[cloudera@quickstart ~]$ hadoop fs -put adrianarmas.txt  
[cloudera@quickstart ~]$ nano adrianarmas.txt  
[cloudera@quickstart ~]$ hadoop fs -cat adrianarmas.txt  
Archivo actividad Adrian _
```

9. Mover un Archivo dentro de HDFS

Mueve un archivo a otra ubicación en HDFS

9.1 Propósito

Practicar la reorganización de archivos en HDFS.

9.2 Comando

```
hadoop fs -mv adrianarmas.txt CarpetaAdrianArmas/
```

9.3 Resultado

Creo de nuevo el directorio:

```
cloudera@quickstart ~]$ hadoop fs -mkdir CarpetaAdrianArmas  
cloudera@quickstart ~]$ hadoop fs -chmod 755 CarpetaAdrianArmas
```

Copio el archivo a mi carpeta

```
[cloudera@quickstart ~]$ hadoop fs -mv adrianarmas.txt CarpetaAdrianArmas/
```



10. Copiar un Archivo dentro de HDFS

Copia un archivo a otra ubicación dentro de HDFS.

10.1 Propósito

Practicar la duplicación de archivos en HDFS.

10.2 Comando

```
hadoop fs -cp adrianarmas.txt CarpetaAdrianArmas/
```

10.3 Resultado

Elimino el txt:

```
[cloudera@quickstart ~]$ hadoop fs -rm -r CarpetaAdrianArmas/adrianarmas.txt
Deleted CarpetaAdrianArmas/adrianarmas.txt
```

Lo creo en raíz y lo pego en mi carpeta (para tener en cuenta los que además se cumplen los permisos de mi carpeta):

```
[cloudera@quickstart ~]$ hadoop fs -put adrianarmas.txt
[cloudera@quickstart ~]$ hadoop fs -cat adrianarmas.txt
Archivo actividad Adrian
```

Lo copio en mi carpeta:

```
[cloudera@quickstart ~]$ hadoop fs -cp adrianarmas.txt CarpetaAdrianArmas/
[cloudera@quickstart ~]$
```

11. Mover un Archivo dentro de HDFS

Encuentra archivos en HDFS usando un patrón de búsqueda

11.1 Propósito

Practicar la búsqueda de archivos en HDFS.

11.2 Comando

```
hadoop fs -find CarpetaAdrianArmas/ -name adrianarmas.txt
```

11.3 Resultado

find de mi txt en mi carpeta:



Práctica con Comandos HDFS

```
[cloudera@quickstart ~]$ hadoop fs -find CarpetaAdrianArmas/ -name adrianarmas.txt  
CarpetaAdrianArmas/adrianarmas.txt
```

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

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.

12.1 Propósito

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

12.2 Comando

```
hadoop fsck /user/cloudera/CarpetaAdrianArmas
```

12.3 Resultado

```
[cloudera@quickstart ~]$ hdfs fsck /user/cloudera/CarpetaAdrianArmas  
Connecting to namenode via http://quickstart.cloudera:50070/fsck?ugi=cloudera&path=%2Fuser%2Fcloudera%2FCarpetaAdrianArmas  
FSCK started by cloudera (auth:SIMPLE) from /127.0.0.1 for path /user/cloudera/CarpetaAdrianArmas at Wed Nov 13 10:55:53 PST 2024  
.Status: HEALTHY  
Total size: 25 B  
Total dirs: 1  
Total files: 1  
Total symlinks: 0  
Total blocks (validated): 1 (avg. block size 25 B)  
Minimally replicated blocks: 1 (100.0 %)  
Over-replicated blocks: 0 (0.0 %)  
Under-replicated blocks: 0 (0.0 %)  
Mis-replicated blocks: 0 (0.0 %)  
Default replication factor: 1  
Average block replication: 1.0  
Corrupt blocks: 0  
Missing replicas: 0 (0.0 %)  
Number of data-nodes: 1  
Number of racks: 1  
FSCK ended at Wed Nov 13 10:55:53 PST 2024 in 1 milliseconds  
  
The filesystem under path '/user/cloudera/CarpetaAdrianArmas' is HEALTHY  
[cloudera@quickstart ~]$
```



12.4 Interpretación del resultado

Este resultado nos dice que pesa un total de 25 B, que contiene un directorio, 1 archivo, 0 enlaces simbólicos, 1 bloque validado, 1 bloque mínimamente replicado, 0 bloques sobre-replicados, 0 bloques bajo-replicados, 0 bloques mal replicados, el tamaño promedio del bloque (25B), 0 bloques corruptos, 0 réplicas faltantes, 1 datanode y 1 rack (donde están todos los datos).

En conclusión analiza el estado de todos los datos tanto en su bloque individual como en el general y su estado (perdidos, corruptos o sanos) y en función a eso nos devuelve si el directorio es “Healthy o no”