

LAB: Installing CDH5 using Cloudera Manager:

Step: 1 Meet the prerequisites: Ensure following are met before installing CDH5.

The hosts in a Cloudera Manager deployment must satisfy the following networking and security requirements:

Cluster hosts must have a working network name resolution system and correctly formatted `/etc/hosts` file.

All cluster hosts must have properly configured forward and reverse host resolution through DNS.

- The `/etc/hosts` files must Contain consistent information about hostnames and IP addresses across all hosts
- Not contain uppercase hostnames
- Not contain duplicate IP addresses.

Step: 2 In our case we have following Ubuntu instances configured in amazon aws as specified in the picture below:(Ask administrator for instance details).

cloudera11	i-1385f23c	m3.large	us-east-1e	running	2/2 checks ...	None
cloudera12	i-6385f24c	m3.large	us-east-1e	running	2/2 checks ...	None
cloudera13	i-6285f24d	m3.large	us-east-1e	running	2/2 checks ...	None

Step:3 Following are the private IP and the edited `/etc/hosts` as specified below, configure the same in your instances using putty or ssh client connecting to the instances as specified below.

Instance 1 configuration:

```
Downloads — ubuntu@ip-172-31-4-103: ~ — ssh — 92x26
ubuntu@ip-172-31-4-103: ~
bash
127.0.0.1 localhost
172.31.4.103 ip-172-31-4-103.ec2.internal
172.31.4.105 ip-172-31-4-105.ec2.internal
172.31.4.104 ip-172-31-4-104.ec2.internal
# The following lines are desirable for IPv6 capable hosts
```

Instance 2 Configuration:

```
ubuntu@ip-172-31-4-103: ~  ubuntu@ip-172-31-4-105: ~  bash
127.0.0.1 localhost
172.31.4.103 ip-172-31-4-103.ec2.internal
172.31.4.105 ip-172-31-4-105.ec2.internal
172.31.4.104 ip-172-31-4-104.ec2.internal
```

Instance 3 Configuration:

```
ubuntu@ip-172-31-4-103: ~  ubuntu@ip-172-31-4-105: ~  ubuntu@ip-172-31-4-104: ~
127.0.0.1 localhost
172.31.4.103 ip-172-31-4-103.ec2.internal
172.31.4.105 ip-172-31-4-105.ec2.internal
172.31.4.104 ip-172-31-4-104.ec2.internal
```

Step:4 Select instance to install CM5 package as specified below and download the utility as specified below:

```

Downloads — ubuntu@ip-172-31-8-130: ~ — ssh — 132x27
ubuntu@ip-172-31-8-130:~$ wget http://archive.cloudera.com/cm5/installer/latest/cloudera-manager-installer.bin
—2015-01-19 13:56:55— http://archive.cloudera.com/cm5/installer/latest/cloudera-manager-installer.bin
Resolving archive.cloudera.com (archive.cloudera.com)... 54.230.19.29, 54.240.160.223, 54.230.16.118, ...
Connecting to archive.cloudera.com (archive.cloudera.com)[54.230.19.29]:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 514295 (502K) [application/octet-stream]
Saving to: 'cloudera-manager-installer.bin'

100%[=====] 514,295  --.-K/s  in 0.02

2015-01-19 13:56:55 (26.5 MB/s) - 'cloudera-manager-installer.bin' saved [514295/514295]

```

Step:5 Once your wget runs properly do ls to get the installation binary.

```

ubuntu@ip-172-31-4-103:~$ ls
cloudera-manager-installer.bin
ubuntu@ip-172-31-4-103:~$

```

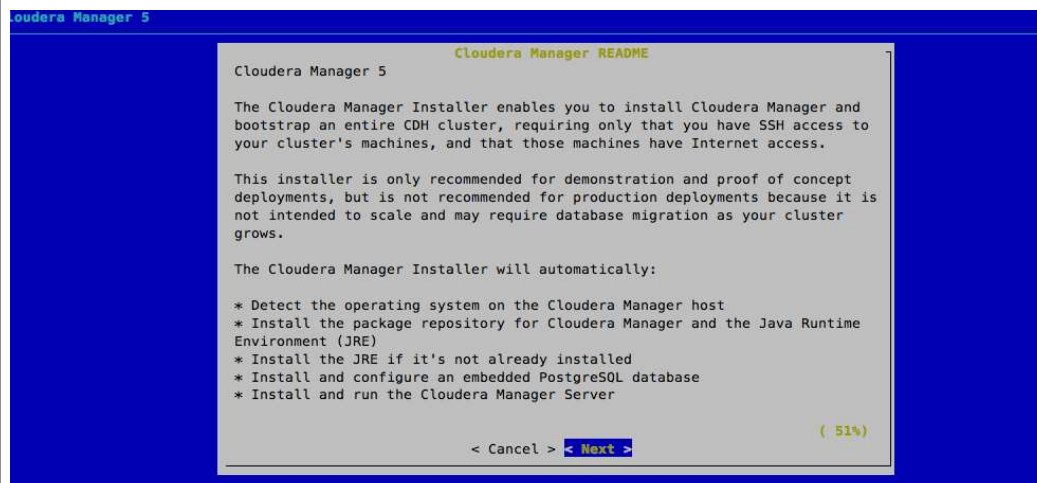
Step:6 Run the installer(ensure as root user)

```

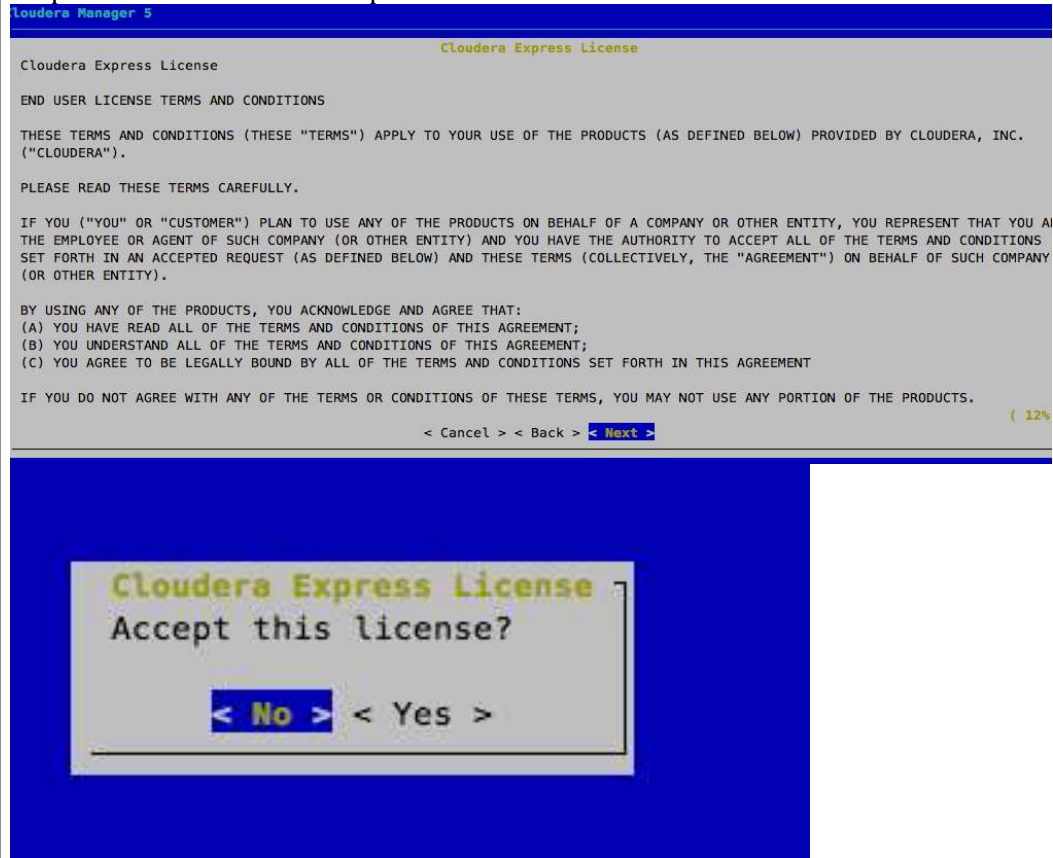
ubuntu@ip-172-31-4-103:~$ sudo ./cloudera-manager-installer.bin

```

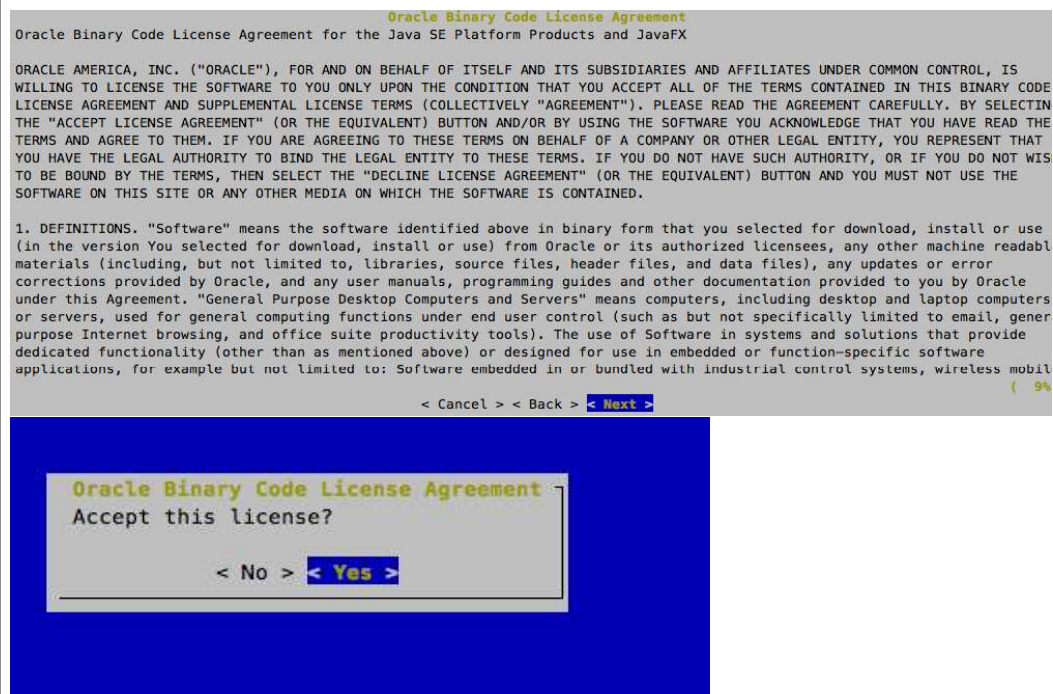
Step:7 From here specify follow the wizard.Follow the wizard as specified in below diagrams:



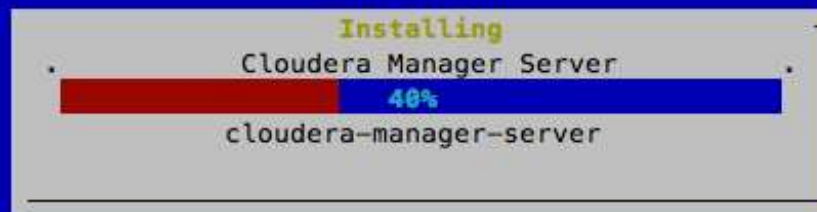
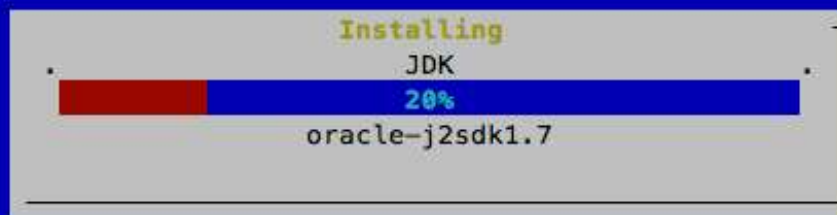
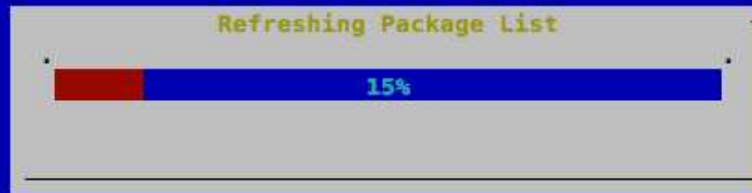
Step:8 Select Next and Accept the license.



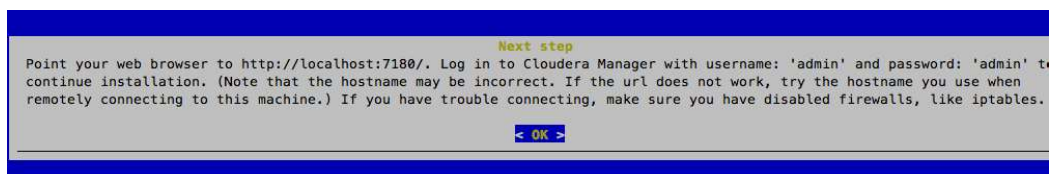
Step: 9 Accept non-cloudera licenses.



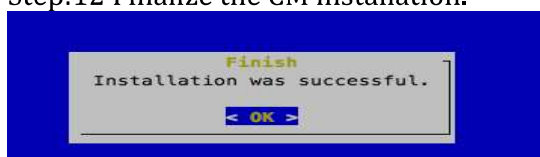
Step: 10 It starts installation with following tasks .



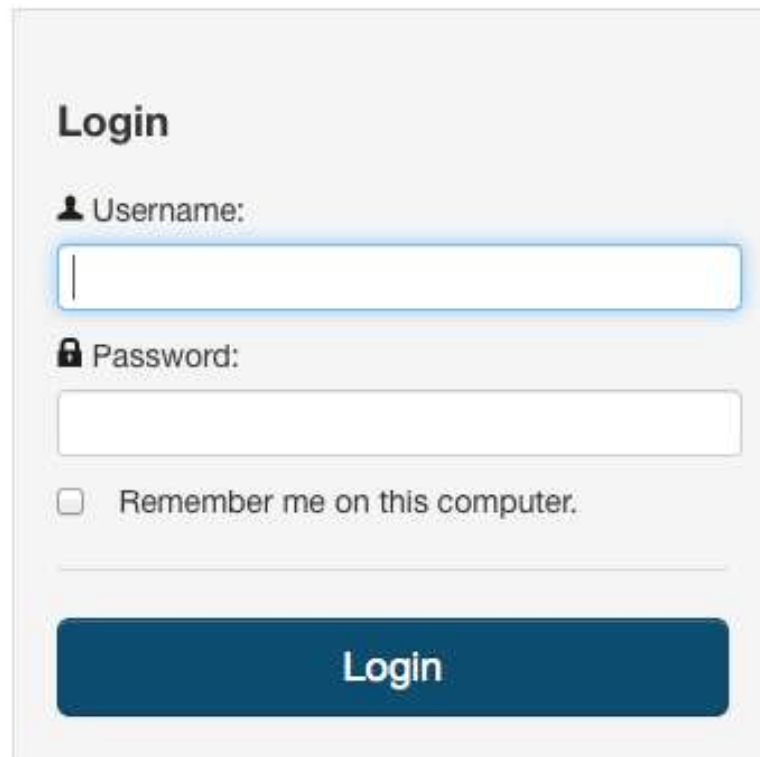
Step:11 Once installation is done browse using 7180 tcp port as indicated in last steps.



Step:12 Finalize the CM installation.



Step:13 Login in the CM using admin as userid and admin as password.



Login

Username:

Password:

☐ Remember me on this computer.

Login

Step:14 Select the right distribution for installation .We would be selecting express since no licenses are required

Welcome to Cloudera Manager. Which edition do you want to deploy?

Upgrading to Cloudera Enterprise Data Hub Edition provides important features that help you manage and monitor your Hadoop clusters in mission-critical environments.

	Cloudera Express	Cloudera Enterprise Data Hub Edition Trial	Cloudera Enterprise
License	Free	60 Days After the trial period, the product will continue to function as Cloudera Express . Your cluster and your data will remain unaffected.	Annual Subscription Upload License Cloudera Enterprise is available in three editions: • Basic Edition • Flex Edition • Data Hub Edition
Node Limit	Unlimited	Unlimited	Unlimited
CDH	✓	✓	✓
Core Cloudera Manager Features	✓	✓	✓
Advanced Cloudera Manager Features	✓	✓	✓

[Continue](#)

Compare the feature and select the right one as per your production need.

Step:15 Click on continue to reach to the products overview page as specified below.

Thank you for choosing Cloudera Manager and CDH.

This installer will install **Cloudera Enterprise Data Hub Edition Trial 5.3.0** and enable you to later choose packages for the services below (there may be some license implications).

- Apache Hadoop (Common, HDFS, MapReduce, YARN)
- Apache HBase
- Apache ZooKeeper
- Apache Oozie
- Apache Hive
- Hue (Apache licensed)
- Apache Flume
- Cloudera Impala (Apache licensed)
- Apache Sentry
- Apache Sqoop
- Cloudera Search (Apache licensed)
- Apache Spark

You are using Cloudera Manager to install and configure your system. You can learn more about Cloudera Manager by clicking on the **Support** menu above.

Continue

Step:16 Select the host as specified below ensure the ip addresses selected are correct as given/specified by administrator.

Specify hosts for your CDH cluster installation.

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

Cloudera recommends including Cloudera Manager Server's host. This will also enable health monitoring for that host.

Hint: Search for hostnames and/or IP addresses using [patterns](#).

3 hosts scanned, 3 running SSH.

New Search

<input checked="" type="checkbox"/>	Expanded Query	Hostname (FQDN)	IP Address	Currently Managed	Result
<input checked="" type="checkbox"/>	172.31.4.103	ip-172-31-4-103.ec2.internal	172.31.4.103	No	✓ Host ready: 1 ms response time.
<input checked="" type="checkbox"/>	172.31.4.104	ip-172-31-4-104.ec2.internal	172.31.4.104	No	✓ Host ready: 1 ms response time.
<input checked="" type="checkbox"/>	172.31.4.105	ip-172-31-4-105.ec2.internal	172.31.4.105	No	✓ Host ready: 8 ms response time.

Back

Continue

Step:17 Mark the progress of cluster installation

Cluster Installation

Installation in progress.

0 of 3 host(s) completed successfully.

Abort Installation

Hostname	IP Address	Progress	Status	
ip-172-31-4-103.ec2.internal	172.31.4.103	<div></div>	Refreshing package metadata...	Details
ip-172-31-4-104.ec2.internal	172.31.4.104	<div></div>	Refreshing package metadata...	Details
ip-172-31-4-105.ec2.internal	172.31.4.105	<div></div>	Refreshing package metadata...	Details

Step:18 Once everything goes fine you will see the below screen marking successful installation of cluster packages .

Cluster Installation

Installation completed successfully.

3 of 3 host(s) completed successfully.

Hostname	IP Address	Progress	Status	
ip-172-31-4-103.ec2.internal	172.31.4.103	<div></div>	✓ Installation completed successfully.	Details
ip-172-31-4-104.ec2.internal	172.31.4.104	<div></div>	✓ Installation completed successfully.	Details
ip-172-31-4-105.ec2.internal	172.31.4.105	<div></div>	✓ Installation completed successfully.	Details

⏪ Back

1 2 3 4 5 6 7 8

⏩ Continue

Step: 19 Click on continue to mark the CDH5 parcel as specified below.

Cluster Installation

Installing Selected Parcels

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

CDH 5.3.0-1.cdh5.3.0.p0.30

<div></div>	Downloaded
<div></div>	Distributed
<div></div>	Activated

⏪ Back

1 2 3 4 5 6 7 8

⏩ Continue

Step:20 select the pack you want to install as specified below.

Cluster Setup

Choose the CDH 5 services that you want to install on your cluster.

Choose a combination of services to install.

- ☒ **Core Hadoop**
HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, and Sqoop
- ☐ **Core with HBase**
HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, Sqoop, and HBase
- ☐ **Core with Impala**
HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, Sqoop, and Impala
- ☐ **Core with Search**
HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, Sqoop, and Solr
- ☐ **Core with Spark**
HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, Sqoop, and Spark
- ☐ **All Services**
HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, Sqoop, HBase, Impala, Solr, Spark, and Key-Value Store Indexer
- ☐ **Custom Services**
Choose your own services. Services required by chosen services will automatically be included. Fume can be added after your initial cluster has been set up.

⏪ Back

1 2 3 4 5 6

⏩ Continue

Step:21 Configure the required databases (don't worry its automatically done)

Configure and test database connections. If using custom databases, create the databases first according to the [Installing and Configuring an External Database](#) section of the [Installation Guide](#).

☐ Use Custom Databases
☒ Use Embedded Database

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

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

Database Host Name:
 Database Type:
 Database Name:
 Username:
 Password:

Activity Monitor ✓ Successful

Currently assigned to run on ip-172-31-4-105.ec2.internal.

Database Host Name:
 Database Type:
 Database Name:
 Username:
 Password:

[Test Connection](#)

[Back](#)
1 2 3 4 5 6
[Continue](#)

Step:22 In the following steps CM configures and starts the services as specified below. Its a 22 step task keep patience till it succeeds.

Command Progress

Completed 2 of 22 steps.

☒ Initializing ZooKeeper Service
 Completed 1 steps successfully.

☒ Starting ZooKeeper Service
 Completed 1 steps successfully.
[Details](#)

☐ Checking if the name directories of the NameNode are empty. Formatting HDFS only if empty.
[Details](#)

Starting HDFS Service

Creating HDFS /tmp directory

Creating MR2 job history directory

Creating NodeManager remote application log directory

Starting YARN (MR2 Included) Service

Creating Hive Metastore Database

Creating Hive Metastore Database Tables

Step:23 once all 22 services are installed and done you get the below message as specified.

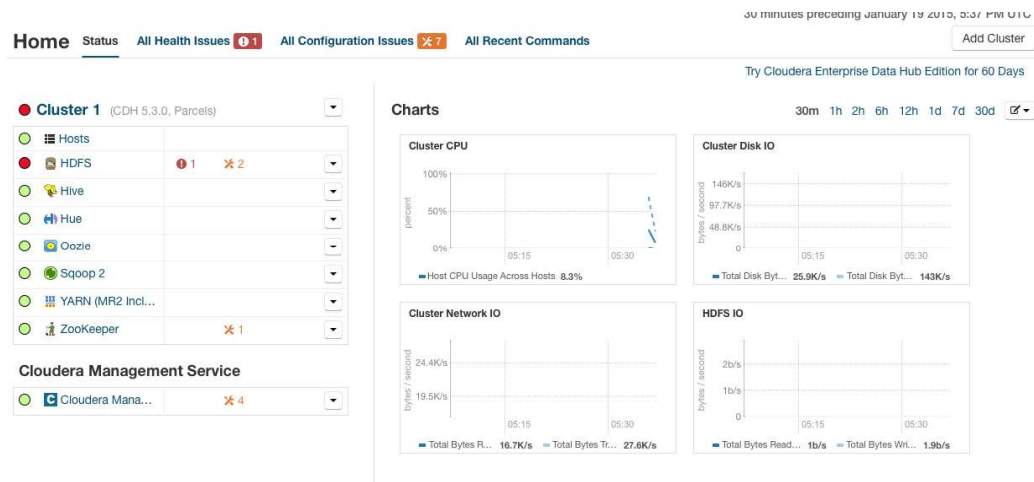
Cloudera manager Support [adm](#)

Cluster Setup

Congratulations!

The services are installed, configured, and running on your cluster.

Step:24 once everything is done click on finish and you get a Dashboard for provisioning and configuring the hadoop services as specified below.



Step:25 you may find some warnings ignore for now.Your cloudera CDH5(Hadoop is up and running).