[概要 6](#_Toc23569)

[安装向导 7](#_Toc16)

[1. 配置操作系统 7](#_Toc26373)

[1.1. 配置所有服务器各自的网络静态IP地址： 7](#_Toc20411)

[1.2. 配置所有服务器的各自的主机名 8](#_Toc14723)

[1.3. 配置所有服务器的网络名称 8](#_Toc23135)

[1.4. 为所有服务器配置集群信息（所有服务器配置统一集群信息） 9](#_Toc16226)

[1.5. 停止所有服务器的防火墙，并关闭所有防火墙 9](#_Toc21252)

[1.6. 为所有服务器配置安全权限信息 9](#_Toc17362)

[1.7. 配置ssh免密码登陆 10](#_Toc21306)

[1.8. 同步时间 11](#_Toc4283)

[2. JDK安装 11](#_Toc8573)

[3. 资源库安装 12](#_Toc11796)

[3.1. 配置一个本地系统yum源，创建资源目录 12](#_Toc15275)

[3.2. 使用自己配置的资源库，配置集群各节点的yum源 13](#_Toc26876)

[3.3. 使用公司统一配置的资源库，配置集群各节点的yum源 14](#_Toc20027)

[4. MYSQL的安装和配置 14](#_Toc18586)

[4.1. 检查操作系统以前是否安装有MySQL 14](#_Toc17939)

[4.2. 安装MySQL 15](#_Toc1816)

[4.3. 创建数据库并赋权 18](#_Toc23088)

[5. 安装集群管理节点 19](#_Toc5923)

[5.1. 通过yum源安装监控管理系统 19](#_Toc4996)

[5.2. 创建一个MySQL的JDBC目录 19](#_Toc17124)

[5.3. 配置集群管理节点 19](#_Toc19627)

[5.4. 配置数据仓库组件所需的MySQL的JDBC包 22](#_Toc5967)

[6. 界面化一键部署集群 22](#_Toc12395)

[6.0.安装smcs-web 22](#_Toc10028)

[6.1.添加license 23](#_Toc13598)

[6.2.新建集群 23](#_Toc19287)

[6.3.版本选择 24](#_Toc17522)

[6.4.选择节点 24](#_Toc29185)

[6.5.确认主机 25](#_Toc31268)

[6.6.选择组件 26](#_Toc23802)

[6.7.启动安装部署 28](#_Toc31140)

[7. 安装MPP大数据引擎 29](#_Toc28496)

[7.1. 修改配置信息 29](#_Toc6594)

[7.2. 安装插件 31](#_Toc27835)

[7.3. 重新登录到海盒大数据平台的界面 33](#_Toc24274)

[非root用户启动ambari 37](#_Toc25017)

[1. 创建用户 37](#_Toc16109)

[2. 添加ambari 用户添加变量 37](#_Toc24905)

[3. 配置ambari 38](#_Toc18742)

[4. 修改ambari-agent.ini 38](#_Toc28061)

[备注 39](#_Toc20557)

[5. 防火墙操作(centos7.x centos6.x) 39](#_Toc19263)

[6. 安装注意事项 39](#_Toc31063)

[7. 组件启动报错解决方案 39](#_Toc14969)

[8. Pxf启动报错解决方案 39](#_Toc23617)

[9. 补充：安装mysql5.7操作。 40](#_Toc12557)

[1) 检查操作系统以前是否安装有MySQL 40](#_Toc1836)

[2) 安装MySQL 40](#_Toc15987)

[10. 补充：安装mvn 42](#_Toc1571)

[11. 同步网络时间 42](#_Toc22326)

[关于我们 42](#_Toc22916)



**海盒大数据平台**

**—数 据 定 义 未 来 —**

**CentOS 7 系统**

**安装手册**

北京东方金信科技有限公司

免责声明

本操作手册依据现有信息制作，其内容如有更改，恕不另行通知。北京东方金信科技有限公司在编写本操作手册时已尽最大努力保证其内容准确可靠，但北京东方金信科技有限公司不对本操作手册中的遗漏、不准确或印刷错误导致的损失和损害承担责任。具体产品使用请以实际使用为准。

注释: Hadoop® 和 SPARK® 是Apache™ 软件基金会在美国和其他国家的商标或注册的商标。 Java® 是Oracle公司在美国和其他国家的商标或注册的商标。 Intel® 和Xeon® 是英特尔公司在美国、中国和其他国家的商标或注册的商标。

版权所有 © 北京东方金信科技有限公司。保留所有权利。

© 北京东方金信科技有限公司版权所有，并保留对本操作手册及本声明的最终解释权和修改权。本操作手册的版权归北京东方金信科技有限公司，未得到北京东方金信科技有限公司的书面许可，任何人不得以任何方式或形式对本说明书内的任何部分进行复制、摘录、备份、修改、传播、翻译成其他语言、或将其全部或部分用于商业用途。

海盒大数据平台版本：2.0v

操作系统版本：CentOS 7

安装手册版本：3.0v

发布日期： 2016-06-22

目录

[概要 6](#_Toc14811)

[安装向导 7](#_Toc9712)

[1. 配置操作系统 7](#_Toc11585)

[1.1. 配置所有服务器各自的网络静态IP地址： 7](#_Toc29840)

[1.2. 配置所有服务器的各自的主机名 8](#_Toc1803)

[1.3. 配置所有服务器的网络名称 8](#_Toc21968)

[1.4. 为所有服务器配置集群信息（所有服务器配置统一集群信息） 9](#_Toc22677)

[1.5. 停止所有服务器的防火墙，并关闭所有防火墙 9](#_Toc15871)

[1.6. 为所有服务器配置安全权限信息 9](#_Toc21442)

[1.7. 配置ssh免密码登陆 10](#_Toc21221)

[1.8. 同步时间 11](#_Toc7404)

[2. JDK安装 11](#_Toc3701)

[3. 资源库安装 12](#_Toc6448)

[3.1. 配置一个本地系统yum源，创建资源目录 12](#_Toc12396)

[3.2. 使用自己配置的资源库，配置集群各节点的yum源 13](#_Toc1059)

[3.3. 使用公司统一配置的资源库，配置集群各节点的yum源 14](#_Toc15787)

[4. MYSQL的安装和配置 14](#_Toc12347)

[4.1. 检查操作系统以前是否安装有MySQL 14](#_Toc899)

[4.2. 安装MySQL 15](#_Toc12142)

[4.3. 创建数据库并赋权 18](#_Toc1622)

[5. 安装集群管理节点 19](#_Toc19923)

[5.1. 通过yum源安装监控管理系统 19](#_Toc15921)

[5.2. 创建一个MySQL的JDBC目录 19](#_Toc22091)

[5.3. 配置集群管理节点 19](#_Toc22863)

[5.4. 配置数据仓库组件所需的MySQL的JDBC包 22](#_Toc6527)

[6. 界面化一键部署集群 22](#_Toc15271)

[6.0.安装smcs-web 22](#_Toc12724)

[6.1.添加license 23](#_Toc27633)

[6.2.新建集群 23](#_Toc4668)

[6.3.版本选择 24](#_Toc25029)

[6.4.选择节点 24](#_Toc10137)

[6.5.确认主机 25](#_Toc1474)

[6.6.选择组件 26](#_Toc13704)

[6.7.启动安装部署 28](#_Toc31712)

[7. 安装MPP大数据引擎 29](#_Toc19085)

[7.1. 修改配置信息 29](#_Toc20822)

[7.2. 安装插件 31](#_Toc7354)

[7.3. 重新登录到海盒大数据平台的界面 33](#_Toc12995)

[非root用户启动ambari 37](#_Toc448)

[1. 创建用户 37](#_Toc13425)

[2. 添加ambari 用户添加变量 37](#_Toc1172)

[3. 配置ambari 38](#_Toc17907)

[4. 修改ambari-agent.ini 38](#_Toc27949)

[备注 39](#_Toc29548)

[5. 防火墙操作(centos7.x centos6.x) 39](#_Toc29467)

[6. 安装注意事项 39](#_Toc27192)

[7. 组件启动报错解决方案 39](#_Toc23060)

[8. Pxf启动报错解决方案 39](#_Toc9178)

[9. 补充：安装mysql5.7操作。 40](#_Toc28582)

[1) 检查操作系统以前是否安装有MySQL 40](#_Toc29016)

[2) 安装MySQL 40](#_Toc23467)

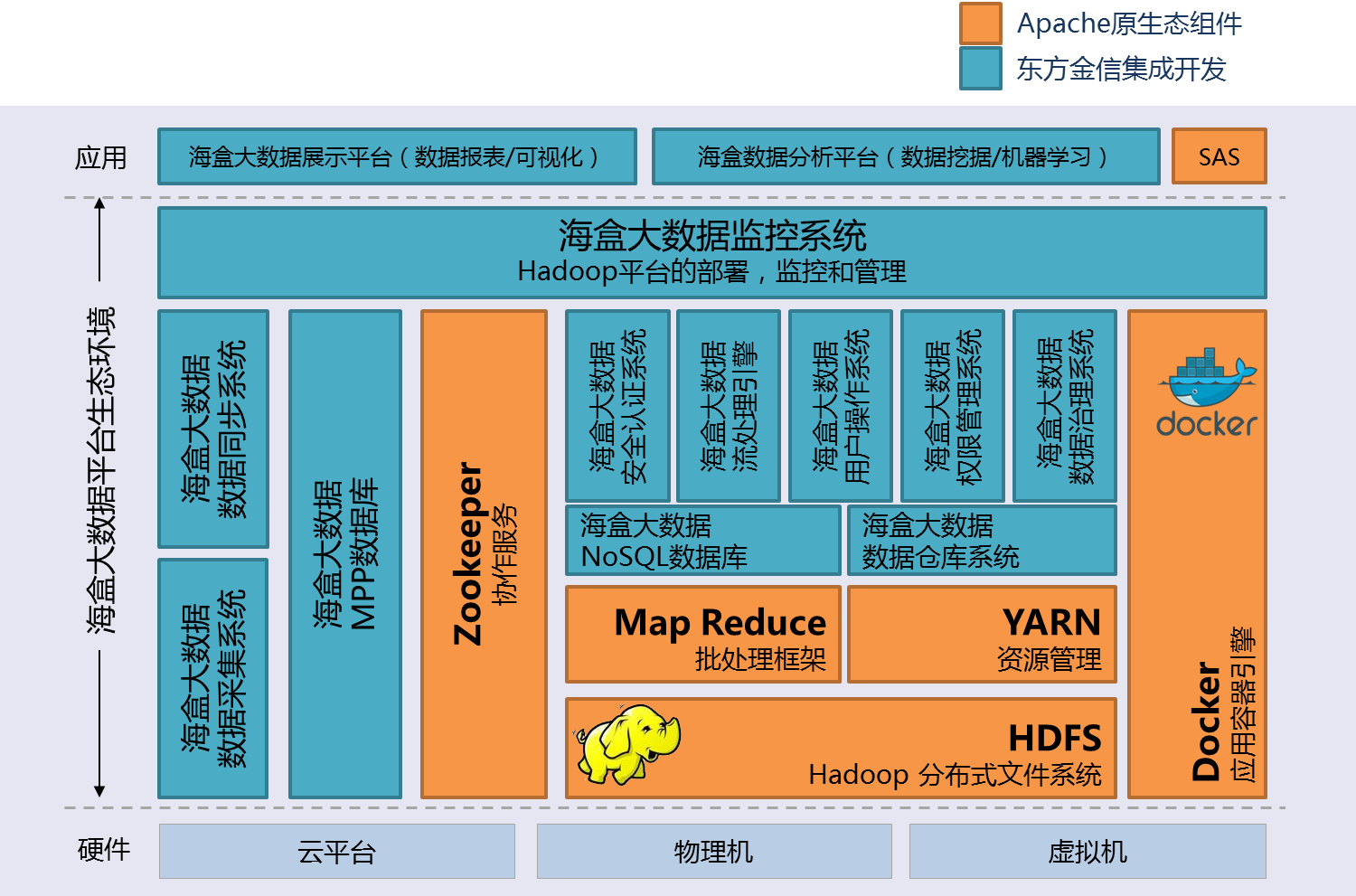
[10. 补充：安装mvn 42](#_Toc12907)

[11. 同步网络时间 42](#_Toc310)

[关于我们 42](#_Toc7974)

# 概要

海盒大数据平台由东方金信科技有限公司开发和管理，是一个集数据实时同步、分布式数据存储、大数据计算和可视化分析等功能为一体的大数据平台。用户可以利用海盒大数据平台对自身的内外部数据进行历史数据和实时数据的存储备份，实现多角度和多维度的数据处理、分析、挖掘、预测、展现和管理。海盒大数据平台主要包括以下组件：



海盒大数据平台监控系统提供了非常直观的用户界面，帮助用户轻松有效地监控并管理大数据集群。该系统具有以下特点：

* 通过逐步的安装向导简化了集群的搭建。
* 预先配置好关键的运维指标（metrics），帮助用户直接查询大数据平台核心组件及拓展组件的健康。
* 支持作业与任务执行的可视化与分析。
* 提供自动集成的运维工具，帮助用户有效的监控并管理集群。

本手册将帮助客户离线搭建基于CentOS7系统的海盒大数据平台。

# 安装向导

## 配置操作系统

操作系统的语言必须是英文环境

**声明：若是手动安装，配置完IP MYSQL JDK之后强烈建议保存快照，安装完海盒正常后另存快照**

**保存快照如下图：**



### 配置所有服务器各自的网络静态IP地址：

其中红色部分为必须修改的内容，其他可选。vi /etc/sysconfig/network-scripts/ifcfg-eno16777736

>vi /etc/sysconfig/network-scripts/ifcfg-enoXXX（该文件可能会因环境而异）

#################################################

TYPE=Ethernet

BOOTPROTO=static

DEFROUTE=yes

IPV4\_FAILURE\_FATAL=no

IPV6INIT=yes

IPV6\_AUTOCONF=yes

IPV6\_DEFROUTE=yes

IPV6\_FAILURE\_FATAL=no

NAME=centos7\_ambari\_demo

ONBOOT=yes

DNS1=<dns1>

DNS2=<dns2>

DNS3=<dns3>

DNS4=<dns4>

IPADDR=对应网段的IP地址(10.1.3.103)

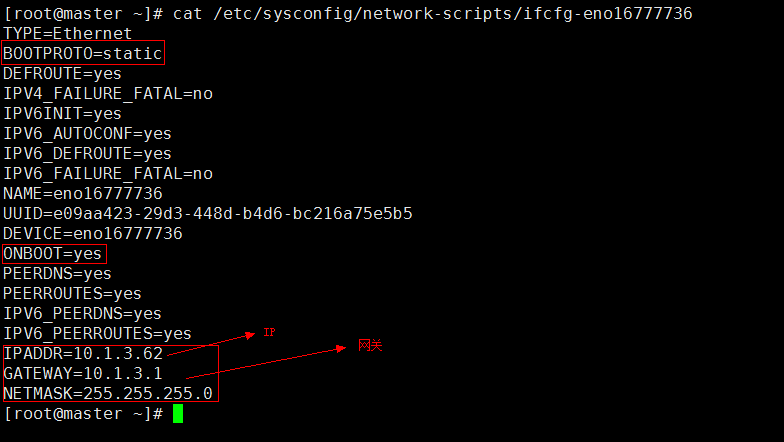
PREFIX=24

GATEWAY=网关

NETMASK=255.255.252.0

IPV6\_PEERDNS=yes

IPV6\_PEERROUTES=yes



配置完成后,重启网络服务：

>service network restart

ifconfig

### 配置所有服务器的各自的主机名

配置每台机器的主机名

注意主机名中可以有横线但是不能有下划线，下划线在scala中为特殊字符，如果mysql数据库对应的主机名有下划线，会导致ambari-serup配置mysql为存储失败，如果集群中的节点主机名字有下划线会导致hdfs集群无法启动。如果主机名字中含有其他特殊字符，如中文句号，逗号，空格，斜线，管道符号，也要立即改正。

方法一：>hostnamectl set-hostname master



方法二：>vi /etc/hostname

##################################################

master

X

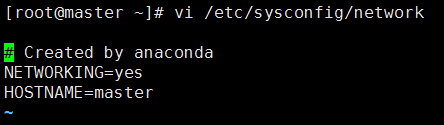
### 配置所有服务器的网络名称

>vi /etc/sysconfig/network

###############################################################

NETWORKING=yes

HOSTNAME=master //master 为你设置的hostname



### 为所有服务器配置集群信息（所有服务器配置统一集群信息）

>vi /etc/hosts

#####################################################

10.1.3.62 master

10.1.3.116 slave1

10.1.3.76slave2

### 停止所有服务器的防火墙，并关闭所有防火墙

>setenforce 0

>vi /etc/selinux/config

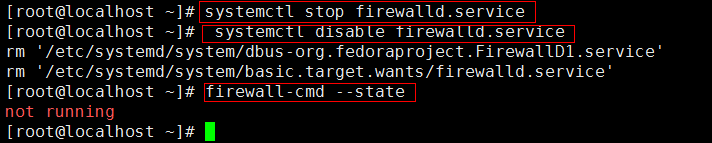
###############################################

SELINUX=disabled #修改内容为disabled

>systemctl stop firewalld.service

>systemctl disable firewalld.service

> firewall-cmd --state #查看防火墙状态(关闭显示notrunning，开启显示running)



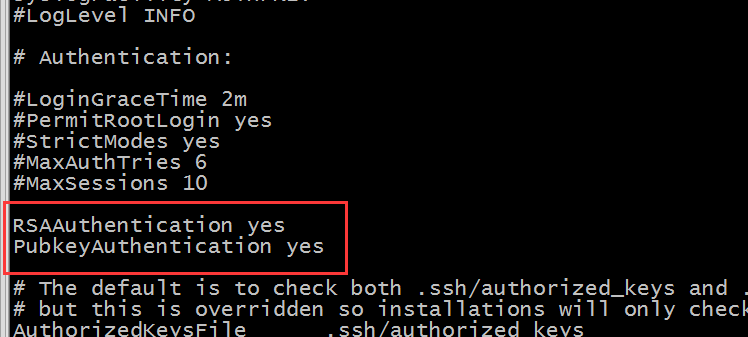
### 为所有服务器配置安全权限信息

>vi /etc/ssh/sshd\_config

去掉注释：

RSAAuthentication yes

PubkeyAuthentication yes

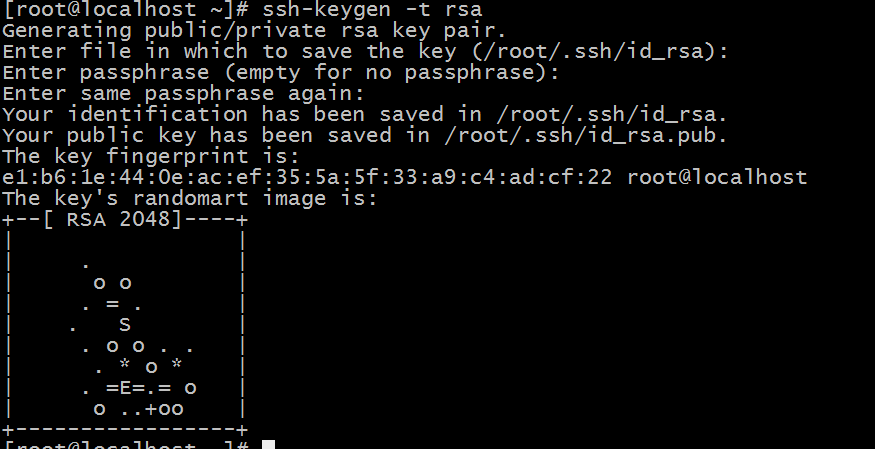


### 配置ssh免密码登陆（必配，否则主机注册时会报错）

为所有服务器配置密码：

>ssh-keygen -t rsa

（过程中输入回车进入下一步即可。）



> cd /root/.ssh

进入.ssh目录，然后分别拷贝其他机器中的id\_rsa.pub到同一机器的目录下,例如：

在master上执行:

>scp id\_rsa.pub slave2:/root/.ssh/id\_rsa-master.pub

在slave1上执行:

>scp id\_rsa.pub slave2:/root/.ssh/id\_rsa-slave1.pub

在slave2上，将所以的id\_rsa\*.pub文件内容追加到authorized\_keys文件中

>cat id\_rsa-master.pub >>authorized\_keys

>cat id\_rsa-slave1.pub >>authorized\_keys

>cat id\_rsa.pub >>authorized\_keys

然后在将生成的authorized\_keys拷贝回其他机器，执行

>scp authorized\_keys master:/root/.ssh/authorized\_keys

>scp authorized\_keysslave1:/root/.ssh/authorized\_keys

确认maste1 和slave1两台机器有authorized\_keys文件,然后删除多余的id\_rsa-master.pub，id\_rsa-slave1.pub文件

用ssh测试结果(首次需要确认，另外测试连接一次退出一次)

>ssh master;

>exit;

chkconfig NetworkManager off

chkconfig network on

service NetworkManager stop

service network start

#### 1.7.1补充单节点配置ssh：

>cd /root/.ssh

>cp id\_rsa.pub authorized\_keys

>ssh-keygen -t rsa  //生成时候不需要密码 直接回车

>ssh-copy-id master//此为主节点主机名

配置完之后检验：

>ssh master//若不需输入密码即可进入，则配置成功

### 同步时间

有时自动安装脚本不能成功安装ntp，在安装完ambari后务必检查集群各个节点的时间，检查是否安装ntp。集群节点的时间不一致会导致hdfs集群启动缓慢或者无法启动。

1. 检查是否安装同步服务软件

rpm -q ntp

1. ambari-server将要安装的服务器上，修改文件/etc/ntp.conf

将一下两行打开注释：



修改/etc/npt.conf

注释以下几行，在行首增加#

server 0.centos.pool.ntp.org iburst

server 1.centos.pool.ntp.org iburst

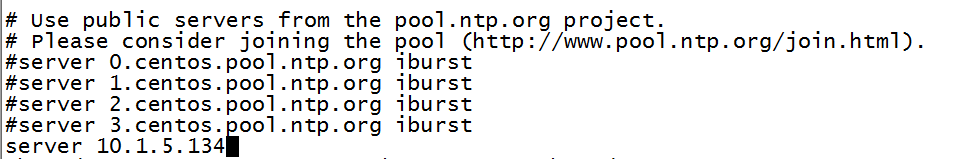
server 2.centos.pool.ntp.org iburst

server 3.centos.pool.ntp.org iburst

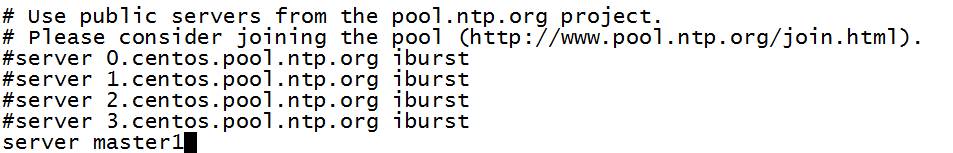
并添加ntp需要同步的服务器的主ntp的IP地址或host主机名

server IP/host

例如



或者



如果这台机器的主机名不在各个节点的/etc/

1. 保存并重启ntpd服务

>/etc/init.d/ntpd restart

1. 等待大约5分钟后（时间服务器自身同步的时间），在客户端与服务器同步时间，可得到一下结果：



## JDK安装

JAVA在lunix上的安装步骤如下流程(本文以安装jdk-7u79-linux-x64.rpm为例)：

1. 注意安装jdk之前，需要把原来有的jdk卸载掉：

>rpm -qa|grep java

>rpm -e --nodeps XXX #卸载掉存在的java程序，保留依赖。

卸载完成后，使用find / -name java 命令查找相关文件和目录，并把这些文件都删除。

1. 下载JDK文件，复制并解压到/usr/lib/jdk：

>mkdir /usr/lib/jdk

>tar -zxvf /tools/jdk-7u79-linux-x64.tar.gz -C /usr/lib/jdk

>ll /usr/lib/jdk

1. 配置环境变量如下

查看并修改文件/etc/profile：

>vim /etc/profile

在该文件添加如下内容：

#JAVA Environment

export JAVA\_HOME=/usr/lib/jdk/jdk1.7.0\_79

export JRE\_HOME=/usr/lib/jdk/jdk1.7.0\_79/jre

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

export CLASSPATH=$CLASSPATH:.:$JAVA\_HOME/lib:$JRE\_HOME/lib

使JDK环境变量生效：

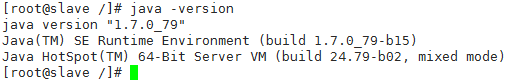
>source /etc/profile

1. 验证

>echo $JAVA\_HOME

>java -version

如果成功输出正确的版本信息，说明安装成功了。



## 资源库安装

建议：配置一个服务器作为资源库服务器，资源库服务器的防火墙必须关闭，设置静态IP地址，在局域网内能访问。

如果选择自己配置yum资源库，请执行3.1以及3.2，并跳过3.3.

如果采用公司统一配置的yum资源库，请跳过3.1以及3.2，直接执行3.3.

### 配置一个本地系统yum源，创建资源目录

解压好的：

<https://10.1.3.186:440/svn/安装包/SDB/当前可用安装包/Release-3.0/Seabox-Centos7.2-release/Centos7.2-Seabox>

>mkdir /sdp-yum-repos

将https://10.1.3.186:440/svn/研发项目/07 IBM与东方金信大数据测试/最终产品软件资源/sdp-yum-repos/Seabox大数据平台/yum\_repos目录下的

centos7-1511-seabox\_repos.tar.gz、seabox-hdp-2.3.2.tar.gz、seabox-hdp-utils-1.1.0.tar.gz、seabox-sdp-2.2.1.tar.gz文件复制到目录/sdp-yum-repos下，然后分别解压：

> cd /sdp-yum-repos

> tar -zxvf centos7-1511-seabox\_repos.tar.gz

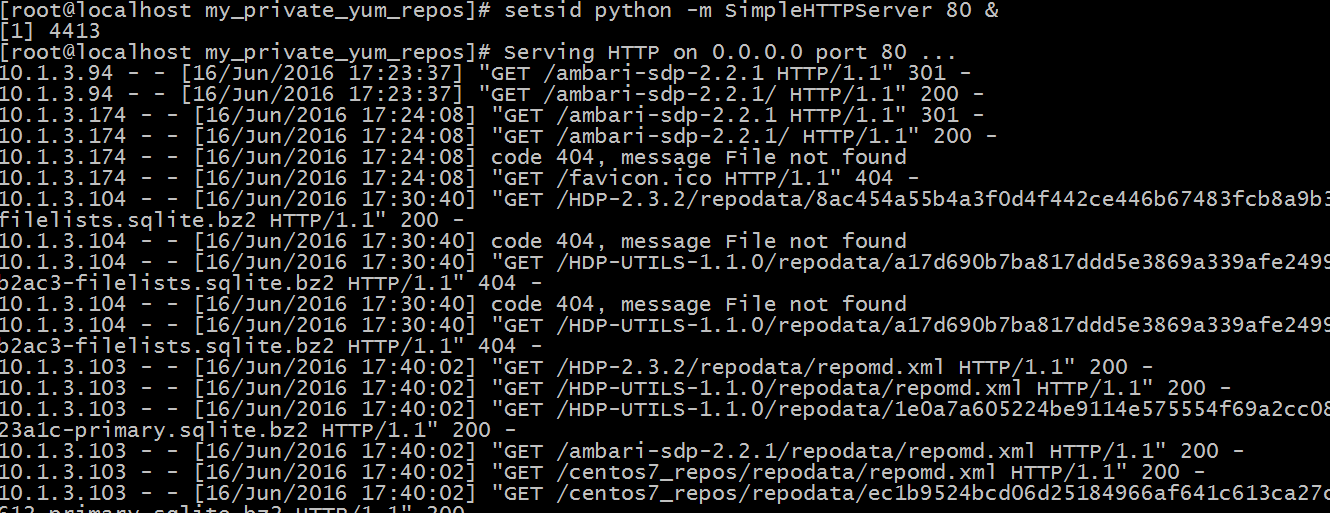
> tar -zxvf seabox-hdp-2.3.2.tar.gz

> tar -zxvf seabox-hdp-utils-1.1.0.tar.gz

> tar -zxvf seabox-sdp-2.2.1.tar.gz

启动服务：

>setsid python -m SimpleHTTPServer 80 &



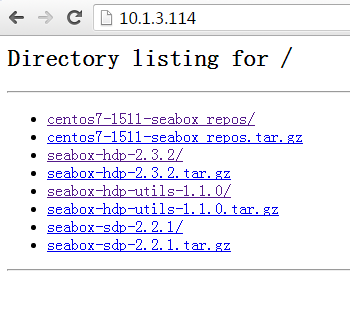
至此资源库创建完成。资源库路径为：

http://<资源服务器IP>/centos7\_repos

http://<资源服务器IP>/ambari-sdp-2.2.1

http://<资源服务器IP>/HDP-2.3.2

http://<资源服务器IP>/HDP-UTILS-1.1.0



### 使用自己配置的资源库，配置集群各节点的yum源

集群的每台机器执行：

删除以前的yum源文件

> cd /etc/yum.repo.d

>rm -rf \*

配置新的yum源：将https://10.1.3.186:440/svn/研发项目/07 IBM与东方金信大数据测试/最终产品软件资源/sdp-yum-repos/Seabox大数据平台/repo\_config\_files目录下的ambari.repo、centos7-1511-seabox.repo、seabox-hdp.repo、seabox-hdp-utils.repo文件复制各个集群的目录/etc/yum.repo.d/下，并进行如下修改：

修改ambari.repo中的baseurl值为：

http://<资源服务器IP>/ seabox-sdp-2.2.1

修改seabox-hdp.repo中的baseurl值为：

http://<资源服务器IP>/seabox-hdp-2.3.2

修改seabox-hdp-utils.repo中的baseurl值为：

http://<资源服务器IP>/seabox-hdp-utils-1.1.0

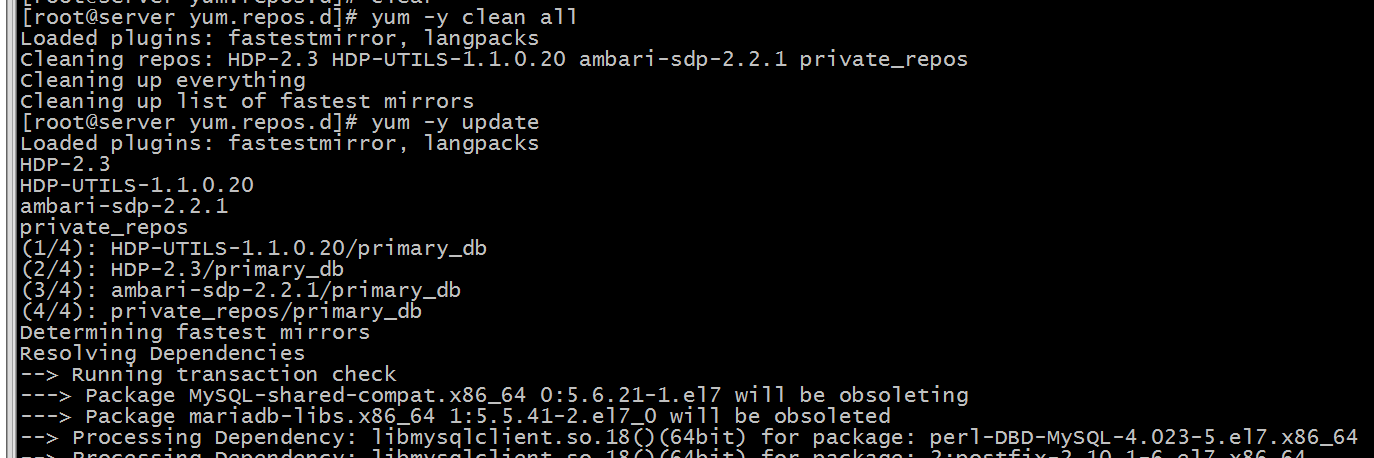
修改centos7-1511-seabox.repo中的baseurl值为：

http://<资源服务器IP>/ centos7-1511-seabox\_repos

修改完后，更新yum缓存：

>yum -y clean all

> yum -y update



### 使用公司统一配置的资源库，配置集群各节点的yum源

**公司公用资源库：10.1.3.50**

集群的每台机器执行：

删除以前的yum源文件

> cd /etc/yum.repo.d/

>rm -rf \*

配置新的yum源：将https://10.1.3.186:440/svn/研发项目/07 IBM与东方金信大数据测试/最终产品软件资源/sdp-yum-repos/Seabox大数据平台/repo\_config\_files目录下的ambari.repo、centos7-1511-seabox.repo、seabox-hdp.repo、seabox-hdp-utils.repo文件复制各个集群的目录/etc/yum.repo.d/下即可，下载完成后，更新yum缓存：

> yum -y clean all

> yum -y update

## MYSQL的安装和配置

MySQL数据库将作为存储平台配置信息以及数据仓库等组件的元数据的默认数据库。

### 检查操作系统以前是否安装有MySQL

rpm -qa|grep -i mysql

rpm -qa|grep mariadb

如果有显示内容，说明已安装过MySQL相关组件，需卸载。

卸载相关的MySQL组件：

yum -y remove XXX 或者rpm -e --nodeps XXX

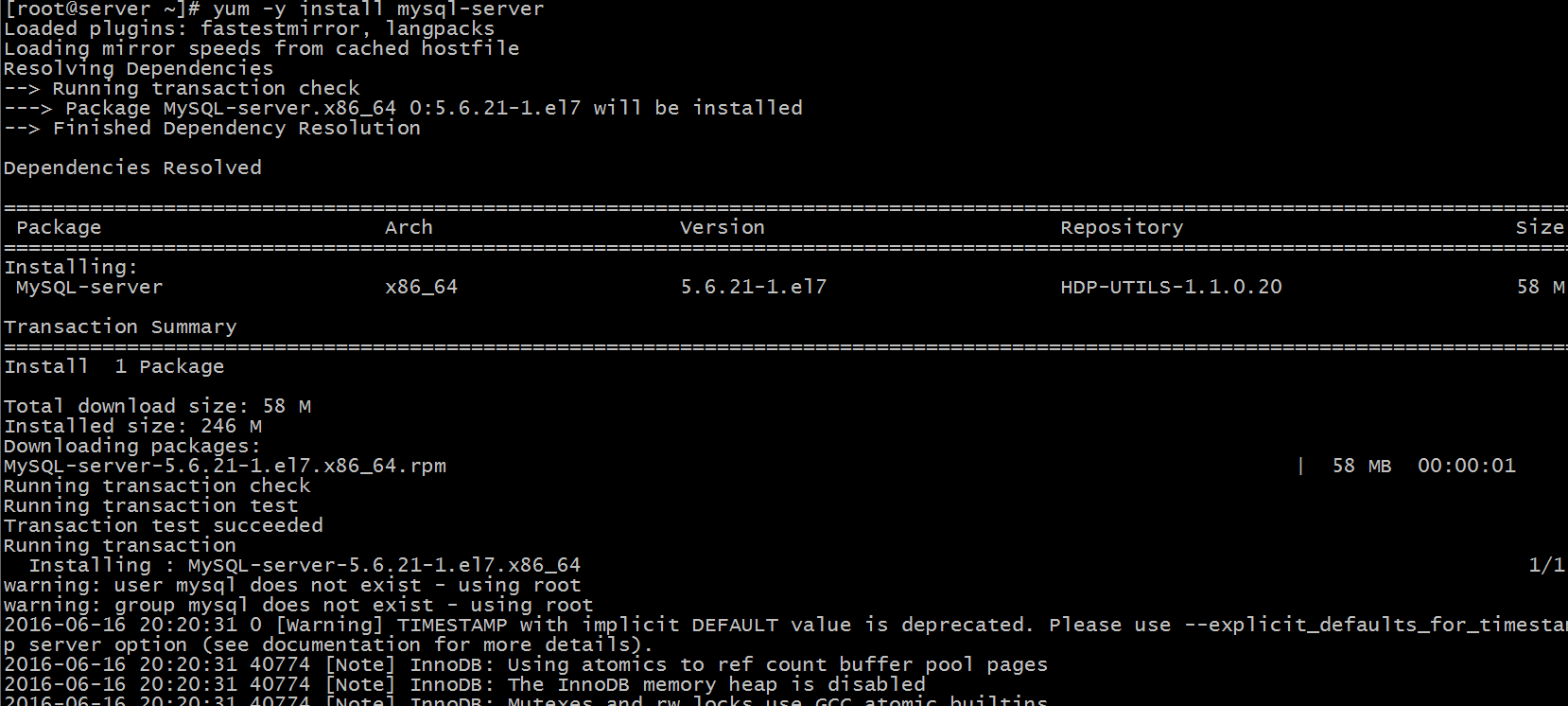
### 安装MySQL

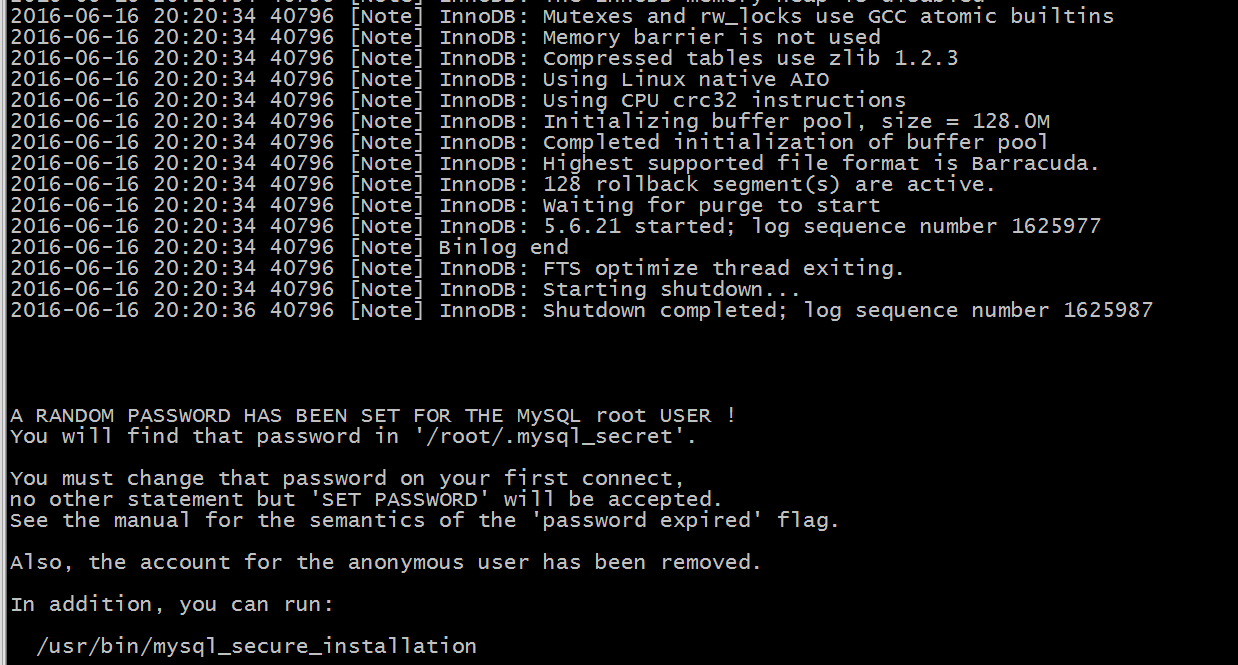
安装依赖包：

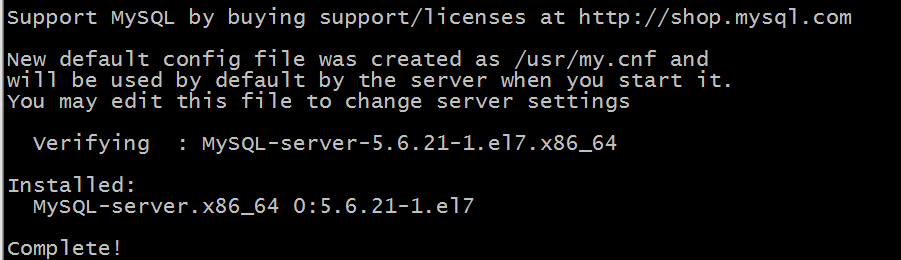
yum install perl-Data-Dumper -y

安装MySQL：

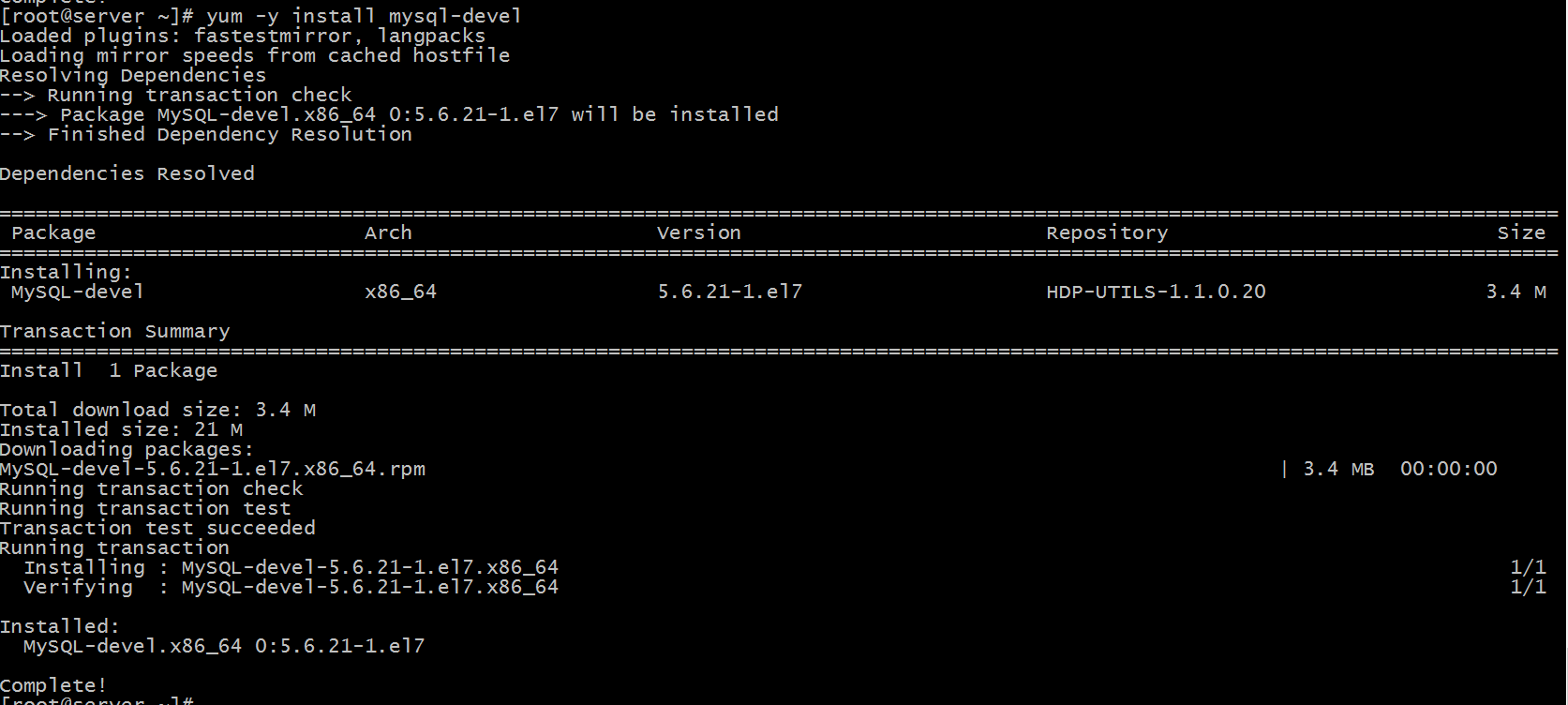
yum -y install mysql-server



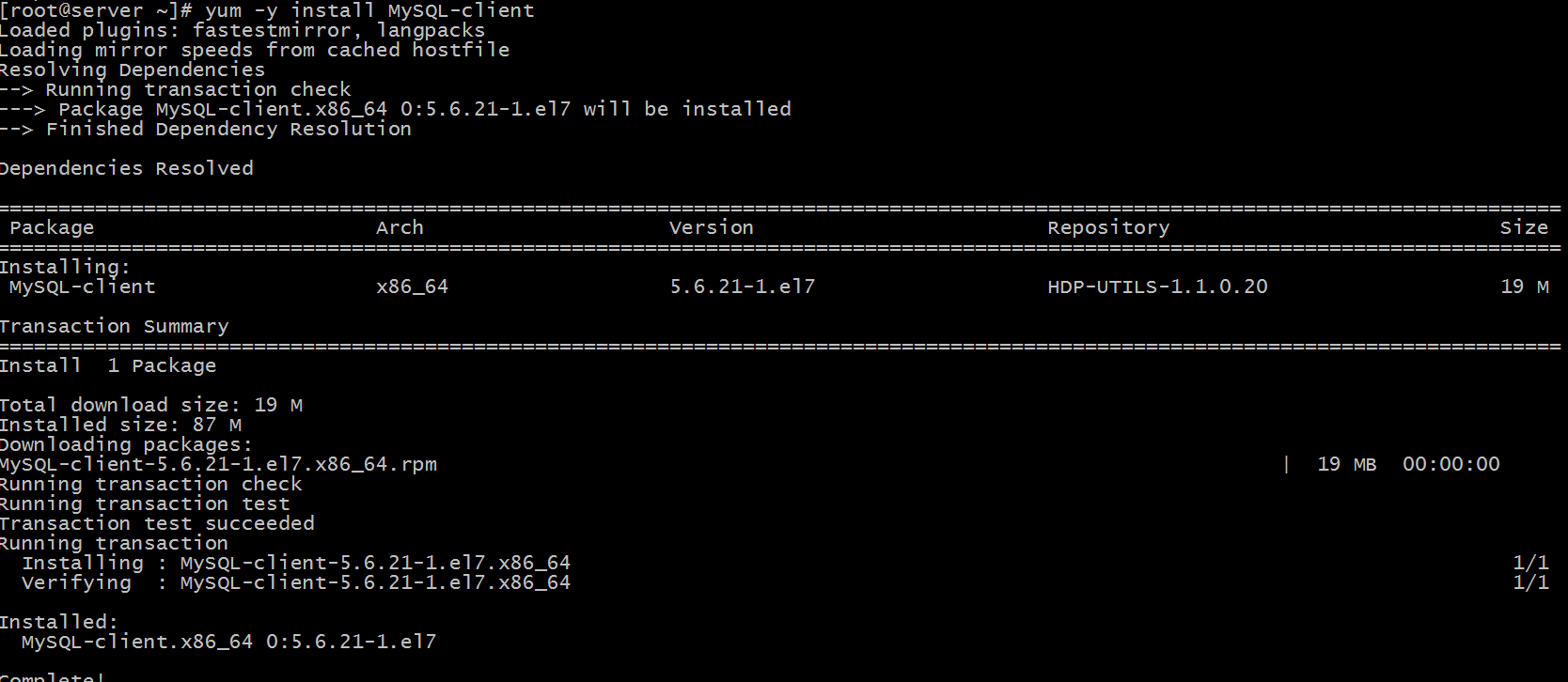




> yum -y install mysql-devel

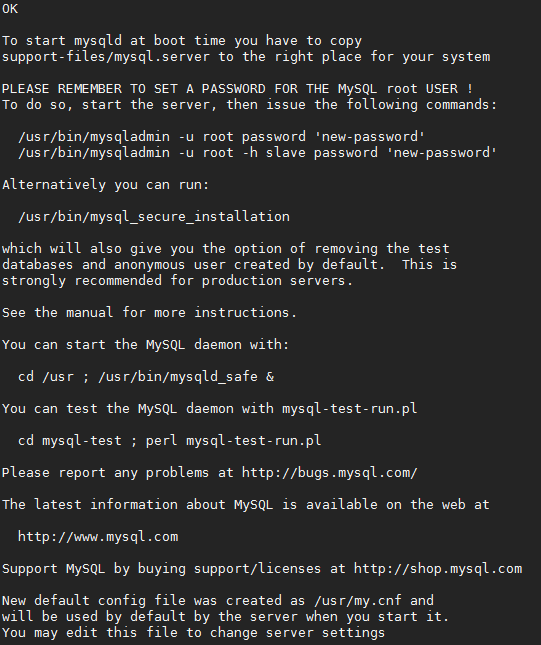


> yum -y install MySQL-client



初始化数据库：

> /usr/bin/mysql\_install\_db



修改配置文件位置及其内容：

>vi /usr/my.cnf

####################################################

#添加编码配置

character-set-server=utf8

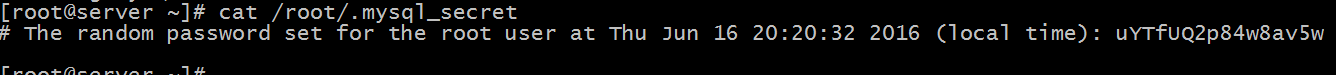
开启服务：

> service mysql start



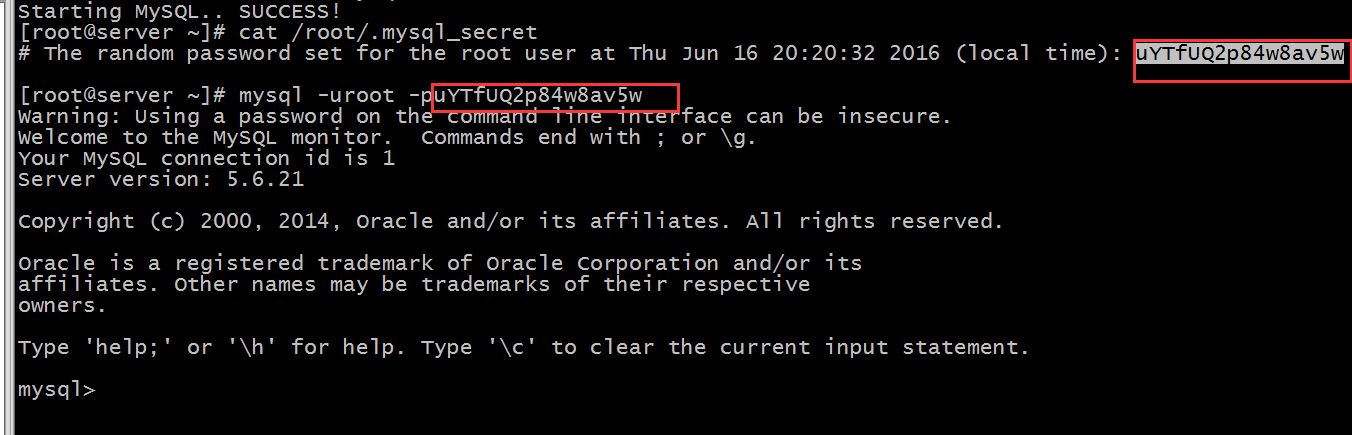
查看初始密码：

> cat /root/.mysql\_secret

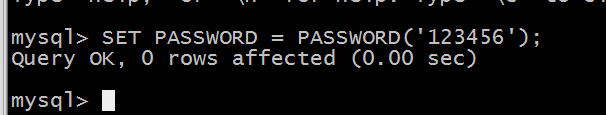


登录MySQL为root用户配置新密码：

mysql -uroot -p uYTfUQ2p84w8av5w



> SET PASSWORD = PASSWORD('123456');



### 创建数据库并赋权denied-

mysql -uroot -p123456

create database ambari;

create database hive;

create database oozie;

创建hive，oozie用户，并赋予权限

grant all privileges on ambari.\* to 'ambari'@'mine'identified by 'ambari';

grant all privileges on hive.\* to 'hive'@'mine'identified by 'hive';

grant all privileges on oozie.\* to 'oozie'@'mine'identified by 'oozie';

flush privileges;

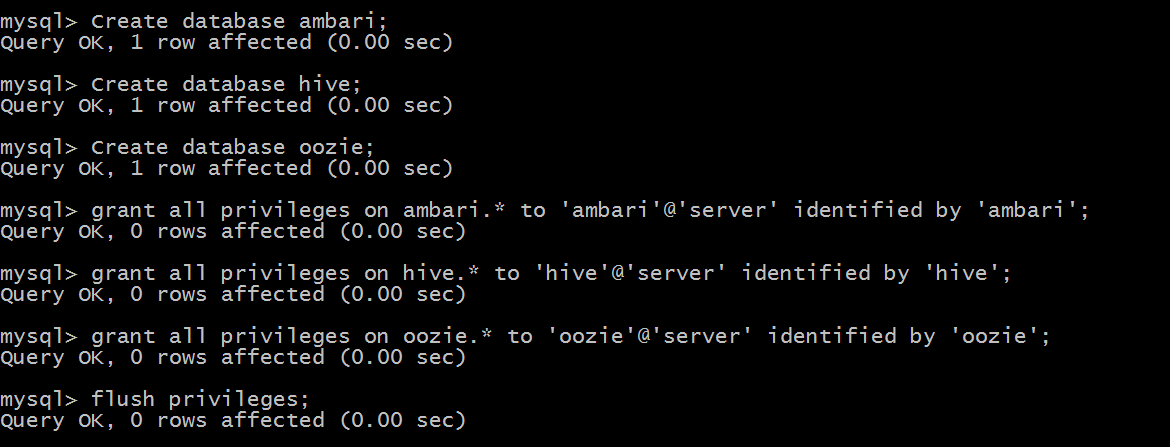
或者执行（开放集群每台机器的mysql访问权限）

grant all privileges on ambari.\* to 'ambari'@'%' identified by 'ambari';

grant all privileges on hive.\* to 'hive'@'%' identified by 'hive';

grant all privileges on oozie.\* to 'oozie'@'%' identified by 'oozie';

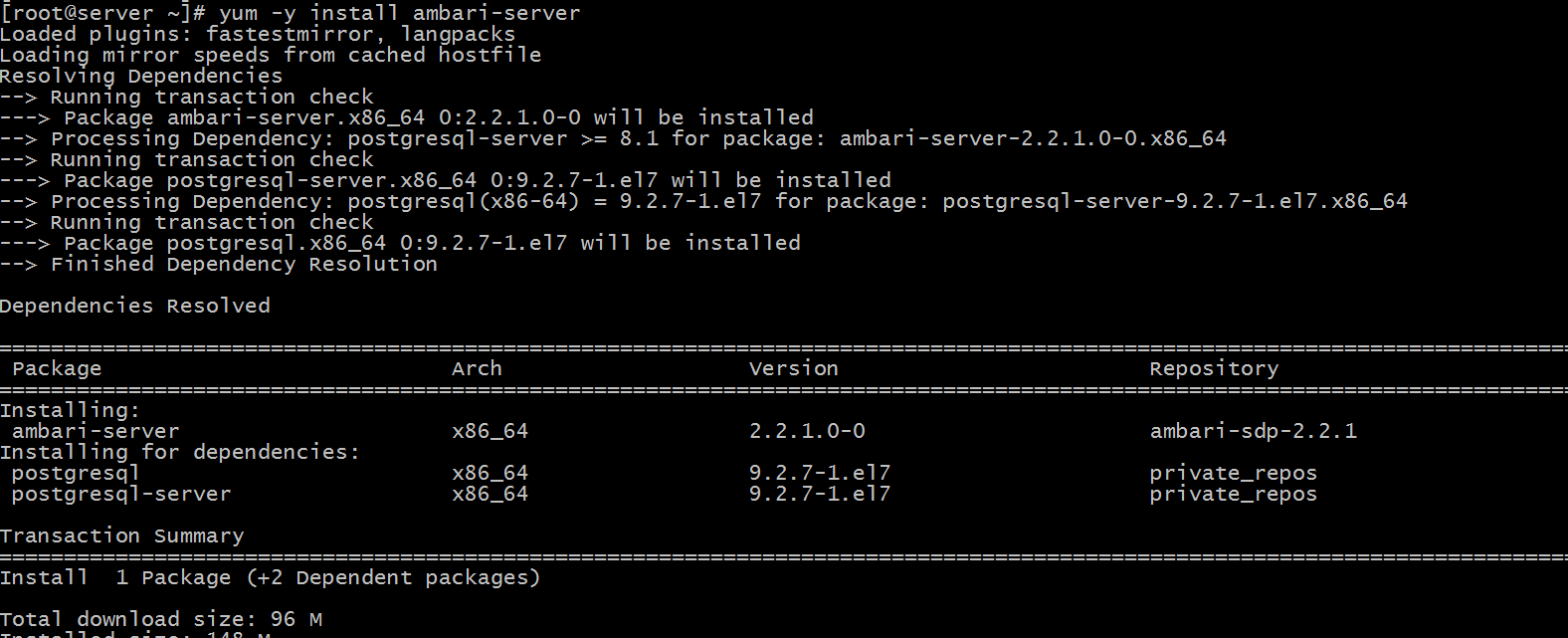
flush privileges;



## 安装集群管理节点

### 通过yum源安装监控管理系统

> yum -y install ambari-server



### 创建一个MySQL的JDBC目录

安装完以后，将MySQL的JDBC jar包复制到ambari-server的默认目录/usr/lib/ambari-server和/usr/share/java下。

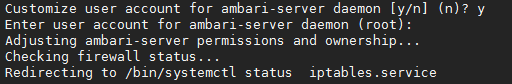
**补充**：连接jar的svn路径:

https://10.1.3.186:440/svn/安装包/传统数据库安装包/Mysql/release-5.6.27-el7/mysql-connector-java-5.1.38-bin.jar

### 配置集群管理节点

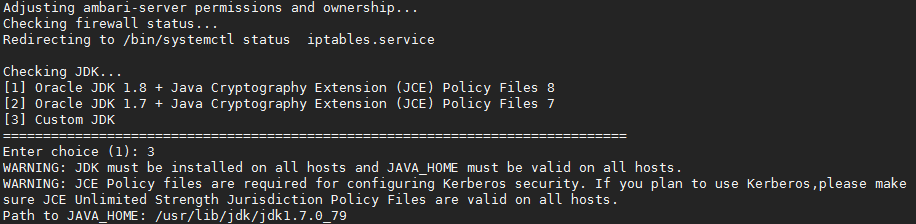
>ambari-server setup

选择ambari-server 运行用户：直接回车或输入root

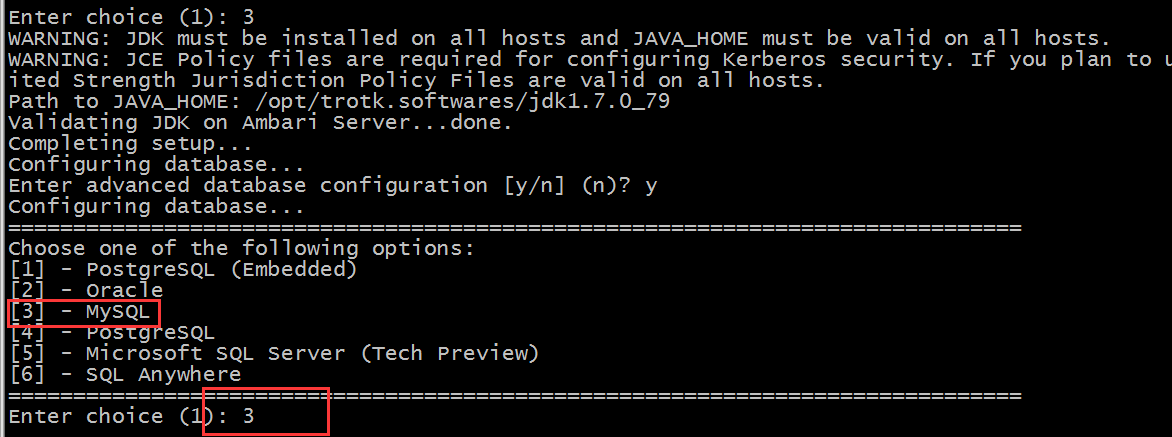


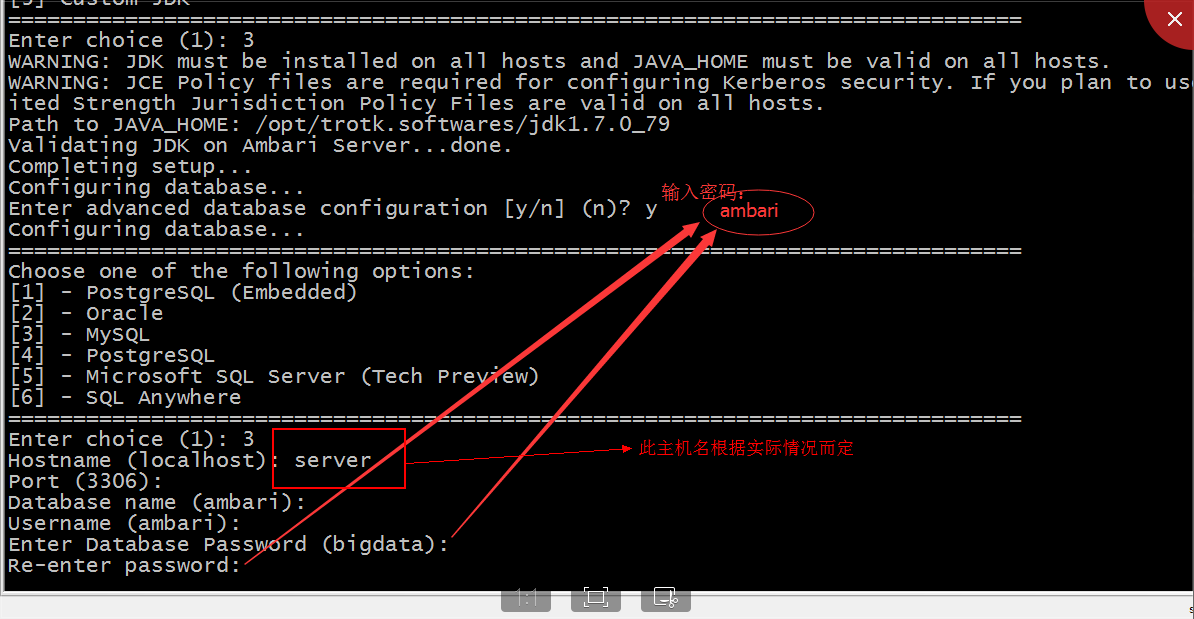
选择jdk：选择3 自定义jdk

输入jdk的目录：/usr/lib/jdk/jdk1.7.0\_79

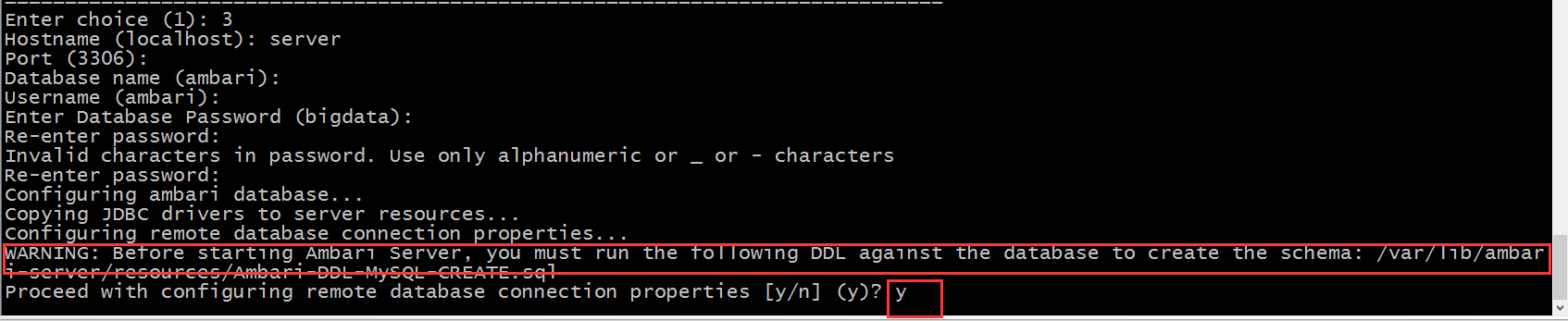


配置数据库：输入y，并选择3 MySQL作为数据库





输入y完成配置



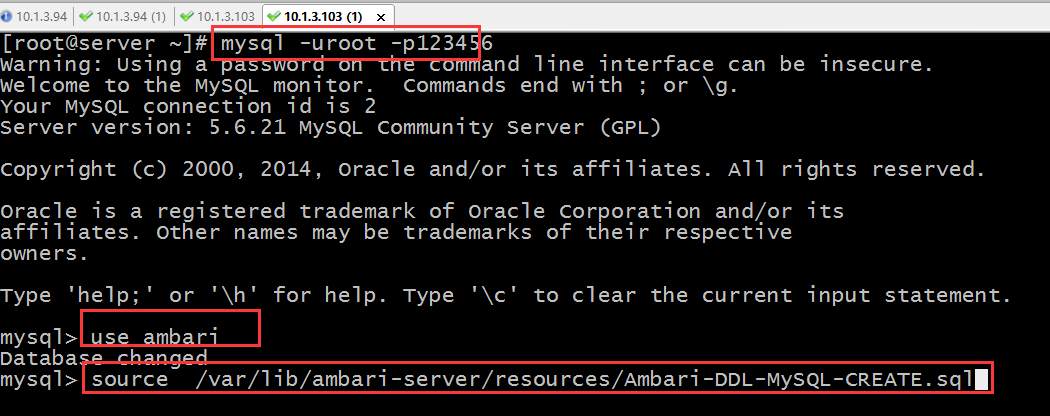
重新开启一个新的窗口：进入数据库

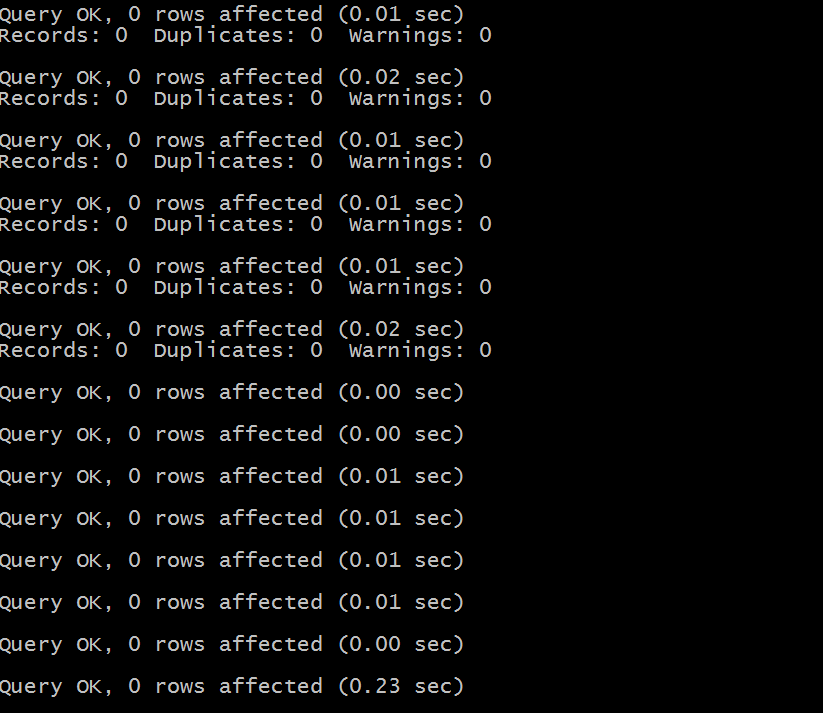
mysql -uroot -p123456

登录MySQL数据库执行初始化SQL脚本：

> use ambari;

>source /var/lib/ambari-server/resources/Ambari-DDL-MySQL-CREATE.sql



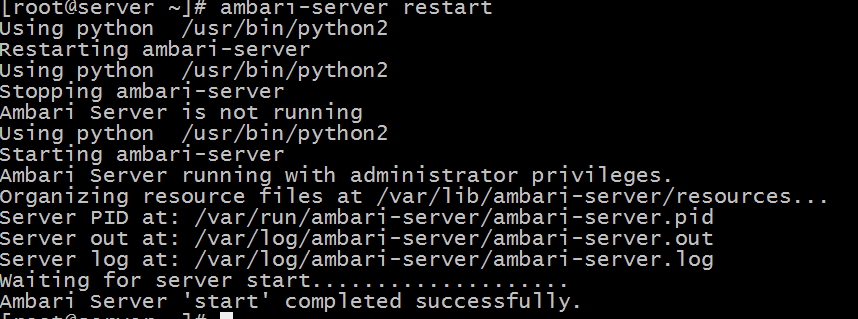


执行成功后退出数据库，关闭该窗口。

启动监控管理系统：

>ambari-server start #启动ambari-server

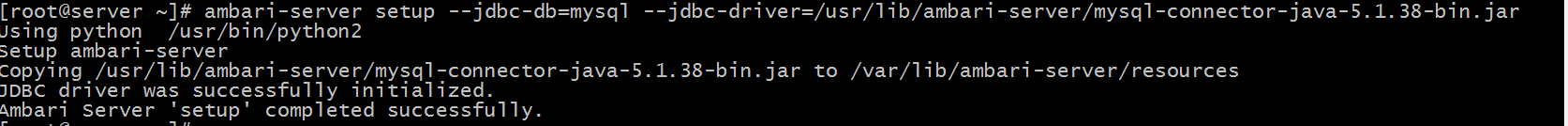
>ambari-server restart #重启ambari-server



### 配置数据仓库组件所需的MySQL的JDBC包

根据JDBC包所在目录，执行：（注意版本信息，版本可更改）

ambari-serversetup --jdbc-db=mysql --jdbc-driver=/usr/lib/ambari-server/mysql-connector-java-5.1.38-bin.jar



## 界面化一键部署集群

### 6.0.安装smcs-web

（1）下载smcs-web安装包

下载地址：[https://10.1.3.186:440/svn//安装包/SMCS/当前可用安装包/《最新版本》/front](https://10.1.3.186:440/svn/安装包/SMCS/当前可用安装包/《最新版本》/front)

（2）安装有两种方法

第一种是：将smcs-web-<版本号>-release.noarch.rpm放入系统中，直接执行

rpm -ivh smcs-web-<版本号>-release.noarch.rpm

第二种是将此安装包放入yum源中，使用下面指令进行安装：

yum -y install smcs-web

安装后修改配置文件进行修改请求server 的IP以及端口号，如下图

cd /usr/SDP/SMCS-WEB/conf

vi smcs-web.properties



配置完后启动smc-web

启动：smcs-web start

停止：smcs-web stop

重启：smcs-web restart

查看运行状态：smcs-web status

卸载：rpm -e smcs-web

1. 访问页面

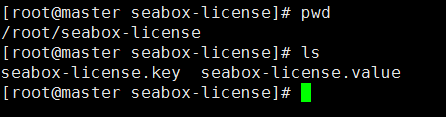
安装成功后

打开浏览器http://<集群管理服务器IP>:9999，admin/admin 登录访问

### 6.1.添加license

将管理员给的seabox-license.key、seabox-license.value文件存在ambari-server的/root/seabox-license文件夹下

（备注：license获取路径：https://10.1.3.186:440/svn//研发项目/海盒大数据平台安装包/Seabox-license）

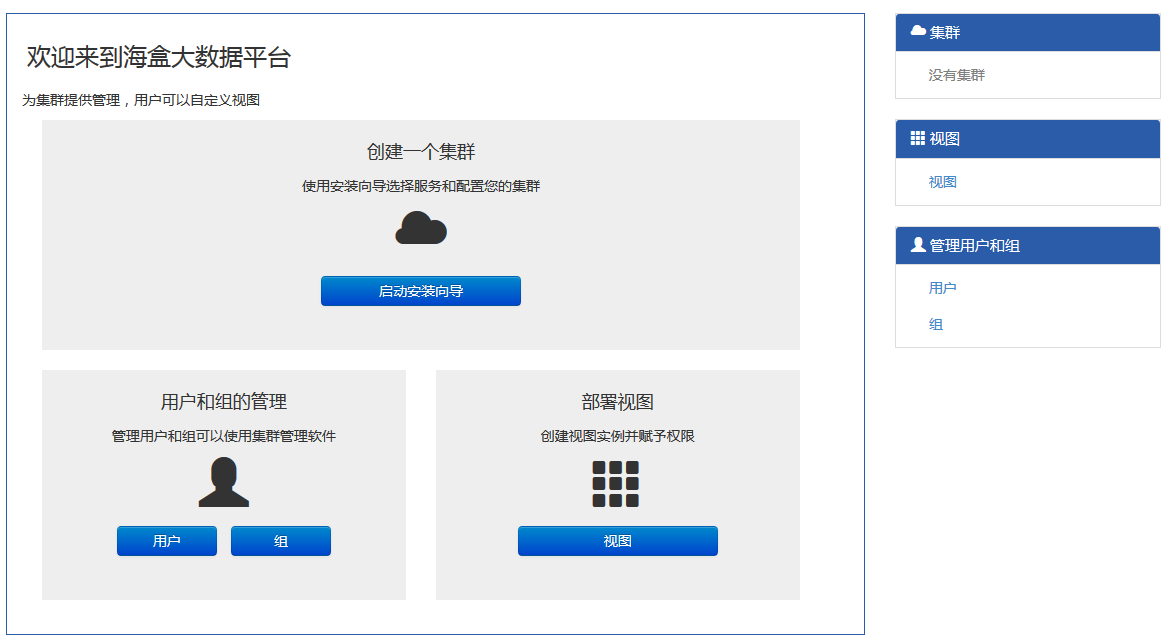


然后重新访问http://<集群管理服务器IP>:9999

### 6.2.新建集群

~~监控管理系统启动后，使用浏览器打开：http://<集群管理服务器IP>:8080。使用用户名和密码admin/admin登录进入安装管理界面。~~

~~启动安装向导：~~



新建集群建议采用英文字母+数字的命名方式，如seabox1



### 6.3.版本选择

选择需要版本，并根据操作系统勾选资源仓库，若有本地库，输入本地库IP：



**根据自己的虚拟机版本选择对应的操作系统：**

**若是Centos7则更新reahat7的资源库地址，**

**若是Contos6则更新redhat6的资源库地址，**

**HDP对应的是本地资源库的SDP-XXX，HDP-UTIL对应的是本地资源库的SDP-UTIL-XXX。**

### 6.4.选择节点

根据集群配置，输入不同服务器的主机名，每个主机名都另起一行，并将主机的私钥（随便添加一台，以添加master节点为优）添加到注册选项中：



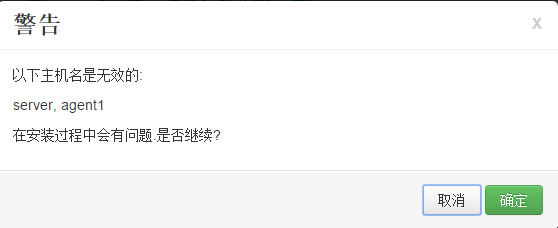
补充：

**授权码：第一次安装可任意填写**

**注册主机：您的主机名**

**密钥：本机的/root/.ssh下的id\_rsa文件，拖到本地即可**

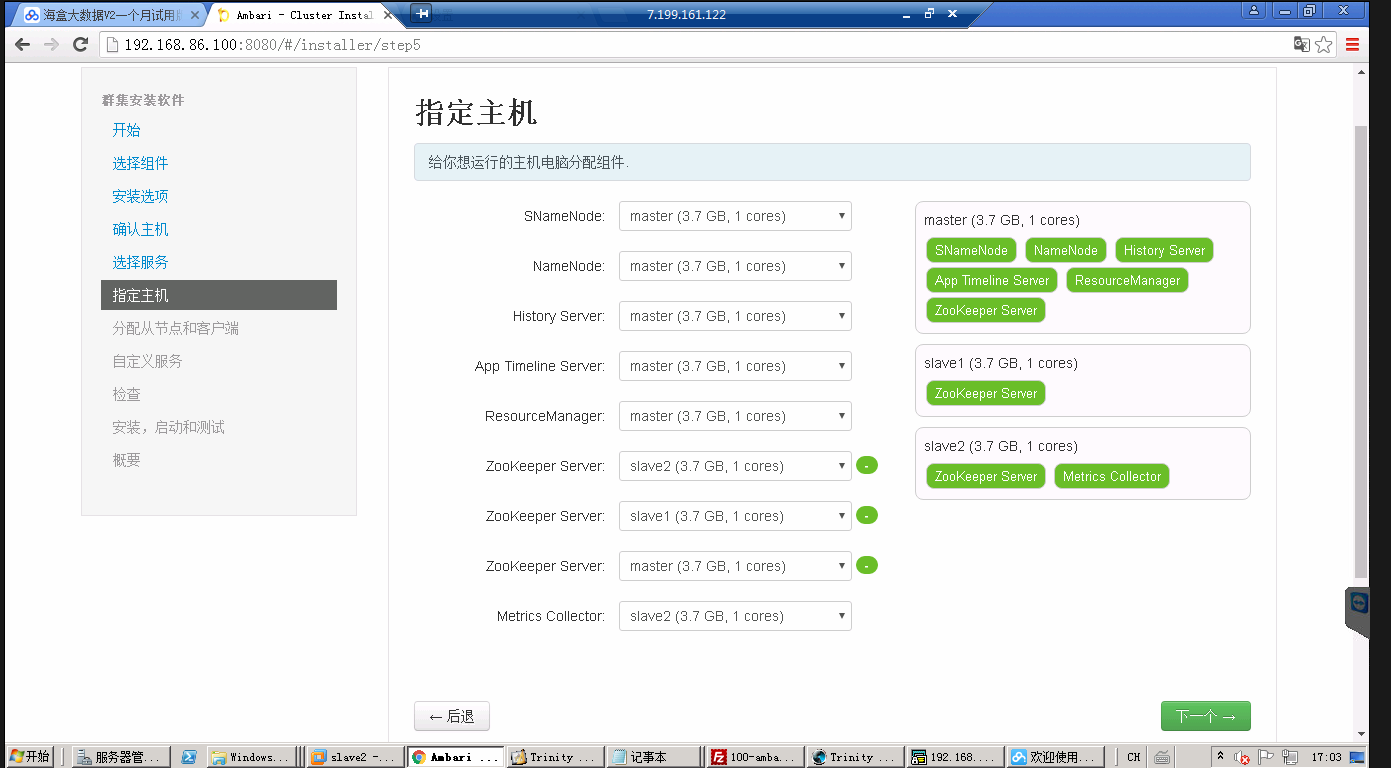
确认时可能出现警告，可以忽略，点击确认



### 6.5.确认主机

等待确认各个节点中的注册日志：







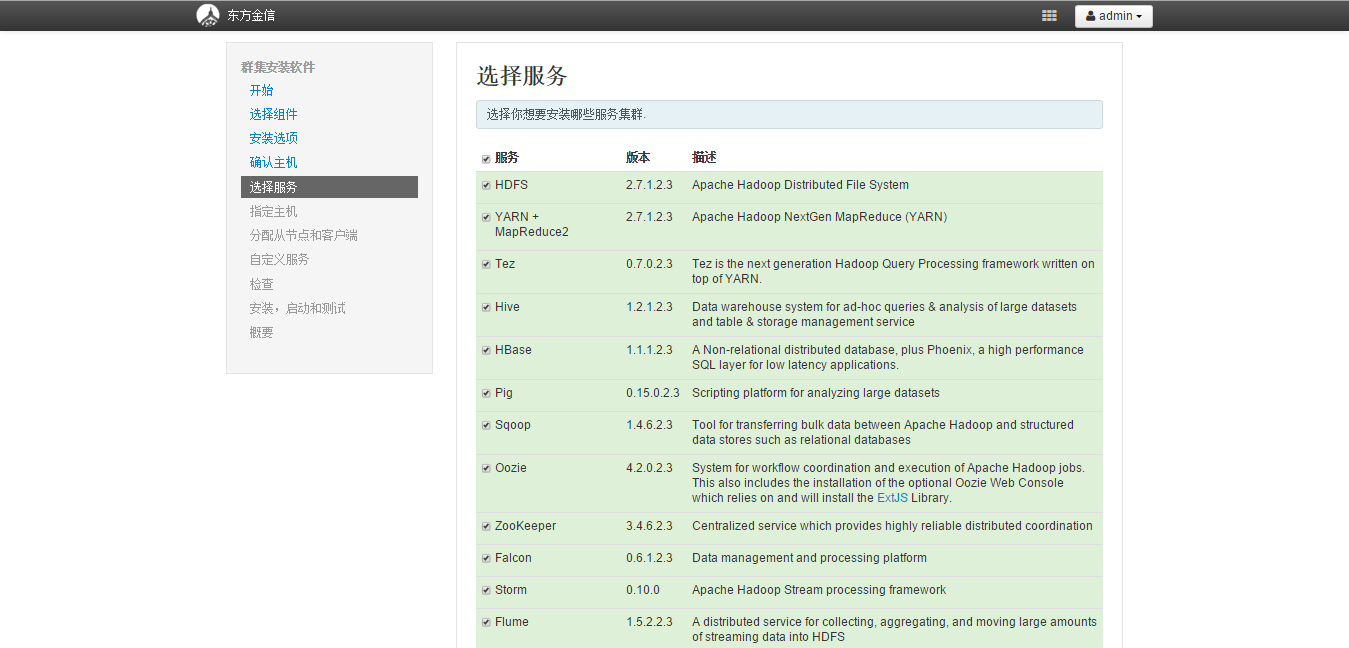
### 6.6.选择组件

根据需要，勾选所需组件：

**首次安装：请选择基础组件，切记勿全选！**

基础组件：hdfs+ yarn+ mapreduce+ zookeeper+ ambari metrics

也可以选择 hive 及其依赖组件



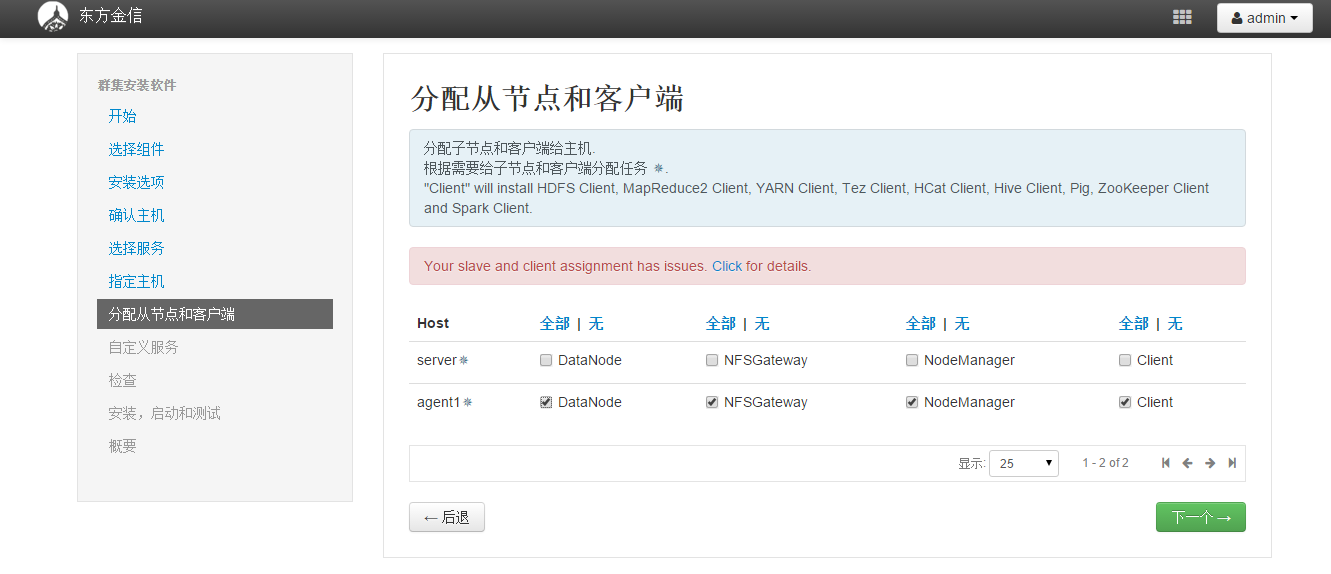
**主节点配置：**

指定各个组件主节点运行的服务器：



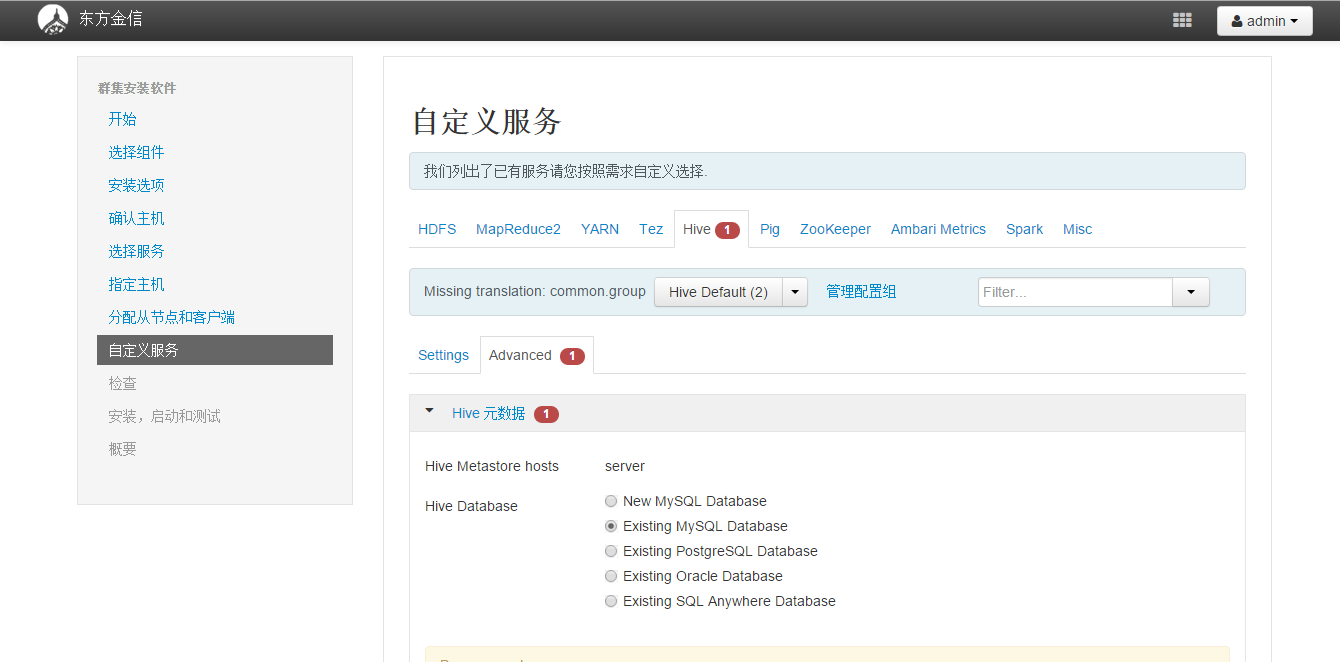
**从节点及客户端配置：**

指定各个组件从节点服务和客户端运行的服务器：

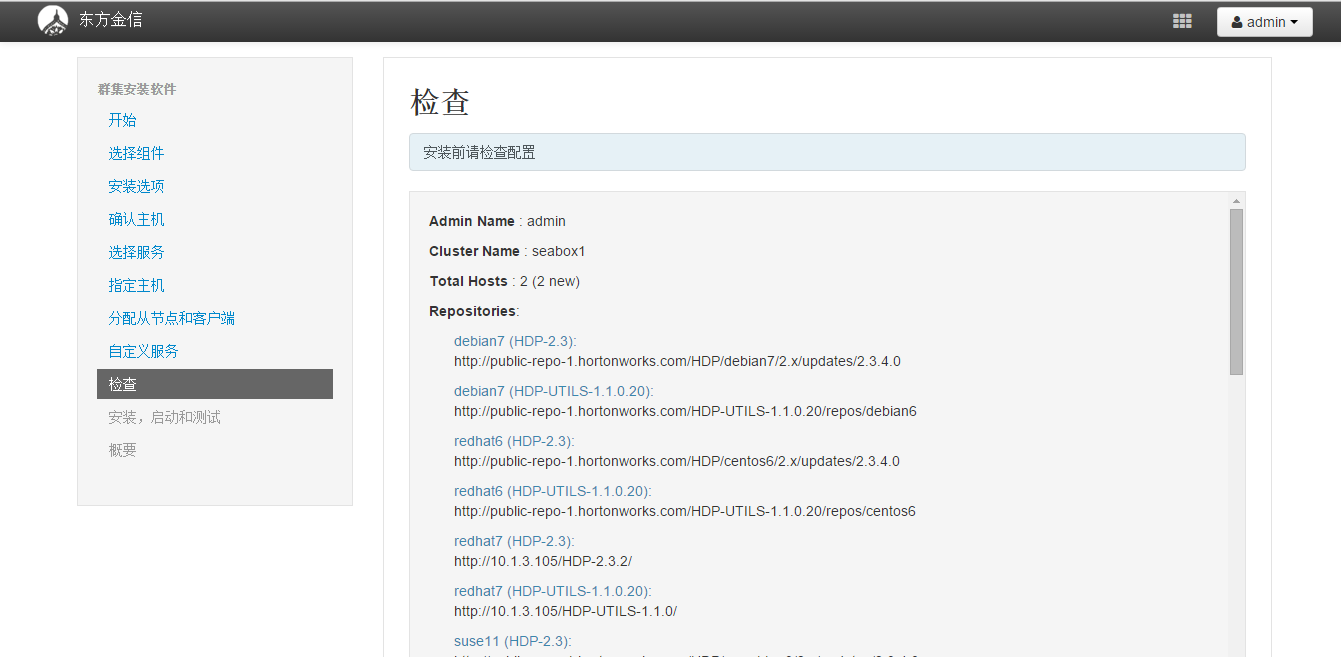


**组件元数据配置：**

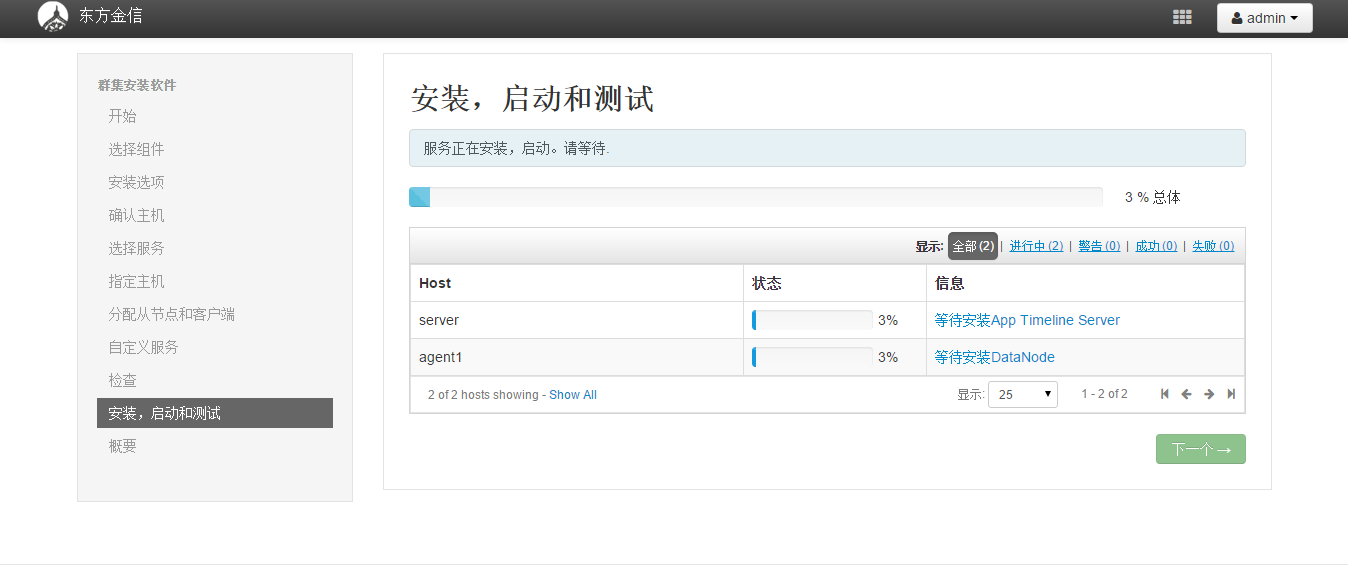
配置数据仓库等组件的元数据存放的数据库：



**检查配置：**



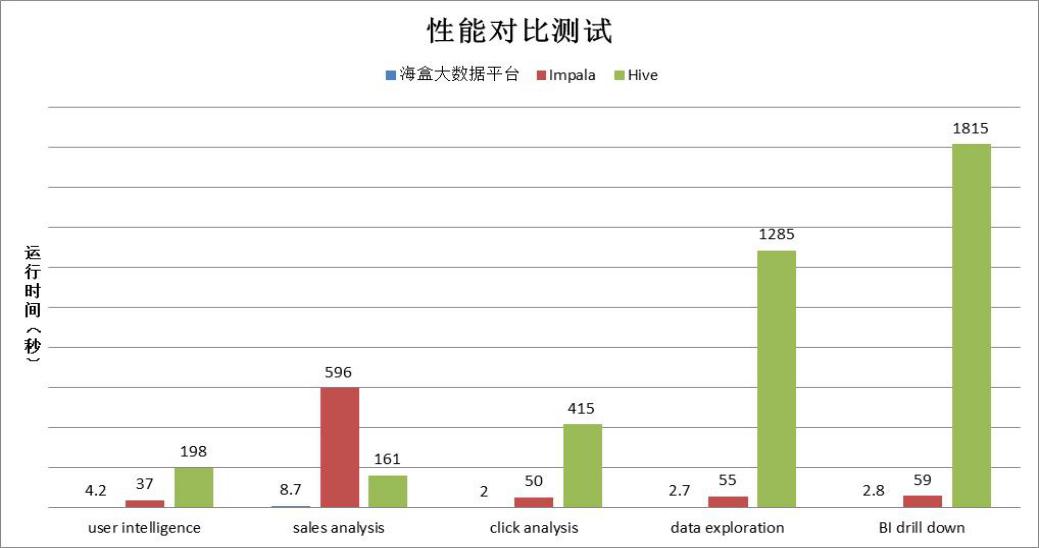
### 6.7.启动安装部署



## 安装MPP大数据引擎

软件特色

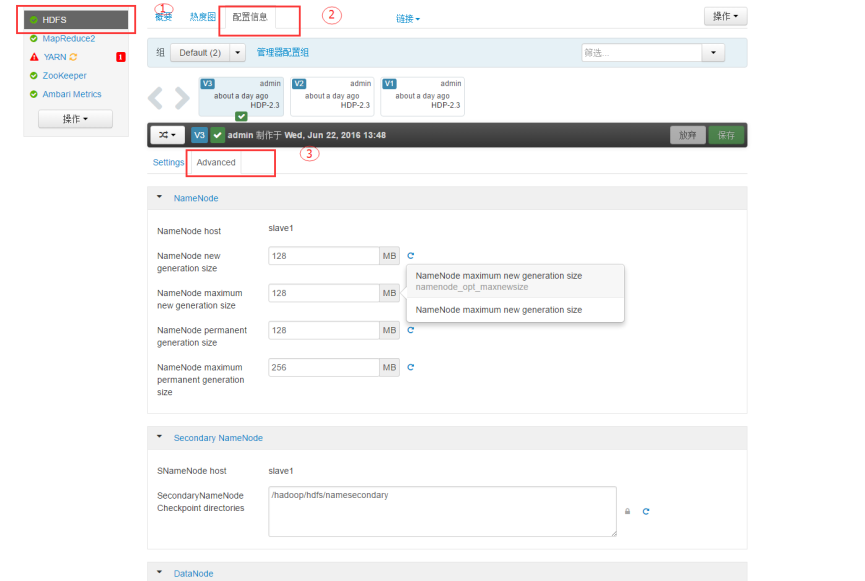
* 全面支持TPC-DS SQL99标准，支持事务完整性
* SQL客户端使用pgADMIN进行连接测试测试，pdADMIN可以在Windows和Linux下运行
* 具有MPP（大规模并行处理系统）的性能，比其他Hadoop里面的SQL引擎快10倍以上。以下为在60节点的hadoop集群中与同类产品的性能对比测试。数据量十亿，时间单为秒



### 修改配置信息

1. 需要在界面配置hawq对hdfs要求的参数，在界面上配置。

安装之前，首先应该在海盒大数据平台的界面修改选择HDFS功能，切换到配置信息选项卡,选在高级选项配置卡（Advanced）:



选择自定义hdfs-site.xml和自定义core-site.xml选项：并点击下面的添加属性：

Add Property

分别添加以下属性：

core-site.xml

ipc.client.connect.timeout=300000

ipc.server.listen.queue.size=3300

hdfs-site.xml

dfs.allow.truncate=true

dfs.block.local-path-access.user=gpadmin

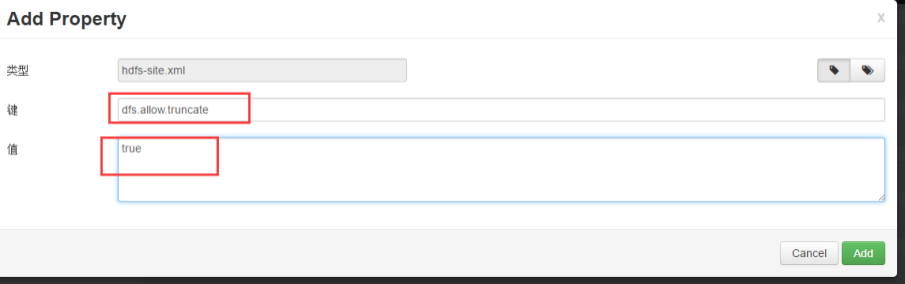
dfs.client.socket-timeout=300000000

dfs.client.use.legacy.blockreader.local=false

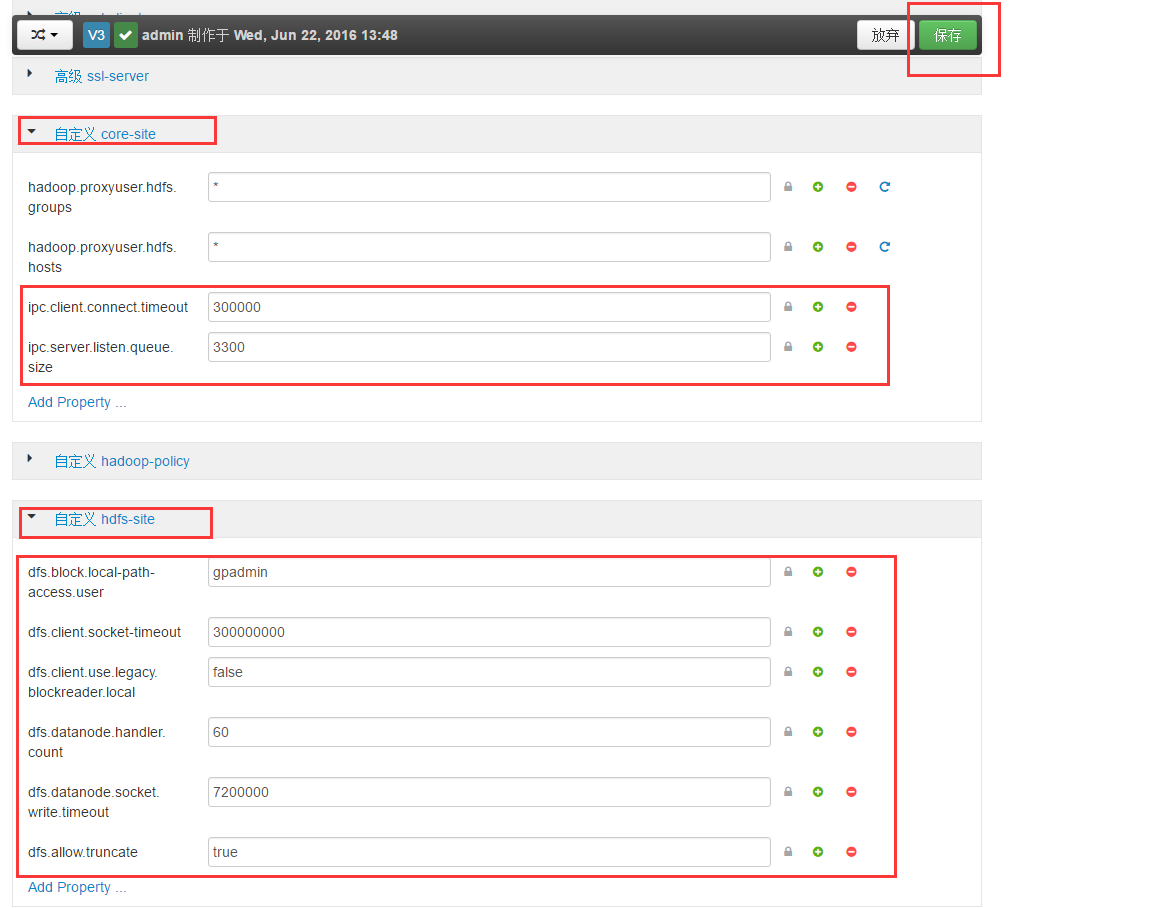
dfs.datanode.handler.count=60

dfs.datanode.socket.write.timeout=7200000

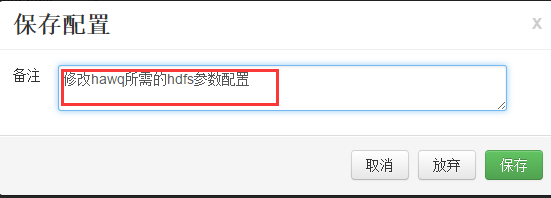
例如：



添加完成后，点击保存：



会弹出一个备注的框，在此记录一下本次修改的记录。



保存即可。

1. 配置完成以后，安装在ambari-server服务器上安装hawq插件，安装命令如下：

>yum -y install hawq-plugin

安装完以后，重启所有ambari-server及其所有的组件。

>ambari-server restart

注意以下限制和要求：

1） pxf必须装在HDFS NameNode以及全部的HDFS DataNodeS上；

2） 若用HBASE 和HIVE与pxf, 首先必须装HBASE 客户端然后HIVE客户端到各个分机。

在https://10.1.3.186:440/svn/研发项目/07 IBM与东方金信大数据测试/最终产品软件资源/sdp-yum-repos/Seabox大数据平台/ext\_rpms目录下拷贝文件：

pxf-service-3.0.0-21030.noarch.rpm

在集群的每台HDFS服务器(nameNode和dataNode)上操作：需要手动安装apache-tomcat和pxf-service

安装命令：

> yum -y install apache-tomcat

> rpm -ivh --force --nodeps pxf-service-3.0.0-21030.noarch.rpm

其中

1） 每个节点必须满足系统要求；

2） 运行其他postgresql实例的节点不能被用来运行默认的hawq-master.

3） 每个hawq-segment必须联合定位于运行了HDFS DataNode的节点上。

4） hawq-master部分必须运行于一个独立的机器上。

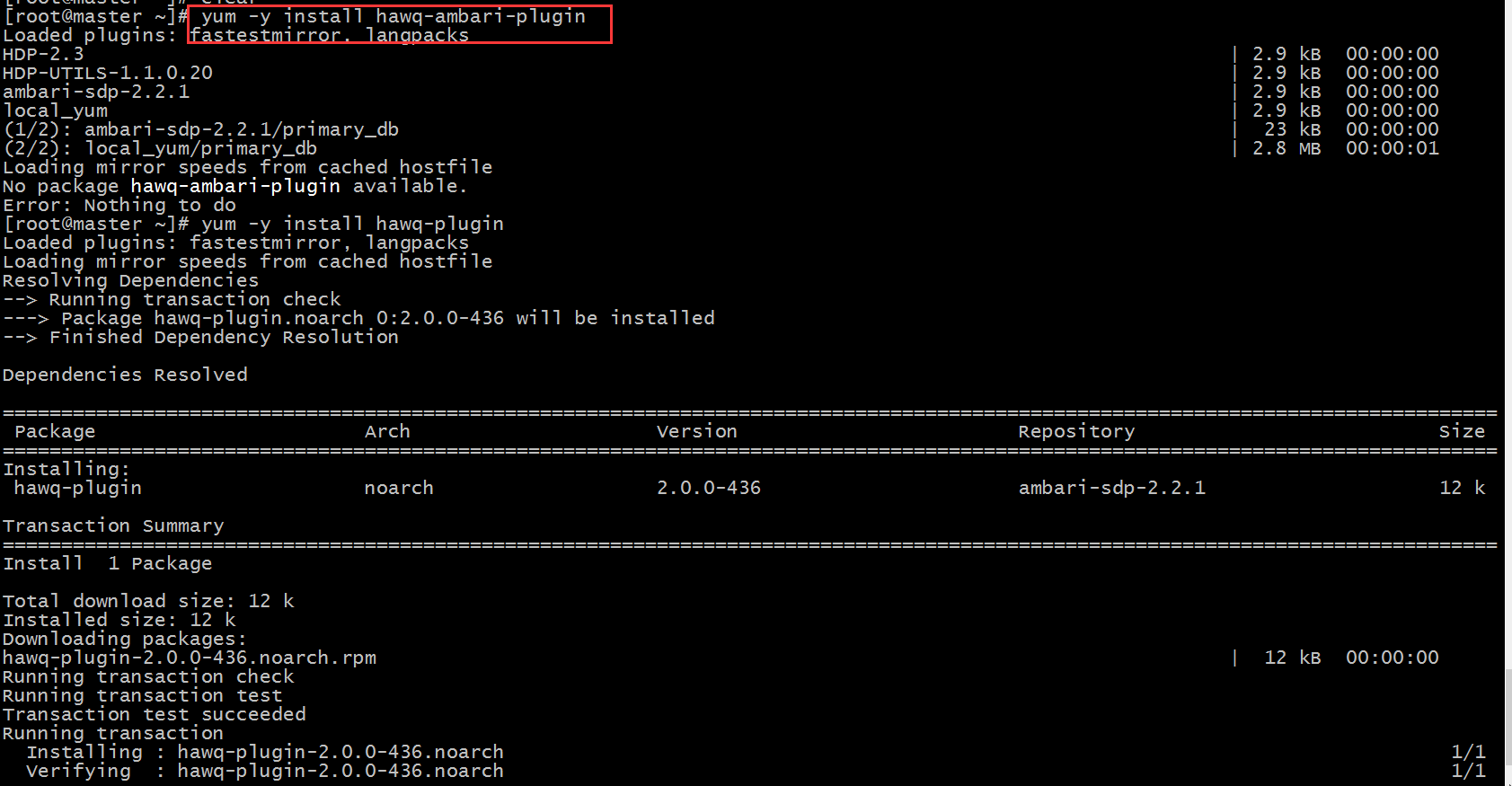
选一个或多个将会被用作储存临时文件的路径。创建当地目录到每个HAWQ节点上，并且监控用户和组的许可于gpadmin.

上面准备完成后就可直接在ambari界面安装hawq和pxf。

### 安装插件

在主机上使用yum安装

>yum -y install hawq-ambari-plugin



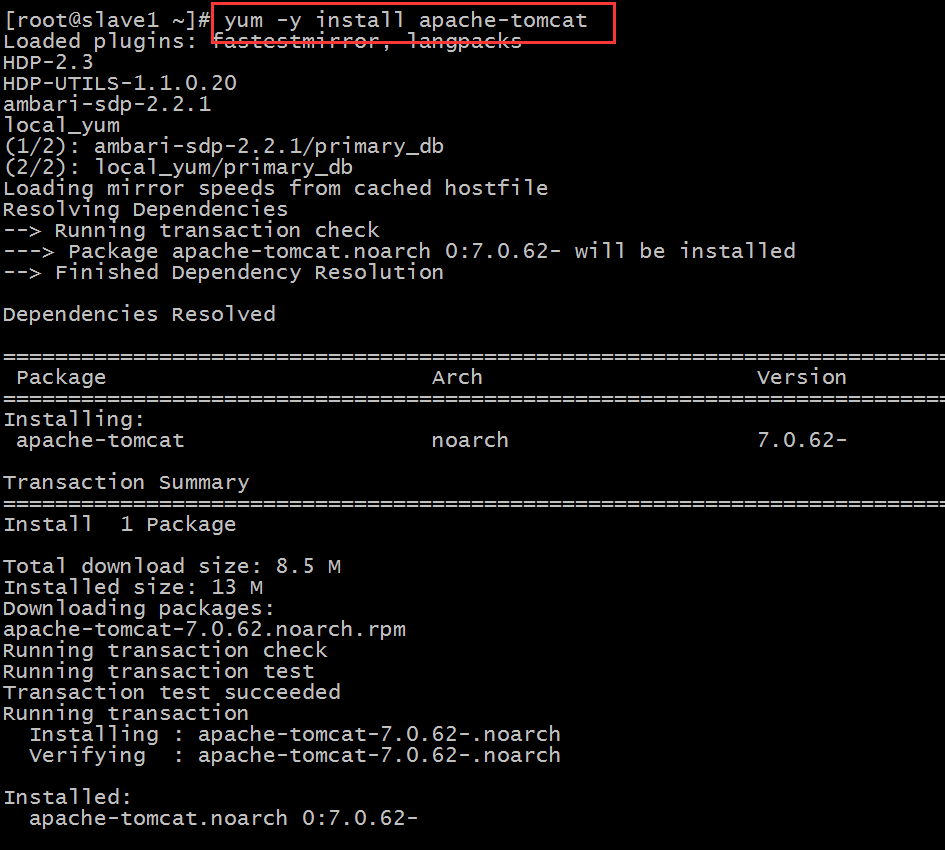
重启ambari-server:

>ambari-server restart

在所有的NameNode和DataNode节点上安装pxf组件及其相关组件:

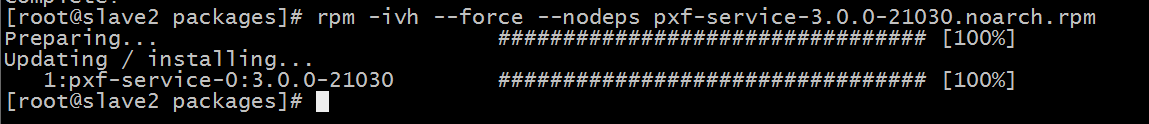
首先需要安装apache-tomcat

>yum -y install apache-tomcat



然后安装pxf-service包：

>rpm -ivh --force --nodeps pxf-service-3.0.0-21030.noarch.rpm

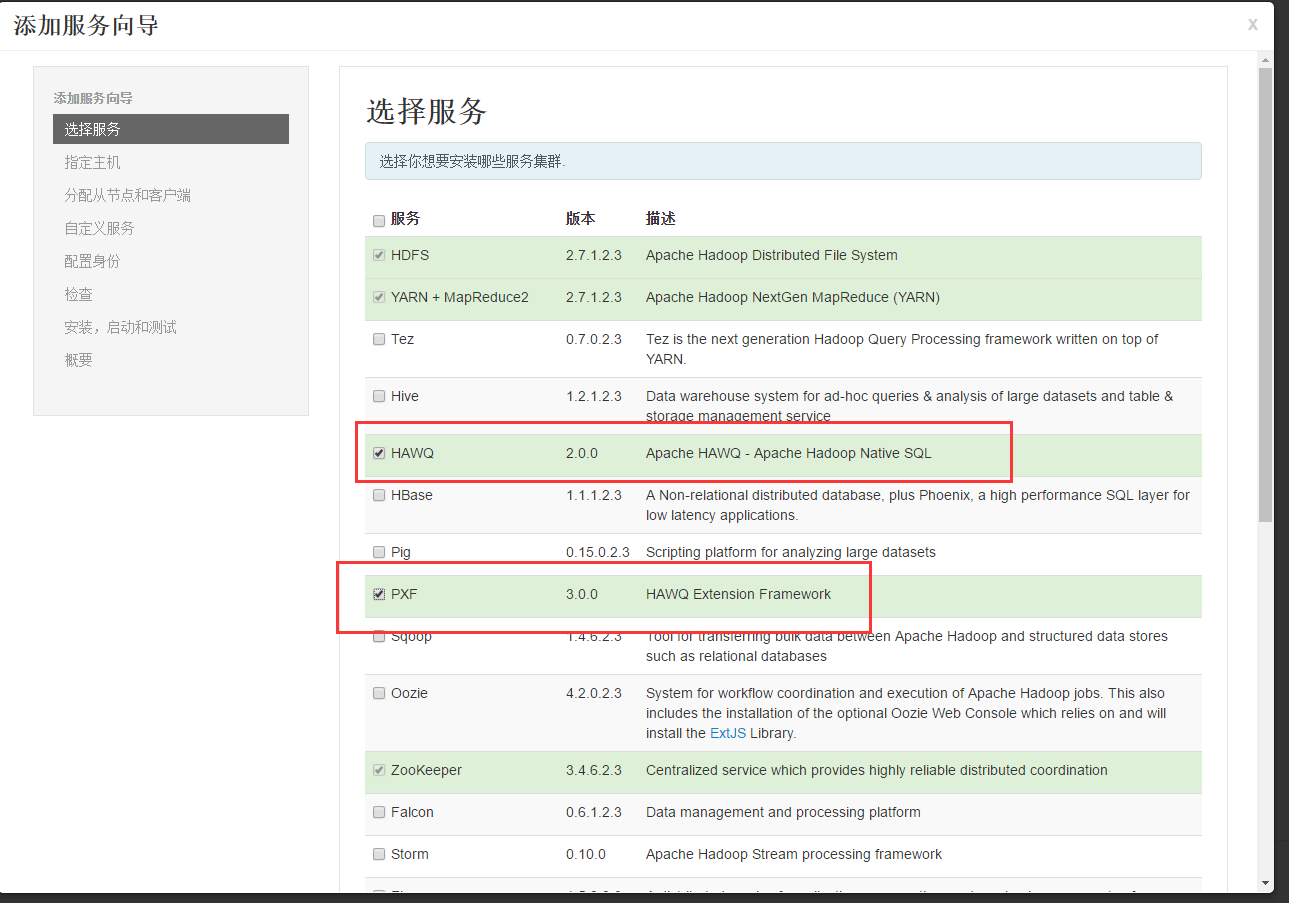


### 重新登录到海盒大数据平台的界面

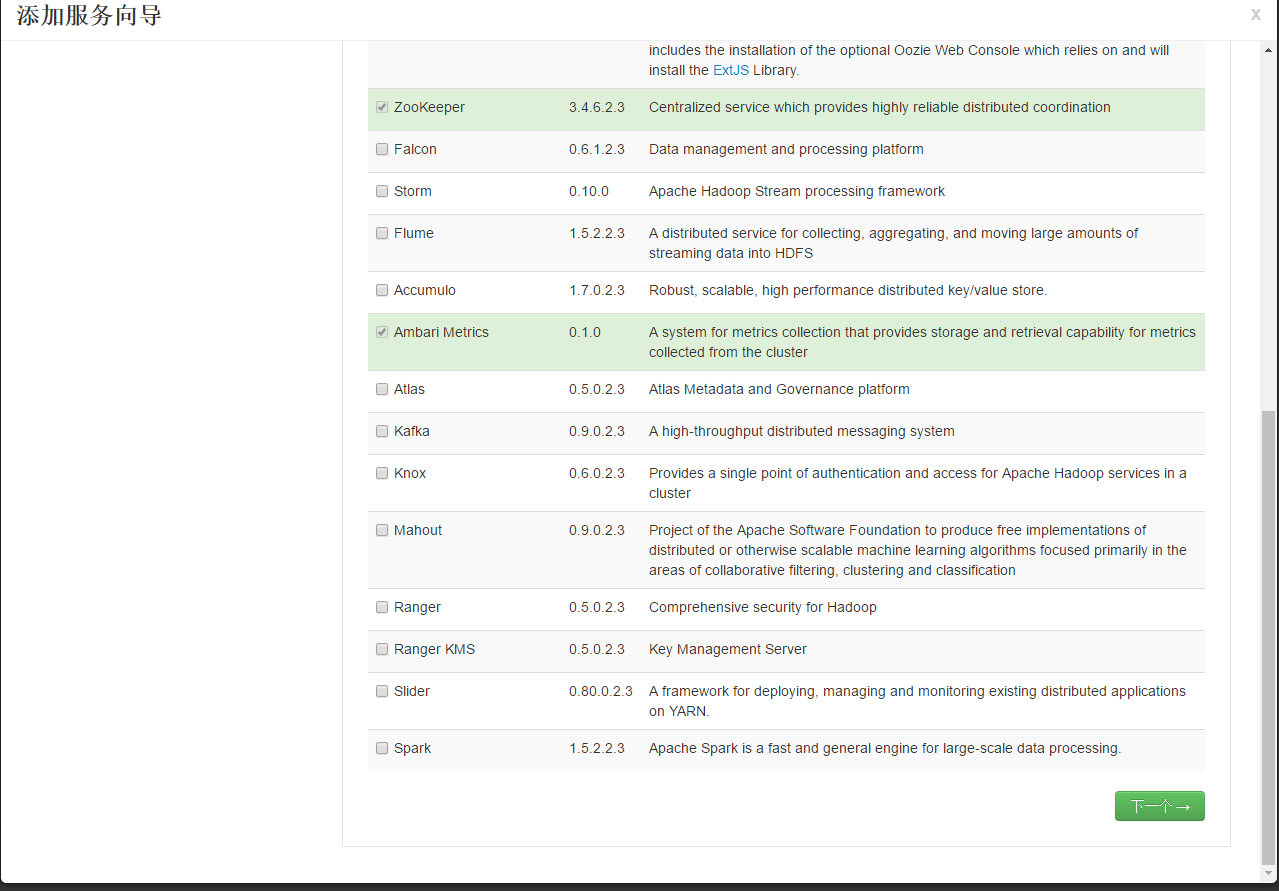
选择操作功能，并选择添加服务：



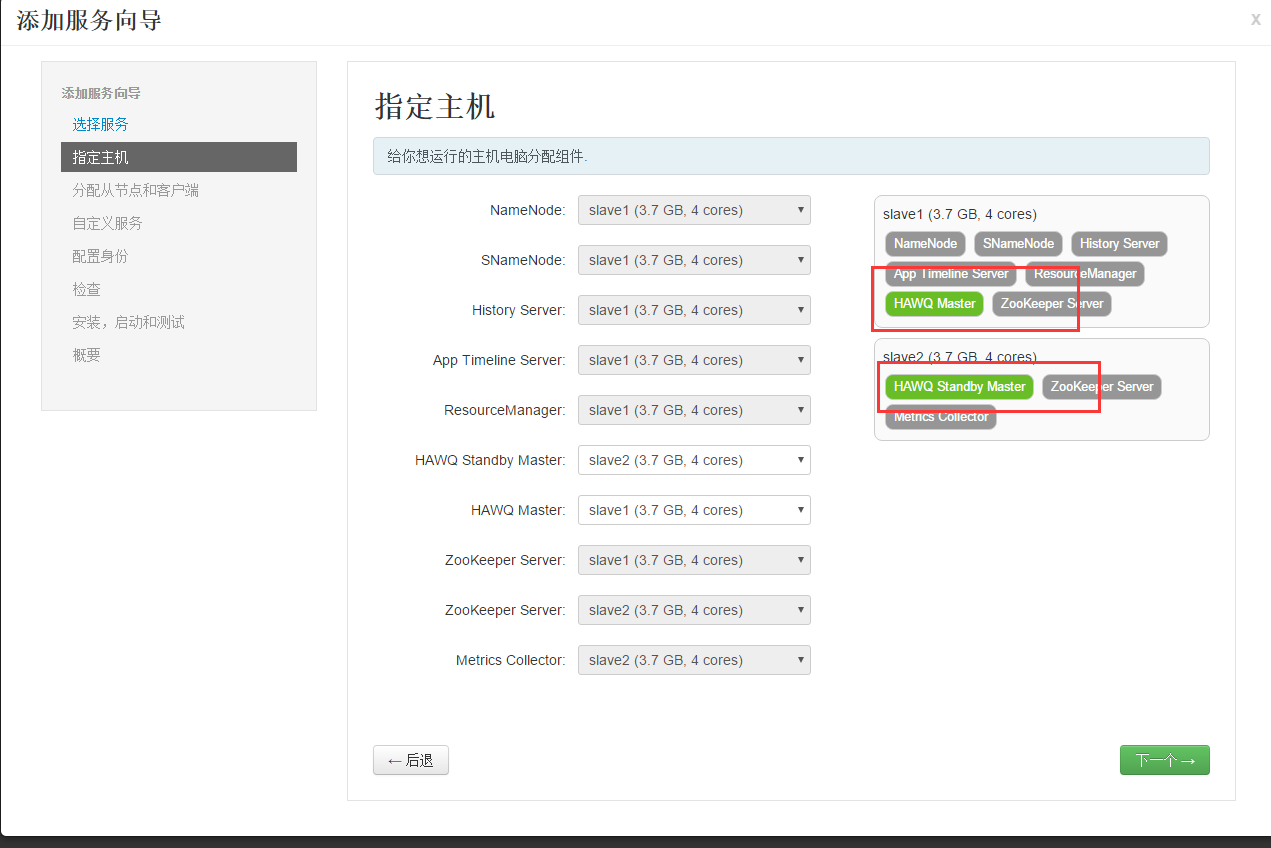
下一步：选择HAWQ和PXF服务：



然后点击下一个：



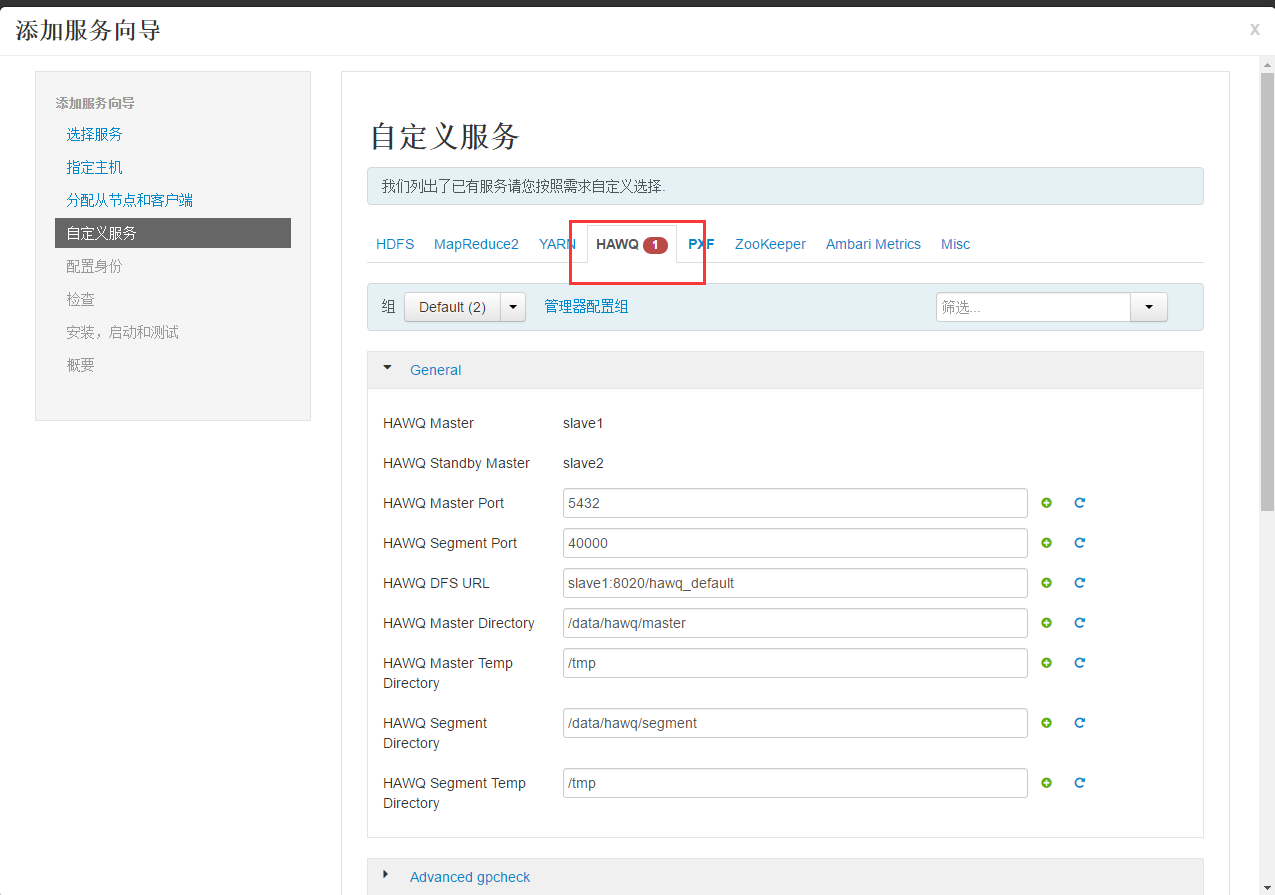
指定主机，将HAWQ安装在NameNode上：



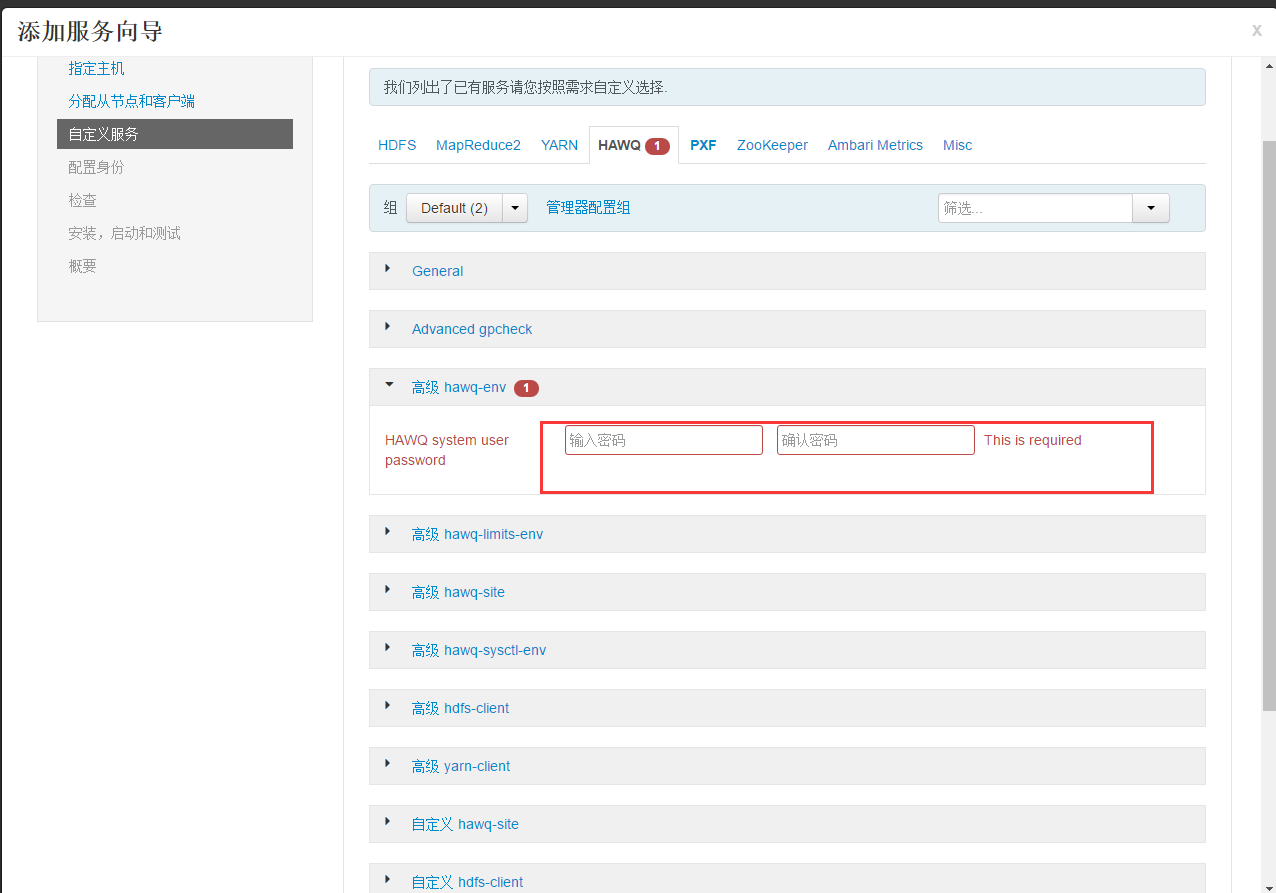
分配从节点和客户端



自定义服务：选择HAWQ选项卡：



设定HAWQ系统管理员的用户密码:

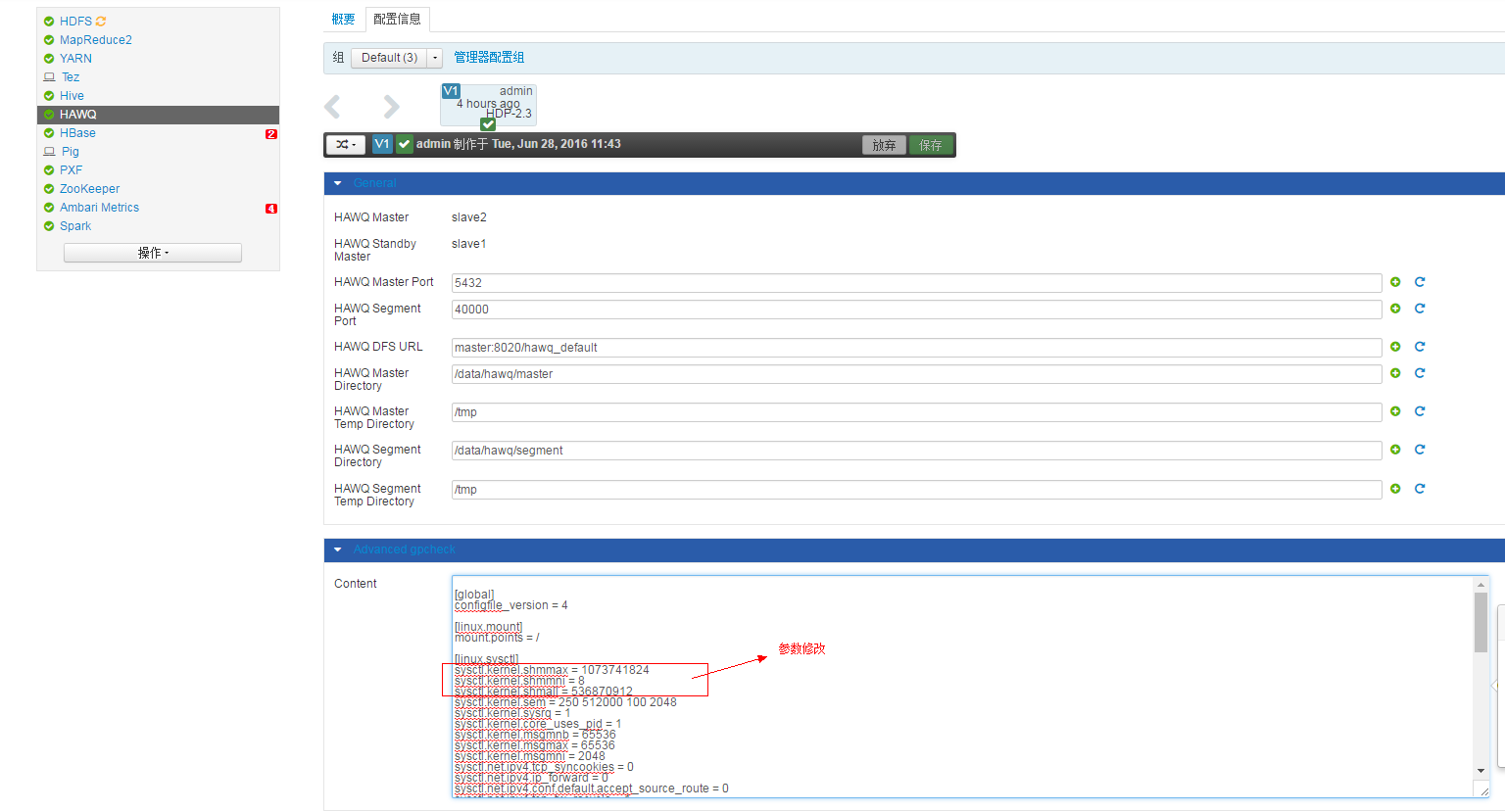


注意：当安装用来hawq的内存>=4g，还需要修改上面的Advancedgpcheck里面的几个参数值。有时可能因为系统内存的原因，无法正常启动，需要在界面上调整参数

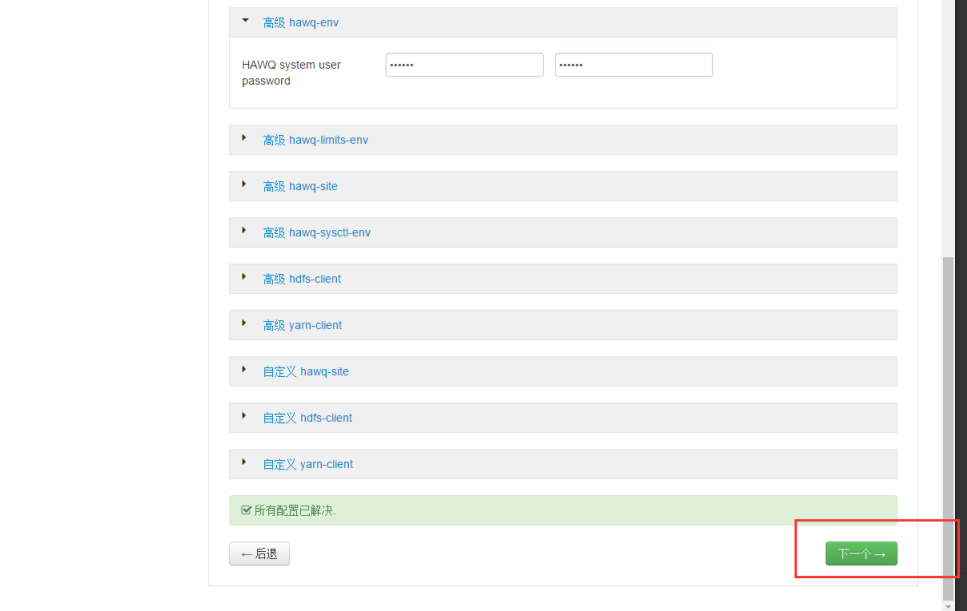
sysctl.kernel.shmmax = 1073741824

sysctl.kernel.shmmni = 8

sysctl.kernel.shmall = 536870912



配置完成后点击下一个：



下一步检查、点击部署，进入安装、启动和测试界面：

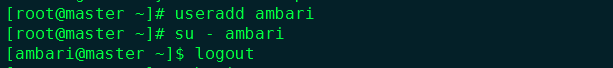


最后都点击默认的下一个，直到安装完成，安装完成后，注意一点，有可能由于内存大小的原因，可能会报警告，不用理会，直接点击完成即可，然后进入abmari界面重启hawq即可。

# 非root用户启动ambari

## 创建用户

Useradd ambari



## 添加ambari 用户添加变量

修改 vim /etc/sudoers 这个文件是只读文件 所以需要强制保存在 root下 添加

ambari ALL=(ALL) NOPASSWD:SETENV: /bin/su hdfs \*, /usr/bin/su hdfs \*, /bin/su ambari-qa \*, /usr/bin/su ambari-qa \*, /bin/su ranger \*, /usr/bin/su ranger \*, /bin/su zookeeper \*, /usr/bin/su zookeeper \*, /bin/su knox \*, /usr/bin/su knox \*, /bin/su falcon \*, /usr/bin/su falcon \*, /bin/su ams \*, /usr/bin/su ams \*, /bin/su flume \*, /usr/bin/su flume \*, /bin/su hbase \*, /usr/bin/su hbase \*, /bin/su spark \*, /usr/bin/su spark \*, /bin/su accumulo \*, /usr/bin/su accumulo \*, /bin/su hive \*, /usr/bin/su hive \*, /bin/su hcat \*, /usr/bin/su hcat \*, /bin/su kafka \*, /usr/bin/su kafka \*, /bin/su mapred \*, /usr/bin/su mapred \*, /bin/su oozie \*, /usr/bin/su oozie \*, /bin/su sqoop \*, /usr/bin/su sqoop \*, /bin/su storm \*, /usr/bin/su storm \*, /bin/su tez \*, /usr/bin/su tez \*, /bin/su atlas \*, /usr/bin/su atlas \*, /bin/su yarn \*, /usr/bin/su yarn \*, /bin/su kms \*, /usr/bin/su kms \*, /bin/su mysql \*, /usr/bin/su mysql \*, /usr/bin/yum, /usr/bin/zypper, /usr/bin/apt-get, /bin/mkdir, /usr/bin/mkdir, /usr/bin/test, /bin/ln, /usr/bin/ln, /bin/chown, /usr/bin/chown, /bin/chmod, /usr/bin/chmod, /bin/chgrp, /usr/bin/chgrp, /usr/sbin/groupadd, /usr/sbin/groupmod, /usr/sbin/useradd, /usr/sbin/usermod, /bin/cp, /usr/bin/cp, /usr/sbin/setenforce, /usr/bin/test, /usr/bin/stat, /bin/mv, /usr/bin/mv, /bin/sed, /usr/bin/sed, /bin/rm, /usr/bin/rm, /bin/kill, /usr/bin/kill, /bin/readlink, /usr/bin/readlink, /usr/bin/pgrep, /bin/cat, /usr/bin/cat, /usr/bin/unzip, /bin/tar, /usr/bin/tar, /usr/bin/tee, /bin/touch, /usr/bin/touch, /usr/bin/hdp-select, /usr/bin/conf-select, /usr/hdp/current/hadoop-client/sbin/hadoop-daemon.sh, /usr/lib/hadoop/bin/hadoop-daemon.sh, /usr/lib/hadoop/sbin/hadoop-daemon.sh, /sbin/chkconfig gmond off, /sbin/chkconfig gmetad off, /etc/init.d/httpd \*, /sbin/service hdp-gmetad start, /sbin/service hdp-gmond start, /usr/sbin/gmond, /usr/sbin/update-rc.d ganglia-monitor \*, /usr/sbin/update-rc.d gmetad \*, /etc/init.d/apache2 \*, /usr/sbin/service hdp-gmond \*, /usr/sbin/service hdp-gmetad \*, /sbin/service mysqld \*, /usr/bin/python2.6 /var/lib/ambari-agent/data/tmp/validateKnoxStatus.py \*, /usr/hdp/current/knox-server/bin/knoxcli.sh \*, /usr/hdp/\*/ranger-usersync/setup.sh, /usr/bin/ranger-usersync-stop, /usr/bin/ranger-usersync-start, /usr/hdp/\*/ranger-admin/setup.sh \*, /usr/hdp/\*/ranger-knox-plugin/disable-knox-plugin.sh \*, /usr/hdp/\*/ranger-storm-plugin/disable-storm-plugin.sh \*, /usr/hdp/\*/ranger-hbase-plugin/disable-hbase-plugin.sh \*, /usr/hdp/\*/ranger-hdfs-plugin/disable-hdfs-plugin.sh \*, /usr/hdp/current/ranger-admin/ranger\_credential\_helper.py, /usr/hdp/current/ranger-kms/ranger\_credential\_helper.py

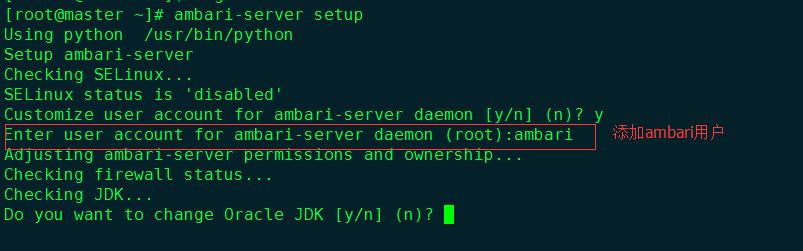
Defaults exempt\_group = ambari

Defaults !env\_reset,env\_delete-=PATH

Defaults: ambari !requiretty

输入后 sudo -i 测试一下是否溢出

## 配置ambari

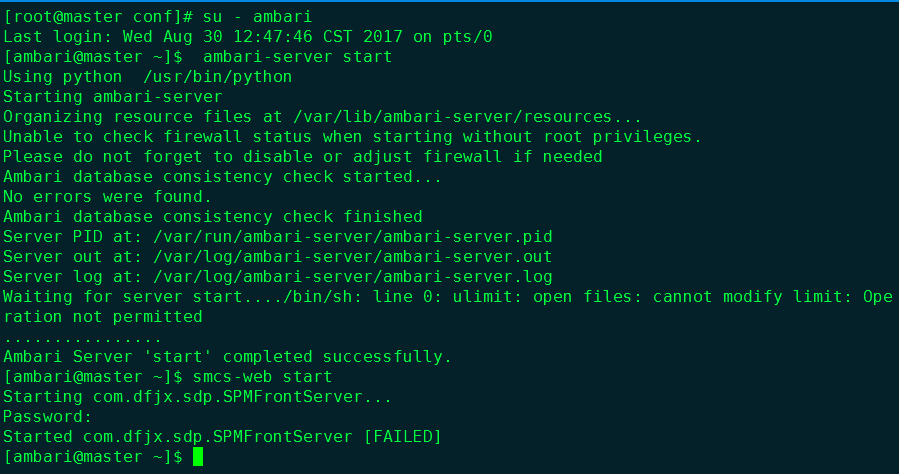


## 修改ambari-agent.ini

修改# vim /etc/ambari-agent/conf/ambari-agent.ini 下的

run\_as\_user=root // 更改root 为 ambari

Su 到ambari 里面 启动 ambari



# 备注

## 防火墙操作(centos7.x centos6.x)

centos7.X操作

systemctl start firewalld.service#启动firewall

systemctl stop firewalld.service#停止firewall

systemctl disable firewalld.service#禁止firewall开机启动

firewall-cmd --state#查看防火墙状态(关闭后显示notrunning，开启后显示running)

centos6.X操作

service iptables start#启动firewall

service iptables stop#停止firewall

chkconfigiptables off#禁止firewall开机启动

service iptables status#查看防火墙状态(关闭后显示notrunning，开启后显示running)

## 安装注意事项

* 如果你的机器之前安装过Hadoop的相关服务，特别是Hbase 里面配置了HBASE\_HOME 的环境变量，就需要删除掉。这个环境变量会影响到组件的安装的路径。
* 注意时间不同服务器的同步问题，以防RegionServer启动报错。

## 组件启动报错解决方案

TEZ配置属性：tez.history.logging.service.class

更改为：org.apache.tez.dag.history.logging.ats.ATSHistoryLoggingService

解决问题：WARNING: Use "yarn jar" to launch YARN applications

YARN配置属性：yarn.timeline-service.store-class

更改为org.apache.hadoop.yarn.server.timeline.LeveldbTimelineStore

解决问题：Yarn app timeline server不能正常启动

## Pxf启动报错解决方案

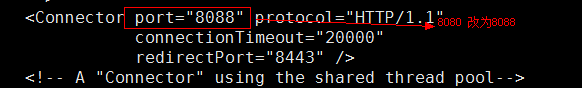
错误：Tomcat is running

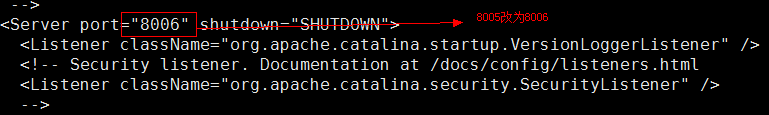
原因：tomcat端口8080被占用

解决方法：

8080端口改为8088

8005端口改为8006





## 补充：安装mysql5.7操作。

### 检查操作系统以前是否安装有MySQL

rpm -qa|grep -i mysql

rpm -qa|grep -i mariadb

如果有显示内容，说明已安装过MySQL相关组件，需卸载。

卸载相关的MySQL组件：

yum -y remove XXX 或者rpm -e --nodeps XXX

### 安装MySQL

~~安装依赖包：可以不执行，正常情况下不会缺少这些依赖包，留作备用。~~

yum -y install autoconf

yum -y install numa\*

yum -y install libaio, libaio-devel

安装MySQL：

yum install -y mysql-community-server

yum install -y mysql-community-devel

替换配置文件：

rm -rf /etc/my.cnf

vi /etc/my.cnf

添加如下内容

[client]

port = 3306

socket = /var/lib/mysql/mysql.sock

default-character-set=utf8

[mysql]

no-auto-rehash

default-character-set=utf8

[mysqld]

datadir=/var/lib/mysql

socket=/var/lib/mysql/mysql.sock

user=mysql

symbolic-links=0

lower\_case\_table\_names=1

character-set-server=utf8

wait\_timeout=31536000

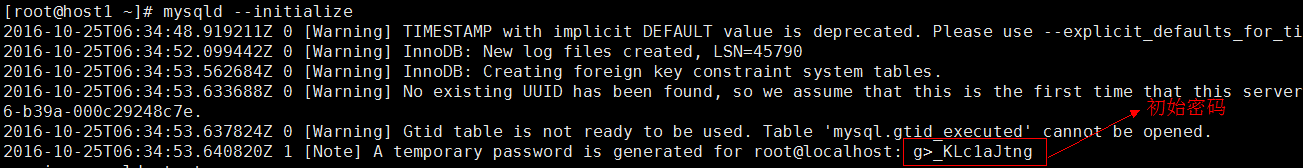
[mysqld\_safe]

log-error=/var/log/mysqld.log

pid-file=/var/run/mysqld/mysqld.pid

初始化数据库：

mysqld --initialize



开启服务：

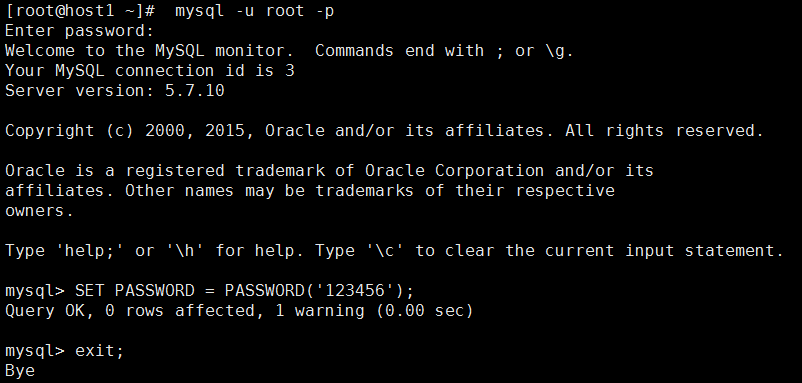
servicemysqld start



登录MySQL为root用户配置新密码：

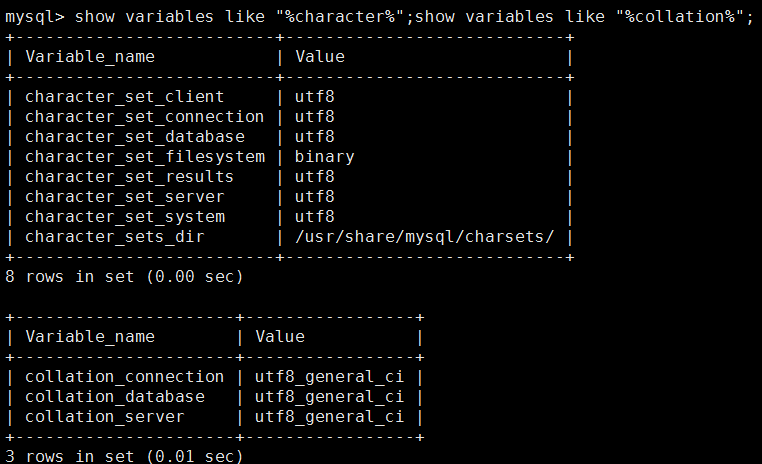
mysql -uroot -p g>\_KLc1aJtng

SET PASSWORD = PASSWORD('123456'); //设置新密码



查看编码格式

show variables like "%character%";show variables like "%collation%";



## 补充：安装mvn

> mkdir /usr/lib/mvn

> unzip /tools/apache-maven-3.0.4-bin.zip -d /usr/lib/mvn

> ll /usr/lib/mvn

> vim /etc/profile

#MAVEN Environment

export MAVEN\_HOME=/usr/lib/mvn/apache-maven-3.0.4

export PATH=${PATH}:${MAVEN\_HOME}/bin

> echo $MAVEN\_HOME

> mvn -version

## 同步网络时间

>ntpdate 202.108.6.95

# 关于我们

北京东方金信科技有限公司（后简称东方金信）成立于2013年2月，核心产品来自Google分布式大数据平台，专注于实现企业级高性能分布式大数据解决方案和云解决方案。公司核心团队来自美国Google、美国银行、阿里、 Teradata、IBM等国内外知名企业的核心数据开发团队，目前公司有员工100多人，研发团队80%为研究生及以上学历。目前公司已经签约华夏银行、光大银行、哈尔滨银行、兰州银行、晋商银行、国华保险、中国移动、中国联通、中材国际、解放军报等几十个大数据平台及应用项目。

公司在2013年获得第一版Hadoop发行版软件著作权。2014年在兰州银行实施落地大数据平台项目，该项目被《银行家》杂志评选为2014年金融创新奖。

2015年被《合作伙伴》评为大数据优秀企业。

2016年东方金信与北方交通大学联合成立铁路大数据研究中心。在交大计算机学院挂牌成立当天，中心就签署了中国铁路大数据的首个案例——朔黄铁路大数据运维管理平台项目。

2016年，东方金信推出基于云和多租户的大数据平台，并获得第二版大数据平台发行版软件著作权，同时参加了工信部大数据平台能力测试，本次测试认证主要针对是商用大数据平台，包括Hadoop和Spark生态圈所有组件。能力认证包括七大项，共38小项的测试用例。东方金信全部通过认证，在中国众多大数据平台厂商中脱颖而出，在最终的专家评审中，通过企业答辩和专家评审对测试结果进行论证，最终成为国内首批通过的六家企业之一。



地址： 北京市海淀区知春路致真大厦C座14层 - 100036

邮箱： dongfangjinxin@seaboxdata.com

电话： 010-56930733

网址：[www.seaboxdata.com](http://www.seaboxdata.com/)