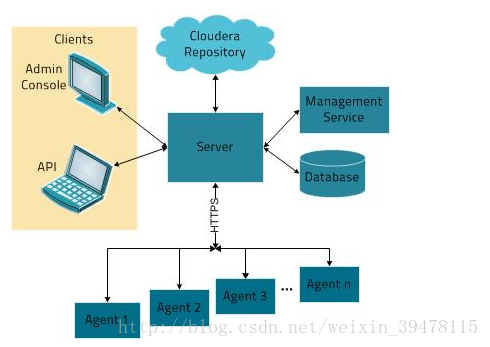
# Cloudera-Manager5.14.0离线安装

Cloudera Manager是cloudera公司提供的一种大数据的解决方案，可以通过ClouderaManager管理界面来对我们的集群进行安装和操作，提供了良好的UI界面交互，使得我们管理集群不用熟悉任何的linux技术，只需要通过网页浏览器就可以实现我们的集群的操作和管理，让我们使用和管理集群更加的方便。

# 1、ClouderaManager整体架构



Cloudera Manager的核心是Cloudera Manager Server。Server托管Admin Console Web Server和应用程序逻辑。它负责安装软件、配置、启动和停止服务以及管理运行服务的群集。

解释：

* Agent：安装在每台主机上。它负责启动和停止进程，解压缩配置，触发安装和监控主机
* Management Service：执行各种监控、报警和报告功能的一组角色的服务。
* Database：存储配置和监控信息
* Cloudera Repository：可供Cloudera Manager分配的软件的存储库（repo库）
* Client：用于与服务器进行交互的接口：   
  + Admin Console：管理员控制台
  + API：开发人员使用 API可以创建自定义的Cloudera Manager应用程序

## ****Cloudera Management Service****

Cloudera Management Service 可作为一组角色实施各种管理功能

* Activity Monitor：收集有关服务运行的活动的信息
* Host Monitor：收集有关主机的运行状况和指标信息
* Service Monitor：收集有关服务的运行状况和指标信息
* Event Server：聚合组件的事件并将其用于警报和搜索
* Alert Publisher ：为特定类型的事件生成和提供警报
* Reports Manager：生成图表报告，它提供用户、用户组的目录的磁盘使用率、磁盘、io等历史视图

## ****信号检测****

默认情况下，Agent 每隔 15 秒向 Cloudera Manager Server 发送一次检测信号。但是，为了减少用户延迟，在状态变化时会提高频率。

## ****状态管理****

* 模型状态捕获什么进程应在何处运行以及具有什么配置
* 运行时状态是哪些进程正在何处运行以及正在执行哪些命令（例如，重新平衡 HDFS 或执行备份/灾难恢复计划或滚动升级或停止）
* 当您更新配置（例如Hue Server Web 端口）时，您即更新了模型状态。但是，如果 Hue 在更新时正在运行，则它仍将使用旧端口。当出现这种不匹配情况时，角色会标记为具有”过时的配置”。要重新同步，您需重启角色（这会触发重新生成配置和重启进程）
* 特殊情况如果要加入一些clouder manager控制台没有的属性时候都在高级里面嵌入

## ****服务器和客户端配置****

* 如使用HDFS，文件 /etc/hadoop/conf/hdfs-site.xml 仅包含与 HDFS 客户端相关的配置
* 而 HDFS 角色实例（例如，NameNode 和 DataNode）会从/var/run/cloudera-scm-agent/process/unique-process-name下的每个进程专用目录获取它们的配置

## ****进程管理****

* 在 Cloudera Manager 管理的群集中，只能通过 Cloudera Manager 启动或停止服务。ClouderaManager 使用一种名为 supervisord的开源进程管理工具，它会重定向日志文件，通知进程失败，为合适用户设置调用进程的有效用户 ID 等等
* Cloudera Manager 支持自动重启崩溃进程。如果一个角色实例在启动后反复失败，Cloudera Manager还会用不良状态标记该实例
* 特别需要注意的是，停止 Cloudera Manager 和 Cloudera Manager Agent 不会停止群集；所有正在运行的实例都将保持运行
* Agent 的一项主要职责是启动和停止进程。当 Agent 从检测信号检测到新进程时，Agent 会在/var/run/cloudera-scm-agent 中为它创建一个目录，并解压缩配置
* Agent 受到监控，属于 Cloudera Manager 的主机监控的一部分：如果 Agent 停止检测信号，主机将被标记为运行状况不良

## ****主机管理****

* Cloudera Manager 自动将作为群集中的托管主机身份：JDK、Cloudera Manager Agent、CDH、Impala、Solr 等参与所需的所有软件部署到主机
* Cloudera Manager 提供用于管理参与主机生命周期的操作以及添加和删除主机的操作
* Cloudera Management Service Host Monitor 角色执行运行状况检查并收集主机度量，以使您可以监控主机的运行状况和性能

## ****安全****

* **身份验证**   
  + Hadoop中身份验证的目的仅仅是证明用户或服务确实是他或她所声称的用户或服务，通常，企业中的身份验证通过单个分布式系统（例如，轻型目录访问协议 (LDAP) 目录）进行管理。LDAP身份验证包含由各种存储系统提供支持的简单用户名/密码服务
  + Hadoop 生态系统的许多组件会汇总到一起来使用 Kerberos 身份验证并提供用于在 LDAP 或 AD 中管理和存储凭据的选项
* **授权**   
  CDH 当前提供以下形式的访问控制：   
  + 适用于目录和文件的传统 POSIX 样式的权限
  + 适用于 HDFS 的扩展的访问控制列表 (ACL)
  + Apache HBase 使用 ACL 来按列、列族和列族限定符授权各种操作 (READ, WRITE, CREATE, ADMIN)
  + 使用 Apache Sentry 基于角色进行访问控制
* **加密**   
  + 需要获得企业版的Cloudera（Cloudera Navigator 许可）

# 2、clouderaManager环境安装前准备

准备两台虚拟机，其中一台作为我们的主节点，安装我们的ClouderaManager Server与ClouderaManager agent，另外一台作为我们的从节点只安装我们的clouderaManager agent

机器规划如下

|  |  |  |
| --- | --- | --- |
| 服务器IP | 192.168.52.100 | 192.168.52.110 |
| 主机名 | node01.hadoop.com | node02.hadoop.com |
| 主机名与IP地址映射 | 是 | 是 |
| 防火墙 | 关闭 | 关闭 |
| selinux | 关闭 | 关闭 |
| jdk | 安装 | 安装 |
| ssh免密码登录 | 是 | 是 |
| mysql数据库 | 否 | 是 |
| 服务器内存 | 16G | 8G |

所有机器统一两个路径

mkdir -p /export/softwares/

mkdir -p /export/servers/

## 2.1、两台机器更改主机名

第一台机器更改主机名

vim /etc/sysconfig/network

NETWORKING=yes

HOSTNAME=node01.hadoop.com

第二台机器更改主机名

vim /etc/sysconfig/network

NETWORKING=yes

HOSTNAME=node02.hadoop.com

## 2.2、更改主机名与IP地址的映射

两台机器更改hosts文件

vim /etc/hosts

192.168.52.100 node01.hadoop.com node01

192.168.52.110 node02.hadoop.com node02

## 2.3、两台机器关闭防火墙

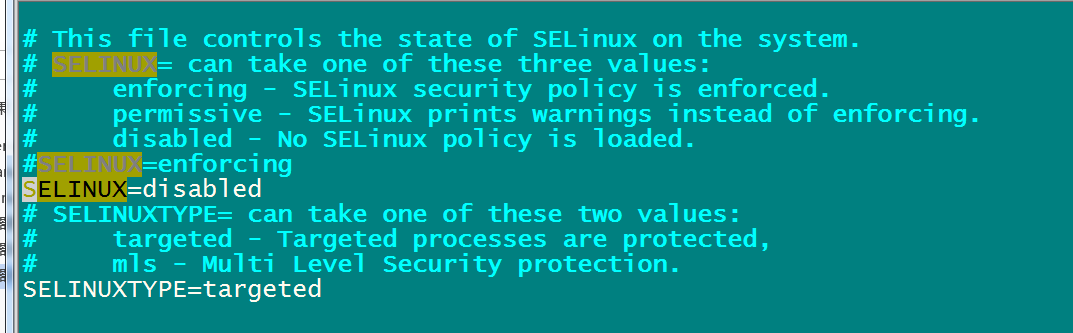
service iptables stop

chkconfig iptables off

## 2.4、两台机器关闭selinux

vim /etc/selinux/config

SELINUX=disabled



## 2.5、两台机器安装jdk

将我们的jdk的压缩包上传到node01.hadoop.com的/export/softwares路径下

cd /export/softwares/

tar -zxvf jdk-8u141-linux-x64.tar.gz -C /export/servers/

配置环境变量

vim /etc/profile

export JAVA\_HOME=/export/servers/jdk1.8.0\_141

export PATH=:$JAVA\_HOME/bin:$PATH

source /etc/profile

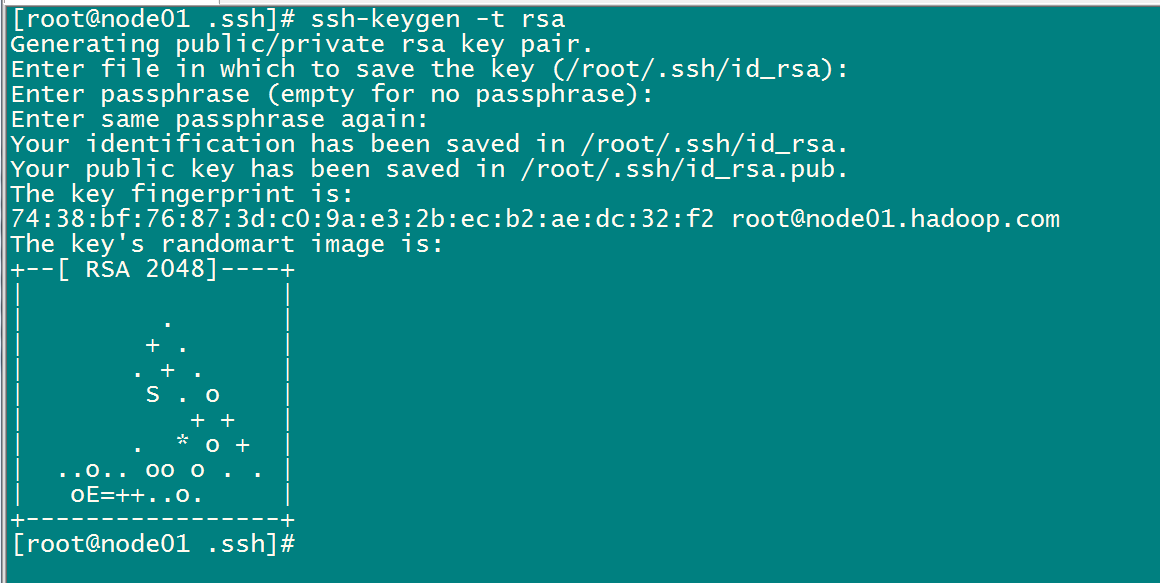
第二台机器同样安装jdk即可

## 2.6、两台机器实现SSH免密码登录

### 第一步：两台器生成公钥与私钥

两台机器上面执行以下命令，然后按下三个回车键即可生成公钥与私钥

ssh-keygen -t rsa



### 第二步：两台机器将公钥拷贝到同一个文件当中去

两台机器执行以下命令

ssh-copy-id node01.hadoop.com

### 第三步：拷贝authorized\_keys到其他机器

第一台机器上将authorized\_keys拷贝到第二台机器

scp /root/.ssh/authorized\_keys node02.hadoop.com:/root/.ssh/

## 2.7、第二台机器安装mysql数据库

通过yum源，在线安装mysql

yum install mysql mysql-server mysql-devel

/etc/init.d/mysqld start

/usr/bin/mysql\_secure\_installation

grant all privileges on \*.\* to 'root'@'%' identified by '123456' with grant option;

flush privileges;



## 2.8、解除linux系统打开文件最大数量的限制

两台机器都需要执行

vi /etc/security/limits.conf

添加以下内容

\* soft noproc 11000

\* hard noproc 11000

\* soft nofile 65535

\* hard nofile 65535

## 2.9、设置linux交换区内存

两台机器都要执行

执行命令

echo 10 > /proc/sys/vm/swappiness

并编辑文件sysctl.conf：

vim /etc/sysctl.conf

添加或修改

vm.swappiness = 0

两台机器都要执行：

echo never > /sys/kernel/mm/transparent\_hugepage/defrag

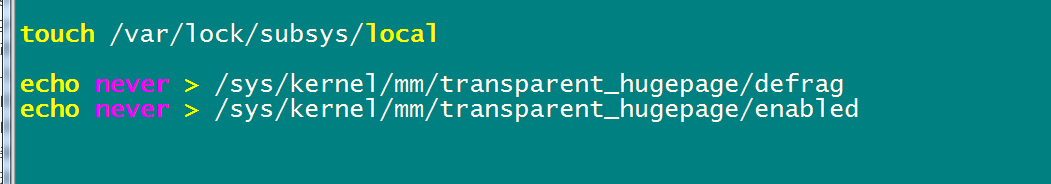
echo never > /sys/kernel/mm/transparent\_hugepage/enabled

并编辑文件rc.local ：

vim /etc/rc.local

echo never > /sys/kernel/mm/transparent\_hugepage/defrag

echo never > /sys/kernel/mm/transparent\_hugepage/enabled



## 2.10、两台机器时钟同步

两台机器需要进行时钟同步操作，保证两台机器时间相同

# 3、clouderaManager安装资源下载

## 第一步：下载安装资源并上传到服务器

我们这里安装CM5.14.0这个版本，需要下载以下这些资源，一共是四个文件即可

下载cm5的压缩包

下载地址：<http://archive.cloudera.com/cm5/cm/5/>

具体文件地址：

<http://archive.cloudera.com/cm5/cm/5/cloudera-manager-el6-cm5.14.0_x86_64.tar.gz>

下载cm5的parcel包

下载地址：

<http://archive.cloudera.com/cdh5/parcels/>

第一个文件具体下载地址：

<http://archive.cloudera.com/cdh5/parcels/5.14.0/CDH-5.14.0-1.cdh5.14.0.p0.24-el6.parcel>

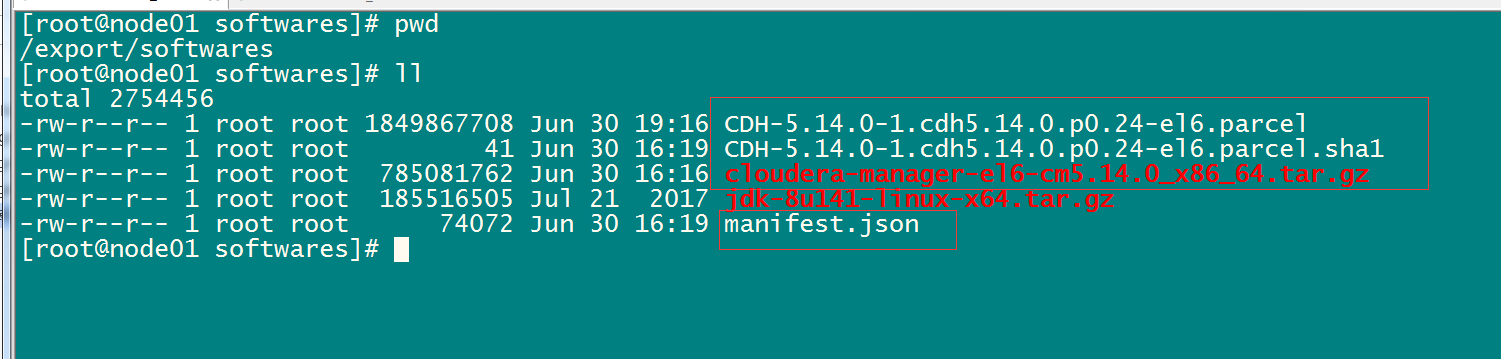
第二个文件具体下载地址：

<http://archive.cloudera.com/cdh5/parcels/5.14.0/CDH-5.14.0-1.cdh5.14.0.p0.24-el6.parcel.sha1>

第三个文件具体下载地址：

<http://archive.cloudera.com/cdh5/parcels/5.14.0/manifest.json>

将这四个安装包都上传到第一台机器的/opt/softwares路径下



## 第二步：解压压缩包到指定路径

解压CM安装包到/opt路径下去

cd /export/softwares

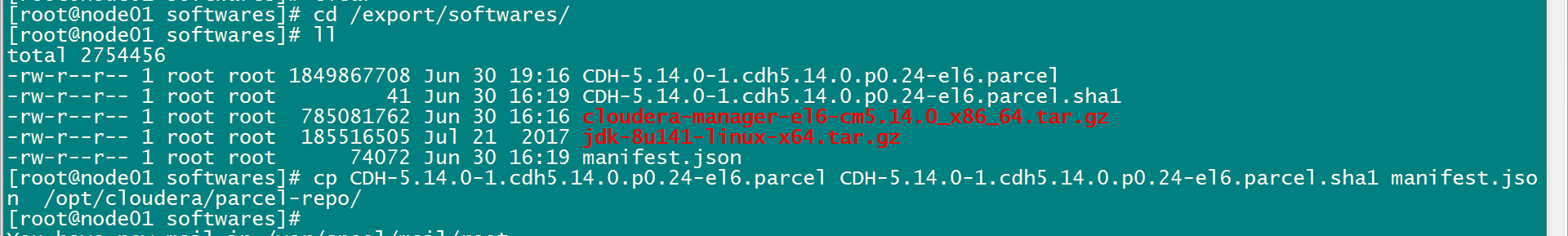
tar -zxvf cloudera-manager-el6-cm5.14.0\_x86\_64.tar.gz -C /opt/

## 第三步：将我们的parcel包的三个文件拷贝到对应路径

将我们的parcel包含三个文件，拷贝到/opt/cloudera/parcel-repo路径下面去，并记得有个文件需要重命名

cd /export/softwares/

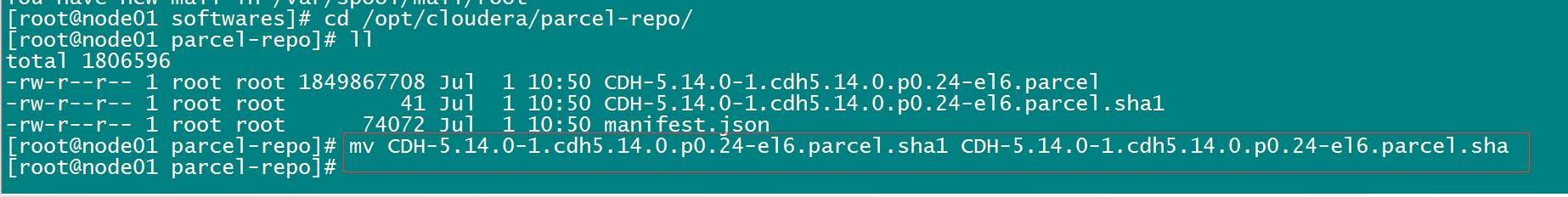
cp CDH-5.14.0-1.cdh5.14.0.p0.24-el6.parcel CDH-5.14.0-1.cdh5.14.0.p0.24-el6.parcel.sha1 manifest.json /opt/cloudera/parcel-repo/



重命名标黄的这个文件

cd /opt/cloudera/parcel-repo/

mv CDH-5.14.0-1.cdh5.14.0.p0.24-el6.parcel.sha1 CDH-5.14.0-1.cdh5.14.0.p0.24-el6.parcel.sha



## 第四步：所有节点添加普通用户并给与sudo权限

在node01机器上面添加普通用户并赋予sudo权限

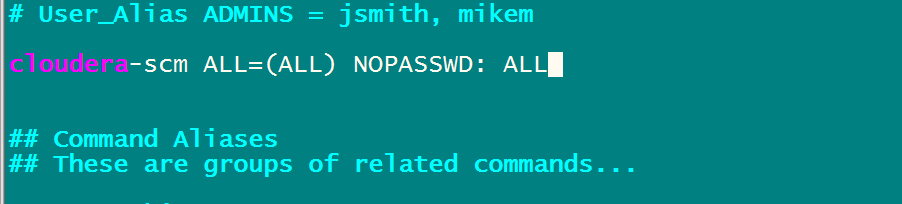
执行以下命令创建普通用户cloudera-scm

useradd --system --home=/opt/cm-5.14.0/run/cloudera-scm-server --no-create-home --shell=/bin/false --comment "Cloudera SCM User" cloudera-scm

赋予cloudera-scm普通用户的sudo权限

vim /etc/sudoers

cloudera-scm ALL=(ALL) NOPASSWD: ALL

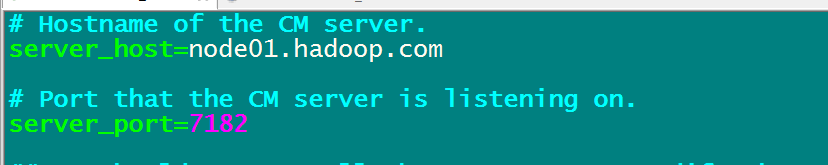


## 第五步：更改主节点的配置文件

node01机器上面更改配置文件

vim /opt/cm-5.14.0/etc/cloudera-scm-agent/config.ini

server\_host=node01.hadoop.com



## 第六步：将/opt目录下的安装包发放到其他机器

将第一台机器的安装包发放到其他机器

cd /opt

scp -r cloudera/ cm-5.14.0/ node02:/opt

## 第七步：创建一些数据库备用

node02机器上面创建数据库

hive 数据库

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

集群监控数据库

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

hue 数据库

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

oozie 数据库

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

## 第八步：准备数据库连接的驱动包

在所有机器上面都准备一份数据库的连接驱动jar包放到/usr/share/java路径下

准备一份mysql的驱动连接包，放到/usr/share/java路径下去

cd /export/softwares/

wget https://dev.mysql.com/get/Downloads/Connector-J/mysql-connector-java-5.1.45.tar.gz

tar -zxvf mysql-connector-java-5.1.45.tar.gz

cd /export/softwares/mysql-connector-java-5.1.45

cp mysql-connector-java-5.1.45-bin.jar /usr/share/java/mysql-connector-java.jar

拷贝驱动包到第二台机器

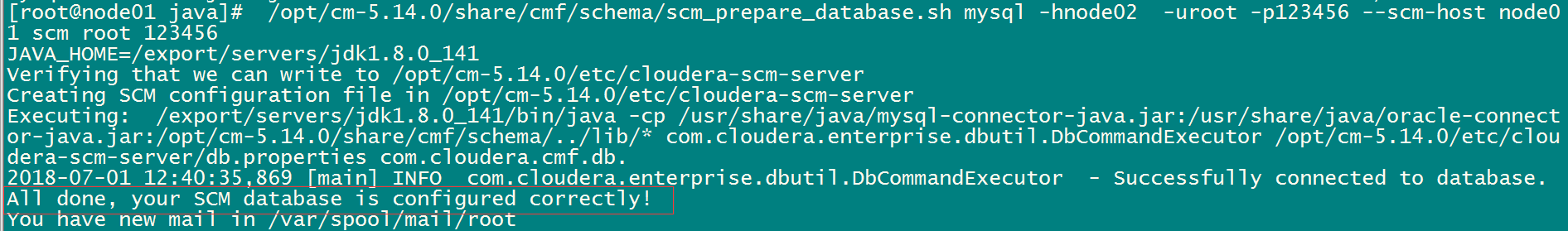
cd /usr/share/java

scp mysql-connector-java.jar node02:$PWD

## 第九步：为clouderaManager创建数据库

/opt/cm-5.14.0/share/cmf/schema/scm\_prepare\_database.sh mysql -hnode02 -uroot -p123456 --scm-host node01 scm root 123456

命令说明：/**opt**/**cm**-5.14.0/share/cmf/schema/scm\_prepare\_database.**sh** 数据库类型 -h数据库主机 –u数据库用户名 –p数据库密码 --scm-host **cm**主机 数据库名称 用户名 密码

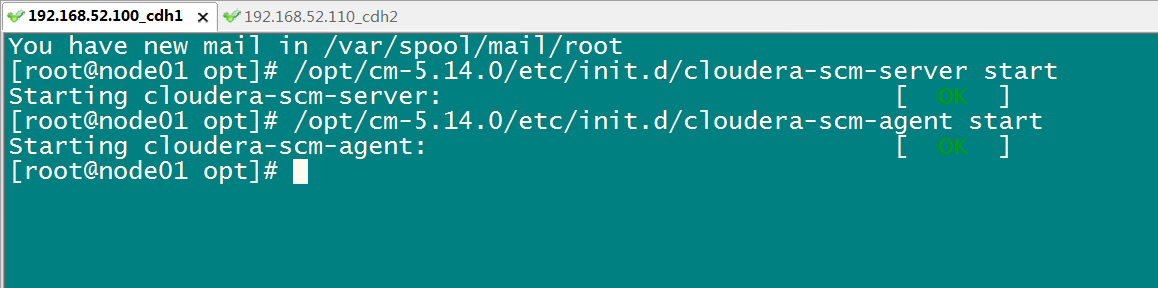


## 第十步：启动服务

主节点启动clouderaManager Server与ClouderaManager agent服务

/opt/cm-5.14.0/etc/init.d/cloudera-scm-server start

/opt/cm-5.14.0/etc/init.d/cloudera-scm-agent start



从节点node02启动ClouderaManager agent服务

## 第十一步：浏览器页面访问

<http://node01:7180/cmf/login>

默认用户名admin

密码 admin