

Hoja de trucos del Módulo 2: Introducción al Ecosistema de Hadoop

Paquete/Método	Descripción	Ejemplo de Código
bin/hadoop	Todos los comandos de Hadoop se invocan mediante el script bin/hadoop. Ejecutar el script de Hadoop sin argumentos imprime la descripción de todos los comandos.	Ejecutando el script de Hadoop sin argumentos: bin/hadoop
cat	Lee cada parámetro de archivo en secuencia y lo escribe en la salida estándar. Si no especificas un nombre de archivo, el comando cat lee de la entrada estándar. También puedes especificar un nombre de archivo de - (guion) para la entrada estándar.	Crea dos archivos de ejemplo. echo "Este es el archivo 1" > file1.txt echo "Este es el archivo 2" > file2.txt Usa el comando cat para leer y mostrar el contenido de ambos archivos cat file1.txt file2.txt Salida de ejemplo (Contenido de file1.txt y file2.txt): Este es el archivo 1 Este es el archivo 2
cd	Se utiliza para moverse de manera eficiente desde el directorio de trabajo actual a diferentes directorios en tu sistema.	Sintaxis básica del comando cd: cd [opciones]... [directorio] Ejemplo 1: Cambiar la ubicación del directorio a "folder1" cd /usr/local/folder1

Paquete/Método	Descripción	Ejemplo de Código
		<p>Ejemplo 2: Volver al directorio de trabajo anterior</p> <pre>cd -</pre> <p>Ejemplo 3: Subir un nivel desde el árbol del directorio de trabajo actual</p> <pre>cd ..</pre>
create table	Se utiliza para crear una nueva tabla en una base de datos	<p>Crea una nueva base de datos (si no se ha creado ya).</p> <pre>CREATE DATABASE your_database;</pre> <p>Usa la base de datos recién creada.</p> <pre>USE your_database;</pre> <p>Crea una nueva tabla llamada "employees" en Hive.</p> <pre>CREATE TABLE employees (id INT, first_name STRING, last_name STRING, email STRING, hire_date DATE) ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' STORED AS TEXTFILE;</pre> <p>Muestra la lista de tablas en la base de datos.</p> <pre>SHOW TABLES;</pre>

Paquete/Método	Descripción	Ejemplo de Código
		<p>Salida de ejemplo (Lista de Tablas):</p> <pre>OK employees</pre>
curl	Una herramienta de línea de comandos (pronunciado "curl") que permite intercambiar datos entre un dispositivo y un servidor a través de un terminal. El usuario especifica la URL del servidor, la ubicación a la que desea enviar la solicitud y los datos que desea enviar a la URL del servidor utilizando esta interfaz de línea de comandos (CLI).	<p>Ejemplo 1: Enviando una solicitud GET y mostrando la respuesta</p> <p>Envía una solicitud GET a un servidor y muestra la respuesta.</p> <pre>curl https://www.example.com</pre> <p>En este ejemplo, usamos el comando curl para enviar una solicitud GET a https://www.example.com y mostrar la respuesta HTML del servidor.</p> <p>-----</p> <p>Ejemplo 2: Enviando datos a un servidor usando Solicitud POST:</p> <p>Envía una solicitud POST con datos a un servidor y muestra la respuesta.</p> <pre>curl -X POST -d "name=John&age=30" https://www.example.com/api</pre> <p>En este ejemplo, usamos el comando curl para enviar una solicitud POST a https://www.example.com/api con datos name=John&age=30 y mostrar la respuesta JSON del servidor.</p>
docker exec	Ejecuta un nuevo comando en un contenedor en ejecución. Solo se ejecuta cuando el proceso principal del contenedor está en funcionamiento, y no se reinicia si el contenedor se reinicia.	<p>Ejecutando un comando en un contenedor Docker en ejecución:</p> <p>Ejecuta un nuevo comando dentro de un contenedor Docker en ejecución.</p> <pre>docker exec -it container_name_or_id ls /app</pre> <p>Salida de ejemplo (Lista de archivos en el directorio '/app' dentro del contenedor):</p> <pre>file1.txt file2.txt subdirectory</pre>

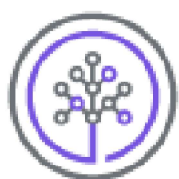
Paquete/Método	Descripción	Ejemplo de Código
		<p>En este ejemplo:</p> <ul style="list-style-type: none">• <code>docker exec</code> se utiliza para ejecutar un nuevo comando (<code>ls /app</code>) dentro de un contenedor Docker en ejecución.• <code>-it</code> habilita una sesión de terminal interactiva, lo que te permite ver la salida del comando.• <code>container_name_or_id</code> es el nombre o ID del contenedor Docker en ejecución en el que deseas ejecutar el comando.• <code>ls /app</code> es el comando que lista los archivos y directorios en el directorio <code>'/app'</code> dentro del contenedor.
docker-compose	Compose es una herramienta para definir y ejecutar aplicaciones Docker de múltiples contenedores. Utiliza el archivo YAML para configurar los servicios y nos permite crear y comenzar todos los servicios desde un solo archivo de configuración.	<p>Iniciando contenedores Docker usando docker-compose:</p> <p>Supongamos que tienes un archivo <code>docker-compose.yml</code> como este:</p> <pre>version: '3' services: web: image: nginx:latest ports: - "80:80" db: image: postgres:latest environment: POSTGRES_PASSWORD: example_password</pre> <p>Puedes usar <code>docker-compose</code> para iniciar los servicios definidos en el archivo <code>docker-compose.yml</code> de la siguiente manera:</p> <p>Navega al directorio que contiene el archivo <code>docker-compose.yml</code>.</p> <pre>cd /path/to/your/docker-compose-project</pre> <p>Inicia los contenedores Docker definidos en el archivo <code>docker-compose.yml</code></p> <pre>docker-compose up</pre>
docker pull	Puedes descargar imágenes de Docker de Internet.	<pre>docker pull [OPTIONS] IMAGE_NAME[:TAG]</pre>

Paquete/Método	Descripción	Ejemplo de Código
docker run	Ejecuta un comando en un nuevo contenedor, obteniendo la imagen y comenzando el contenedor si es necesario.	<pre>docker run [OPTIONS] IMAGE [COMMAND] [ARG...]</pre>
git clone	Puedes crear una copia de un repositorio específico o de una rama dentro de un repositorio.	<pre>git clone REPOSITORY_URL [DESTINATION_DIRECTORY]</pre>
hdfs dfs	Apache Hadoop <code>hadoop fs</code> o <code>hdfs dfs</code> son comandos del sistema de archivos para interactuar con HDFS. Estos comandos son muy similares a los comandos de Unix. Hadoop proporciona dos tipos de comandos para interactuar con el sistema de archivos: <code>hadoop fs</code> o <code>hdfs dfs</code> . La principal diferencia es que los comandos de Hadoop son compatibles con múltiples sistemas de archivos como S3, Azure y muchos más.	<p>Ejemplo-1:</p> <p>Listando archivos y directorios en HDFS:</p> <p>Lista archivos y directorios en el directorio raíz de HDFS.</p> <pre>hdfs dfs -ls /</pre> <p>Ejemplo-2: En este ejemplo, usamos el <code>hdfs dfs -ls</code> comando para listar archivos y directorios en el directorio raíz de HDFS.</p> <pre>hdfs dfs -ls /</pre> <p>Salida de ejemplo:</p> <pre>drwxr-xr-x - hdfs hduser 0 2023-09-13 10:00 /user drwxrwxrwx - hdfs hduser 0 2023-09-13 10:05 /tmp drwxrwxrwx - mapred hduser 0 2023-09-13 10:10 /mapred</pre> <p>Crea un nuevo directorio llamado "mydata" en HDFS.</p> <pre>hdfs dfs -mkdir /user/your_username/mydata</pre>

Paquete/Método	Descripción	Ejemplo de Código
hdfs dfs -cat	Muestra el contenido de un archivo.	<p>Muestra el contenido de un archivo en HDFS.</p> <pre>hdfs dfs -cat /path/to/file.txt</pre>
hdfs dfs -mkdir	Crea un directorio llamado path en HDFS	<p>Crea un directorio en HDFS.</p> <pre>hdfs dfs -mkdir /user/username/mydirectory</pre>
hdfs dfs -put	Sube un archivo o carpeta desde el disco local a HDFS.	<p>Sube un archivo desde el sistema de archivos local a HDFS.</p> <pre>hdfs dfs -put localfile.txt /user/username/hdfsfile.txt</pre>
LOAD DATA INPATH	Hive proporciona la funcionalidad para cargar entidades de tabla precreadas ya sea desde el sistema de archivos local o desde HDFS. Este comando se utiliza para cargar datos en la tabla de Hive.	<p>Carga datos desde HDFS en una tabla de Hive.</p> <pre>LOAD DATA INPATH '/user/username/hdfsfile.txt' INTO TABLE mytable;</pre>
ls	Escribe en la salida estándar los contenidos de cada parámetro de directorio especificado o el nombre de cada parámetro de archivo especificado, junto con cualquier otra información que solicites con las banderas. Si no especificas un parámetro de archivo o directorio, el comando ls muestra el contenido del directorio actual.	<p>Sintaxis básica del comando</p> <pre>ls [opciones] [archivo/directorio]</pre> <p>Ejemplo 1: Ordena los nombres de archivos mostrados en el orden del último tiempo de modificación. 'r' es para mostrar en orden inverso</p> <pre>ls -lt ls -ltr</pre>

Paquete/Método	Descripción	Ejemplo de Código
		<p>Ejemplo 2: Muestra archivos ocultos</p> <pre>ls -a</pre>
mkdir	Se utiliza para crear uno o más directorios especificados por el parámetro de Directorio. Cada nuevo directorio contiene las entradas estándar punto (.) y punto punto (..). Puedes especificar los permisos para los nuevos directorios con la bandera -m Modo.	<p>Crea un nuevo directorio llamado "myfolder."</p> <pre>mkdir myfolder</pre>
SELECT * FROM	Lista todas las filas de la tabla para verificar si los datos se han cargado desde el archivo.	<p>Selecciona todas las filas de una tabla.</p> <pre>SELECT * FROM tablename;</pre>
show tables	Se utiliza para ver todas las tablas en la base de datos que han sido seleccionadas.	<p>Muestra todas las tablas en la base de datos seleccionada.</p> <pre>SHOW TABLES;</pre>
tar	Busca archivos comprimidos en el dispositivo predeterminado (generalmente cinta) a menos que especifiques otro dispositivo. Al escribir en un archivo comprimido, el comando tar utiliza un archivo temporal (el archivo /tmp/tar*) y mantiene en memoria una tabla de archivos con varios enlaces.	<p>Crea un archivo comprimido tar de un directorio.</p> <pre>tar -cvf archive.tar /path/to/directory</pre>
wget	Significa web get. El wget es un comando descargador de archivos gratuito y no interactivo. No interactivo significa que puede funcionar en segundo plano cuando el usuario no está conectado.	<p>Sintaxis básica del comando wget; las opciones comúnmente usadas son [-v], [-h], [-b], [-e], [-o], [-a], [-q]</p> <pre>wget [opciones]... [URL]...</pre>

Paquete/Método	Descripción	Ejemplo de Código
		<p>Ejemplo 1: Especifica descargar file.txt a través de la URL del sitio web HTTP en el directorio de trabajo.</p> <pre>wget http://example.com/file.txt</pre> <p>Ejemplo 2: Especifica descargar archive.zip a través de la URL del sitio web HTTP en segundo plano y te devuelve al símbolo del sistema en el ínterin.</p> <pre>wget -b http://www.example.org/files/archive.zip</pre>



Skills Network