# CM以及CDH安装文档

## 一.关于CDH和Cloudera Manager(CM)

CDH (Cloudera's Distribution, including Apache Hadoop)，是Hadoop众多分支中的一种，由Cloudera维护，基于稳定版本的Apache Hadoop构建，并集成了很多补丁，可直接用于生产环境。

Cloudera Manager则是为了便于在集群中进行Hadoop等大数据处理相关的服务安装和监控管理的组件，对集群中主机、Hadoop、Hive、Spark等服务的安装配置管理做了极大简化。

## 二.系统环境

* 操作系统: CentOS 6.6
* Cloudera Manager: 5.6.0
* CDH: 5.6.0
* Jdk: 1.7.0\_79

## 系统环境搭建

### 1.相关包下载

CM下载地址：

http://archive.cloudera.com/cm5/cm/5/[cloudera-manager-el6-cm5.7.0\_x86\_64.tar.gz](http://archive.cloudera.com/cm5/cm/5/cloudera-manager-el6-cm5.6.0_x86_64.tar.gz)

CDH安装包地址:<http://archive.cloudera.com/cdh5/parcels/5.6.0/>

由于我们的操作系统为CentOS6.6，需要下载以下文件：

* CDH-5.6.0-1.cdh5.7.0.p0.45-el6.parcel
* CDH-5.6.0-1.cdh5.7.0.p0.45-el6.parcel.sha1
* manifest.json

### 2. 网络配置(所有节点)

vi /etc/sysconfig/network 修改hostname:



通过service network restart重启网络服务生效。

vi /etc/hosts,修改ip与主机名的对应关系



**注意：**本文档以3台主机为样本,实际安装时需注意将所有的主机ip和对应的主机名写进去,本机的也要写进去，否则启动Agent的时候会提示hostname解析错误。

### 打通SSH,配置ssh免密码登录

在主节点上执行ssh-keygen -t rsa一路回车，生成无密码的密钥对。

将公钥添加到认证文件中：cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys，并设置authorized\_keys的访问权限：chmod 600 ~/.ssh/authorized\_keys。

scp文件到所有datenode节点：

scp ~/.ssh/authorized\_keys root@n2:~/.ssh/

**测试：**在主节点上ssh n2，正常情况下，不需要密码就能直接登陆进去了。

### 4.安装Oracle的Java（所有节点）

CentOS，自带OpenJdk，不过运行CDH5需要使用Oracle的Jdk，需要Java7的支持。

卸载自带的OpenJdk，使用rpm -qa | grep java查询java相关的包，使用rpm -e --nodeps 包名卸载之。

去Oracle的官网下载jdk的rpm安装包，并使用rpm -ivh 包名安装之。

由于是rpm包并不需要我们来配置环境变量，我们只需要配置一个全局的JAVA\_HOME变量即可，执行命令：

echo "JAVA\_HOME=/usr/java/latest/" >> /etc/environment

### 5.安装配置MySql（主节点）

通过yum install mysql-server安装mysql服务器。chkconfig mysqld on设置开机启动，并service mysqld start启动mysql服务，并根据提示设置root的初试密码：mysqladmin -u root password 'xxxx'。

mysql -uroot -pxxxx进入mysql命令行，创建以下数据库：

#hive

create database hive DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

#activity monitor

create database amon DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

设置root授权访问以上所有的数据库：

#授权root用户在主节点拥有所有数据库的访问权限

flush privileges;

grant all privileges on \*.\* to 'root'@'n1' identified by 'xxxx' with grant option;

### 6.关闭防火墙和SELinux

**注意：** 需要在所有的节点上执行，因为涉及到的端口太多了，临时关闭防火墙是为了安装起来更方便，安装完毕后可以根据需要设置防火墙策略，保证集群安全。

关闭防火墙：

service iptables stop （临时关闭）

chkconfig iptables off （重启后生效）

关闭SELINUX:

setenforce 0 （临时生效）

修改 /etc/selinux/config 下的 SELINUX=disabled （重启后永久生效）

### 7.所有节点配置NTP服务

集群中所有主机必须保持时间同步，如果时间相差较大会引起各种问题。 具体思路如下：

master节点作为ntp服务器与外界对时中心同步时间，随后对所有datanode节点提供时间同步服务。

所有datanode节点以master节点为基础同步时间。

所有节点安装相关组件：yum install ntp。完成后，配置开机启动：chkconfig ntpd on,检查是否设置成功：chkconfig --list ntpd其中2-5为on状态就代表成功。

**主节点配置**

在配置之前，先使用ntpdate手动同步一下时间，免得本机与对时中心时间差距太大，使得ntpd不能正常同步。这里选用65.55.56.206作为对时中心,

ntpdate -u 65.55.56.206。

ntp服务只有一个配置文件(/etc/ntp.conf)，配置好了就OK。 这里只给出有用的配置，不需要的配置都用#注掉，这里就不在给出：

driftfile /var/lib/ntp/drift

restrict 127.0.0.1

restrict -6 ::1

restrict default nomodify notrap

server 65.55.56.206 prefer

includefile /etc/ntp/crypto/pw

keys /etc/ntp/keys

配置文件完成，保存退出，启动服务，执行如下命令：service ntpd start

检查是否成功，用ntpstat命令查看同步状态，出现以下状态代表启动成功：

synchronised to NTP server () at stratum 2

time correct to within 74 ms

polling server every 128 s

如果出现异常请等待几分钟，一般等待5-10分钟才能同步。

**配置ntp客户端（所有datanode节点）**

driftfile /var/lib/ntp/drift

restrict 127.0.0.1

restrict -6 ::1

restrict default kod nomodify notrap nopeer noquery

restrict -6 default kod nomodify notrap nopeer noquery

#这里是主节点的主机名或者ip

server n1

includefile /etc/ntp/crypto/pw

keys /etc/ntp/keys

ok保存退出，请求服务器前，请先使用ntpdate手动同步一下时间：ntpdate -u n1 (主节点ntp服务器)

这里可能出现同步失败的情况，请不要着急，一般是本地的ntp服务器还没有正常启动，一般需要等待5-10分钟才可以正常同步。启动服务：service ntpd start

因为是连接内网，这次启动等待的时间会比master节点快一些，但是也需要耐心等待一会儿。

8.配置普通用户的sudo权限

/etc/sudoers

>>> 写权限

# chmod u+w /etc/sudoers

>>> 添加

# vi /etc/sudoers

内容：

beifeng ALL=(root)NOPASSWD:ALL

>>> 收回权限

# chmod u-w /etc/sudoers

## 正式安装

### 安装Cloudera Manager Server 和Agent

**主节点解压安装**

cloudera manager的目录默认位置在/opt下，解压：

tar xzvf cloudera-manager\*.tar.gz将解压后的cm-5.6.0和cloudera目录放到/opt目录下。

**为Cloudera Manager 5建立数据库**

首先需要去MySql的官网下载JDBC驱动，

<http://dev.mysql.com/downloads/connector/j/>

解压后，找到相对应的jar包，放到/opt/cm-5.6.0/share/cmf/lib/中。

在主节点初始化CM5的数据库：

/opt/cm-5.6.0/share/cmf/schema/scm\_prepare\_database.sh mysql cm -hlocalhost -uroot -pxxxx --scm-host localhost scm scm scm

**Agent配置**

修改/opt/cm-5.1.3/etc/cloudera-scm-agent/config.ini中的server\_host为主节点的主机名。

**同步Agent到其他节点**

scp -r /opt/cm-5.6.0 root@n2:/opt/

**在所有节点创建cloudera-scm用户**

useradd --system --home=/opt/cm-5.6.0/run/cloudera-scm-server/ --no-create-home --shell=/bin/false --comment "Cloudera SCM User" cloudera-scm

**准备Parcels，用以安装CDH5**

将CHD5相关的Parcel包放到主节点的/opt/cloudera/parcel-repo/目录中。

相关的文件如下：

* CDH-5.6.0-1.cdh5.6.0.p0.45-el6.parcel
* CDH-5.6.0-1.cdh5.6.0.p0.45-el6.parcel.sha1
* manifest.json
* 最后将CDH-5.6.0-1.cdh5.6.0.p0.45-el6.parcel.sha1，重命名为CDH-5.6.0-1.cdh5.6.0.p0.45-el6.parcel.sha，这点必须注意，否则，系统会重新下载CDH-5.6.0-1.cdh5.6.0.p0.45-el6.parcel文件。

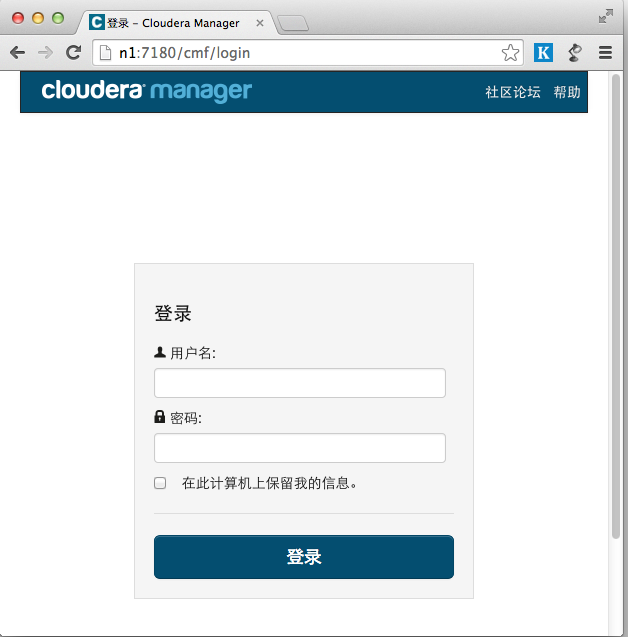
**相关启动脚本**

通过/opt/cm-5.6.0/etc/init.d/cloudera-scm-server start启动服务端。

通过/opt/cm-5.6.0/etc/init.d/cloudera-scm-agent start启动Agent服务。

我们启动的其实是个service脚本，需要停止服务将以上的start参数改为stop就可以了，重启是restart。

Cloudera Manager Server和Agent都启动以后,这时可以通过浏览器访问主节点的7180端口测试一下了（由于CM Server的启动需要花点时间，这里可能要等待一会才能访问），默认的用户名和密码均为admin：如果能够看到如下界面,则安装成功.



>>>>>>>>>>>>>>>>>>>>>>>>>>>>yum本地源>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>

1) Apache 服务器

[beifeng@bigdata-cdh01 ~]$ sudo service httpd status

httpd is stopped

[beifeng@bigdata-cdh01 ~]$ sudo service httpd start

Starting httpd: [ OK ]

[beifeng@bigdata-cdh01 ~]$ sudo chkconfig httpd on

2) 搭建本地源

[beifeng@bigdata-cdh01 html]$ pwd

/var/www/html

[beifeng@bigdata-cdh01 html]$ sudo mkdir -p cm5/redhat/6/x86\_64/cm/5.3.6/RPMS/x86\_64/

[beifeng@bigdata-cdh01 x86\_64]$ pwd

/var/www/html/cm5/redhat/6/x86\_64/cm/5.3.6/RPMS/x86\_64

[beifeng@bigdata-cdh01 x86\_64]$ sudo mv /opt/softwares/cm-5.3.6/\* .

