**EXP NO: 5**                          **INSTALLATION OF HIVE**
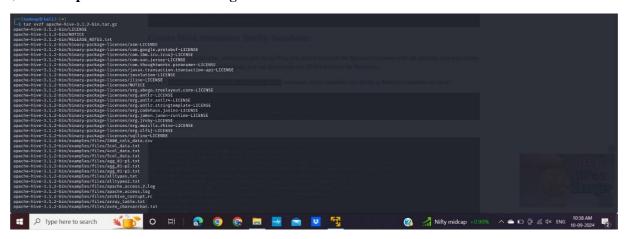
**$wget https://archive.apache.org/dist/hive/hive-3.1.2/apache-hive-3.1.2-bin.tar.gz**

```
karthickragav@fedora:~$ wget https://archive.apache.org/dist/hive/hive-3.1.2/apache-hive-3.1.2-bin.tar.gz
apache-hive-3.1.2-bi    5% [==>                                                    ]   14.67M  857.89KB/s
                        [Files: 0  Bytes: 0  [0 B/s] Re]
```

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

```
┌──(hadoop㉿kali)-[~]
└─$ 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/asm-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.slf4j-LICENSE
apache-hive-3.1.2-bin/binary-package-licenses/sqlline-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/agg_01-p1.txt
apache-hive-3.1.2-bin/examples/files/agg_01-p2.txt
apache-hive-3.1.2-bin/examples/files/agg_01-p3.txt
apache-hive-3.1.2-bin/examples/files/alltypes.txt
apache-hive-3.1.2-bin/examples/files/alltypes2.txt
apache-hive-3.1.2-bin/examples/files/apache.access.2.log
apache-hive-3.1.2-bin/examples/files/apache.access.log
apache-hive-3.1.2-bin/examples/files/archive_corrupt.rc
apache-hive-3.1.2-bin/examples/files/array_table.txt
apache-hive-3.1.2-bin/examples/files/avro_charvarchar.txt
```

**$mv apache-hive-3.1.2-bin hive**

**$nano ~/.bashrc**

**$source ~/.bashrc**

**$cd hive**

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

```
karthickragav@fedora:~ — nano conf/hive-site.xml

karthickragav@fedora:~$ nano conf/hive-site.xml
karthickragav@fedora:~$
```

**$nano conf/hive-site.xml**

**$start-all.sh**



```
karthickragav@fedora:~$ nano conf/hive-site.xml
karthickragav@fedora:~$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as karthickragav 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 [fedora]
Starting resourcemanager
Starting nodemanagers
karthickragav@fedora:~$
```

**$ jps**

```
karthickragav@fedora:~$ jps
4130 SecondaryNameNode
4611 NodeManager
5030 Jps
3705 NameNode
3899 DataNode
4479 ResourceManager
karthickragav@fedora:~$
```

**$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**

```
karthickragav@fedora:~$ hdfs dfs -ls /user
Found 2 items
drwxr-xr-x   - karthickragav supergroup          0 2024-09-12 12:16 /user/bin
drwxr-xr-x   - karthickragav supergroup          0 2024-09-18 12:29 /user/hive
karthickragav@fedora:~$
```
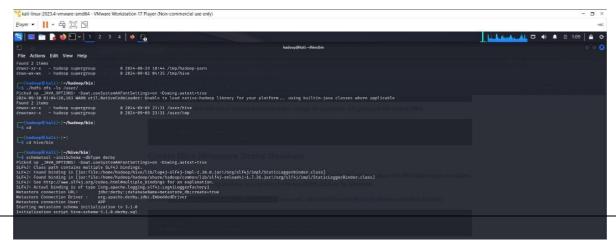
**$cd $HIVE_HOME/bin**

**$schematool -initSchema –dbType derby**

```
karthickragav@fedora:~/hive/bin$ hive
which: no hbase in (/home/karthickragav/.local/bin:/home/karthickragav/bin:/usr/local/bin:/usr/local/sbin:/us
bin:/usr/sbin:/home/karthickragav/pig/bin:/home/karthickragav/hadoop/sbin:/home/karthickragav/hadoop/bin:/hom
karthickragav/hive/bin:/home/karthickragav/pig/bin:/home/karthickragav/hadoop/sbin:/home/karthickragav/hadoop
in:/home/karthickragav/hive/bin)
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/karthickragav/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/St
icLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/karthickragav/hadoop/share/hadoop/common/lib/slf4j-reload4j-1.7.36.ja
/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]
Hive Session ID = e738ec85-4469-46db-a955-e21e70debcdd

Logging initialized using configuration in jar:file:/home/karthickragav/hive/lib/hive-common-3.1.2.jar!/hive-
g4j2.properties Async: true
```

**$hive**

**>show databases;**

```
karthickragav@fedora:~/hive/bin$ hive
which: no hbase in (/home/karthickragav/.local/bin:/home/karthickragav/bin:/usr/local/bin:/usr/local/sbin:/usr/bin:/usr/sbin:/home/karthickrag
av/pig/bin:/home/karthickragav/hadoop/sbin:/home/karthickragav/hadoop/bin:/home/karthickragav/hive/bin:/home/karthickragav/pig/bin:/home/karth
ickragav/hadoop/sbin:/home/karthickragav/hadoop/bin:/home/karthickragav/hive/bin)
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/karthickragav/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/karthickragav/hadoop/share/hadoop/common/lib/slf4j-reload4j-1.7.36.jar!/org/slf4j/impl/StaticLoggerBin
der.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]
Hive Session ID = e738ec85-4469-46db-a955-e21e70debcdd

Logging initialized using configuration in jar:file:/home/karthickragav/hive/lib/hive-common-3.1.2.jar!/hive-log4j2.properties Async: true
Hive Session ID = 02954020-f3d1-40c5-a16f-d647eb366b17
Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, t
ez) or using Hive 1.X releases.
hive> show databases
    > financials
```

**> CREATE DATABASE financials;**

```
hive> CREATE DATABASE financials;
OK
Time taken: 0.591 seconds
```

**> use financials;**

```
hive> use financials;
OK
Time taken: 0.229 seconds
```

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

```
hive> CREATE TABLE finance_table( id INT, name STRING );
OK
Time taken: 2.536 seconds
```

**> INSERT INTO finance_tableVALUES (1, 'Alice'), (2, 'Bob'), (3, 'Charlie');**

```
hive> INSERT INTO finance_table VALUES (1, 'Alice'), (2, 'Bob'), (3, 'Charlie');
Query ID = hadoop_20240910011318_759e97f4-dbcc-4020-a345-37b6aac0dbbf
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_1725938931910_0001, Tracking URL = http://kali:8088/proxy/application_1725938931910_0001/
Kill Command = /home/hadoop/hadoop/bin/mapred job  -kill job_1725938931910_0001
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2024-09-10 01:14:33,477 Stage-1 map = 0%,   reduce = 0%
2024-09-10 01:15:17,392 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 20.45 sec
2024-09-10 01:15:43,645 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 33.61 sec
MapReduce Total cumulative CPU time: 33 seconds 610 msec
Ended Job = job_1725938931910_0001
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to directory hdfs://localhost:9000/user/hive/warehouse/financials.db/finance_table/.hive-staging_hive_2024-09-10_01-13-18_205_31425232155846852-1/-ext-10000
Loading data to table financials.finance_table
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 33.61 sec   HDFS Read: 15668 HDFS Write: 291 SUCCESS
Total MapReduce CPU Time Spent: 33 seconds 610 msec
OK
Time taken: 151.732 seconds
```

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

```
hive> CREATE VIEW myview AS SELECT name, id FROM finance_table;
OK
Time taken: 0.82 seconds
```

**> SELECT * FROM myview;**

```
hive> SELECT * FROM myview;
OK
Alice    1
Bob      2
Charlie 3
Time taken: 0.716 seconds, Fetched: 3 row(s)
hive>
```

**> DESCRIBE finance_table;**

```
hive> DESCRIBE finance_table;
OK
id                      int
name                    string
Time taken: 0.348 seconds, Fetched: 2 row(s)
```

**> ALTER TABLE finance_table ADD COLUMNS (age INT);**

**>quit;**

```
hive> ALTER TABLE finance_table ADD COLUMNS (age INT);
OK
Time taken: 0.457 seconds
hive> quit;
```