Centos 7+CDH5.7.2全部署流程 原 荐

my.oschina.net 2018/05/26 19:15

八年Java开发的感悟: 什么才是程序员的立身之本 >>>

这个配置是我在网上看到的, 我就借用了这个配置:

主节点: 8g内存、硬盘80g

从节点: 2g内存、硬盘80g

安装系统的时候,我建议将IP和主机名都给配置好,这样就省的在系统中进行配置了,当然下面也有在系统中配置的方法。虚拟机系统安装可以参见:虚拟机:CentOS 7系统安装

所以如果你要使用虚拟机来完成这个配置的话,那么,你的主机的配置内存就不能低于16G了。而且还要找一个盘符较大的磁盘空间,不然到后面操作CDH的时候会很尴尬,我第一次就被卡在了硬盘空间不足上,然后又重新找了一个较大的盘符进行了重新安装,包括虚拟机系统。

1.JDK版本: jdk-8u131-linux-x64.tar.gz

2.Scala版本: scala-2.11.0.tgz

以上两个包建议使用rpm或者yum安装,这样可以省点事,不过我是使用的这两个包。

3.jdbc连接包:mysql-connector-java-5.1.38-bin.jar

4.MySQL相关包:

 $My SQL\text{-}client\text{-}5.6.29\text{-}1.linux_glibc2.5.x86_64.rpm$

 $My SQL-server-5.6.29-1.linux_glibc 2.5.x 86_64.rpm$

5.CDH安装相关的包:

cloudera manager包: cloudera-manager-centos7-cm5.7.2_x86_64.tar.gz

CDH包:

CDH-5.7.2-1.cdh5.7.2.p0.11-el7.parcel

CDH-5.7.2-1.cdh5.7.2.p0.11-el7.parcel.sha1

注意centos要下载el7的。

6.系统使用Linux的CentOS 7版本,CentOS 7和CentOS 6.5在命令上有一些区别,这里会进行说明。

IP地址 主机名 说明 192.168.75.41 cdh01主节点、从节点 192.168.75.42 cdh02 从节点 192.168.75.43 cdh03 从节点

IP地址

主机名

说明

192.168.75.41

cdh01

主节点、从节点

192.168.75.42

cdh02

从节点

192.168.75.43

cdh03

从节点

修改每台节点服务器的有关配置,包括IP、hostname、selinux,防火墙。

如果你在安装CentOS7系统的时候已经指定了主机名和IP地址,那么IP修改和hostname修改就可以省略。但是这里面要强调一点的是,如果你想使用yum安装,那么DNS一定要配置。

这里要使用固定IP,而centOS 7已经没有setup命令,所以图形界面的设置已经没有了,这里只能修改配置文件,而且配置文件的名称还不固定,每个人的都可能不一样,这里我说一下位置,IP配置文件的位置:/etc/sysconfig/network-scripts/,如下图:

这里我的配置文件名称叫做ifcfg-ens33。而CentOS 7系统竟然连vim也没有了,只有vi,所以就是如下:

```
INTELLEMENT IN THE PROOF IN TH
```

如上图: BOOTPROTO、ONBOOT是需要修改的,剩下四项是需要手动添加的。因为后面要使用yum安装,所以这里必须配置DNS,测试环境可以将DNS配置成网关的地址。

如上图,重启网络服务,这个命令没有更改,然后查看ip。

每台都进行这样的配置。

分别对三台都进行更改,并且注意每台名称和ip,每台都要配上。

此处修改完成需要重启才能生效,也可以使用如下方法,立即生效:

CentOS 7.0默认使用的是firewall作为防火墙,相关操作变动很大,和6.5版本的一点也不一样,操作如下:

所有节点都需要关闭,这是官方文档要求的。

重启才能生效, 重启后检查。

安装过程中master需要各个节点的root免登录密码, 先在master上生成公钥:

分别对每台节点进行以上操作,操作完了可以通过ssh ip进行验证是否可以无需输入密码就能直接登录到节点服务器上去。

操作如下:

内容如上,保存退出,然后可以使用scp命令将这个配置拷贝给其他主机。

网上很多博客都提到这个配置,说是让各个节点统一时间,但是我没有配置,我也将这个配置放到这里:

注释掉所有server *.*.*的指向,新添加一条可连接的ntp服务器(百度一下ntp服务器)

server ntp.sjtu.edu.cn iburst

在其他节点上把ntp指向master服务器地址即可(/etc/ntp.conf下)

server 192.168.160.130 iburst 所有节点。

也可参见: NTP服务搭建

针对这个步骤,你可以看网上很多的博客都是将这个步骤放到了MySQL安装的后面,但是经过本人的测试,这一步还是放在前面比较好,因为后面安装的软件,但凡有依赖的,你在这里已经做好了,所以在这里操作安装依赖包,并且是所有节点都安装。

注意这个地方依赖包一定要安装完整了,会经过三个y/N的过程,全部选择Y,这个安装的过程完全看你的网速了,基本敲完这个命令你就可以休息一会了。

这个环节只需要在主节点上进行即可。

在cdh01上准备mysql的jar包:

修改jar包的名字,并拷贝到/usr/share/java/目录:

注意上述操作,是修改了jar包的名字,因为后面有操作会寻找这个jar包的这个固定的名字, 所以这里要修改一下名字。

这个是所有服务的基础,每个节点都需要安装。

安装之前首先检查一下,你的服务器是否默认安装了OpenJDK,如果系统自带了,就需要卸载,步骤如下:
如果没有安装,则可以跳过此步骤,我的系统没有自带jdk。
如下两个方法,任选其一,不要都做。

直接使用命令进行安装即可,这里不做过多解释。

将准备的好的安装包上传,我习惯的目录是:/home/software,上传完成,进行解压:

如果你还要安装spark,可以将环境变量和Scala一起配置。

配置环境变量:

在文件中追加一下内容:

保存退出之后,这里还需要设置一个软连接:

如果你使用的是yum或者rpm安装的jdk那么这些步骤就可以省略。因为CDH平台安装的时候 默认寻找的jdk路径为/usr/java。

如果你要安装Spark一定不能省略Scala的安装,否则,你在web端操作CDH的时候,会给你报错,网上的解决方法,基本都是说的jdk没人说到Scala。

这个操作同样也是在所有的节点上。

目录还是在/home/software。操作如下:

然后配置环境变量:

同样在末尾追加以下内容:

MySQL的安装只需要在主节点进行即可。

安装MySQL也同样需要检查,系统是否自带了,centos7自带的是mariadb,首先检查,命令如下:

如上的情况是存在,如果存在就使用以下命令进行卸载:

将下载好的MySQL rpm包拷贝到服务器上,然后解压、安装。

将mysqld加入系统服务,并随机启动,命令如下:

首先获取mysql安装时root用户的随机密码: 也可使用cat命令查看: 此密码只能用来修改密码使用。 必须要修改root用户的密码才可以使用mysql,否则只能连接不能操作 如果出现没有权限的问题,在mysql授权(在安装mysql的机器上执行),执行下面的语句,进 行授权: *.*:所有库下的所有表。 %: 任何IP地址或主机都可以连接。如果%配置不生效,就配置具体的主机名称。 如下命令是创建部署各个服务所需的数据库,我本人倾向用不用先创建好,用的时候就可以 直接部署服务了,不必再来数据库进行创建。 将Cloudera相关的四个包,如下: cloudera-manager-centos7-cm5.7.2_x86_64.tar.gz CDH-5.7.2-1.cdh5.7.2.p0.11-el7.parcel CDH-5.7.2-1.cdh5.7.2.p0.11-el7.parcel.sha1 manifest.json 上传到所有的服务器上, 在所有的服务器上执行以下操作。 解压cm tar包到指定目录,先创建目录,命令操作如下: 创建cloudera-scm用户 配置从节点cloudera-manger-agent指向主节点服务器,我现在的集群规划,是需要在每台 服务器上都进行如下配置: 将server_host改为CMS所在的主机名即cdh01

此操作只需在主节点进行即可,在主节点中创建parcel-repo仓库目录,命令如下:

注意: 其中CDH-5.7.2-1.cdh5.7.2.p0.18-el5.parcel.sha1 后缀要把1去掉,操作如下:

启动mysqld的命令如下:

在所有的节点上创建parcels目录,操作如下:

解释: Clouder-Manager将CDH从主节点的/opt/cloudera/parcel-repo目录中抽取出来,分发解压激活到各个节点的/opt/cloudera/parcels目录中。

此操作在主节点上进行,初始脚本配置数据库scm_prepare_database.sh,操作命令如下:

说明:这个脚本就是用来创建和配置CMS需要的数据库的脚本。各参数是指:

mysql:数据库用的是mysql,如果安装过程中用的oracle,那么该参数就应该改为oracle。

-hcdh01:数据库建立在cdh01主机上面。也就是主节点上面。

-uroot: root身份运行mysql。-proot: mysql的root密码是root。

--scm-host cdh01: CMS的主机,一般是和mysql安装的主机是在同一个主机上。

最后三个参数是:数据库名,数据库用户名,数据库密码。

执行完成命令正常如下:

在这个地方就可以解释上面为什么要改jar名了。

此命令只需在主节点执行即可。

进入/opt/cloudera-manager/cm-5.7.2/etc/init.d/目录中,执行如下命令:

启动cloudera-scm-agent,需要在所有的从节点上启动。

进入/opt/cloudera-manager/cm-5.7.2/etc/init.d/目录中,执行如下命令:

启动如下图所示:

```
[root@cdh01 software]# /opt/cloudera-manager/cm-5.7.2/etc/init.d/cloudera-scm-se rver start
Starting cloudera-scm-server: [ 端定 ]
[root@cdh01 software]# /opt/cloudera-manager/cm-5.7.2/etc/init.d/cloudera-scm-ag ent start
Starting cloudera-scm-agent: [ 端定 ]
```

这里要说明一下,基本上上述两项启动都会配置随机启动,但是我实验了很多个随机启动的 方式,都不管用,这里暂时不做,后期我找到可以随机启动的方式再补充。

注:以上环节只要操作的没有问题,以下的步骤将不会出现任何问题。

在浏览器中输入192.168.75.41:7180。

出现下图登录界面,默认用户名和密码: admin

admin	
	<u> </u>
登录	
□ 在此计算机上保留我的信息。	

出现这个界面说明CM已经安装成功了,下面就在这个web界面中部署各种服务了。

此过程中没有提到的界面, 默认选择继续。

欢迎使用 Cloudera Manager

节点限制

Cloudera Manager 核心功能

CDH

返回

当登录之后,会进入选择express版本的界面,在此界面选择免费即可,然后继续。

无限制

紫想要部署哪个版本?

在服务器的各个从节点已经安装并启动了agent,还在各个节点都配置了server指向,所以各个节点的agent就会给server发消息报告,这里可以在"当前管理的主机"中看到三个主机,全部勾选并继续。

1 2

无限制

无限制

注意如果cloudera-scm-agent没有设为开机启动,如果以上有重启这里可能会检测不到其他服务器。之前配置的时候,我就没有配置随机启动。

为 CDH 群集安装指定主机。



这里选择自己需要的版本即可。这里我选择了两项,其一选择方式使用Parcel;其二CDH版本选择的是CDH-5.7.2-1 cdh5.7.2p0 18;其他均选择无。

群集安装



此步是分发parcels到各个节点,上一步继续,这一步是自动进行的不需要什么操作,唯一需要的就是等待。当全部完成点击继续。

群集安装

正在安装选定 Parcel

选定的 Parcel 正在下载并安装在群集的所有主机上。



此步是对主机配置的正确性进行检测。

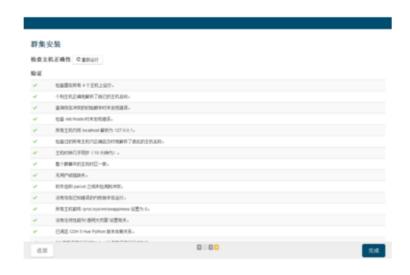
群集安装



这个地方要注意这个地方有两项没有检查通过,但是我没做处理,我在网上找到了一个方法,但是不知道具体是否管用,方式如下:

可以在集群中使用以下命令

然后再点击上面的重新运行会发现这次全部检查通过了。



我没有做上面的处理, 但是后面的安装也没有影响。

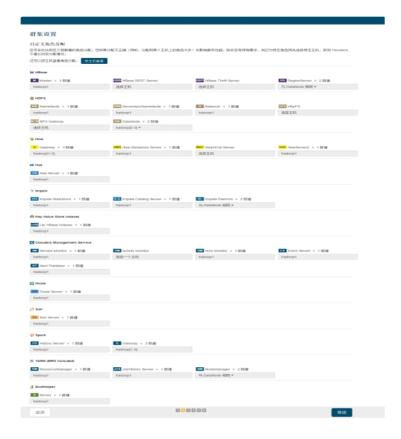
重头戏来了,这里要选择要安装使用的服务,CDH5默认提供了如图的六种搭配,当然那也可以选择则自定义,这里我选择的就是自定义。

群集设置

选择您要在群集上安装的 CDH 5 服务。



这里就是进行各个服务角色分配给那个节点的,可以默认,也可以自己做选择。

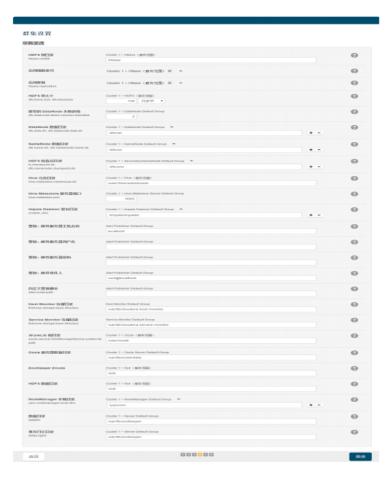


这里就用到了前面创建的各个数据库,根据你选择的服务,这里会让你填写每个服务用到的 数据库,以及用户名和密码。



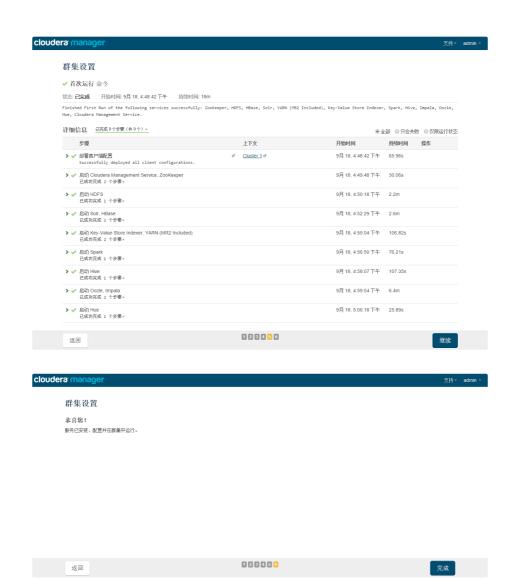
- 创建数据库时,**数据库主机名称**字段中的值必须与度用于主机名的的值匹配。 <u>了越更多</u>。 加数据库未在其默认端山运行,请使用 **数据库主机名称**字段中的 host:port 指定端口号。 强烈建议将各个数据库与相应角色实例常于同一主机 上。

这里就都是默认的就好,除非你有规划,可以在这里进行更改。



如果web引导安装会出问题,那么问题都是出在这一步,这里会出现的问题,我遇到的问题,基本都在前面做了介绍,就是安装jdk和Scala的时候。基本上其他的不会出什么问题。





这个时候安装完成了,点击完成,就会进入如下界面。

查看集群情况,我这里有挺多报警,大概查看下基本都是内存或者存储空间使用阈值的报 警,由于我是本地虚拟机的,所以这些条件都有限,这里暂时不处理这些报警。



以上就是安装的全部过程。

下一篇: CDH的坑之Sqoop导出数据到MySQL