Exp. No : 5

HIVE Installation

1. Download and Extract apache hive 3.1.2

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

2. Update HIVE Configurations in .bashrc

```
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/harithaah/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/harithaah/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

```
lksh@fedora:~/hive$ cd conf
lksh@fedora:~/hive/conf$ ls
beeline-log4j2.properties.template     hive-site.xml
hive-default.xml.template              ivysettings.xml
hive-env.sh                            llap-cli-log4j2.properties.template
hive-env.sh.template                   llap-daemon-log4j2.properties.template
hive-exec-log4j2.properties.template   parquet-logging.properties
hive-log4j2.properties.template
lksh@fedora:~/hive/conf$
```

4. Create hive-env.sh

```
GNU nano 7.2                                                              hive-env.sh
# if [ "$SERVICE" = "cli" ]; then
#   if [ -z "$DEBUG" ]; then
#     export HADOOP_OPTS="$HADOOP_OPTS -XX:NewRatio=12 -Xms10m -XX:MaxHeapFreeRatio=40 -XX:MinHeapFreeRatio=15 -XX:+UseParNewGC -XX:-UseGCOverheadLimit"
#   else
#     export HADOOP_OPTS="$HADOOP_OPTS -XX:NewRatio=12 -Xms10m -XX:MaxHeapFreeRatio=40 -XX:MinHeapFreeRatio=15 -XX:-UseGCOverheadLimit"
#   fi
# fi

# The heap size of the jvm stared 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=$HADOOP_HOME
export HIVE_CONF_DIR=$HIVE_CONF_DIR
```

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>
```

5. 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
```

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

```
lksh@fedora:~/hive/conf$ nano hive-env.sh
lksh@fedora:~/hive/conf$ cd ..
lksh@fedora:~/hive$ cd lib
lksh@fedora:~/hive/lib$ ls
accumulo-core-1.7.3.jar                 datanucleus-api-jdo-4.2.4.jar            hive-hcatalog-server-extensions-3.1.2.jar  javax.jdo-3.2.0-m3.jar
accumulo-fate-1.7.3.jar                 datanucleus-core-4.1.17.jar             hive-hplsql-3.1.2.jar                   javax.servlet-api-3.1.0.jar
accumulo-start-1.7.3.jar                datanucleus-rdbms-4.1.19.jar            hive-jdbc-3.1.2.jar                     javax.servlet.jsp-2.3.2.jar
accumulo-trace-1.7.3.jar                derby-10.14.1.0.jar                     hive-jdbc-handler-3.1.2.jar             javax.servlet.jsp-api-2.3.1.jar
aircompressor-0.10.jar                  disruptor-3.3.6.jar                     hive-kryo-registrator-3.1.2.jar         javax.ws.rs-api-2.0.1.jar
ant-1.9.1.jar                           dropwizard-metrics-hadoop-metrics2-reporter-0.1.2.jar  hive-llap-client-3.1.2.jar  javolution-5.5.1.jar
ant-launcher-1.9.1.jar                  druid-hdfs-storage-0.12.0.jar           hive-llap-common-3.1.2.jar              jaxb-api-2.2.11.jar
antlr4-runtime-4.5.jar                  ecj-4.4.2.jar                           hive-llap-common-3.1.2-tests.jar        jcodings-1.0.18.jar
antlr-runtime-3.5.2.jar                 esri-geometry-api-2.0.0.jar             hive-llap-ext-client-3.1.2.jar          jcommander-1.32.jar
aopalliance-repackaged-2.5.0-b32.jar    findbugs-annotations-1.3.9-1.jar        hive-llap-server-3.1.2.jar              jdo-api-3.0.1.jar
apache-curator-2.12.0.pom               flatbuffers-1.2.0-3f79e055.jar          hive-llap-tez-3.1.2.jar                 jersey-client-2.25.1.jar
apache-jsp-9.3.20.v20170531.jar         groovy-all-2.4.11.jar                   hive-metastore-3.1.2.jar                jersey-common-2.25.1.jar
apache-jstl-9.3.20.v20170531.jar        gson-2.2.4.jar                          hive-serde-3.1.2.jar                    jersey-container-servlet-core-2.25.1.ja
arrow-format-0.8.0.jar                  guava-19.0.jar                          hive-service-3.1.2.jar                  jersey-guava-2.25.1.jar
arrow-memory-0.8.0.jar                  hbase-client-2.0.0-alpha4.jar           hive-service-rpc-3.1.2.jar              jersey-media-jaxb-2.25.1.jar
arrow-vector-0.8.0.jar                  hbase-common-2.0.0-alpha4.jar           hive-shims-0.23-3.1.2.jar               jersey-server-2.25.1.jar
asm-5.0.1.jar                           hbase-common-2.0.0-alpha4-tests.jar     hive-shims-3.1.2.jar                    jettison-1.1.jar
asm-commons-5.0.1.jar                   hbase-hadoop2-compat-2.0.0-alpha4.jar   hive-shims-common-3.1.2.jar             jetty-annotations-9.3.20.v20170531.jar
asm-tree-5.0.1.jar                      hbase-hadoop2-compat-2.0.0-alpha4-tests.jar  hive-shims-scheduler-3.1.2.jar     jetty-client-9.3.20.v20170531.jar
audience-annotations-0.5.0.jar          hbase-hadoop-compat-2.0.0-alpha4.jar    hive-standalone-metastore-3.1.2.jar     jetty-http-9.3.20.v20170531.jar
avatica-1.11.0.jar                      hbase-http-2.0.0-alpha4.jar             hive-storage-api-2.7.0.jar              jetty-io-9.3.20.v20170531.jar
avro-1.7.7.jar                          hbase-mapreduce-2.0.0-alpha4.jar        hive-streaming-3.1.2.jar                jetty-jaas-9.3.20.v20170531.jar
bonecp-0.8.0.RELEASE.jar                hbase-metrics-2.0.0-alpha4.jar          hive-testutils-3.1.2.jar                jetty-jndi-9.3.20.v20170531.jar
calcite-core-1.16.0.jar                 hbase-metrics-api-2.0.0-alpha4.jar      hive-upgrade-acid-3.1.2.jar             jetty-plus-9.3.20.v20170531.jar
calcite-druid-1.16.0.jar                hbase-prefix-tree-2.0.0-alpha4.jar      hive-vector-code-gen-3.1.2.jar          jetty-rewrite-9.3.20.v20170531.jar
calcite-linq4j-1.16.0.jar               hbase-procedure-2.0.0-alpha4.jar        hk2-api-2.5.0-b32.jar                   jetty-runner-9.3.20.v20170531.jar
commons-cli-1.2.jar                     hbase-protocol-2.0.0-alpha4.jar         hk2-locator-2.5.0-b32.jar               jetty-schemas-3.1.jar
commons-codec-1.7.jar                   hbase-protocol-shaded-2.0.0-alpha4.jar  hk2-utils-2.5.0-b32.jar                 jetty-security-9.3.20.v20170531.jar
commons-collections4-4.1.jar            hbase-replication-2.0.0-alpha4.jar      hppc-0.7.2.jar                          jetty-server-9.3.20.v20170531.jar
commons-compiler-2.7.6.jar              hbase-server-2.0.0-alpha4.jar           htrace-core-3.2.0-incubating.jar        jetty-servlet-9.3.20.v20170531.jar
commons-compress-1.9.jar                hbase-shaded-miscellaneous-1.0.1.jar    httpclient-4.5.2.jar                    jetty-util-9.3.20.v20170531.jar
commons-crypto-1.0.0.jar                hbase-shaded-netty-1.0.1.jar            httpcore-4.4.4.jar                      jetty-webapp-9.3.20.v20170531.jar
commons-dbcp-1.4.jar                    hbase-shaded-protobuf-1.0.1.jar         ivy-2.4.0.jar                           jetty-xml-9.3.20.v20170531.jar
commons-io-2.4.jar                      HikariCP-2.6.1.jar                      jackson-annotations-2.9.5.jar           jline-2.12.jar
commons-lang-2.6.jar                    hive-accumulo-handler-3.1.2.jar         jackson-core-2.9.5.jar                  joda-time-2.9.9.jar
commons-lang3-3.2.jar                   hive-beeline-3.1.2.jar                  jackson-core-asl-1.9.13.jar             joni-2.1.11.jar
commons-logging-1.0.4.jar               hive-classification-3.1.2.jar           jackson-databind-2.9.5.jar              jpam-1.1.jar
commons-math-2.1.jar                    hive-cli-3.1.2.jar                      jackson-dataformat-smile-2.9.5.jar      json-1.8.jar
commons-math3-3.6.1.jar                 hive-common-3.1.2.jar                   jackson-mapper-asl-1.9.13.jar           jsr305-3.0.0.jar
commons-pool-1.5.4.jar                  hive-contrib-3.1.2.jar                  jamon-runtime-2.3.1.jar                 jta-1.1.jar
commons-vfs2-2.1.jar                    hive-druid-handler-3.1.2.jar            janino-2.7.6.jar                        libfb303-0.9.3.jar
curator-client-2.12.0.jar               hive-exec-3.1.2.jar                     javassist-3.20.0-GA.jar                 libthrift-0.9.3.jar
curator-framework-2.12.0.jar            hive-hbase-handler-3.1.2.jar            javax.annotation-api-1.2.jar            log4j-1.2-api-2.10.0.jar
curator-recipes-2.12.0.jar              hive-hcatalog-core-3.1.2.jar            javax.inject-2.5.0-b32.jar              log4j-api-2.10.0.jar
lksh@fedora:~/hive/lib$
```

7. 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
```

8. Start hive