

Install pig in linux(ubuntu)

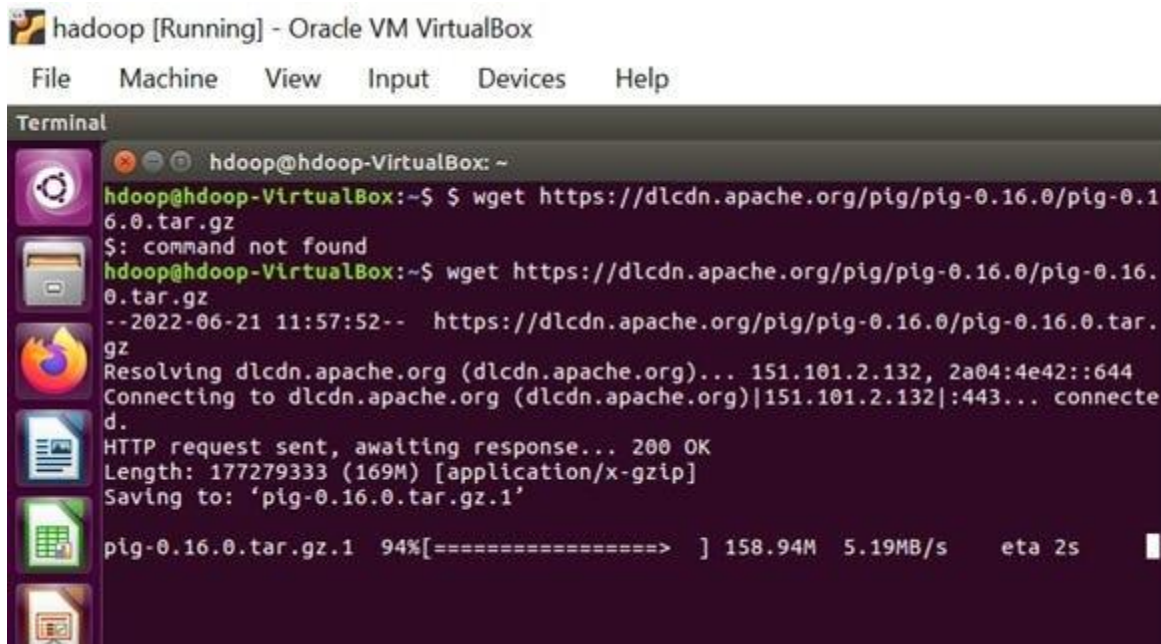
Step 1: Login into Hadoop user you used while installing Hadoop , here we use hdoop user

```
hdoop@NuvobookV1: ~$
```

\$ su – hdoop

Step 2: Go to <https://pig.apache.org/releases.html> and copy the path of the latest version of pig that you want to install. Run the following command to download Apache Pig in Ubuntu:

\$ wget <https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz>



```
hadoop [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Terminal
hdoop@hdoop-VirtualBox: ~
hdoop@hdoop-VirtualBox:~$ $ wget https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz
$: command not found
hdoop@hdoop-VirtualBox:~$ wget https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz
--2022-06-21 11:57:52-- https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz
Resolving dlcdn.apache.org (dlcdn.apache.org)... 151.101.2.132, 2a04:4e42::644
Connecting to dlcdn.apache.org (dlcdn.apache.org)|151.101.2.132|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 177279333 (169M) [application/x-gzip]
Saving to: 'pig-0.16.0.tar.gz.1'
pig-0.16.0.tar.gz.1 94%[=====] 158.94M 5.19MB/s eta 2s
```

Step 3: To untar pig-0.16.0.tar.gz file run the following command:

\$ tar xvfz pig-0.16.0.tar.gz

Step 4: Now open the .bashrc file to edit the path and variables/settings for pig. Run the following command:

```
$ sudo nano .bashrc
```

```
#PIG settings
```

```
export PIG_HOME=/home/hadoop/pig-0.17.0
```

```
export PATH=$PATH:$PIG_HOME/bin
```

```
export PIG_CLASSPATH=$PIG_HOME/conf:$HADOOP_INSTALL/etc/hadoop/
```

```
export PIG_CONF_DIR=$PIG_HOME/conf
```

```
export PIG_CLASSPATH=$PIG_CONF_DIR:$PATH
```

```
#PIG settings
export PIG_HOME=/home/hadoop/pig-0.17.0
export PATH=$PATH:$PIG_HOME/bin
export PIG_CLASSPATH=$PIG_HOME/conf:$HADOOP_INSTALL/etc/hadoop/
export PIG_CONF_DIR=$PIG_HOME/conf
export PIG_CLASSPATH=$PIG_CONF_DIR:$PATH
```

Step 5: Run the following command to make the changes effective in the .bashrc file:

```
$ source .bashrc
```

Step 6: To start all Hadoop daemons, navigate to the hadoop-3.2.1/sbin folder and run the following commands:

```
$ cd hadoop-3.4.0/sbin/
```

```
$ ./start-dfs.sh
```

```
$ ./start-yarn
```

```
$ jps
```

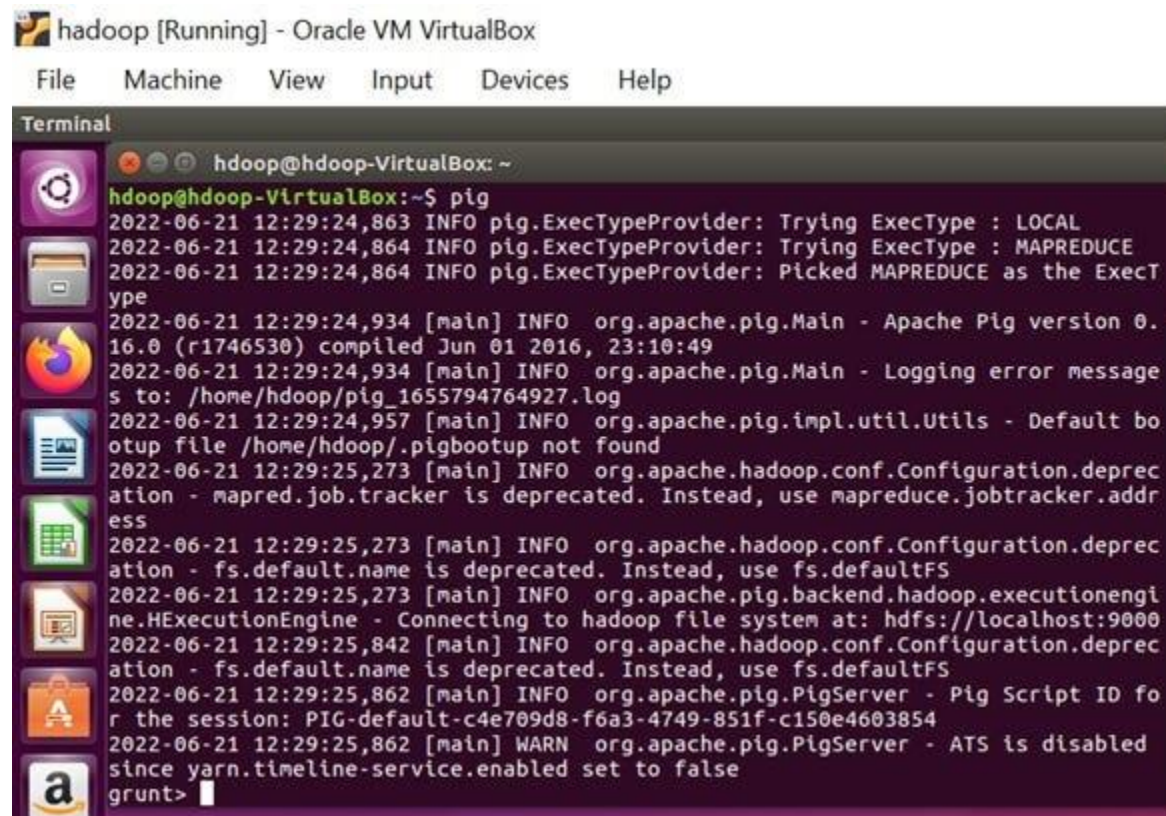
```

hdoop@hdoop-VirtualBox:~$ cd hadoop-3.2.1/sbin
hdoop@hdoop-VirtualBox:~/hadoop-3.2.1/sbin$ ./start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [hdoop-VirtualBox]
hdoop@hdoop-VirtualBox:~/hadoop-3.2.1/sbin$ ./start-yarn.sh
Starting resourcemanager
Starting nodemanagers
hdoop@hdoop-VirtualBox:~/hadoop-3.2.1/sbin$ jps
4817 DataNode
5298 ResourceManager
5000 SecondaryNameNode
5450 NodeManager
4683 NameNode
5982 Jps
hdoop@hdoop-VirtualBox:~/hadoop-3.2.1/sbin$ █

```

Step 7: Now you can launch pig by executing the following command:

\$ pig



```

hadoop [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Terminal
hdoop@hdoop-VirtualBox: ~
hdoop@hdoop-VirtualBox:~$ pig
2022-06-21 12:29:24,863 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
2022-06-21 12:29:24,864 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE
2022-06-21 12:29:24,864 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType
2022-06-21 12:29:24,934 [main] INFO org.apache.pig.Main - Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:49
2022-06-21 12:29:24,934 [main] INFO org.apache.pig.Main - Logging error message s to: /home/hdoop/pig_1655794764927.log
2022-06-21 12:29:24,957 [main] INFO org.apache.pig.impl.util.Utils - Default bootstrap file /home/hdoop/.pigbootstrap not found
2022-06-21 12:29:25,273 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2022-06-21 12:29:25,273 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2022-06-21 12:29:25,273 [main] INFO org.apache.pig.backend.hadoop.executionengine.HExecutionEngine - Connecting to hadoop file system at: hdfs://localhost:9000
2022-06-21 12:29:25,842 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2022-06-21 12:29:25,862 [main] INFO org.apache.pig.PigServer - Pig Script ID for the session: PIG-default-c4e709d8-f6a3-4749-851f-c150e4603854
2022-06-21 12:29:25,862 [main] WARN org.apache.pig.PigServer - ATS is disabled since yarn.timeline-service.enabled set to false
grunt> █

```

Step 8: Now you are in pig and can perform your desired tasks on pig. You can come out of the pig by the quit command:

> quit;