大数据技术之Ambari

版本：V1.0

第1章 Ambari简述

1.1 什么是Ambari

Apache Ambari项目旨在通过开发用于配置，管理和监控Apache Hadoop集群的软件来简化Hadoop管理。Ambari提供了一个直观，易用的Hadoop管理Web UI。

1.2 Ambari的功能

提供了跨任意数量的主机安装Hadoop服务的分步向导。

处理群集的Hadoop服务配置。

提供集中管理，用于在整个集群中启动，停止和重新配置Hadoop服务。

提供了一个仪表板，用于监控Hadoop集群的运行状况和状态。

利用Ambari指标系统进行指标收集。

利用Ambari Alert Framework进行系统警报，并在需要您注意时通知您（例如，节点出现故障，剩余磁盘空间不足等）。

第2章 环境准备

注意：以下操作三台机器都需要进行

2.1 虚拟机准备

克隆三台虚拟机（hadoop102、hadoop103、hadoop104），配置好对应主机的网络IP、主机名称、关闭防火墙。

[root@hadoop102 ~]# chkconfig iptables off

[root@hadoop102 ~]# service iptables stop

[root@hadoop102 ~]# chkconfig --list iptables

iptables 0:关闭 1:关闭 2:关闭 3:关闭 4:关闭 5:关闭 6:关闭

2.2 关闭SELINUX

[root@hadoop102 ~]# vim /etc/sysconfig/selinux  
将SELINUX=enforcing改为SELINUX=disabled

执行该命令后重启机器生效

2.3 安装JDK

1）在hadoop102的/opt目录下创建module和software文件夹

[root@hadoop102 opt]# mkdir module

[root@hadoop102 opt]# mkdir software

2）用SecureCRT将jdk-8u144-linux-x64.tar.gz导入到hadoop102的/opt/software目录下

3）在Linux系统下的opt目录中查看软件包是否导入成功

[root@hadoop102 software]$ ls

jdk-8u144-linux-x64.tar.gz

4）解压JDK到/opt/module目录下

[root@hadoop102 software]$ tar -zxvf jdk-8u144-linux-x64.tar.gz -C /opt/module/

5）配置JDK环境变量

（1）先获取JDK路径

[root@hadoop102 jdk1.8.0\_144]$ pwd

/opt/module/jdk1.8.0\_144

（2）打开/etc/profile文件

[root@hadoop102 software]$ vi /etc/profile

在profile文件末尾添加JDK路径

#JAVA\_HOME

export JAVA\_HOME=/opt/module/jdk1.8.0\_144

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

（3）保存后退出

:wq

（4）让修改后的文件生效

[root@hadoop102 jdk1.8.0\_144]$ source /etc/profile

6）测试JDK是否安装成功

[root@hadoop102 jdk1.8.0\_144]# java -version

java version "1.8.0\_144"

7）将hadoop102中的JDK和环境变量分发到hadoop103、hadoop104两台主机

[root@hadoop102 opt]# xsync /opt/module/

[root@hadoop102 opt]# xsync /etc/profile

分别在hadoop103、hadoop104上source一下

[root@hadoop103 ~]$ source /etc/profile

[root@hadoop104 ~]# source /etc/profile

2.4 SSH免密登录

配置hadoop102对hadoop102、hadoop103、hadoop104三台主机的免密登陆。

（1）生成公钥和私钥：

[root@hadoop102 .ssh]$ ssh-keygen -t rsa

然后敲（三个回车），就会生成两个文件id\_rsa（私钥）、id\_rsa.pub（公钥）

（2）将公钥拷贝到要免密登录的目标机器上

[root@hadoop102 .ssh]$ ssh-copy-id hadoop102

[root@hadoop102 .ssh]$ ssh-copy-id hadoop103

[root@hadoop102 .ssh]$ ssh-copy-id hadoop104

2.5 修改yum源为阿里云镜像

[root@hadoop102 yum.repos.d]# mv /etc/yum.repos.d/CentOS-Base.repo /etc/yum.repos.d/CentOS-Base.repo.bk

[root@hadoop102 yum.repos.d]# wget -O /etc/yum.repos.d/CentOS-Base.repo http://mirrors.aliyun.com/repo/Centos-6.repo

//把服务器的包下载到本地缓存

[root@hadoop102 yum.repos.d]# yum makecache

2.6 安装ntp

[root@hadoop102 ~]# yum install -y ntp

[root@hadoop102 ~]# chkconfig --list ntpd

[root@hadoop102 ~]# chkconfig ntpd on

[root@hadoop102 ~]# service ntpd start

2.7 关闭Linux的THP服务

如果不关闭transparent\_hugepage，HDFS会因为这个性能严重受影响。

关闭transparent\_hugepage方法是：

[root@hadoop102 ~]# vim /etc/grub.conf

添加 transparent\_hugepage=never

[root@hadoop102 ~]# vim /etc/rc.local

添加：

if test -f /sys/kernel/mm/transparent\_hugepage/defrag; then

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

fi

if test -f /sys/kernel/mm/transparent\_hugepage/enabled; then

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

fi

exit 0

重启之后，用下面的命令检查：

[root@hadoop102 yum.repos.d]#

cat /sys/kernel/mm/redhat\_transparent\_hugepage/enabled

always madvise [never]

有 [never]则表示THP被禁用

2.8 配置UMASK

设定用户所创建目录的初始权限

[root@hadoop102 ~]# umask 0022

2.9 禁止离线更新

vim /etc/yum/pluginconf.d/refresh-packagekit.conf

修改：enabled=0

第3章 安装Ambari集群

注意：以下操作主节点操作即可

3.1 制作本地源

制作本地源是因为在线安装Ambari太慢。制作本地源只需在主节点上进行。

3.1.1 配置HTTPD 服务

配置HTTPD 服务到系统层使其随系统自动启动

[root@hadoop102 ~]# chkconfig httpd on

[root@hadoop102 ~]# service httpd start

3.1.2 安装工具

安装本地源制作相关工具

[root@hadoop102 ~]# yum install yum-utils createrepo yum-plugin-priorities -y

[root@hadoop102 ~]# vim /etc/yum/pluginconf.d/priorities.conf

添加gpgcheck=0

3.1.3 将下载的3个tar包解压

[root@hadoop102 software]# tar -zxvf /opt/software/ambari-2.5.0.3-centos6.tar.gz -C /var/www/html/

[root@hadoop102 software]mkdir /var/www/html/hdp

[root@hadoop102 software]# tar -zxvf /opt/software/HDP-2.6.0.3-centos6-rpm.tar.gz -C /var/www/html/hdp

[root@hadoop102 software]# tar -zxvf /opt/software/HDP-UTILS-1.1.0.21-centos6.tar.gz -C /var/www/html/hdp

3.1.4 创建本地源

[root@hadoop102 software]# cd /var/www/html/

[root@hadoop102 html]# createrepo ./

3.1.5 将Ambari存储库文件下载到安装主机上的目录中

[root@hadoop102 yum.repos.d]# wget -nv http://public-repo-1.hortonworks.com/ambari/centos6/2.x/updates/2.6.1.5/ambari.repo -O /etc/yum.repos.d/ambari.repo

3.1.6 修改配置文件，配置为本地源

修改ambari.repo，配置为本地源

[root@hadoop102 html]# vim /etc/yum.repos.d/ambari.repo

#VERSION\_NUMBER=2.6.1.5-3

[ambari-2.6.1.5]

name=ambari Version - ambari-2.6.1.5

baseurl=http://hadoop102/ambari/centos6/

gpgcheck=0

gpgkey=http://hadoop102/ambari/centos6/RPM-GPG-KEY/RPM-GPG-KEY-Jenkins

enabled=1

priority=1

修改hdp-util.repo，配置为本地源

[root@hadoop102 hdp]# vim /var/www/html/hdp/hdp-util.repo

[HDP-UTILS-1.1.0.21]

name=Hortonworks Data Platform Version - HDP-UTILS-1.1.0.21

baseurl=http://hadoop102/hdp/

gpgcheck=0

enabled=1

priority=1

修改hdp.repo，配置为本地源

[root@hadoop102 centos6]#

vim /var/www/html/hdp/HDP/centos6/hdp.repo

#VERSION\_NUMBER=2.6.0.3-8

[HDP-2.6.0.3]

name=HDP Version - HDP-2.6.0.3

baseurl=http://hadoop102/hdp/HDP/centos6/

gpgcheck=0

gpgkey=http://hadoop102/hdp/HDP/centos6/RPM-GPG-KEY/RPM-GPG-KEY-Jenkins

enabled=1

priority=1

[HDP-UTILS-1.1.0.21]

name=HDP-UTILS Version - HDP-UTILS-1.1.0.21

baseurl=http://hadoop102/hdp/

gpgcheck=0

gpgkey=http://hadoop102/hdp/RPM-GPG-KEY/RPM-GPG-KEY-Jenkins

enabled=1

priority=1

[root@hadoop102 ~]# yum clean all

[root@hadoop102 ~]# yum makecache

查看是否有Ambari

[root@hadoop102 ~]# yum repolist

查看Ambari 与 HDP 资源的资源库。

也可以打开浏览器查看一下：

<http://hadoop102/ambari/centos6/>

<http://hadoop102/hdp/HDP/centos6/>

<http://hadoop102/hdp/>

3.2 安装MySQL

Ambari使用的默认数据库是PostgreSQL，用于存储安装元数据，可以使用自己安装MySQL数据库作为Ambari元数据库。

注意：一定要用root用户操作如下步骤；先卸载MySQL再安装

**1）安装包准备**

（1）查看MySQL是否安装

[root@hadoop102 桌面]# rpm -qa|grep mysql

mysql-libs-5.1.73-7.el6.x86\_64

（2）如果安装了MySQL，就先卸载

[root@hadoop102 桌面]# rpm -e --nodeps mysql-libs-5.1.73-7.el6.x86\_64

（3）解压mysql-libs.zip文件到当前目录

[root@hadoop102 software]# unzip mysql-libs.zip

[root@hadoop102 software]# ls

mysql-libs.zip

mysql-libs

（4）进入到mysql-libs文件夹下

[root@hadoop102 mysql-libs]# ll

总用量 76048

-rw-r--r--. 1 root root 18509960 3月 26 2015 MySQL-client-5.6.24-1.el6.x86\_64.rpm

-rw-r--r--. 1 root root 3575135 12月 1 2013 mysql-connector-java-5.1.27.tar.gz

-rw-r--r--. 1 root root 55782196 3月 26 2015 MySQL-server-5.6.24-1.el6.x86\_64.rpm

**2）安装MySQL服务器**

（1）安装MySQL服务端

[root@hadoop102 mysql-libs]# rpm -ivh MySQL-server-5.6.24-1.el6.x86\_64.rpm

（2）查看产生的随机密码

[root@hadoop102 mysql-libs]# cat /root/.mysql\_secret

XrBwJEENp2yhVe4O

（3）查看MySQL状态

[root@hadoop102 mysql-libs]# service mysql status

（4）启动MySQL

[root@hadoop102 mysql-libs]# service mysql start

**3）安装MySQL客户端**

（1）安装MySQL客户端

[root@hadoop102 mysql-libs]# rpm -ivh MySQL-client-5.6.24-1.el6.x86\_64.rpm

（2）链接MySQL

[root@hadoop102 mysql-libs]# mysql -uroot -pXrBwJEENp2yhVe4O

（3）修改密码

mysql>SET PASSWORD=PASSWORD('000000');

（4）退出MySQL

mysql>exit

**4）MySQL中user表中主机配置**

配置只要是root用户+密码，在任何主机上都能登录MySQL数据库。

（1）进入MySQL

[root@hadoop102 mysql-libs]# mysql -uroot -p000000

（2）显示数据库

mysql>show databases;

（3）使用MySQL数据库

mysql>use mysql;

（4）展示MySQL数据库中的所有表

mysql>show tables;

（5）展示user表的结构

mysql>desc user;

（6）查询user表

mysql>select User, Host, Password from user;

（7）修改user表，把Host表内容修改为%

mysql>update user set host='%' where host='localhost';

（8）删除root用户的其他host

mysql>

delete from user where Host='hadoop102';

delete from user where Host='127.0.0.1';

delete from user where Host='::1';

（9）刷新

mysql>flush privileges;

（10）退出

mysql>quit;

3.3 安装Ambari

1）安装ambari-server

[root@hadoop102 hdp]# yum install ambari-server

2） 拷贝mysql驱动

将mysql-connector-java.jar复制到/usr/share/java目录下并改名为mysql-connector-java.jar

[root@hadoop102 hdp]# mkdir /usr/share/java

[root@hadoop102 hdp]# cp /opt/software/mysql-libs/mysql-connector-java-5.1.27/mysql-connector-java-5.1.27-bin.jar /usr/share/java/mysql-connector-java.jar

将mysql-connector-java.jar复制到/var/lib/ambari-server/resources目录下并改名为mysql-jdbc-driver.jar

[root@hadoop102 hdp]# cp /usr/share/java/mysql-connector-java.jar /var/lib/ambari-server/resources/mysql-jdbc-driver.jar

修改ambari.properties文件

[root@hadoop102 hdp]#

vim /etc/ambari-server/conf/ambari.properties

添加

server.jdbc.driver.path=/usr/share/java/mysql-connector-java.jar

3.4 在MySQL中创建数据库

1）创建ambari库

[root@hadoop102 hdp]# mysql -u root -p000000

mysql >create database ambari;

2）使用Ambari自带脚本创建表

mysql >use ambari;

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

3）赋予用户root权限：

mysql> grant all privileges on \*.\* to 'root'@'%' identified by '000000';

4）刷新

mysql> flush privileges;

3.5 配置Ambari

执行

[root@hadoop102 hdp]# ambari-server setup

下面是配置执行流程，按照提示操作

1） 提示是否自定义设置。输入：y

Customize user account for ambari-server daemon [y/n] (n)? y

2）ambari-server 账号。

Enter user account for ambari-server daemon (root):

如果直接回车就是默认选择root用户

3）设置JDK。输入：3

Checking JDK...

Do you want to change Oracle JDK [y/n] (n)? y

[1] Oracle JDK 1.8 + Java Cryptography Extension (JCE) Policy Files 8

[2] Oracle JDK 1.7 + Java Cryptography Extension (JCE) Policy Files 7

[3] Custom JDK

==============================================================================

Enter choice (1): 3

如果上面选择3自定义JDK,则需要设置JAVA\_HOME。输入：/opt/module/jdk1.8.0\_144

WARNING: JDK must be installed on all hosts and JAVA\_HOME must be valid on all hosts.

WARNING: JCE Policy files are required for configuring Kerberos security. If you plan to use Kerberos,please make sure JCE Unlimited Strength Jurisdiction Policy Files are valid on all hosts.

Path to JAVA\_HOME: /opt/module/jdk1.8.0\_144

Validating JDK on Ambari Server...done.

Completing setup...

4）数据库配置。选择：y

Configuring database...

Enter advanced database configuration [y/n] (n)? y

5）选择数据库类型。输入：3

Configuring database...

==============================================================================

Choose one of the following options:

[1] - PostgreSQL (Embedded)

[2] - Oracle

[3] - MySQL

[4] - PostgreSQL

[5] - Microsoft SQL Server (Tech Preview)

[6] - SQL Anywhere

==============================================================================

Enter choice (3): 3

6）设置数据库的具体配置信息，根据实际情况输入，如果和括号内相同，则可以直接回车。如果想重命名，就输入。

Hostname (localhost):hadoop102

Port (3306):

Database name (ambari):

Username (ambari):root

Enter Database Password (bigdata):000000（这里不显示）

Re-Enter password:000000

7）将Ambari数据库脚本导入到数据库

WARNING: Before starting Ambari Server, you must run the following DDL against the database to create the schema: /var/lib/ambari-server/resources/Ambari-DDL-MySQL-CREATE.sql

Proceed with configuring remote database connection properties [y/n] (y)?

如果使用自己定义的数据库，必须在启动Ambari服务之前导入Ambari的sql脚本。

3.6 启动Ambari

启动命令为：

ambari-server start

停止命令为：

ambari-server stop

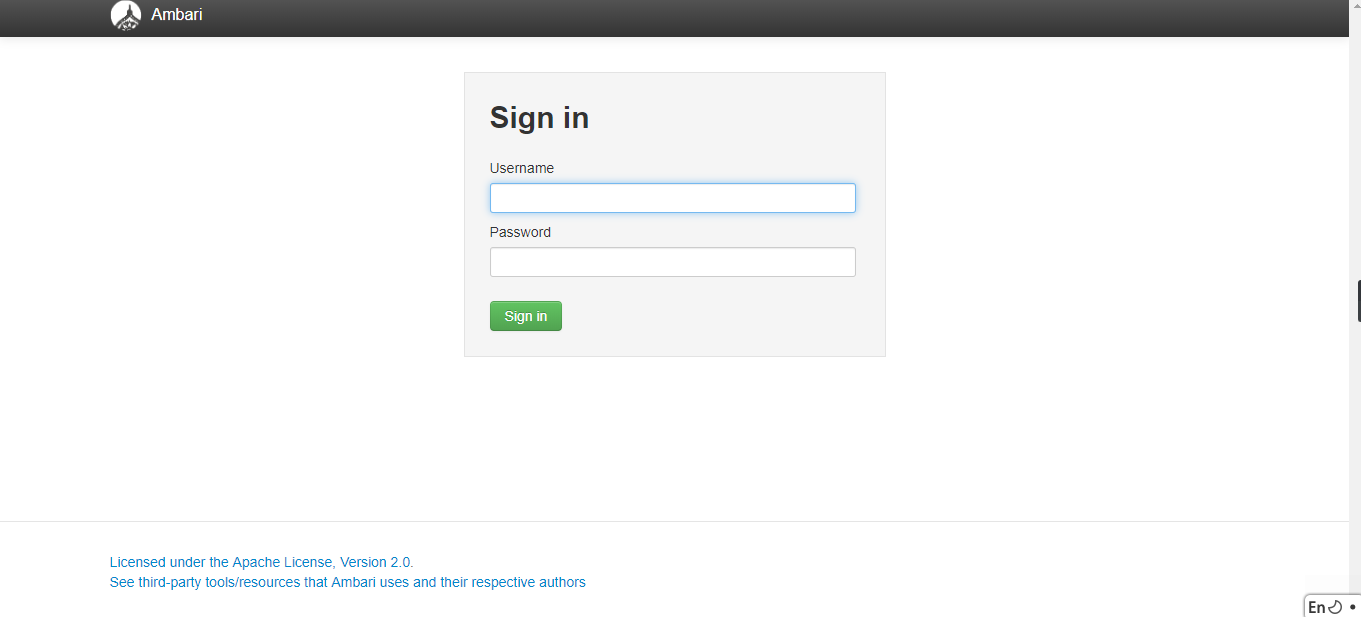
第4章 HDP集群部署

4.1 集群搭建

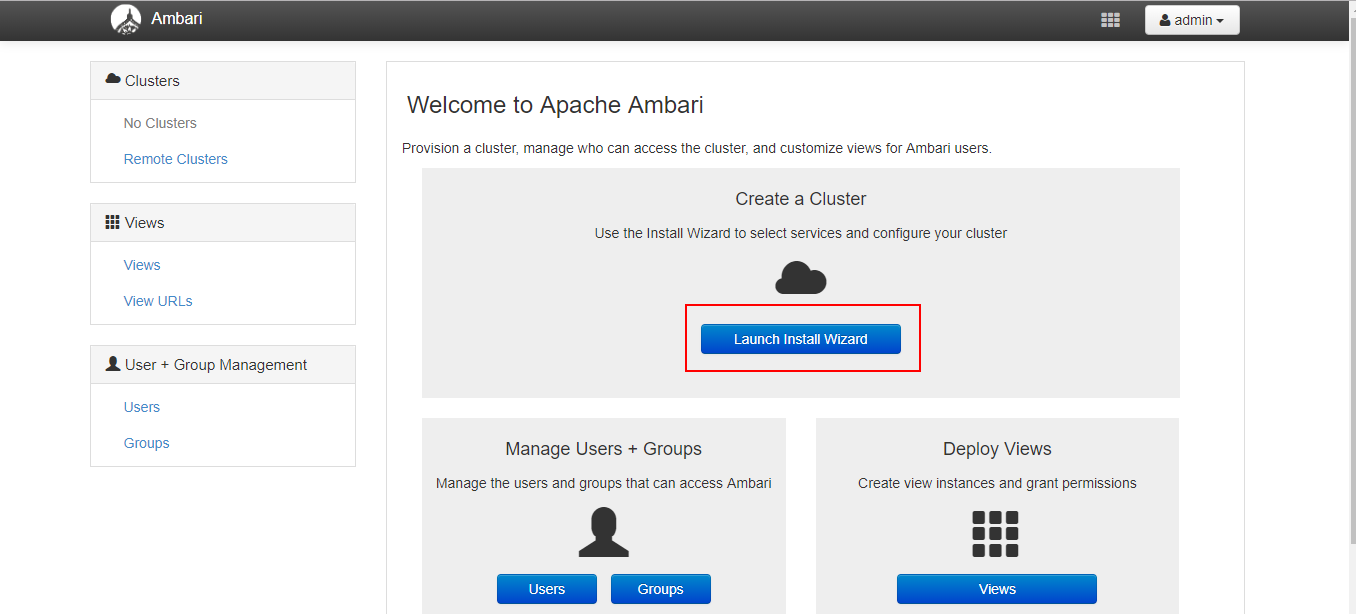
4.1.1进入登录页面

浏览器输入<http://hadoop102:8080/>

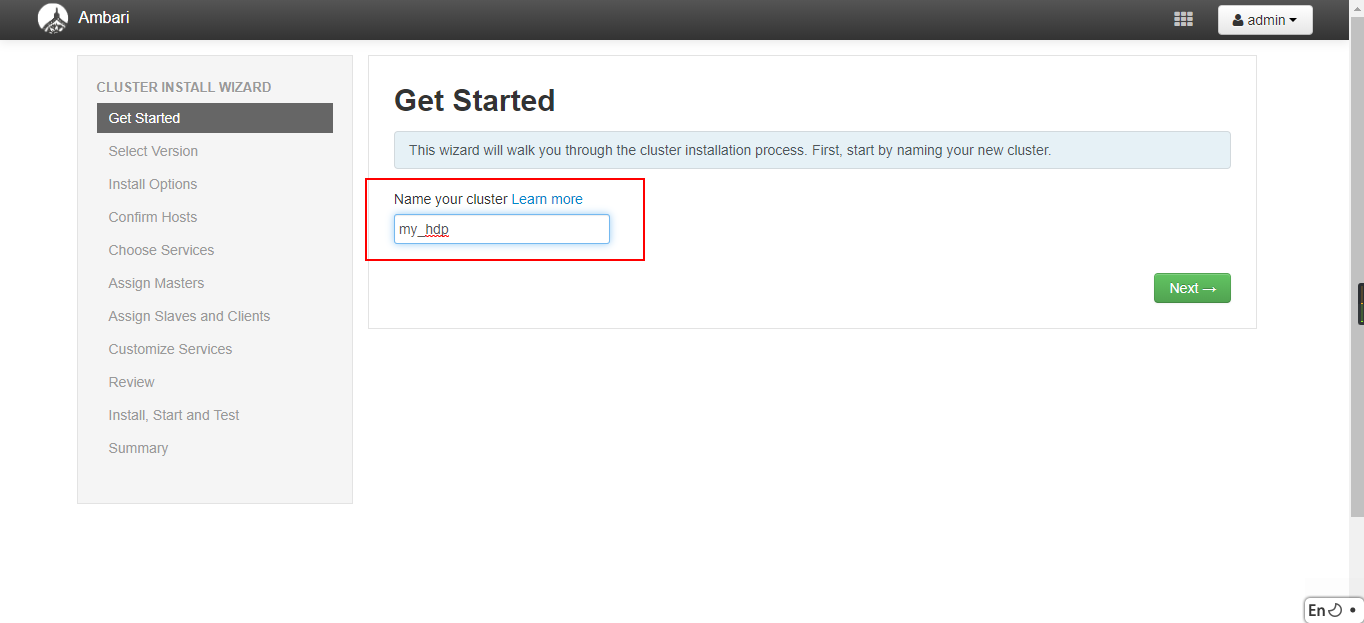
默认管理员账目密码：admin



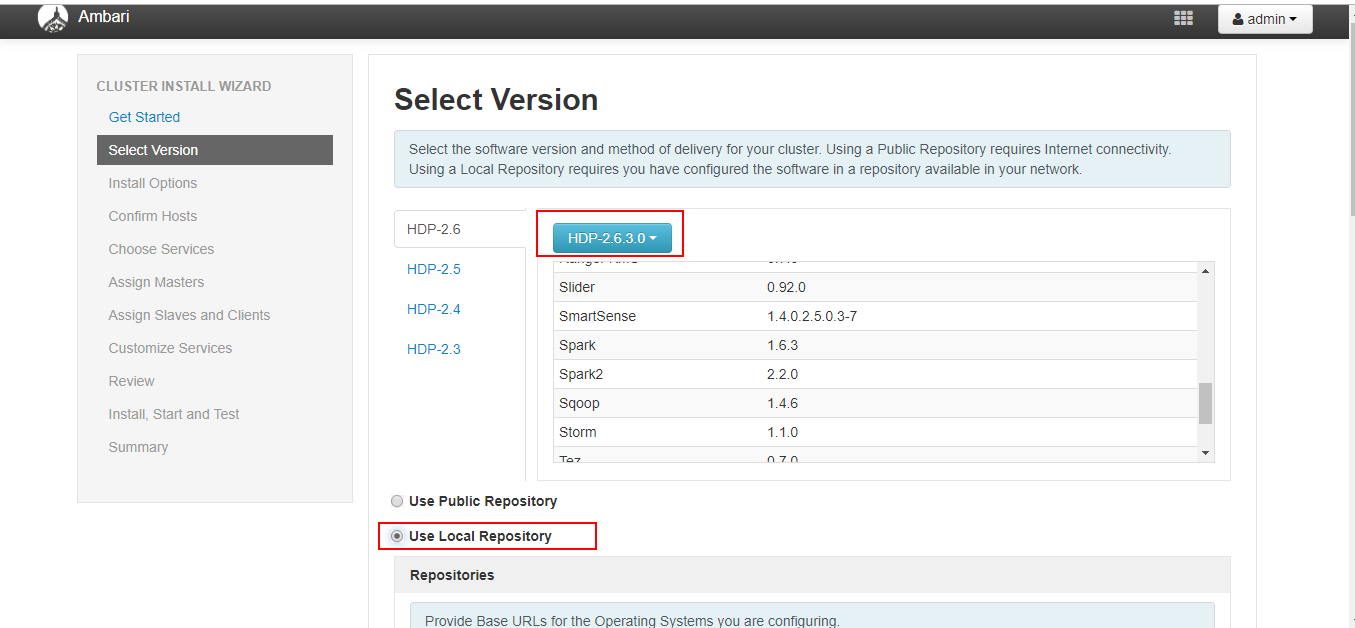
4.1.2 点击Launch Install Wizard



4.1.3 设置集群名称



4.1.4选择版本和存储库

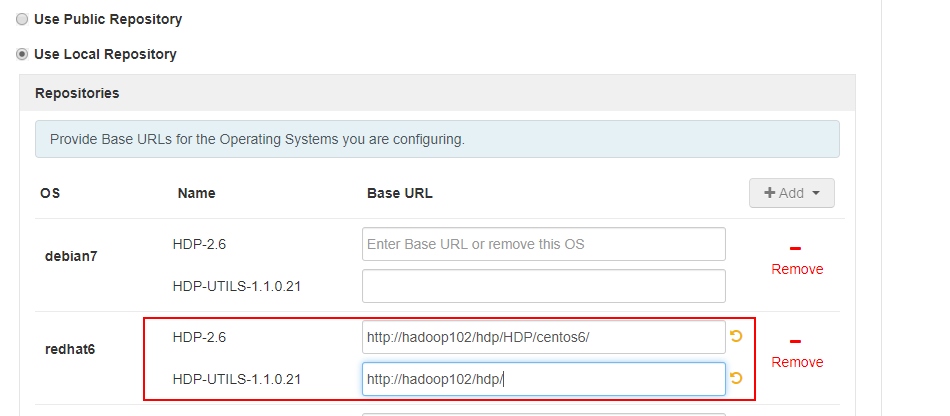


4.1.5 填写本地库地址

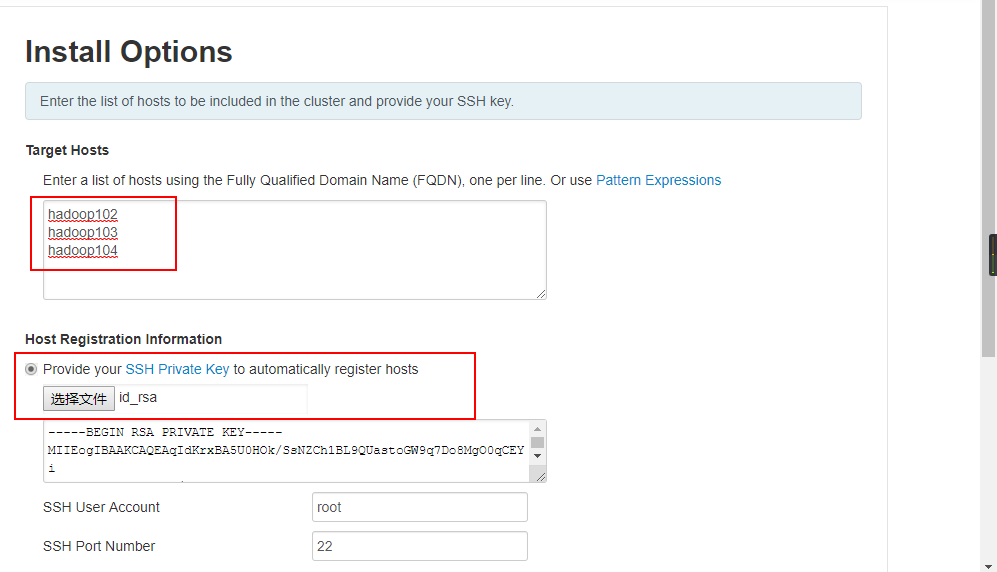
在redhat6后面分别填写

http://hadoop102/hdp/HDP/centos6/

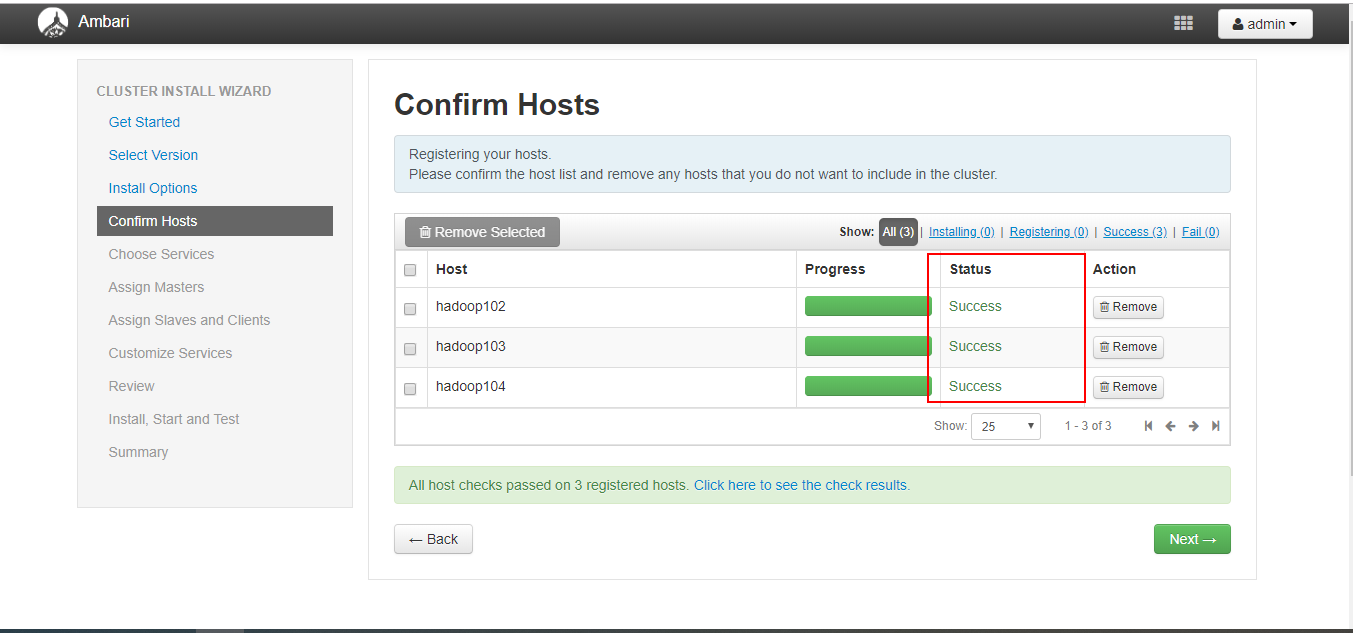
http://hadoop102/hdp/



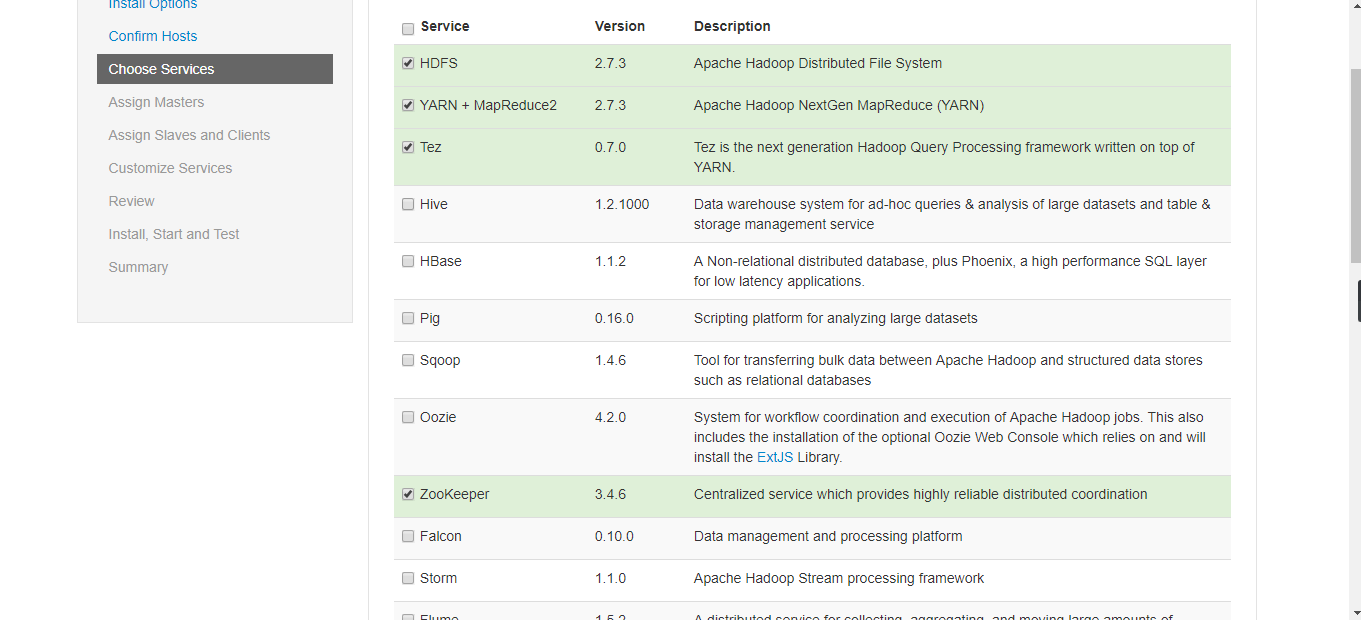
4.1.6 填写主机地址以及主节点的id.rsa文件



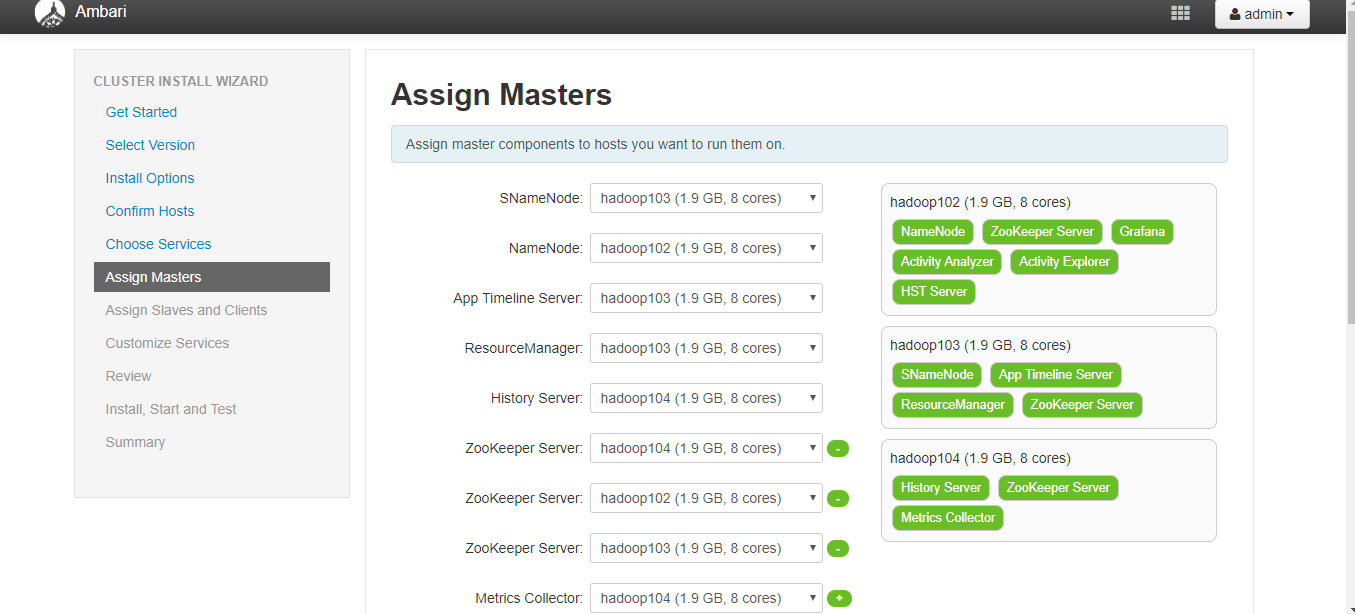
4.1.7 等待安装



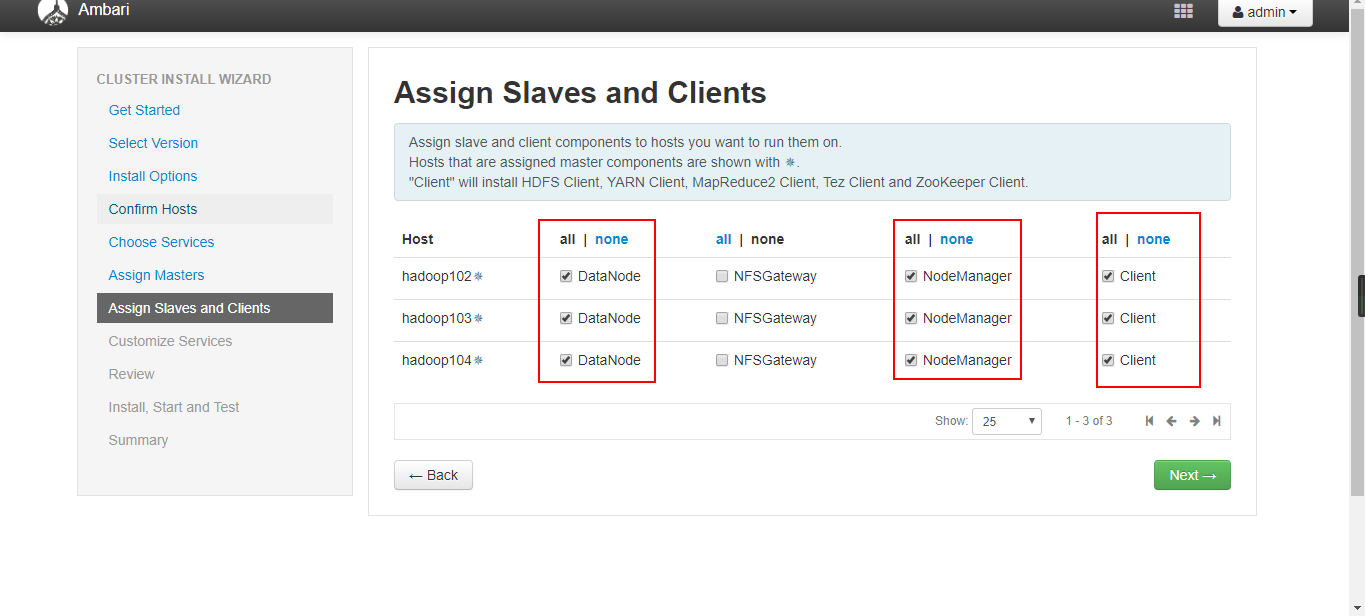
4.1.8 选择服务



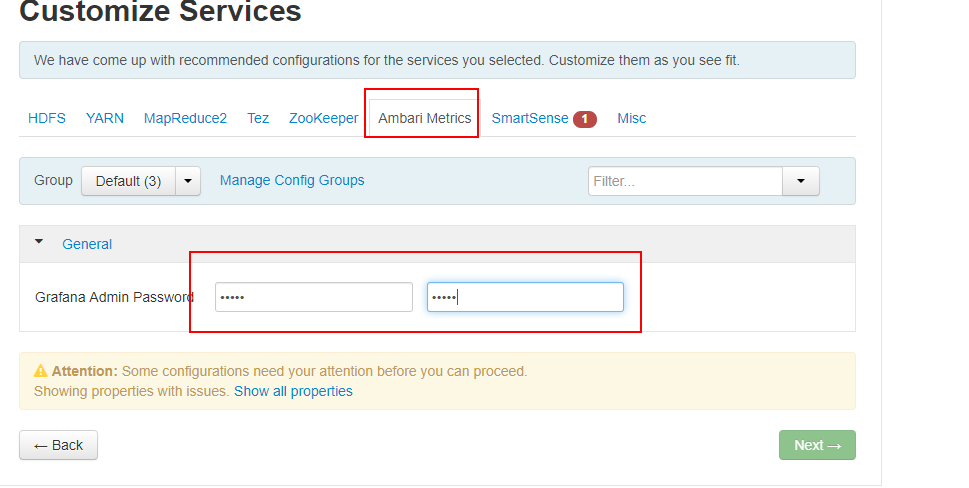
4.1.9 选择每台机器的角色

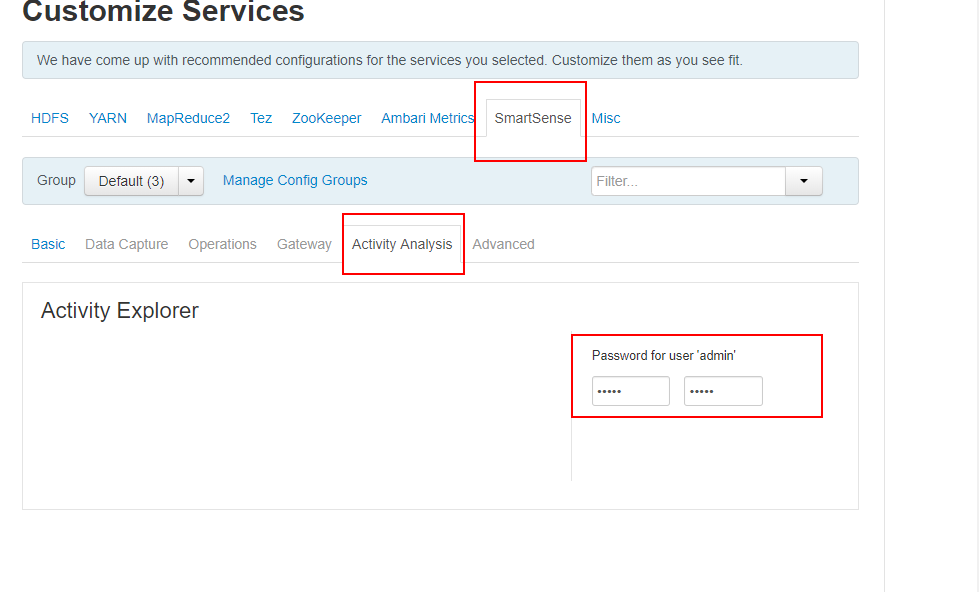


4.1.10 设置从节点

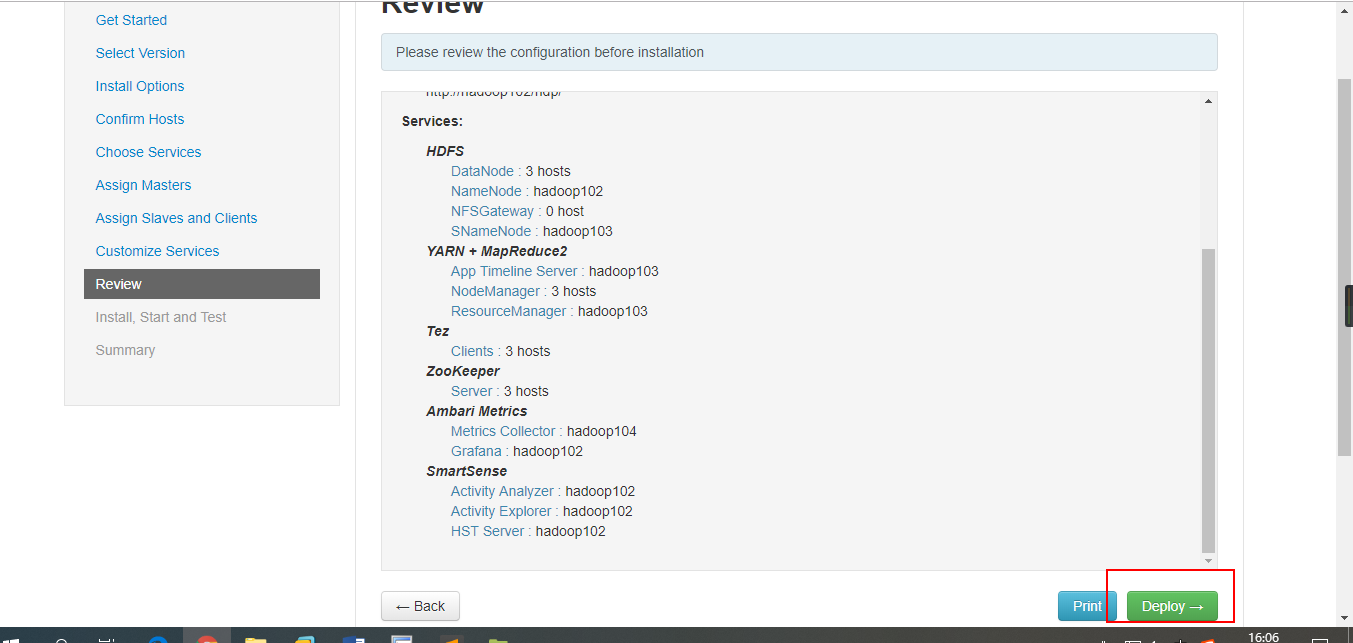


4.1.11 输入两次admin



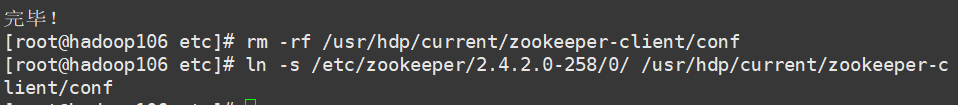


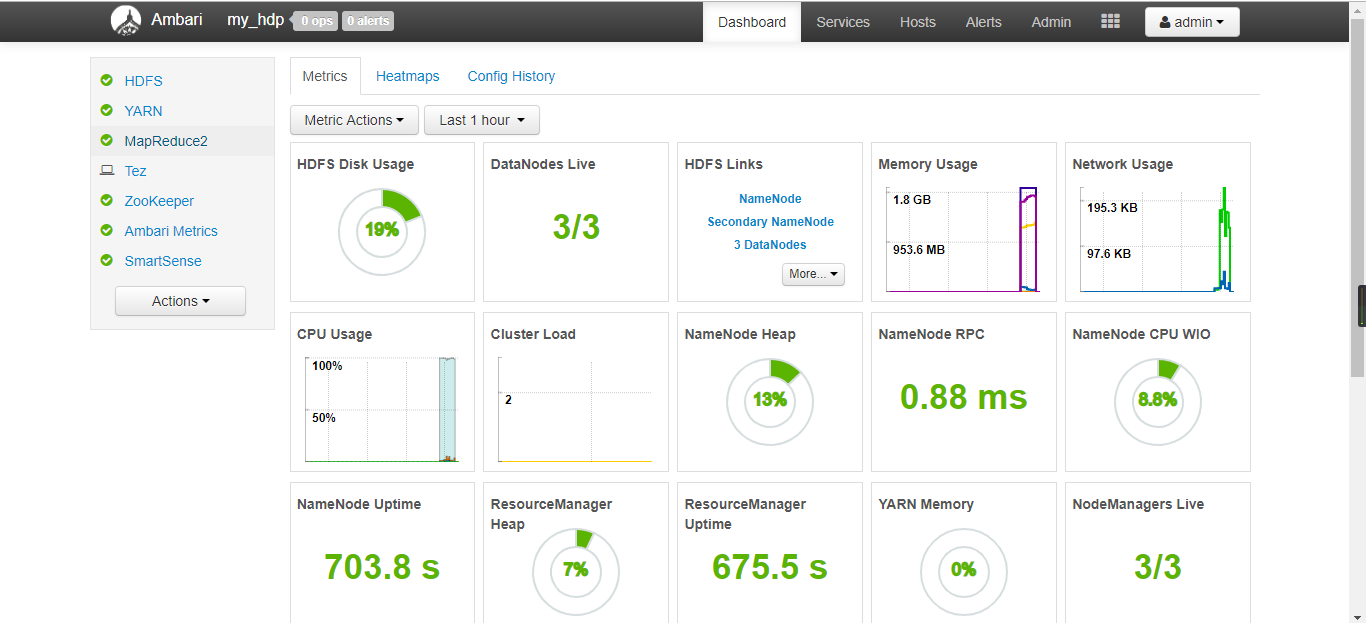
4.1.12 继续



4.1.13 等待服务安装和启动

yum -y erase hdp-select





4.2 安装Hive

4.2.1 在/path/to/mysql/目录下放置Mysql驱动，并重命名为：mysql-connector-java.jar

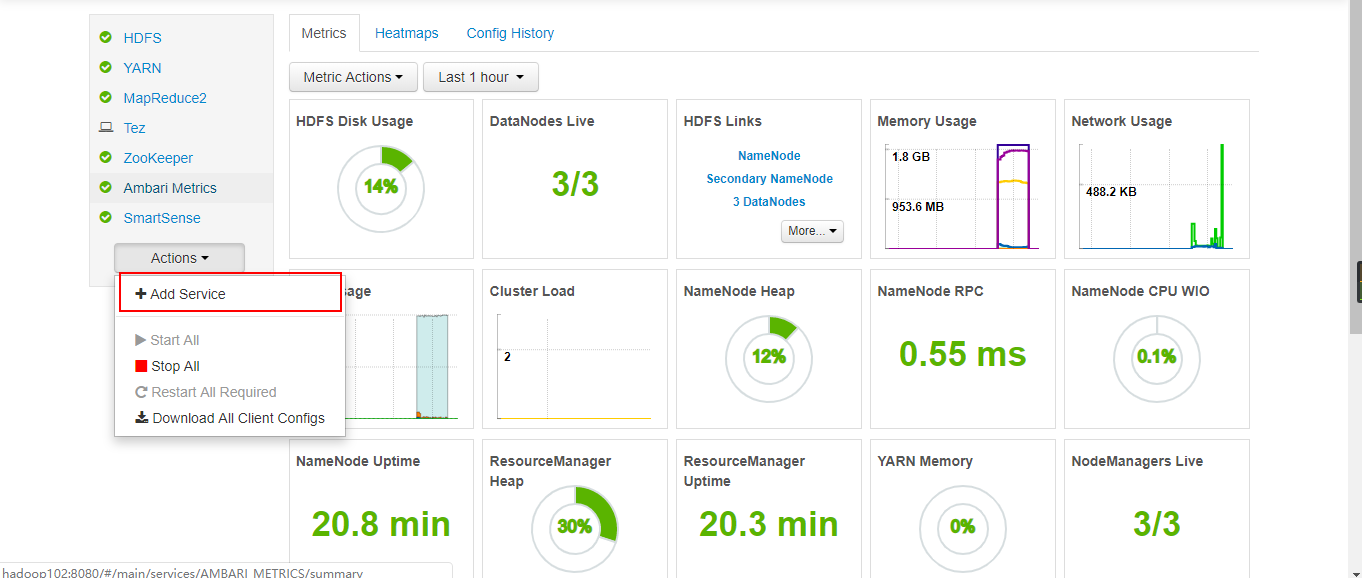
[root@hadoop102 software]# mkdir -p /path/to/mysql/

[root@hadoop102 software]# cp /opt/software/mysql-libs/mysql-connector-java-5.1.27/mysql-connector-java-5.1.27-bin.jar /path/to/mysql/mysql-connector-java.jar

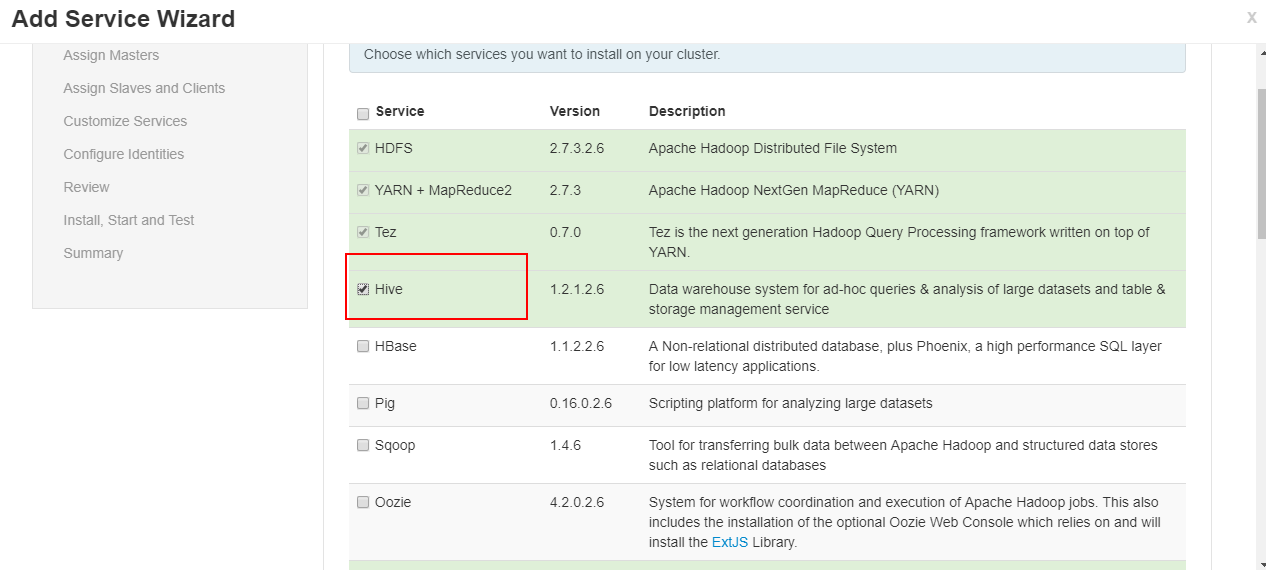
4.2.2 执行ambari-server setup

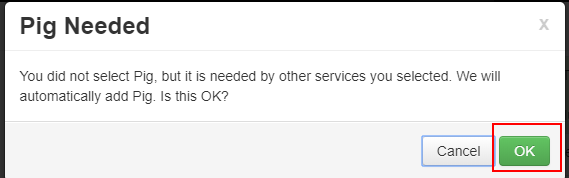
[root@hadoop102 software]# ambari-server setup --jdbc-db=mysql --jdbc-driver=/path/to/mysql/mysql-connector-java.jar

4.2.3 点击添加服务

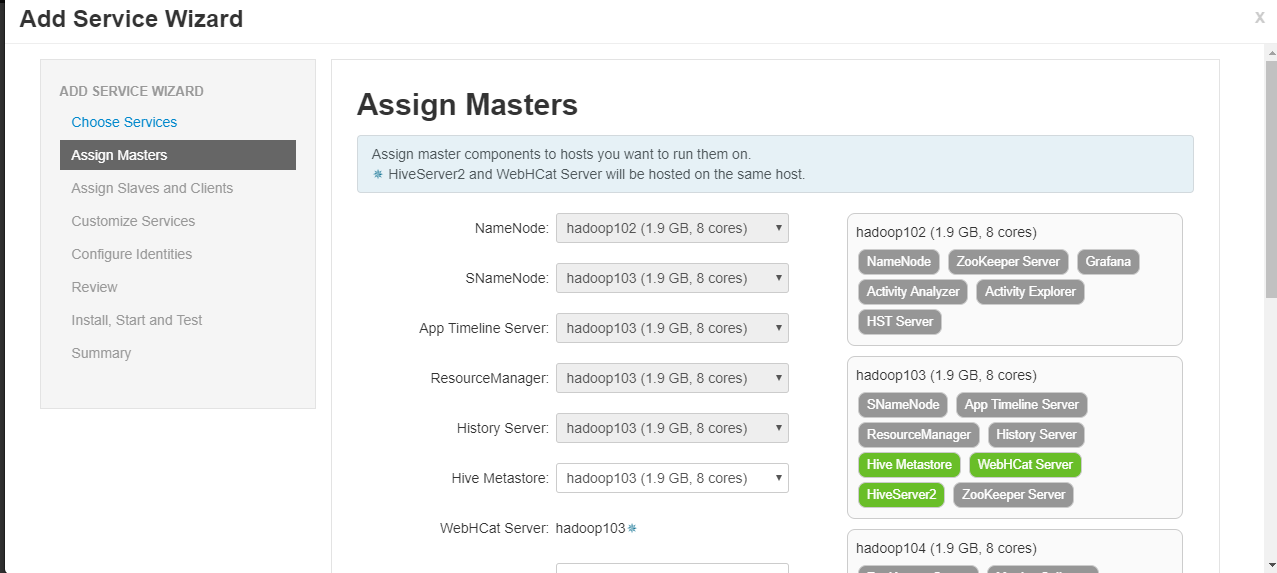


4.3.4 选择hive

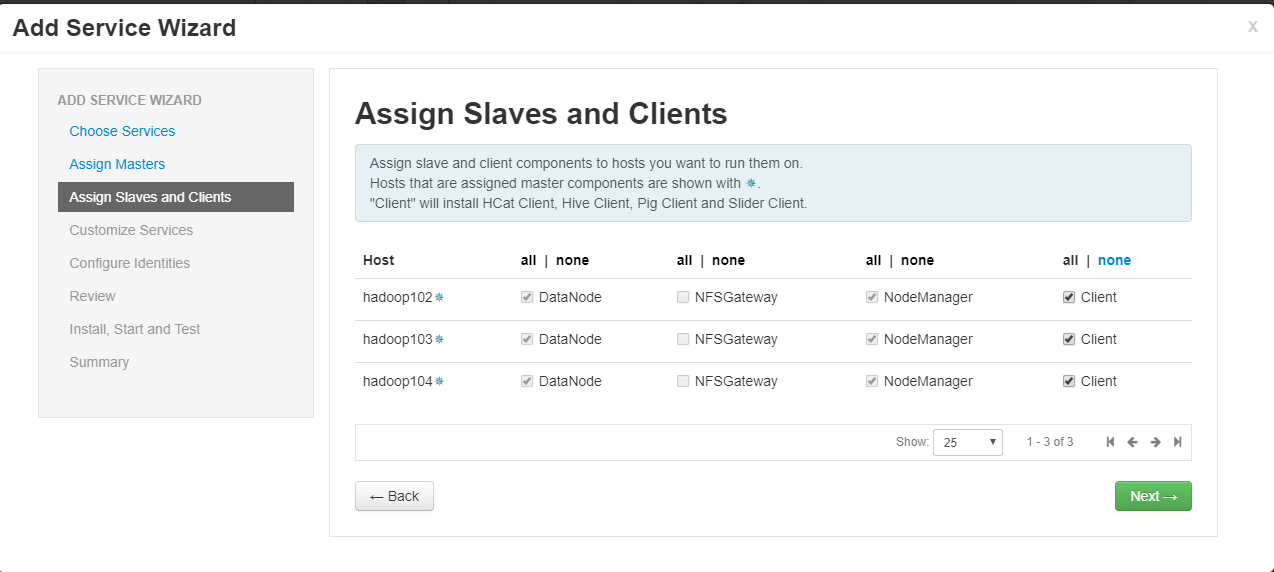




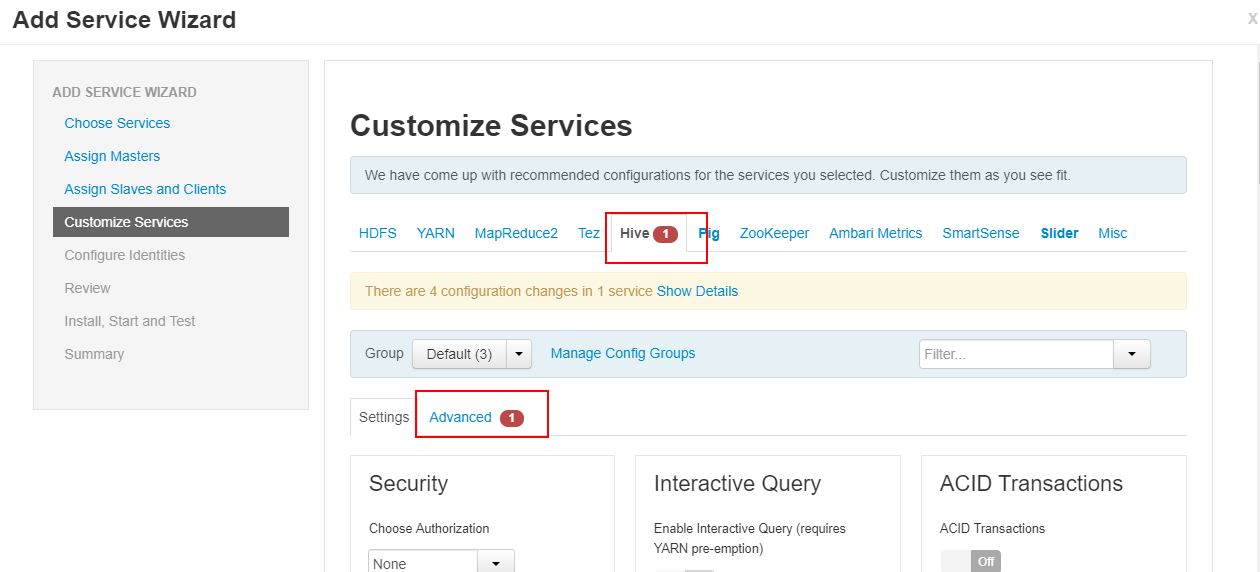
4.3.5 分配节点

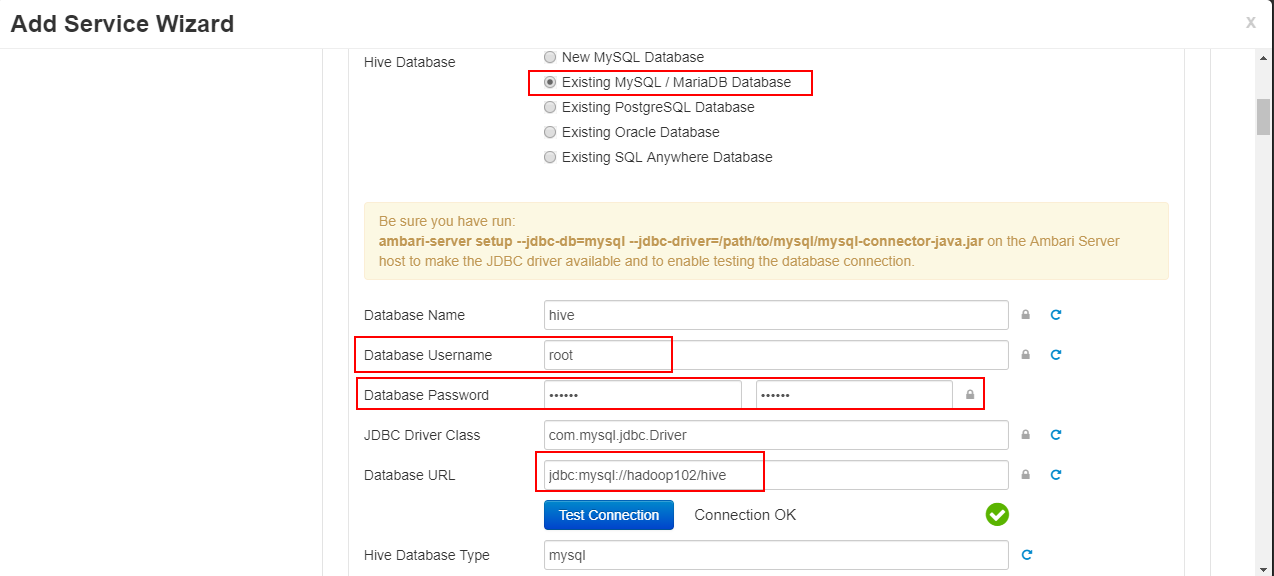


4.3.6 分配从属和客户端

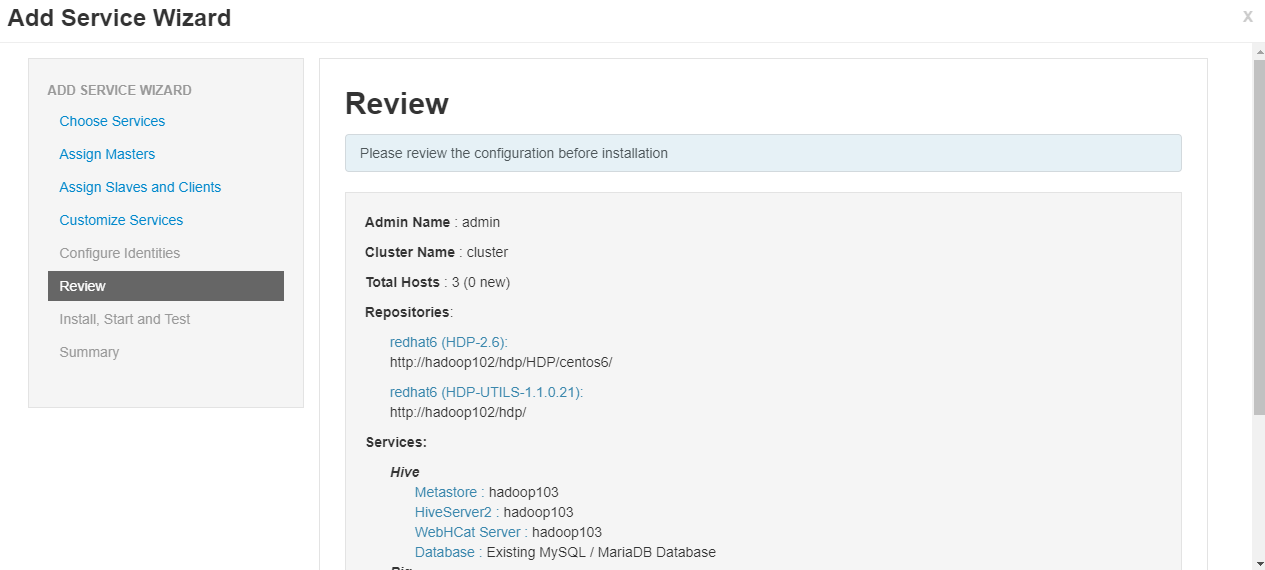


4.3.7 配置hive数据库连接

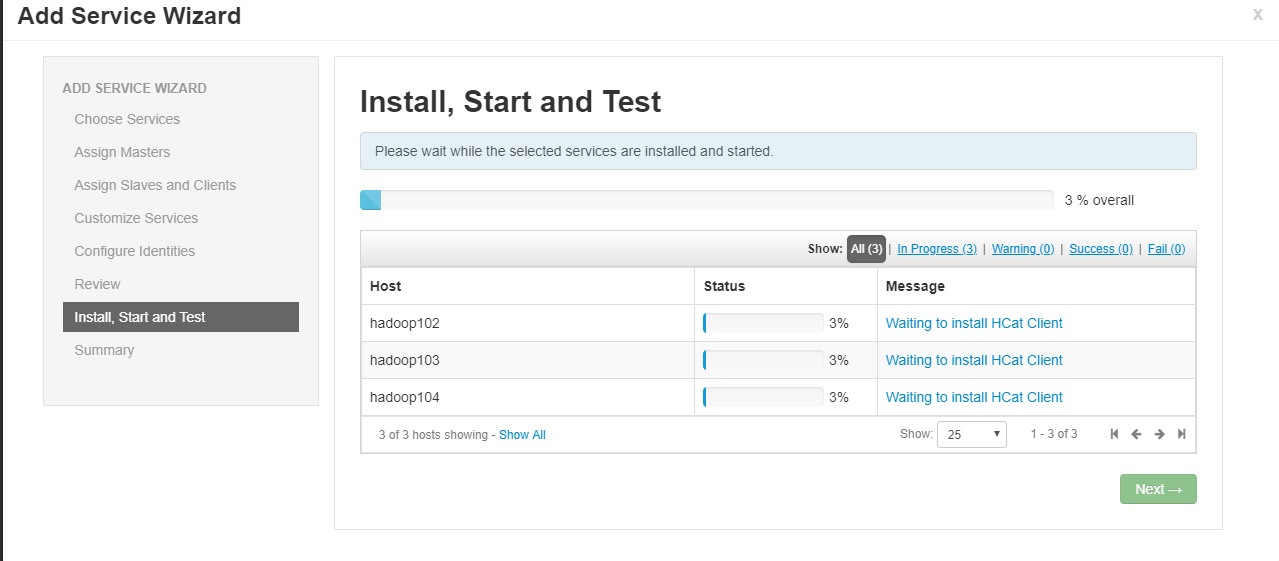




4.3.8 点击下一步

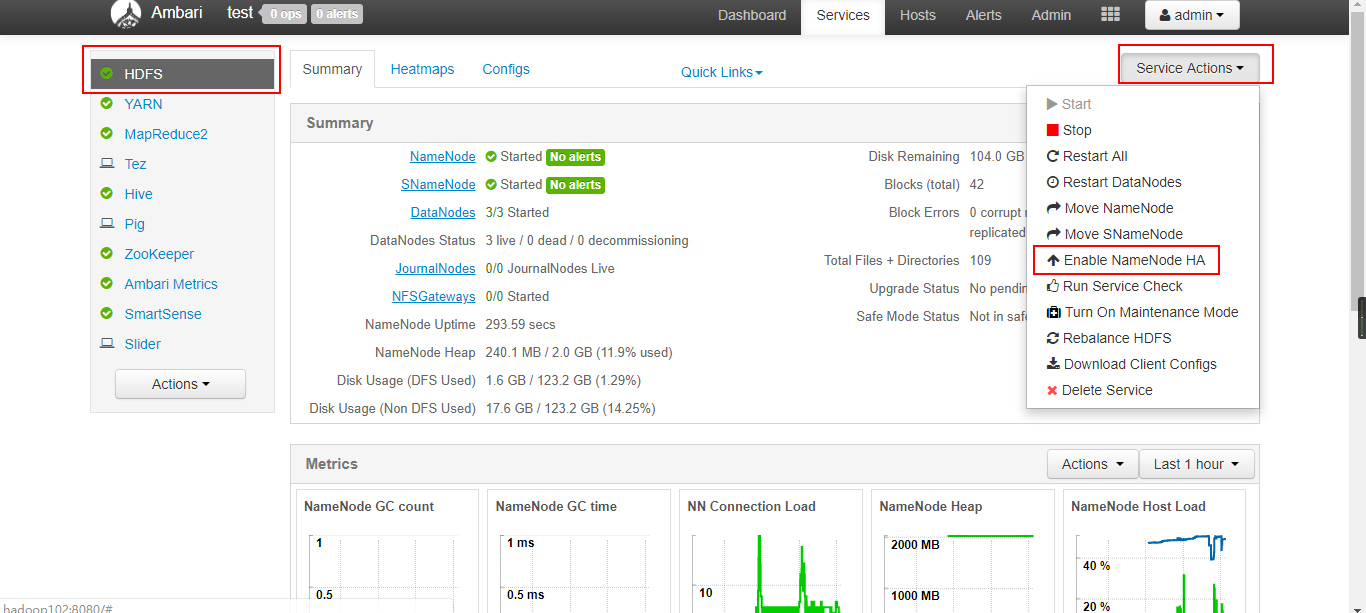


4.3.9 等待安装

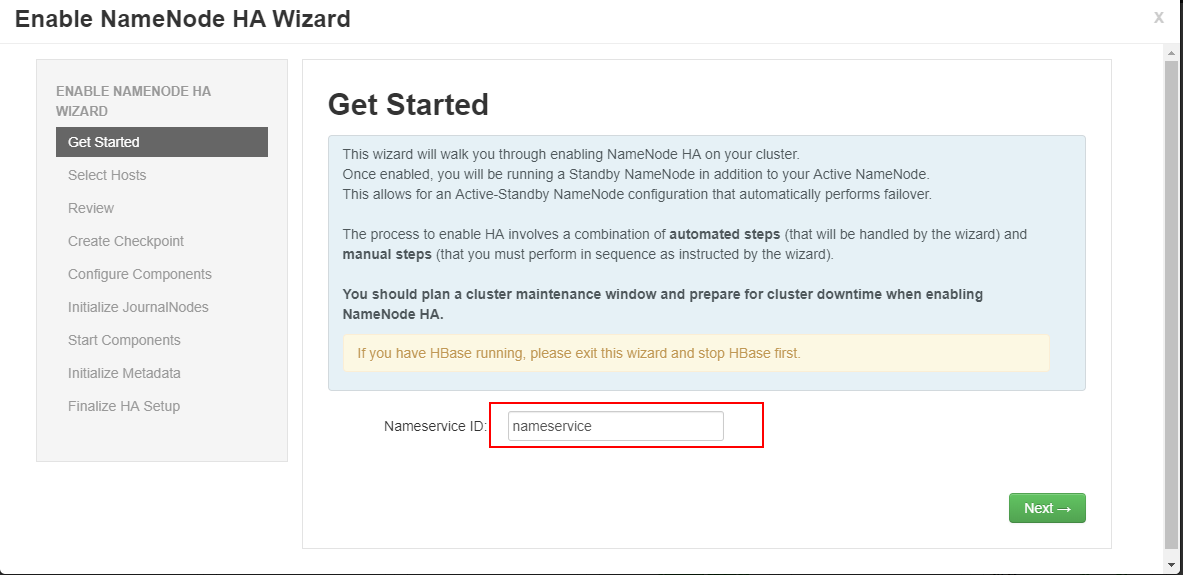


4.3 配置HDFS-HA

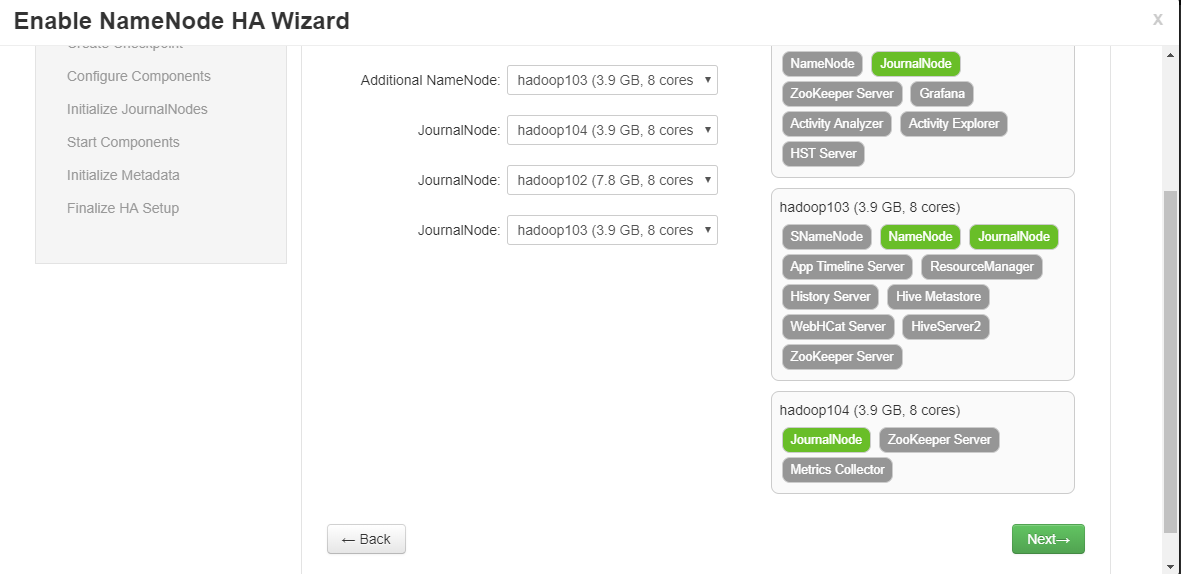
4.3.1 添加服务



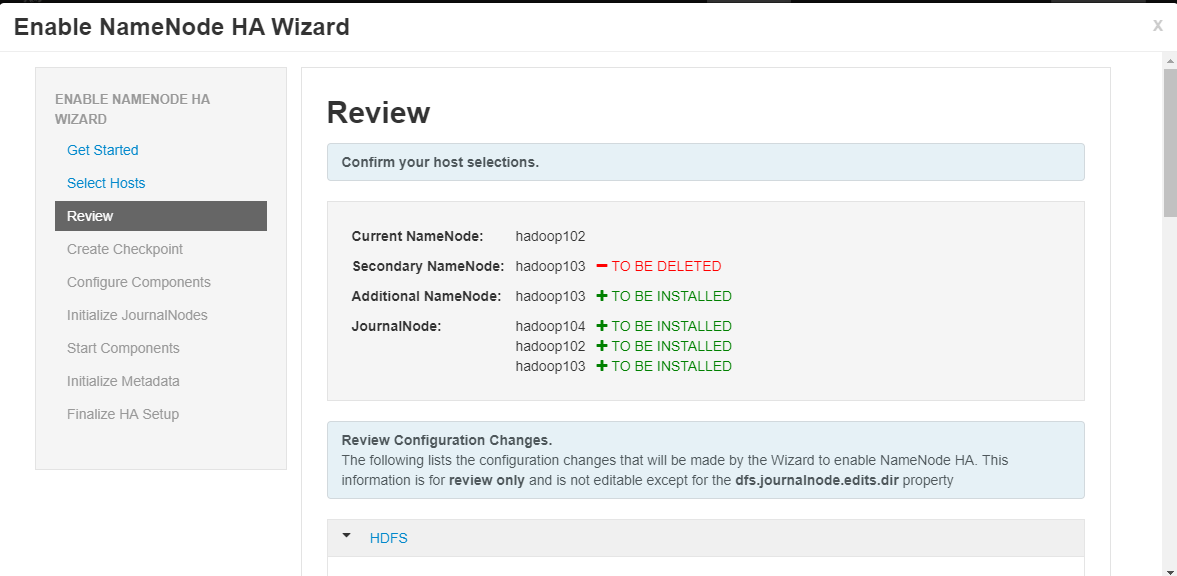
4.3.2 填写nameservice名称



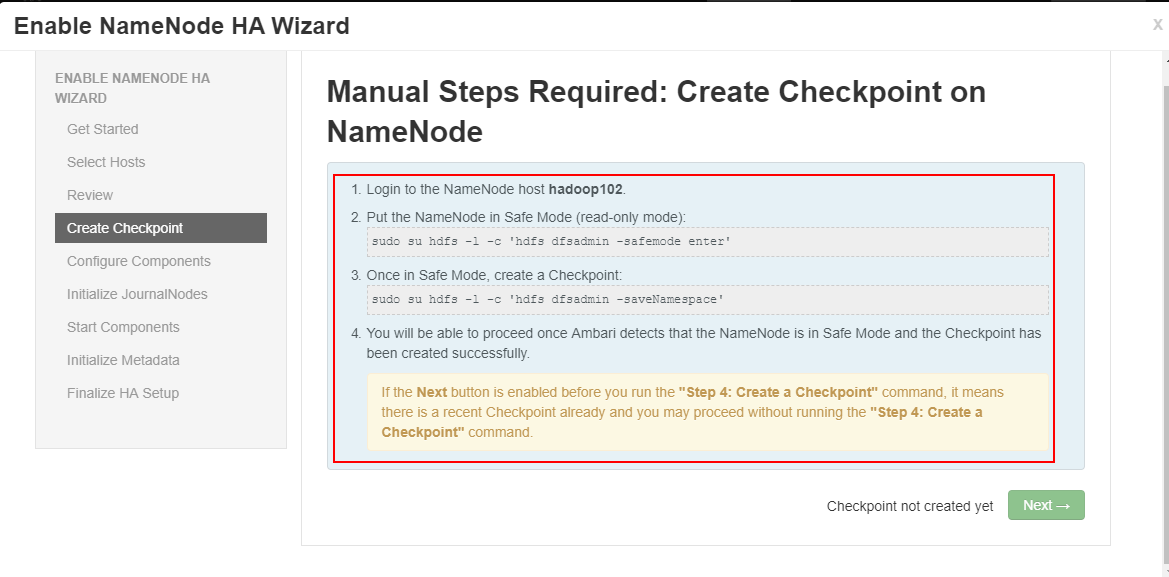
4.3.3角色分配



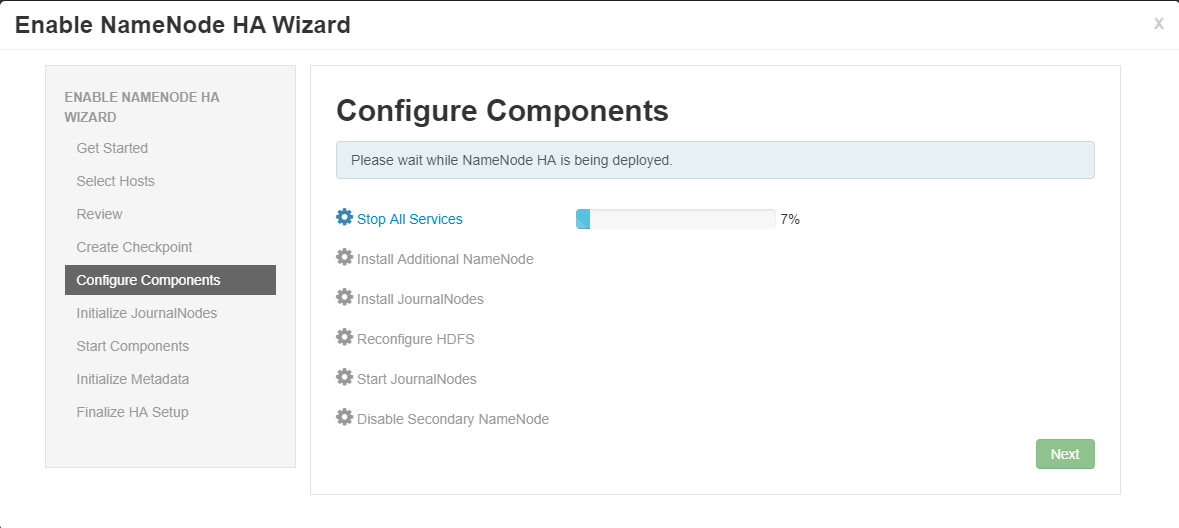
4.3.4 直接下一步



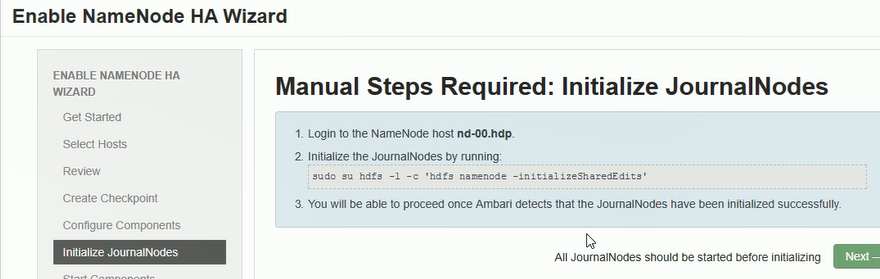
4.3.5 安装提示执行命令



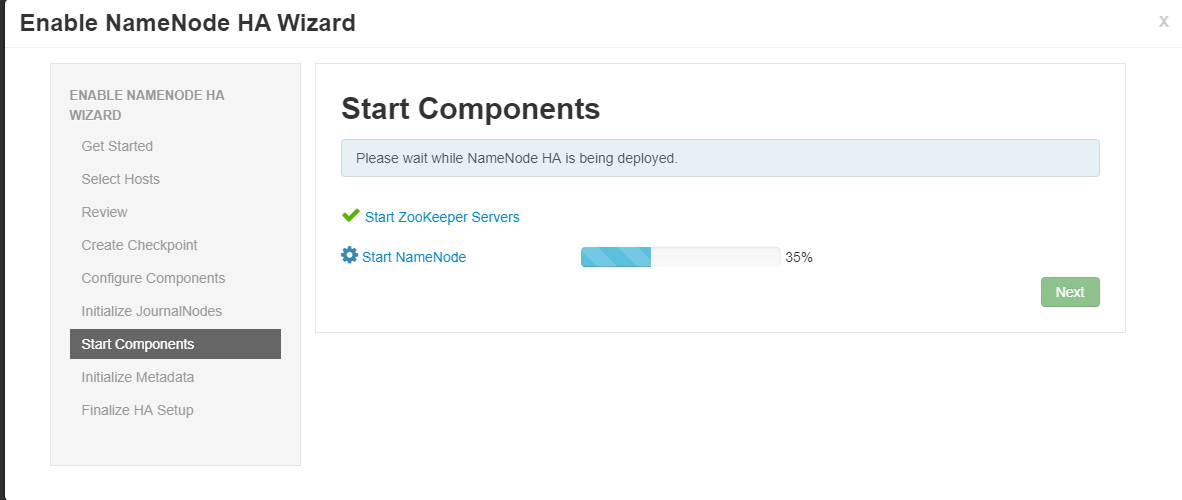
4.3.6 安装组件



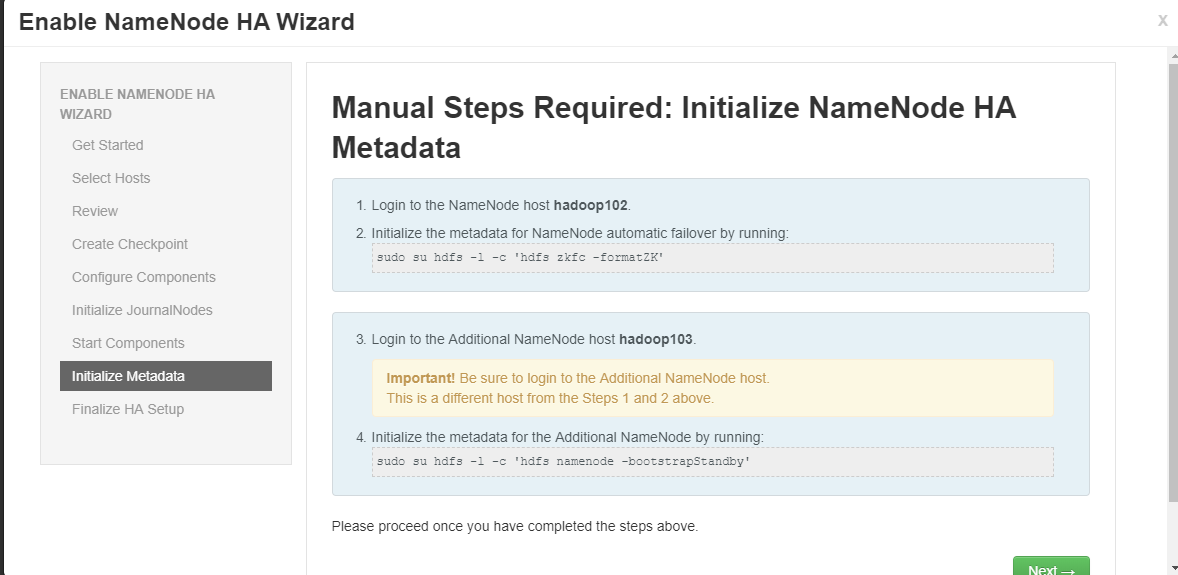
4.3.7 根据页面提示操作



4.3.8 等待启动组件



4.3.9 安装页面提示初始化NameNode HA



4.3.10 等待完成

