

Apache Hive Installation Steps

Step 1: Prerequisites

- Install Java (JDK 8 or 11)
- Install Hadoop (single-node or cluster)
- Enable SSH if cluster setup

Step 2: Download Hive

```
cd /usr/local/  
wget https://dlcdn.apache.org/hive/hive-4.0.0/apache-hive-4.0.0-bin.tar.gz  
tar -xvzf apache-hive-4.0.0-bin.tar.gz  
mv apache-hive-4.0.0-bin hive
```

Step 3: Set Environment Variables

Edit `~/.bashrc` and add:

```
export HIVE_HOME=/usr/local/hive  
export PATH=$PATH:$HIVE_HOME/bin  
export HADOOP_HOME=/usr/local/hadoop  
export PATH=$PATH:$HADOOP_HOME/bin:$HADOOP_HOME/sbin  
source ~/.bashrc
```

Step 4: Configure Hive

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

```
cd $HIVE_HOME/conf  
cp hive-default.xml.template hive-site.xml
```

Edit hive-site.xml with:

```
javax.jdo.option.ConnectionURL = jdbc:derby::databaseName=metastore_db;create=true  
hive.metastore.warehouse.dir = /user/hive/warehouse
```

Step 5: Configure Metastore (Optional MySQL)

```
sudo apt-get install libmysql-java  
cp /usr/share/java/mysql-connector-java.jar $HIVE_HOME/lib/
```

Update hive-site.xml with MySQL configs:

URL: jdbc:mysql://localhost:3306/hive_metastore?createDatabaseIfNotExist=true

Driver: com.mysql.jdbc.Driver

User: hiveuser

Password: hivepassword

Initialize schema:

```
schematool -initSchema -dbType mysql
```

Step 6: Start Hive

start-dfs.sh

start-yarn.sh

hive

Step 7: Test Hive

```
CREATE DATABASE testdb;  
USE testdb;  
CREATE TABLE emp(id INT, name STRING);  
INSERT INTO emp VALUES (1, 'Praveen'), (2, 'Amit');  
SELECT * FROM emp;
```

Hive installation is complete!