**Instalación de HIVE**

**Creado**: 19 Junio

**Última modificación**: 19 Junio

Para conseguir una instalación de HIVE congruente con la anterior instalación de HADOOP, se ha mantenido el hilo de tutoriales de instalación desde el portal de Edureka [1], contrastando la información con la página oficial de HIVE [2]. Además, se ha hecho uso de la ayuda proporcionada en Stack Overflow [3] puesto que la instalación según Edureka fallaba.

**Paso 1.** Se ha descargado la versión estable 2.1.1 de HIVE, desde la página de lanzamientos de HIVE [4]

**¿Cómo?**

**1.1:** Acceder a la página de los lanzamientos de HIVE.

**1.2:** Elegir el enlace llamado “stable-2”.

**1.3:** Descargar el archivo comprimido “apache-hive-2.1.1-bin.tar.gz”

**Paso 2.** Se ha descomprimido el fichero descargado

**¿Cómo?**

**2.1:** *tar -xvf apache-hive-2.1.1-bin.tar.gz*

**Paso 3.** Se ha movido el directorio descomprimido en una carpeta deseada (a partir de ahora dicha carpeta se llamará HIVE\_HOME)

**¿Cómo?**

**3.1:** *cp apache-hive-2.1.1-bin HIVE\_HOME*

Ahora que HIVE está descomprimido, hay que tener en cuenta 4 pasos para poder montarlo sobre HADOOP:

* **Primero:** Actualizar las variables de entorno en el fichero .*bashrc* para integrar HIVE.
* **Segundo:** Crear un directorio para HIVE dentro del sistema de ficheros de HADOOP.
* **Tercero:** Indicar la ruta hacia HADOOP dentro de los ficheros de configuración de HIVE.
* **Cuarto:** Indicar la conexión con la base de datos MySQL en los ficheros de configuración de HIVE.

**Paso 4.** Se han añadido las variables de entorno para HIVE

**¿Cómo?**

**4.1:** Se ha editado el fichero .bashrc de la carpeta HOME de Ubuntu y se han añadido las variable de entorno para el HOME de HIVE:

* **HIVE\_HOME** – Ruta a la carpeta de instalación de HIVE (Ruta a la carpeta HIVE\_HOME).
* **4.2:** Se ha actualizado el fichero .bashrc con el comando *source .bashrc*

**Paso 5.** Se ha creado un directorio para HIVE dentro del sistema de ficheros de HADOOP y se le ha dado el permiso de escritura a dicho directorio

**¿Cómo?**

**5.1:** Se ha ejecutado el comando *hadoop fs -mkdir -p /user/hive/warehouse* para crear el directorio dentro de HADOOP

**5.2:** Se ha ejecutado el comando *hadoop fs -chmod g+w /user/hive/warehouse* para darle permisos de escritura al grupo sobre esa carpeta.

**¿Como comprobarlo?**

Para comprobar lo anterior, basta con arrancar el SingleNode de HADOOP (*$HADOOP\_HOME/sbin/start-all.sh*), acceder a localhost:50070 y explorar el sistema de ficheros desde el menú superior derecho llamado *Utilities → Browse the file system*. Podremos observar que tenemos la carpeta *user*, dentro de ella *hive* y dentro de hive está *warehouse*.

**Paso 6.** Se le ha dado permiso de escritura para el grupo en el directorio /tmp de HADOOP

**¿Cómo?**

**6.1:** Se ha ejecutado el comando *hadoop fs -chmod g+w /tmp* para darle permisos de escritura al grupo sobre esa carpeta.

**Paso 7.** Se ha actualizado la ruta a HADOOP en los ficheros de configuración de HIVE:

**¿Cómo?**

**7.1:** Se ha editado el fichero HIVE\_HOME/bin/hive-conf.sh añadiendo la ruta a HADOOP\_HOME. Esto se ha hecho añadiendo una linea en dicho fichero: *export HADOOP\_HOME= <ruta a HADOOP\_HOME>.*

**Paso 8.** Se ha creado el fichero *hive-site.xml* en el directorio de configuración de HIVE:

**¿Cómo?**

**8.1:** Se ha creado el fichero *hive-site.xml* en el directorio HIVE\_HOME/conf con la siguiente información:

*<?xml version="1.0" encoding="UTF-8" standalone="no"?>*

*<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>*

*<configuration>*

*<property>*

*<name>javax.jdo.option.ConnectionURL</name>*

*<value>jdbc:mysql://localhost:3306/hive</value>*

*</property>*

*<property>*

*<name>javax.jdo.option.ConnectionDriverName</name>*

*<value>com.mysql.jdbc.Driver</value>*

*</property>*

*<property>*

*<name>javax.jdo.option.ConnectionUserName</name>*

*<value>root</value>*

*</property>*

*<property>*

*<name>javax.jdo.option.ConnectionPassword</name>*

*<value><<****password****>></value>*

*</property>*

*<property>*

*<name>datanucleus.autoCreateSchema</name>*

*<value>true</value>*

*</property>*

*<property>*

*<name>datanucleus.fixedDatastore</name>*

*<value>true</value>*

*</property>*

*<property>*

*<name>datanucleus.autoCreateTables</name>*

*<value>True</value>*

*</property>*

*</configuration>*

Como se puede observar, en este fichero se han configurado aquellos ajustes para conectarse con una base de datos MySQL a través del usuario *root.*

**Paso 9.** Se ha descargado y copiado el paquete jar necesario para la conexión con MySQL

**¿Cómo?**

**9.1:** Se ha accedido a la página de los repositorios Maven para el conector necesario de MySQL [5] y se ha descargado el archivo jar.

**9.2:** Se ha copiado dicho archivo en la carpeta HIVE\_HOME/lib

**Paso 10.** Se ha arrancado HIVE

**¿Cómo?**

**10.1:** Se han arrancado los demonios de HADOOP con el comando *HADOOP\_HOME/sbin/start\_all.sh*

**10.1:** Se ha arrancado HIVE con el comando *HIVE\_HOME/bin/hive*

[1] Portal Edureka: <https://www.edureka.co/blog/apache-hive-installation-on-ubuntu?utm_source=youtube&utm_campaign=hive-installation-051216-wr&utm_medium=description>

[2] Página oficial Apache HIVE: <https://hive.apache.org/>

[3] Ayuda de Stack Overflow: <https://stackoverflow.com/questions/35449274/java-lang-runtimeexception-unable-to-instantiate-org-apache-hadoop-hive-ql-meta>

[4] Página de lanzamientos de Apache HIVE: <http://mirror.netcologne.de/apache.org/hive/>

[5] Página del jar de MySQL: <https://mvnrepository.com/artifact/mysql/mysql-connector-java/5.1.28>