

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>

```
kali-linux-2023.4-vmware-amd64 - VMware Workstation 17 Player (Non-commercial use only)
Player
File Actions Edit View Help
[hadoop@kali: ~]$ wget https://archive.apache.org/dist/hive/hive-3.1.2/apache-hive-3.1.2-bin.tar.gz
--2024-09-10 08:48:34-- 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:a08::2, 65.180.265.189
Connecting to archive.apache.org (archive.apache.org)|2a01:4f9:1a:a08::2|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 276813748 (266M) [application/gzip]
Saving to: 'apache-hive-3.1.2-bin.tar.gz'

apache-hive-3.1.2-bin.tar.gz 100% [====>] 265.98M 1.67MB/s in 4m 21s

2024-09-10 08:44:57 (1.02 MB/s) - 'apache-hive-3.1.2-bin.tar.gz' saved [276813748/276813748]
```

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

```
[hadoop@kali: ~]$ tar xvfz 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.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/java.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/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.chimino-LICENSE
apache-hive-3.1.2-bin/binary-package/licenses/org.sifaj-LICENSE
apache-hive-3.1.2-bin/binary-package/licenses/sqlite-LICENSE
apache-hive-3.1.2-bin/examples/files/2000_colt_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/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

```
[hadoop@kali: ~]$ cd hive
[hadoop@kali: ~/hive]$ $ {system:java.io.tmpdir}
[hadoop@kali: ~/hive]$ cp conf/hive-default.xml.template conf/hive-site.xml
[hadoop@kali: ~/hive]$ nano conf/hive-site.xml
```

\$nano conf/hive-site.xml

\$start-all.sh

```

(hadoop@kali)~$ 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 [kali]
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-11 04:59:16,429 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting resourcemanager
Starting nodemanagers

```

\$ jps

```

(hadoop@kali)~$ jps
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
14436 NodeManager
16772 Jps
13830 SecondaryNameNode
14311 ResourceManager
13597 DataNode
13471 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

```

(hadoop@kali)~/hadoop/bin$ ./hdfs dfs -ls /user/
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-10 01:04:28,163 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
Found 2 items
drwxr-xr-x - hadoop supergroup 0 2024-09-09 23:31 /user/hive
drwxrwxr-x - hadoop supergroup 0 2024-09-09 23:31 /user/tmp

```

\$cd \$HIVE_HOME/bin

\$schematool -initSchema -dbType derby

```

kali-linux-2023.4-vmware-amd64 - VMware Workstation 17 Player (Non-commercial use only)
Player
(hadoop@kali)~/hive/bin$ ./schematool -initSchema -dbType derby
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
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-reloadable-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: root
Starting metastore schema initialization to 3.1.0
Initialization script hive-schema-3.1.0.derby.sql

```

```
>show databases;
```

```

[hadoop@kali:~/hadoop]$ ./bin/hive
hive
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
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/apache/log4j/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/hadoop/hadoop/share/hadoop/common/lib/slf4j-log4j-1.7.36.jar/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://maven.apache.org/dependency-resolving.html for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Hive Session ID = e9f8de15-ab67-4906-87c6-16d66db8c54

Logging initialized using configuration in jar:file:/home/hadoop/hive/lib/hive-common-3.12.2.jar!/hive-loggtj.properties Async: true
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, tez) or using Hive 1.X releases.
Hive Session ID = e21299ac-dfed-4c43-b2db-98719db2a5a
hive> show databases;
OK
default
Time taken: 3.928 seconds, Fetched: 1 row(s)
hive>

```

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

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

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

```
hive> INSERT INTO finance_table VALUES (1, 'Alice'), (2, 'Bob'), (3, 'Charlie');
Query ID = hadoop_20240910011318_759e97fa-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.reducers=<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 01:16:33,477 Stage-1 map = 0%, reduce = 0%
2024-09-10 01:16:37,992 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-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;
```