

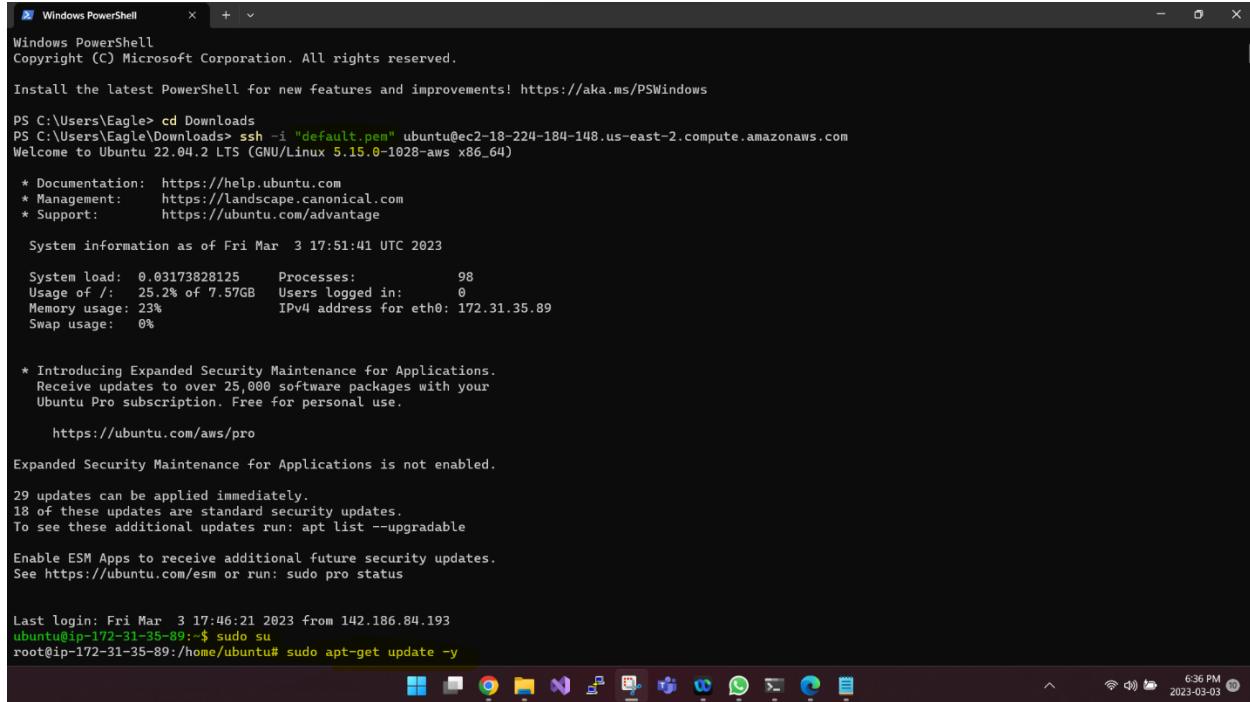
[Type here]

Contents

| | |
|--|---|
| In figure one I have used update command which run some files. Figure 1 | 2 |
| In this second figure I used upgarde command to run all upgradation. Moreover I installed java 8. Figure 2 | 3 |
| I used keygen command and further some commands. | 3 |
| I showd hdoop is running all services. | 5 |
| IN this picture I have installed sqoop where I highlighted | 5 |
| In this figure I installed hive. | 6 |
| I showed all paths hive, sqoop and hadoop | 7 |
| Figure 7 all path | 7 |

[Type here]

In figure one I have used update command which run some files. Figure 1



```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and improvements! https://aka.ms/PSWindows

PS C:\Users\Eagle> cd Downloads
PS C:\Users\Eagle\Downloads> ssh -i "default.pem" ubuntu@ec2-18-224-184-148.us-east-2.compute.amazonaws.com
Welcome to Ubuntu 22.04.2 LTS (GNU/Linux 5.15.0-1028-aws x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

System information as of Fri Mar  3 17:51:41 UTC 2023

System load:  0.03173828125   Processes:            98
Usage of /:   25.2% of 7.57GB Users logged in:        0
Memory usage: 23%           IPv4 address for eth0: 172.31.35.89
Swap usage:   0%

 * Introducing Expanded Security Maintenance for Applications.
   Receive updates to over 25,000 software packages with your
   Ubuntu Pro subscription. Free for personal use.

   https://ubuntu.com/aws/pro

Expanded Security Maintenance for Applications is not enabled.

29 updates can be applied immediately.
18 of these updates are standard security updates.
To see these additional updates run: apt list --upgradable

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status

Last login: Fri Mar  3 17:46:21 2023 from 142.186.84.193
ubuntu@ip-172-31-35-89:~$ sudo su
root@ip-172-31-35-89:/home/ubuntu# sudo apt-get update -y
```

Figure 1update command

[Type here]

In this second figure I used upgrade command to run all upgradation. Moreover I installed java 8. Figure 2

```
Windows PowerShell
root@ip-172-31-35-89:/home/ubuntu# sudo apt-get update -y
Hit:1 http://us-east-2.ec2.archive.ubuntu.com/ubuntu jammy InRelease
Hit:2 http://us-east-2.ec2.archive.ubuntu.com/ubuntu jammy-updates InRelease
Hit:3 http://us-east-2.ec2.archive.ubuntu.com/ubuntu jammy-backports InRelease
Hit:4 http://security.ubuntu.com/ubuntu jammy-security InRelease
Reading package lists... Done
root@ip-172-31-35-89:/home/ubuntu# sudo apt-get upgrade -y
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
Calculating upgrade... Done
The following packages have been kept back:
  grub-efi-amd64-bin grub-efi-amd64-signed libmbim-glib4 libmbim-proxy libqmi-glib5 libqmi-proxy libssl2-2
  libssl2-modules libssl2-modules-db linux-aws linux-headers-aws linux-image-aws python3-software-properties
  software-properties-common tcpdump ubuntu-advantage-tools
0 upgraded, 0 newly installed, 0 to remove and 16 not upgraded.
root@ip-172-31-35-89:/home/ubuntu# sudo apt-get install openjdk-8-jdk
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  adwaita-icon-theme alsa-topology-conf alsa-ucm-conf at-spi2-core ca-certificates-java dconf-gsettings-backend
  dconf-service fontconfig fontconfig-config fonts-dejavu-core fonts-dejavu-extra gsettings-desktop-schemas
  gtk-update-icon-cache hicolor-icon-theme humanity-icon-theme java-common libasound2 libasound2-data libasound2-plugins
  libatk-bridge2.0-0 libatk-wrapper-java libatk-wrapper-java-jni libatk1.0-0 libatk1.0-data libatspi2.0-0
  libavahi-client3 libavahi-common-data libavahi-common3 libcairo-gobject2 libcairo2 libcups2 libdatrie1 libdconf1
  libdeflate0 libdrm-amdgpu1 libdrm-intel1 libdrm-nouveau2 libdrm-radeon1 libflac8 libfontconfig1 libfontenc1
  libgail-common libgail18 libgdk-pixbuf2.0-0 libgdk-pixbuf2.0-bin libgdk-pixbuf2.0-common libgif7 libgl1
  libgl1-amd-glx libgl1-mesa-dri libgl1-mesa-glx libglapi-mesa libglvnd0 libglx-mesa0 libglx0 libgraphite2-3
  libgtk2.0-0 libgtk2.0-bin libgtk2.0-common libharfbuzz0b libice-dev libice6 libjpeg-turbo8 libjpeg8
  liblcms2-2 libllvm15 libogg0 libopus0 libpango-1.0-0 libpangocairo-1.0-0 libpangoft2-1.0-0 libpciaccess0
  libpcsc-lite1 libpixmap1-0 libpthread-stubs0-dev libpulse0 librsync2-2 librsync2-common libsensors-config libsensors5
  libsm-dev libsm6 libsndfile1 libthai-data libthai0 libtiff5 libvorbis0a libvorbisenc2 libwebp7 libx11-dev
  libx11-xcb1 libxau-dev libxaw7 libxcb-dri2-0 libxcb-dri3-0 libxcb-glx0 libxcb-present0 libxcb-render0 libxcb-shape0
  libxcb-shm0 libxcb-sync1 libxcb-xfixes0 libxcb1-dev libxcomposite1 libxcursor1 libxdamage1 libxdmcp-dev libxfixes3
  libxft2 libxi6 libxinerama1 libxkbfile1 libxmu6 libxpm4 libxrandr2 libxrender1 libxshmfence1 libxt-dev libxt6
  libxtst6 libxv1 libxxf86dga1 libxxf86vm1 openjdk-8-jdk-headless openjdk-8-jre openjdk-8-jre-headless
  session-migration ubuntu-mono x11-common x11-utils x11proto-dev xorg-sgml-doctools xtrans-dev
Suggested packages:
  default-jre libasound2-plugins alsa-utils cups-common gvfs libice-doc liblcms2-utils opus-tools pscd pulseaudio
```

Figure 2upgrade

I used keygen command and further some commands.

[Type here]

```
Windows PowerShell

ssh-keygen -Y verify -f allowed_signers_file -I signer_identity
-n namespace -s signature_file [-r krl_file] [-O option]
hadoop@ip-172-31-35-89:/home/ubuntu$ ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa
Generating public/private rsa key pair.
Created directory '/home/hadoop/.ssh'.
Your identification has been saved in /home/hadoop/.ssh/id_rsa
Your public key has been saved in /home/hadoop/.ssh/id_rsa.pub
The key fingerprint is:
SHA256:2caUmRM8KtTyNx0vUA30T75mxdGqzt8rY9Tw4HnQaU hadoop@ip-172-31-35-89
The key's randomart image is:
+----[RSA 3072]-----+
| .. ooooooo |
| .+X...o. |
| .oB + E. |
| o . o* . o + |
| . o .So+o . * o |
| . o o o o . +o |
| . . . +. . |
| o+ . |
| ..+o |
+----[SHA256]-----+
hadoop@ip-172-31-35-89:/home/ubuntu$ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
hadoop@ip-172-31-35-89:/home/ubuntu$ chmod 0600 ~/.ssh/authorized_keys
hadoop@ip-172-31-35-89:/home/ubuntu$
hadoop@ip-172-31-35-89:/home/ubuntu$ cd
hadoop@ip-172-31-35-89:~$
hadoop@ip-172-31-35-89:~$ pwd
/home/hadoop
hadoop@ip-172-31-35-89:~$ wget https://archive.apache.org/dist/hadoop/common/hadoop-3.3.4/hadoop-3.3.4.tar.gz
--2023-03-03 19:33:19-- https://archive.apache.org/dist/hadoop/common/hadoop-3.3.4/hadoop-3.3.4.tar.gz
Resolving archive.apache.org (archive.apache.org)... 138.201.131.134, 2a01:4f8:172:2ec5::2
Connecting to archive.apache.org (archive.apache.org)[138.201.131.134]:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 695457782 (663M) [application/x-gzip]
Saving to: 'hadoop-3.3.4.tar.gz'

hadoop-3.3.4.tar.gz      100%[=====] 663.24M  14.7MB/s   in 45s

2023-03-03 19:34:05 (14.6 MB/s) - 'hadoop-3.3.4.tar.gz' saved [695457782/695457782]
```

Figure 3 keygen

```
hadoop@ip-172-31-18-136: ~$
hadoop@ip-172-31-18-136:~/hadoop-2.10.1/sbin$ ./stop-dfs.sh
Stopping namenodes on [ec2-3-19-237-51.us-east-2.compute.amazonaws.com]
The authenticity of host 'ec2-3-19-237-51.us-east-2.compute.amazonaws.com (172.31.18.136)' can't be established.
ECDSA key fingerprint is SHA256:pY00byniGvGssdEI+6+erdM/9k03XFc9IceneD5z7Z8.
Are you sure you want to continue connecting (yes/no)? yes
ec2-3-19-237-51.us-east-2.compute.amazonaws.com: Warning: Permanently added 'ec2-3-19-237-51.us-east-2.compute.amazonaws.com' (ECDSA) to the list of known hosts.
ec2-3-19-237-51.us-east-2.compute.amazonaws.com: no namenode to stop
localhost: no datanode to stop
Stopping secondary namenodes [0.0.0.0]
0.0.0.0: no secondarynamenode to stop
hadoop@ip-172-31-18-136:~/hadoop-2.10.1/sbin$ ./stop-yarn.sh
stopping yarn daemons
no resourcemanager to stop
localhost: no nodemanager to stop
no proxyserver to stop
hadoop@ip-172-31-18-136:~/hadoop-2.10.1/sbin$ jps
8499 Jps
hadoop@ip-172-31-18-136:~/hadoop-2.10.1/sbin$ ./start-dfs.sh
Starting namenodes on [ec2-3-19-237-51.us-east-2.compute.amazonaws.com]
ec2-3-19-237-51.us-east-2.compute.amazonaws.com: starting namenode, logging to /home/hadoop/hadoop-2.10.1/logs/hadoop-hadoop-namenode-ip-172-31-18-136.out
localhost: starting datanode, logging to /home/hadoop/hadoop-2.10.1/logs/hadoop-hadoop-datanode-ip-172-31-18-136.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /home/hadoop/hadoop-2.10.1/logs/hadoop-hadoop-secondarynamenode-ip-172-31-18-136.out
hadoop@ip-172-31-18-136:~/hadoop-2.10.1/sbin$ ./start-yarn.sh
starting yarn daemons
starting resourcemanager, logging to /home/hadoop/hadoop-2.10.1/logs/yarn-hadoop-resourcemanager-ip-172-31-18-136.out
localhost: starting nodemanager, logging to /home/hadoop/hadoop-2.10.1/logs/yarn-hadoop-nodemanager-ip-172-31-18-136.out
hadoop@ip-172-31-18-136:~/hadoop-2.10.1/sbin$ jps
9586 Jps
9435 NodeManager
9180 SecondaryNameNode
8878 DataNode
9262 ResourceManager
hadoop@ip-172-31-18-136:~/hadoop-2.10.1/sbin$ ll
total 136
drwxr-xr-x 3 hadoop hadoop 4096 Mar 13 02:49 ./
drwxr-xr-x 10 hadoop hadoop 4096 Mar 13 01:20 ../
drwxr-xr-x 4 hadoop hadoop 4096 Sep 14 2020 FederationStateStore/
```

[Type here]

I showd hdoop is running all services.

The screenshot displays the Hadoop cluster management interface. The top navigation bar includes 'Instance details | EC2 Management' and 'Nodes of the cluster'. The main header shows the Hadoop logo and the title 'Nodes of the cluster'. A sidebar on the left lists navigation options: Cluster, About, Nodes, Node Labels, Applications, NEW, NEW SAVING, SUBMITTED, ACCEPTED, RUNNING, FINISHED, FAILED, KILLED, Scheduler, and Tools. The main content area is divided into several sections: Cluster Metrics, Cluster Nodes Metrics, Scheduler Metrics, and a table of node details. The Cluster Metrics section shows various counts and resource usage. The Cluster Nodes Metrics section shows the status of nodes. The Scheduler Metrics section shows the scheduler type and resource allocation. The table of node details lists the node labels, rack, node state, node address, node HTTP address, last health update, health report, containers, memory used, memory available, vcores used, vcores available, GPUs used, GPUs available, and version.

| Node Labels | Rack | Node State | Node Address | Node HTTP Address | Last health update | Health-report | Containers | Mem Used | Mem Avail | Vcores Used | Vcores Avail | GPUs Used | GPUs Avail | Version |
|---------------|------|------------|------------------------|-----------------------|--------------------------------|---------------|------------|----------|-----------|-------------|--------------|-----------|------------|---------|
| /default-rack | | RUNNING | ip-172-31-18-136:41273 | ip-172-31-18-136:8042 | Tue Mar 14 02:47:15 +0000 2023 | | 0 | 0 B | 8 GB | 0 | 8 | 0 | 0 | 2.10.1 |

Figure 4 hadoop running

IN this picture I have installed sqoop where I highlighted .

[Type here]

```
hadoop@ip-172-31-2-87: ~  
New release '20.04.5 LTS' available.  
Run 'do-release-upgrade' to upgrade to it.  
  
Last login: Tue Mar 14 01:31:28 2023 from 72.141.31.74  
ubuntu@ip-172-31-2-87:~$ sudo su hadoop  
hadoop@ip-172-31-2-87:~/home/ubuntu$ cd  
hadoop@ip-172-31-2-87:~$ wget http://archive.apache.org/dist/sqoop/1.99.7/sqoop-1.99.7-bin-hadoop200.tar.gz  
--2023-03-14 03:00:38-- http://archive.apache.org/dist/sqoop/1.99.7/sqoop-1.99.7-bin-hadoop200.tar.gz  
Resolving archive.apache.org (archive.apache.org)... 138.201.131.134, 2a01:4f8:172:2ec5::2  
Connecting to archive.apache.org (archive.apache.org)[138.201.131.134]:80... connected.  
HTTP request sent, awaiting response... 200 OK  
Length: 102436055 (98M) [application/x-gzip]  
Saving to: 'sqoop-1.99.7-bin-hadoop200.tar.gz.1'  
  
sqoop-1.99.7-bin-hadoop200.tar.gz.1 100%[=====] 97.69M 370KB/s in 4m 36s  
  
2023-03-14 03:05:13 (363 KB/s) - 'sqoop-1.99.7-bin-hadoop200.tar.gz.1' saved [102436055/102436055]  
  
hadoop@ip-172-31-2-87:~$ tar -xzf sqoop-1.99.7-bin-hadoop200.tar.gz  
hadoop@ip-172-31-2-87:~$ cd sqoop-1.99.7-bin-hadoop200  
hadoop@ip-172-31-2-87:~/sqoop-1.99.7-bin-hadoop200$ pwd  
/home/hadoop/sqoop-1.99.7-bin-hadoop200  
hadoop@ip-172-31-2-87:~/sqoop-1.99.7-bin-hadoop200$ cd  
hadoop@ip-172-31-2-87:~$ vi .bashrc  
hadoop@ip-172-31-2-87:~$ source .bashrc  
hadoop@ip-172-31-2-87:~$ cd $SQOOP_HOME/conf  
hadoop@ip-172-31-2-87:~/sqoop-1.99.7-bin-hadoop200/conf$ ls server/lib/  
ls: cannot access 'server/lib/': No such file or directory  
hadoop@ip-172-31-2-87:~/sqoop-1.99.7-bin-hadoop200/conf$ cd ..  
hadoop@ip-172-31-2-87:~/sqoop-1.99.7-bin-hadoop200$ ls server/lib/  
avro-1.7.7.jar          jackson-mapper-asl-1.9.13.jar  mysql-connector-java-5.1.48.jar  sqoop-core-1.99.7.jar  
commons-cli-1.2.jar     javax.servlet-api-3.1.0.jar    paranamer-2.3.jar               sqoop-execution-mapreduce-1.99.7.jar  
commons-codec-1.9.jar  jcommander-1.27.jar           postgresql-9.1-901.jdbc4.jar   sqoop-execution-common-1.99.7.jar  
commons-compress-1.9.jar  jetty-http-9.2.13.v20150730.jar  slf4j-api-1.6.1.jar            sqoop-repository-common-1.99.7.jar  
commons-dbc-1.4.jar     jetty-io-9.2.13.v20150730.jar  snappy-java-1.0.5.jar          sqoop-repository-derby-1.99.7.jar  
commons-io-2.4.jar      jetty-security-9.2.13.v20150730.jar  sqoop-common-1.99.7.jar        sqoop-repository-mysql-1.99.7.jar  
commons-lang-2.5.jar    jetty-server-9.2.13.v20150730.jar  sqoop-connector-ftp-1.99.7.jar  sqoop-repository-postgresql-1.99.7.jar  
commons-pool-1.5.4.jar  jetty-servlet-9.2.13.v20150730.jar  sqoop-connector-generic-jdbc-1.99.7.jar  sqoop-security-1.99.7.jar  
sqoop-server-1.99.7.jar
```

Figure 5 sqoop

In this figure I installed hive.

```
hadoop@ip-172-31-2-87: ~  
commons-dbc-1.4.jar      jetty-io-9.2.13.v20150730.jar  snappy-java-1.0.5.jar          sqoop-repository-mysql-1.99.7.jar  
commons-io-2.4.jar      jetty-security-9.2.13.v20150730.jar  sqoop-common-1.99.7.jar        sqoop-repository-postgresql-1.99.7.jar  
commons-lang-2.5.jar    jetty-server-9.2.13.v20150730.jar  sqoop-connector-ftp-1.99.7.jar  sqoop-security-1.99.7.jar  
commons-pool-1.5.4.jar  jetty-servlet-9.2.13.v20150730.jar  sqoop-connector-generic-jdbc-1.99.7.jar  sqoop-server-1.99.7.jar  
connector-sdk-1.99.7.jar  jetty-util-9.2.13.v20150730.jar  sqoop-connector-hdfs-1.99.7.jar  sqoop-submission-mapreduce-1.99.7.jar  
connector-sdk-hadoop-1.99.7.jar  joda-time-2.4.jar  sqoop-connector-kafka-1.99.7.jar  sqoop-tools-1.99.7.jar  
derby-10.8.2.2.jar      json-simple-1.1.jar  sqoop-connector-kite-1.99.7.jar  
guava-11.0.2.jar        jsr305-1.3.9.jar    sqoop-connector-oracle-jdbc-1.99.7.jar  
jackson-core-asl-1.9.13.jar  log4j-1.2.16.jar  sqoop-connector-sftp-1.99.7.jar  
hadoop@ip-172-31-2-87:~/sqoop-1.99.7-bin-hadoop200$ cd  
hadoop@ip-172-31-2-87:~$ wget https://downloads.apache.org/hive/hive-3.1.2/apache-hive-3.1.2-bin.tar.gz  
--2023-03-14 03:11:36-- https://downloads.apache.org/hive/hive-3.1.2/apache-hive-3.1.2-bin.tar.gz  
Resolving downloads.apache.org (downloads.apache.org)... 88.99.95.219, 135.181.214.104, 2a01:4f9:3a:2c57::2, ...  
Connecting to downloads.apache.org (downloads.apache.org)[88.99.95.219]:443... connected.  
HTTP request sent, awaiting response... 200 OK  
Length: 278813748 (266M) [application/x-gzip]  
Saving to: 'apache-hive-3.1.2-bin.tar.gz.1'  
  
apache-hive-3.1.2-bin.tar.gz.1 100%[=====] 265.90M 295KB/s in 12m 25s  
  
2023-03-14 03:24:01 (366 KB/s) - 'apache-hive-3.1.2-bin.tar.gz.1' saved [278813748/278813748]  
  
hadoop@ip-172-31-2-87:~$ tar xzf apache-hive-3.1.2-bin.tar.gz  
hadoop@ip-172-31-2-87:~$ cd  
hadoop@ip-172-31-2-87:~$ vi .bashrc  
hadoop@ip-172-31-2-87:~$ source .bashrc  
hadoop@ip-172-31-2-87:~$ ls  
apache-hive-3.1.2-bin  hadoop-2.10.1  pig-0.17.0  sqoop-1.99.7-bin-hadoop200.tar.gz.1  
apache-hive-3.1.2-bin.tar.gz  hadoop-2.10.1.tar.gz  pig-0.17.0.tar.gz  tmpdata  
apache-hive-3.1.2-bin.tar.gz.1  mysql-connector-java-5.1.48  sqoop-1.99.7-bin-hadoop200  sqoop-1.99.7-bin-hadoop200.tar.gz  
dfsdata  mysql-connector-java-5.1.48.tar.gz  sqoop-1.99.7-bin-hadoop200.tar.gz  
hadoop@ip-172-31-2-87:~$ cd apache-hive-3.1.2-bin  
hadoop@ip-172-31-2-87:~/apache-hive-3.1.2-bin$ cd  
hadoop@ip-172-31-2-87:~$ vi .bashrc  
hadoop@ip-172-31-2-87:~$ source .bashrc  
hadoop@ip-172-31-2-87:~$  
hadoop@ip-172-31-2-87:~$ sudo vi $HIVE_HOME/bin/hive-config.sh  
[sudo] password for hadoop:  
hadoop@ip-172-31-2-87:~$
```

Figure 6hive

[Type here]

I showed all paths hive, sqoop and hadoop

```
hadoop@ip-172-31-2-87: ~  
# Alias definitions.  
# You may want to put all your additions into a separate file like  
# ~/.bash_aliases, instead of adding them here directly.  
# See /usr/share/doc/bash-doc/examples in the bash-doc package.  
  
if [ -f ~/.bash_aliases ]; then  
    . ~/.bash_aliases  
fi  
  
# enable programmable completion features (you don't need to enable  
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile  
# sources /etc/bash.bashrc).  
if ! shopt -oq posix; then  
    if [ -f /usr/share/bash-completion/bash_completion ]; then  
        . /usr/share/bash-completion/bash_completion  
    elif [ -f /etc/bash_completion ]; then  
        . /etc/bash_completion  
    fi  
fi  
  
export HADOOP_HOME=/home/hadoop/hadoop-2.10.1  
export HADOOP_INSTALL=$HADOOP_HOME  
export HADOOP_MAPRED_HOME=$HADOOP_HOME  
export HADOOP_COMMON_HOME=$HADOOP_HOME  
export HADOOP_HDFS_HOME=$HADOOP_HOME  
export YARN_HOME=$HADOOP_HOME  
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native  
export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin  
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib/native"  
export JAVA_HOME="/usr/lib/jvm/java-8-openjdk-amd64"  
export PATH=$PATH:/home/hadoop/pig-0.17.0/bin  
export PIG_HOME=/home/hadoop/pig-0.17.0  
export PIG_CLASSPATH=$HADOOP_HOME/conf  
export SQOOP_HOME=/home/hadoop/sqoop-1.99.7-bin-hadoop200  
export PATH=$PATH:$SQOOP_HOME/bin  
export HIVE_HOME=/home/hadoop/apache-hive-3.1.2-bin  
export PATH=$PATH:$HIVE_HOME/bin  
export PATH=$PATH:/path/to/hadoop/bin  
~  
".bashrc" 135L, 4563C
```

Figure 7 all path