

EXP NO: 5

INSTALLATION OF HIVE

\$wget <https://archive.apache.org/dist/hive/hive-3.1.2/apache-hive-3.1.2-bin.tar.gz>

\$tar xvzf apache-hive-3.1.2-bin.tar.gz

```
hadoop@ubuntu:~$ cd
hadoop@ubuntu:~$ wget https://archive.apache.org/dist/hive/hive-3.1.2/apache-hive-3.1.2-bin.tar.gz
--2024-09-18 10:46:01-- https://archive.apache.org/dist/hive/hive-3.1.2/apache-hive-3.1.2-bin.tar.gz
Resolving archive.apache.org (archive.apache.org)... 2a01:4f9:1a:a084::2, 65.108.204.189
Connecting to archive.apache.org (archive.apache.org)|2a01:4f9:1a:a084::2|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 278813740 (266M) [application/x-gzip]
Saving to: 'apache-hive-3.1.2-bin.tar.gz'

apache-hive-3.1.2-bin.tar.gz      100%[=====] 265.90M   801KB/s   in 7m 7s

2024-09-18 10:53:11 (637 KB/s) - 'apache-hive-3.1.2-bin.tar.gz' saved [278813740/278813740]

hadoop@ubuntu:~$ ^[[260-tar xvzf apache-hive-3.1.2-bin.tar.gz-
tar: command not found
hadoop@ubuntu:~$ tar xvzf apache-hive-3.1.2-bin.tar.gz
apache-hive-3.1.2-bin/LICENSE
apache-hive-3.1.2-bin/NOTICE
apache-hive-3.1.2-bin/RELEASE_NOTES.txt
apache-hive-3.1.2-bin/binary-package-licenses/asn-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/com.google.protobuf-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/com.ibm.icu.icu4j-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/com.sun.jersey-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/com.thoughtworks.paranamer-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/javax.transaction.transaction-api-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/javolution-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/jline-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/NOTICE
apache-hive-3.1.2-bin/binary-package-licenses/org.abego.treelayout.core-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org antlr-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org antlr/antlr4-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org antlr/stringtemplate-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org.codehaus.janino-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org.jamon.jamon-runtime-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org.jruby-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org.mozilla.rhino-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org.sif4j-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/org.solline-LICENSE
apache-hive-3.1.2-bin/examples/files/2000_cols_data.csv
apache-hive-3.1.2-bin/examples/files/3col_data.txt
apache-hive-3.1.2-bin/examples/files/4col_data.txt
apache-hive-3.1.2-bin/examples/files/5col_data.txt
apache-hive-3.1.2-bin/examples/files/aggr_01-p1.txt
apache-hive-3.1.2-bin/examples/files/aggr_01-p2.txt
apache-hive-3.1.2-bin/examples/files/aggr_01-p3.txt
```

\$mv apache-hive-3.1.2-bin hive

\$nano ~/.bashrc

```
export HIVE_HOME="/home/haresh/apache-hive-3.1.2-bin"
export HIVE_CONF_DIR=$HIVE_HOME/conf
export PATH=$PATH:$HIVE_HOME/bin
export HADOOP_USER_CLASSPATH_FIRST=true
```

\$source ~/.bashrc

\$cd hive

\$cp conf/hive-default.xml.template conf/hive-site.xml

\$nano conf/hive-site.xml

```
export HADOOP_HOME=$HADOOP_HOME
export HIVE_CONF_DIR=$HIVE_CONF_DIR
```

\$start-all.sh

\$ jps

```

apache-hive-3.1.2-bin/hcatalog/share/website/java-client/hive-weibcat-java-client-3.1.2.jar
hadoop@ubuntu:~$ mv apache-hive-3.1.2 hive
mv: cannot stat 'apache-hive-3.1.2': No such file or directory
hadoop@ubuntu:~$ ls
dataset.txt  demo_pig.pig  Downloads  example.pig  hadoop-3.3.6  mapper_data.py  Pictures  Public  sample.txt  uppercase.py
hadoop@ubuntu:~$ ls
dataset.txt  demo_pig.pig  Downloads  example.pig  hadoop-3.3.6  mapper_data.py  Pictures  Public  sample.txt  uppercase.py
hadoop@ubuntu:~$ mv apache-hive-3.1.2-bin.tar.gz Desktop
hadoop@ubuntu:~$ mv apache-hive-3.1.2-bin.tar.gz Documents
hadoop@ubuntu:~$ mv /path/to/apache-hive-3.1.2 /path/to/hive
mv: cannot stat '/path/to/apache-hive-3.1.2': No such file or directory
hadoop@ubuntu:~$ ls
dataset.txt  demo_pig.pig  Downloads  example.pig  hadoop-3.3.6  mapper_data.py  Pictures  Public  sample.txt  uppercase.py
hadoop@ubuntu:~$ mv apache-hive-3.1.2-bin.tar.gz Desktop
hadoop@ubuntu:~$ mv apache-hive-3.1.2-bin.tar.gz Documents
hadoop@ubuntu:~$ mv apache-hive-3.1.2-bin hive
mv: cannot stat 'apache-hive-3.1.2': No such file or directory
hadoop@ubuntu:~$ mv apache-hive-3.1.2-bin hive
hadoop@ubuntu:~$ nano ~/.bashrc
hadoop@ubuntu:~$ source ~/.bashrc
hadoop@ubuntu:~$ cd hive
hadoop@ubuntu:~/hive$ cp conf/hive-default.xml.template conf/hive-site.xml
hadoop@ubuntu:~/hive$ nano conf/hive-site.xml
hadoop@ubuntu:~/hive$ hdfs dfs -mkdir /user
mkdir: call From ubuntu/127.0.1.1 to localhost:9000 failed on connection exception: java.net.ConnectException: Connection refused; For more details see: http://wiki.apache.org/hadoop/ConnectionRefused
hadoop@ubuntu:~/hive$ hdfs dfs -mkdir /user/hive
mkdir: call From ubuntu/127.0.1.1 to localhost:9000 failed on connection exception: java.net.ConnectException: Connection refused; For more details see: http://wiki.apache.org/hadoop/ConnectionRefused
hadoop@ubuntu:~/hive$ jps
5238 Jps
hadoop@ubuntu:~/hive$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [ubuntu]
Starting resourcemanager
Starting nodemanagers
hadoop@ubuntu:~/hive$ jps
6648 Jps
5635 DataNode
5927 SecondaryNameNode
6285 NodeManager
6464 NameNode

```

\$hdfs dfs -mkdir /user

\$hdfs dfs -mkdir /user/hive

\$hdfs dfs -mkdir /user/hive/warehouse

\$hdfs dfs -mkdir /user/tmp

\$hdfs dfs -chmod g+w /user/tmp

\$hdfs dfs -chmod g+w /user/hive/warehouse

\$cd \$HIVE_HOME/bin

\$schematool -initSchema -dbType derby

```

WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [ubuntu]
Starting resourcemanager
Starting nodemanagers
hadoop@ubuntu:~/hive$ jps
6648 Jps
5635 DataNode
5927 SecondaryNameNode
6285 NodeManager
5454 NameNode
6142 ResourceManager
hadoop@ubuntu:~/hive$ hdfs dfs -mkdir /user
hadoop@ubuntu:~/hive$ hdfs dfs -mkdir /user/hive
hadoop@ubuntu:~/hive$ hdfs dfs -mkdir /user/hive/warehouse
hadoop@ubuntu:~/hive$ hdfs dfs -mkdir /user/tmp
hadoop@ubuntu:~/hive$ hdfs dfs -chmod g+w /user/tmp
hadoop@ubuntu:~/hive$ hdfs dfs -chmod g+w /user/hive/warehouse
hadoop@ubuntu:~/hive$ cd $HIVE_HOME/bin
hadoop@ubuntu:~/hive/bin$ schematool -initSchema -dbType derby
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/hadoop/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/hadoop/hadoop/share/hadoop/common/lib/slf4j-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:derby:databaseName=metastore_db;create=true
Metastore Connection Driver : org.apache.derby.jdbc.EmbeddedDriver
Metastore connection User: APP
Starting metastore schema initialization to 3.1.0
Initialization script hive-schema-3.1.0.derby.sql

```

\$hive

>show databases;

> CREATE DATABASE financials;

```

hive> CREATE DATABASE financials;
OK
Time taken: 0.41 seconds

```

> use financials;

```

hive> use financials;
OK
Time taken: 0.031 seconds

```

> CREATE TABLE finance_table(id INT, name STRING);

```

hive> CREATE TABLE finance_table (id INT, name STRING);
OK
Time taken: 0.552 seconds

```

> INSERT INTO finance_table VALUES (1, 'Alice'), (2, 'Bob'), (3, 'Charlie');

```

hive> INSERT INTO finance_table VALUES (1, 'Alice'), (2, 'Bob'), (3, 'Charlie');
Query ID = haresh_20240912212701_f51e1fcd-d01a-4b83-8d6e-bf0388e6e714
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1726155698102_0001, Tracking URL = http://fedora:8088/proxy
2_0001/

```

> CREATE VIEW myview AS SELECT name, id FROM finance_table;

```
hive> CREATE VIEW myview AS SELECT name, id FROM finance_table;  
OK  
Time taken: 0.256 seconds
```

> SELECT * FROM myview;

```
hive> SELECT * FROM myview;  
OK  
Alice    1  
Bob      2  
Charlie  3  
Time taken: 0.213 seconds, Fetched: 3 row(s)
```

> DESCRIBE finance_table;

```
hive> DESCRIBE finance_table;  
OK  
id                int  
name              string  
Time taken: 0.075 seconds, Fetched: 2 row(s)
```

> ALTER TABLE finance_table ADD COLUMNS (age INT);

```
hive> ALTER TABLE finance_table ADD COLUMNS (age INT);  
OK  
Time taken: 0.126 seconds  
hive> quit;
```

>quit;