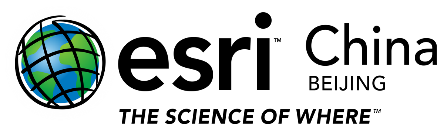
Esri大数据平台搭建指南之Ambari



易智瑞中国信息技术有限公司

2017 年 11 月

**——制定及修订记录——**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **版本** | **完成日期** | **编写/修订纪要** | **编写者** | **备注** |
| **V1.0** | **2017-11-6** | **初稿** | **胡源** |  |
| **V1.1** | **2017-11-7** | **修订** | **胡源** | **内容完善** |
| **V1.2** | **2017-11-9** | **修订** | **胡源** | **写总结** |
|  |  |  |  |  |

目录

[Esri大数据平台搭建指南之Ambari 1](#_Toc498034651)

[**1.** **Ambari简介** 4](#_Toc498034652)

[**2.** **搭建环境概览** 4](#_Toc498034653)

[2.1硬件环境 4](#_Toc498034654)

[2.2 操作系统 5](#_Toc498034655)

[2.3 软件介质 5](#_Toc498034656)

[**3.** **Ambari搭建** 5](#_Toc498034657)

[3.1 Ambari Yum源搭建 5](#_Toc498034658)

[3.2 Ambari Server搭建 6](#_Toc498034659)

[3.3 Ambari Agent搭建 6](#_Toc498034660)

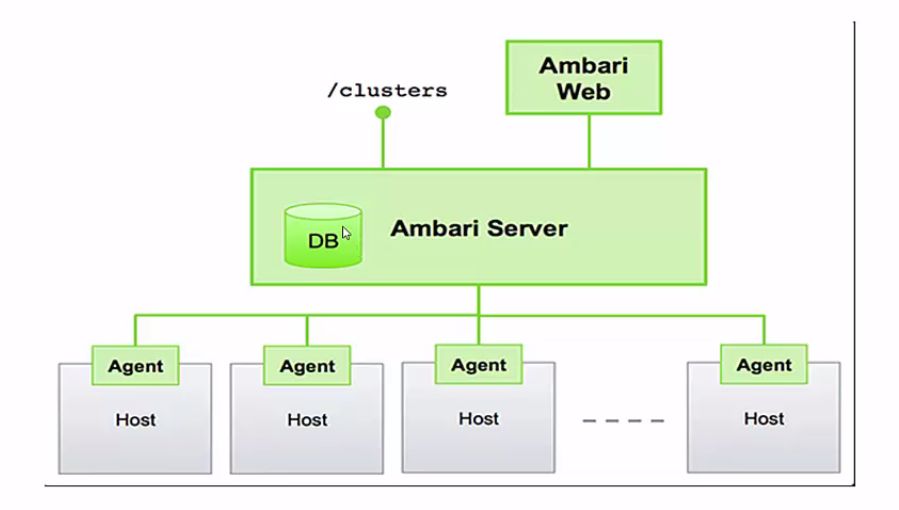
[**4.** **Ambari中创建Hadoop集群** 7](#_Toc498034661)

[**5.** **总结** 8](#_Toc498034662)

1. **Ambari简介**

Apache Ambari是一种基于Web的工具，支持Hadoop（这里指广义的Hadoop生态圈，包括HDFS、MapReduce、Spark、Hive、Hbase、Zookeeper等）集群的创建、管理和监控。

Ambari架构图：



本文介绍了如何搭建Ambari，进而使用Ambari搭建HDFS、Hive等Hadoop生态圈大数据组件，最后写了自己的一些思考。

1. **搭建环境概览**

2.1硬件环境

六台CentOS服务器：

CPU：8核/结点

内存：16G /结点

磁盘：60G/结点

*（提示：yum源机器配置可适当降低）*

2.2 操作系统

CentOS版本: [CentOS-6.8-x86\_64-minimal](http://centos.ustc.edu.cn/centos/6.8/isos/x86_64/CentOS-6.8-x86_64-minimal.iso)（点击下载）

2.3 软件介质

[ambari-2.4.2.0-centos6.tar.gz](http://public-repo-1.hortonworks.com/ambari/centos6/2.x/updates/2.4.2.0/ambari-2.4.2.0-centos6.tar.gz)

[HDP-2.5.3.0-centos6-rpm.tar.gz](http://public-repo-1.hortonworks.com/HDP/centos6/2.x/updates/2.5.3.0/HDP-2.5.3.0-centos6-rpm.tar.gz)

[HDP-UTILS-1.1.0.21-centos6.tar.gz](http://public-repo-1.hortonworks.com/HDP-UTILS-1.1.0.21/repos/centos6/HDP-UTILS-1.1.0.21-centos6.tar.gz)

1. **Ambari搭建**

*提示：可通过直接编译ambari源码的方式获得安装包进行搭建，也可搭建本地yum源来搭建ambari。直接编译的方式费时费力，对带宽和网络稳定性要求也比较高，而且会存在在线依赖更新所导致的版本不一致等诸多问题，故此文档采用搭建本地yum源的方式来搭建ambari。亲测：直接编译的方式会出现各种坑，编译成功前后共花费近一星期的业余时间。*

搭建概览：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 |
| IP地址 | 10.10.10.110 | 10.10.10.111 | 10.10.10.112 | 10.10.10.113 | 10.10.10.114 | 10.10.10.115 |
| 全域名 | yum.bigdata.com | server.bigdata.com | agent1.bigdata.com | agent2.bigdata.com | agent3.bigdata.com | agent4.bigdata.com |
| 搭建角色 | yum源 | ambari server | ambari agent | ambari agent | ambari agent | ambari agent |

3.1 Ambari Yum源搭建

1. 关闭防火墙并安装http服务

service iptables stop

chkconfig iptables off

yum install -y httpd

service httpd start

chkconfig httpd on

1. 拷贝并解压ambari所需yum源

cd /var/www/html

将ambari拷贝至html目录，在html目录下创建目录hdp，将HDP、HDP-UTILS包拷贝至此目录。

tar zxvf ambari-2.4.2.0-centos6.tar.gz

cd hdp

tar zxvf HDP-2.5.3.0-centos6-rpm.tar.gz

tar zxvf HDP-UTILS-1.1.0.21-centos6.tar.gz

3.2 Ambari Server搭建

1）安装ntp、yum-plugin-priorities

yum install -y ntp yum-plugin-priorities

2）关闭防火墙和selinux防火墙，设置ntp服务自动启动

service iptables stop

chkconfig iptables off

service ntpd start

chkconfig ntpd on

setenforce 0

vi /etc/selinux/config 设置SELINUX=disabled

reboot

3）下载[ambari.repo](http://public-repo-1.hortonworks.com/ambari/centos6/2.x/updates/2.4.2.0/ambari.repo)文件，将baseurl和gpgkey中默认值改为搭建的本地yum源地址，并拷贝至/etc/yum.repos.d/目录，设置好hosts文件（包含这几台机器IP与全域名的映射）拷贝至/etc目录。

4）安装ambari-server

yum install -y ambari-server

5）网上下载jdk-7u67-linux-x64.tar.gz、jce\_policy-8.zip、jdk-8u77-linux-x64.tar.gz、UnlimitedJCEPolicyJDK7.zip四个文件并拷贝至/var/lib/ambari-server/resources目录。

6）设置ambari server

ambari-server setup

如无特殊需求，一路回车即可。

7）启动ambari server

ambari-server start

3.3 Ambari Agent搭建

*提示：通过生成SSH秘钥，创建集群时选择密钥文件也可以，这里我选择的是在每一台机器上安装ambari agent。*

1）安装ntp、yum-plugin-priorities

yum install -y ntp yum-plugin-priorities

2）关闭防火墙和selinux防火墙，设置ntp服务自动启动

service iptables stop

chkconfig iptables off

service ntpd start

chkconfig ntpd on

setenforce 0

vi /etc/selinux/config 设置SELINUX=disabled

echo never > /sys/kernel/mm/redhat\_transparent\_hugepage/defrag

echo never > /sys/kernel/mm/redhat\_transparent\_hugepage/enabled

并在/etc/rc.d/rc.local文件中加入上面两行命令

3）下载[ambari.repo](http://public-repo-1.hortonworks.com/ambari/centos6/2.x/updates/2.4.2.0/ambari.repo)文件，将baseurl和gpgkey中默认值改为搭建的本地yum源地址，并拷贝至/etc/yum.repos.d/目录，设置好hosts文件（包含这几台机器IP与全域名的映射）拷贝至/etc目录。

4）安装ambari-agent

yum install -y ambari-agent

5）设置ambari-agent中参数

vi /etc/ambari-agent/conf/ambari-agent.ini

修改server中hostname为ambari-server全域名

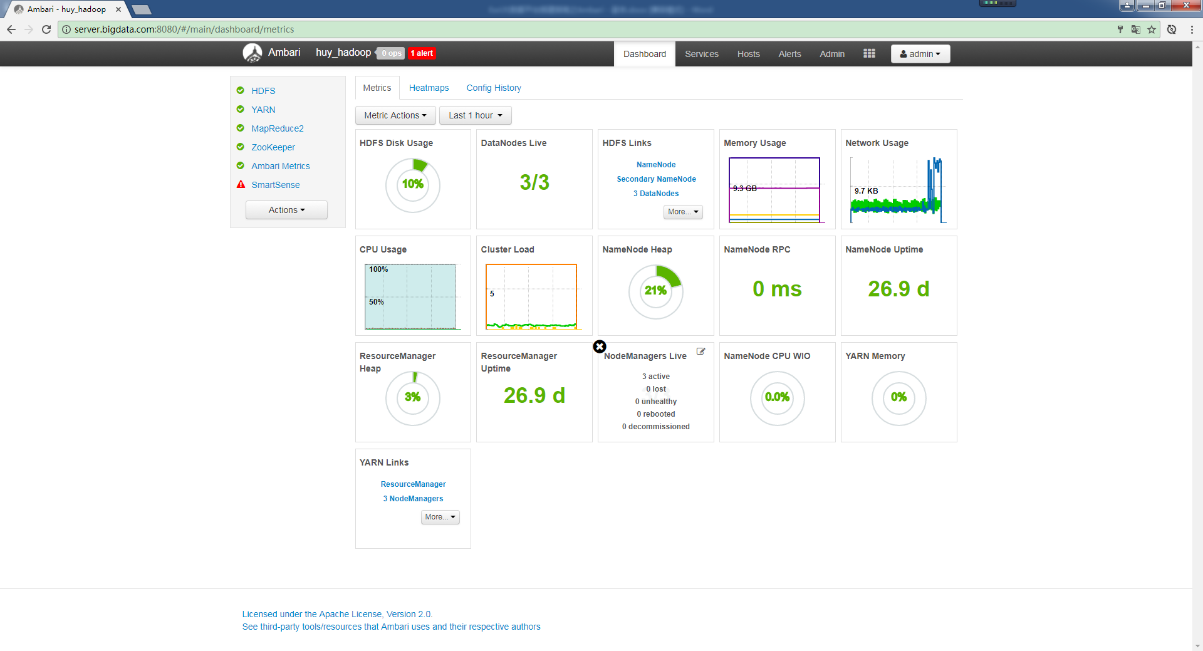
6）启动ambari-agent

ambari-agent start

1. **Ambari中创建Hadoop集群**

在浏览器中打开ambari server IP地址（或全域名）:8080，输入默认账号admin，密码admin，进入ambari主页面。然后根据自己的应用需要，设置集群中机器、机器角色、修复报错和警告信息，一步一步创建集群。下图是因为之前已经创建了集群，所以自动进入了集群的Dashboard页面。

创建过程很简单，大家可以自己动手感受一下，这里就不赘述了。



从图中可以看出，ambari可以很方便地监控整个Hadoop集群的运行状况，使Hadoop管理人员对集群情况有个直观的了解。

同时，ambari也可监控集群中每一台机器的运行情况和资源占用情况，并按需动态地开启、关闭、添加、移除大数据组件服务和更改服务配置，着实是大数据集群管理利器。

1. **总结**

从整个搭建过程来看，确实是比分开搭建Hadoop生态中单独的组件要简单地多，而且最关键的两点：一劳永逸，方便管理。当然，事物都有它的两面性，通过Ambari的方式搭建会消耗更多的资源，需要更高配置的硬件资源。

搭建体验上，感觉有点儿像高级版的ArcGIS Enterprise Builder（能跨机器的、能选择安装内容的，这样比喻还是不太恰当，意会就行），HDFS、Hive、HBase这些组件就相当于Server、Portal、DataStore。