**Cloudera Installation Procedure**

This document shows how to install Cloudera distribution(CDH 5.0) on a multi node cluster hosted on India.

# **Cloudera Manager Installation**

Cloudera distribution is installed using Cloudera Manager(CM) on one node. CM can be used to install other Hadoop ecosystems like Hadoop, MapReduce, HBase, HIVE and others using web based application.

Prerequisites for Installing CM.

1. All nodes in the cluster should be able to do password less SSH among each other.
2. One node should be used for only Cloudera Manager and other for Namenode and datanodes likewise.
3. One VM hosted on India for installing Cloudera Manager
   1. **Cloudera Manager Requirements**

Cloudera Manager Requirements. All these requirements should be satisfied for the VMs hosted on India

<http://www.cloudera.com/content/cloudera/en/documentation/cloudera-manager/v5-1-x/Cloudera-Manager-Installation-Guide/cm5ig_cm_requirements.html>

* 1. **MySQL database Installation**

The below commands mentioned in the URL should be executed only using ROOT User id

Cloudera manager needs a database for all its services to run. We can install MySQL on one of the nodes where we are installing Cloudera Manager. Only the respective Ubuntu commands should be executed if we are working with VMs hosted on India

<http://www.cloudera.com/content/cloudera/en/documentation/cloudera-manager/v5-1-x/Cloudera-Manager-Installation-Guide/cm5ig_mysql.html>

* 1. **Installation**

Cloudera manager on a node can be installed using the below link

<http://www.cloudera.com/content/cloudera/en/documentation/cloudera-manager/v5-1-x/Cloudera-Manager-Installation-Guide/cm5ig_install_path_B.html?scroll=cmig_topic_6_6>

Only the below commands from the above URL should be executed for Cloudera Manager,

1. Adding the repository in the machine
2. cloudera.list file should be created with the below text and copied to /etc/apt/sources.list.d/

# Packages for Cloudera's Distribution for Hadoop, Version 4, on Ubuntu 12.04 x86\_64

deb [arch=amd64] http://archive.cloudera.com/cm5/ubuntu/precise/amd64/cm precise-cm5 contrib

deb-src http://archive.cloudera.com/cm5/ubuntu/precise/amd64/cm precise-cm5 contrib

1. Install Oracle Java

sudo apt-get install oracle-j2sdk1.7

1. Installing cloudera manager server packages

sudo apt-get install cloudera-manager-daemons cloudera-manager-server

1. Installing Cloudera manager agent packages

sudo apt-get install cloudera-manager-agent cloudera-manager-daemons

1. Start the cloudera manager server

sudo service cloudera-scm-server start

1. Hosts file should be updated with the datanodes host details. /etc/hosts

Like below,

10.39.1.50 vmadhire-0.novalocal

10.39.1.51 vmadhire-1.novalocal

1. Cloudera Manager is ready for use and it can be accessed using the below URL.

<http://host:7182>

Host is the server address where cloudera manager is hosted.

1. We can select cloudera manager edition and the hosts on which we need to install all the Hadoop ecosystems

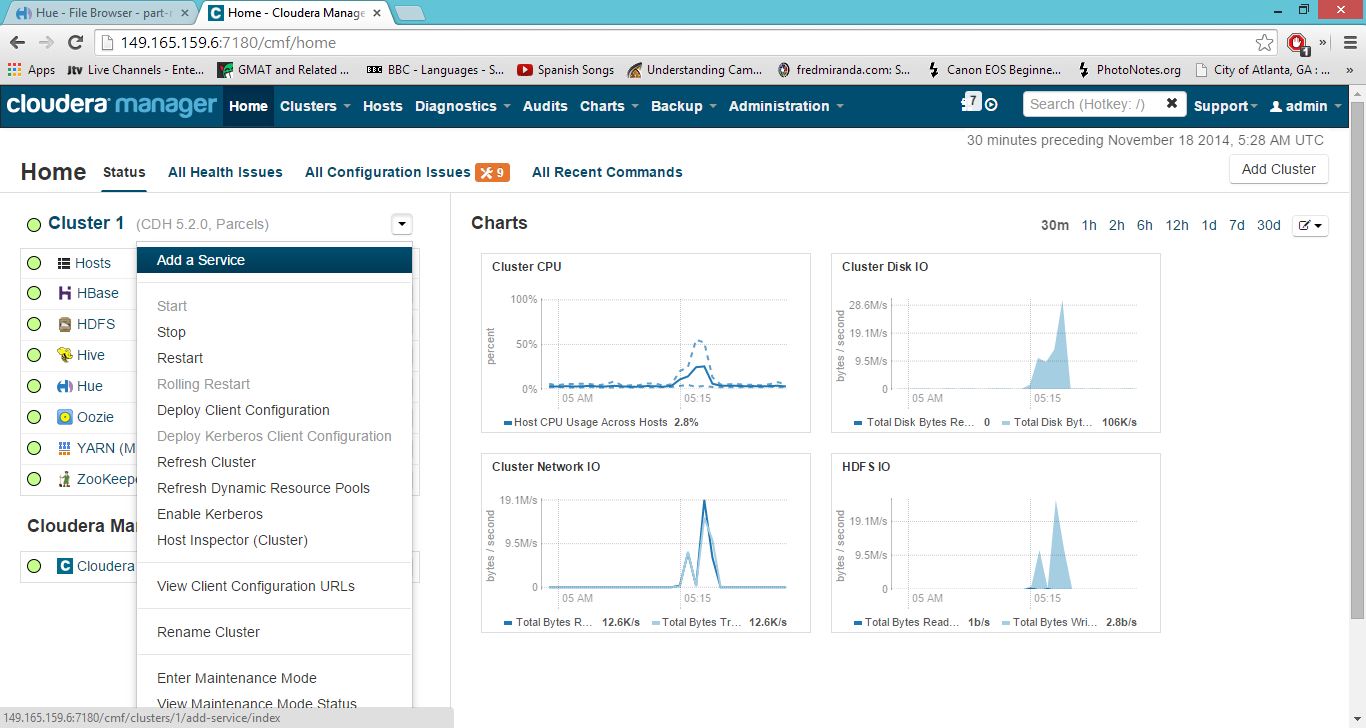
<http://www.cloudera.com/content/cloudera/en/documentation/cloudera-manager/v5-1-x/Cloudera-Manager-Installation-Guide/cm5ig_install_path_B.html#concept_qzf_gyt_fm_unique_1>

We can install Cloudera Enterprise Data hub edition trial version which we can use for 60 days. After the trial period we can use the same cluster but with few limited services.

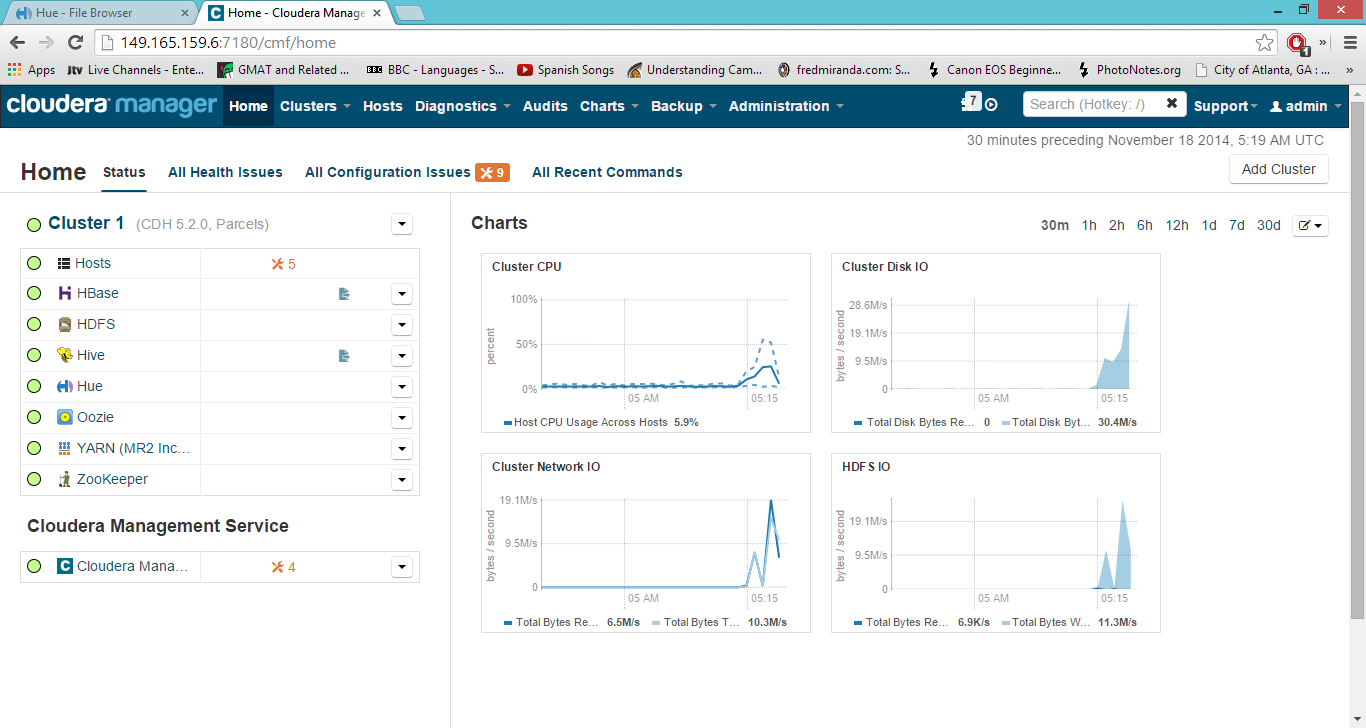
After the installation of Cloudera manager. We can use Add service option to install Hadoop, MapReduce, Hive, HBase with just a click of button and also using the below link.

http://www.cloudera.com/content/cloudera/en/documentation/cloudera-manager/v5-1-x/Cloudera-Manager-Installation-Guide/cm5ig\_install\_path\_B.html#cmig\_topic\_6\_6\_10\_unique\_1

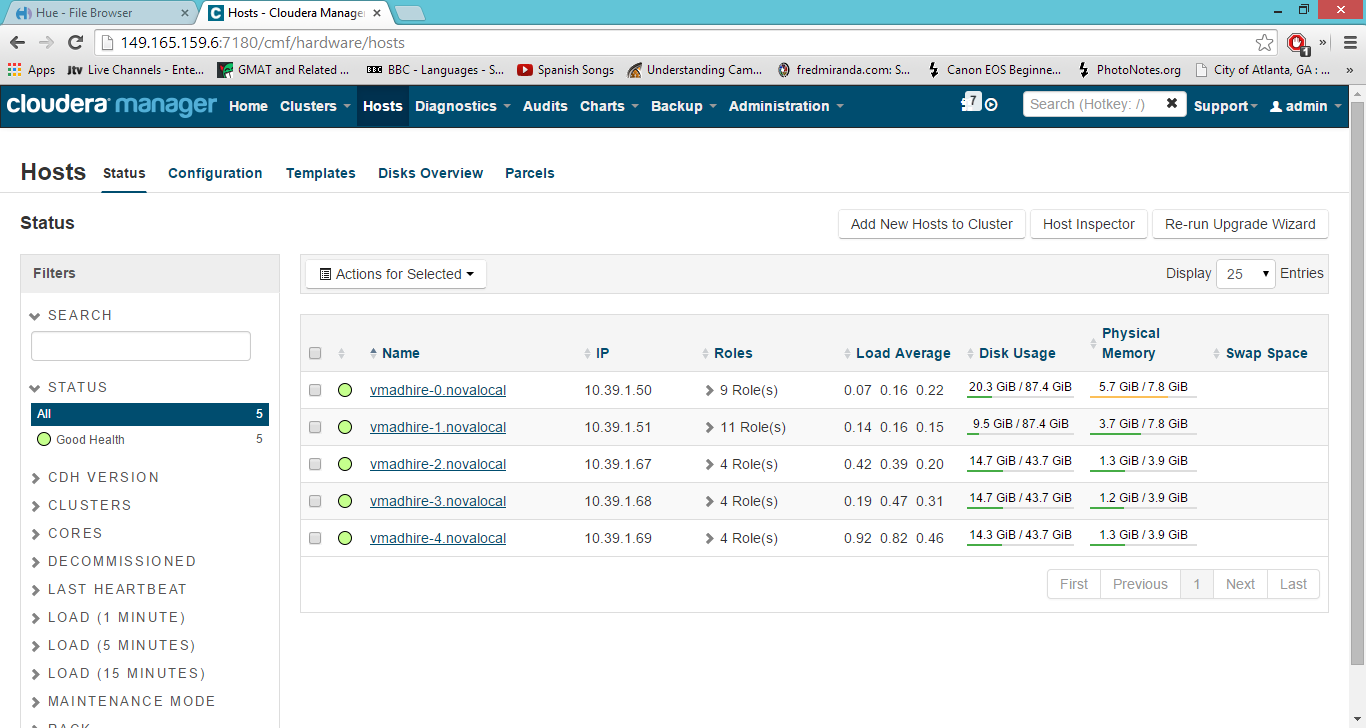
Add Service option is shown below,



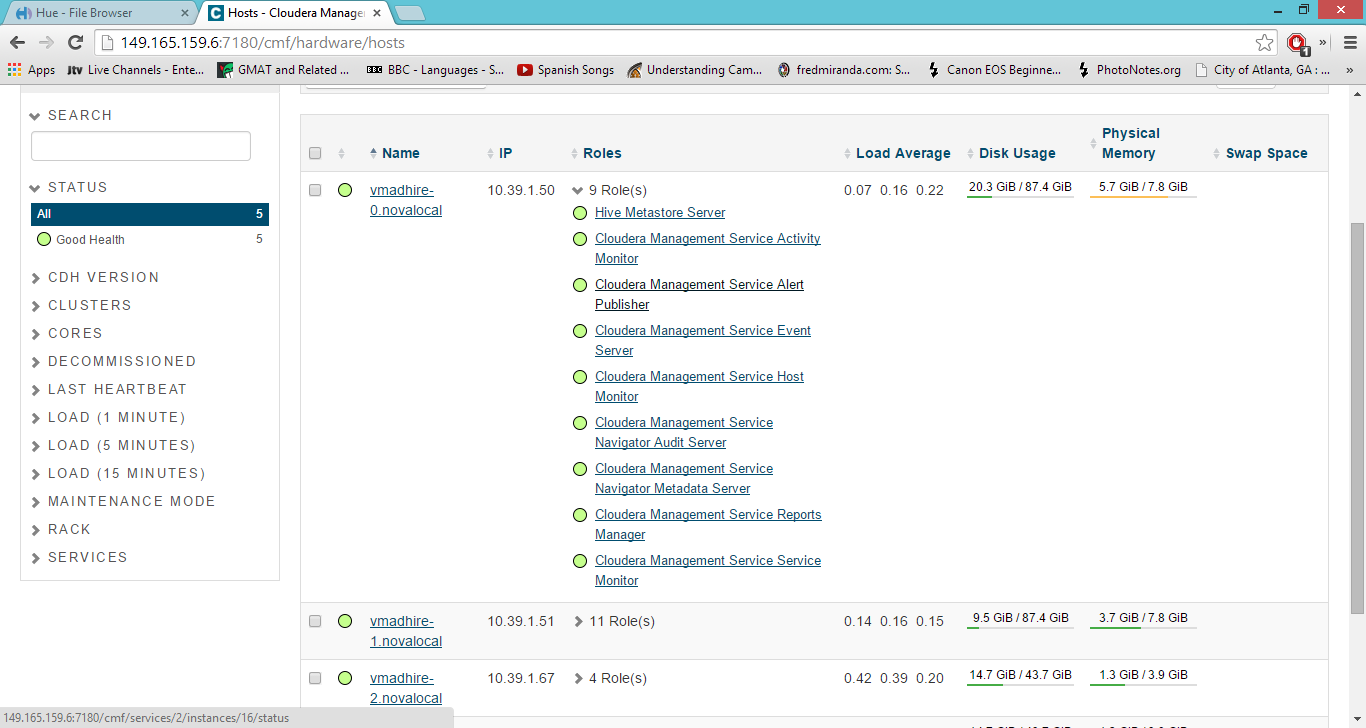
After the installation Cloudera Manager looks like this,

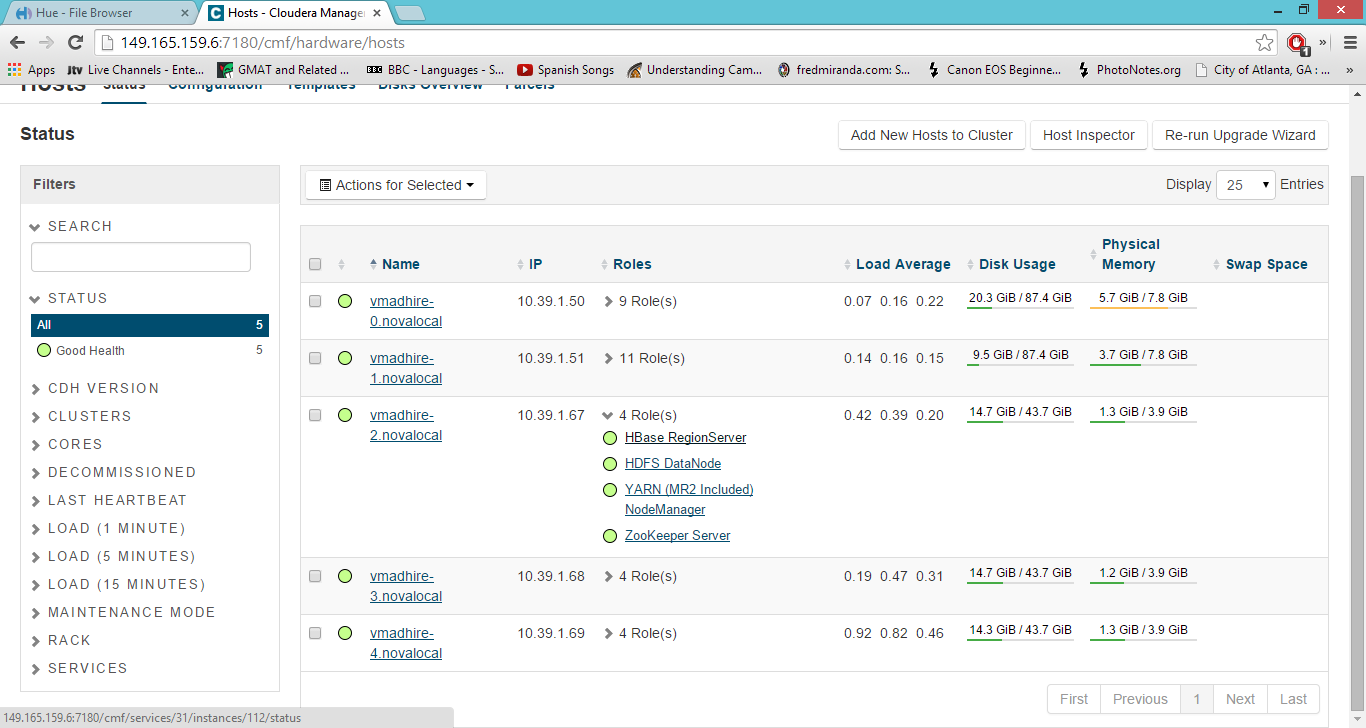


Hosts of a cluster can be as seen below,



Services on each node can be seen as below,





Hue (Web Based application) for managing jobs and viewing HDFS directory structure

After installing Hue, it can be accessed using <http://host:8888>

Example

As an example Cloudera (CDH 5.2) Hadoop has been installed on 5 node cluster running on India.

One Node used for Cloudera Manager

One Node used for NameNode

3 Nodes used for Datanodes

Cloudera Manager can be accessed by using the URL mentioned in the table,

A user has been created to access the Cloudera(CDH 5.2) cluster and to check about different settings present in Cloudera.

|  |  |  |  |
| --- | --- | --- | --- |
| **Task** | **Link** | **User** | **Password** |
| Cloudera Manager | <http://149.165.159.6:7180/cmf/home> | reviewuser | Same as User |
| Hue | [http://149.165.159.41:8888](http://149.165.159.41:8888/) | reviewuser | Same as User |