

**Exp. No : 5****HIVE Installation**

1. Download and Extract apache hive 3.1.2

```
harish@fedora:~/hive$ ls
bin  binary-package-licenses  conf  examples  hcatalog  jdbc  lib  LICENSE  NOTICE  RELEASE_NOTES.txt  scripts
harish@fedora:~/hive$
```

2. Update HIVE Configurations in .bashrc

```
harish@fedora:~/hive
GNU nano 7.2 .bashrc
export YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib/native"
export HADOOP_CONF_DIR=$HADOOP_HOME/etc/hadoop
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export HADOOP_STREAMING=$HADOOP_HOME/share/hadoop/tools/lib/hadoop-streaming-3.4.0.jar
export HADOOP_LOG_DIR=$HADOOP_HOME/logs
export PDSH_RCMD_TYPE=ssh

#PIG settings
export PIG_HOME=/home/harish/pig
export PATH=$PATH:$PIG_HOME/bin
export PIG_CLASSPATH=$PIG_HOME/conf:$HADOOP_INSTALL/etc/hadoop/
export PIG_CONF_DIR=$PIG_HOME/conf
export JAVA_HOME=/usr/lib/jvm/jdk1.8.0_202
export PATH=$PATH:$JAVA_HOME/bin
export PIG_CLASSPATH=$PIG_CONF_DIR:$PATH
#PIG setting ends

export HIVE_HOME=/home/harish/hive
export HIVE_CONF_DIR=$HIVE_HOME/conf
export PATH=$PATH:$HIVE_HOME/bin
export HADOOP_USER_CLASSPATH_FIRST=true
export CLASSPATH=$CLASSPATH:$HIVE_CONF_DIR:$HIVE_HOME/lib/*
```

3. Change directory to apache-hive-3.1.2-bin/conf

```
harish@fedora:~/hive$ ls
bin  binary-package-licenses  conf  examples  hcatalog  jdbc  lib  LICENSE  NOTICE  RELEASE_NOTES.txt  scripts
harish@fedora:~/hive$ cd conf/
harish@fedora:~/hive/conf$ ls
beeline-log4j2.properties.template  hive-env.sh.template  hive-site.xml  llap-daemon-log4j2.properties.template
hive-default.xml.template  hive-exec-log4j2.properties.template  ivysettings.xml  parquet-logging.properties
hive-env.sh  hive-log4j2.properties.template  llap-cli-log4j2.properties.template
harish@fedora:~/hive/conf$
```

4. Create hive-env.sh

```
harish@fedora:~/hive/conf
GNU nano 7.2 hive-env.sh

if [ "$SERVICE" = "cli" ]; then
  if [ -z "$DEBUG" ]; then
    export HADOOP_OPTS="$HADOOP_OPTS -XX:NewRatio=12 -Xms10m -XX:MaxHeapFreeRatio=40 -XX:MinHeapFreeRatio=10"
  else
    export HADOOP_OPTS="$HADOOP_OPTS -XX:NewRatio=12 -Xms10m -XX:MaxHeapFreeRatio=40 -XX:MinHeapFreeRatio=10"
  fi
fi

The heap size of the jvm started by hive shell script can be controlled via:

export HADOOP_HEAPSIZE=1024

Larger heap size may be required when running queries over large number of files or partitions.
By default hive shell scripts use a heap size of 256 (MB). Larger heap size would also be appropriate for hive server.

Set HADOOP_HOME to point to a specific hadoop install directory
HADOOP_HOME=${bin}/../../hadoop

Hive Configuration Directory can be controlled by:
export HIVE_CONF_DIR=

Folder containing extra libraries required for hive compilation/execution can be controlled by:
export HIVE_AUX_JARS_PATH=

export HADOOP_HOME=/home/harish/hadoop
export HIVE_CONF_DIR=/home/harish/hive/conf
```

## 5. Install and Change mysql root password

```
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 8
Server version: 8.0.39 Source distribution

Copyright (c) 2000, 2024, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> SELECT user, host, plugin FROM mysql.user WHERE user = 'root';
+-----+-----+-----+
| user | host      | plugin                |
+-----+-----+-----+
| root | localhost | mysql_native_password |
+-----+-----+-----+
1 row in set (0.00 sec)

mysql> ALTER USER 'root'@'localhost' IDENTIFIED WITH 'mysql_native_password' BY 'hayag
reevan';
Query OK, 0 rows affected (0.62 sec)

mysql> FLUSH PRIVILEGES;
Query OK, 0 rows affected (0.12 sec)

mysql>
```

## 6. Create and configure hive-site.xml

```
GNU nano 7.2      hive-site.xml

<property>
<name>javax.jdo.option.ConnectionURL</name>
<value>jdbc:mysql://localhost/metastore?createDatabaseIfNotExist=true</value>
</property>
<property>
<name>javax.jdo.option.ConnectionDriverName</name>
<value>com.mysql.cj.jdbc.Driver</value>
</property>
<property>
<name>javax.jdo.option.ConnectionUserName</name>
<value>root</value>
</property>

<property>
<name>javax.jdo.option.ConnectionPassword</name>
<value>hayagreevan</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>

^G Help      ^O Write Out  ^W Where Is   ^K Cut        ^T Execute    ^C Location
^X Exit      ^R Read File  ^\ Replace    ^U Paste      ^J Justify    ^_ Go To Line
```

## 7. Download and Move mysql java connector to apache-hive-3.1.2-bin/lib

```

harish@fedora:~/hive/lib
commons-vfs2-2.1.jar
curator-client-2.12.0.jar
curator-framework-2.12.0.jar
curator-recipes-2.12.0.jar
datanucleus-api-jdo-4.2.4.jar
datanucleus-core-4.1.17.jar
datanucleus-rdbms-4.1.19.jar
derby-10.14.1.0.jar
disruptor-3.3.6.jar
dropwizard-metrics-hadoop-metrics2-reporter-0.1.2.jar
druid-hdfs-storage-0.12.0.jar
ecj-4.4.2.jar
esri-geometry-api-2.0.0.jar
findbugs-annotations-1.3.9-1.jar
flatbuffers-1.2.0-3f79e055.jar
groovy-all-2.4.11.jar
gson-2.2.4.jar
guava-19.0.jar
hbase-client-2.0.0-alpha4.jar
hbase-common-2.0.0-alpha4.jar
hk2-api-2.5.0-b32.jar
hk2-locator-2.5.0-b32.jar
hk2-utils-2.5.0-b32.jar
hppc-0.7.2.jar
htrace-core-3.2.0-incubating.jar
httpclient-4.5.2.jar
httpcore-4.4.4.jar
ivy-2.4.0.jar
jackson-annotations-2.9.5.jar
jackson-core-2.9.5.jar
jackson-core-asl-1.9.13.jar
jackson-databind-2.9.5.jar
jackson-dataformat-smile-2.9.5.jar
jackson-mapper-asl-1.9.13.jar
jamon-runtime-2.3.1.jar
janino-2.7.6.jar
javassist-3.20.0-GA.jar
javax.annotation-api-1.2.jar
javax.inject-2.5.0-b32.jar
javax.jdo-3.2.0-m3.jar
netty-buffer-4.1.17.Final.jar
netty-common-4.1.17.Final.jar
opencsv-2.3.jar
orc-core-1.5.6.jar
orc-shims-1.5.6.jar
org.abego.treelayout.core-1.0.1.jar
osgi-resource-locator-1.0.1.jar
paranamer-2.3.jar
parquet-hadoop-bundle-1.10.0.jar
php
postgresql-9.4.1208.jre7.jar
postgresql-metadata-storage-0.12.0.jar
protobuf-java-2.5.0.jar
py
sketches-core-0.9.0.jar
snappy-java-1.1.4.jar
sqlline-1.3.0.jar
ST4-4.0.4.jar
stax-api-1.0.1.jar
super-csv-2.2.0.jar

```

## 8. Execute schematool -initSchema -dbType mysql

```

lf4j-reload4j-1.7.36.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Metastore connection URL:      jdbc:mysql://localhost/metastore?createDatabaseIfNo
tExist=true
Metastore Connection Driver :   com.mysql.cj.jdbc.Driver
Metastore connection User:     root
Starting metastore schema initialization to 3.1.0
Initialization script hive-schema-3.1.0.mysql.sql

```

## 9. Start hive

```
harish@fedora:~/hive/lib$ hive
which: no hbase in (/home/harish/.local/bin:/home/harish/bin:/home/harish/pig/bin:/usr/lib/jvm/jdk1.8.0_202/bin:/home/harish/hive/bin:/home/harish/hive/bin)
Hive Session ID = 72f6d993-0190-4d45-bd88-71acc8c52db4

Logging initialized using configuration in jar:file:/home/harish/hive/lib/hive-logging.properties
Hive Session ID = 6d0c44f6-0f4e-4f77-8b7b-783351811c7d
Hive-on-MR is deprecated in Hive 2 and may not be available in future releases.
hive> 
```