

---

# 使用 Ambari 部署 HDP 平台

Name : 曲中岭

Email: zlingqu@126.com

Q Q : 441869115

日期 : 2021 年 08 月

---

# 目录

<b>1</b>	<b>介绍.....</b>	<b>4</b>
1.1	AMBARI 介绍 .....	4
1.2	规划.....	4
1.3	ANSIBLE 脚本 .....	4
<b>2</b>	<b>AMBARI-SERVER 初始化.....</b>	<b>5</b>
2.1	准备相关包.....	5
2.2	安装 NGINX 服务.....	7
2.3	添加 YUM 私有 REPO.....	8
2.4	设置免密登录 .....	9
<b>3</b>	<b>AMBARI-AGENT 初始化.....</b>	<b>10</b>
3.1	修改主机名 .....	10
3.2	HOSTS 配置 .....	10
3.3	关闭防火墙 .....	10
3.4	关闭 SELINUX.....	10
3.5	配置 LIMITS 参数 .....	10
3.6	禁用交换分区 .....	11
3.7	禁用大内存页功能 .....	11
3.8	时间同步 .....	11
3.9	安装 JDK.....	11
3.10	添加 AMBARI 源 .....	11
3.11	安装 AMBARI-AGENT .....	11
<b>4</b>	<b>AMBARI 部署 .....</b>	<b>12</b>
4.1	MYSQL 准备 .....	12
4.2	部署 AMBARI-SERVER.....	12
4.2.1	安装服务 .....	12
4.2.2	数据库驱动配置 .....	12
4.2.3	配置初始化 .....	13
4.2.4	数据库初始化 .....	14
4.2.5	启动服务 .....	15
<b>5</b>	<b>部署 HADOOP 集群.....</b>	<b>16</b>
5.1	登录 AMBARI 界面 .....	16
5.2	启动 AMBARI 部署向导 .....	16
5.3	选择版本 .....	17
5.4	安装选项.....	19
5.5	确认主机 .....	19
5.6	选择服务 .....	23
5.7	分配主要组件 .....	24
5.8	从属组件配置 .....	25

---

5.9	定制服务 .....	25
5.10	REVIEW 概要 .....	28
5.11	安装、启动和检测 .....	29
5.12	汇总 .....	30
<b>6</b>	<b>添加服务 .....</b>	<b>32</b>
6.1	添加 YARN+MAPREDUCE2 服务 .....	32
6.2	添加 HIVE 服务 .....	34
6.3	添加 SPARK 服务 .....	35
6.4	添加 SQOOP 服务 .....	36
6.5	添加 OOZIE 服务 .....	36
6.6	添加 KAFKA 服务 .....	37
6.7	概览 .....	38
<b>7</b>	<b>排错 .....</b>	<b>40</b>
7.1	HIVE 安装报警 .....	40
7.2	KAFKA 安装报警 .....	40

# 1 介绍

## 1.1 ambari 介绍

Ambari 是 Apache Software Foundation 中的一个顶级项目，它可以创建、管理、监视 Hadoop 整个生态圈（例如 Hive、HBase、Sqoop、ZooKeeper 等）的集群，使得 Hadoop 以及相关的大数据软件更容易使用。

Ambari 是一个分布式架构的软件，由 Ambari Server 和 Ambari Agent 两部分组成，用户可通过 Ambari Server 通知 Ambari Agent 安装对应的软件，Ambari Agent 会定时地发送各个机器每个软件模块的状态给 Ambari Server，最终这些状态信息会呈现在 Ambari 的 GUI 界面上，方便用户了解到集群的各种状态，并进行相应的维护。

## 1.2 规划

序号	IP	hostname	OS	数据盘	角色
0	10.12.19.31	devops-10-12-19-31	CentOS7.8		ambari-server,yum-repo
1	10.11.16.1	10-11-16-1.dm-ai.com	CentOS7.8	/data	ambari-agent,hadoop
2	10.11.16.2	10-11-16-2.dm-ai.com	CentOS7.8	/data	ambari-agent,hadoop
3	10.11.16.3	10-11-16-3.dm-ai.com	CentOS7.8	/data	ambari-agent,hadoop
4	10.11.16.4	10-11-16-4.dm-ai.com	CentOS7.8	/data	ambari-agent,hadoop
5	10.11.16.5	10-11-16-5.dm-ai.com	CentOS7.8	/data	ambari-agent,hadoop
6	10.11.16.6	10-11-16-6.dm-ai.com	CentOS7.8	/data	ambari-agent,hadoop
7	10.11.16.7	10-11-16-7.dm-ai.com	CentOS7.8	/data	ambari-agent,hadoop
8	10.11.16.8	10-11-16-8.dm-ai.com	CentOS7.8	/data	ambari-agent,hadoop
9	10.11.16.9	10-11-16-9.dm-ai.com	CentOS7.8	/data	ambari-agent,hadoop
10	10.11.16.10	10-11-16-10.dm-ai.com	CentOS7.8	/data	ambari-agent,hadoop

说明：

- 第 0 台作为主控服务器，部署 ambari-server，搭建私有 yum 源
- 第 1-10 台作为 slave，部署 ambari-agent，也是 hadoop 运行的地方。
- mysql 数据库位于 192.168.3.199，使用 docker 启动的，后面会说。实际生产中，一般归 DBA 管理，这里不再单独列出。

## 1.3 ansible 脚本

本文中的内容有一部分可以使用 ansible 方便的管理，地址：

<https://github.com/zlingqu/ansible/tree/master/hdp>

## 2 Ambari-server 初始化

本章节的操作针对 **Ambari-server** 节点操作！

### 2.1 准备相关包

在未订阅的情况下，Cloudera 官方已经不提供相关包的下载了（），商业化进程进一步加快。官方公告：<https://cn.cloudera.com/downloads/paywall-expansion.html>。

我自己在百度网盘保存了一份：

链接：<https://pan.baidu.com/s/1bRtnpyYMIK8N1zXj61Y86Q>

提取码：qjs3

下载好后，放到 `/data/hdp` 目录，会得到如下四个 `tgz` 压缩文件

```
# ls -lh *gz
-rw-r--r-- 1 root root 2.0G Jul 27 23:43 ambari-2.7.5.0-centos7.tar.gz
-rw-r--r-- 1 root root 9.2G Jul 29 00:55 HDP-3.1.5.0-centos7-rpm.tar.gz
-rw-r--r-- 1 root root 159K Jul 29 00:55 HDP-GPL-3.1.5.0-centos7-gpl.tar.gz
-rw-r--r-- 1 root root 87M Jul 29 01:10 HDP-UTILS-1.1.0.22-centos7.tar.gz
# pwd
/data/hdp
```

分别解压四个文件，解压后的目录结构如下类似：

```
# ls -lh
total 12G
drwxr-xr-x 3 root root 21 Aug 2 18:03 ambari
-rw-r--r-- 1 root root 2.0G Jul 27 23:43 ambari-2.7.5.0-centos7.tar.gz
drwxr-xr-x 3 1001 users 21 Dec 18 2019 HDP
-rw-r--r-- 1 root root 9.2G Jul 29 00:55 HDP-3.1.5.0-centos7-rpm.tar.gz
drwxr-xr-x 3 1001 users 21 Dec 18 2019 HDP-GPL
-rw-r--r-- 1 root root 159K Jul 29 00:55 HDP-GPL-3.1.5.0-centos7-gpl.tar.gz
drwxr-xr-x 3 1001 users 21 Aug 13 2018 HDP-UTILS
-rw-r--r-- 1 root root 87M Jul 29 01:10 HDP-UTILS-1.1.0.22-centos7.tar.gz

# tree -L 4
.
├── ambari
│   ├── centos7
│   │   └── 2.7.5.0-72
│   │       ├── ambari
│   │       ├── ambari.repo
│   │       ├── artifacts.txt
│   │       ├── build.id
│   │       ├── build_metadata.txt
│   │       ├── hotfix_index.html
│   │       └── index.html
```

```
|      |—— private_index.html
|      |—— public_index.html
|      |—— repodata
|      |—— RPM-GPG-KEY
|      |—— smartsense
|      |—— tars
|—— ambari-2.7.5.0-centos7.tar.gz
|—— HDP
|  |—— centos7
|  |  |—— 3.1.5.0-152
|  |    |—— accumulo
|  |    |—— artifacts.txt
|  |    |—— atlas
|  |    |—— bigtop-jsvc
|  |    |—— bigtop-tomcat
|  |    |—— build.id
|  |    |—— build_metadata.txt
|  |    |—— datafu
|  |    |—— druid
|  |    |—— hadoop
|  |    |—— hbase
|  |    |—— HDP-3.1.5.0-152-MAINT.xml
|  |    |—— HDP-3.1.5.0-152.xml
|  |    |—— hdp.repo
|  |    |—— hdp-select
|  |    |—— hive
|  |    |—— hive_warehouse_connector
|  |    |—— hotfix_index.html
|  |    |—— index.html
|  |    |—— kafka
|  |    |—— knox
|  |    |—— livy
|  |    |—— oozie
|  |    |—— phoenix
|  |    |—— pig
|  |    |—— private_index.html
|  |    |—— public_index.html
|  |    |—— ranger
|  |    |—— repodata
|  |    |—— RPM-GPG-KEY
|  |    |—— shc
|  |    |—— spark2
|  |    |—— spark_atlas_connector
|  |    |—— spark_schema_registry
```

```

|       |—— sqoop
|       |—— ssl_hdp.repo
|       |—— storm
|       |—— superset
|       |—— tez
|       |—— vrpms
|       |—— zeppelin
|       |—— zookeeper
|—— HDP-3.1.5.0-centos7-rpm.tar.gz
|—— HDP-GPL
|   |—— centos7
|   |   |—— 3.1.5.0-152
|   |       |—— hadoopplzo
|   |       |—— hdp.gpl.repo
|   |       |—— repodata
|   |       |—— RPM-GPG-KEY
|   |       |—— ssl_hdp.gpl.repo
|   |       |—— vrpms
|—— HDP-GPL-3.1.5.0-centos7-gpl.tar.gz
|—— HDP-UTILS
|   |—— centos7
|   |   |—— 1.1.0.22
|   |       |—— hdp-utils.repo
|   |       |—— openblas
|   |       |—— repodata
|   |       |—— RPM-GPG-KEY
|   |       |—— snappy
|—— HDP-UTILS-1.1.0.22-centos7.tar.gz

```

## 2.2 安装 nginx 服务

```
yum install nginx -y
```

配置 nginx

```

# cat /etc/nginx/conf.d/yum.conf
server {
    listen    80;
    server_name localhost;
    large_client_header_buffers 4 128k;
    client_max_body_size 300m;
    client_body_buffer_size 128k;
    proxy_connect_timeout 600;
    proxy_read_timeout 600;
    proxy_send_timeout 600;
    proxy_buffer_size 64k;
}

```

```
proxy_buffers 4 32k;
proxy_busy_buffers_size 64k;
proxy_temp_file_write_size 64k;

#charset koi8-r;

#access_log logs/host.access.log main;

location / {
    root /data/hdp;
    index index.html index.htm;
    autoindex on;
    autoindex_exact_size off;
    autoindex_localtime on;
    proxy_set_header X-Real-IP $remote_addr;
    proxy_set_header REMOTE-HOST $remote_addr;
    proxy_set_header X-Forwarded-For $proxy_add_x_forwarded_for;
}
}
```

启动 nginx

```
service nginx start
```

浏览器访问，看到以下类似内容表示正常

Index of /			
<hr/>			
<a href="#">../</a>			
<a href="#">HDP/</a>	18-Dec-2019 05:21	-	
<a href="#">HDP-GPL/</a>	18-Dec-2019 04:12	-	
<a href="#">HDP-UTILS/</a>	13-Aug-2018 18:28	-	
<a href="#">ambari/</a>	02-Aug-2021 18:03	-	
<a href="#">HDP-3.1.5.0-centos7-rpm.tar.gz</a>	29-Jul-2021 00:55	9G	
<a href="#">HDP-GPL-3.1.5.0-centos7-gpl.tar.gz</a>	29-Jul-2021 00:55	158K	
<a href="#">HDP-UTILS-1.1.0.22-centos7.tar.gz</a>	29-Jul-2021 01:10	86M	
<a href="#">ambari-2.7.5.0-centos7.tar.gz</a>	27-Jul-2021 23:43	2G	
<hr/>			

这一步是制作了一个私有的 yum 源，提供相关服务。

### 2.3 添加 yum 私有 repo

```
# cat /etc/yum.repos.d/ambari.repo
[ambari]
baseurl = http://10.12.19.31/ambari/centos7/2.7.5.0-72/
enabled = 1
```



---

```
gpgcheck = 0  
name = ambari server
```

## 2.4 设置免密登录

Ambari Server 访问 Ambari Agent 需要配置 ssh 无密互访，Ambari Server 通过 SSH 公钥授权远程登录其他主机。

这里不再演示，可用 `ansible` 实现。

---

## 3 Ambari-agent 初始化

本章节的操作，针对 10 台 Ambari-agent 进行

### 3.1 修改主机名

根据规划，修改 11 台服务器的主机名，这里不再演示，可用 `ansible` 实现。

### 3.2 hosts 配置

ambari-server 的 hosts 添加如下配置。

```
# Ansible inventory hosts BEGIN
10.11.16.1 10-11-16-1.dm-ai.com
10.11.16.2 10-11-16-2.dm-ai.com
10.11.16.3 10-11-16-3.dm-ai.com
10.11.16.4 10-11-16-4.dm-ai.com
10.11.16.5 10-11-16-5.dm-ai.com
10.11.16.6 10-11-16-6.dm-ai.com
10.11.16.7 10-11-16-7.dm-ai.com
10.11.16.8 10-11-16-8.dm-ai.com
10.11.16.9 10-11-16-9.dm-ai.com
10.11.16.10 10-11-16-10.dm-ai.com
# Ansible inventory hosts END
10.12.19.31 devops-10-12-19-31
```

可用 `ansible` 实现

### 3.3 关闭防火墙

进行安装任务，集群所有主机都必须关闭防火墙

```
systemctl disable firewalld
systemctl stop firewalld
```

如果用了 `iptables`，也需要清空相关规则

```
iptables -F && iptables -t nat -F && iptables -t mangle -F && iptables -X
```

可用 `ansible` 实现

### 3.4 关闭 SELinux

关闭 `SeLinux` 功能，实现细节这里不再说。

可用 `ansible` 实现。

### 3.5 配置 limits 参数

编辑 `/etc/security/limits.conf` 文件：

保证文件中有如下内容：

```
* hard nproc 65535
* soft nproc 65535
* hard nofile 65535
* soft nofile 65535
```

可用 **ansible** 实现。

### 3.6 禁用交换分区

集群内所有主机禁用交换分区。

```
sysctl vm.swappiness=0
echo vm.swappiness=0 >> /etc/sysctl.conf
```

可用 **ansible** 实现。

### 3.7 禁用大内存页功能

```
echo never > /sys/kernel/mm/transparent_hugepage/defrag
echo never > /sys/kernel/mm/transparent_hugepage/enabled
```

可用 **ansible** 实现。

### 3.8 时间同步

**hadoop** 集群中各节点时间要同步，方法有很多，但都大同小异，这里不再细说。

可用 **ansible** 实现。

### 3.9 安装 JDK

我选用的 **oracle** 的 **jdk**，**JAVA\_HOME=/usr/local/jdk**，这里不再细说。

可使用 **ansible** 实现。

### 3.10 添加 ambari 源

```
# cat /etc/yum.repos.d/ambari.repo
[ambari]
baseurl = http://10.12.19.31/ambari/centos7/2.7.5.0-72/
enabled = 1
gpgcheck = 0
name = ambari server
```

可以使用 **ansible** 实现。

### 3.11 安装 ambari-agent

```
yum install ambari-agent -y
service ambari-agent start
```

可以使用 **ansible** 实现。

---

## 4 Ambari 部署

### 4.1 mysql 准备

ambari 可使用多种数据库存储数据，我这里选择最常用的 mysql。mysql 的搭建这里不再细说，如果只是测试可以使用 docker 部署，比较方便，这里提供 docker-compose 文件。

```
# cat docker-compose.yml
version: "3"
services:
  mysql:
    image: mysql:5.7.27
    ports:
      - "3306:3306"
    environment:
      MYSQL_ROOT_PASSWORD: 123456
      MYSQL_DATABASE: test
      TZ: Asia/Shanghai
    volumes:
      - /data/mysql-3306/data:/var/lib/mysql
```

### 4.2 部署 Ambari-server

在规划中的第 0 台服务器操作！

#### 4.2.1 安装服务

```
yum install ambari-server -y
```

#### 4.2.2 数据库驱动配置

保证环境有 mysql 的 java 驱动，比如/usr/share/java/mysql-connector-java-5.1.49.jar

```
# ambari-server setup --jdbc-db=mysql --jdbc-driver=/usr/share/java/mysql-connector-java-5.1.49.jar
Using python /usr/bin/python
Setup ambari-server
Copying /usr/share/java/mysql-connector-java-5.1.49.jar to /var/lib/ambari-server/resources/mysql-connector-java-5.1.49.jar
Creating symlink /var/lib/ambari-server/resources/mysql-connector-java-5.1.49.jar to /var/lib/ambari-server/resources/mysql-connector-java.jar
```

If you are updating existing jdbc driver jar for mysql with mysql-connector-java-5.1.49.jar. Please remove the old driver jar, from all hosts. Restarting services that need the driver, will automatically copy the new jar to the hosts.

JDBC driver was successfully initialized.

Ambari Server 'setup' completed successfully.

设置后，配置文件会产生变化，最重要的变化是添加了

custom.mysql.jdbc.name=mysql-connector-java-5.1.49.jar

### 4.2.3 配置初始化

ambari-server setup

如下图所示，图中箭头标识处需要进行选择或者输入相关内容

```
[root@devops-10-12-19-31 hdp]# ambari-server setup
Using python /usr/bin/python
Setup ambari-server
Checking SELinux...
SELinux status is 'disabled'
Customize user account for ambari-server daemon [y/n] (n)? y
Enter user account for ambari-server daemon (root):
Adjusting ambari-server permissions and ownership...
Checking firewall status...
Checking JDK...
[1] Oracle JDK 1.8 + Java Cryptography Extension (JCE) Policy Files 8
[2] Custom JDK
Enter choice (1): 2
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: /usr/local/jdk
Validating JDK on Ambari Server...done.
Check JDK version for Ambari Server...
JDK version found: 8
Hint: New JDK version is 8 for Ambari. Skipping to setup different JDK for Ambari Server.
Checking GPL software agreement...
GPL license for LZ0: https://www.gnu.org/licenses/old-licenses/gpl-2.0.en.html
Enable Ambari Server to download and install GPL Licensed LZ0 packages [y/n] (n)? y
Completing setup...
Configuring database...
Enter advanced database configuration [y/n] (n)? y
Configuring database...
Choose one of the following options:
[1] - PostgreSQL (Embedded)
[2] - Oracle
[3] - MySQL / MariaDB
[4] - PostgreSQL
[5] - Microsoft SQL Server (Tech Preview)
[6] - SQL Anywhere
[7] - RDB
Enter choice (1): 3
Hostname (local host): 192.168.3.199
Port (3306):
Database name (ambari):
Username (ambari): root
Enter Database Password (bigdata):
Re-enter password:
Configuring ambari database...
Enter full path to custom jdbc driver: /usr/share/java/mysql-connector-java-5.1.49.jar
Configuring remote database connection properties...
WARNING: Before starting Ambari Server, you must run the following DDL directly from the database shell to create the schema: /var/lib/ambari-server/resources/Ambari-DDL-MySQL-CREATE.sql
Proceed with configuring remote database connection properties [y/n] (y)? y
Extracting system views...
Ambari repo file doesn't contain latest json url, skipping repoinfos modification
Adjusting ambari-server permissions and ownership...
Ambari Server 'setup' completed successfully.
[root@devops-10-12-19-31 hdp]#
```

这一步也会自动修改配置文件，比如添加了如下内容

server.jdbc.user.name=root

server.jdbc.url=jdbc:mysql://192.168.3.199:3306/ambari

server.jdbc.driver.path=/usr/share/java/mysql-connector-java-5.1.49.jar

java.home=/usr/local/jdk

java.releases=jdk1.8

custom.mysql.jdbc.name=mysql-connector-java-5.1.49.jar

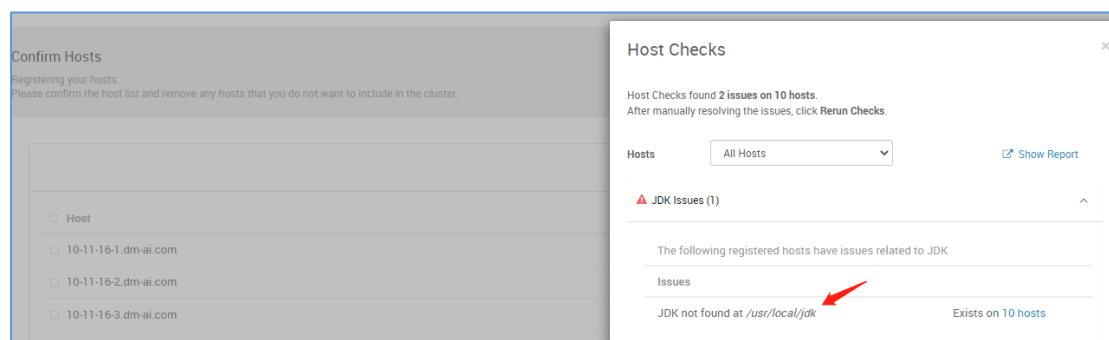
stack.java.home=/usr/local/jdk

## 4.2.4 数据库初始化

如前图 mysql 初始化脚本在：/var/lib/ambari-server/resources/Ambari-DDL-MySQL-CREATE.sql  
可以使用 navicat 或者命令行等方式，对 ambari 库进行初始化。  
初始化完成后，会有如下表，共 111 个：

名	自动递增	修改日期	数据长度	表类型	行	注释
adminpermission		2021-08-03 17:16:54	16 KB	InnoDB	7	
adminprincipal		2021-08-03 17:16:54	16 KB	InnoDB	8	
adminprincipaltype		2021-08-03 17:16:54	16 KB	InnoDB	3	
adminprivilege		2021-08-03 17:16:54	16 KB	InnoDB	1	
adminresource		2021-08-03 17:16:54	16 KB	InnoDB	0	
adminresourcetype		2021-08-03 17:16:54	16 KB	InnoDB	3	
alert_current			16 KB	InnoDB	0	
alert_definition			16 KB	InnoDB	0	
alert_group			16 KB	InnoDB	0	
alert_group_target			16 KB	InnoDB	0	
alert_grouping			16 KB	InnoDB	0	
alert_history			16 KB	InnoDB	0	
alert_notice			16 KB	InnoDB	0	
alert_target			16 KB	InnoDB	0	
alert_target_states			16 KB	InnoDB	0	
ambari_configuration			16 KB	InnoDB	0	
ambari_operation_history			16 KB	InnoDB	0	
ambari_sequences		2021-08-03 17:16:54	16 KB	InnoDB	50	
artifact			16 KB	InnoDB	0	
blueprint			16 KB	InnoDB	0	
blueprint_configuration			16 KB	InnoDB	0	
blueprint_setting			16 KB	InnoDB	0	
clusterconfig			16 KB	InnoDB	0	
clusterhostmapping			16 KB	InnoDB	0	
clusters			16 KB	InnoDB	0	
clusterservices			16 KB	InnoDB	0	
clusterstate			16 KB	InnoDB	0	
configgroupclusterconfigmapping			16 KB	InnoDB	0	
configgroup			16 KB	InnoDB	0	
configgrouphostmapping			16 KB	InnoDB	0	
execution_command			16 KB	InnoDB	0	
extension			16 KB	InnoDB	0	
extensionlink			16 KB	InnoDB	0	
groups			16 KB	InnoDB	0	
host_role_command			16 KB	InnoDB	0	
host_version			16 KB	InnoDB	0	
hostcomponentdesiredstate			16 KB	InnoDB	0	
hostcomponentstate			16 KB	InnoDB	0	
hostconfigmapping			16 KB	InnoDB	0	

注意：图中自定义了 jdk 路径 /usr/local/jdk, 经测试发现，对于 ambari-agent 也会从这里路径查找 jdk，否则在主机检测时会报错如下



## 4.2.5 启动服务

```
# service ambari-server start
```

```
[root@devops-10-12-19-31 hdp]# service ambari-server start
Using python /usr/bin/python
Starting ambari-server
Ambari Server running with administrator privileges.
Organizing resource files at /var/lib/ambari-server/resources...
Ambari database consistency check started...
Server PID at: /var/run/ambari-server/ambari-server.pid
Server out at: /var/log/ambari-server/ambari-server.out
Server log at: /var/log/ambari-server/ambari-server.log
Waiting for server start.....
Server started listening on 8080

DB configs consistency check: no errors and warnings were found.
Ambari Server 'start' completed successfully.
```

默认监听在 8080 端口。

已自动设置了开机自启。

```
[root@devops-10-12-19-31 hdp]# chkconfig
Note: This output shows SysV services only and does not include native
systemd services. SysV configuration data might be overridden by native
systemd configuration.

If you want to list systemd services use 'systemctl list-unit-files'.
To see services enabled on particular target use
'systemctl list-dependencies [target]'.

ambari-server 0:off 1:off 2:on 3:on 4:on 5:on 6:off
netconsole    0:off 1:off 2:off 3:off 4:off 5:off 6:off
network       0:off 1:off 2:on 3:on 4:on 5:on 6:off
[root@devops-10-12-19-31 hdp]#
[root@devops-10-12-19-31 hdp]# service ambari-server stop
```

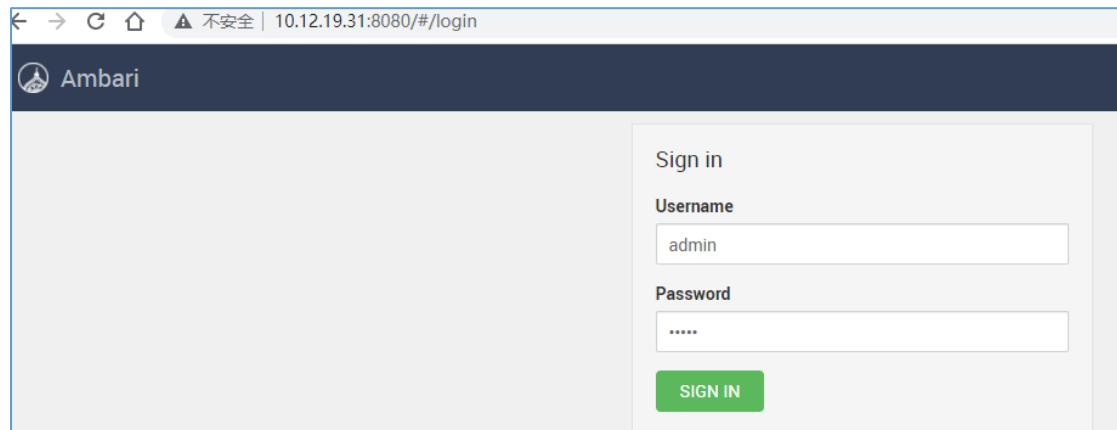
## 5 部署 Hadoop 集群

### 5.1 登录 ambari 界面

启动 ambari-server 后，通过客户端浏览器登录 Ambari 的界面。

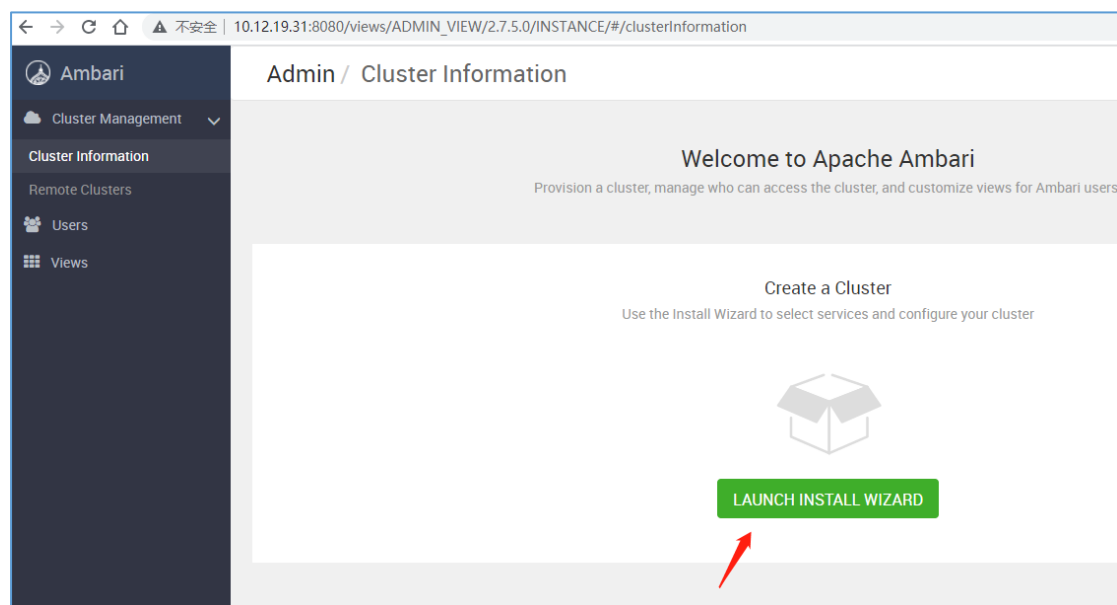
浏览器网址：<http://10.12.19.31:8080>,默认监听 8080

默认登录用户名/密码为：admin/admin。



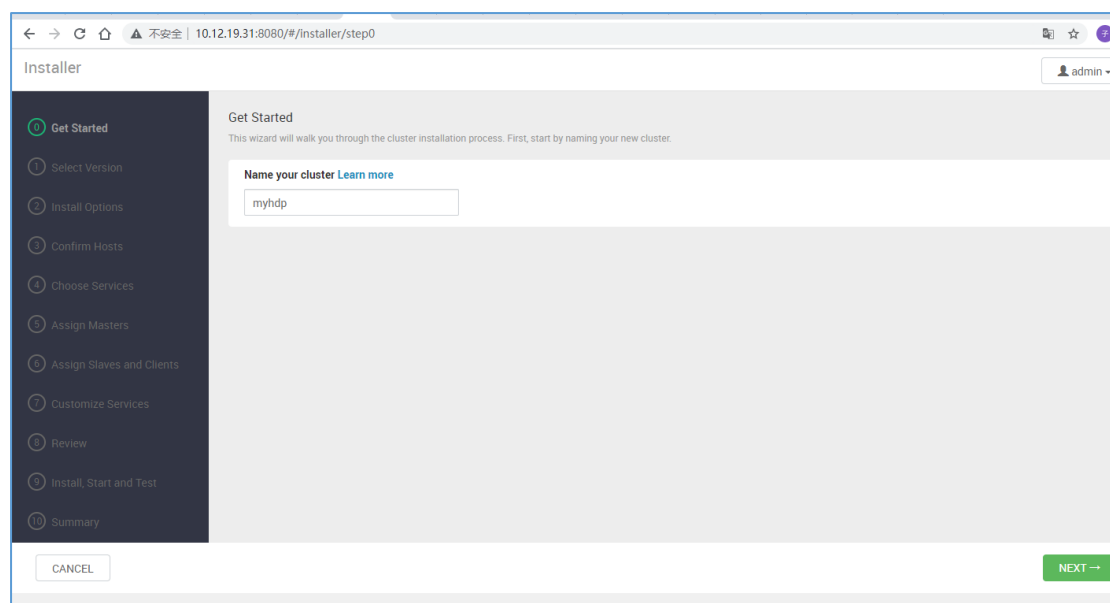
### 5.2 启动 Ambari 部署向导

点击 **Launch Install Wizard** 搭建集群



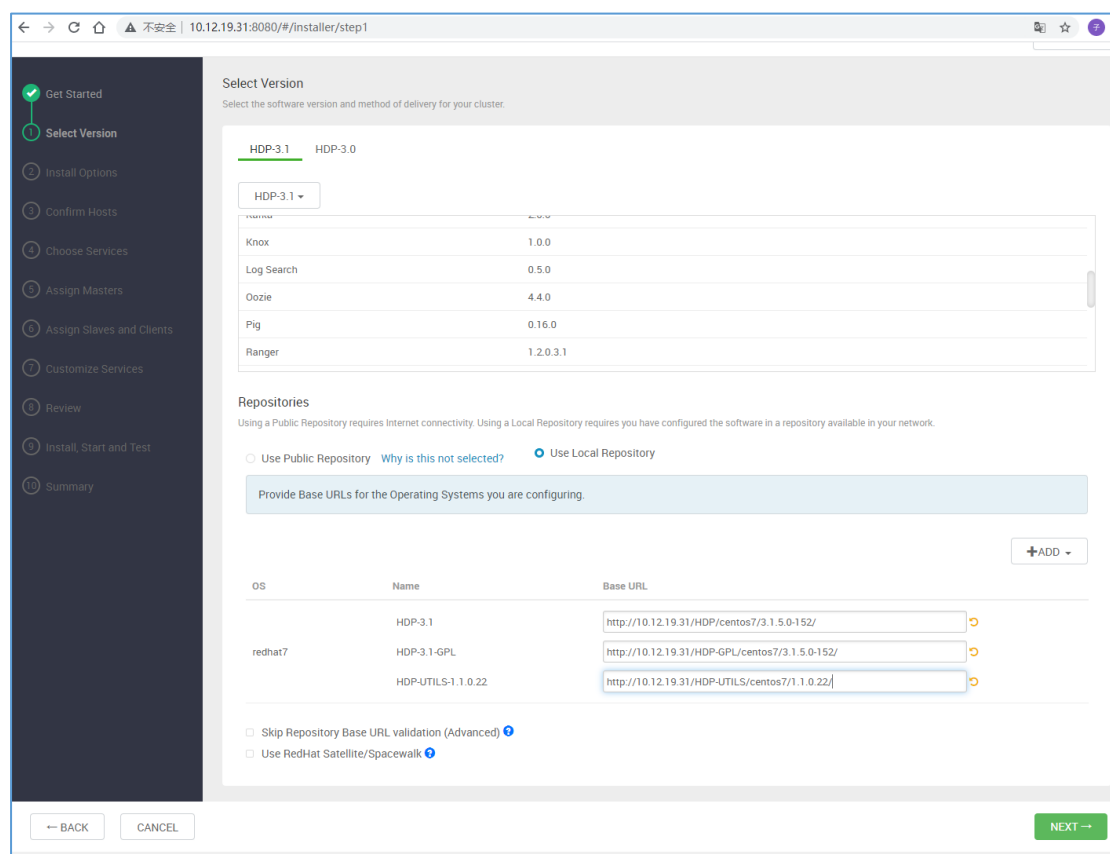
输入集群的名字，不能使用空格和其他特殊字符，我这里使用 myhdp





## 5.3 选择版本

这一步是配置各种 os 的源，我这里全都是 CentOS，所以只保留 redhat7 即可，其他的进行移除，当然有需要后续也可以添加。



输入的三个地址，是前面创建好的。

http://10.12.19.31/HDP/centos7/3.1.5.0-152/

http://10.12.19.31/HDP-GPL/centos7/3.1.5.0-152/

<http://10.12.19.31/HDP-UTILS/centos7/1.1.0.22/>

这一步配置，完成后会在各 agent 节点生成 yum 的 repo 文件，如下

```
# cat /etc/yum.repos.d/ambari-hdp-1.repo
[HDP-3.1-repo-1]
name=HDP-3.1-repo-1
baseurl=http://10.12.19.31/HDP/centos7/3.1.5.0-152/

path=/
enabled=1
gpgcheck=0
[HDP-3.1-GPL-repo-1]
name=HDP-3.1-GPL-repo-1
baseurl=http://10.12.19.31/HDP-GPL/centos7/3.1.5.0-152/

path=/
enabled=1
gpgcheck=0
[HDP-UTILS-1.1.0.22-repo-1]
name=HDP-UTILS-1.1.0.22-repo-1
baseurl=http://10.12.19.31/HDP-UTILS/centos7/1.1.0.22/

path=/
enabled=1
gpgcheck=0
```

```
[root@10-11-16-3 ~]# cat /etc/yum.repos.d/ambari-hdp-1.repo
[HDP-3.1-repo-1]
name=HDP-3.1-repo-1
baseurl=http://10.12.19.31/HDP/centos7/3.1.5.0-152/

path=/
enabled=1
gpgcheck=0
[HDP-3.1-GPL-repo-1]
name=HDP-3.1-GPL-repo-1
baseurl=http://10.12.19.31/HDP-GPL/centos7/3.1.5.0-152/

path=/
enabled=1
gpgcheck=0
[HDP-UTILS-1.1.0.22-repo-1]
name=HDP-UTILS-1.1.0.22-repo-1
baseurl=http://10.12.19.31/HDP-UTILS/centos7/1.1.0.22/

path=/
enabled=1
gpgcheck=0
```

## 5.4 安装选项

选择管理的主机、ssh 私钥

10-11-16-[1-10].dm-ai.com

**Install Options**  
Enter the list of hosts to be included in the cluster and provide your SSH key.

**Target Hosts**  
Enter a list of hosts using the Fully Qualified Domain Name (FQDN), one per line. Or use [Pattern Expressions](#)

10-11-16-[1-10].dm-ai.com

**Host Registration Information**

☒ Provide your **SSH Private Key** to automatically register hosts ☐ Perform **manual registration** on hosts and do not use SSH

CHOOSE FILE No file selected

```
-----BEGIN RSA PRIVATE KEY-----
MIIEowIBAAKCAQEA0v5+wo4L/CXtQFkcnLNU9zf+okkGunM1a40EGPrD7Ao7VzC0
LwP0YicFSJkQcQ69q0Nf4QCD2pMbZ4eJRTdaTJ1h+Al6N4bD7eq01Z60ahckMdvA
p3eM/bzC+k6/vS0ZIN1zt1wFQldy6tObxtCbFkVrY7A8Z7CjnGYXiyl1HiES8LqB
Bs+qGVWHP4k3slhqaqStXSMFkseeSoVI7f1j6rSW85e3NgJWFm5DQ+xtVtDAQE+x
vtJZM6QNJfv7WKz1hgSfSWHZdrHohzNVLj1RvbvhdSVe9QeFwermtGUo2xnHOSDO
8Jj2LEz7v0P4vRF9FSPf6pKu5yaoqST14tU5QIDAQABaoIBAC500iWapRi2f1Yk
KHc+zmxfNpwJKNfck3bSIPfWec8gw7sz7ldvUND41VkcU0b3MDsY8z7fLPdJCS21
CkdN1ZzL1qgn9FGHwV1pBHXvSasK3x+uG07n24cIWUyz6DKs82DB1TaNaZLVHFxb
CpIMkENBO8zKDce3SYBa14c4FAMDW5eJUpK7KwaCu5/g5I4dn6d3K3uJ4Dgha0Bg
hJMLNCUgR2YknUFBW0+MgtusQW7ht3rMq1tOSiNYpBn+YZ45x0IZbSAhocrSMKs
y/tCjqpR6dTKbbEj3wseNfGqYmXhc7G+JyJ/di1UhmFT1aRg12vb4XC8pMBYt4r
D8f6rjUCgYEA/XWuV01s5XdmCP/N/ZIDKj0Amy+EjIc2Rbybqg+r110Yv+43FXF2
0xxq8n1zkSFGS1x7sZGJX8c+n8gD71v9GcR1Ne118sz30lh4jRbFMnvL+1XTo7Y
HcgQF5HVMWd9MMfP1wCxd0nzrvEDSG7d4cPnqAHU0FI/F9VOS51c8CgYEA1Rvb
eVJ210gHpZbsCptjmgGnTS+2i4KR3Uxy/80XePd5I1Wd2NRTf40c9jEYFLIA33T
/w2G3icjUgcvfhzx4z02EuhRIICvwrEw15dAHTgeEHPjPp2F/kaK8I+7+CJW8s
apTDqgcV3z50U6YTI0aJfYdz9Qvqc+9j+28yweCgYBXW4Qm1THPNQF91mztBLJu
Z8d1tU7+Ed1HC+adQUF8pD2vkS8z2Mm1aaWw/1s/AowHxxce8I2km1uhDpMsSEt
0mhg4KXWt0cZ0aj3CgyY0xfrdh89WEZnVFO19j4445mjcBw7z2do7KyXSIWBS5p
n3AkqJwXFUBUWYRkHfP8SQK8gQc0c+18se60x0Smg/H5QHwV70sLvFdzV7atUyzj
/t478IOA5s8rEQeflU0c0ZChA060yPhhBULjaad2Fj/+S1JavoufzciGOpSpJn35c
GSbY9E/qPzssK1ziaIn30TpEpVDe/k4Bwa4YzK20rFONFdtQ+wtbCsiZ73NYSta
jHXbmQKBgDJ0MC2+aiH8dppEaShsJQ/QuGoqyuca19dAgL/LAJO5NS5OegpjdsBe
wxJhnHd1Y7p5ySGpcc23V8mJSHh1Fbp9Ss1S2AEVxktmzp28ARcCJBCoym1VZhP/
eaY13YZ+iuNjrfq/e4xP0DBEioYbQCe74FpOoa5JZCraJaU1bXtS
-----END RSA PRIVATE KEY-----
```

**SSH User Account** root

## 5.5 确认主机

主机名的正则表达式会自动展开，10-11-16-[1-10].dm-ai.com

**Host name pattern expressions**

10-11-16-1.dm-ai.com  
10-11-16-2.dm-ai.com  
10-11-16-3.dm-ai.com  
10-11-16-4.dm-ai.com  
10-11-16-5.dm-ai.com  
10-11-16-6.dm-ai.com  
10-11-16-7.dm-ai.com  
10-11-16-8.dm-ai.com  
10-11-16-9.dm-ai.com  
10-11-16-10.dm-ai.com

CANCEL OK

<input type="checkbox"/> Host	Progress	Status
<input type="checkbox"/> 10-11-16-1.dm-ai.com	<div></div>	Success
<input type="checkbox"/> 10-11-16-2.dm-ai.com	<div></div>	Success
<input type="checkbox"/> 10-11-16-3.dm-ai.com	<div></div>	Installing
<input checked="" type="checkbox"/> 10-11-16-4.dm-ai.com	<div></div>	Installing
<input type="checkbox"/> 10-11-16-5.dm-ai.com	<div></div>	Success
<input type="checkbox"/> 10-11-16-6.dm-ai.com	<div></div>	Installing
<input type="checkbox"/> 10-11-16-7.dm-ai.com	<div></div>	Success
<input type="checkbox"/> 10-11-16-8.dm-ai.com	<div></div>	Installing
<input type="checkbox"/> 10-11-16-9.dm-ai.com	<div></div>	Success
<input type="checkbox"/> 10-11-16-10.dm-ai.com	<div></div>	Installing

环境检测

Confirm Hosts			
Registering your hosts. Please confirm the host list and remove any hosts that you do not want to include in the cluster.			
<input type="checkbox"/> Host	Progress	Status	Action
<input type="checkbox"/> 10-11-16-1.dm-ai.com	<div></div>	Success	
<input type="checkbox"/> 10-11-16-2.dm-ai.com	<div></div>	Success	
<input type="checkbox"/> 10-11-16-3.dm-ai.com	<div></div>	Success	
<input type="checkbox"/> 10-11-16-4.dm-ai.com	<div></div>	Success	
<input type="checkbox"/> 10-11-16-5.dm-ai.com	<div></div>	Success	
<input checked="" type="checkbox"/> 10-11-16-6.dm-ai.com	<div></div>	Success	
<input type="checkbox"/> 10-11-16-7.dm-ai.com	<div></div>	Success	
<input type="checkbox"/> 10-11-16-8.dm-ai.com	<div></div>	Success	
<input type="checkbox"/> 10-11-16-9.dm-ai.com	<div></div>	Success	
<input type="checkbox"/> 10-11-16-10.dm-ai.com	<div></div>	Success	
Some warnings were encountered while performing checks against the 10 registered hosts above <a href="#">Click here to see the warnings.</a>			

检测后如果出现一些警告，比如下图，则需要处理下

## Host Checks

Host Checks found **2 issues on 10 hosts**.  
After manually resolving the issues, click **Rerun Checks**.

Hosts

All Hosts

Show Report

JDK Issues (1)

The following registered hosts have issues related to JDK

Issues
JDK not found at <code>/usr/local/jdk</code> Exists on <b>10 hosts</b>

Package Issues (1)

The following packages should be uninstalled

Package
oracle-j2sdk1.81.8.0+update181-1Installed on <b>8 hosts</b>

Transparent Huge Pages Issues (0)

Disk Issues (0)

Repository Issues (0)

Firewall Issues (0)

Process Issues (0)

File and Folder Issues (0)

Service Issues (0)

User Issues (0)

Misc Issues (0)

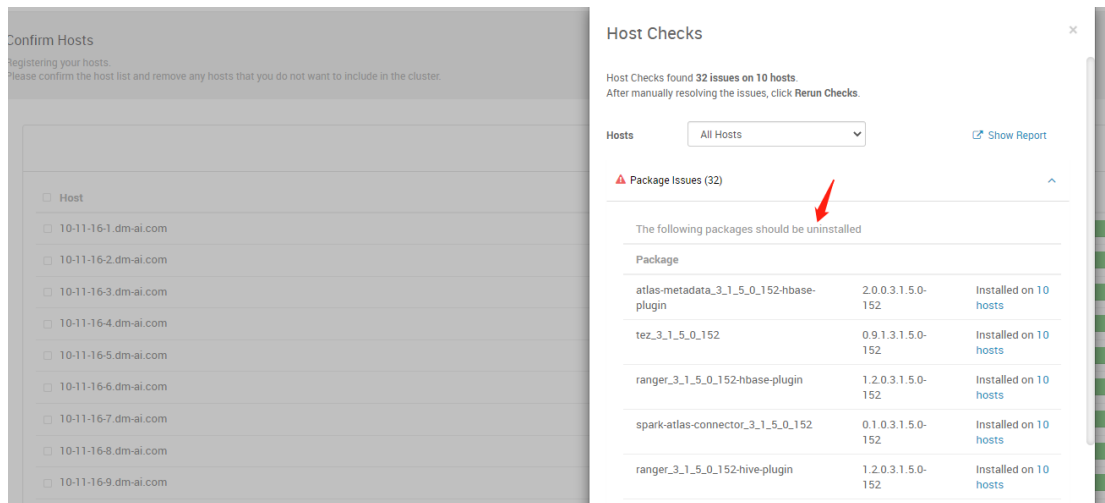
Alternatives Issues (0)

Reverse Lookup Issues (0)

Hostname Resolution Issues (0)

RERUN CHECKS

CLOSE



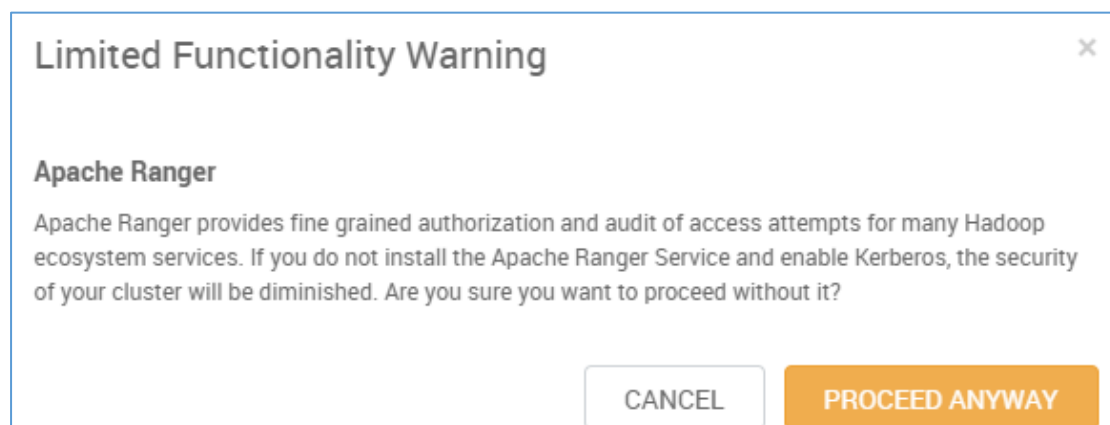
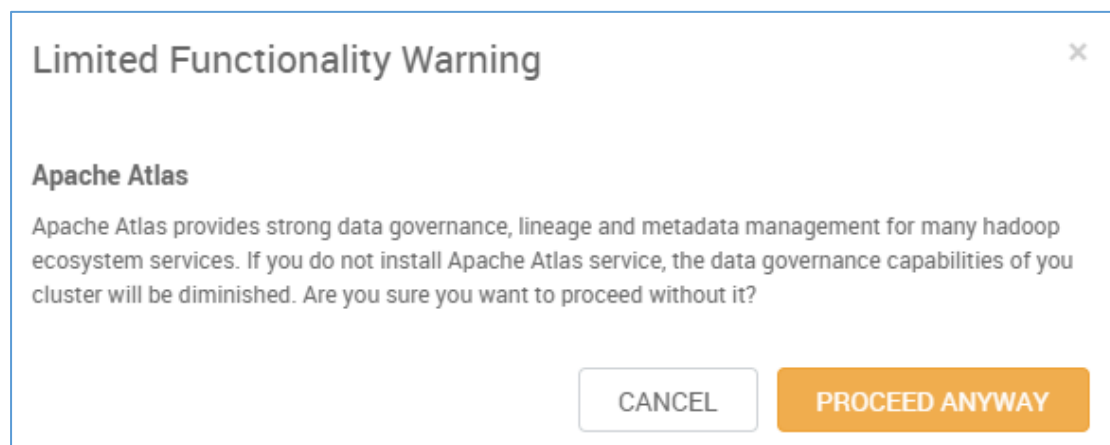
出现图中类似警告，是应为之前的 **hadoop** 环境没有清理干净，使用 **yum** 卸载即可。

比如：

```
ansible ambari-agent -m shell -a "yum remove atlas-metadata_3_1_5_0_152-hbase-plugin tez_3_1_5_0_152 \
ranger_3_1_5_0_152-hbase-plugin spark-atlas-connector_3_1_5_0_152 \
ranger_3_1_5_0_152-hive-plugin livy2_3_1_5_0_152 hadoop_3_1_5_0_152-hdfs \
hive_warehouse_connector_3_1_5_0_152 spark2_3_1_5_0_152 hadoop_3_1_5_0_152-mapreduce \
atlas-metadata_3_1_5_0_152-hive-plugin hive_3_1_5_0_152 hadoop_3_1_5_0_152 hadoop_3_1_5_0_152-yarn \
hive_3_1_5_0_152-jdbc ambari-metrics-grafana spark2_3_1_5_0_152-python spark2_3_1_5_0_152-yarn-shuffle \
hive_3_1_5_0_152-hcatalog hadoop_3_1_5_0_152-client bigtop-jsvc zookeeper_3_1_5_0_152 \
ranger_3_1_5_0_152-hdfs-plugin ambari-metrics-hadoop-sink ambari-metrics-monitor smartsense-hst hbase_3_1_5_0_152 \
ranger_3_1_5_0_152-yarn-plugin hadoop_3_1_5_0_152-libhdfs hdp-select zookeeper_3_1_5_0_152-server \
ambari-metrics-collector -y"
```

全部检测通过，出现 **passed** 字样，纳入管理后会出现 **success** 字样，如下图

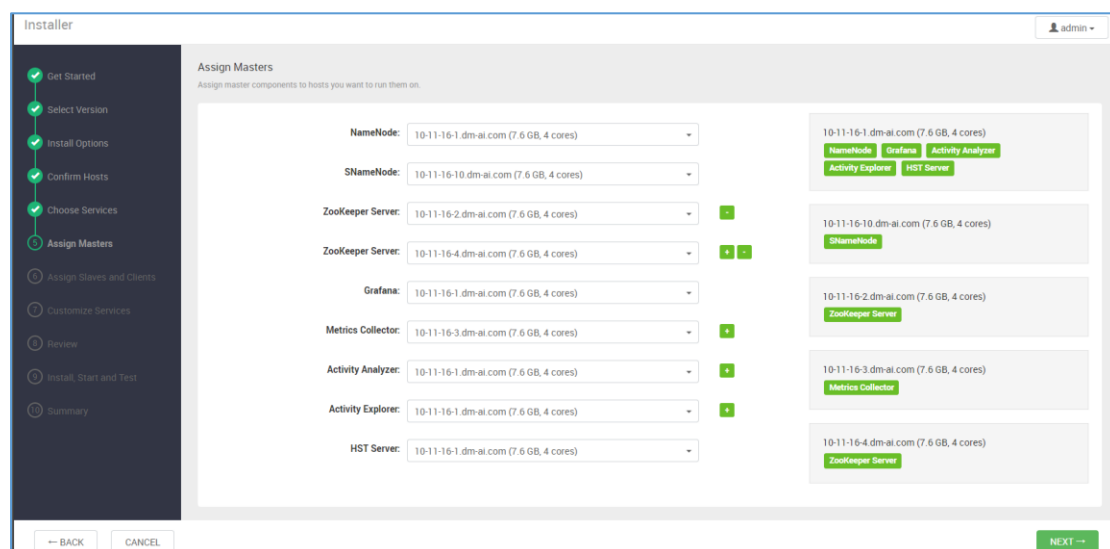




有可能会出现以下几种情况，都选择 continue 或 proceed anyway。

## 5.7 分配主要组件

分配主要组件，会有一个默认的划分，可以自己调整。中间是以组件为视角，右侧以主机为视角。





## 5.8 从属组件配置

如下图所示，前面 5 台已经安装了主要组件，会有\*号标记，剩余的几台默认勾选了 DataNode 角色，可根据自己的需要进行更改。Client 角色是指 HDFS、ZK 的客户端。

Installer

admin

Get Started

Select Version

Install Options

Confirm Hosts

Choose Services

Assign Masters

Assign Slaves and Clients

Customize Services

Review

Install, Start and Test

Summary

Assign Slaves and Clients

Assign slave and client components to hosts you want to run them on.  
Hosts that are assigned master components are shown with \*.  
\*Client\* will install HDFS Client and ZooKeeper Client.

Host	all	none	all	none	all	none
10-11-16-1.dm-ai.com*	<input type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-2.dm-ai.com*	<input type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-3.dm-ai.com*	<input type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-4.dm-ai.com*	<input type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-10.dm-ai.com*	<input type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-5.dm-ai.com	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input checked="" type="checkbox"/>	Client
10-11-16-6.dm-ai.com	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-7.dm-ai.com	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-8.dm-ai.com	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-9.dm-ai.com	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input checked="" type="checkbox"/>	Client

Items per page: 25 1 - 10 of 10

BACK

CANCEL

NEXT

我进行了一个变更，如下

Installer

admin

Get Started

Select Version

Install Options

Confirm Hosts

Choose Services

Assign Masters

Assign Slaves and Clients

Customize Services

Review

Install, Start and Test

Summary

Assign Slaves and Clients

Assign slave and client components to hosts you want to run them on.  
Hosts that are assigned master components are shown with \*.  
\*Client\* will install HDFS Client and ZooKeeper Client.

Host	all	none	all	none	all	none
10-11-16-1.dm-ai.com*	<input type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-2.dm-ai.com*	<input type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input checked="" type="checkbox"/>	Client
10-11-16-3.dm-ai.com*	<input type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input checked="" type="checkbox"/>	Client
10-11-16-4.dm-ai.com*	<input type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input checked="" type="checkbox"/>	Client
10-11-16-10.dm-ai.com*	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-5.dm-ai.com	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-6.dm-ai.com	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-7.dm-ai.com	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-8.dm-ai.com	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client
10-11-16-9.dm-ai.com	<input checked="" type="checkbox"/>	DataNode	<input type="checkbox"/>	NFSGateway	<input type="checkbox"/>	Client

Items per page: 25 1 - 10 of 10

BACK

CANCEL

NEXT

## 5.9 定制服务

输入相关密码，我这里为了方便，都设置为 123456

	Username*	Password*	Confirm Password*
Grafana Admin	admin	*****	*****
Activity Explorer's Admin	N/A	*****	*****

下一步配置相关目录，以下是默认情况，其中 **datanode**、**namenode** 使用了 **directories**（其他是 **directory**），所以可以配置多个路径，

Directory Name	Default Path	Actions
DataNode directories	/hadoop/hdfs/data/data/hadoop/hdfs/data	+ C 🛑
NameNode directories	/hadoop/hdfs/namenode/data/hadoop/hdfs/namenode	C 🛑
SecondaryNameNode Checkpoint directories	/hadoop/hdfs/namesecondary	C 🛑
NFSGateway dump directory	/tmp/hdfs-nfs	+ C 🛑
NameNode Backup directory	/tmp/upgrades	+ C
JournalNode Edits directory	/hadoop/hdfs/journalnode	+ C 🛑
NameNode Checkpoint Edits directory	\$(if's namenode checkpoint dir)	+ C 🛑

可根据实际情况进行修改，下图中红色部分是我修改过的

```

/data/hadoop/hdfs/datanode
/data/hadoop/hdfs/namenode
/data/hadoop/hdfs/namesecondary
/data/hadoop/hdfs/journalnode
/data/hadoop/zookeeper

```

DATA DIRS

DataNode directories

/data/hadoop/hdfs/datanode

NameNode directories

/data/hadoop/hdfs/namenode

SecondaryNameNode Checkpoint directories

/data/hadoop/hdfs/namesecondary

NFSGateway dump directory

/tmp/.hdfs-nfs

NameNode Backup directory

/tmp/upgrades

JournalNode Edits directory

/data/hadoop/hdfs/journalnode

NameNode Checkpoint Edits directory

\$(dfs.namenode.checkpoint.dir)

同时我也修改了 zk 的数据目录

Get Started

Select Version

Install Options

Confirm Hosts

Choose Services

Assign Masters

Assign Slaves and Clients

Customize Services

Review

Install, Start and Test

Summary

CREDENTIALS

DATABASES

DIRECTORIES

ACCOUNTS

ALL CONFIGURATIONS

HDFS

ZOOKEEPER

AMBARI METRICS

DATA DIRS

ZooKeeper directory

/data/hadoop/zookeeper

LOG DIRS

ZooKeeper Log Dir

/var/log/zookeeper

PID DIRS

ZooKeeper PID Dir

/var/run/zookeeper

27

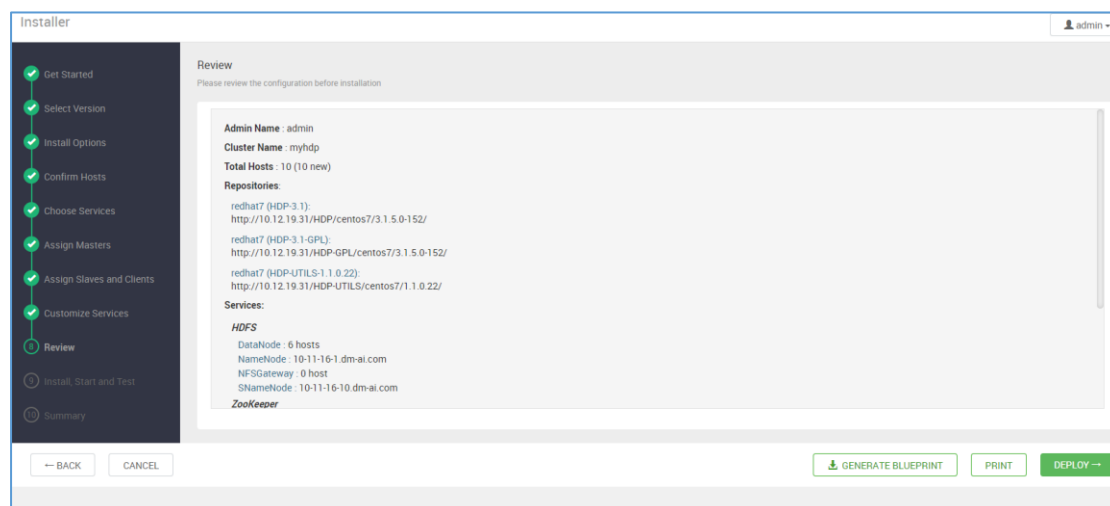
用户和组管理，这里我使用了默认配置，没有修改

Users/Groups	Usernames
Smoke User	ambari-qa
Hadoop Group	hadoop
Ambari Metrics User	ams
HDFS User	hdfs
Proxy User Group	users
ZooKeeper User	zookeeper

所有配置，包含前面几步的配置，也有一些新的配置，可以根据需要定制。我这里没有修改。

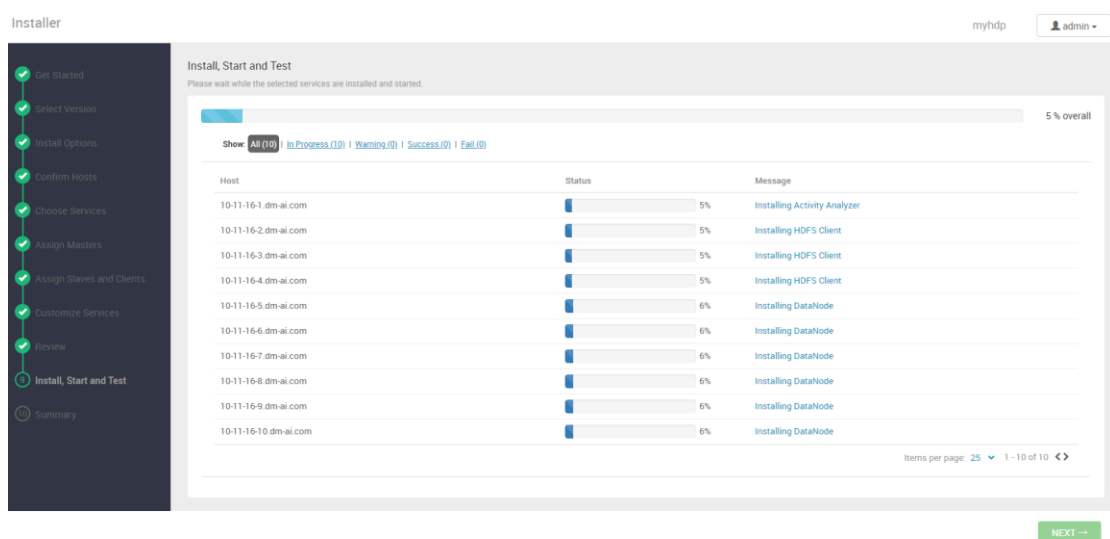
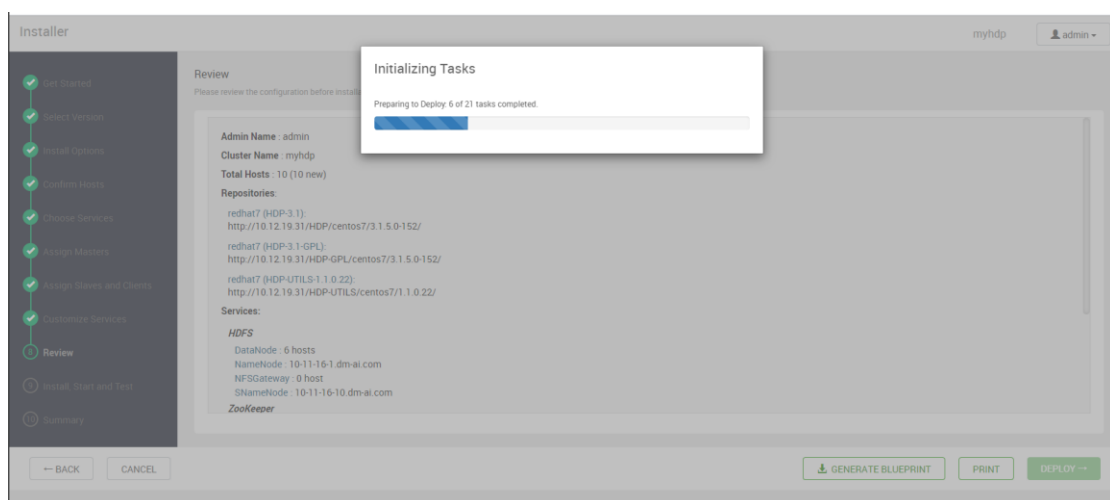
Parameter	Value
NameNode directories	/data/hadoop/hdfs/namedata
NameNode Java heap size	10GB
NameNode Server threads	100
Minimum replicated blocks %	100%
DataNode directories	/data/hadoop/hdfs/datanode
DataNode failed disk tolerance	0
DataNode maximum Java heap size	10GB
DataNode max data transfer threads	4096

## 5. 10 review 概要



没有问题点击 **deploy** 进行部署。

## 5.11 安装、启动和检测



Installer

myhdpadmin

Get Started

Select Version

Install Options

Confirm Hosts

Choose Services

Assign Masters

Assign Slaves and Clients

Customize Services

Review

Install, Start and Test

Summary

Install, Start and Test

Please wait while the selected services are installed and started.

94 % overall

Show All (10) | In Progress (2) | Warnings (0) | Success (7) | Fail (0)

Host	Status	Message
10-11-16-1.dm-ai.com	74%	Preparing to start Activity Analyzer
10-11-16-2.dm-ai.com	100%	Success
10-11-16-3.dm-ai.com	91%	Starting Metrics Collector
10-11-16-4.dm-ai.com	100%	Success
10-11-16-5.dm-ai.com	100%	Success
10-11-16-6.dm-ai.com	83%	Preparing to execute Ambari Metrics Service Check
10-11-16-7.dm-ai.com	100%	Success
10-11-16-8.dm-ai.com	100%	Success
10-11-16-9.dm-ai.com	100%	Success
10-11-16-10.dm-ai.com	100%	Success

Items per page: 25 | 1 - 10 of 10

NEXT

Installer

myhdpadmin

Get Started

Select Version

Install Options

Confirm Hosts

Choose Services

Assign Masters

Assign Slaves and Clients

Customize Services

Review

Install, Start and Test

Summary

Install, Start and Test

Please wait while the selected services are installed and started.

100 % overall

Show All (10) | In Progress (0) | Warnings (0) | Success (10) | Fail (0)

Host	Status	Message
10-11-16-1.dm-ai.com	100%	Success
10-11-16-2.dm-ai.com	100%	Success
10-11-16-3.dm-ai.com	100%	Success
10-11-16-4.dm-ai.com	100%	Success
10-11-16-5.dm-ai.com	100%	Success
10-11-16-6.dm-ai.com	100%	Success
10-11-16-7.dm-ai.com	100%	Success
10-11-16-8.dm-ai.com	100%	Success
10-11-16-9.dm-ai.com	100%	Success
10-11-16-10.dm-ai.com	100%	Success

Items per page: 25 | 1 - 10 of 10

Successfully installed and started the services.

NEXT

## 5.12 汇总

节点 1 安装了 NameNode,节点 10 安装了 SNameNode，这也是前面规划好的。

Installer

myhdpadmin

Get Started

Select Version

Install Options

Confirm Hosts

Choose Services

Assign Masters

Assign Slaves and Clients

Customize Services

Review

Install, Start and Test

Summary

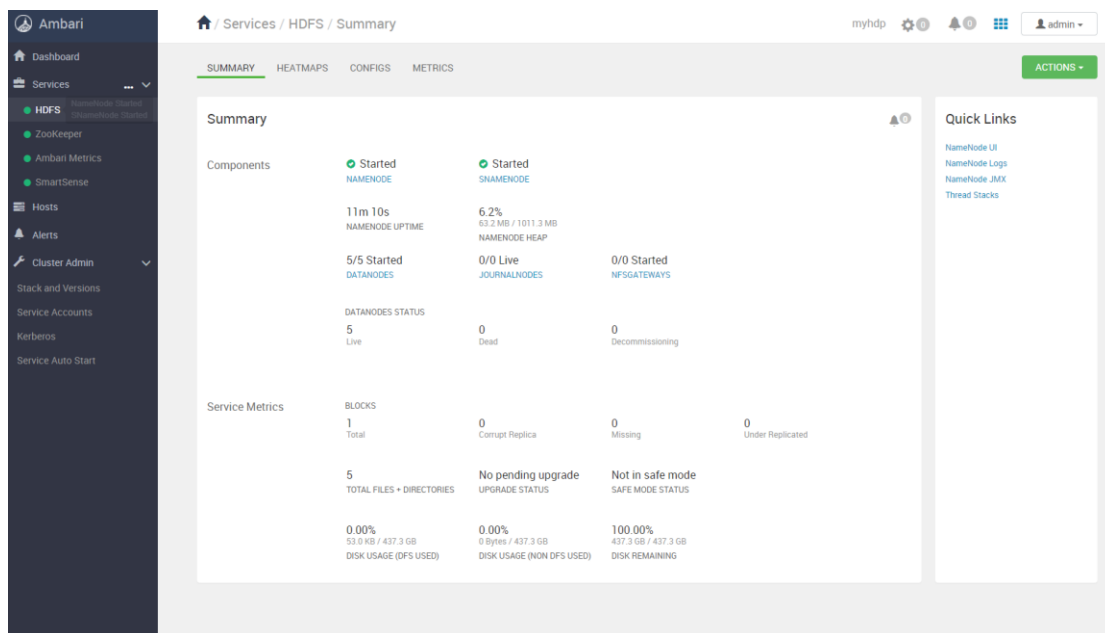
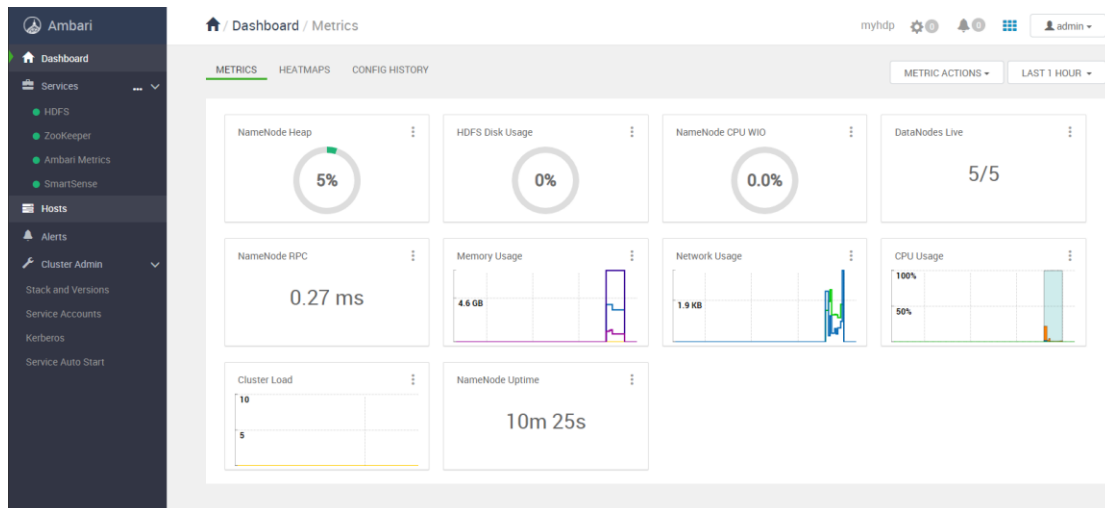
Summary

Here is the summary of the install process.

The cluster consists of 10 hosts  
Installed and started services successfully on 10 new hosts  
Master services installed  
NameNode installed on 10-11-16-1.dm-ai.com  
SNameNode installed on 10-11-16-10.dm-ai.com  
All services started  
All tests passed  
Install and start completed in 8 minutes and 20 seconds

COMPLETE

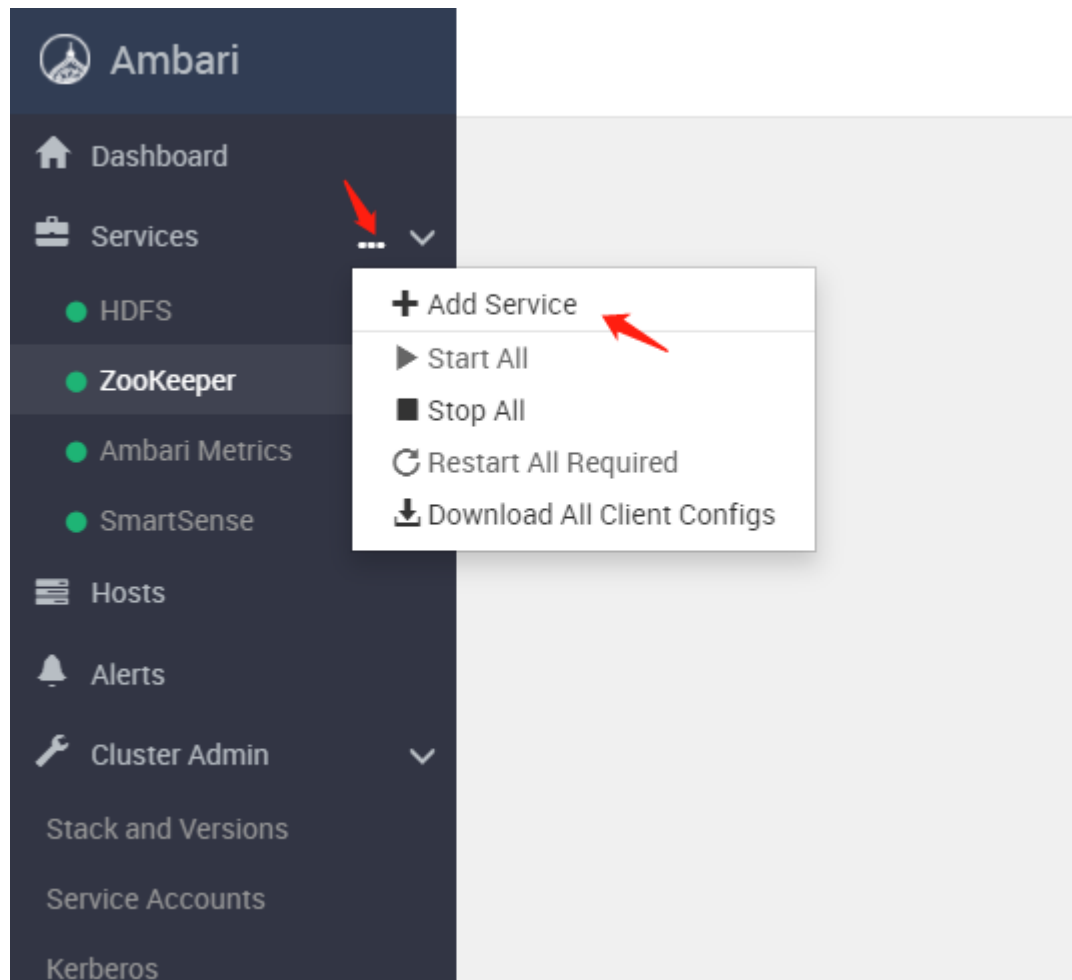
30



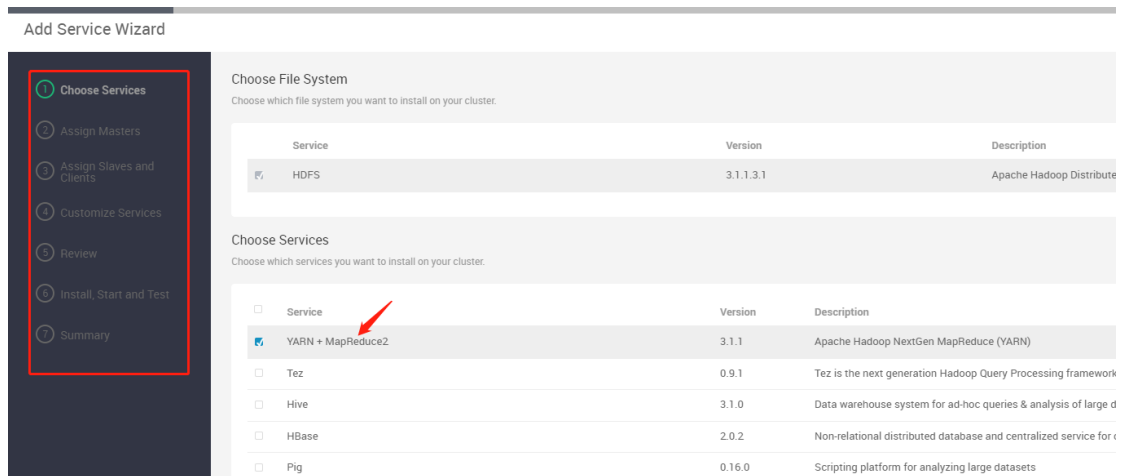
## 6 添加服务

### 6.1 添加 yarn+mapreduce2 服务

点击 add service



选择 YARN+MapReduce2



从图中也可以看到，服务添加和前一章节的 **hadoop** 部署中后面的几步是相同的。



**Add Service Wizard**

**Assign Masters**  
Assign master components to hosts you want to run them on.

Choose Services  
Assign Masters  
Assign Slaves and Clients  
Customize Services  
Review  
Install, Start and Test  
Summary

SNameNode: 10-11-16-10.dm-ai.com (7.6 GB, 4 cores)  
 NameNode: 10-11-16-1.dm-ai.com (7.6 GB, 4 cores)  
 Timeline Service V1.5: 10-11-16-10.dm-ai.com (7.6 GB, 4 cores)  
 YARN Registry DNS: 10-11-16-2.dm-ai.com (7.6 GB, 4 cores)  
 ResourceManager: 10-11-16-3.dm-ai.com (7.6 GB, 4 cores)  
 Timeline Service V2.0 Reader: 10-11-16-4.dm-ai.com (7.6 GB, 4 cores)  
 History Server: 10-11-16-10.dm-ai.com (7.6 GB, 4 cores)  
 ZooKeeper Server: 10-11-16-2.dm-ai.com (7.6 GB, 4 cores)  
 ZooKeeper Server: 10-11-16-4.dm-ai.com (7.6 GB, 4 cores)  
 Metrics Collector: 10-11-16-3.dm-ai.com (7.6 GB, 4 cores)  
 Grafana: 10-11-16-1.dm-ai.com (7.6 GB, 4 cores)  
 Activity Analyzer: 10-11-16-1.dm-ai.com (7.6 GB, 4 cores)  
 Activity Explorer: 10-11-16-1.dm-ai.com (7.6 GB, 4 cores)  
 HST Server: 10-11-16-1.dm-ai.com (7.6 GB, 4 cores)

10-11-16-10.dm-ai.com (7.6 GB, 4 cores)  
 NameNode Grafana Activity Analyzer  
 Activity Explorer HST Server  
 10-11-16-10.dm-ai.com (7.6 GB, 4 cores)  
 NameNode Timeline Service V1.5 History Server  
 10-11-16-2.dm-ai.com (7.6 GB, 4 cores)  
 YARN Registry DNS ZooKeeper Server  
 10-11-16-3.dm-ai.com (7.6 GB, 4 cores)  
 ResourceManager Metrics Collector  
 10-11-16-4.dm-ai.com (7.6 GB, 4 cores)  
 Timeline Service V2.0 Reader ZooKeeper Server  
 5 hosts not running master services

如下图配置 NodeManager 节点和 Client 节点，这里的 client 值 yarn、mapreduce2 的 client，client 的意思可看图中的英文说明，后面章节不再单独说。

**Add Service Wizard**

**Assign Slaves and Clients**  
Assign slave and client components to hosts you want to run them on.  
Hosts that are assigned master components are shown with \*.  
"Client" will install YARN Client and MapReduce2 Client.

Choose Services  
Assign Masters  
Assign Slaves and Clients  
Customize Services  
Review  
Install, Start and Test  
Summary

Host	<input checked="" type="checkbox"/> all   <input type="checkbox"/> none	<input checked="" type="checkbox"/> all   <input type="checkbox"/> none	<input type="checkbox"/> all   <input type="checkbox"/> none	<input type="checkbox"/> all   <input type="checkbox"/> none
10-11-16-10.dm-ai.com *	<input checked="" type="checkbox"/> DataNode	<input checked="" type="checkbox"/> NFSGateway	<input type="checkbox"/> NodeManager	<input checked="" type="checkbox"/> Client
10-11-16-10.dm-ai.com *	<input checked="" type="checkbox"/> DataNode	<input checked="" type="checkbox"/> NFSGateway	<input type="checkbox"/> NodeManager	<input type="checkbox"/> Client
10-11-16-2.dm-ai.com *	<input checked="" type="checkbox"/> DataNode	<input checked="" type="checkbox"/> NFSGateway	<input checked="" type="checkbox"/> NodeManager	<input checked="" type="checkbox"/> Client
10-11-16-3.dm-ai.com *	<input checked="" type="checkbox"/> DataNode	<input checked="" type="checkbox"/> NFSGateway	<input checked="" type="checkbox"/> NodeManager	<input checked="" type="checkbox"/> Client
10-11-16-4.dm-ai.com *	<input checked="" type="checkbox"/> DataNode	<input checked="" type="checkbox"/> NFSGateway	<input checked="" type="checkbox"/> NodeManager	<input checked="" type="checkbox"/> Client
10-11-16-5.dm-ai.com	<input type="checkbox"/> DataNode	<input checked="" type="checkbox"/> NFSGateway	<input type="checkbox"/> NodeManager	<input type="checkbox"/> Client
10-11-16-6.dm-ai.com	<input type="checkbox"/> DataNode	<input checked="" type="checkbox"/> NFSGateway	<input type="checkbox"/> NodeManager	<input type="checkbox"/> Client
10-11-16-7.dm-ai.com	<input type="checkbox"/> DataNode	<input checked="" type="checkbox"/> NFSGateway	<input type="checkbox"/> NodeManager	<input type="checkbox"/> Client
10-11-16-8.dm-ai.com	<input type="checkbox"/> DataNode	<input checked="" type="checkbox"/> NFSGateway	<input type="checkbox"/> NodeManager	<input type="checkbox"/> Client
10-11-16-9.dm-ai.com	<input type="checkbox"/> DataNode	<input checked="" type="checkbox"/> NFSGateway	<input type="checkbox"/> NodeManager	<input checked="" type="checkbox"/> Client

Items per page: 25 1 - 10 of 10

← BACK NEXT →

出现如下报错

**Configurations**

Highly Recommended Configurations 1  
Please review the following recommended changes, and click on the property name to change its value.

Type	Service	Property	Current Value	Description
Warning	Ambari Metrics	hbase_master_heapsize	768	Value is less than the recommended default of 896 HBase Master Heap Size. In embedded mode, total heap size is sum of master and regionserver heap sizes.

CANCEL PROCEED ANYWAY

找到如下位置，修改为 1024

Customize Services

We have come up with recommended configurations for the services you selected. Customize them as you see fit.

HDFS YARN MAPREDUCE2 ZOOKEEPER **AMBARI METRICS** SMARTSENSE MISC

Group: Default (10)

General

Metric Collector

Advanced ams-env

Advanced ams-grafana-env

Advanced ams-grafana-ini

Advanced ams-hbase-env

hbase\_classpath\_additional

HBase Log Dir Prefix

HBase Master Maximum Memory

hbase\_master\_maxperm\_size

HBase Master maximum value for Xmn

HBase PID Dir

HBase RegionServer Maximum Memory

HBase RegionServer shutdown timeout

768 MB

128 MB

192 MB

512 MB

30

后面的步骤不再单独列出。

## 6.2 添加 hive 服务

依赖 Tez 服务，一并安装下

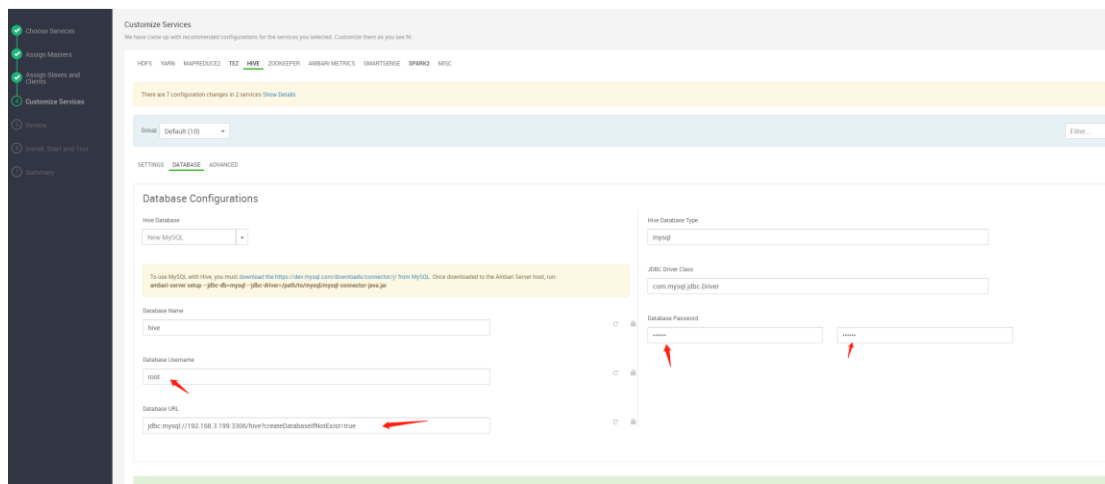
Tez Needed

You did not select Tez, but it is needed by other services you selected. We will automatically add TEZ. Is this OK?

CANCEL OK

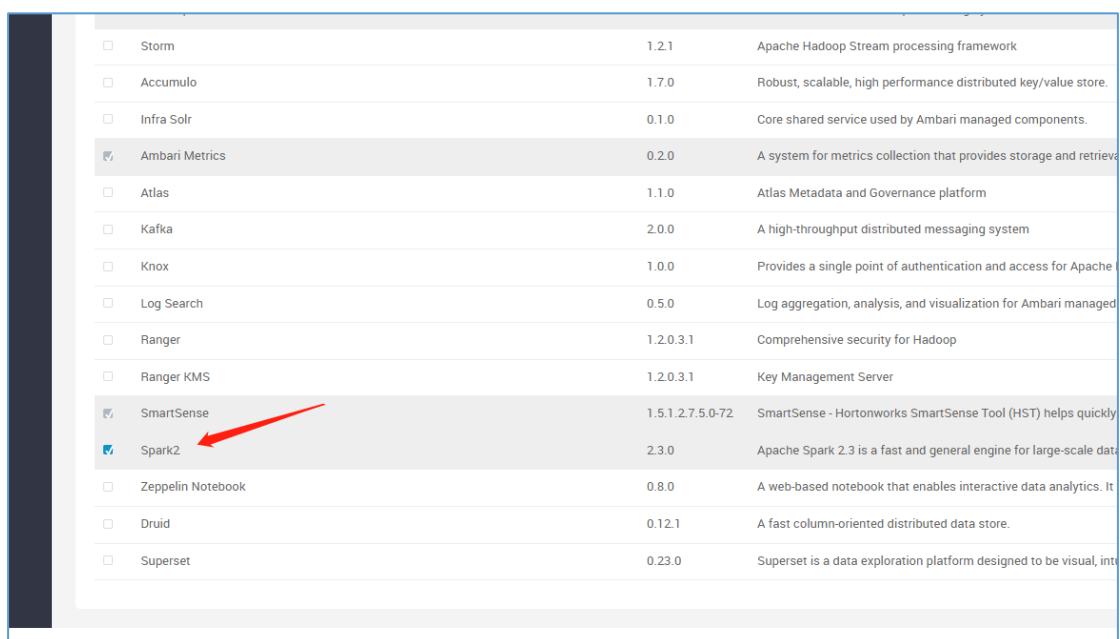
注意：Hive Metastore 和 HiveServer2 要在同一节点。

关键步骤，配置 hive 的数据库



## 6.3 添加 spark 服务

点击 add service, 选择 Spark2



我这里选了 2 个节点安装 spark2 history server，依赖于 hdfs client、mapreduecs2 client、yarn client、hive client 服务

Hive Metastore: 10-11-16-6 dm-ai.com (7.6 GB, 4 cores)

HiveServer2: 10-11-16-6 dm-ai.com (7.6 GB, 4 cores)

Zookeeper Server: 10-11-16-2 dm-ai.com (7.6 GB, 4 cores)

Zookeeper Server: 10-11-16-4 dm-ai.com (7.6 GB, 4 cores)

Metrics Collector: 10-11-16-3 dm-ai.com (7.6 GB, 4 cores)

Grafana: 10-11-16-1 dm-ai.com (7.6 GB, 4 cores)

Activity Analyzer: 10-11-16-1 dm-ai.com (7.6 GB, 4 cores)

Activity Explorer: 10-11-16-1 dm-ai.com (7.6 GB, 4 cores)

HST Server: 10-11-16-1 dm-ai.com (7.6 GB, 4 cores)

Spark2 History Server: 10-11-16-6 dm-ai.com (7.6 GB, 4 cores)

Spark2 History Server: 10-11-16-7 dm-ai.com (7.6 GB, 4 cores)

10-11-16-4 dm-ai.com (7.6 GB, 4 cores)  
Timeline Service V2.0 Header Zookeeper Server

10-11-16-6 dm-ai.com (7.6 GB, 4 cores)  
Hive Metastore HiveServer2 Spark2 History Server

10-11-16-7 dm-ai.com (7.6 GB, 4 cores)  
Spark2 History Server

3 hosts not running master services

注意：spark2 thrift server 依赖 spark2 history server，前者要装到后者所在节点上。

后续步骤略

## 6.4 添加 sqoop 服务

Add Service Wizard

Assign Slaves and Clients

Assign slave and client components to hosts you want to run them on. Hosts that are assigned master components are shown with a . Client will install Sqoop Client.

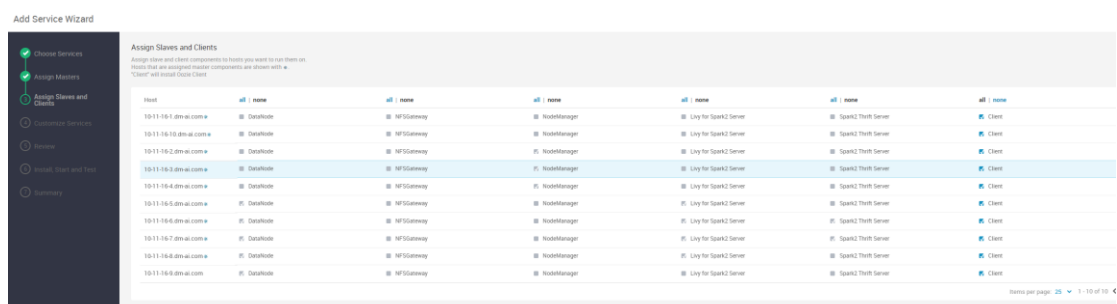
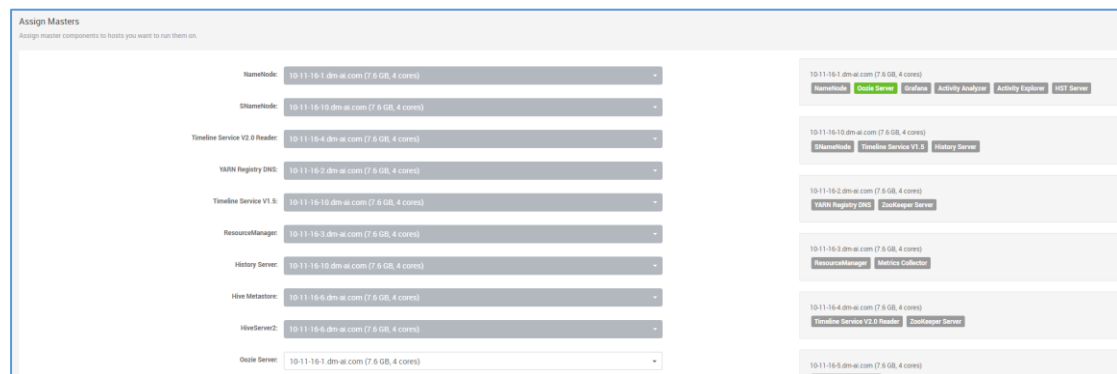
Host	all	name	all	name	all	name	all	name
10-11-16-1 dm-ai.com		DataNode		NFSGateway		NodeManager		Spark2 Thrift Server
10-11-16-10 dm-ai.com		DataNode		NFSGateway		NodeManager		Spark2 Thrift Server
10-11-16-2 dm-ai.com		DataNode		NFSGateway		NodeManager		Spark2 Thrift Server
10-11-16-3 dm-ai.com		DataNode		NFSGateway		NodeManager		Spark2 Thrift Server
10-11-16-4 dm-ai.com		DataNode		NFSGateway		NodeManager		Spark2 Thrift Server
10-11-16-5 dm-ai.com		DataNode		NFSGateway		NodeManager		Spark2 Thrift Server
10-11-16-6 dm-ai.com		DataNode		NFSGateway		NodeManager		Spark2 Thrift Server
10-11-16-7 dm-ai.com		DataNode		NFSGateway		NodeManager		Spark2 Thrift Server
10-11-16-8 dm-ai.com		DataNode		NFSGateway		NodeManager		Spark2 Thrift Server
10-11-16-9 dm-ai.com		DataNode		NFSGateway		NodeManager		Spark2 Thrift Server

Items per page: 25 1 - 10 of 10

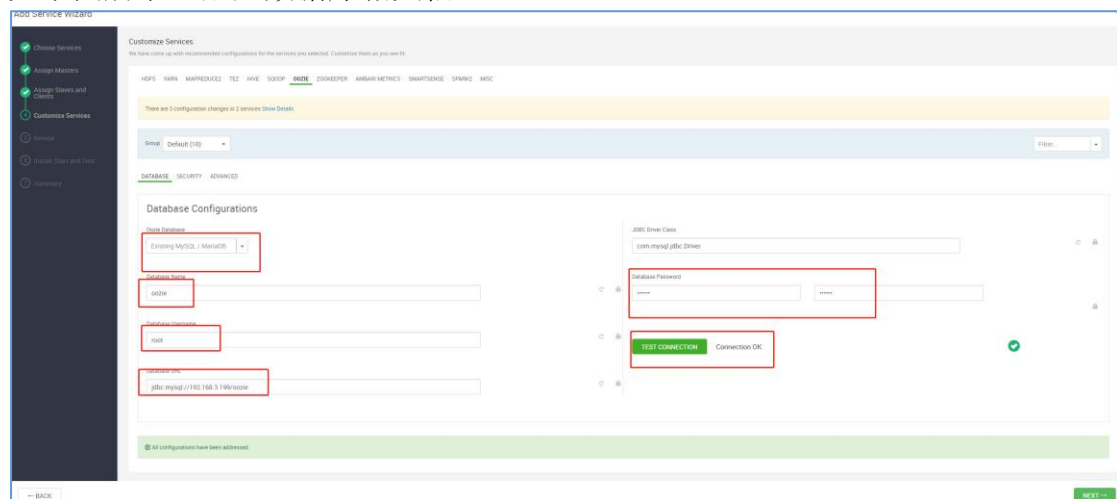
## 6.5 添加 oozie 服务

提前创建好 oozie 库

```
mysql>
mysql> create database oozie;
Query OK, 1 row affected
```



如下图所示，配置数据库相关信息

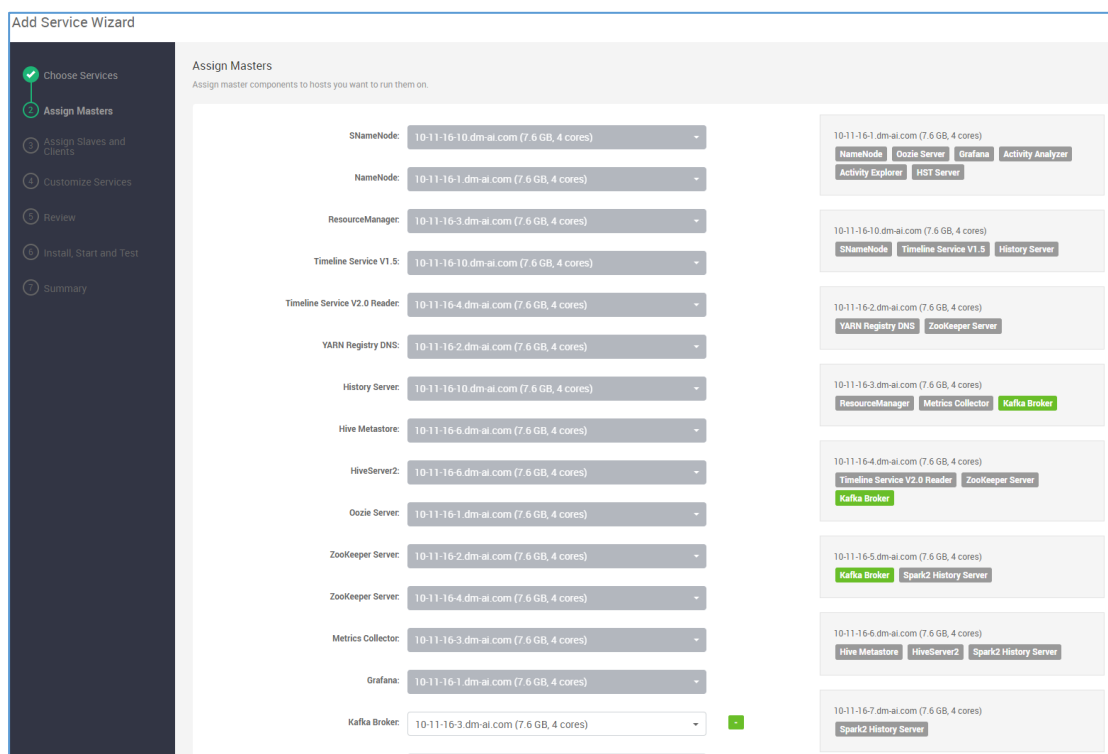


jdbc:mysql://192.168.3.199/oozie

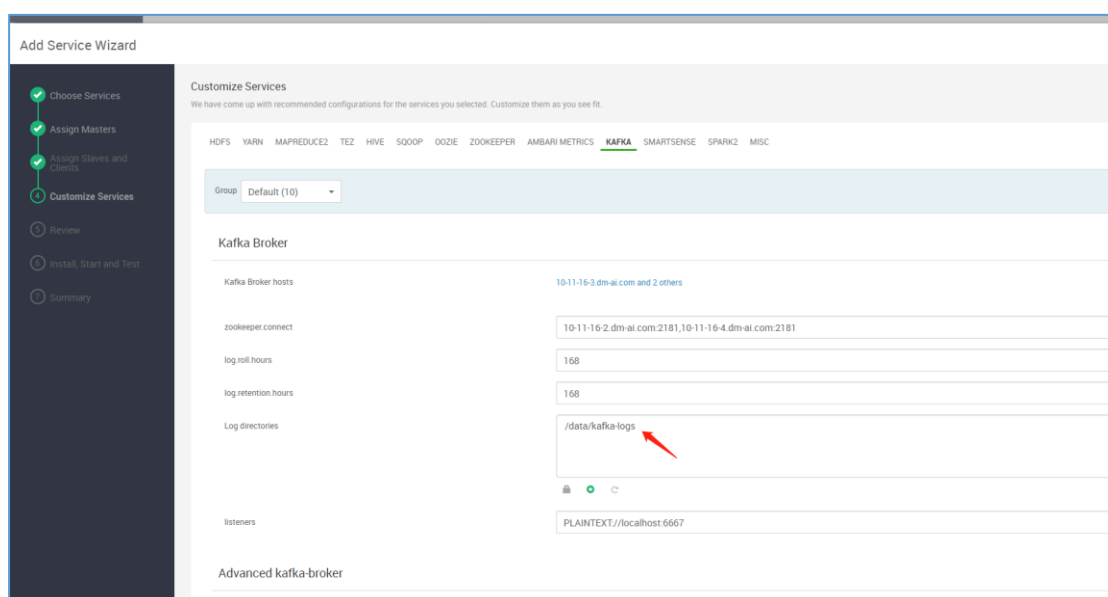
需要测试通过。

下面的步骤略。

## 6.6 添加 kafka 服务



这里我修改了 kafka 的 logs 目录地址



## 6.7 概览

添加完成以上几个服务之后，控制台查看效果如下

Ambari

Dashboard

Services

YARN

Resource Manager

State

HDFS

MapReduce2

Tez

Hive

Sqoop

Oozie

ZooKeeper

Ambari Metrics

Kafka

SmartSense

Spark2

Hosts

Alerts

Cluster Admin

Stack and Versions

Service Accounts

Kerberos

Service Auto Start

Hosts

myhdp

admin

Hosts

Name

IP Address

Rack

Cores

RAM

Disk Usage

Load Avg

Versions

Components

☐

☒

10-11-16-1-dm-ai.com

10.11.16.1

/default-rack

4 (4)

7.60GB

0.18

HDP-3.1

15 Components

☐

☒

10-11-16-10-dm-ai.com

10.11.16.10

/default-rack

4 (4)

7.60GB

0.08

HDP-3.1

12 Components

☐

☒

10-11-16-2-dm-ai.com

10.11.16.2

/default-rack

4 (4)

7.60GB

0.27

HDP-3.1

14 Components

☐

☒

10-11-16-3-dm-ai.com

10.11.16.3

/default-rack

4 (4)

7.60GB

0.37

HDP-3.1

15 Components

☐

☒

10-11-16-4-dm-ai.com

10.11.16.4

/default-rack

4 (4)

7.60GB

0.19

HDP-3.1

15 Components

☐

☒

10-11-16-5-dm-ai.com

10.11.16.5

/default-rack

4 (4)

7.60GB

0.14

HDP-3.1

13 Components

☐

☒

10-11-16-6-dm-ai.com

10.11.16.6

/default-rack

4 (4)

7.60GB

0.09

HDP-3.1

18 Components

☐

☒

10-11-16-7-dm-ai.com

10.11.16.7

/default-rack

4 (4)

7.60GB

0.12

HDP-3.1

13 Components

☐

☒

10-11-16-8-dm-ai.com

10.11.16.8

/default-rack

4 (4)

7.60GB

0.09

HDP-3.1

12 Components

☐

☒

10-11-16-9-dm-ai.com

10.11.16.9

/default-rack

4 (4)

7.60GB

0.03

HDP-3.1

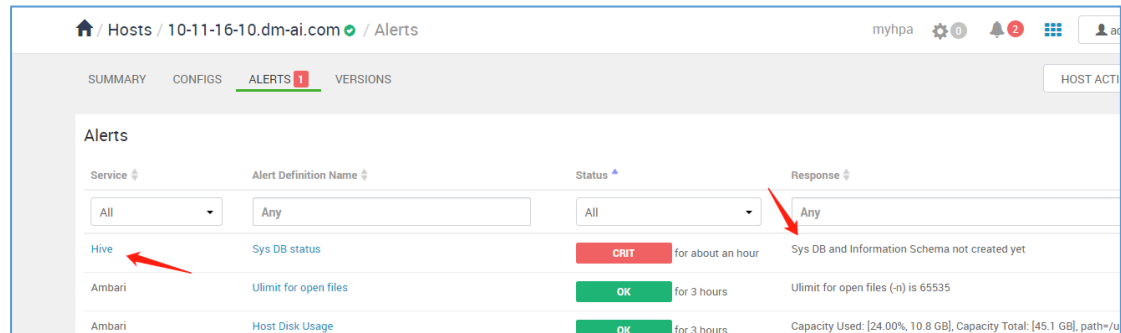
9 Components

Items per page: 101 - 10 of 10

## 7 排错

### 7.1 hive 安装报警

有如下警告：Sys DB and Information Schema not created yet

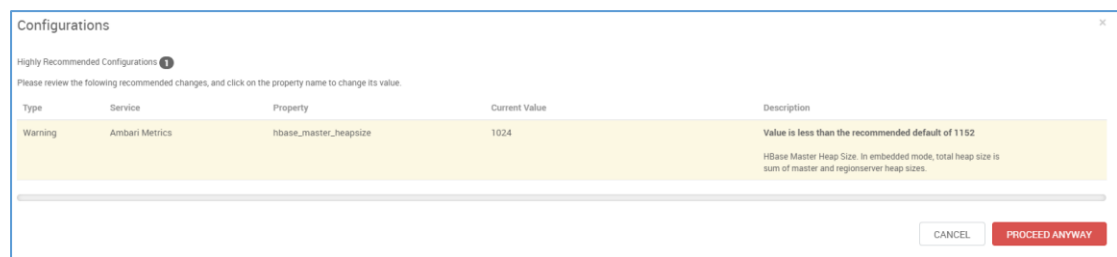


解决方法：进入 hive metastore 所在的节点，执行如下命令即可：

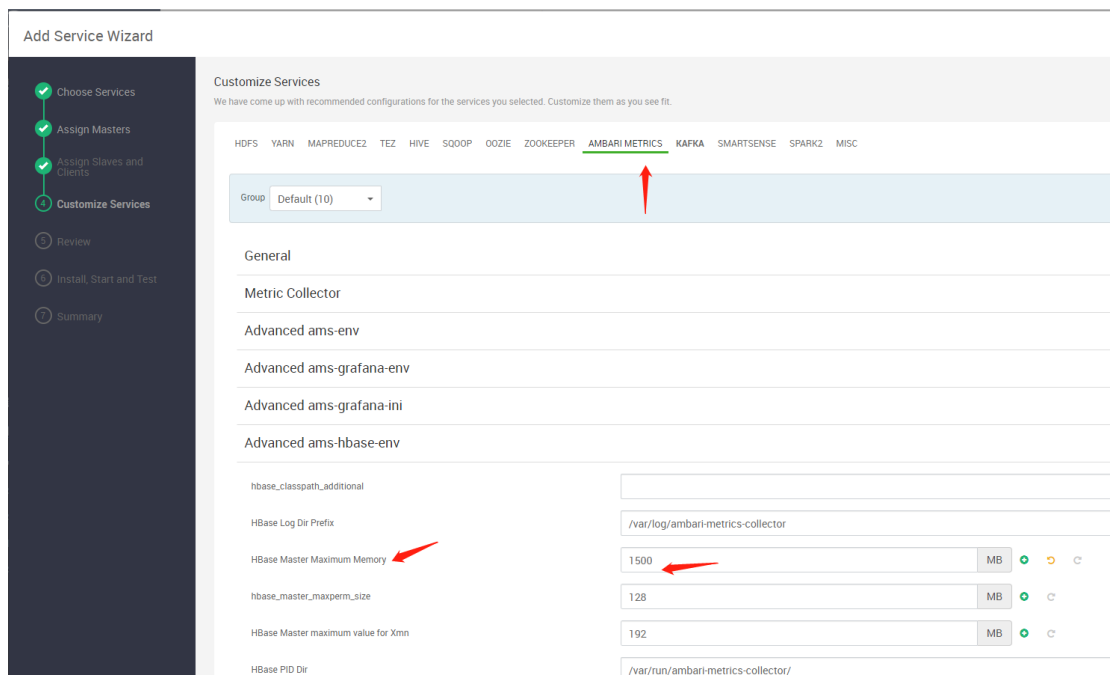
```
touch /etc/hive/sys.db.created
```

### 7.2 kafka 安装报警

如果出现以下警告



如下图，将这个配置改为 1500





以下告警：

Value is lesser than the recommended minimum Xmn size of 242 (12% of hbase\_master\_heapsize + hbase\_regionserver\_heapsize)  
HBase Master maximum value for young generation heap size.

Configurations

Highly Recommended Configurations 1  
Please review the following recommended changes, and click on the property name to change its value.

Type	Service	Property	Current Value	Description
Warning	Ambari Metrics	hbase_master_xmn_size	192	Value is lesser than the recommended minimum Xmn size of 242 (12% of hbase_master_heapsize + hbase_regionserver_heapsize) HBase Master maximum value for young generation heap size.

CANCEL

PROCEED ANYWAY

1500+512 的 12%等于 242，也就是 hbase master maximum value for xmn 至少要是 242，我这里改为 250

hbase\_classpath\_additional

HBase Log Dir Prefix

HBase Master Maximum Memory

hbase\_master\_maxperm\_size

HBase Master maximum value for Xmn

HBase PID Dir

HBase RegionServer Maximum Memory

HBase RegionServer shutdown timeout

hbase\_regionserver\_xmn\_ratio

max\_open\_files\_limit

HBase RegionServer maximum value for Xmn

/var/log/ambari-metrics-collector

1500

128

250

/var/run/ambari-metrics-collector/

512

30

0.2

32768

128

MB

MB

MB

MB

MB

MB

MB

MB