

*****Cloudera Installation*****

Before Installation Stop firewall and SELinux.

Change Hostname

#hostnamectl set-hostname master

Disable SELINUX

vim /etc/selinux/config

SELINUX=disabled

:wq

savefile

#getenforce

#init 6

Disable Firewall

#systemctl stop firewalld.service

#systemctl disable firewalld.service

#systemctl status firewalld.service

/* Cloudera Installation Use below command on master node. */

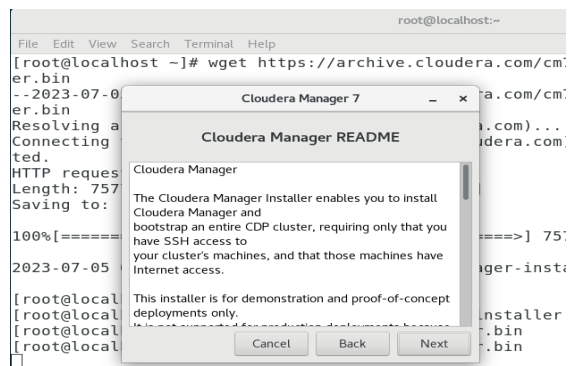
wget https://archive.cloudera.com/cm7/7.4.4/cloudera-manager-installer.bin

chmod u+x cloudera-manager-installer.bin

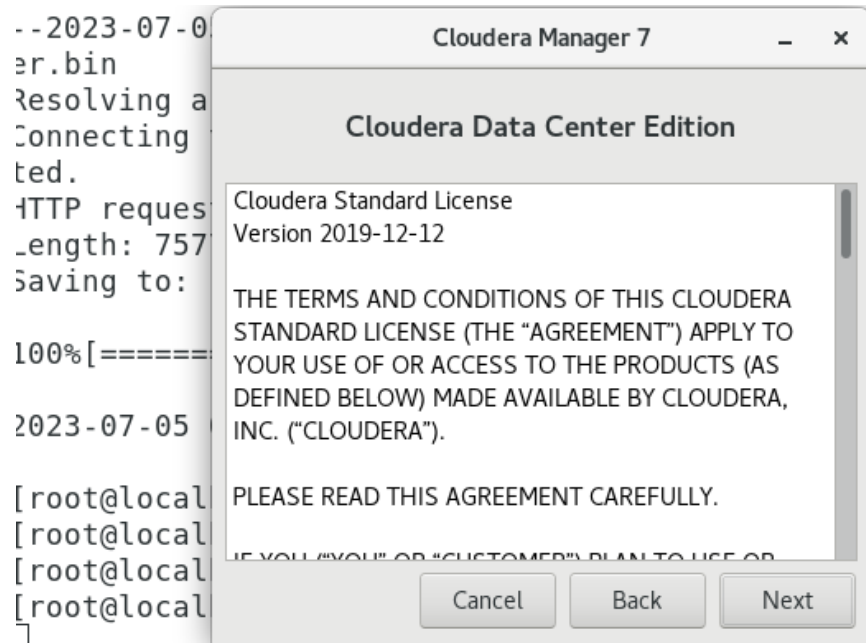
sudo ./cloudera-manager-installer.bin

```
root@localhost:~  
File Edit View Search Terminal Help  
[root@localhost ~]# wget https://archive.cloudera.com/cm7/7.4.4/cloudera-manager-installer.bin  
--2023-07-05 09:06:58-- https://archive.cloudera.com/cm7/7.4.4/cloudera-manager-installer.bin  
Resolving archive.cloudera.com (archive.cloudera.com)... 199.232.252.167  
Connecting to archive.cloudera.com (archive.cloudera.com)|199.232.252.167|:443... connected.  
HTTP request sent, awaiting response... 200 OK  
Length: 757794 (740K) [application/octet-stream]  
Saving to: 'cloudera-manager-installer.bin'  
  
100%[=====>] 757,794 --.-K/s in 0.07s  
2023-07-05 09:06:58 (10.6 MB/s) - 'cloudera-manager-installer.bin' saved [757794/757794]
```

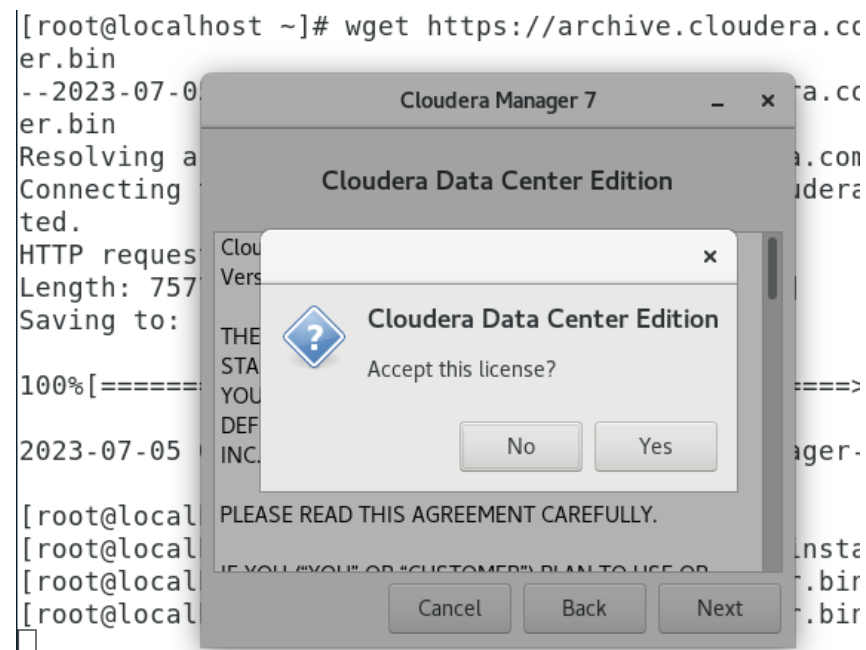
After popup this window Click On Next.



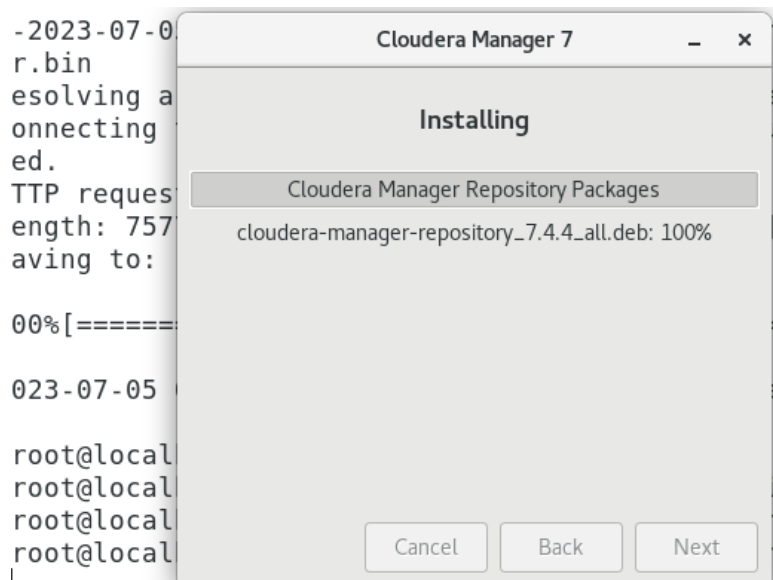
Then Click Next.



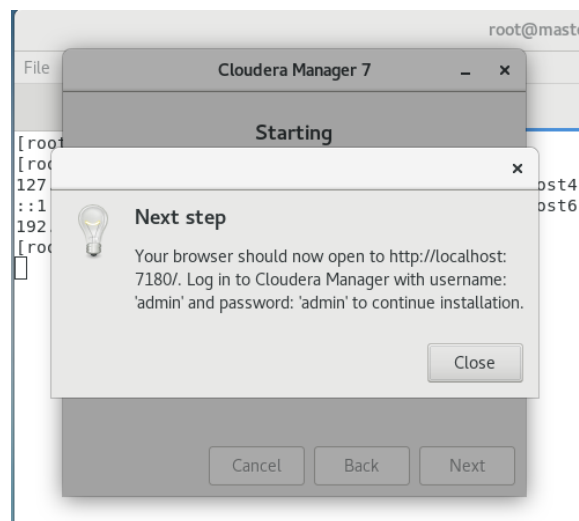
Then Accept Licenses. Click on yes.



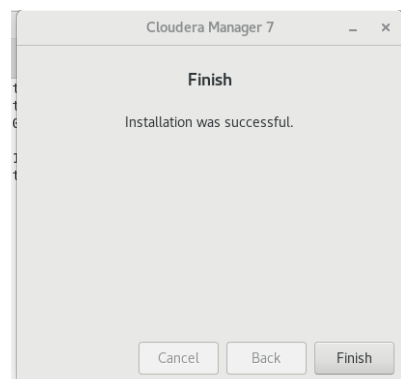
Installation process going on.



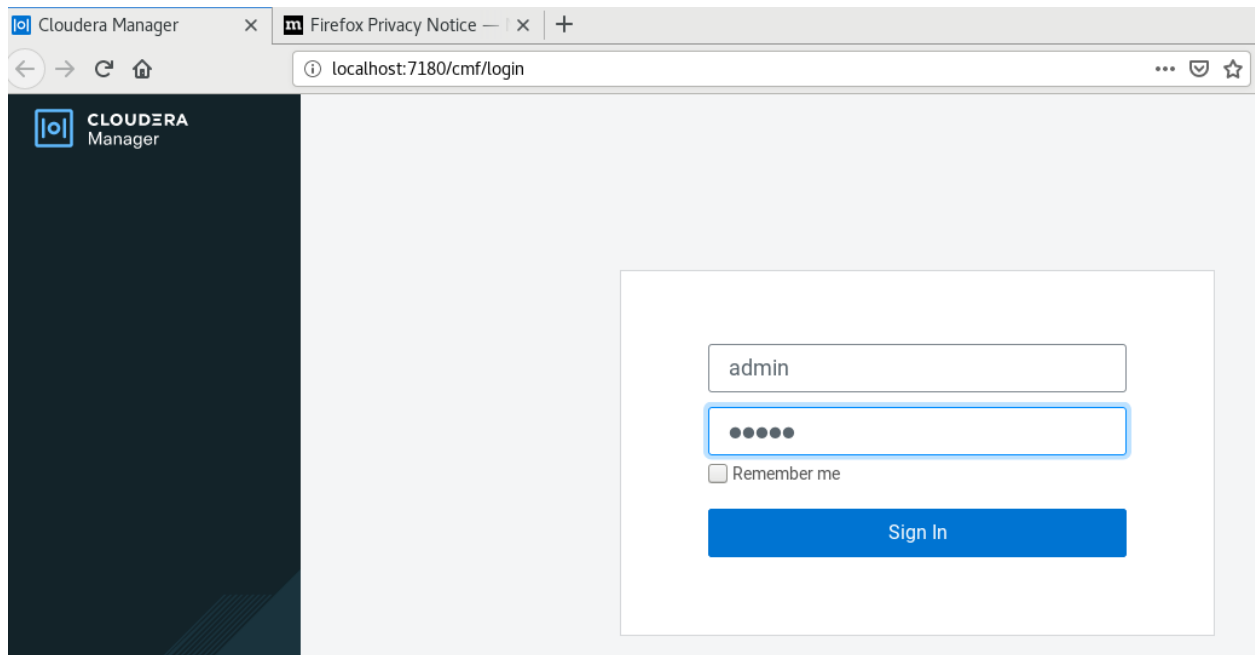
After that click on Close.



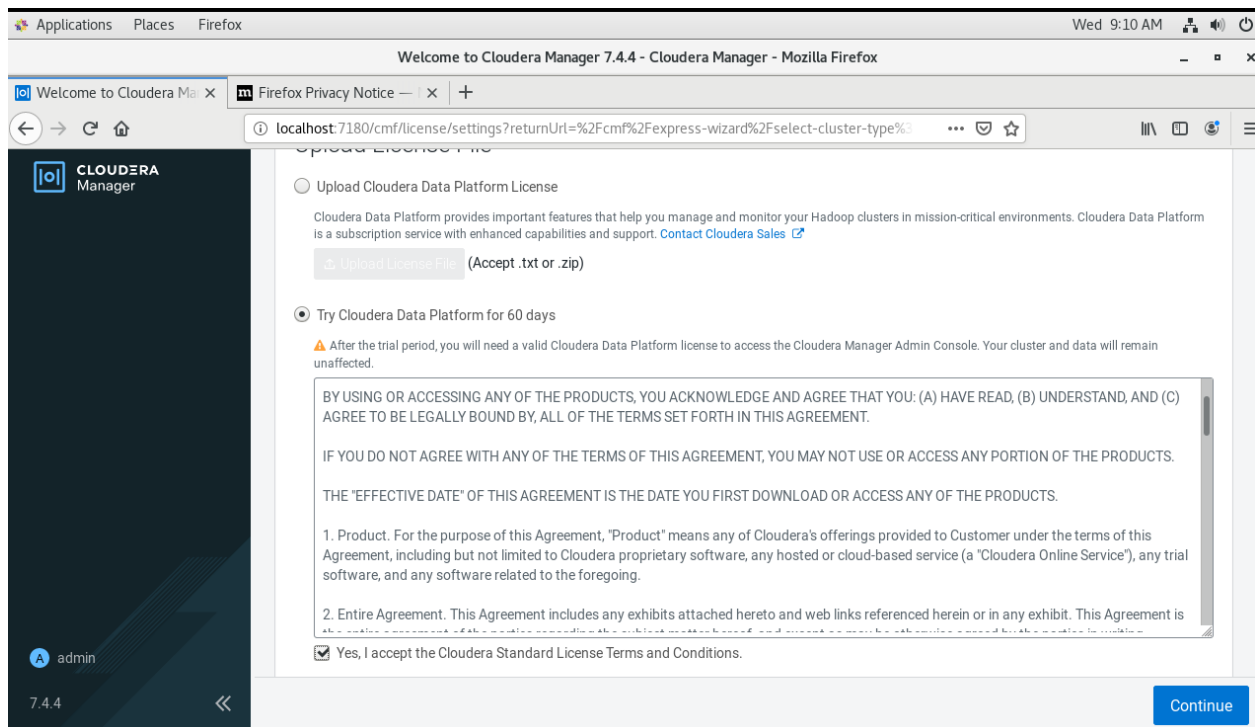
Then click On Finish.



After opening Cludera portal put there user: - admin and password :admin



Then sign in and then select try cludera data and Continue



Then continue and write there cluster name:-HadoopExam

Applications Places Firefox Wed 9:10 AM

Add Traditional Bare Metal Cluster - Cloudera Manager - Mozilla Firefox

localhost:7180/cmf/express-wizard/wizard?allowResume=false&clusterType=BASE_CLUSTER

Add Traditional Bare Metal Cluster

1 Cluster Basics


2 Specify Hosts

3 Select Repository

4 Install Parcels

5 Inspect Cluster

Cluster Name



Base Cluster

A Base Cluster contains storage nodes, compute nodes, and other services such as metadata and security collocated in a single cluster.

Parcels

Running Commands

Support

admin

7.4.4

Cancel

Back Continue

Then continue and put there machine name and click search, select master and continue.

Add Traditional Bare Metal Cluster

Cluster Basics

2 Specify Hosts

3 Select Repository

4 Select JDK

5 Enter Login Credentials

6 Install Agents

7 Install Parcels

8 Inspect Cluster

Specify Hosts

Hosts should be specified using the same hostname (FQDN) that they will identify themselves with.

Hostname

Hint: Search for hostnames or IP addresses using [pattern](#)

SSH Port Search

1 hosts scanned, 1 running SSH.

<input checked="" type="checkbox"/>	Expanded Query	Hostname (FQDN) ↑	IP Address	Currently Managed	Result
<input checked="" type="checkbox"/>	master	master	192.168.82.102	No	Host was successfully scanned.

1 - 1 of 1

Cancel

Back Continue

Again continue.

Add Traditional Bare Metal Cluster

✓ Cluster Basics

✓ Specify Hosts

3 Select Repository

4 Select JDK

5 Enter Login Credentials

6 Install Agents

7 Install Parcels

8 Inspect Cluster

Select Repository

Cloudera Manager Agent

Cloudera Manager Agent **7.4.4** (#15850731) needs to be installed on all new hosts.

Repository Location ☒ Cloudera Repository (Requires direct Internet access on all hosts.) ☐ Custom Repository

Example: `http://LOCAL_SERVER/cloudera-repos/cm7/7.4.4`

Do not include operating system-specific paths in the URL. The path will be automatically derived.

Learn more at [How to set up a custom repository.](#)

Other Software

Cloudera recommends the use of parcels for installation over packages, because parcels enable Cloudera Manager to easily manage the software on your cluster, automating the deployment and upgrade of service binaries. Electing not

[Cancel](#) [← Back](#) [Continue →](#)

Then select install a Cloudera and continue.

✓ Specify Hosts

✓ Select Repository

4 Select JDK

5 Enter Login Credentials

6 Install Agents

7 Install Parcels

8 Inspect Cluster

Select JDK

Selected Version	Cloudera Runtime 7.1
Supported JDK Version	OpenJDK 8, 11 or Oracle JDK 8, 11

[More details on supported JDK version.](#)

If you plan to use JDK 11, you will need to install it manually on all hosts and then select the **Manually manage JDK** option below.

☐ Manually manage JDK

Please ensure that a supported JDK is **already installed on all hosts. You will need to manage installing the unlimited strength JCE policy file, if necessary.**

☒ **Install a Cloudera-provided version of OpenJDK**

By proceeding, Cloudera will install a supported version of OpenJDK version 8.

☐ Install a system-provided version of OpenJDK

By proceeding, Cloudera will install the default version of OpenJDK version 8 provided by the Operating System.

[Cancel](#) [Select install a cloudera, then Continue](#)

[← Back](#) [Continue →](#)

zelRunningCommandsPopup

Put password and continue.

✓ Cluster Basics

✓ Specify Hosts

✓ Select Repository

✓ Select JDK

5 Enter Login Credentials


6 Install Agents

7 Install Parcels

8 Inspect Cluster

Enter Login Credentials

Root access to your hosts is required to install the Cloudera packages. This installer will connect to your hosts via SSH and log in either directly as root or as another user with password-less sudo/pbrun privileges to become root.

SSH Username  ☒ root ☐ Another user

Authentication Method ☒ All hosts accept same password ☐ All hosts accept same private key

Password

SSH Port

Simultaneous Installations
(Running a large number of installations at once can consume large amounts of network bandwidth and other system resources)

[Cancel](#)[< Back](#)[Continue >](#)

Then install agents.

✓ Cluster Basics

✓ Specify Hosts

✓ Select Repository

✓ Select JDK

✓ Enter Login Credentials

6 Install Agents

7 Install Parcels

8 Inspect Cluster

Add Traditional Bare Metal Cluster

Install Agents

Installation completed successfully.

1 of 1 host(s) completed successfully.

Hostname	IP Address	Progress	Status
master	192.168.82.102		✓ Installation completed successfully. Details

1 - 1 of 1

[Cancel](#)[< Back](#)[Continue >](#)

Then click continue

Add Traditional Bare Metal Cluster

✓ Cluster Basics

✓ Specify Hosts

✓ Select Repository

✓ Select JDK

✓ Enter Login Credentials

✓ Install Agents

7 Install Parcels

8 Inspect Cluster

Install Parcels

The selected parcels are being downloaded and installed on all the hosts in the cluster.

> Cloudera Runtime 7.1.7

Downloaded: 100%

Distributed: ...

Unpacked: 1/1

Activated: 1/1

Cancel

← Back

Continue →

Then Click on I understand and Finish.

✓ Specify Hosts

✓ Select Repository

✓ Select JDK

✓ Enter Login Credentials

✓ Install Agents

✓ Install Parcels

8 Inspect Cluster

Inspect Cluster

You have created a new empty cluster. Cloudera recommends that you run the following inspections. For accurate measurements, Cloudera recommends that they are performed sequentially.

⌚ Inspect Network Performance

Once the inspection is complete, review the inspector results before proceeding.

> Advanced Options

Inspect Network Performance

⌚ Inspect Hosts

Once the inspection is complete, review the inspector results before proceeding.

Inspect Hosts

☐ Fix the issues and run the inspection tools again.

☐ Quit the wizard and Cloudera Manager will delete the temporarily created cluster.

☒ I understand the risks of not running the inspections or the detected issues, let me continue with cluster setup.

Cancel

← Back

Finish →

Then select Custom services. And select services.

5

Review Changes

6

Command Details

7

Summary

Data Mart

Browse, query, and explore your data in an interactive way.
Services: HDFS, Ranger, Atlas, Hive, Hive on Tez, Impala, and Hue

Operational Database

Real-time insights for modern data-driven business.
Services: HDFS, Ranger, Atlas, and HBase

Custom Services

Choose your own services. Services required by chosen services will automatically be included.

Service Type	Description
<div><input type="checkbox"/></div> Atlas	Apache Atlas provides a set of metadata management and governance services that enable you to find, organize, and manage data assets. This service requires

70/express-add-services/index?proceed=true#

← Back

Continue →

Select services.

optimize and eliminate your data with this powerful interface.

☐

HBase

Apache HBase is a highly scalable, highly resilient NoSQL OLTP database that enables applications to leverage big data.

☒

HDFS

Apache Hadoop Distributed File System (HDFS) is the primary storage system used by Hadoop applications. HDFS creates multiple replicas of data blocks and distributes them on compute hosts throughout a cluster to enable reliable, extremely rapid computations.

☒

Hive

Apache Hive is a SQL based data warehouse system. In CDH 6 and earlier, this service includes Hive Metastore and HiveServer2. In Cloudera Runtime 7.0 and later, this service only includes the Hive Metastore; HiveServer2 and other components of the Hive execution engines are part of the Hive on Tez service.

☒

Hive on Tez

Hive on Tez is a SQL query engine using Apache Tez.

☐

Hue

Hue is the leading SQL Workbench for optimized, interactive query design and data exploration.

☐

Impala







Apache Impala provides a real-time SQL query interface for data stored in HDFS

← Back

Continue →

Then continue.

Optimize and administer your data with this powerful interface.

<input type="checkbox"/>	 HBase	Apache HBase is a highly scalable, highly resilient NoSQL OLTP database that enables applications to leverage big data.
<input checked="" type="checkbox"/>	 HDFS	Apache Hadoop Distributed File System (HDFS) is the primary storage system used by Hadoop applications. HDFS creates multiple replicas of data blocks and distributes them on compute hosts throughout a cluster to enable reliable, extremely rapid computations.
<input checked="" type="checkbox"/>	 Hive	Apache Hive is a SQL based data warehouse system. In CDH 6 and earlier, this service includes Hive Metastore and HiveServer2. In Cloudera Runtime 7.0 and later, this service only includes the Hive Metastore; HiveServer2 and other components of the Hive execution engines are part of the Hive on Tez service.
<input checked="" type="checkbox"/>	 Hive on Tez	Hive on Tez is a SQL query engine using Apache Tez.
<input type="checkbox"/>	 Hue	Hue is the leading SQL Workbench for optimized, interactive query design and data exploration.
<input type="checkbox"/>	 Impala	Apache Impala provides a real-time SQL query interface for data stored in HDFS

[< Back](#) [Continue →](#)

Select all hosts and continue.

Add Cluster - Configuration

☒ Select Services

☒ **Assign Roles**

☐ Setup Database

☐ Enter Required Parameters

☐ Review Changes


☐ Command Details

☐ Summary

Assign Roles

You can customize the role assignments for your new cluster here, but if assignments are made incorrectly, such as assigning too many roles to a single host, this can impact the performance of your services. Cloudera does not recommend altering assignments unless you have specific requirements, such as having pre-selected a specific host for a specific role.

You can also view the role assignments by host. [View By Host](#)

 HDFS

NameNode × 1 New


SecondaryNameNode × 1 New

Balancer × 1 New

HttpFS × 1 New

NFS Gateway

DataNode × 1 New

 Hive

[< Back](#) [Continue →](#)

Select master hosts and click on Ok.

1 Hosts Selected

Select hosts for a new or existing role. The host list is filtered to remove hosts that are not valid candidates; these include hosts that are unhealthy, members of other clusters, or have an incompatible version of the software installed on them.

Q Enter hostnames: host01, IP addresses or rack

<input checked="" type="checkbox"/>	Hostname	IP Address	Rack	Cores	Physical Memory	Existing Roles	Added Roles
<input checked="" type="checkbox"/>	master	192.168.82.102	/default	4	15.5 GIB		B DN HFS NN NF... SNN G HMS HS2 AP ES HM RM SM YA... G JHS NM S

1 - 1 of 1

Cancel OK

Then continue

4 Enter Required Parameters

5 Review Changes

6 Command Details

7 Summary

for a specific role.

You can also view the role assignments by host. [View By Host](#)

HDFS

NameNode × 1 New

SecondaryNameNode × 1 New

Balancer × 1 New

HttpFS × 1 New

NFS Gateway × 1 New

DataNode × 1 New

Hive

Gateway × 1 New

Metastore Server × 1 New

WebHCat Server × 1 New

HiveServer2

[Back](#) [Continue](#)

After selecting all hosts node do the continue.

4

Enter Required Parameters

5

Review Changes

6

Command Details

7

for a specific role.

You can also view the role assignments by host. [View By Host](#)

HDFS

NameNode × 1 New	SecondaryNameNode × 1 New	Balancer × 1 New
<input type="text" value="master"/>	<input type="text" value="master"/>	<input type="text" value="master"/>
HttpFS × 1 New	NFS Gateway × 1 New	DataNode × 1 New
<input type="text" value="master"/>	<input type="text" value="master"/>	<input type="text" value="All Hosts"/>

Hive

Gateway × 1 New	Metastore Server × 1 New	WebHCat Server × 1 New
<input type="text" value="master"/>	<input type="text" value="master"/>	<input type="text" value="master"/>
HiveServer2 × 1 New		
<input type="text" value="master"/>		

← Back

Continue →

Then test connection and take user id and password.

5

Review Changes

6

Command Details

7

embedded database, passwords are automatically generated. Please copy them down.

Hive

✓ Skipped. Cloudera Manager will create this database in a later step.

Type	Database Hostname	Database Name
<input type="text" value="PostgreSQL"/>	master:7432	hive
Username	Password	
hive	FZm48LeHFc	

Activity Monitor

✓ Successful

Currently assigned to run on **master**.

Type	Database Hostname	Database Name
<input type="text" value="PostgreSQL"/>	master:7432	amon
Username	Password	
amon	QouQpL3g8K	

← Back

Continue →

Currently assigned to run on **master**.

Type	Database Hostname	Database Name
PostgreSQL	master:7432	amon
Username	Password	
amon	QouQpL3g8K	

Reports Manager

Currently assigned to run on **master**.

Type	Database Hostname	Database Name
PostgreSQL	master:7432	rman
Username	Password	
rman	IEfZRYnu1h	

Test Connection

← Back

Continue →

Continue .

Add Cluster - Configuration

✓ Select Services

✓ Assign Roles

✓ Setup Database

✓ Enter Required Parameters

5 Review Changes

6 Command Details

7 Summary

Review Changes

HDFS Block Size

HadoopExam > HDFS (Service-Wide)

dfs.blocksize

dfs_block_size

128

MiB

DataNode Failed Volumes Tolerated

HadoopExam > DataNode Default Group

dfs.datanode.failed.volumes.tolerated

dfs_datanode_failed_volumes_tolerated

0

DataNode Data Directory

HadoopExam > DataNode Default Group

dfs.datanode.data.dir

dfs_data_dir_list

/dfs/dn

NameNode Data Directories

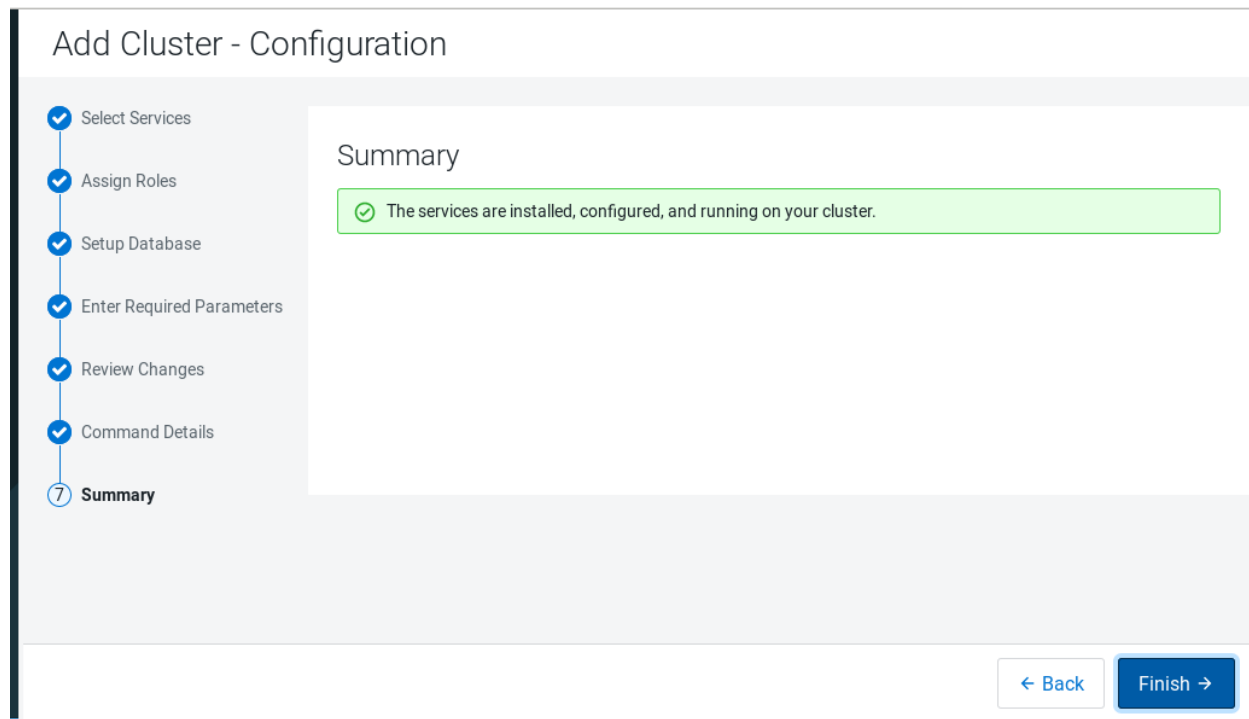
HadoopExam > NameNode Default Group

dfs.namenode.name.dir

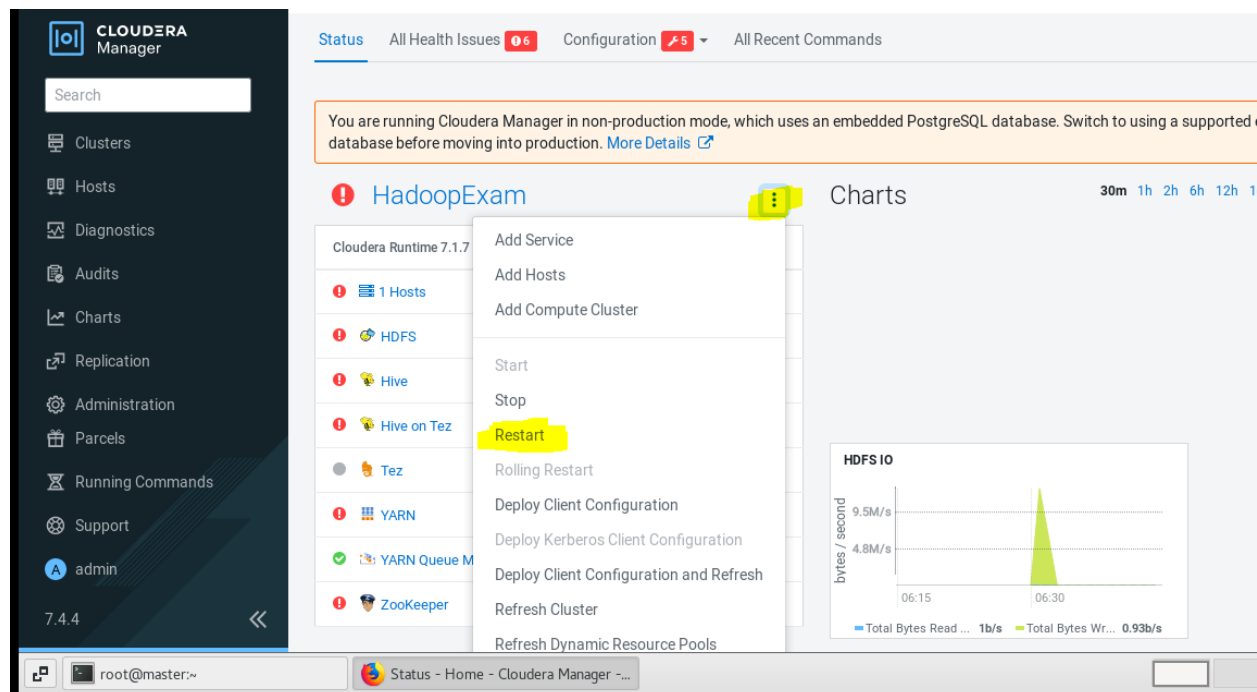
← Back

Continue →

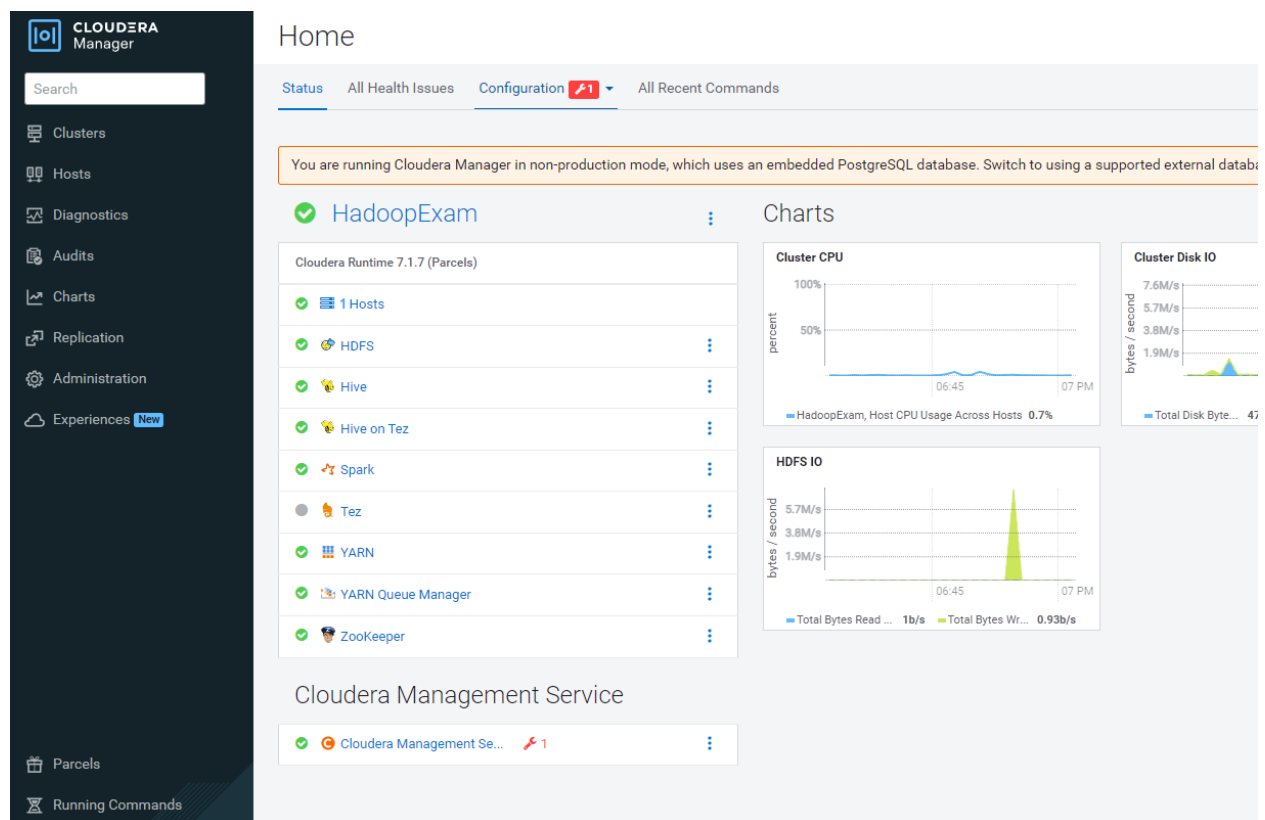
After completed steps click on Finish.



Then click your Cluster name three dot and restart.



After restart your cluster service getting green and working fine.



Now our cluster successfully installed.