# 大数据环境部署文档

|  |  |  |
| --- | --- | --- |
| **编写人** | **时间** | **版本** |
| 王英楠 | 2019/08/30 | v1.0.0 |
|  |  |  |

# 1、机器配置与服务分配



# 系统要求

|  |  |
| --- | --- |
| 系统及其组件 | 系统要求 |
| 操作系统 | centos 7.2 |
| jdk环境 | 1.8版本 |
| 主机名 | 不能带符号 ‘-’ |
| 系统参数 | Ulimit 一定要设置 |
| 时间同步 | √ |

# 服务版本说明

|  |  |
| --- | --- |
| 服务 | 版本 |
| Ambari | 2.4.2 |
| Jdk | 1.8.0\_161 |
| Mysql | 5.7.16 |
| HDFS | 2.7.1 |
| YARN | 2.7.1 |
| MapReduce2 | 2.7.1 |
| Tez | 0.7.0 |
| Hive | 1.2.1 |
| HBase | 1.1.2 |
| Sqoop | 1.4.6 |
| ZooKeeper | 3.4.6 |
| Flume | 1.8.0 |
| Ambari Metrics | 0.1.0 |
| Kafka | 0.10.0 |
| Spark2 | 2.0 |
| Presto | 0.191 |
| Nginx | 1.10.2 |
| phoeinx | 4.7.0.2.5 |
| kettle | 8.2 |

# 环境准备

## 4.1配置SSH免密码登录

修改ssh配置

vim /etc/ssh/sshd\_config

#被连接的机器此属性要开启

PasswordAuthentication yes

ssh重启

systemctl restart sshd.service

ssh开机自启

systemctl enable sshd.service

主节点里root用户登录执行如下步骤

ssh-keygen

cd ~/.ssh/

cat id\_rsa.pub >>authorized\_keys

chmod -R 700 ~/.ssh

先在从节点登录root执行命令

mkdir ~/.ssh/

分发主节点里配置好的authorized\_keys到各从节点

scp /root/.ssh/authorized\_keys [root@:/root/.ssh/authorized\_keys](mailto:root@bmweb01:/root/.ssh/authorized_keys)

## 4.2创建ambari系统用户和用户组

只在主节点操作

添加ambari安装、运行用户和用户组，也可以不创建新用户，直接使用root或者系统其他账号

adduser ambari

passwd ambari

## 4.3开启NTP服务

所有集群上节点都需要操作

Centos 7 命令

yum install ntp

systemctl is-enabled ntpd

systemctl enable ntpd

systemctl start ntpd

## 4.4设置hostname、检查DNS和NSCD

设置hostname

hostnamectl set-hostname xxx

所有节点都要设置

ambari在安装时需要配置全域名，所以需要检查DNS。为了减轻DNS的负担, 建议在节点里用 Name Service Caching Daemon (NSCD)

vi /etc/hosts

每台节点里配置FQDN，如下以主节点为例

vi /etc/sysconfig/network

NETWORKING=yes

HOSTNAME=web01

## 4.5关闭防火墙

所有节点都要设置

Centos 7 命令

systemctl disable firewalld

systemctl stop firewalld

## 4.6关闭SELinux

所有节点都要设置

查看SELinux状态：

sestatus

如果SELinux status参数为enabled即为开启状态

SELinux status: enabled

临时关闭，不用重启机器：

setenforce 0

修改配置文件需要重启机器：

vi /etc/sysconfig/selinux

SELINUX=disabled

## 4.7禁用IPv6

如果Centos 7的Ipv6模块没有关闭，Ambari在启动时会先绑定到Ipv6的地址，这会导致无法通过ipv4访问。禁用的方式为：

在 /etc/sysctl.conf 中加入

net.ipv6.conf.all.disable\_ipv6=1

net.ipv6.conf.default.disable\_ipv6=1

运行 sysctl -p 应用更改。运行 netstata -anp |grep 8080 查看ambari绑定的协议和端口号是否为ipv4和8080

## 4.8数据盘挂载

### 4.8.1挂载2TB以下数据盘

首先检查是否有数据盘：

fdisk -l

看到数据盘的路径是/dev/vdb

然后给硬盘分区：

fdisk /dev/vdb

依次输入 n 、p、 1、 回车、回车、wq

这里的VDB是上面看到数据硬盘的名称，如果你不是这个需要根据你真实的盘名称替换，如果是和我一样，那就直接复制。

格式化分区

mkfs.ext3 /dev/vdb1

挂载分区

mount /dev/vdb1 /要挂载的目录

最后设置开机自动挂载：

echo '/dev/vdb1 /要挂载的目录 ext3 defaults 0 0' >> /etc/fstab

至此腾讯云的数据盘就挂在好了，如果需要验证下的话执行df -h命令即可查看。

df -h

### 4.8.2挂载2TB以上数据盘

parted用于对磁盘（或RAID磁盘）进行分区及管理，与fdisk分区工具相比，支持2TB以上的磁盘分区，并且允许调整分区的大小。

对磁盘进行分区

parted /dev/vdb mklabel gpt mkpart /dev/vdb1 ext3 1 4295G

格式化分区

mkfs.ext3 /dev/vdb1

挂载分区

mount /dev/vdb1 /要挂载的目录

最后设置开机自动挂载：

echo '/dev/vdb1 /要挂载的目录 ext3 defaults 0 0' >> /etc/fstab

# 安装服务

## 5.1 MySQL安装

安装ambari之前要先安装好MySQL

### 5.1.1安装新版mysql前，需将系统自带的mariadb-lib卸载

[root@slave mytmp]# rpm -qa|grep mariadb

mariadb-libs-5.5.44-2.el7.centos.x86\_64

[root@slave mytmp]# rpm -e --nodeps mariadb-libs-5.5.44-2.el7.centos.x86\_64

### 5.1.2解压MySQL

root@slave mytmp]# tar -zxf mysql-5.7.16-1.el7.x86\_64.rpm-bundle.tar

[root@slave mytmp]# ls

mysql-5.7.16-1.el7.x86\_64.rpm-bundle.tar mysql-community-libs-5.7.16-1.el7.x86\_64.rpm

mysql-community-client-5.7.16-1.el7.x86\_64.rpm mysql-community-libs-compat-5.7.16-1.el7.x86\_64.rpm

mysql-community-common-5.7.16-1.el7.x86\_64.rpm mysql-community-minimal-debuginfo-5.7.16-1.el7.x86\_64.rpm

mysql-community-devel-5.7.16-1.el7.x86\_64.rpm mysql-community-server-5.7.16-1.el7.x86\_64.rpm

mysql-community-embedded-5.7.16-1.el7.x86\_64.rpm mysql-community-server-minimal-5.7.16-1.el7.x86\_64.rpm

mysql-community-embedded-compat-5.7.16-1.el7.x86\_64.rpm mysql-community-test-5.7.16-1.el7.x86\_64.rpm

mysql-community-embedded-devel-5.7.16-1.el7.x86\_64.rpm

### 5.1.3 使用rpm -ivh命令依次进行安装

rpm -ivh mysql-community-common-5.7.16-1.el7.x86\_64.rpm

rpm -ivh mysql-community-libs-5.7.16-1.el7.x86\_64.rpm

rpm -ivh mysql-community-client-5.7.16-1.el7.x86\_64.rpm

安装mysql-community-server-5.7.16-1.el7.x86\_64.rpm 前需要安装libaio-0.3.107-10.el6.x86\_64.rpm

下载地址：

http://mirror.centos.org/centos/6/os/x86\_64/Packages/libaio-0.3.107-10.el6.x86\_64.rpm

安装libaio库：

rpm -ivh libaio-0.3.107-10.el6.x86\_64.rpm（若在有网情况下可执行yum install libaio）

安装mysql-community-server：

rpm -ivh mysql-community-server-5.7.16-1.el7.x86\_64.rpm（此步骤若报错，可参考http://blog.csdn.net/yoon0205/article/details/50605584）

### 5.1.4 初始化数据库

// 指定datadir, 执行后会生成~/.mysql\_secret密码文件

[root@slave mytmp]# mysql\_install\_db --datadir=/var/lib/m

// 初始化，执行生会在/var/log/mysqld.log生成随机密码

[root@slave mytmp]# mysqld --initialize

### 5.1.5 更改mysql数据库目录的所属用户及其所属组，并启动mysql数据库

[root@slave mytmp]# chown mysql:mysql /var/lib/mysql -R

[root@slave mytmp]# systemctl start mysqld.service

### 5.1.6 登录到mysql，更改root用户的密码

// password 通过 cat ~/.mysql\_secret 命令可以查看初始密码

[root@slave mytmp]# mysql -uroot -p

Enter password:

mysql> set password=password('1234');

### 5.1.7 创建用户，及作权限分配

安装完成后创建ambari数据库及用户，登录root用户执行下面语句：

create database ambari character set utf8 ;

CREATE USER 'ambari'@'%'IDENTIFIED BY 'Ambari-123';

GRANT ALL PRIVILEGES ON \*.\* TO 'ambari'@'%';

FLUSH PRIVILEGES;

如果要安装Hive，再创建Hive数据库和用户 再执行下面的语句：

create database hive character set utf8 ;

CREATE USER 'hive'@'%'IDENTIFIED BY 'Hive-123';

GRANT ALL PRIVILEGES ON \*.\* TO 'hive'@'%';

FLUSH PRIVILEGES;

### 5.1.8 远程登陆授权

mysql> grant all privileges on \*.\* to 'root'@'%' identified by '1234' with grant option;

mysql> flush privileges;

### 5.1.9 设置mysql开机启动

// 检查是否已经是开机启动

systemctl list-unit-files | grep mysqld

// 开机启动

systemctl enable mysqld.service

### 5.1.10 安装mysql jdbc 驱动

yum install mysql-connector-java

## 5.2 Ambari安装

### 5.2.1本地源制作

制作本地源只需在主节点上进行即可

相关准备工作

#### 5.2.1.1安装 Apache HTTP 服务器

安装HTTP 服务器，允许 http 服务通过防火墙(永久,若防火墙已关闭，则不设置)

yum install httpd

firewall-cmd --add-service=http

firewall-cmd --permanent --add-service=http

添加 Apache 服务到系统层使其随系统自动启动

systemctl start httpd.service

systemctl enable httpd.service

#### 5.2.1.2 安装本地源制作相关工具

yum install yum-utils createrepo

#### 5.2.1.3 下载安装资源

Ambari 2.4.2 资源下载（官网http://public-repo-1.hortonworks.com/）

在httpd网站根目录,默认是即/var/www/html/，创建目录ambari，

并且将下载的压缩包解压到/var/www/html/ambari目录

cd /var/www/html/

mkdir ambari

cd /var/www/html/ambari/

tar -zxvf ambari-2.4.2.0-centos7.tar.gz

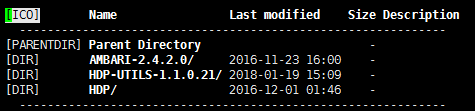
tar -zxvf HDP-2.5.3.0-centos7-rpm.tar.gz

tar -zxvf HDP-UTILS-1.1.0.21-centos7.tar.gz

验证httd网站是否可用，可以使用links 命令或者浏览器直接访问下面的地址：

links http://10.135.79.152/ambari/

结果如下：



#### 5.2.4配置ambari、HDP、HDP-UTILS的本地源

首先下载上面资源列表中的相应repo文件，修改其中的URL为本地的地址，相关配置如下：

ambari.repo

#VERSION\_NUMBER=2.4.2.0-136

[Updates-ambari-2.4.2.0]

name=ambari-2.4.2.0 - Updates

baseurl=http://10.135.79.152/ambari/AMBARI-2.4.2.0/centos7/2.4.2.0-136

gpgcheck=1

gpgkey=http://10.135.79.152/ambari/AMBARI-2.4.2.0/centos7/2.4.2.0-136/RPM-GPG-KEY/RPM-GPG-KEY-Jenkins

enabled=1

priority=1

hdp.repo

#VERSION\_NUMBER=2.5.3.0-37

[HDP-2.5.3.0]

name=HDP Version - HDP-2.5.3.0

baseurl=http://10.135.79.152/ambari/HDP/centos7/

gpgcheck=1

gpgkey=http://10.135.79.152/ambari/HDP/centos7/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://10.135.79.152/ambari/HDP-UTILS-1.1.0.21/

gpgcheck=1

gpgkey=http://10.135.79.152/ambari/HDP-UTILS-1.1.0.21/RPM-GPG-KEY/RPM-GPG-KEY-Jenkins

enabled=1

priority=1

将上面的修改过的源放到/etc/yum.repos.d/下面后执行以下命令

yum clean all

yum list update

yum makecache

yum repolist

### 5.2.2进行安装Ambari

#### 5.2.2.1安装Ambari

yum install ambari-server

#### 5.2.2.2配置Ambari

ambari-server setup

下面是配置执行流程，按照提示操作

1.检查SELinux是否关闭，如果关闭不用操作

Using python /usr/bin/python

Setup ambari-server

Checking SELinux...

SELinux status is 'disabled'

2.提示是否自定义设置。输入：y

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

3.ambari-server 账号。输入：ambari

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

Adjusting ambari-server permissions and ownership...

4.检查防火墙，如果关闭则不用操作

Checking firewall status...

Redirecting to /bin/systemctl status iptables.service

5.设置JDK。输入：3

Checking JDK...

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

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

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

[] Custom JDK

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

Enter choice (1): 3

6.如果上面选择3自定义JDK,则需要设置JAVA\_HOME。输入：/opt/java/jdk1.8.0\_151

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/java/jdk1.8.0\_151

Validating JDK on Ambari Server...done.

Completing setup...

7.数据库配置。选择：y

Configuring database...

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

8.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

9.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。

Hostname (localhost):

Port ():

Database name (ambari):

Username (ambari):

Enter Database Password (Ambari-123):

10.提示必须安装MySQL JDBC，回车结束ambari配置

WARNING: Before starting Ambari Server, you must copy the MySQL JDBC driver JAR file to /usr/share/java.

Press <enter> to continue.

11.将Ambari数据库脚本导入到数据库

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

用Ambari用户（上面设置的用户）登录mysql

mysql -u ambari -p

use ambari

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

#### 5.2.2.3启动Amabri

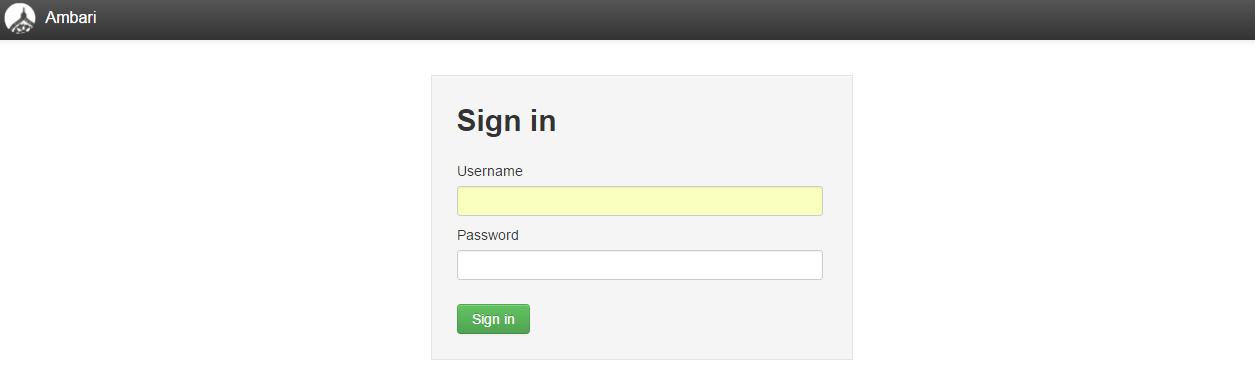
执行启动命令，启动Ambari服务

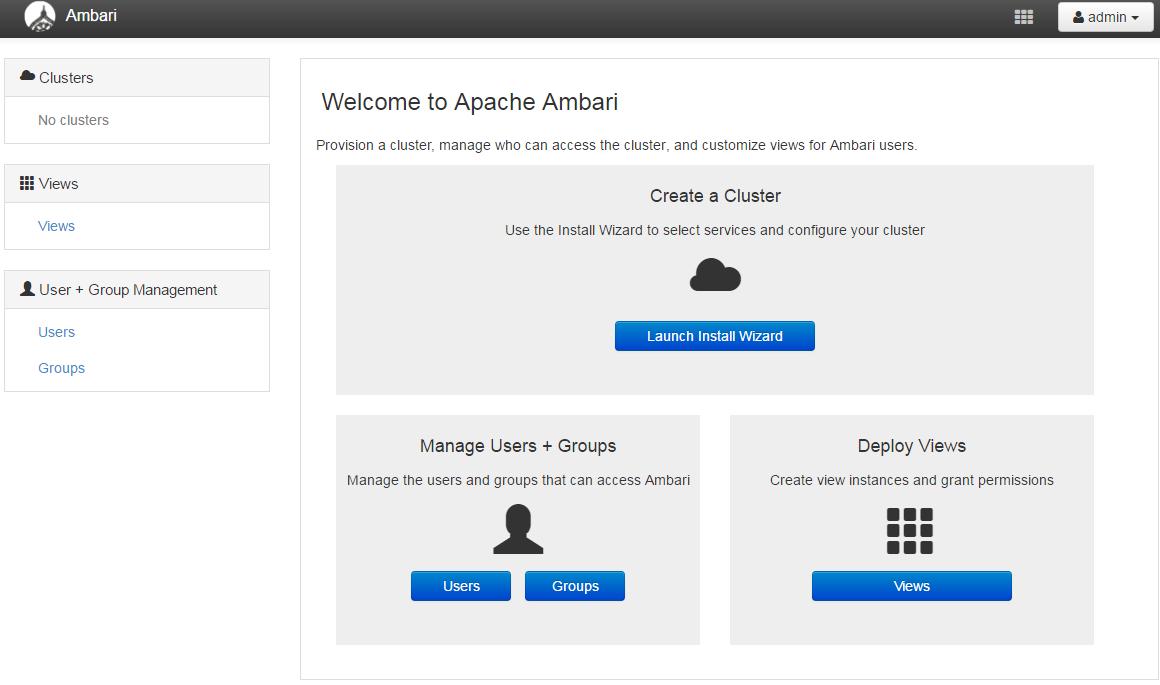
ambari-server start

成功启动后在浏览器输入Ambari地址：

<http://172.16.1.101:8080/>

出现登录界面，默认管理员账户登录， 账户：admin 密码：admin

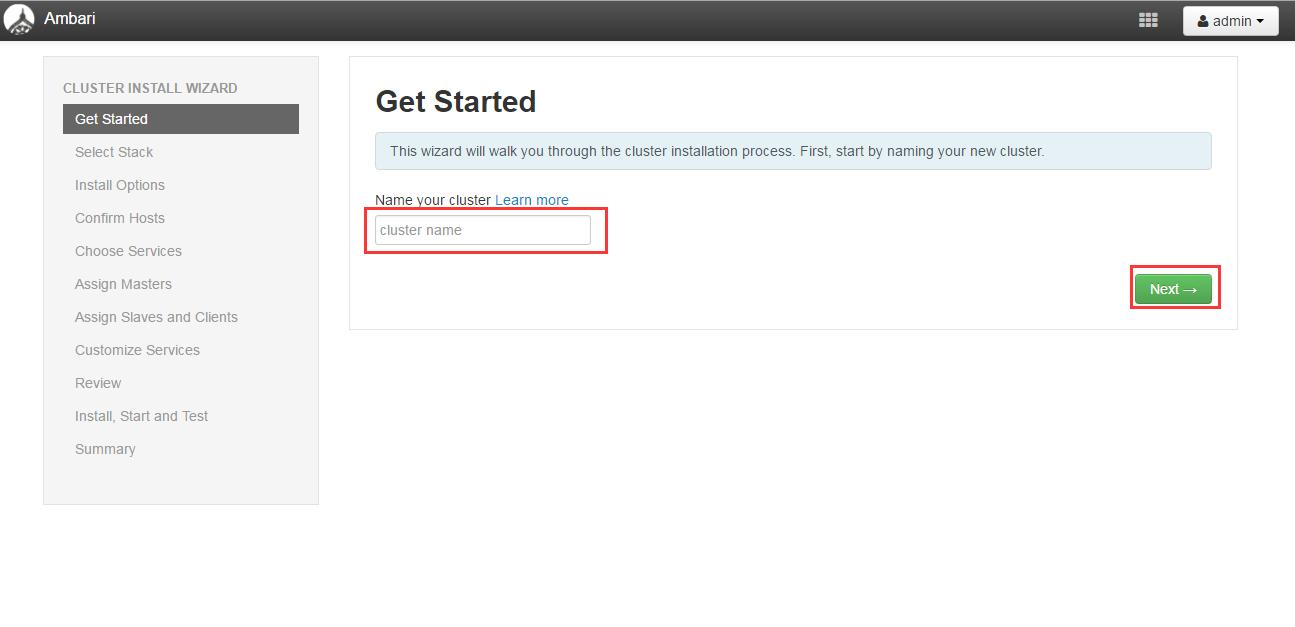


登录成功后出现下面的界面，至此Ambari的安装成功   


#### 5.2.2.4配置集群

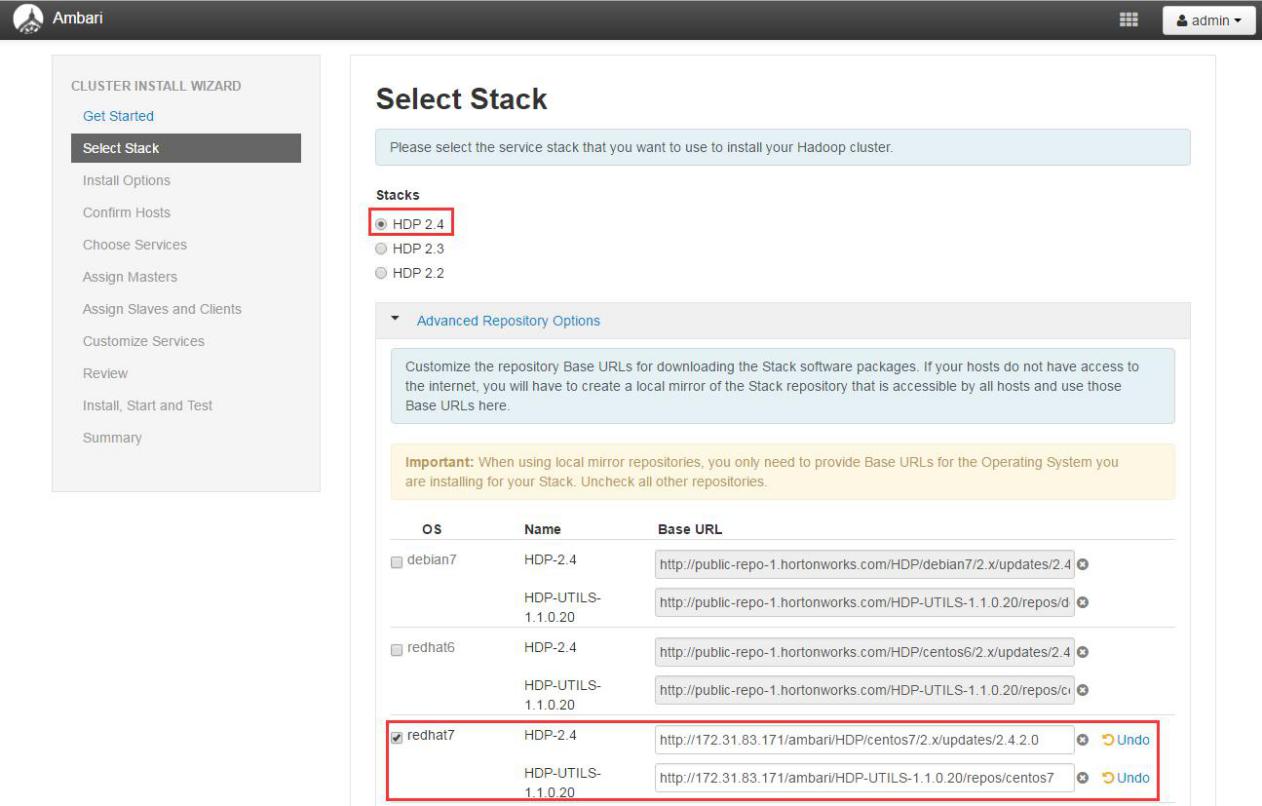
点击上面登录成功页面的Launch Install Wizard 按钮进行集群配置

##### 5.2.2.4.1设置集群名称

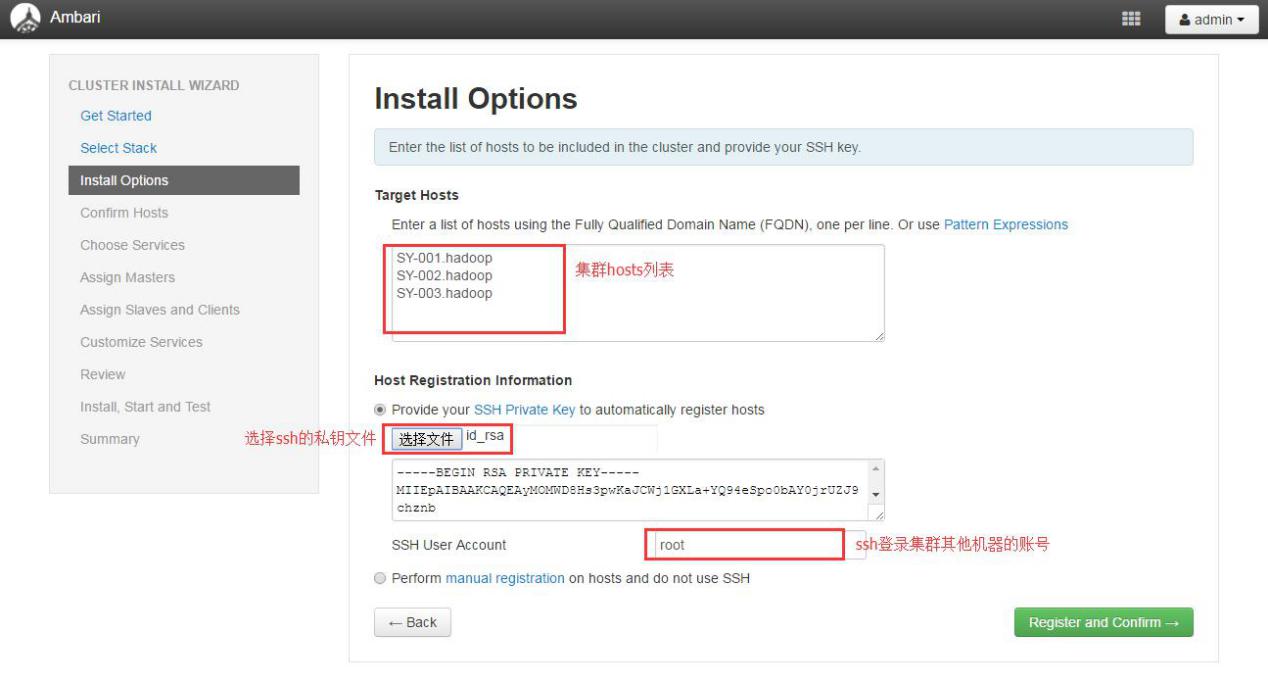


##### 5.2.2.4.2设置HDP 安装源

选择HDP2.5 ,并且设置Advanced Repository Options 的信息，本次使用本地源，所以修改对用系统的安装源为本地源地址。

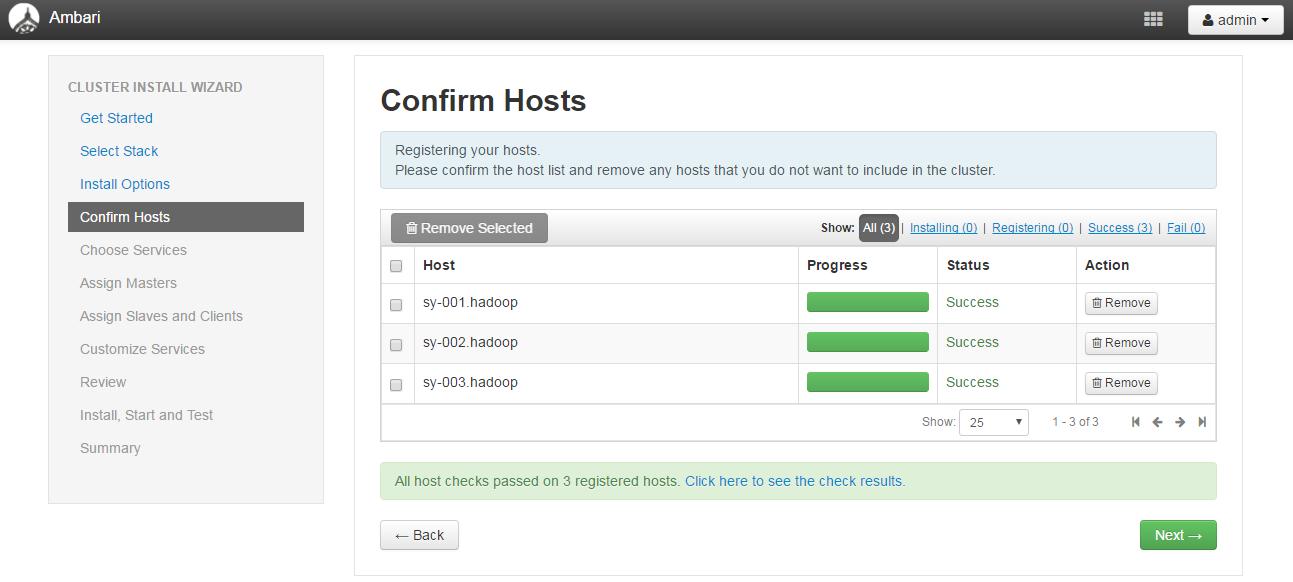


##### 5.2.2.4.3设置集群机器

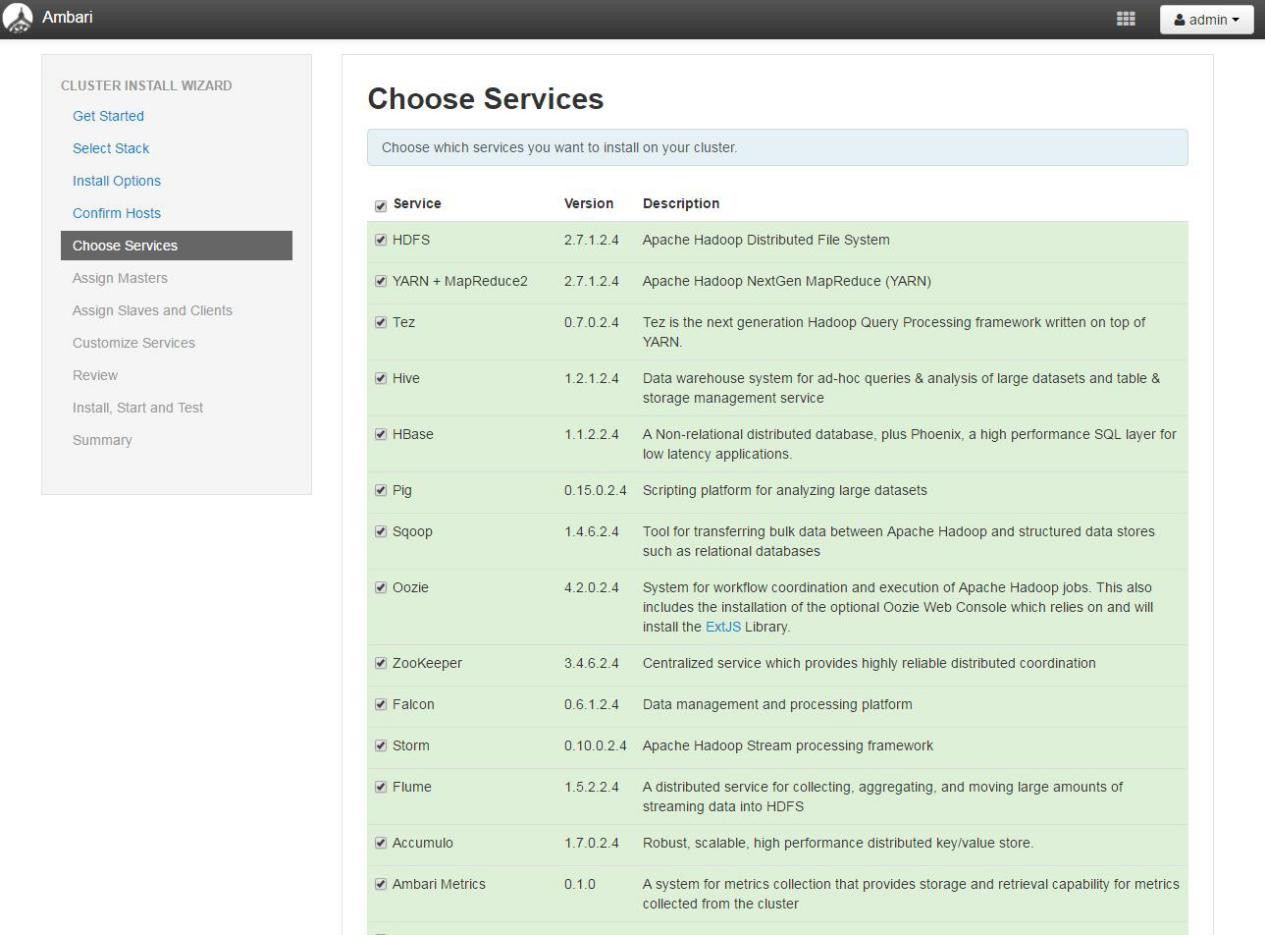


##### 5.2.2.4.4 Host 确认

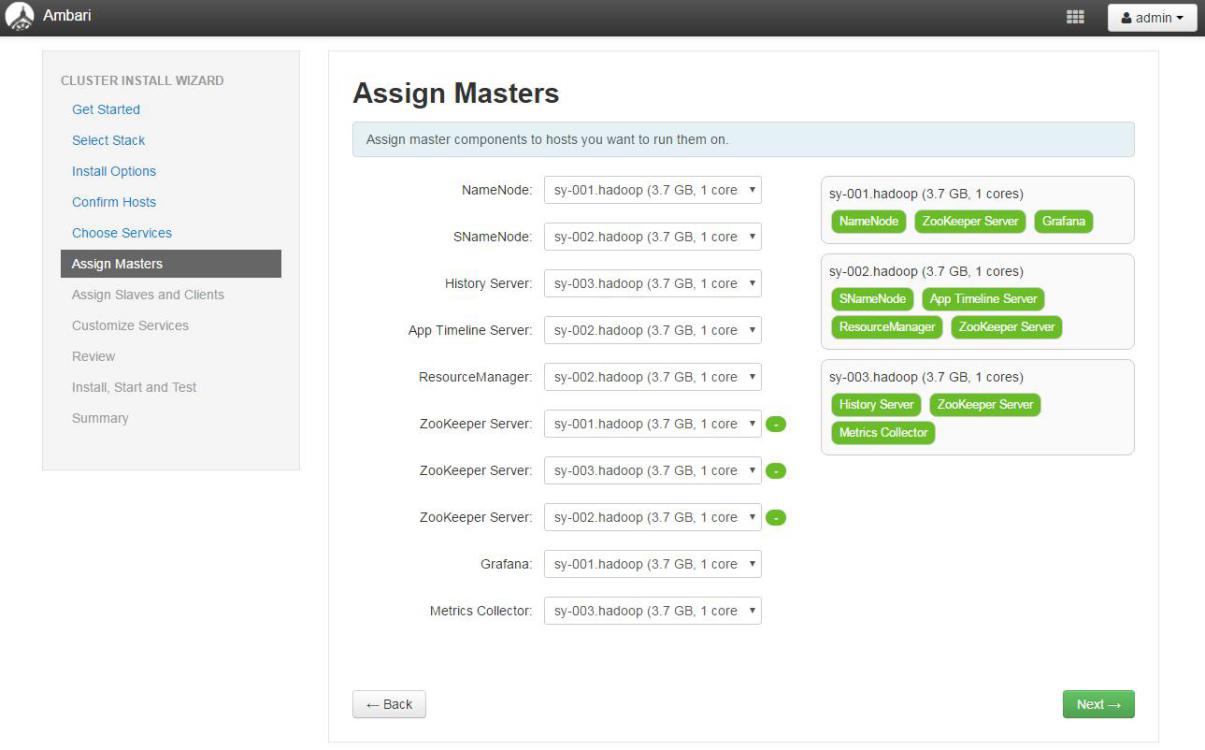
确认前面配置集群中hosts列表 中的机器是否都可用，也可以移除相关机器，集群中机器Success后进行下一步操作。



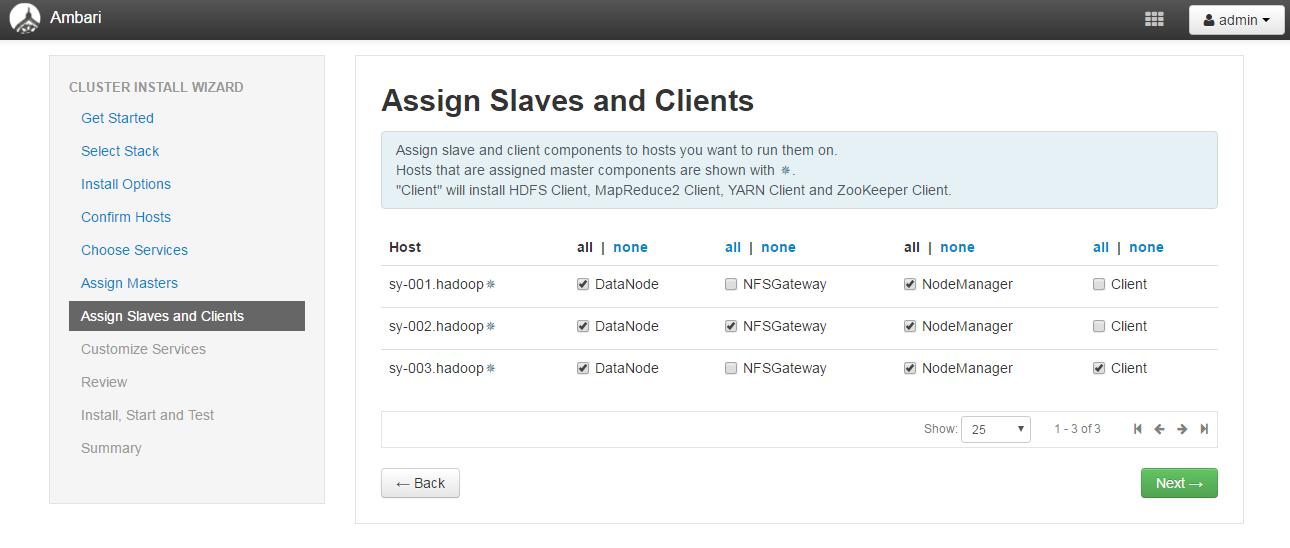
##### 5.2.2.4.5选择要安装的服务



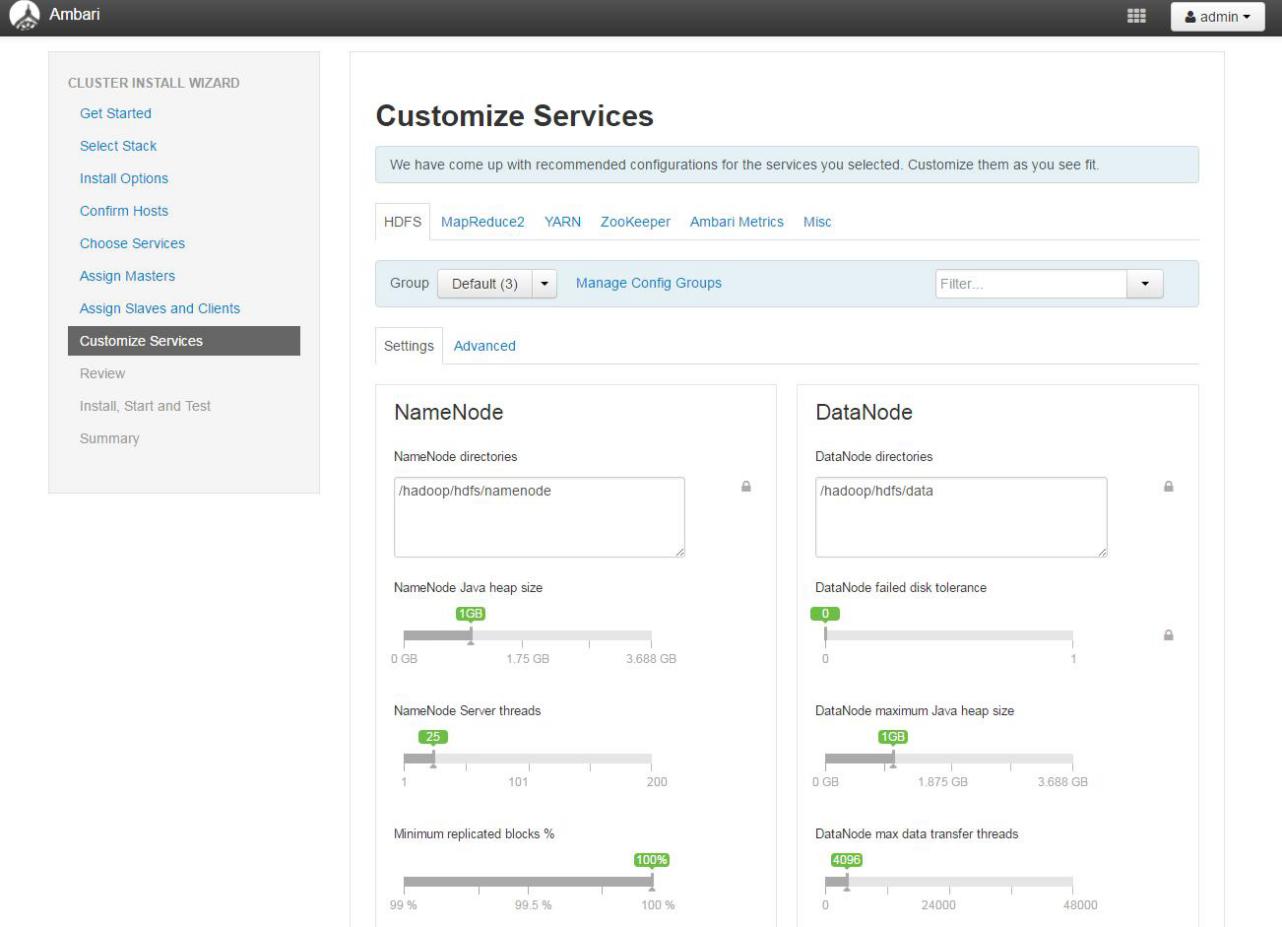
##### 5.2.2.4.6各个服务Master配置



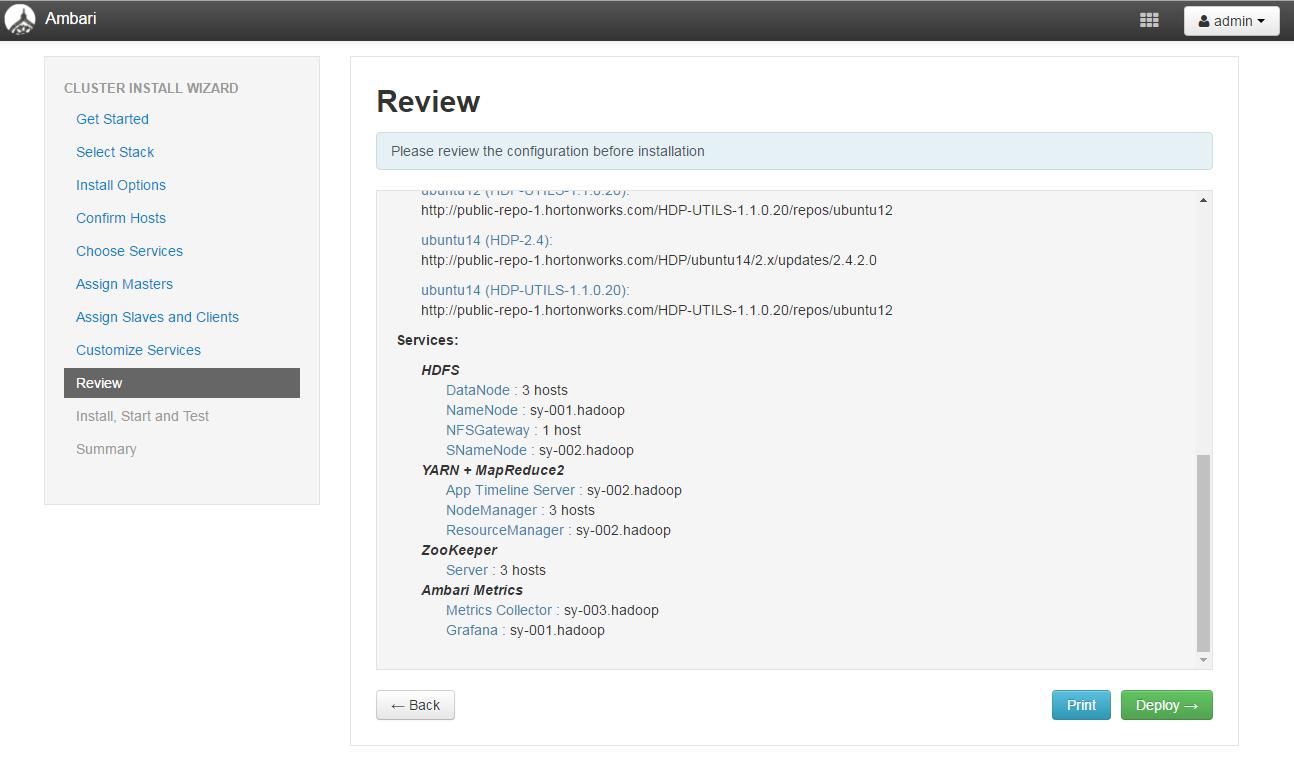
##### 5.2.2.4.7服务的Slaves 和 Clients节配置



##### 5.2.2.4.8服务自定义配置

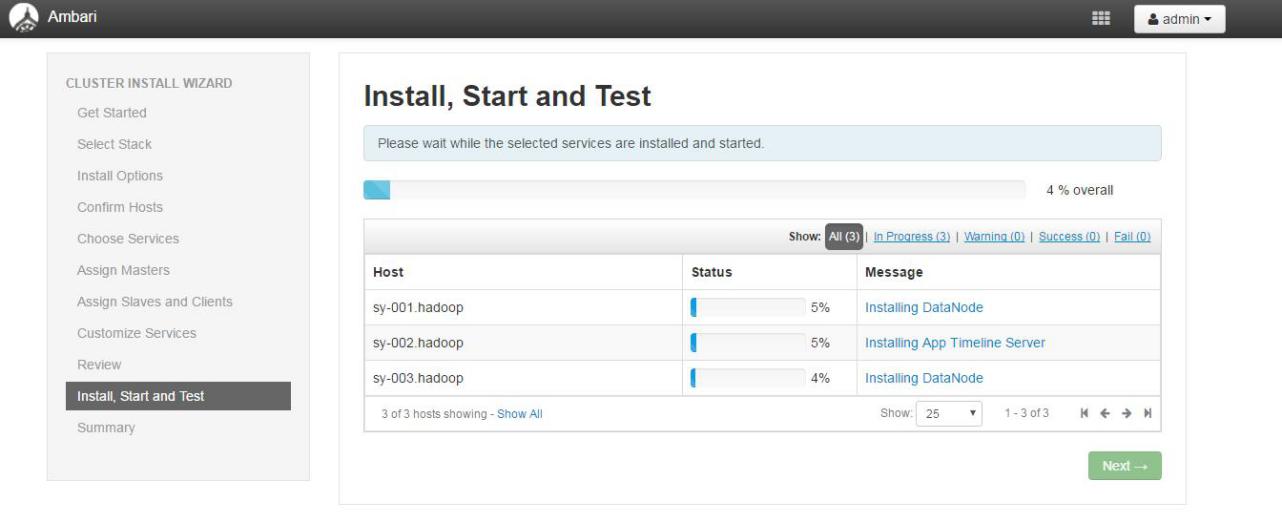


##### 5.2.2.4.9显示配置信息



##### 5.2.2.4.10开始安装

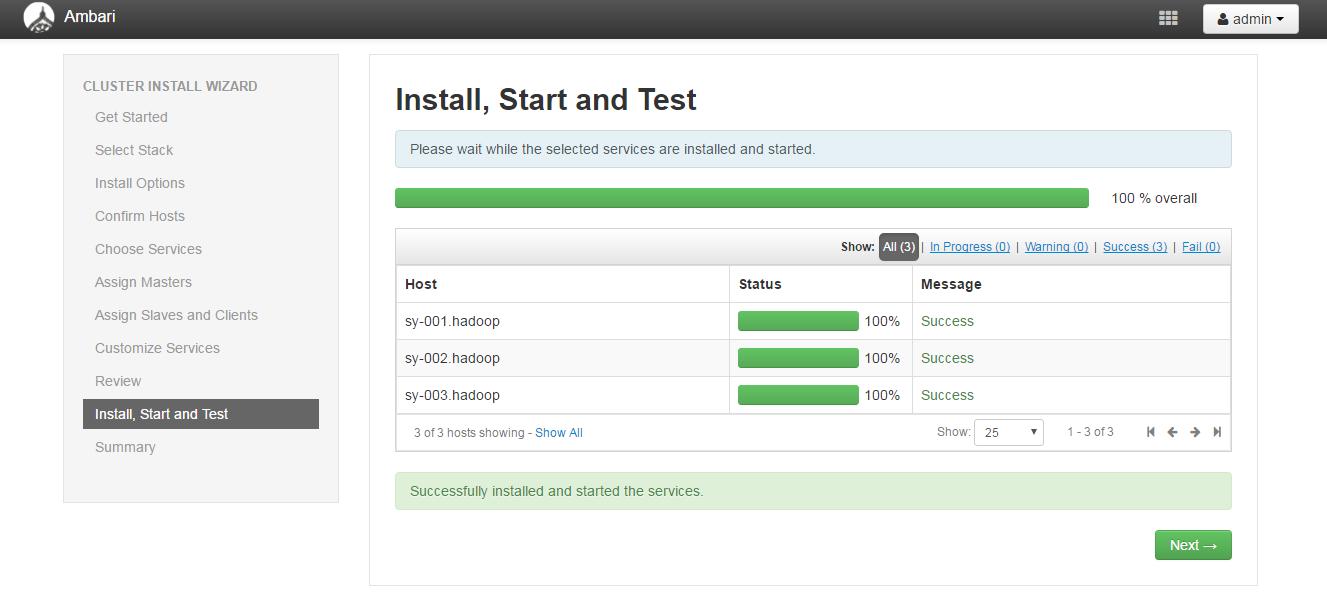
安装各个服务，并且完成安装后会启动相关服务，安装过程比较长，如果中途出现错误，请根据具体提示或者log进行操作。



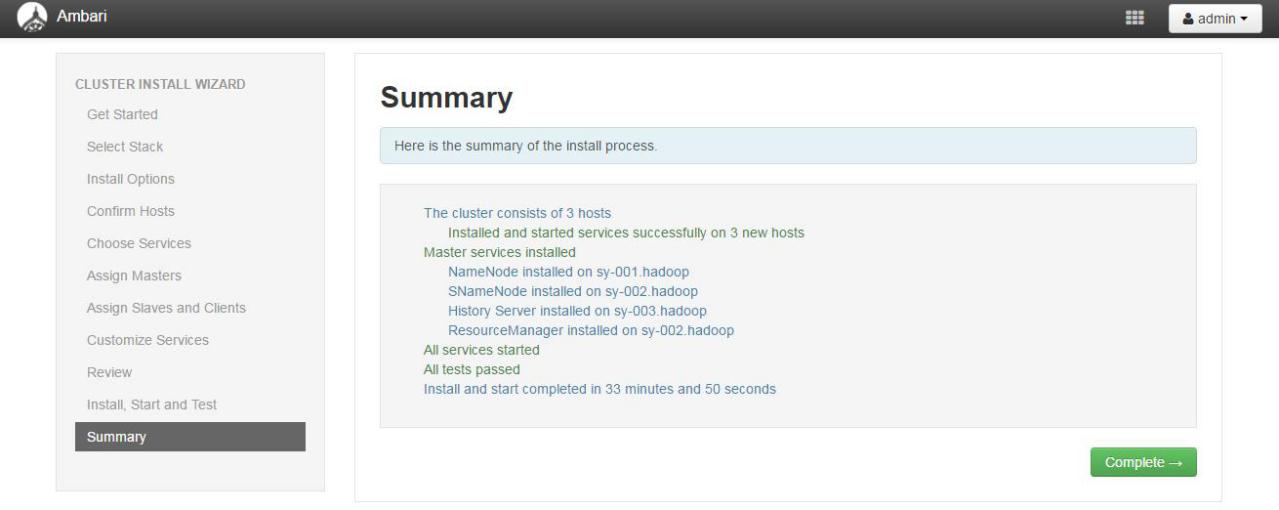
安装过程可以随时查看每个节点的安装进度及日志

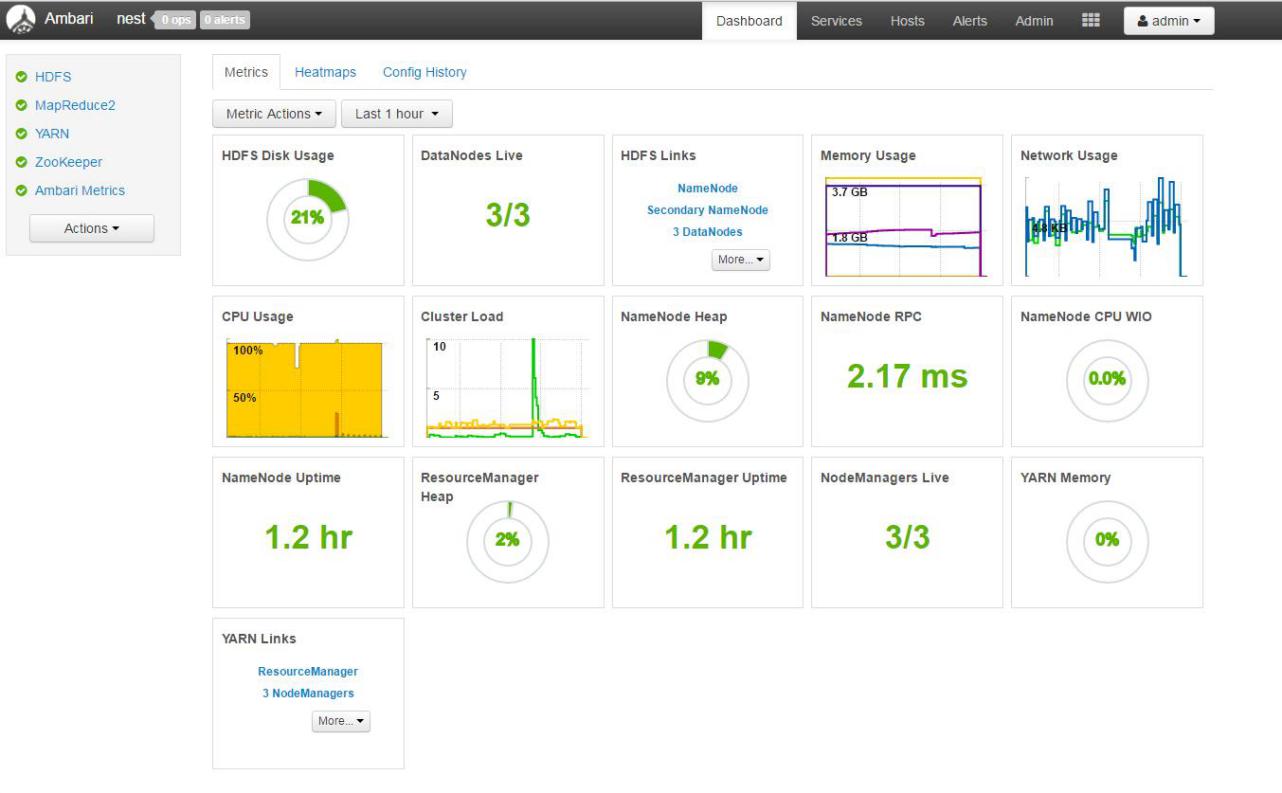


全部安装成功界面如下



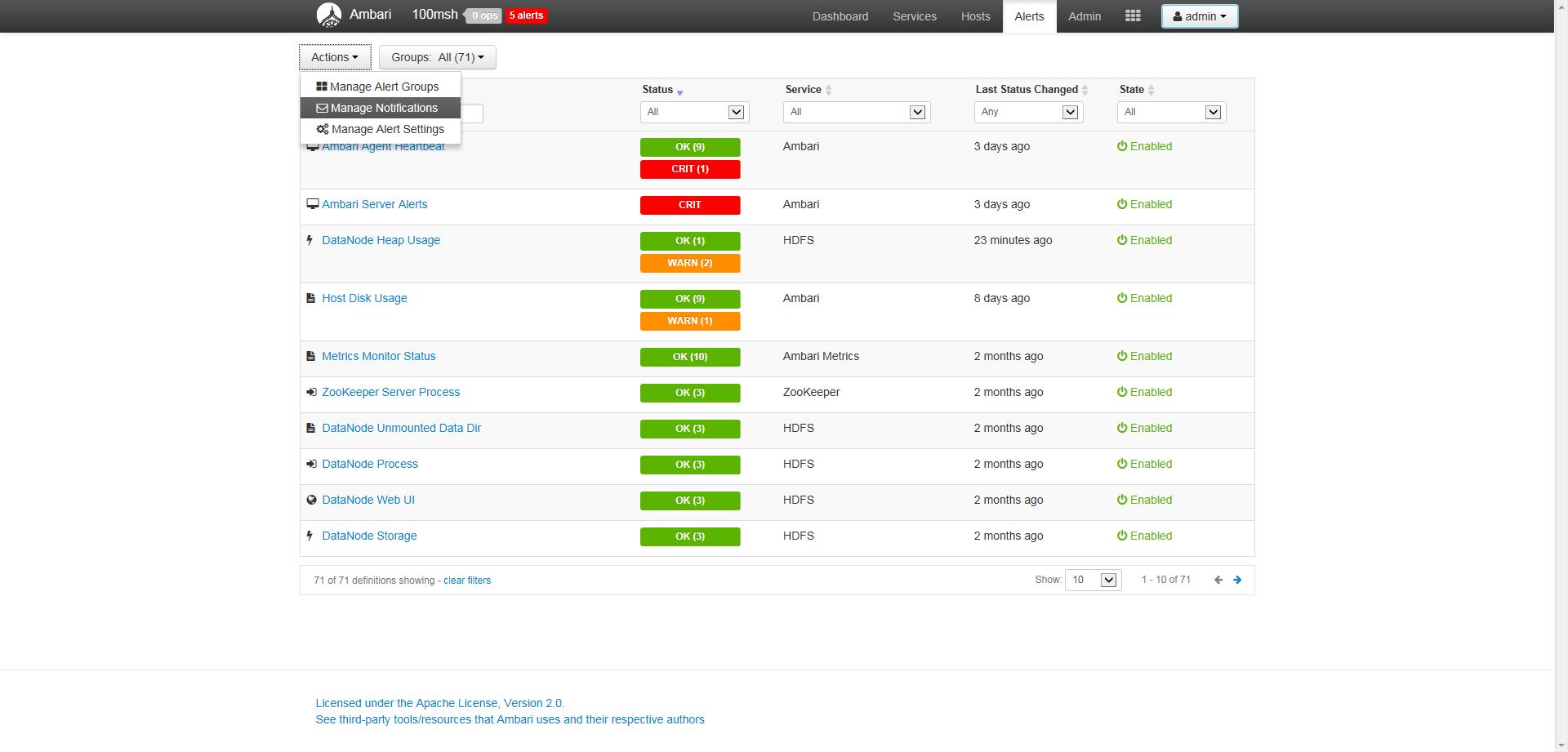
##### 5.2.2.4.11安装完成

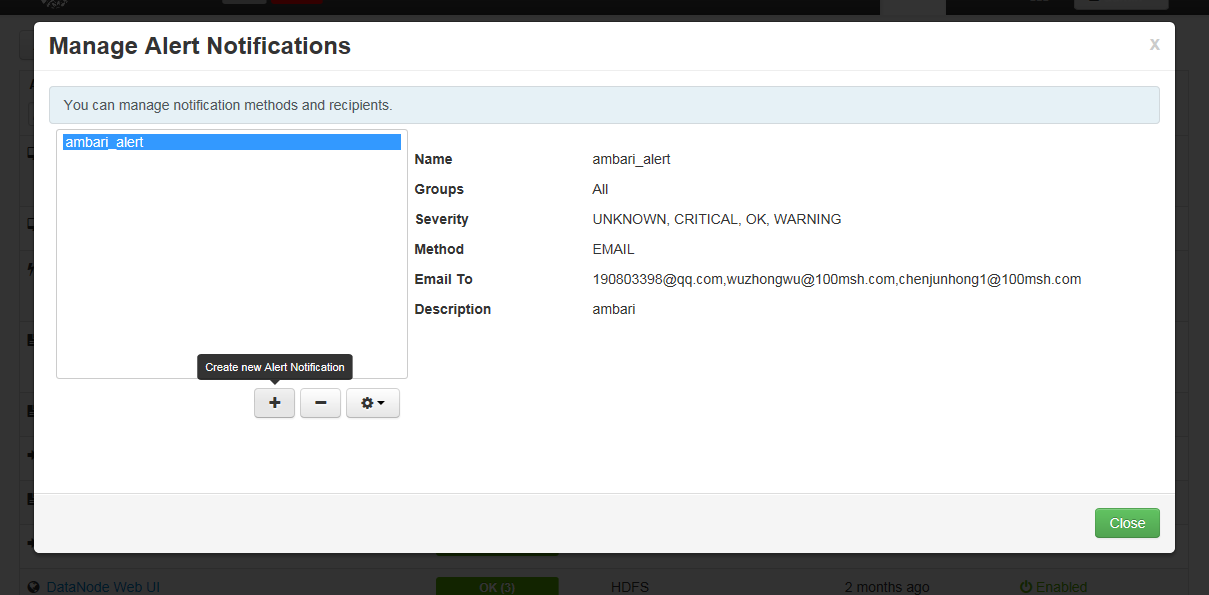




#### 5.2.2.5 Ambari邮件告警配置

配置Manage Notifications



创建告警通知

配置告警



163邮箱授权码获取

1、打开设置àPOP3/SMTP/IMAP



2、开启POP3/SMTP/IMAP服务



3、获取授权密码



定义发送邮件格式

1、下载 alert-templates.xml：<https://github.com/apache/ambari/blob/branch-2.1/ambari-server/src/main/resources/alert-templates.xml>

2、  将alert-templates.xml移动到 /var/lib/ambari-server/resources/

3、 编辑/etc/ambari-server/conf/ambari.properties文件

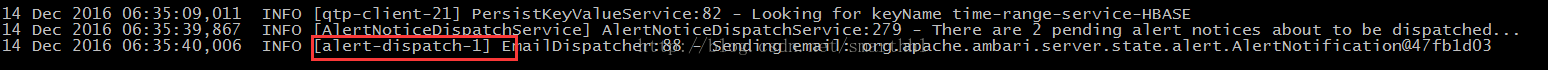
在文件最后一行增加 alerts.template.file=/var/lib/ambari-server/resources/alert-templates.xml

4、 重启ambariserver

ambari-server restart

5、 查看ambariserver日志

tail -f /var/log/ambari-server/ambari-server.log



## 5.3 Nginx安装

### 5.3.1下载Nginx及相关组件

进入用户目录下载程序

IMG_256

下载相关组件

wget http://nginx.org/download/nginx-1.10.2.tar.gz

wget http://www.openssl.org/source/openssl-fips-2.0.10.tar.gz

wget http://zlib.net/zlib-1.2.11.tar.gz

wget ftp://ftp.csx.cam.ac.uk/pub/software/programming/pcre/pcre-8.40.tar.gz

安装c++编译环境

yum -y install gcc-c++

### 5.3.2安装Nginx及相关组件

openssl安装

tar zxvf openssl-fips-2.0.10.tar.gz

cd openssl-fips-2.0.10

./config && make && make install

pcre安装

tar zxvf pcre-8.40.tar.gz

cd pcre-8.40

./configure && make && make install

zlib安装

tar zxvf zlib-1.2.11.tar.gz

cd zlib-1.2.11

./configure && make && make install

nginx安装

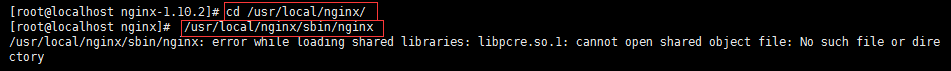
tar zxvf nginx-1.10.2.tar.gz

cd nginx-1.10.2

./configure && make && make install

### 5.3.3 Nginx启动

进入nginx目录并执行启动命令



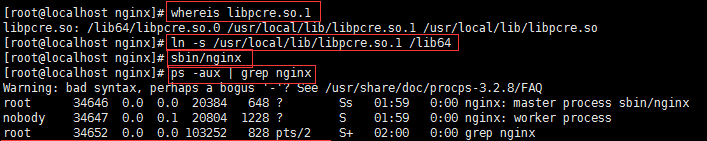
如果报错了，error while loading shared libraries: libpcre.so.1: cannot open shared object file: No such file or directory，按照下面方式解决

1.用whereis libpcre.so.1命令找到libpcre.so.1在哪里

2.用ln -s /usr/local/lib/libpcre.so.1 /lib64命令做个软连接就可以了

3.用sbin/nginx启动Nginx

4.用ps -aux | grep nginx查看状态



打开浏览器输入localhost会看到下图，说明nginx启动成功



## 5.4安装NFS

### 5.4.1 安装

yum -y install nfs-utils rpcbind

### 5.4.2 服务端配置

在NFS服务端上创建共享目录/data/test并设置权限

mkdir -p /data/test

chmod 666 /data/test/

编辑export文件

vim /etc/exports

/data/lys 172.16.1.101/24(rw,no\_root\_squash,no\_all\_squash,sync)

配置生效

exportfs -r

启动rpcbind、nfs服务

service rpcbind start

service nfs start

### 5.4.3客户端配置

安装nfs-utils客户端

yum -y install nfs-utils

创建挂载目录

mkdir /test

查看服务器抛出的共享目录信息

showmount -e 172.16.1.101

为了提高NFS的稳定性，使用TCP协议挂载，NFS默认用UDP协议

mount -t nfs 172.16.1.101:/data/test /test-o proto=tcp -o nolock

### 5.4.4 目录所有者问题解决

服务端与客户端访问该共享目录的用户的uid，gid必须相同，否则会出现其中一台机器给该目录授权a用户后，另一台机器显示该目录所有者为其它用户或者显示数字。

需手动修改uid和gid

## 5.5 安装httpd (配置正向代理)

### 5.5.1 安装

安装HTTP 服务器，允许 http 服务通过防火墙(永久,若防火墙已关闭，则不设置)

yum install httpd

firewall-cmd --add-service=http

firewall-cmd --permanent --add-service=http

firewall-cmd --add-port=9090/tcp --permanent

firewall-cmd --reload

添加 Apache 服务到系统层使其随系统自动启动

systemctl start httpd.service

systemctl enable httpd.service

### 5.5.2 正向代理服务端配置

vim /etc/httpd/conf/httpd.conf

添加以下内容

Listen 9090

#正向代理设置

ProxyRequests On

ProxyVia On

<Proxy \*>

Order deny,allow

Deny from all

Allow from 10.135.79.152 //放通客户端ip

Allow from 10.135.95.118

Allow from 10.135.80.73

</Proxy>

重启httpd服务

/bin/systemctl restart httpd.service

### 5.5.3 客户端配置

vim /etc/profile

添加以下内容

export http\_proxy=http://服务端ip:服务端端口

source /etc/profile