**PREWORK**

Sesión 04

**Comandos básicos de HDFS**

HDFS básico

**🎯 Objetivos**

* Entender los comandos básicos de HDFS en Hadoop: HDFS es el componente principal del ecosistema de Hadoop, que es responsable de almacenar grandes conjuntos de datos estructurados o no estructurados en varios nodos y, por lo tanto, mantener los metadatos en forma de archivos de registro. Para usar los comandos de HDFS, primero debe iniciar los servicios de Hadoop con el siguiente comando:

### 

### 🚀 **Desarrollo**

### **Parte 1: Los comandos elementales en HDFS**

Para usar los comandos de HDFS, primero debe iniciar los servicios de Hadoop con el siguiente comando:

sbin/start-all.sh

Para comprobar que los servicios de Hadoop están en funcionamiento, utilice el siguiente comando:

jps

Lista de comandos

ls: este comando se usa para listar todos los archivos. Utilice lsr para un enfoque recursivo. Es útil cuando queremos una jerarquía de una carpeta.

Sintáxis:

bin/hdfs dfs -ls <path>

Imprimirá todos los directorios presentes en HDFS. El directorio bin contiene ejecutables, por lo que bin / hdfs significa que queremos los ejecutables de hdfs, particularmente los comandos dfs (sistema de archivos distribuido).

**mkdir**: To create a directory. In Hadoop *dfs* there is no home directory by default. So let’s first create it.

**Syntax:**

**bin/hdfs dfs -mkdir <folder name>**

**creating home directory:**

**hdfs/bin -mkdir /user**

**hdfs/bin -mkdir /user/username -> write the username of your computer**

**touchz**: It creates an empty file.

**Syntax:**

**bin/hdfs dfs -touchz <file\_path>**

**copyFromLocal (o) put:** para copiar archivos / carpetas desde el sistema de archivos local a la tienda hdfs. Este es el comando más importante. Sistema de archivos local significa los archivos presentes en el sistema operativo.

**Syntax:**

bin/hdfs dfs -copyFromLocal <local file path> <dest(present on hdfs)>

**Cat:** Para imprimir el contenido del archivo.

Sintaxis:

bin/hdfs dfs -cat <path>

**copyToLocal (or) get:** Para copiar archivos / carpetas desde la tienda hdfs al sistema de archivos local.

**Syntax:**

bin/hdfs dfs -copyToLocal <<srcfile(on hdfs)> <local file dest>

**Listado de directorios y ficheros:**

| hdfs dfs -ls / | Lista todos los ficheros y directorios para el path / |
| --- | --- |
| hdfs dfs -ls -h / | Lista los ficheros con su tamaño en formato legible |
| hdfs dfs -ls -R / | Lista todos los ficheros y directorios recursivamente (con subdirectorios) |
| hdfs dfs -ls /file\* | Lista todos los ficheros que cumplen el patrón (ficheros que comienzan con ‘file’) |

**Lectura y escritura de archivos**

| COMANDO | DESCRIPCIÓN |
| --- | --- |
| hdfs dfs -text /app.log | Imprime el fichero en modo texto por la terminal |
| hdfs dfs -cat /app.log | Muestra el contenido del fichero en la salida estándar |
| hdfs dfs -appendToFile /home/file1 /file2 | Añade el contenido del fichero local ‘file1’ al fichero en hdfs ‘file2’ |

Cargas y descargar archivos

| hdfs dfs -put /home/file1 /hadoop | Copia el fichero ‘file1’ del sistema de ficheros local a hdfs |
| --- | --- |
| hdfs dfs -put -f /home/file1 /hadoop | Copia el fichero ‘file1’ del sistema de ficheros local a hdfs y lo sobreescribe en el caso de que ya exista |
| hdfs dfs -put -l /home/file1 /hadoop | Copia el fichero ‘file1’ del sistema de ficheros local a hdfs. Fuerza replicación 1 y permite al DataNode persistir los datos de forma perezosa. |
| hdfs dfs -put -p /home/file1 /hadoop | Copia el fichero ‘file1’ del sistema de ficheros local a hdfs. Mantiene los tiempos de acceso, de modificación y propietario original |
| hdfs dfs -get /file1 /home/ | Copia el fichero ‘file1’ de hdfs al sistema de ficheros local |
| hdfs dfs -moveFromLocal /home/file1 /hadoop | Copia el fichero ‘file1’ del sistema de ficheros local a hdfs y luego lo borra del sist. ficheros local |

Gestión de ficheros

| hdfs dfs -cp /hadoop/file1 /hadoop1 | Copia el fichero al directorio destino en hdfs |
| --- | --- |
| hdfs dfs -cp -p /hadoop/file1 /hadoop1 | Copia el fichero al directorio destino en hdfs conservando tiempos de acceso y de modificación, propietario y modo |
| hdfs dfs -rm /hadoop/file1 | Elimina el fichero ‘file1’ de hdfs y lo envía a la papelera |
| hdfs dfs -rm -r /hadoop  hdfs dfs -rm -R /hadoop  hdfs dfs -rmr /hadoop | Elimina el directorio y su contenido en hdfs |
| hdfs dfs -rm -skipTrash /file1 | Elimina el fichero sin dejarlo en la papelera |
| hdfs dfs -mkdir /hadoop2 | Crea un directorio en hdfs |
| hdfs dfs -touchz /hadoop3 | Crea un fichero en hdfs con tamaño 0 |

Comandos de administración

| hdfs dfs -df /hadoop | Muestra la capacidad y el espacio libre y usado del sistema de ficheros |
| --- | --- |
| hdfs dfs -df -h /hadoop | Muestra la capacidad y el espacio libre y usado del sistema de ficheros en formato legible |
| hadoop version | Muestra la versión de hadoop |
| hdfs fsck / | Comprueba el estado de salud del sistema de ficheros |
| hdfs dfsadmin -safemode leave | Deshabilita el modo seguro del NameNode |
| hdfs namenode -format | Formatea el NameNode |

Estos son uno de varios comandos que puedes usar en Hadoop. Durante la sesión aprenderás a aplicar algunos de ellos.