An easy to go guide for installing Hive in Windows 10.

A yellow honeycomb with a bee and a black text

Description automatically generated

Image taken from Google images

**1. Prerequisites**

1. Hardware Requirement  
   \* RAM — Min. 8GB, if you have SSD in your system then 4GB RAM would also work.  
   \* CPU — Min. Quad core, with at least 1.80GHz
2. JRE 1.8 — Offline installer for JRE
3. Java Development Kit — 1.8
4. A Software for Un-Zipping like 7Zip or Win Rar  
   \* I will be using a 64-bit windows for the process, please check and download the version supported by your system x86 or x64 for all the software.
5. Hadoop  
   \* I am using Hadoop-2.9.2, you can also use any other STABLE version for Hadoop.  
   \* If you don’t have Hadoop, you can refer installing it from Hadoop : How to install in 5 Steps in Windows 10.
6. MySQL Query Browser
7. Download Hive zip  
   \* I am using Hive-3.1.2, you can also use any other STABLE version for Hive.

A screenshot of a computer

Description automatically generated

Fig 1:- Download Hive-3.1.2

**2. Unzip and Install Hive**

After Downloading the Hive, we need to Unzip the apache-hive-3.1.2-bin.tar.gz file.

A screenshot of a computer

Description automatically generated

Fig 2:- Extracting Hive Step-1

Once extracted, we would get a new file apache-hive-3.1.2-bin.tar  
Now, once again we need to extract this tar file.

A screenshot of a computer

Description automatically generated

Fig 3:- Extracting Hive Step-2

* Now we can organize our Hive installation, we can create a folder and move the final extracted file in it. For Eg. :-

A screenshot of a computer

Description automatically generated

Fig 4:- Hive Directory

* Please note while creating folders, DO NOT ADD SPACES IN BETWEEN THE FOLDER NAME.(it can cause issues later)
* I have placed my Hive in D: drive you can use C: or any other drive also.

**3. Setting Up Environment Variables**

Another important step in setting up a work environment is to set your Systems environment variable.

To edit environment variables, go to Control Panel > System > click on the “Advanced system settings” link  
Alternatively, We can Right click on This PC icon and click on Properties and click on the “Advanced system settings” link  
Or, easiest way is to search for Environment Variable in search bar and there you GO…😉

A screenshot of a computer

Description automatically generated

Fig. 5:- Path for Environment Variable

A screenshot of a computer program

Description automatically generated

Fig. 6:- Advanced System Settings Screen

**3.1 Setting HIVE\_HOME**

* Open environment Variable and click on “New” in “User Variable”

A screenshot of a computer

Description automatically generated

Fig. 7:- Adding Environment Variable

* On clicking “New”, we get below screen.

A screenshot of a computer

Description automatically generated

Fig. 8:- Adding HIVE\_HOME

* Now as shown, add HIVE\_HOME in variable name and path of Hive in Variable Value.
* Click OK and we are half done with setting HIVE\_HOME.

**3.2 Setting Path Variable**

* Last step in setting Environment variable is setting Path in System Variable.

A screenshot of a computer program

Description automatically generated

Fig. 9:- Setting Path Variable

* Select Path variable in the system variables and click on “Edit”.

A screenshot of a computer

Description automatically generated

Fig. 10:- Adding Path

* Now we need to add these paths to Path Variable :-  
  \* %HIVE\_HOME%\bin
* Click OK and OK. & we are done with Setting Environment Variables.

**3.4 Verify the Paths**

* Now we need to verify that what we have done is correct and reflecting.
* Open a **NEW** Command Window
* Run following commands

echo %HIVE\_HOME%

**4. Editing Hive**

Once we have configured the environment variables next step is to configure Hive. It has 7 parts:-

**4.1 Replacing bins**

First step in configuring the hive is to download and replace the bin folder.  
\* Go to this GitHub Repo and download the bin folder as a zip.  
\* Extract the zip and replace all the files present under bin folder to %HIVE\_HOME%\bin

**Note:- If you are using different version of HIVE then please search for its respective bin folder and download it.**

**4.2 Creating File Hive-site.xml**

Now we need to create the Hive-site.xml file in hive for configuring it :-  
(We can find these files in Hive -> conf -> hive-default.xml.template)

We need to copy the hive-default.xml.template file and paste it in the same location and rename it to hive-site.xml. This will act as our main Config file for Hive.

A screenshot of a computer

Description automatically generated

Fig. 11:- Creating Hive-site.xml

**4.3 Editing Configuration Files**

**4.3.1 Editing the Properties**

Now Open the newly created Hive-site.xml and we need to edit the following properties

<property>  
 <name>hive.metastore.uris</name>  
 <value>thrift://<Your IP Address>:9083</value>  
   
 <property>  
 <name>hive.downloaded.resources.dir</name>  
 <value><Your drive Folder>/${hive.session.id}\_resources</value>  
   
<property>  
 <name>hive.exec.scratchdir</name>  
 <value>/tmp/mydir</value>

Replace the value for <Your IP Address> with the IP Address of your System and replace <Your drive Folder> with the Hive folder Path.

**4.3.2 Removing Special Characters**

This is a short step and we need to remove all the &#8 character present in the hive-site.xml file.

**4.3.3 Adding few More Properties**

Now we need to add the following properties as it is in the hive-site.xml File.

<property>  
 <name>hive.querylog.location</name>  
 <value>$HIVE\_HOME/iotmp</value>  
 <description>Location of Hive run time structured log file</description>  
 </property><property>  
 <name>hive.exec.local.scratchdir</name>  
 <value>$HIVE\_HOME/iotmp</value>  
 <description>Local scratch space for Hive jobs</description>  
</property><property>  
 <name>hive.downloaded.resources.dir</name>  
 <value>$HIVE\_HOME/iotmp</value>  
 <description>Temporary local directory for added resources in the remote file system.</description>  
 </property>

Great..!!! We are almost done with the Hive part, for configuring MySQL database as Metastore for Hive, we need to follow below steps:-

**4.4 Creating Hive User in MySQL**

The next important step in configuring Hive is to create users for MySQL.  
These Users are used for connecting Hive to MySQL Database for reading and writing data from it.

*Note:- You can skip this step if you have created the hive user while SQOOP installation.*

* Firstly, we need to open the MySQL Workbench and open the workspace(default or any specific, if you want). We will be using the default workspace only for now.

A screenshot of a computer

Description automatically generated

Fig 12:- Open MySQL Workbench

* Now Open the Administration option in the Workspace and select Users and privileges option under Management.

A screenshot of a computer

Description automatically generated

Fig 13:- Opening Users and Privileges

* Now select Add Account option and Create an new user with Login Name as hive and Limit to Host Mapping as the localhost and Password of your choice.

A screenshot of a computer screen

Description automatically generated

Fig 14:- Creating Hive User

* Now we have to define the roles for this user under Administrative Roles and select DBManager ,DBDesigner and BackupAdmin Roles

A screenshot of a computer

Description automatically generated

Fig 15:- Assigning Roles

* Now we need to grant schema privileges for the user by using Add Entry option and selecting the schemas we need access to.

A screenshot of a computer

Description automatically generated

Fig 16:- Schema Privileges

I am using schema matching pattern as %\_bigdata% for all my bigdata related schemas. You can use other 2 options also.

* After clicking OK we need to select All the privileges for this schema.

A screenshot of a computer

Description automatically generated

Fig 17:- Select All privileges in the schema

* Click Apply and we are done with the creating Hive user.

**4.5 Granting permission to Users**

Once we have created the user hive the next step is to Grant All privileges to this user for all the Tables in the previously selected Schema.

* Open the MySQL cmd Window. We can open it by using the Window’s Search bar.

A screenshot of a computer

Description automatically generated

Fig 18:- MySQL cmd

* Upon opening it will ask for your root user password(created while setting up MySQL).
* Now we need to run the below command in the cmd window.

grant all privileges on test\_bigdata.\* to 'hive'@'localhost';

where test\_bigdata will be you schema name and hive@localhost will be the user name @ Host name.

**4.6 Creating Metastore**

Now we need to create our own metastore for Hive in MySQL..

Firstly, we need to create a database for metastore in MySQL OR we can use the one which used in previous step test\_bigdata in my case.

Now Navigate to the below path

hive -> scripts -> metastore -> upgrade -> mysql and execute the file hive-schema-3.1.0.mysql in MySQL in your database.

*Note:- If you are using different Database, select the folder for same in upgrade folder and execute the hive-schema file.*

**4.7 Adding Few More Properties(Metastore related Properties)**

Finally, we need to open our hive-site.xml file once again and make some changes their, these are related to Hive metastore that’s why did not add them in starting so as to distinguish between the different set of properties.

<property>  
 <name>javax.jdo.option.ConnectionUserName</name>  
 <value>hive</value>  
 <description>Username to use against metastore database</description>  
 </property>  
   
<property>  
 <name>javax.jdo.option.ConnectionURL</name>  
 <value>jdbc:mysql://localhost:3306/<Your Database>?createDatabaseIfNotExist=true</value>  
 <description>  
 JDBC connect string for a JDBC metastore.  
 To use SSL to encrypt/authenticate the connection, provide database-specific SSL flag in the connection URL.  
 For example, jdbc:postgresql://myhost/db?ssl=true for postgres database.  
 </description>  
 </property>  
   
 <property>  
 <name>hive.metastore.warehouse.dir</name>  
 <value>hdfs://localhost:9000/user/hive/warehouse</value>  
 <description>location of default database for the warehouse</description>  
 </property>  
   
 <property>  
 <name>javax.jdo.option.ConnectionPassword</name>  
 <value><Hive Password></value>  
 <description>password to use against metastore database</description>  
 </property>  
   
 <property>  
 <name>datanucleus.schema.autoCreateSchema</name>  
 <value>true</value>  
</property>  
<property>  
 <name>datanucleus.schema.autoCreateTables</name>  
 <value>True</value>  
 </property>  
   
 <property>  
 <name>datanucleus.schema.validateTables</name>  
 <value>true</value>  
 <description>validates existing schema against code. turn this on if you want to verify existing schema</description>  
 </property>

Replace the value for *<Hive Password>* with the hive user password that we created in MySQL user creation. And <Your Database> with the database that we used for metastore in MySQL.

**5. Starting Hive**

**5.1 Starting Hadoop**

Now we need to start a new Command Prompt remember to run it as administrator to avoid permission issues and execute below commands

***start-all.cmd***

A screenshot of a computer program

Description automatically generated

Fig. 19:- start-all.cmd

All the 4 daemons should be UP and running.

**5.2 Starting Hive Metastore**

Open a cmd window, run below command to start the Hive metastore.

hive --service metastore

A black background with white text

Description automatically generated

Fig 20:- Starting Hive Metastore

**5.3 Starting Hive**

Now open a new cmd window and run the below command to start Hive

hive

**6. Common Issues**

**6.1 Unable to export or import data in hive**

The 1st common issue that we face after starting Hive is that we are unable to import Or Export

Sol:- We need to edit the below property and make it as false

<property> <name>hive.metastore.event.db.notification.api.auth</name> <value>false</value>   
<description>  
 Should metastore do authorization against database notification related APIs such as get\_next\_notification.  
 If set to true, then only the superusers in proxy settings have the permission  
 </description>  
 </property>

**6.2 Join not Working**

We need to run below commands before running the join query if we face an issue while running a join query:-

set hive.auto.convert.join=false;  
set hive.auto.convert.join.noconditionaltask=false;

Because without these hive tries a map-side join which fails, for normal join set these param as false.