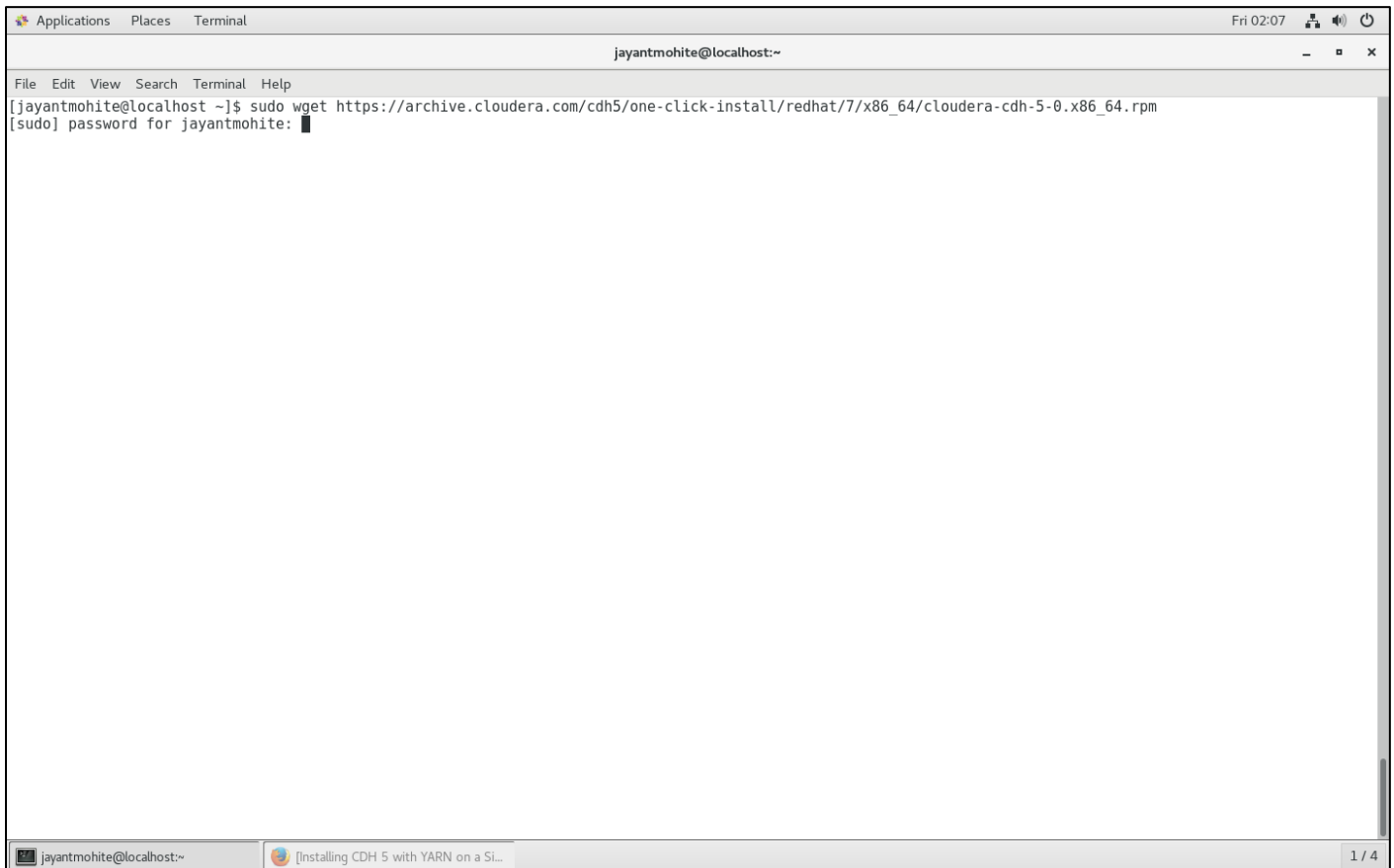


## Installation of Hadoop Single Node Cluster

Step1:

Download the Cloudera Package of Hadoop Version 5



A terminal window titled "Applications Places Terminal" with a status bar showing "Fri 02:07". The terminal prompt is "jayantmohite@localhost:~". The command entered is "sudo wget https://archive.cloudera.com/cdh5/one-click-install/redhat/7/x86\_64/cloudera-cdh-5-0.x86\_64.rpm". The prompt changes to "[sudo] password for jayantmohite:" followed by a cursor. The terminal has a menu bar with "File Edit View Search Terminal Help". At the bottom, there is a taskbar with a terminal icon labeled "jayantmohite@localhost:~" and a window icon labeled "[Installing CDH 5 with YARN on a SL...]" with a "1 / 4" indicator.

```
Applications Places Terminal
jayantmohite@localhost:~
File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ sudo wget https://archive.cloudera.com/cdh5/one-click-install/redhat/7/x86_64/cloudera-cdh-5-0.x86_64.rpm
[sudo] password for jayantmohite: █
```

Step 2:

## Install the Downloaded Package

```
Applications Places Terminal Fri 02:08
jayantmohite@localhost:~

File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ sudo wget https://archive.cloudera.com/cdh5/one-click-install/redhat/7/x86_64/cloudera-cdh-5-0.x86_64.rpm
[sudo] password for jayantmohite:
--2018-03-02 02:07:40-- https://archive.cloudera.com/cdh5/one-click-install/redhat/7/x86_64/cloudera-cdh-5-0.x86_64.rpm
Resolving archive.cloudera.com (archive.cloudera.com)... 151.101.8.167
Connecting to archive.cloudera.com (archive.cloudera.com)|151.101.8.167|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 9152 (8.9K) [application/x-redhat-package-manager]
Saving to: 'cloudera-cdh-5-0.x86_64.rpm'

100%[=====] 9,152 --.-K/s in 0s

2018-03-02 02:07:40 (49.0 MB/s) - 'cloudera-cdh-5-0.x86_64.rpm' saved [9152/9152]

[jayantmohite@localhost ~]$ sudo yum --nogpgcheck localinstall cloudera-cdh-5-0.x86_64.rpm
```

```
Applications Places Terminal Fri 02:09
jayantmohite@localhost:~

File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ sudo wget https://archive.cloudera.com/cdh5/one-click-install/redhat/7/x86_64/cloudera-cdh-5-0.x86_64.rpm
[sudo] password for jayantmohite:
--2018-03-02 02:07:40-- https://archive.cloudera.com/cdh5/one-click-install/redhat/7/x86_64/cloudera-cdh-5-0.x86_64.rpm
Resolving archive.cloudera.com (archive.cloudera.com)... 151.101.8.167
Connecting to archive.cloudera.com (archive.cloudera.com)|151.101.8.167|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 9152 (8.9K) [application/x-redhat-package-manager]
Saving to: 'cloudera-cdh-5-0.x86_64.rpm'

100%[=====] 9,152 --.-K/s in 0s

2018-03-02 02:07:40 (49.0 MB/s) - 'cloudera-cdh-5-0.x86_64.rpm' saved [9152/9152]

[jayantmohite@localhost ~]$ sudo yum --nogpgcheck localinstall cloudera-cdh-5-0.x86_64.rpm
Loaded plugins: fastestmirror, langpacks
Examining cloudera-cdh-5-0.x86_64.rpm: cloudera-cdh-5-0.x86_64
Marking cloudera-cdh-5-0.x86_64.rpm to be installed
Resolving Dependencies
--> Running transaction check
--> Package cloudera-cdh.x86_64 0:5-0 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

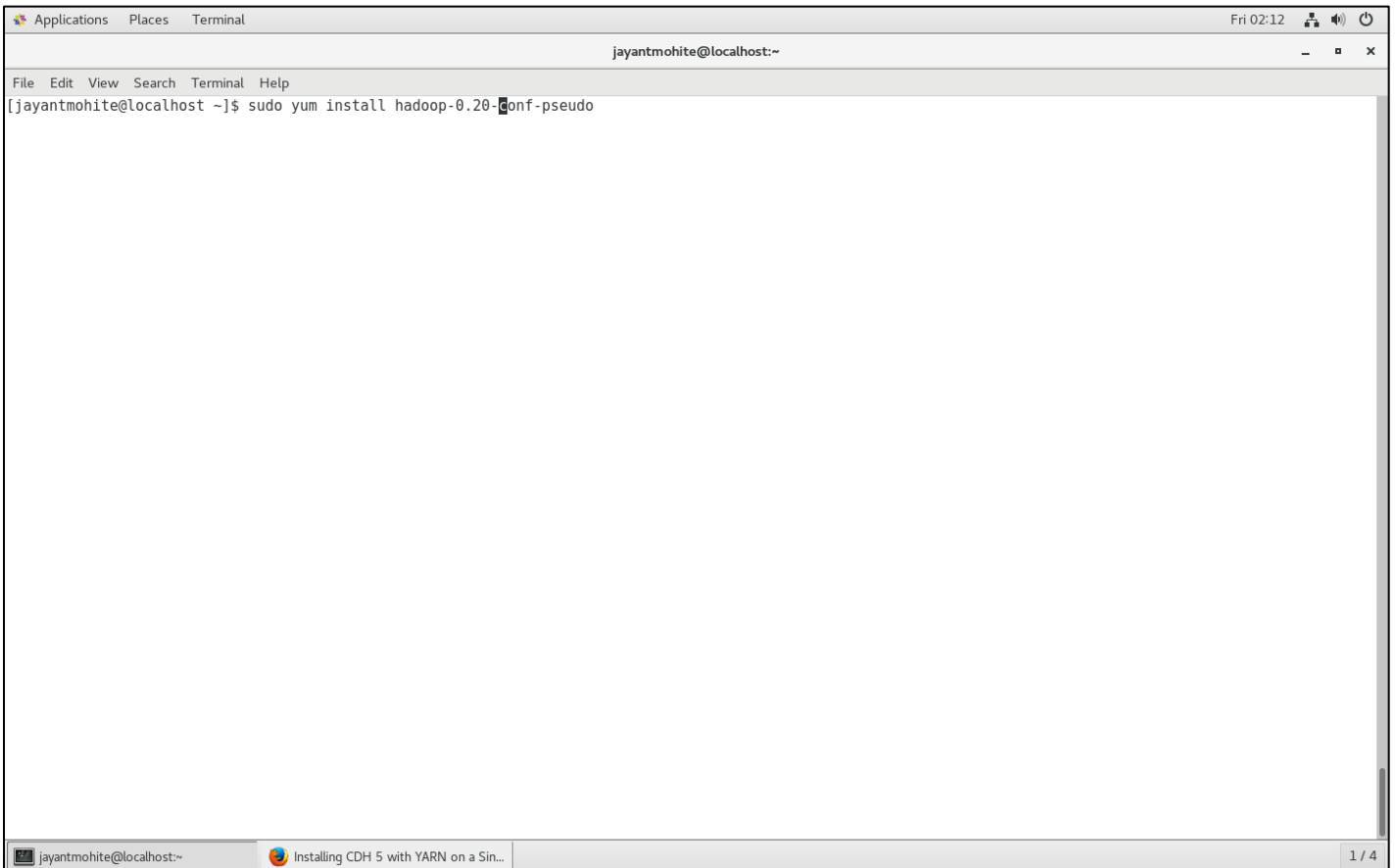
=====
Package Arch Version Repository Size
=====
Installing:
cloudera-cdh x86_64 5-0 /cloudera-cdh-5-0.x86_64 13 k
Transaction Summary
Install 1 Package

Total size: 13 k
Installed size: 13 k
Is this ok [y/d/N]:
```

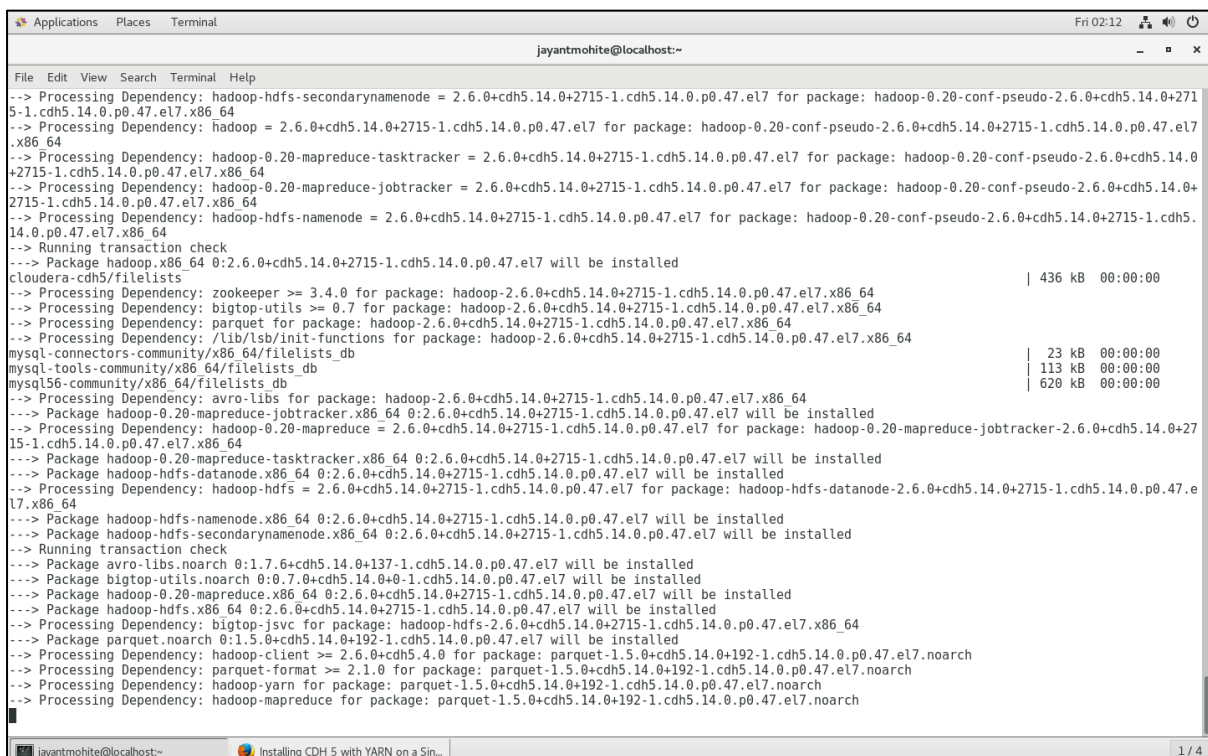
type y

## Step 3:

## Install Hadoop Generation 1 Framework



```
Applications  Places  Terminal  Fri 02:12  [jayantmohite@localhost:~]
jayantmohite@localhost:~
File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ sudo yum install hadoop-0.20-conf-pseudo
```



```
Applications  Places  Terminal  Fri 02:12  [jayantmohite@localhost:~]
jayantmohite@localhost:~
File Edit View Search Terminal Help
--> Processing Dependency: hadoop-hdfs-secondarynamenode = 2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 for package: hadoop-0.20-conf-pseudo-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64
--> Processing Dependency: hadoop = 2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 for package: hadoop-0.20-conf-pseudo-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64
--> Processing Dependency: hadoop-0.20-mapreduce-tasktracker = 2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 for package: hadoop-0.20-conf-pseudo-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64
--> Processing Dependency: hadoop-0.20-mapreduce-jobtracker = 2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 for package: hadoop-0.20-conf-pseudo-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64
--> Processing Dependency: hadoop-hdfs-namenode = 2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 for package: hadoop-0.20-conf-pseudo-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64
--> Running transaction check
--> Package hadoop.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 will be installed
cloudera-cdh5/filelists
--> Processing Dependency: zookeeper >= 3.4.0 for package: hadoop-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64 | 436 kB 00:00:00
--> Processing Dependency: bigtop-utils >= 0.7 for package: hadoop-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64
--> Processing Dependency: parquet for package: hadoop-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64 | 113 kB 00:00:00
--> Processing Dependency: /lib/lsb/init-functions for package: hadoop-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64 | 620 kB 00:00:00
mysql-connectors-community/x86_64/filelists.db
mysql-tools-community/x86_64/filelists.db
mysql56-community/x86_64/filelists.db
--> Processing Dependency: avro-libs for package: hadoop-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64
--> Package hadoop-0.20-mapreduce-jobtracker.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: hadoop-0.20-mapreduce = 2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 for package: hadoop-0.20-mapreduce-jobtracker-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64
--> Package hadoop-0.20-mapreduce-tasktracker.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 will be installed
--> Package hadoop-hdfs-datanode.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: hadoop-hdfs = 2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 for package: hadoop-hdfs-datanode-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64
--> Package hadoop-hdfs-namenode.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 will be installed
--> Package hadoop-hdfs-secondarynamenode.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 will be installed
--> Running transaction check
--> Package avro-libs.noarch 0:1.7.6+cdh5.14.0+137-1.cdh5.14.0.p0.47.el7 will be installed
--> Package bigtop-utils.noarch 0:0.7.0+cdh5.14.0+0-1.cdh5.14.0.p0.47.el7 will be installed
--> Package hadoop-0.20-mapreduce.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 will be installed
--> Package hadoop-hdfs.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: bigtop-jsvc for package: hadoop-hdfs-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64
--> Package parquet.noarch 0:1.5.0+cdh5.14.0+192-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: hadoop-client >= 2.6.0+cdh5.4.0 for package: parquet-1.5.0+cdh5.14.0+192-1.cdh5.14.0.p0.47.el7.noarch
--> Processing Dependency: parquet-format >= 2.1.0 for package: parquet-1.5.0+cdh5.14.0+192-1.cdh5.14.0.p0.47.el7.noarch
--> Processing Dependency: hadoop-yarn for package: parquet-1.5.0+cdh5.14.0+192-1.cdh5.14.0.p0.47.el7.noarch
--> Processing Dependency: hadoop-mapreduce for package: parquet-1.5.0+cdh5.14.0+192-1.cdh5.14.0.p0.47.el7.noarch
```

```

Applications  Places  Terminal
Fri 02:12

jayantmohite@localhost:~
File Edit View Search Terminal Help

--> Package parquet-format.noarch 0:2.1.0+cdh5.14.0+19-1.cdh5.14.0.p0.47.el7 will be installed
--> Package redhat-lsb-submod-security.x86_64 0:4.1-27.el7.centos.1 will be installed
--> Package spax.x86_64 0:1.5.2-13.el7 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

=====
Package                               Arch           Version                               Repository      Size
=====
Installing:
hadoop-0.20-conf-pseudo                x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  9.7 k
Installing for dependencies:
avro-libs                              noarch         1.7.6+cdh5.14.0+137-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  94 M
bigtop-jsvc                            x86_64         0.6.0+cdh5.14.0+917-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  28 k
bigtop-utils                           noarch         0.7.0+cdh5.14.0+0-1.cdh5.14.0.p0.47.el7    cloudera-cdh5  11 k
hadoop                                  x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  75 M
hadoop-0.20-mapreduce                  x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  84 M
hadoop-0.20-mapreduce-jobtracker       x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  5.2 k
hadoop-0.20-mapreduce-tasktracker      x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  5.1 k
hadoop-client                          x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  34 k
hadoop-hdfs                             x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  24 M
hadoop-hdfs-datanode                   x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  5.2 k
hadoop-hdfs-namenode                   x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  5.4 k
hadoop-hdfs-secondarynamenode         x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  5.2 k
hadoop-mapreduce                       x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  32 M
hadoop-yarn                            x86_64         2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  15 M
m4                                      x86_64         1.4.16-10.el7                                base           256 k
parquet                                noarch         1.5.0+cdh5.14.0+192-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  76 M
parquet-format                         noarch         2.1.0+cdh5.14.0+19-1.cdh5.14.0.p0.47.el7    cloudera-cdh5  465 k
redhat-lsb-core                        x86_64         4.1-27.el7.centos.1                          base           38 k
redhat-lsb-submod-security             x86_64         4.1-27.el7.centos.1                          base           15 k
spax                                    x86_64         1.5.2-13.el7                                base          260 k
zookeeper                              x86_64         3.4.5+cdh5.14.0+136-1.cdh5.14.0.p0.47.el7  cloudera-cdh5  4.2 M

Transaction Summary
=====
Install 1 Package (+21 Dependent packages)

Total download size: 405 M
Installed size: 488 M
Is this ok [y/d/N]: █

type y

```

```

Applications  Places  Terminal
Fri 02:14

jayantmohite@localhost:~
File Edit View Search Terminal Help

Transaction Summary
=====
Install 1 Package (+21 Dependent packages)

Total download size: 405 M
Installed size: 488 M
Is this ok [y/d/N]: y
Downloading packages:
warning: /var/cache/yum/x86_64/7/cloudera-cdh5/packages/bigtop-jsvc-0.6.0+cdh5.14.0+917-1.cdh5.14.0.p0.47.el7.x86_64.rpm: Header V4 DSA/SHA1 Signature, key ID
e8f86acd: NOKEY
Public key for bigtop-jsvc-0.6.0+cdh5.14.0+917-1.cdh5.14.0.p0.47.el7.x86_64.rpm is not installed
(1/22): bigtop-jsvc-0.6.0+cdh5.14.0+917-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 28 kB 00:00:00
(2/22): bigtop-utils-0.7.0+cdh5.14.0+0-1.cdh5.14.0.p0.47.el7.noarch.rpm | 11 kB 00:00:00
(3/22): hadoop-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 75 MB 00:00:33
(4/22): hadoop-0.20-conf-pseudo-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 9.7 kB 00:00:00
(5/22): avro-libs-1.7.6+cdh5.14.0+137-1.cdh5.14.0.p0.47.el7.noarch.rpm | 94 MB 00:00:41
(6/22): hadoop-0.20-mapreduce-jobtracker-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 5.2 kB 00:00:00
(7/22): hadoop-0.20-mapreduce-tasktracker-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 5.1 kB 00:00:00
(8/22): hadoop-client-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 34 kB 00:00:00
(9/22): hadoop-hdfs-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 24 MB 00:00:08
(10/22): hadoop-hdfs-datanode-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 5.2 kB 00:00:00
(11/22): hadoop-hdfs-namenode-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 5.4 kB 00:00:00
(12/22): hadoop-hdfs-secondarynamenode-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 5.2 kB 00:00:00
(13/22): hadoop-0.20-mapreduce-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 84 MB 00:00:30
(14/22): m4-1.4.16-10.el7.x86_64.rpm | 256 kB 00:00:01
(15/22): hadoop-mapreduce-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 32 MB 00:00:18
(16/22): hadoop-yarn-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 15 MB 00:00:03
(17/22): redhat-lsb-core-4.1-27.el7.centos.1.x86_64.rpm | 38 kB 00:00:00
(18/22): parquet-format-2.1.0+cdh5.14.0+19-1.cdh5.14.0.p0.47.el7.noarch.rpm | 465 kB 00:00:00
(19/22): spax-1.5.2-13.el7.x86_64.rpm | 260 kB 00:00:00
(20/22): redhat-lsb-submod-security-4.1-27.el7.centos.1.x86_64.rpm | 15 kB 00:00:01
(21/22): zookeeper-3.4.5+cdh5.14.0+136-1.cdh5.14.0.p0.47.el7.x86_64.rpm | 4.2 MB 00:00:01
(22/22): parquet-1.5.0+cdh5.14.0+192-1.cdh5.14.0.p0.47.el7.noarch.rpm | 76 MB 00:00:16

Total 4.7 MB/s | 405 MB 00:01:26
Retrieving key from https://archive.cloudera.com/cdh5/redhat/7/x86_64/cdh/RPM-GPG-KEY-cloudera
Importing GPG key 0xE8F86ACD:
  Userid : "Yum Maintainer <webmaster@cloudera.com>"
  Fingerprint: 5f14 d39e f068 laca 6f04 4a43 f90c 0d8f e8f8 6acd
  From : https://archive.cloudera.com/cdh5/redhat/7/x86_64/cdh/RPM-GPG-KEY-cloudera
Is this ok [y/N]: █

type y

```

```

Applications  Places  Terminal  Fri 02:16
jayantmohite@localhost:~

File Edit View Search Terminal Help

Verifying : hadoop-client-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64 10/22
Verifying : bigtop-utils-0.7.0+cdh5.14.0+0-1.cdh5.14.0.p0.47.el7.noarch 11/22
Verifying : hadoop-hdfs-namenode-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64 12/22
Verifying : hadoop-0.20-mapreduce-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64 13/22
Verifying : bigtop-jsvc-0.6.0+cdh5.14.0+917-1.cdh5.14.0.p0.47.el7.x86_64 14/22
Verifying : hadoop-hdfs-secondarynamenode-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64 15/22
Verifying : hadoop-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64 16/22
Verifying : parquet-format-2.1.0+cdh5.14.0+19-1.cdh5.14.0.p0.47.el7.noarch 17/22
Verifying : hadoop-hdfs-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64 18/22
Verifying : zookeeper-3.4.5+cdh5.14.0+136-1.cdh5.14.0.p0.47.el7.x86_64 19/22
Verifying : spax-1.5.2-13.el7.x86_64 20/22
Verifying : redhat-lsb-core-4.1-27.el7.centos.1.x86_64 21/22
Verifying : hadoop-mapreduce-2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7.x86_64 22/22

Installed:
hadoop-0.20-conf-pseudo.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7

Dependency Installed:
avro-lbns.noarch 0:1.7.6+cdh5.14.0+137-1.cdh5.14.0.p0.47.el7
bigtop-jsvc.x86_64 0:0.6.0+cdh5.14.0+917-1.cdh5.14.0.p0.47.el7
bigtop-utils.noarch 0:0.7.0+cdh5.14.0+0-1.cdh5.14.0.p0.47.el7
hadoop.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
hadoop-0.20-mapreduce.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
hadoop-0.20-mapreduce-jobtracker.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
hadoop-0.20-mapreduce-tasktracker.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
hadoop-client.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
hadoop-hdfs.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
hadoop-hdfs-datanode.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
hadoop-hdfs-namenode.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
hadoop-hdfs-secondarynamenode.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
hadoop-mapreduce.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
hadoop-yarn.x86_64 0:2.6.0+cdh5.14.0+2715-1.cdh5.14.0.p0.47.el7
m4.x86_64 0:1.4.16-10.el7
parquet.noarch 0:1.5.0+cdh5.14.0+192-1.cdh5.14.0.p0.47.el7
parquet-format.noarch 0:2.1.0+cdh5.14.0+19-1.cdh5.14.0.p0.47.el7
redhat-lsb-core.x86_64 0:4.1-27.el7.centos.1
redhat-lsb-submod-security.x86_64 0:4.1-27.el7.centos.1
spax.x86_64 0:1.5.2-13.el7
zookeeper.x86_64 0:3.4.5+cdh5.14.0+136-1.cdh5.14.0.p0.47.el7

Complete!
[jayantmohite@localhost ~]$ █

```

Step 4:

Verify if Execution Happened Properly

```

Applications  Places  Terminal  Fri 02:21
jayantmohite@localhost:~

File Edit View Search Terminal Help

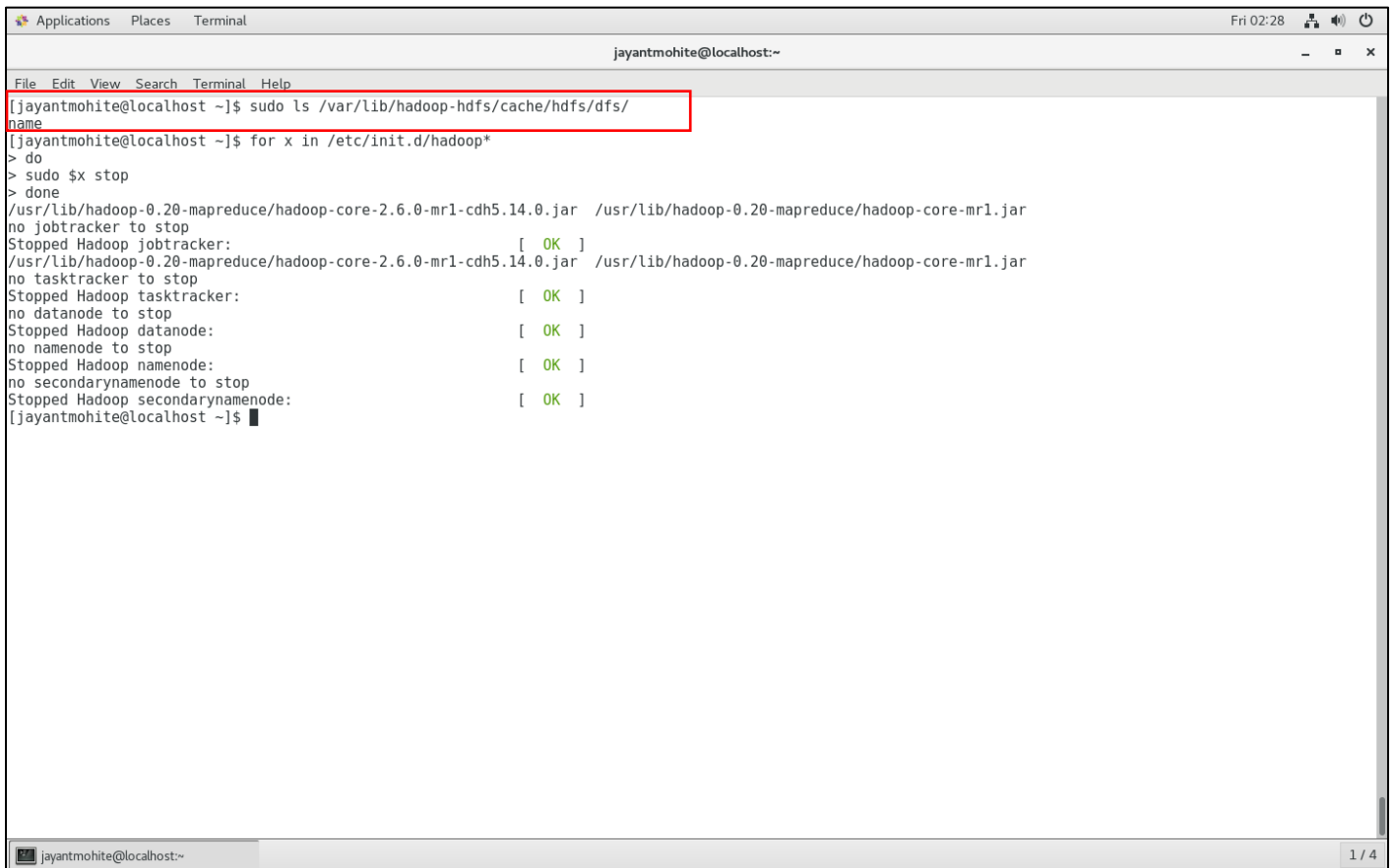
[jayantmohite@localhost ~]$ rpm -ql hadoop-0.20-conf-pseudo
/etc/hadoop/conf.pseudo.mr1
/etc/hadoop/conf.pseudo.mr1/README
/etc/hadoop/conf.pseudo.mr1/core-site.xml
/etc/hadoop/conf.pseudo.mr1/fair-scheduler.xml
/etc/hadoop/conf.pseudo.mr1/hadoop-env.sh
/etc/hadoop/conf.pseudo.mr1/hadoop-metrics.properties
/etc/hadoop/conf.pseudo.mr1/hdfs-site.xml
/etc/hadoop/conf.pseudo.mr1/log4j.properties
/etc/hadoop/conf.pseudo.mr1/mapred-site.xml
[jayantmohite@localhost ~]$ █

```



Step 6:

## Stop All Hadoop Services



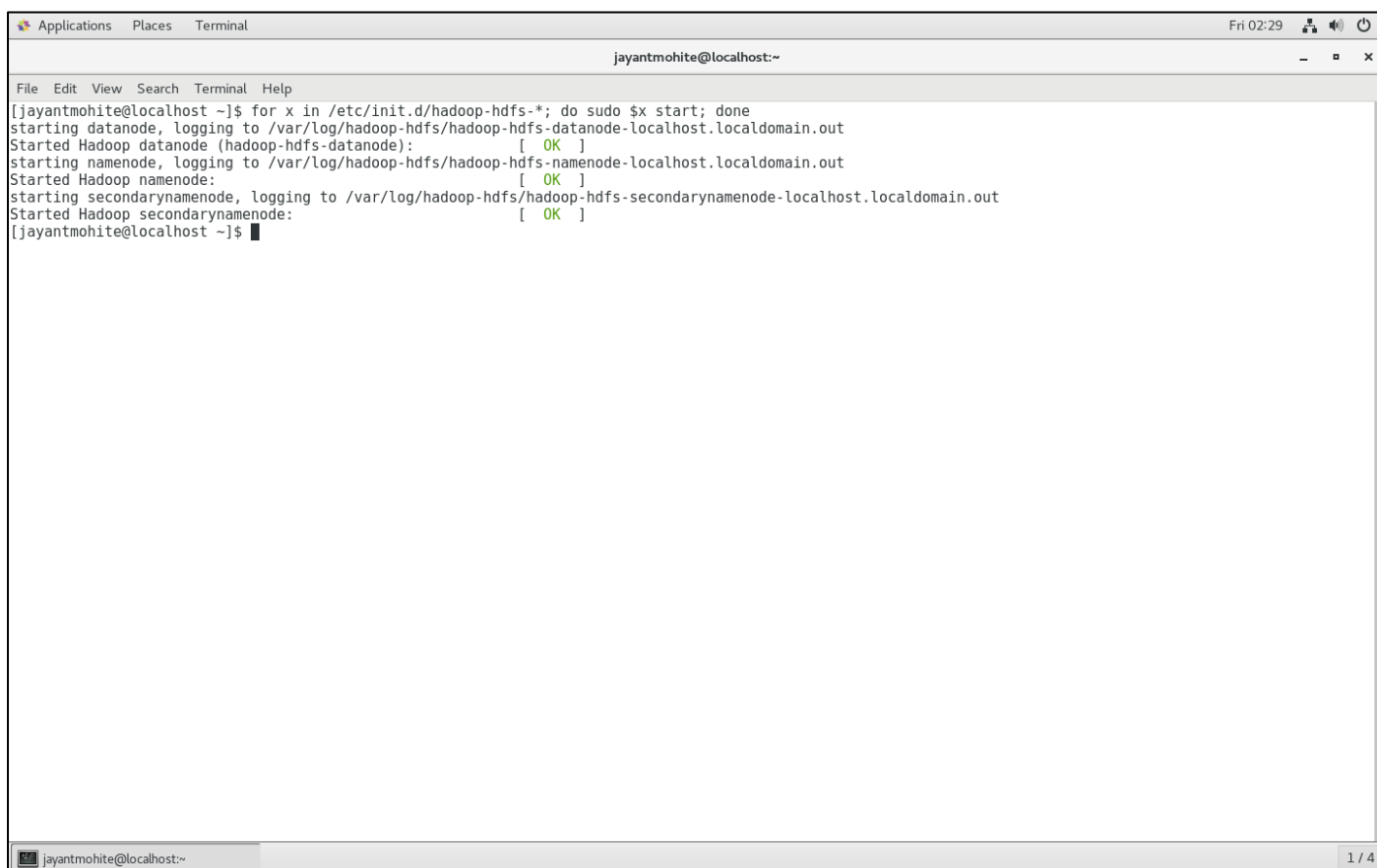
A terminal window titled 'Applications Places Terminal' with a status bar showing 'Fri 02:28'. The terminal shows the user 'jayantmohite@localhost' running a series of commands to stop Hadoop services. A red rectangle highlights the command 'sudo ls /var/lib/hadoop-hdfs/cache/hdfs/dfs/name'. The output shows that all Hadoop services (jobtracker, tasktracker, datanode, namenode, and secondarynamenode) have been successfully stopped, each with a green '[ OK ]' status.

```
[jayantmohite@localhost ~]$ sudo ls /var/lib/hadoop-hdfs/cache/hdfs/dfs/
name
[jayantmohite@localhost ~]$ for x in /etc/init.d/hadoop*
> do
> sudo $x stop
> done
/usr/lib/hadoop-0.20-mapreduce/hadoop-core-2.6.0-mr1-cdh5.14.0.jar /usr/lib/hadoop-0.20-mapreduce/hadoop-core-mr1.jar
no jobtracker to stop
Stopped Hadoop jobtracker: [ OK ]
/usr/lib/hadoop-0.20-mapreduce/hadoop-core-2.6.0-mr1-cdh5.14.0.jar /usr/lib/hadoop-0.20-mapreduce/hadoop-core-mr1.jar
no tasktracker to stop
Stopped Hadoop tasktracker: [ OK ]
no datanode to stop
Stopped Hadoop datanode: [ OK ]
no namenode to stop
Stopped Hadoop namenode: [ OK ]
no secondarynamenode to stop
Stopped Hadoop secondarynamenode: [ OK ]
[jayantmohite@localhost ~]$
```

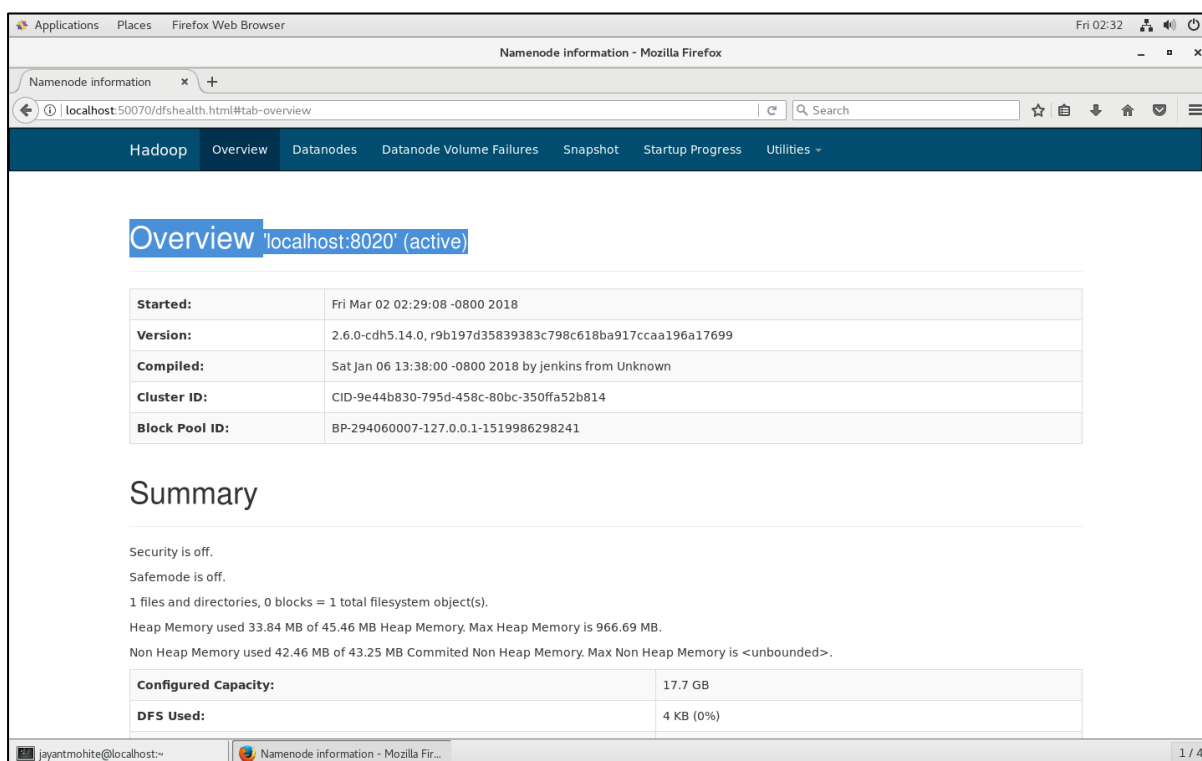


Step 7:

## Start Only HDFS Services



```
Applications  Places  Terminal  Fri 02:29
jayantmohite@localhost:~$ for x in /etc/init.d/hadoop-hdfs-*; do sudo $x start; done
starting datanode, logging to /var/log/hadoop-hdfs/hadoop-hdfs-datanode-localhost.localdomain.out
Started Hadoop datanode (hadoop-hdfs-datanode): [ OK ]
starting namenode, logging to /var/log/hadoop-hdfs/hadoop-hdfs-namenode-localhost.localdomain.out
Started Hadoop namenode: [ OK ]
starting secondarynamenode, logging to /var/log/hadoop-hdfs/hadoop-hdfs-secondarynamenode-localhost.localdomain.out
Started Hadoop secondarynamenode: [ OK ]
jayantmohite@localhost ~]$
```



Namenode information - Mozilla Firefox

localhost:50070/dfshealth.html#tab-overview

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

### Overview 'localhost:8020' (active)

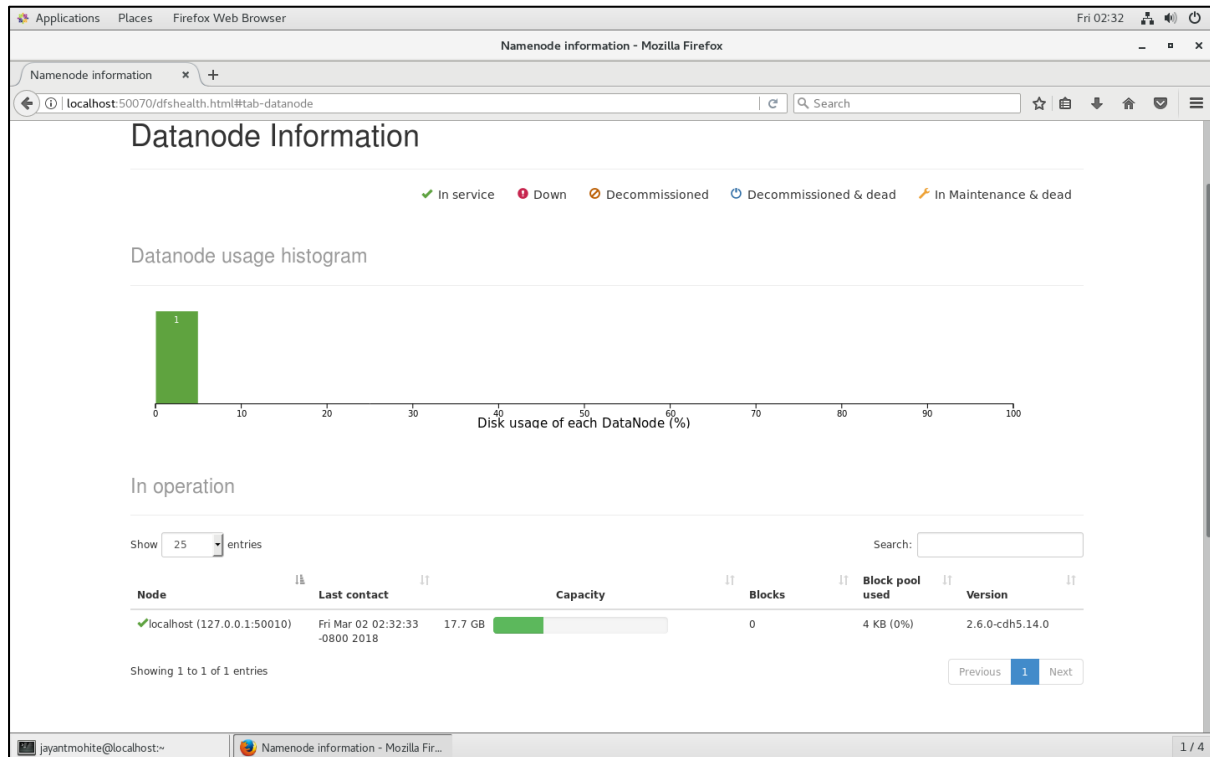
Started:	Fri Mar 02 02:29:08 -0800 2018
Version:	2.6.0-cdh5.14.0, r9b197d35839383c798c618ba917ccea196a17699
Compiled:	Sat Jan 06 13:38:00 -0800 2018 by jenkins from Unknown
Cluster ID:	CID-9e44b830-795d-458c-80bc-350ffa52b814
Block Pool ID:	BP-294060007-127.0.0.1-1519986298241

### Summary

Security is off.  
Safemode is off.  
1 files and directories, 0 blocks = 1 total filesystem object(s).  
Heap Memory used 33.84 MB of 45.46 MB Heap Memory. Max Heap Memory is 966.69 MB.  
Non Heap Memory used 42.46 MB of 43.25 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	17.7 GB
DFS Used:	4 KB (0%)





Applications Places Terminal

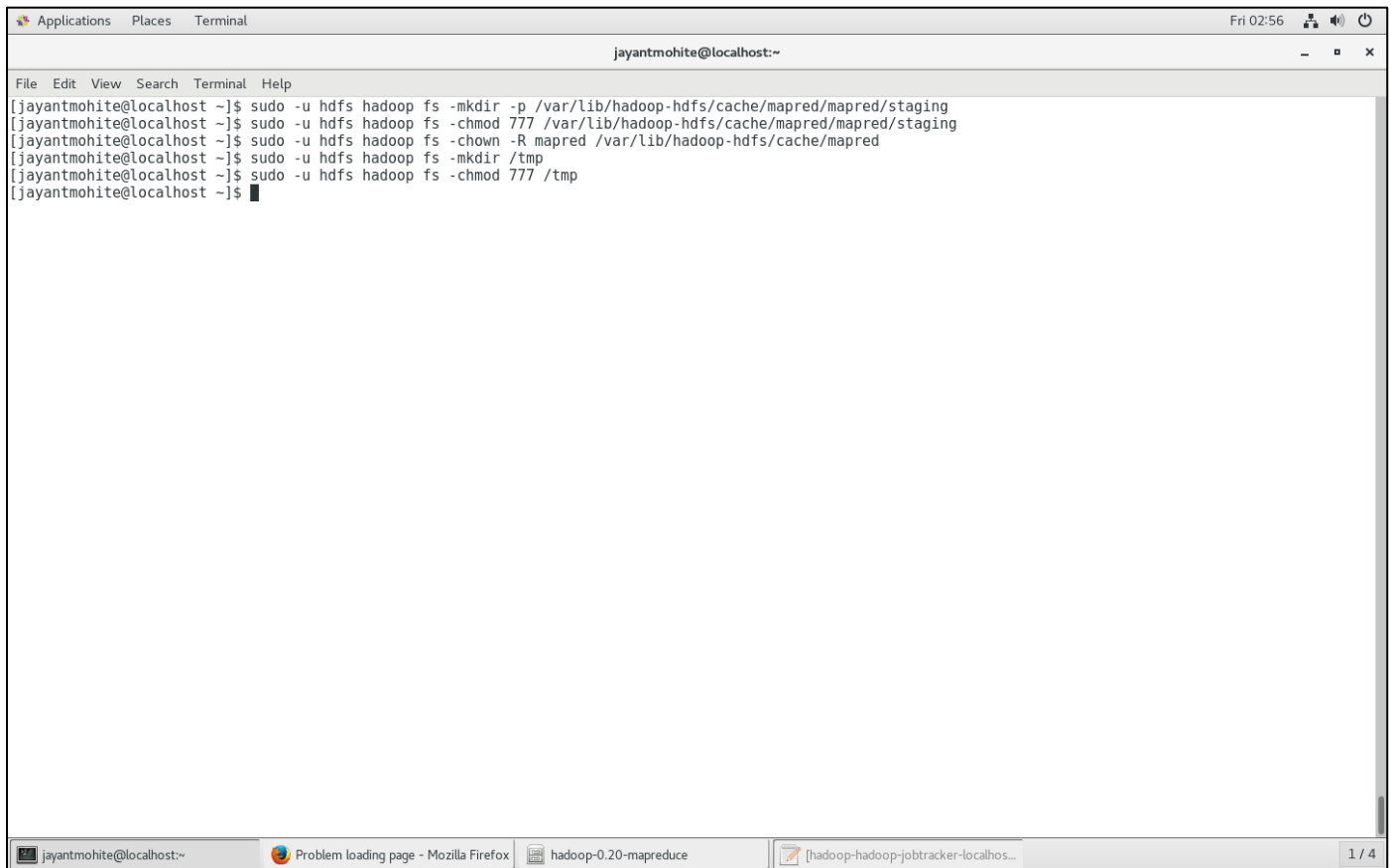
jayantmohite@localhost:~

```
File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ sudo ls /var/lib/hadoop-hdfs/cache/hdfs/dfs/
data name namessecondary
[jayantmohite@localhost ~]$
```

jayantmohite@localhost:~    Namenode information - Mozilla Fir...    1 / 4

Step 8:

Create the Required Directory Structure with Proper Permissions



A terminal window titled 'jayantmohite@localhost:~' showing the execution of four Hadoop commands to create a directory structure. The commands are: 1) 'sudo -u hdfs hadoop fs -mkdir -p /var/lib/hadoop-hdfs/cache/mapred/mapred/staging', 2) 'sudo -u hdfs hadoop fs -chmod 777 /var/lib/hadoop-hdfs/cache/mapred/mapred/staging', 3) 'sudo -u hdfs hadoop fs -chown -R mapred /var/lib/hadoop-hdfs/cache/mapred', and 4) 'sudo -u hdfs hadoop fs -mkdir /tmp'. The prompt returns to '~\$' after each command. The terminal window has a menu bar (File, Edit, View, Search, Terminal, Help) and a status bar at the bottom showing the user 'jayantmohite@localhost:~' and four open tabs: 'Problem loading page - Mozilla Firefox', 'hadoop-0.20-mapreduce', 'hadoop-hadoop-jobtracker-localhos...', and '1 / 4'.

```
jayantmohite@localhost ~]$ sudo -u hdfs hadoop fs -mkdir -p /var/lib/hadoop-hdfs/cache/mapred/mapred/staging
[jayantmohite@localhost ~]$ sudo -u hdfs hadoop fs -chmod 777 /var/lib/hadoop-hdfs/cache/mapred/mapred/staging
[jayantmohite@localhost ~]$ sudo -u hdfs hadoop fs -chown -R mapred /var/lib/hadoop-hdfs/cache/mapred
[jayantmohite@localhost ~]$ sudo -u hdfs hadoop fs -mkdir /tmp
[jayantmohite@localhost ~]$ sudo -u hdfs hadoop fs -chmod 777 /tmp
[jayantmohite@localhost ~]$
```

## Step 9:

## Start All Map Reduce Services

```

Applications  Places  Terminal  Fri 02:56
jayantmohite@localhost:~

File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ for x in /etc/init.d/hadoop-0.20-mapreduce-*; do sudo $x start; done
/usr/lib/hadoop-0.20-mapreduce/hadoop-core-2.6.0-mr1-cdh5.14.0.jar /usr/lib/hadoop-0.20-mapreduce/hadoop-core-mr1.jar
starting jobtracker, logging to /var/log/hadoop-0.20-mapreduce/hadoop-hadoop-jobtracker-localhost.localdomain.out
/usr/lib/hadoop-0.20-mapreduce/hadoop-core-2.6.0-mr1-cdh5.14.0.jar /usr/lib/hadoop-0.20-mapreduce/hadoop-core-mr1.jar
Started Hadoop jobtracker: [ OK ]
/usr/lib/hadoop-0.20-mapreduce/hadoop-core-2.6.0-mr1-cdh5.14.0.jar /usr/lib/hadoop-0.20-mapreduce/hadoop-core-mr1.jar
starting tasktracker, logging to /var/log/hadoop-0.20-mapreduce/hadoop-hadoop-tasktracker-localhost.localdomain.out
/usr/lib/hadoop-0.20-mapreduce/hadoop-core-2.6.0-mr1-cdh5.14.0.jar /usr/lib/hadoop-0.20-mapreduce/hadoop-core-mr1.jar
Started Hadoop tasktracker: [ OK ]
[jayantmohite@localhost ~]$

```

Applications Places Firefox Web Browser Fri 02:56

localhost Hadoop Map/Reduce Administration - Mozilla Firefox

Namenode information x / localhost Hadoop Ma... x +

localhost:50030/jobtracker.jsp

### localhost Hadoop Map/Reduce Administration

State: RUNNING  
 Started: Fri Mar 02 02:56:25 PST 2018  
 Version: 2.6.0-cdh5.14.0, 9b197d35839383c798c618ba917ccaa196a17699  
 Compiled: 2018-01-06T21:38Z by Jenkins from Unknown  
 Identifier: 201803020256

**Cluster Summary (Heap Size is 15.56 MB/966.69 MB)**

Running Map Tasks	Running Reduce Tasks	Total Submissions	Nodes	Occupied Map Slots	Occupied Reduce Slots	Reserved Map Slots	Reserved Reduce Slots	Map Task Capacity	Reduce Task Capacity	Avg. Tasks/Node	Blacklisted Nodes	Excluded Nodes
0	0	0	1	0	0	0	0	2	2	4.00	0	0

**Scheduling Information**

Queue Name	State	Scheduling Information
default	running	N/A

Filter (Jobid, Priority, User, Name)   
 Example: 'user:smith 3200' will filter by 'smith' only in the user field and '3200' in all fields

**Running Jobs**

**Retired Jobs**

**Local Logs**  
[Log directory, Job Tracker History](#)  
[Hadoop, 2018.](#)

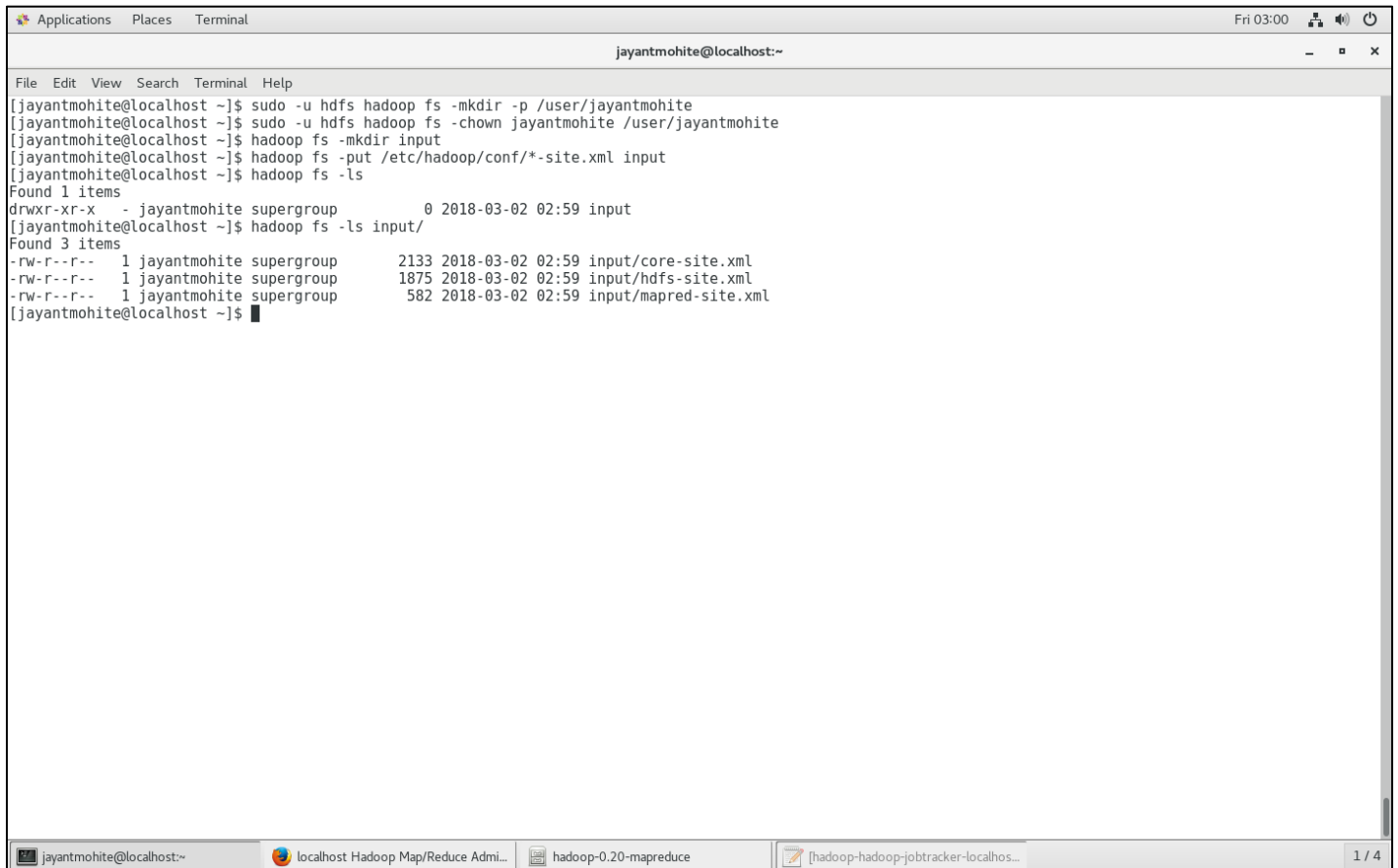
jayantmohite@localhost:~ localhost Hadoop Map/Reduce Admi... hadoop-0.20-mapreduce [hadoop-hadoop-jobtracker-localhos... 1 / 4

Step 10:

Create Directory for User with Proper Permissions

Step 11:

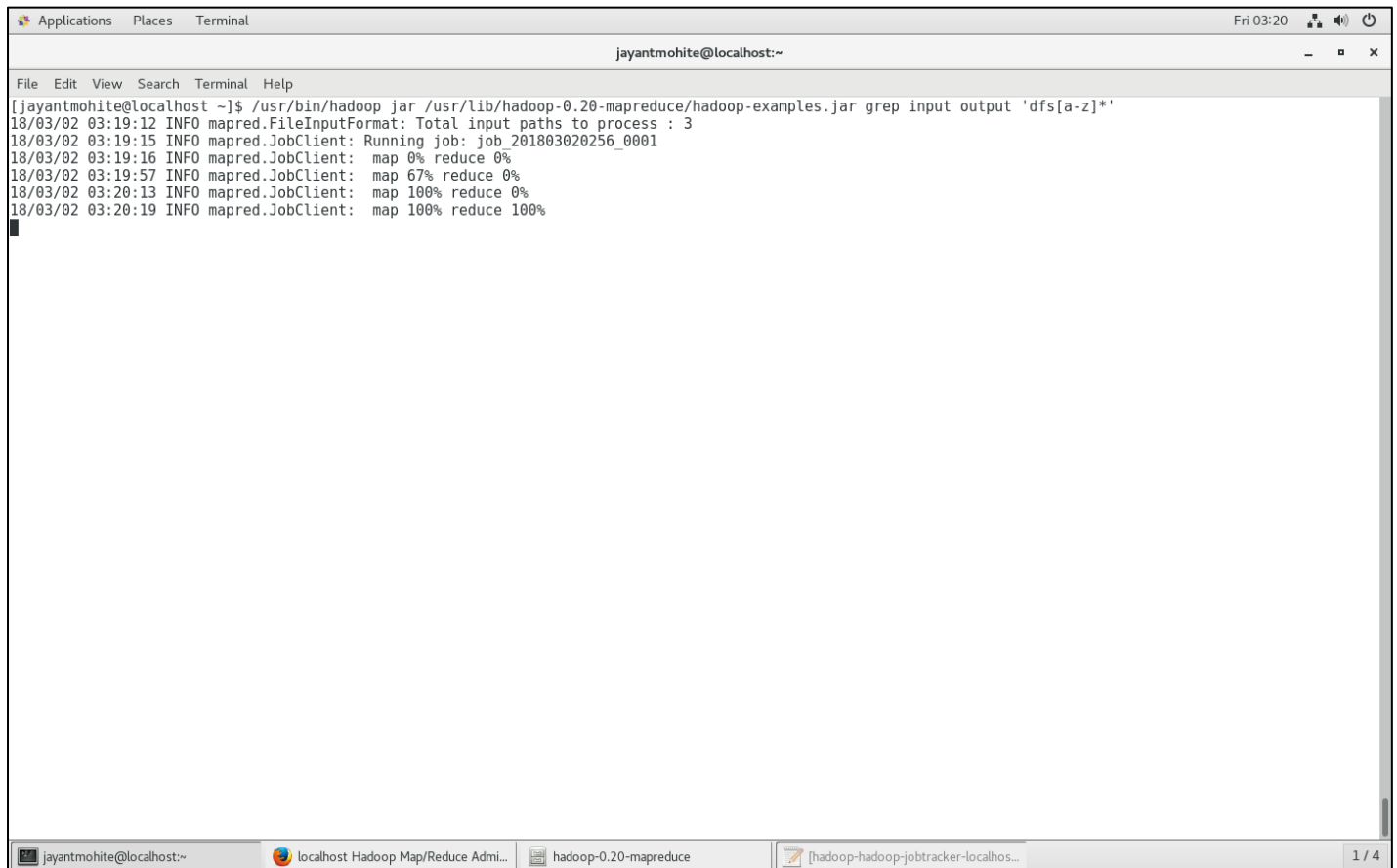
Create Sample Input



```
Applications Places Terminal Fri 03:00
jayantmohite@localhost:~
File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ sudo -u hdfs hadoop fs -mkdir -p /user/jayantmohite
[jayantmohite@localhost ~]$ sudo -u hdfs hadoop fs -chown jayantmohite /user/jayantmohite
[jayantmohite@localhost ~]$ hadoop fs -mkdir input
[jayantmohite@localhost ~]$ hadoop fs -put /etc/hadoop/conf/*-site.xml input
[jayantmohite@localhost ~]$ hadoop fs -ls
Found 1 items
drwxr-xr-x - jayantmohite supergroup 0 2018-03-02 02:59 input
[jayantmohite@localhost ~]$ hadoop fs -ls input/
Found 3 items
-rw-r--r-- 1 jayantmohite supergroup 2133 2018-03-02 02:59 input/core-site.xml
-rw-r--r-- 1 jayantmohite supergroup 1875 2018-03-02 02:59 input/hdfs-site.xml
-rw-r--r-- 1 jayantmohite supergroup 582 2018-03-02 02:59 input/mapred-site.xml
[jayantmohite@localhost ~]$
```

Step 12:

Execute Sample Program



```
File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ /usr/bin/hadoop jar /usr/lib/hadoop-0.20-mapreduce/hadoop-examples.jar grep input output 'dfs[a-z]*'
18/03/02 03:19:12 INFO mapred.FileInputFormat: Total input paths to process : 3
18/03/02 03:19:15 INFO mapred.JobClient: Running job: job_201803020256_0001
18/03/02 03:19:16 INFO mapred.JobClient: map 0% reduce 0%
18/03/02 03:19:57 INFO mapred.JobClient: map 67% reduce 0%
18/03/02 03:20:13 INFO mapred.JobClient: map 100% reduce 0%
18/03/02 03:20:19 INFO mapred.JobClient: map 100% reduce 100%
```

## Step 13:

## Install Sqoop

```

Applications  Places  Terminal
Fri 03:22
jayantmohite@localhost:~$ sudo yum install sqoop
[sudo] password for jayantmohite:
Loaded plugins: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
 * base: ftp.iitm.ac.in
 * extras: ftp.iitm.ac.in
 * updates: ftp.iitm.ac.in
Resolving Dependencies
--> Running transaction check
--> Package sqoop.noarch 0:1.4.6+cdh5.14.0+127-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: kite for package: sqoop-1.4.6+cdh5.14.0+127-1.cdh5.14.0.p0.47.el7.noarch
--> Running transaction check
--> Package kite.noarch 0:1.0.0+cdh5.14.0+146-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: hive-jdbc >= 1.1.0+cdh5.1.0 for package: kite-1.0.0+cdh5.14.0+146-1.cdh5.14.0.p0.47.el7.noarch
--> Processing Dependency: solr >= 4.4.0+cdh5.1.0 for package: kite-1.0.0+cdh5.14.0+146-1.cdh5.14.0.p0.47.el7.noarch
--> Running transaction check
--> Package sentry.noarch 0:1.5.1+cdh5.14.0+432-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: hive-jdbc >= 1.1.0+cdh5.4.0 for package: sentry-1.5.1+cdh5.14.0+432-1.cdh5.14.0.p0.47.el7.noarch
--> Processing Dependency: hive >= 1.1.0+cdh5.4.0 for package: sentry-1.5.1+cdh5.14.0+432-1.cdh5.14.0.p0.47.el7.noarch
--> Package solr.noarch 0:4.10.3+cdh5.14.0+522-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: bigtop-tomcat for package: solr-4.10.3+cdh5.14.0+522-1.cdh5.14.0.p0.47.el7.noarch
--> Running transaction check
--> Package bigtop-tomcat.noarch 0:0.7.0+cdh5.14.0+0-1.cdh5.14.0.p0.47.el7 will be installed
--> Package hive.noarch 0:1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7 will be installed
--> Package hive-jdbc.noarch 0:1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

=====
Package                                Arch                                Version                                Repository                                Size
=====
Installing:
sqoop                                  noarch                              1.4.6+cdh5.14.0+127-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                18 M
Installing for dependencies:
bigtop-tomcat                         noarch                              0.7.0+cdh5.14.0+0-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                7.4 M
hive                                  noarch                              1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                47 M
hive-jdbc                             noarch                              1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                45 M
kite                                  noarch                              1.0.0+cdh5.14.0+146-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                104 M
sentry                                noarch                              1.5.1+cdh5.14.0+432-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                57 M
solr                                  noarch                              4.10.3+cdh5.14.0+522-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                75 M
=====
Transaction Summary
Install 1 Package (+6 Dependent packages)
Total download size: 354 M
Installed size: 447 M
Is this ok [y/d/N]:

```

```

Applications  Places  Terminal
Fri 03:22
jayantmohite@localhost:~$ sudo yum install sqoop
[sudo] password for jayantmohite:
Loaded plugins: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
 * base: ftp.iitm.ac.in
 * extras: ftp.iitm.ac.in
 * updates: ftp.iitm.ac.in
Resolving Dependencies
--> Running transaction check
--> Package sqoop.noarch 0:1.4.6+cdh5.14.0+127-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: kite for package: sqoop-1.4.6+cdh5.14.0+127-1.cdh5.14.0.p0.47.el7.noarch
--> Running transaction check
--> Package kite.noarch 0:1.0.0+cdh5.14.0+146-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: hive-jdbc >= 1.1.0+cdh5.1.0 for package: kite-1.0.0+cdh5.14.0+146-1.cdh5.14.0.p0.47.el7.noarch
--> Processing Dependency: solr >= 4.4.0+cdh5.1.0 for package: kite-1.0.0+cdh5.14.0+146-1.cdh5.14.0.p0.47.el7.noarch
--> Running transaction check
--> Package sentry.noarch 0:1.5.1+cdh5.14.0+432-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: hive-jdbc >= 1.1.0+cdh5.4.0 for package: sentry-1.5.1+cdh5.14.0+432-1.cdh5.14.0.p0.47.el7.noarch
--> Processing Dependency: hive >= 1.1.0+cdh5.4.0 for package: sentry-1.5.1+cdh5.14.0+432-1.cdh5.14.0.p0.47.el7.noarch
--> Package solr.noarch 0:4.10.3+cdh5.14.0+522-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: bigtop-tomcat for package: solr-4.10.3+cdh5.14.0+522-1.cdh5.14.0.p0.47.el7.noarch
--> Running transaction check
--> Package bigtop-tomcat.noarch 0:0.7.0+cdh5.14.0+0-1.cdh5.14.0.p0.47.el7 will be installed
--> Package hive.noarch 0:1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7 will be installed
--> Package hive-jdbc.noarch 0:1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

=====
Package                                Arch                                Version                                Repository                                Size
=====
Installing:
sqoop                                  noarch                              1.4.6+cdh5.14.0+127-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                18 M
Installing for dependencies:
bigtop-tomcat                         noarch                              0.7.0+cdh5.14.0+0-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                7.4 M
hive                                  noarch                              1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                47 M
hive-jdbc                             noarch                              1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                45 M
kite                                  noarch                              1.0.0+cdh5.14.0+146-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                104 M
sentry                                noarch                              1.5.1+cdh5.14.0+432-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                57 M
solr                                  noarch                              4.10.3+cdh5.14.0+522-1.cdh5.14.0.p0.47.el7  cloudera-cdh5                                75 M
=====
Transaction Summary
Install 1 Package (+6 Dependent packages)
Total download size: 354 M
Installed size: 447 M
Is this ok [y/d/N]:

```

type y

Step 14:

## Install Pig

```

Applications  Places  Terminal
jayantmohite@localhost:~

File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ sudo yum install pig
Loaded plugins: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
 * base: ftp.iitm.ac.in
 * extras: ftp.iitm.ac.in
 * updates: ftp.iitm.ac.in
Resolving Dependencies
--> Running transaction check
--> Package pig.noarch 0:0.12.0+cdh5.14.0+112-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: hive-hcatalog for package: pig-0.12.0+cdh5.14.0+112-1.cdh5.14.0.p0.47.el7.noarch
--> Running transaction check
--> Package hive-hcatalog.noarch 0:1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7 will be installed
--> Finished Dependency Resolution

```

```

Applications  Places  Terminal
jayantmohite@localhost:~

File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ sudo yum install pig
Loaded plugins: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
 * base: ftp.iitm.ac.in
 * extras: ftp.iitm.ac.in
 * updates: ftp.iitm.ac.in
Resolving Dependencies
--> Running transaction check
--> Package pig.noarch 0:0.12.0+cdh5.14.0+112-1.cdh5.14.0.p0.47.el7 will be installed
--> Processing Dependency: hive-hcatalog for package: pig-0.12.0+cdh5.14.0+112-1.cdh5.14.0.p0.47.el7.noarch
--> Running transaction check
--> Package hive-hcatalog.noarch 0:1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

=====
Package                Arch          Version                                Repository      Size
=====
Installing:
pig                    noarch        0.12.0+cdh5.14.0+112-1.cdh5.14.0.p0.47.el7    cloudera-cdh5   58 M
Installing for dependencies:
hive-hcatalog          noarch        1.1.0+cdh5.14.0+1330-1.cdh5.14.0.p0.47.el7    cloudera-cdh5   393 k
=====

Transaction Summary
-----
Install 1 Package (+1 Dependent package)

Total download size: 59 M
Installed size: 124 M
Is this ok [y/d/N]:

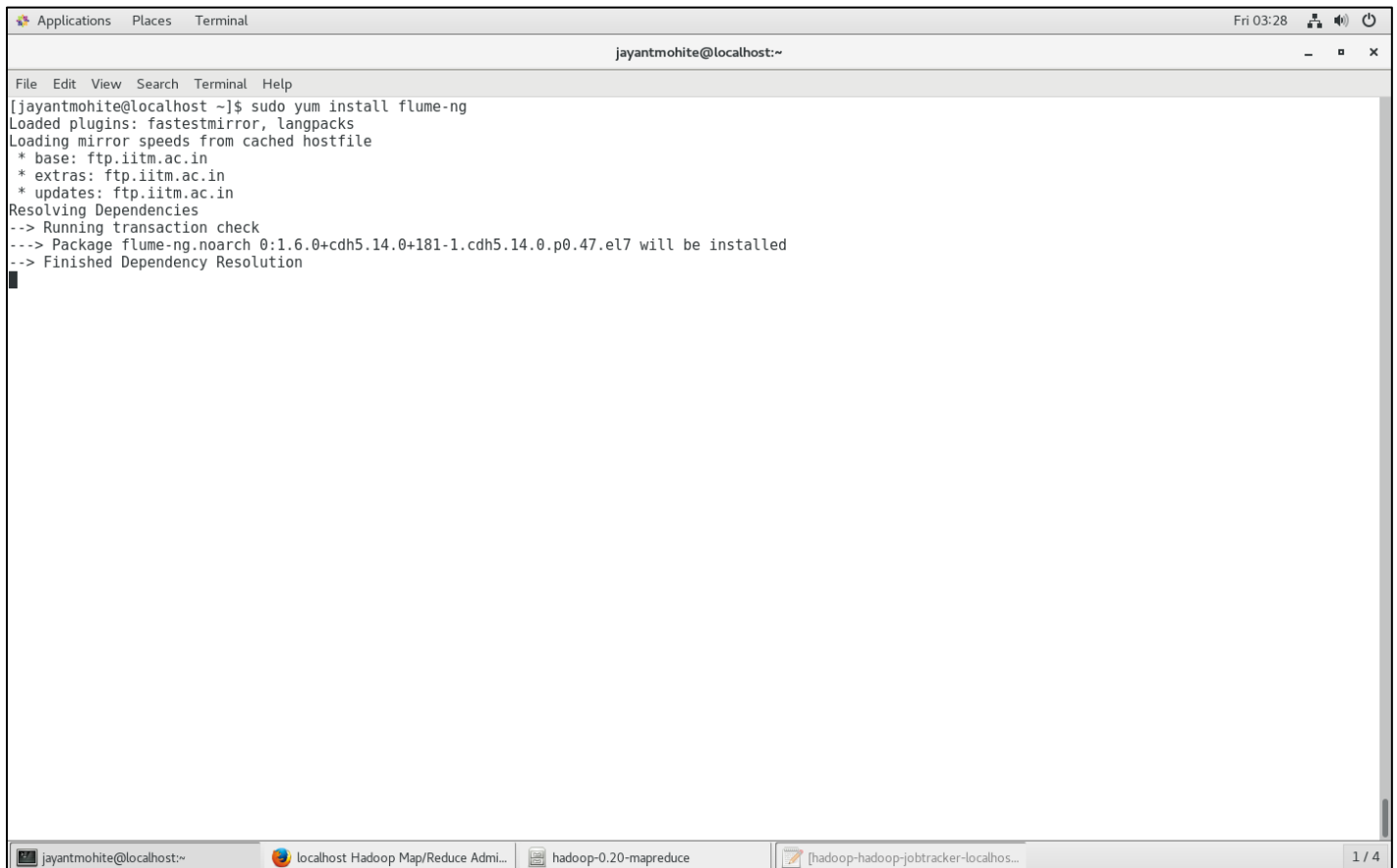
```

type y

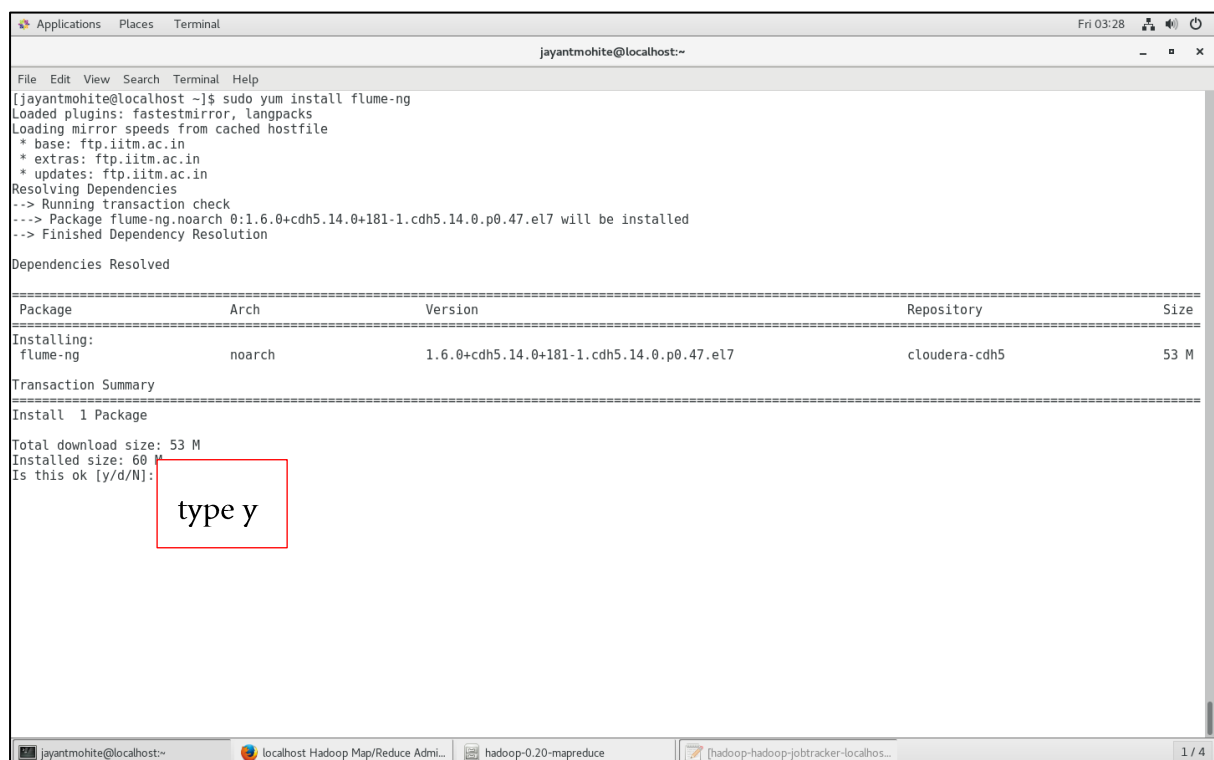


Step 15:

## Install Flume-ng



```
File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ sudo yum install flume-ng
Loaded plugins: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
* base: ftp.iitm.ac.in
* extras: ftp.iitm.ac.in
* updates: ftp.iitm.ac.in
Resolving Dependencies
--> Running transaction check
--> Package flume-ng.noarch 0:1.6.0+cdh5.14.0+181-1.cdh5.14.0.p0.47.el7 will be installed
--> Finished Dependency Resolution
```



```
File Edit View Search Terminal Help
[jayantmohite@localhost ~]$ sudo yum install flume-ng
Loaded plugins: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
* base: ftp.iitm.ac.in
* extras: ftp.iitm.ac.in
* updates: ftp.iitm.ac.in
Resolving Dependencies
--> Running transaction check
--> Package flume-ng.noarch 0:1.6.0+cdh5.14.0+181-1.cdh5.14.0.p0.47.el7 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

=====
Package                Arch             Version                               Repository          Size
=====
Installing:
flume-ng                noarch           1.6.0+cdh5.14.0+181-1.cdh5.14.0.p0.47.el7  cloudera-cdh5       53 M
=====
Transaction Summary
-----
Install 1 Package

Total download size: 53 M
Installed size: 60 M
Is this ok [y/d/N]:
```

type y

Now your Hadoop Cluster is ready and you can start performing your practice tasks on it.