

## Hive Installation

- step 0: Install jdk 1.8 (Java 11 is not supported) and Hadoop 3.x.
- step 1: Download hive 3.1.2 in home directory from hive website.
- step 2: Extract hive archive into home directory (e.g. /home/sunbeam).
- step 3: In ~/.bashrc

```
export HIVE_HOME=$HOME/apache-hive-3.1.2-bin
export PATH=$HIVE_HOME/bin:$PATH
```

- step 4: Create hive-site.xml in \$HIVE\_HOME/conf.

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<configuration>
  <property>
    <name>javax.jdo.option.ConnectionURL</name>
    <value>jdbc:derby:;databaseName=/home/sunbeam/apache-hive-
3.1.2-bin/metastore_db;create=true</value>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionDriverName</name>
    <value>org.apache.derby.jdbc.EmbeddedDriver</value>
  </property>
  <property>
    <name>hive.metastore.uris</name>
    <value>thrift://0.0.0.0:9083</value>
  </property>
  <property>
    <name>javax.jdo.PersistenceManagerFactoryClass</name>
    <value>org.datanucleus.api.jdo.JDOPersistenceManagerFactory</value>
  </property>
  <property>
    <name>hive.metastore.warehouse.dir</name>
    <value>hdfs://localhost:9000/user/hive/warehouse</value>
  </property>
  <property>
    <name>hive.server2.enable.doAs</name>
    <value>>false</value>
  </property>
  <property>
    <name>hive.support.concurrency</name>
    <value>true</value>
  </property>
  <property>
    <name>hive.enforce.bucketing</name>
```

```
<value>true</value>
</property>
<property>
  <name>hive.exec.dynamic.partition.mode</name>
  <value>nonstrict</value>
</property>
<property>
  <name>hive.txn.manager</name>
  <value>org.apache.hadoop.hive.ql.lockmgr.DbTxnManager</value>
</property>
<property>
  <name>hive.compactor.initiator.on</name>
  <value>true</value>
</property>
<property>
  <name>hive.compactor.worker.threads</name>
  <value>2</value>
</property>
</configuration>
```

- step 5: Create metastore.

```
terminal> cd $HIVE_HOME
terminal> schematool -dbType derby -initSchema -verbose
```

- step 6: Start Hadoop.

```
terminal> start-dfs.sh
terminal> start-yarn.sh
```

- step 7: Create Hive warehouse directory.

```
terminal> hadoop fs -mkdir -p /user/hive/warehouse
terminal> hadoop fs -chmod g+w /user/hive/warehouse
```

- step 8: Start Hive metastore service.

```
terminal> hive --service metastore
```

- Then test on new terminal.

```
terminal> netstat -tln | grep "9083"
terminal> jps
```

- step 10: Start hive CLI

```
terminal> hive
```

```
SHOW DATABASES;

CREATE DATABASE edbda;

USE edbda;

SHOW TABLES;

CREATE TABLE students(id INT, name CHAR(50), marks DOUBLE)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE;

SHOW TABLES;

DESCRIBE students;

INSERT INTO students VALUES (1, 'Nitin', 96.23);
INSERT INTO students VALUES (2, 'Vishal', 86.43), (3, 'Rahul', 76.54),
(4, 'Amit', 97.43);

SELECT * FROM students;

EXIT;
```

## Beeline Installation

- step 11. Add these settings in hadoop core-site.xml --> Use current username at \${user}

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

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

- step 13. Restart the Hadoop

```
stop-yarn.sh  
stop-dfs.sh  
start-dfs.sh  
start-yarn.sh
```

- step 14. Go to \$HIVE\_HOME (on terminal) and start the ThriftService (if not already started):

```
hive --service metastore
```

- step 15. Go to \$HIVE\_HOME (on another terminal) and start the HiveServer:

```
hive --service hiveserver2
```

```
terminal> netstat -tln | grep "10000"  
terminal> jps
```

- step 16. Open another terminal and give "beeline" command.

```
beeline -u jdbc:hive2://localhost:10000/default -n $USER
```

```
SHOW DATABASES;  
  
USE edbda;  
  
SHOW TABLES;  
  
SELECT * FROM students;
```