Ex.no:	HIVE INSTALLATION
Date:	

Aim:

To install hive on top of a Hadoop single node cluster and setup hive shell.

Procedure:

1. Download the hive tar file and extract it .

```
edureka@localhost:~$ tar -xzf apache-hive-2.1.0-bin.tar.gz
edureka@localhost:~$ ls
apache-hive-2.1.0-bin          Documents          Music      Templates
apache-hive-2.1.0-bin.tar.gz   Downloads          Pictures   Videos
Desktop                        examples.desktop   Public
edureka@localhost:~$
```

2. Edit the bashrc file and update the Hive and Hadoop environment variables.

```
# Set Hadoop-related environment variables

export HADOOP_HOME=/home/edureka/hadoop-2.7.3
export HADOOP_CONF_DIR=/home/edureka/hadoop-2.7.3/etc/hadoop
export HADOOP_MAPRED_HOME=/home/edureka/hadoop-2.7.3
export HADOOP_COMMON_HOME=/home/edureka/hadoop-2.7.3
export HADOOP_HDFS_HOME=/home/edureka/hadoop-2.7.3
export YARN_HOME=/home/edureka/hadoop-2.7.3
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib"

# Set JAVA_HOME

export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-i386
export PATH=$PATH:/usr/lib/jvm/java-8-openjdk-i386/bin

# Add Hadoop bin/ directory to PATH

export PATH=$PATH:/home/edureka/hadoop-2.7.3/bin
export HADOOP_PID_DIR=/home/edureka/hadoop-2.7.3/hadoop2_data/hdfs/pid
```

3. Use source .bashrc command to apply the changes in the environment variables.
4. Check the hive version

```
edureka@localhost:~$ hive --version
Hive 2.1.0
Subversion git://jcamachguezrMBP/Users/jcamachorodriguez/src/workspaces/
hive/HIVE-release2/hive -r 9265bc24d75ac945bde9ce1a0999fddd8f2aae29
Compiled by jcamachorodriguez on Fri Jun 17 01:03:25 BST 2016
From source with checksum 1f896b8fae57fbd29b047d6d67b75f3c
edureka@localhost:~$
```

5. Create Hive directories within HDFS . The directory ‘warehouse’ is the location to store the table or data related to hive.

   hdfs dfs -mkdir -p /user/hive/warehouse

   hdfs dfs -mkdir /tmp

6. Set read and write permissions for the created directories.

   hdfs dfs -chmod g+w /user/hive/warehouse
   hdfs dfs -chmod g+w /tmp

7. Set the Hadoop path in hive-env.sh file

   cd apache-hive-2.1.0-bin/

   gedit conf/hive-env.sh

```
# Set HADOOP_HOME to point to a specific hadoop install directory
export HADOOP_HOME=/home/edureka/hadoop-2.7.3

export HADOOP_HEAPSIZE=512

# Hive Configuration Directory can be controlled by:
export HIVE_CONF_DIR=/home/edureka/apache-hive-2.1.0-bin/conf
```

8. Edit the hive-site.xml file

9. Initialise the derby database

   bin/schematool -initSchema -dbType derby

```
edureka@localhost:~$ cd apache-hive-2.1.0-bin/
edureka@localhost:~/apache-hive-2.1.0-bin$ bin/schematool -initSchema -d
bType derby
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/edureka/apache-hive-2.1.0-bin/li
b/log4j-slf4j-impl-2.4.1.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/edureka/hadoop-2.7.3/share/hadoo
p/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder
.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an expl
anation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFa
ctory]
Metastore connection URL:        jdbc:derby:;databaseName=/home/edureka/
apache-hive-2.1.0-bin/metastore_db;create=true
Metastore Connection Driver :    org.apache.derby.jdbc.EmbeddedDriver
Metastore connection User:       APP
Starting metastore schema initialization to 2.1.0
Initialization script hive-schema-2.1.0.derby.sql
Initialization script completed
schemaTool completed
edureka@localhost:~/apache-hive-2.1.0-bin$
```

10. Launch hive

```
edureka@localhost:~/apache-hive-2.1.0-bin$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/edureka/apache-hive-2.1.0-bin/li
b/log4j-slf4j-impl-2.4.1.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/edureka/hadoop-2.7.3/share/hadoo
p/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder
.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an expl
anation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFa
ctory]

Logging initialized using configuration in jar:file:/home/edureka/apache
-hive-2.1.0-bin/lib/hive-common-2.1.0.jar!/hive-log4j2.properties Async:
 true
Hive-on-MR is deprecated in Hive 2 and may not be available in the futur
e versions. Consider using a different execution engine (i.e. spark, tez
) or using Hive 1.X releases.
hive>
```

Result:

Thus the hive was installed on top of Hadoop single node cluster and hive shell was successfully launched.