

一.Hive安装包下载

二.Hive安装

三.验证

一.Hive安装包下载

下载地址<http://mirrors.hust.edu.cn/apache/>

选择合适的Hive版本进行下载，进到stable-2文件夹可以看到稳定的2.x的版本是2.3.3

二.Hive安装

使用MySQL做为Hive的元数据库，所以先安装MySQL。参考

<https://www.runoob.com/mysql/mysql-install.html>

1. 解压

```
tar -zxvf apache-hive-2.3.3-bin.tar.gz -C /usr/software
```

2. 修改 hive-env.sh

```
cp -p hive-env.sh.template hive-env.sh && vim hive-env.sh
```

并添加如下内容：

```
HADOOP_HOME=/usr/software/hadoop-2.7.3
```

```
export HIVE_CONF_DIR=/usr/software/apache-hive-2.1.1-bin/conf
```

3. 修改hive-site.xml

```
cp -p hive-default.xml.template hive-site.xml && vim hive-site.xml
```

修改如下内容：

```
<property>
  <name>hive.cli.print.header</name>
  <value>true</value> <!-- false 改为true -->
```

```
<description>Whether to print the names of the columns in query output.
</description>
</property>
```

```
<property>
  <name>hive.cli.print.current.db</name>
  <value>true</value> <!-- false 改为true -->
  <description>Whether to include the current database in the Hive prompt.
</description>
</property>
```

添加元数据 存放的数据库(mysql)配置

```
<property>
  <name>javax.jdo.option.ConnectionURL</name>
  <value>jdbc:mysql://namenode01.hadoop.com:3306/hivedb?
createDatabaseIfNotExist=true</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>123456</value>
</property>
```

4. 修改log日志路径

```
cp -p hive-log4j.properties.template hive-log4j.properties && vim hive-
log4j.properties
```

修改如下内容:

```
property.hive.log.dir = /usr/software/apache-hive-2.1.1-bin/logs
```

5. 添加mysql连接包

将 mysql-connector-java-5.1.39.jar 包放到 hive的lib目录下

6. 添加环境变量

```
vim .bash_profile
```

添加内容

```
export HIVE_HOME=/home/hadoop/apps/apache-hive-2.3.3-bin
export PATH=$PATH:$HIVE_HOME/bin
```

source .bash_profile

三.验证

hive --help

出现hive相关信息

初始化元数据库

schematool -dbType mysql -initSchema

注意：当使用的 hive 是 2.x 之前的版本，不做初始化也是 OK 的，当 hive 第一次启动的时候会自动进行初始化，只不过会不会生成足够多的元数据库中的表。在使用过程中会慢慢生成。但最后进行初始化。如果使用的 2.x 版本的 Hive，那么就必须手动初始化元数据库。使用命令：

使用 hive 命令进入hive

https://www.cnblogs.com/qingyunzong/p/8708057.html#_label0

<https://blog.51cto.com/flyfish225/2096888>

<https://www.cnblogs.com/zhoadng/p/11911127.html>

四: 连接

4.1 cli方式连接

使用hive命令进入,,用于linux平台命令行查询, 查询语句基本跟mysql查询语句类似

通过 **hive --service serviceName --help** 可以查看某个具体命令的使用方式

4.2 HiveServer2/beeline

用java等程序实现通过jdbc等驱动的访问hive就用这种起动方式了, 这个是程序员最需要的方式了

1. 修改hadoop中 hdfs-site.xml 配置文件

加入一条配置信息, 表示启用 webhdfs

```
<property>
<name>dfs.webhdfs.enabled</name>
<value>true</value>
</property>
```

2. 修改hadoop中 core-site.xml 配置文件

```
<property>
<name>hadoop.proxyuser.hadoop.hosts</name>
<value>*</value>
</property>
<property>
<name>hadoop.proxyuser.hadoop.groups</name>
<value>*</value>
</property>
```

以上操作做好了之后（最好重启一下HDFS集群）

hadoop.proxyuser.hadoop.hosts 配置成*的意义，表示任意节点使用 hadoop 集群的代理用户 hadoop 都能访问 hdfs 集群，
hadoop.proxyuser.hadoop.groups 表示代理用户的组所属

3. 启动hiveserver2服务

```
nohup hiveserver2 >/dev/null 2>&1 &
```

会启动一个进程： runJar

4. 启动 beeline 客户端去连接：

```
beeline -u jdbc:hive2://hadoop3:10000 -n hadoop （这种好像要单独设置密码才能使用）
```

或者

```
beeline
```

```
!connect jdbc:hive2://localhost:10000/hivedb
```

再输数据库用户和密码

-u：指定元数据库的连接信息
-n：指定mysql数据库用户名
-p：密码

<https://www.iteye.com/blog/491569462-qq-com-1948436>

https://www.cnblogs.com/qingyunzong/p/8715925.html#_label1_1

<https://www.cnblogs.com/lenmom/p/11218807.html>

