Because Hive metastore default use Derby embed

This is only one session can connect, no connection pool, so use postgresql instead.

https://www.digitalocean.com/community/tutorials/how-to-install-and-use-postgresql-on-ubuntu-14-04

sudo apt-get install postgresql postgresql-contrib

sudo -i -u postgres

sudo usermod -a -G sudo postgres

In postgres shell

su – postgres

type command

* psql –U postgres
* create a hive user
* Create user hive;
* Create database metastore\_db with owner=hive;
* ALTER ROLE postgres WITH PASSWORD 'postgres'; // superuser set passwd the same
* ALTER ROLE hive WITH PASSWORD 'hive';
* GRANT ALL PRIVILEGES ON DATABASE metastore\_db TO hive;
* change postgresql listen ip address

vim /etc/postgresql/9.3/main/postgresql.conf

**listen\_addresses = ‘\*’**

vim /etc/postgresql/9.3/main/pg\_hba.conf

add line

**host all all <ip\_address>/32 md5**

sudo /etc/init.d/postgresql restart

Hive install

download postgres jdbc driver

<http://jdbc.postgresql.org/download.html>

<https://jdbc.postgresql.org/download/postgresql-9.3-1102.jdbc41.jar>

put this JDBC jar to **hive lib folder**

in hive-site.xml

<property>

<name>javax.jdo.option.ConnectionURL</name>

<value>jdbc:postgresql://db\_ip:db\_port/db\_name</value>

<description>JDBC connect string for a JDBC metastore</description>

</property>

<property>

<name>javax.jdo.option.ConnectionDriverName</name>

<value>org.postgresql.Driver</value>

<description>Driver class name for a JDBC metastore</description>

</property>

<property>

<name>javax.jdo.option.ConnectionUserName</name>

<value>db\_user[hive]</value>

<description>username to use against metastore database</description>

</property>

<property>

<name>javax.jdo.option.ConnectionPassword</name>

<value>db\_pwd[hive]</value>

<description>password to use against metastore database</description>

</property>

<property>

<name>hive.exec.local.scratchdir</name>

<value>/home/hadoopuser/apache-hive-0.14.0-bin/tmp</value>

<description>Local scratch space for Hive jobs</description>

</property>

<property>

<name>hive.downloaded.resources.dir</name>

<value>/home/hadoopuser/apache-hive-0.14.0-bin/tmp/${hive.session.id}\_resources</value>

<description>Temporary local directory for added resources in the remote file system.</description>

</property>

<property>

<name>hive.querylog.location</name>

<value>/home/hadoopuser/apache-hive-0.14.0-bin/tmp</value>

<description>Location of Hive run time structured log file</description>

</property>

<property>

<name>hive.server2.logging.operation.log.location</name>

<value>/home/hadoopuser/apache-hive-0.14.0-bin/tmp/operation\_logs</value>

<description>Top level directory where operation logs are stored if logging functionality is enabled</description>

</property>

<property>

<name>datanucleus.autoCreateSchema</name>

<value>false</value>

<description>creates necessary schema on a startup if one doesn't exist. set this to false, after creating it once</description>

</property>

<property>

<name>hive.exec.dynamic.partition.mode</name>

<value>nonstrict</value>

<description>

In strict mode, the user must specify at least one static partition

in case the user accidentally overwrites all partitions.

In nonstrict mode all partitions are allowed to be dynamic.

</description>

</property>

<property>

<name>hive.optimize.index.filter</name>

<value>true</value>

<description>Whether to enable automatic use of indexes</description>

</property>

<property>

<name>hive.optimize.index.autoupdate</name>

<value>true</value>

<description>Whether to update stale indexes automatically</description>

</property>

<!—Attention for test only we use 0 instead to direct use index- default is 5G ->

**<property>**

**<name>hive.optimize.index.filter.compact.minsize</name>**

**<value>0</value>**

**<!--<value>5368709120</value>-->**

**<description>Minimum size (in bytes) of the inputs on which a compact index is automatically used.</description>**

**</property>**

<property>

<name>hive.optimize.index.groupby</name>

<value>true</value>

<description>Whether to enable optimization of group-by queries using Aggregate indexes.</description>

</property>

cp conf /log4j.properties.template to log4j.porperties

modify log dir = /home/hadoopuser/apache-hive-0.14.0-bin/logs

Find file:

apache-hive-0.14.0-bin/scripts/metastore/upgrade/postgres/hive-schema-0.14.0.postgres.sql

type bellow command

psql -U hive -h master -d metastore\_db -f hive-schema-0.14.0.postgres.sql

In bashrc

export HIVE\_HOME=/home/hadoopuser/apache-hive-0.14.0-bin

export HIVE\_CONF\_DIR=$HIVE\_HOME/conf

export PATH=$HIVE\_HOME/bin:$PATH

type command

hive

test hive status

create table test(id INT,str STRING);

show tables;

drop table test;

if ok it is write

exit command line hive

Launch hive server 2

command:

**nohup hive --service hiveserver2 &**

Update and delete section

If we want to use update and delete feature

<https://cwiki.apache.org/confluence/display/Hive/Hive+Transactions#HiveTransactions-NewConfigurationParametersforTransactions>

Need to follow above steps

* [hive.support.concurrency](https://cwiki.apache.org/confluence/display/Hive/Configuration+Properties#ConfigurationProperties-hive.support.concurrency) – true
* [hive.enforce.bucketing](https://cwiki.apache.org/confluence/display/Hive/Configuration+Properties#ConfigurationProperties-hive.enforce.bucketing) – true
* [hive.exec.dynamic.partition.mode](https://cwiki.apache.org/confluence/display/Hive/Configuration+Properties#ConfigurationProperties-hive.exec.dynamic.partition.mode) – nonstrict
* [hive.txn.manager](https://cwiki.apache.org/confluence/display/Hive/Configuration+Properties#ConfigurationProperties-hive.txn.manager) – org.apache.hadoop.hive.ql.lockmgr.DbTxnManager
* [hive.compactor.initiator.on](https://cwiki.apache.org/confluence/display/Hive/Configuration+Properties#ConfigurationProperties-hive.compactor.initiator.on) – true (for exactly one instance of the Thrift metastore service)
* [hive.compactor.worker.threads](https://cwiki.apache.org/confluence/display/Hive/Configuration+Properties#ConfigurationProperties-hive.compactor.worker.threads) – a positive number on at least one instance of the Thrift metastore service

<property>

<name>hive.zookeeper.quorum</name>

<value>master,slave1,idoharness</value>

</property>