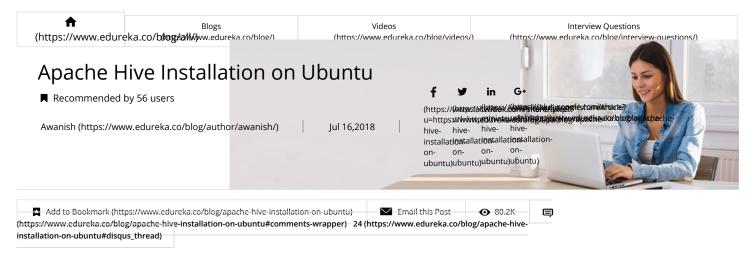
Home (/) > Big Data (/Big-data-and-analytics) > Big Data Analytics (https://www.edureka.co/blog/category/big-data-analytics/) > Apache Hive Installation o...



In this blog, we will learn about Apache Hive and Hive installation on Ubuntu.

What is Apache Hive?

Apache Hive is a data warehouse infrastructure that facilitates querying and managing large data sets which resides in distributed storage system. It is built on top of Hadoop and developed by Facebook. **Hive** provides a way to query the data using a SQL-like query language called **HiveQL(Hive query Language)**.

Internally, a compiler translates HiveQL statements into MapReduce jobs, which are then submitted to Hadoop framework for execution.

Difference between Hive and SQL:

Hive looks very much similar like traditional database with **SQL** access. However, because **Hive** is based on **Hadoop** and **MapReduce** operations, there are several key differences:

As Hadoop is intended for long sequential scans and **Hive** is based on **Hadoop**, you would expect queries to have a very high latency. Itmeans that **Hive** would not be appropriate for those applications that need very fast response times, as you can expect with a traditional RDBMS database.

Finally, Hive is read-based and therefore not appropriate for transaction processing that typically involves a high percentage of write operations.

Hive Installation on Ubuntu:

Please follow the below steps to install **Apache Hive** on Ubuntu:

Step 1: Download Hive tar.

Command: wget http://archive.apache.org/dist/hive/hive-2.1.0/apache-hive-2.1.0-bin.tar.gz

Step 2: Extract the tar file.

Command: tar -xzf apache-hive-2.1.0-bin.tar.gz

Command: Is

```
edureka@localhost:~$ tar -xzf apache-hive-2.1.0-bin.tar.gz
edureka@localhost:~$ ls
apache-hive-2.1.0-bin
apache-hive-2.1.0-bin.tar.gz
Downloads
Desktop
edureka@localhost:~$ Downloads
Pictures
Public
edureka@localhost:~$
```

Step 3: Edit the ".bashrc" file to update the environment variables for user.

Command: sudo gedit .bashrc

Add the following at the end of the file:

Set HIVE_HOME

export HIVE_HOME=/home/edureka/apache-hive-2.1.0-bin export PATH=\$PATH:/home/edureka/apache-hive-2.1.0-bin/bin

Also, make sure that hadoop path is also set.

Add Hadoop bin/ directory to PATH

export PATH=SPATH: /home/edureka/hadoop-2.7.3/bin
export PATH=SPATH: /home/edureka/hadoop-2.7.3/bin
export HADOOP_PID_DIR=/home/edureka/hadoop-2.7.3/hadoop2_data/hdfs/pid

Run below command to make the changes work in same terminal.

Command: source .bashrc

Step 4: Check hive version.

```
edureka@localhost:~$ hive --version
Hive 2.1.0
Subversion git://jcamachguezrMBP/Users/jcamachorodriguez/src/workspaces/
hive/HIVE-release2/hive -r 9265bc24d75ac945bde9ce1a0999fddd8f2aae29
Compiled by jcamachorodriguez on Fri Jun 17 01:03:25 BST 2016
From source with checksum 1f896b8fae57fbd29b047d6d67b75f3c
edureka@localhost:~$
```

Step 5: Create Hive directories within HDFS. The directory 'warehouse' is the location to store the table or data related to hive.

Command:

- hdfs dfs -mkdir -p /user/hive/warehouse
- hdfs dfs -mkdir /tmp

Step 6: Set read/write permissions for table.

Command:

In this command, we are giving write permission to the group:

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

Step 7: Set Hadoop path in hive-env.sh

Command: cd apache-hive-2.1.0-bin/

Command: gedit conf/hive-env.sh

edureka@localhost:~\$ cd apache-hive-2.1.0-bin/
edureka@localhost:~/apache-hive-2.1.0-bin\$ cp conf/hive-env.sh.template
conf/hive-env.sh
edureka@localhost:~/apache-hive-2.1.0-bin\$ gedit conf/hive-env.sh

Set the parameters as shown in the below snapshot.

Set HADDOP_HOME to point to a specific hadoop install directory export HADDOP_HOME=/home/edureka/hadoop-2.7.3

export HADDOP_HEAPSIZE=512

Hive Configuration Directory can be controlled by:
export HIVE_CONF_DIRE_Phome/edureka/apache-hive-2.1.0-bin/conf

Step 8: Edit hive-site.xml

Command: gedit conf/hive-site.xml

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?><!--</pre>
Licensed to the Apache Software Foundation (ASF) under one or more
contributor license agreements. See the NOTICE file distributed with
this work for additional information regarding copyright ownership.
The ASF licenses this file to You under the Apache License, Version 2.0
(the "License"); you may not use this file except in compliance with
the License. You may obtain a copy of the License at
http://www.apache.org/licenses/LICENSE-2.0
Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License.
-->
<configuration>
cproperty>
<name>javax.jdo.option.ConnectionURL</name>
<value>jdbc:derby:;databaseName=/home/edureka/apache-hive-2.1.0-bin/metastore_db;create=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>
cproperty>
<name>hive.metastore.warehouse.dir</name>
<value>/user/hive/warehouse</value>
<description>location of default database for the warehouse</description>
</property>
cproperty>
```

Step 9: By default, Hive uses Derby database. Initialize Derby database.

Command: bin/schematool -initSchema -dbType derby

Step 10: Launch Hive.

Command: hive

```
edureka@localhost:~/apache-hive-2.1.0-bin$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/edureka/apache-hive-2.1.0-bin/lib/log4j-slf4j-impl-2.4.1.jar!/org/slf4j/impl/StaticloggerBinder.class]
SLF4J: Found binding in [jar:file:/home/edureka/hadoop-2.7.3/share/hadoo
/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticloggerBinder
.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an expl
anation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFa
totry]
Logging initialized using configuration in jar:file:/home/edureka/apache
-hive-2.1.0-bin/lib/hive-common-2.1.0.jar!/hive-log4j2.properties Async:
true
true
versions. Consider using a different execution engine (i.e. spark, tez
) or using Hive 1.X releases.
```

(https://www.edureka.co/blog/apache-hive-installation-on-ubuntu)

Step 11: Run few queries in Hive shell.

Command: show databases;

Command: create table employee (id string, name string, dept string) row format delimited fields terminated by '\t' stored as textfile;

Command: show tables;

```
hive> show databases;

OK

default

Time taken: 1.742 seconds, Fetched: 1 row(s)

hive> create table employee (id string, name string, dept string) row format delimited fields terminated by '\t' stored as textfile;

OK

Time taken: 1.396 seconds

hive> show tables;

OK

employee

Time taken: 0.228 seconds, Fetched: 1 row(s)

hive>
```

Step 12: To exit from Hive:

Command: exit;

Now that you are done with Hive installation, the next step forward is to try out Hive commands on Hive shell. Hence, our next blog "Top Hive Commands with Examples in HQL (https://www.edureka.co/blog/hive-commands-with-examples)" will help you to master Hive commands.

Related Posts:

 ${\tt Get\ started\ with\ Big\ Data\ and\ Hadoop\ (https://www.edureka.co/big-data-and-hadoop)}}$

HIVE Commands (https://www.edureka.co/piog/nive-commands-with-examples)

Introduction to Apache Hive (https://www.edureka.co/blog/introduction-to-apache-hive/)



About Awanish (19 Posts (https://www.edureka.co/blog/author/awanish/))

Awanish is a Sr. Research Analyst at Edureka. He has rich expertise in Big Data technologies like Hadoop, Spark, Storm, Kafka, Flink. Awanish also holds good knowledge on Devops tools like Git, Jenkins, Docker, Puppet, Nagios. He is passionate about new technologies which provide realistic solutions.



G+ tildherting og lestome/Ahacle? (https://www.pasfactorsek.com

www.gluediaeto/lolo/blagaspeche u=https://www.wpeninekweenhitegenine

hivehiveinstallationtallationtallationtallation-

on-Share on ubuntu)ubuntu)ubuntu)



NEXT >

Got your brain cells running? Stay tuned to latest technology updates

Enter your Email Address

SUBSCRIBE

Related Posts









Top Hive Commands with Examples in HQL **⊙** 173.2K

Map Side Join Vs. Join **⊙** 63.6K

How to Run Hive Scripts? **⊙** 63.5K

How to Set Up Hadoop Cluster with HDFS High Availability ◆ 46.8K

commands-withexamples)

side-join-vs-join/)

to-run-hive-scripts/)

(https://www.edureka.co/t/http to-set-up-hadoopcluster-with-hdfs-highavailability/)

Browse Categories

Big Data NoSQL (https://www.edureka.co/blog/category/big-data-nosql/)

Blockchain (https://www.edureka.co/blog/category/blockchain/)

Business Intelligence (https://www.edureka.co/blog/category/business-intelligence/)

Cloud Computing (https://www.edureka.co/blog/category/cloud-computing/)

Cyber Security (https://www.edureka.co/blog/category/cyber-security/)

Deep Learning (https://www.edureka.co/blog/category/deep-learning/)

Finance (https://www.edureka.co/blog/category/finance/)

Frameworks (https://www.edureka.co/blog/category/frameworks/)

Marketing (https://www.edureka.co/blog/category/marketing/)

Mobile Development (https://www.edureka.co/blog/category/mobile-development/)

Operations (https://www.edureka.co/blog/category/operations/)

Programming (https://www.edureka.co/blog/category/programming/)

Project Management (https://www.edureka.co/blog/category/project-management/)

Robotic Process Automation (https://www.edureka.co/blog/category/robotic-process-automation/)

Success Story (https://www.edureka.co/blog/category/success-story/)

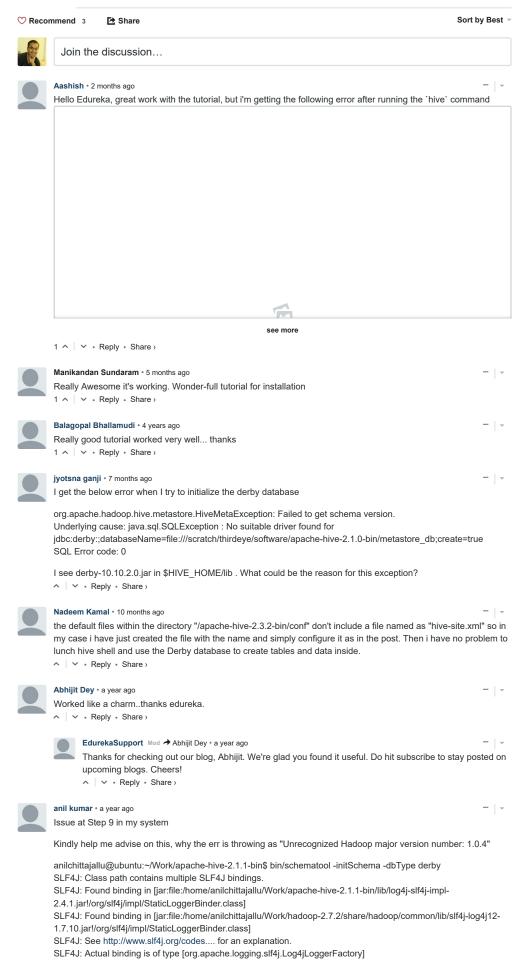
Systems & Architecture (https://www.edureka.co/blog/category/systems-architecture/)

Systems Engineering (https://www.edureka.co/blog/category/systems-engineering/)

Testing (https://www.edureka.co/blog/category/testing/)

Comments

25 Comments



Announcing the Ridiculously Committed Mentor Award at org.apache.hadoop.hive.shims.ShimLoader.loadShims(ShimLoader.java:136) at org.apache.hadoop.hive.shims.ShimLoader.getHadoopShims(ShimLoader.java:95) at org.apache.hive.beeline.HiveSchemaTool.<init>(HiveSchemaTool.java:81) at org.apache.hive.beeline.HiveSchemaTool.<init>(HiveSchemaTool.java:68) see more ∧ V • Reply • Share > EdurekaSupport Mod → anil kumar • a year ago Hey Anil, thanks for checking out our blog. Here's what you need to do: Please delete these jar files binding between Hadoop and Hive rm lib/hive-jdbc-2.0.0-standalone.jar rm lib/log4j-slf4j-impl-2.4.1.jar Now execute your commands and Let us know if you are still getting the same error. Hope this helps. Cheers! Aman Kapri • 3 years ago followed the same steps as mentioned above, but, getting this error : Naina Johari • 3 years ago I'm a new bie to hadoop, please resolve this issue. screenshot attached !! ∧ V • Reply • Share > EdurekaSupport Mod → Naina Johari • 3 years ago Hi Naina. Thank you for reaching out to us. You can raise a support ticket for the same by logging into your LMS, and our support team will surely help You can get in touch with us for further clarification by contacting our sales team on +91-8880862004 (India) or 1800 275 9730 (US toll free). You can also mail us on sales@edureka.co. ^ V • Reply • Share > Naina Johari • 3 years ago I'm unable to create hive directories within HDFS. The command mentioned in step 4 is returning: - No such file or directory ∧ | ∨ • Reply • Share › Naina Johari → Naina Johari • 3 years ago Just like unix we'll need the "-p" flag to create the parent directories as well unless you have already created them. Then command will work hadoop fs -mkdir -p /user/hive/warehouse ^ V • Reply • Share > What is the server domain and the port on which hive is installed and where can i modify it? ✓ • Reply • Share › EdurekaSupport Mod → Sricharan • 4 years ago Hi Sricharan, it takes the domain server of the machine on which hive is installed, by default. You can modify the Hive server domain and port number by editing the below property in hive-site.xml or hivedefault.xml file. property> <name>hive.metastore.uris</name> <value>thrift://\$metastore.server.full.hostname:\$portnumber</value> <description>URI for client to contact metastore server. To enable HiveServer2, leave the property value empty. </description> </property> Hope this helps!!

```
0.13.1.jar!/hive-log4j.properties
Exception in thread "main" java.lang.RuntimeException: java.lang.RuntimeException: Unable to instantiate
org.apache.hadoop.hive.metastore.HiveMetaStoreClient
at org.apache.hadoop.hive.ql.session.SessionState.start(SessionState.java:346)
at org.apache.hadoop.hive.cli....(CliDriver.java:681)
at org.apache.hadoop.hive.cli.CliDriver.main(CliDriver.java:625)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImp...:43)
at java.lang.reflect.Method.invoke(Method.java:622)
at org.apache.hadoop.util.RunJar.main(RunJar.java:156)
Caused by: java.lang.RuntimeException: Unable to instantiate
org.apache.hadoop.hive.metastore.HiveMetaStoreClient
at org.apache.hadoop.hive.metastore.MetaStoreUtils.newInstance(MetaStoreUtils.java:1412)
at org.apache.hadoop.hive.metastore.RetryingMetaStoreClient.<init>(RetryingMetaStoreClient.java:62)
at org.apache.hadoop.hive.metastore.RetryingMetaStoreClient.getProxy(RetryingMetaStoreClient.java:72)
at org.apache.hadoop.hive.ql.metadata.Hive.createMetaStoreClient(Hive.java:2453)
  ✓ • Reply • Share ›
      VishnuVardhan Reddy → Mubina Nadaf • a year ago
      any one
      solve this issue?
      ^ | ∨ • Reply • Share >
Murali • 4 years ago
I followed this steps, When I entered sudo hive in ubuntu, I see the following exception.
sudo: hive: command not found
I have also tried with out sudo,i see the following exception.
No command 'hive' found, did you mean:
Command 'hime' from package 'hime' (universe)
Command 'jive' from package 'filters' (universe)
hive: command not found
Note: I am in this folder.
/usr/lib/hive/apache-hive-0.13.1-bin/conf$
Please help me in solving this issue.!!Thanks in advance..!!
EdurekaSupport Mod → Murali • 4 years ago
      Hi Murali, please try with the following command and check
      Command: cd /usr/lib/hive/apache-hive-0.13.1-bin/
      Command: bin/hive
      Hope this helps!!
       Mitesh → EdurekaSupport • 2 years ago
             I am still getting the same error which is mentioned above by murali.... please suggest what to do..

✓ • Reply • Share ›
                    mahesh gupta → Mitesh • 2 years ago
                    you should correct the path of hive in either bashrc or profile where you have save your
                    environment setting @Murali
                    bala • 4 years ago
Hi.
I followed the steps exactly outlined here. Using Hive 0.8.0 I am getting the error
Exception in thread "main" java.lang.NoClassDefFoundError: org/apache/hadoop/hive/conf/HiveConf
at java.lang.Class.forName0(Native Method)
at java.lang.Class.forName(Class.java:249)
```

at org.apache.hadoop.util.RunJar.main(RunJar.iava:149)

at java.net.URLClassLoader\$1.run(URLClassLoader.java:202) at java.security.AccessController.doPrivileged(Native Method) at java.net.URLClassLoader.findClass(URLClassLoader.java:190) at java.lang.ClassLoader.loadClass(ClassLoader.java:306) at java.lang.ClassLoader.loadClass(ClassLoader.java:247) ... 3 more EdurekaSupport Mod > bala • 4 years ago

Hi Bala, Seems like the HADOOP HOME path has not been set. I would suggest to set the HADOOP_HOME in ~/.bashrc or /etc/profile file. To do that please follow the below steps.

Step1: Open a new terminal and give: sudo gedit ~/.bashrc

Step 2:Go to end of the file and add the below line:

export HADOOP_HOME=/path/to/Hadoop

Step 3: In the command, give the path where Hadoop is installed on your Ubuntu and then try to run Hive.

ALSO ON HTTPS://WWW.EDUREKA.CO/BLOG/

Kubernetes vs Docker: Which One Would You Pick?

1 comment • 9 months ago

Ashish Grover — As per this link from Docker, it does

Avatarseem to support automatic roll-backs.

https://docs.docker.com/v17...

Top 55 Blockchain Interview Questions You Must Prepare In 2018

Maximilian Fischer — A blockchain (database) only Avatarcontains a record of all changes that were ever made to it. Value is ascribed to the fact of actually being ...

7 Reasons to Choose Edureka Online Courses

Gokhan Dedeoglu — My registered e-mail adress is Avatargdedeoglu@gmail.com

Top 10 Reasons Why You Should Learn Selenium

Partho — Since there is no Global Certification in AvatarSelenium what is the benchmark then...

Subscribe to our newsletter

Enter your Email Address

SUBSCRIBE

Related Blogs

"HIVE. (https://www.

Top Hive Commands with Examples in HQL (https://www.edureka.co/blog/hive-commands-with-examples)

commandswithexamples)

□

Map Side Join Vs. Join (https://www.edureka.co/blog/map-side-join-vs-join/)

(https://www. side-join-vsjoin/)

https://www.edureka.co/blog/apache-hive-installation-on-ubuntu?utm source=hd&utm campaign=lms hd 120616&utm medium=lms



to-run-hivescripts/)



Top 50 Hadoop Interview Questions You Must Prepare In 2018 (https://www.edureka.co/blog/interview-questions/top-50hadoop-interview-questions-2016/)

(https://www. questions/top 50-hadoopinterviewquestions-2016/)

Big Data Analytics Courses



Big Data Hadoop Certification Training (/big-data-and-hadoop)

(/big-dataand-hadoop)



Apache Spark and Scala Certification Training (/apache-spark-scala-training)

(/apachespark-scalatraining)



Hadoop Administration Certification Training (/hadoop-admin)

(/hadoopadmin)



Splunk Power User & Admin Certification Training (/splunk)

Edureka	Work with us	Usefull Links	Follow us on
About us	Careers	Reviews	f ¥ in □
(https://www.edureka.co/abou(https://www.edureka.co/care(https://www.edureka.co/reviews) (https://titps://titps://titpsedialage/abinehtalage.comp.enge/a/educe/akjalN)			
us)	Become an Instructor	Terms & conditions	(Treeps (an expectament particular production and an expectation and an expectati
News & Media	(https://www.edureka.co/instr(https://add)w.edureka.co/terms-		
(https://www.edureka.co/all	m &dio)me an Affiliate	and-conditions)	
Contact us	(https://www.edureka.co/aff	iateivacy policy Learn on the GO!	
(https://www.edureka.co/contantogram)		(https://www.edureka.co/priva	tile do:
us)	Hire from Edureka	policy)	
Blog	(https://www.edureka.co/hir	e-Sitemap	(https://itunes.apple.com/in/app/edureka/id1033145415? mt=8)
(https://www.edureka.co/blog/fatb/n-edureka)		(https://www.edureka.co/sitem	nap)
Community			(https://play.google.com/store/apps/details?
(https://www.edureka.co/community)			id=co.edureka.app)

(Hittps://www.edureka.co) © 2014 Brain4ce Education Solutions Pvt. Ltd. All rights Reserved.

"PMP®", "PMI®", "PMI-ACP®" and "PMBOK®" are registered marks of the Project Management Institute, Inc. MongoDB®, Mongo and the leaf logo are the registered trademarks of MongoDB, Inc.