**[CDH6.2离线安装（史上最全安装文档）](https://www.cnblogs.com/swordfall/p/10816797.html)-Lyer**

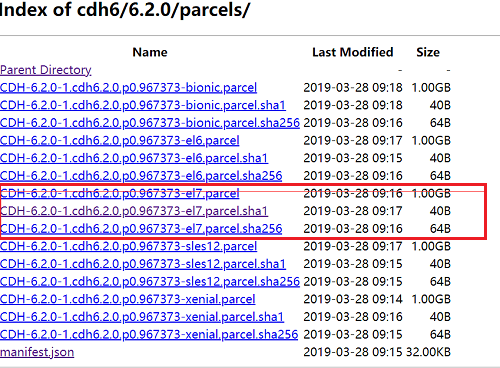
## 1 CDH准备

CDH分为Cloudera Manager管理平台和CDH parcel（parcel包含各种组件的安装包）。这里采用CDH6.2.0。

小知识：官方下载地址现在增加了收费墙,现在CDH包不能再使用在线安装的方式，全部使用离线方式，离线要搭建自己的源，收费墙这个是有史以来最坑的地方，关于CDH资源包，现在都是自己网上找或者群内大家共享。

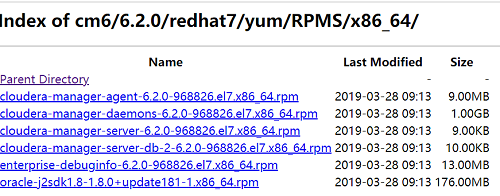
真实线上案例：增扩容节点，使用半离线方式，无自有源，因为收费墙问题，一直造成失败，怀疑人生好几天，没想到CDH官方21年增加收费墙。

资源目录如下：



CDH6.2.0安装包地址为：[https://archive.cloudera.com/cdh6/6.2.0/parcels/](https://archive.cloudera.com/cdh6/6.2.0/parcels/" \t "_blank)

由于操作系统为CentOS7，需要下载以下文件：



上述文件整理资料百度云下载地址为：

链接: https://pan.baidu.com/s/1Dm5Elf9uQqn14BUbgU3AFQ 提取码: mws3 (不一定能用，还是要靠你自己找)

说明：以下操作都是在root用户下进行的

## 2 安装

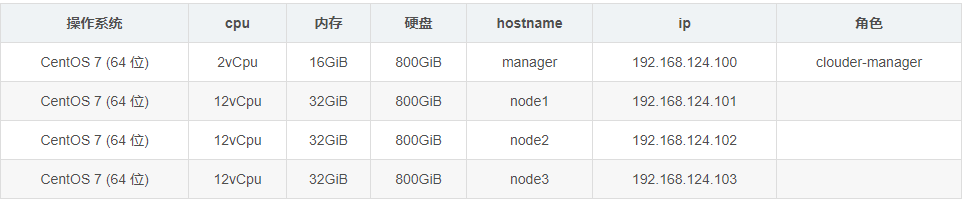
## 2.1 环境准备

### 2.1.1. 准备虚拟机（根据自己的系统资源分配虚拟机资源）

系统版本：centos7+

小知识：系统不要装乱七八糟的东西，RAID最好是0，增加读写性能，hadoop环境天生高可用，不用考虑硬盘损坏，反而使用RAID5的方式，磁盘损坏后，会影响整个集群性能。

真实线上案例：集群莫名降速，某节点磁盘损坏



### 2.1.2. 静态IP设置（每个节点）

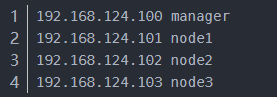
vim /etc/sysconfig/network-scripts/ifcfg-eth0

service network restart 重启网络生效

yum install -y net-tools ifconfig查看设置

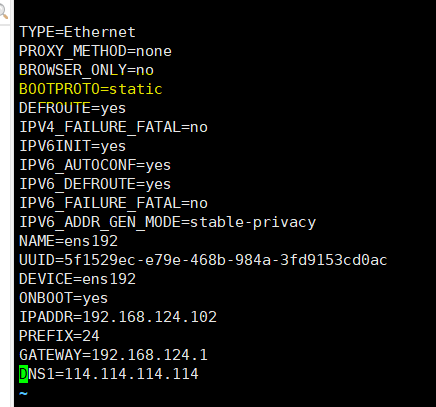
### 2.1.3. 编辑/etc/hosts文件（每个节点）

vim /etc/hosts



### 2.1.4. 关闭防火墙、禁止防火墙开机自启（每个节点）

* systemctl stop firewalld 关闭防火墙
* systemctl disable firewalld 禁止防火墙开机自启
* vim /etc/selinux/config —> SELINUX=disabled (修改)



### 2.1.5. ssh无密码登录

* manager节点执行ssh-keygen -t rsa 一路回车到结束，在/root/.ssh/下面会生成一个公钥文件id\_rsa.pub
* cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys 将公钥追加到authorized\_keys
* chmod 600 ~/.ssh/authorized\_keys 修改权限
* 将~/.ssh从当前节点分发到其他各个节点。如：scp -r ~/.ssh/ root@node1:~/.ssh/
* ssh 各个节点互相登陆

### 2.1.6. 配置NTP服务（所有节点）

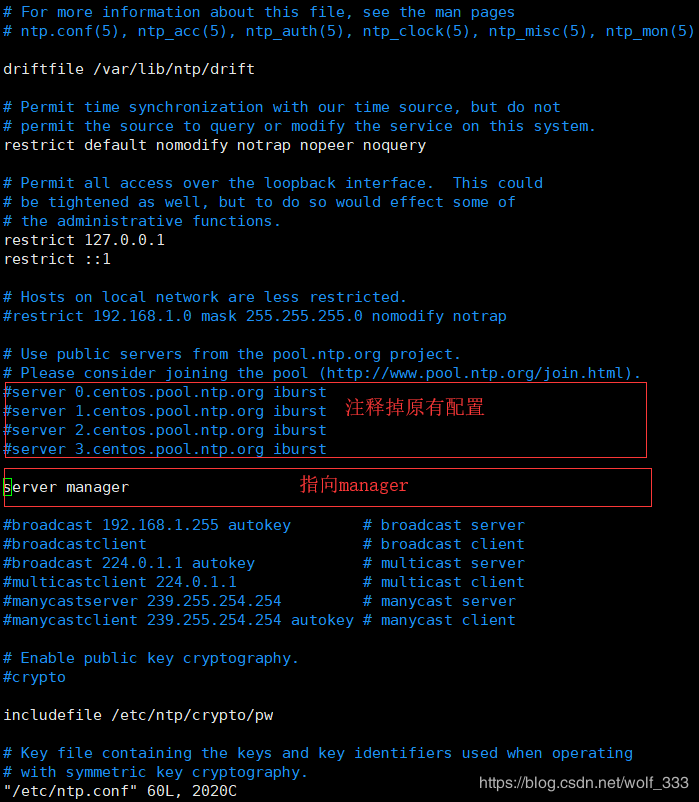
小知识：hadoop环境时间一致性特别重要

真实线上案例：虚拟机重启后，时间和其它节点不一致，各种组件问题。

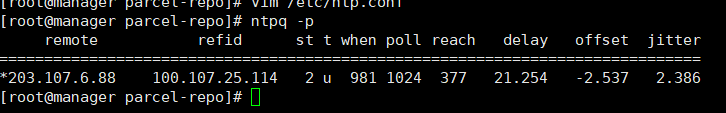
* 修改时区（改为中国标准时区）ln -sf /usr/share/zoneinfo/Asia/Shanghai /etc/localtime
* 安装ntp yum -y install ntp
* ntp主机配置 vim /etc/ntp.conf
* manager节点

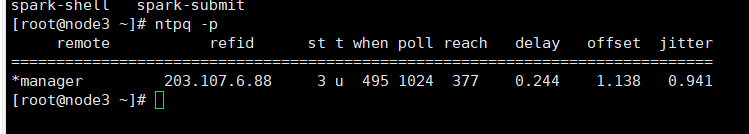


其余节点



* 重新启动 ntp 服务：service ntpd restart
* 设置开机自启：systemctl enable ntpd.service
* ntpdc -c loopinfo #查看与时间同步服务器的时间偏差
* ntpq -p #查看当前同步的时间服务器
* ntpstat #查看状态
* 配置成功状态（服务开启后前面出现\*说明成功）：

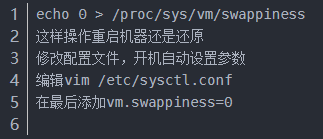




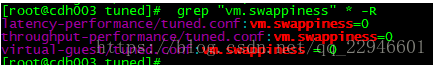
### 2.1.7. 修改Linux swappiness参数(所有节点）

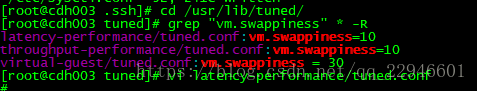
为了避免服务器使用swap功能而影响服务器性能，一般都会把vm.swappiness修改为0（cloudera建议10以下）

* 上述方法rhel6有效，rhel7.2中:tuned服务会动态调整系统参数
* 查找tuned中配置，直接修改配置
* cd /usr/lib/tuned/
* grep “vm.swappiness” \* -R 查询出后依次修改



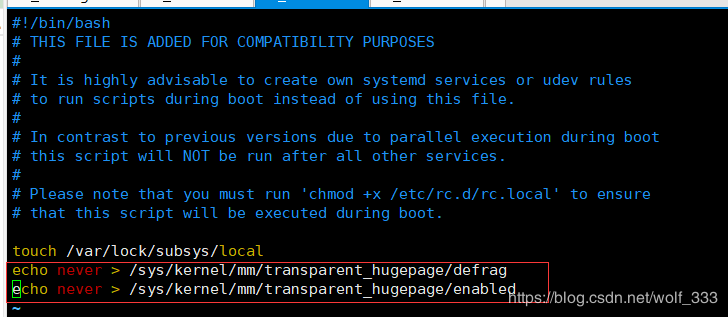
* 上述方法rhel6有效，rhel7.2中:tuned服务会动态调整系统参数
* 查找tuned中配置，直接修改配置
* cd /usr/lib/tuned/
* grep “vm.swappiness” \* -R 查询出后依次修改





### 2.1.8. 禁用透明页(所有节点）

echo never > /sys/kernel/mm/transparent\_hugepage/defrag  
echo never > /sys/kernel/mm/transparent\_hugepage/enabled  
永久生效 在/etc/rc.local 添加上面命令



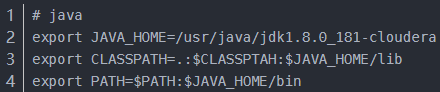
给与可执行权限：chmod +x /etc/rc.d/rc.local

### 2.1.9. JDK安装（所有节点）

小知识：不要用你自己的JDK，用这个JDK，能避免很多很多的问题，你如果非要用，那我也没办法~

真实线上案例：某节点扩容，为了省略这步，用yum随便装了一个，节点集成CDH后，各种版本不兼容问题。

* rpm -qa | grep java # 查询已安装的java
* yum remove java\* # 卸载
* rpm -ivh oracle-j2sdk1.8-1.8.0+update181-1.x86\_64.rpm
* vi /etc/profile 末尾添加



* source /etc/profile
* java -version验证

### 2.1.10. 创建/usr/share/java目录，将mysql-jdbc包放过去（所有节点）

* mkdir -p /usr/share/java
* mv /opt/mysql-j/mysql-connector-java-5.1.34.jar /usr/share/java/
* mysql-connector-java-5.1.34.jar 一定要命名为mysql-connector-java.jar

### 2.1.11. 为保证防火墙、虚拟机参数修改后生效，各节点机器需要重启 reboot

### 2.1.12. Mysql安装

卸载原生的mariadb，安装mysql：

* rpm -qa|grep mariadb
* rpm -e --nodeps mariadb-libs-5.5.60-1.el7\_5.x86\_64
* cd /opt/mysql/
* tar -xvf ./mysql-5.7.19-1.el7.x86\_64.rpm-bundle.tar
* rpm -ivh mysql-community-common-5.7.19-1.el7.x86\_64.rpm
* rpm -ivh mysql-community-libs-5.7.19-1.el7.x86\_64.rpm
* rpm -ivh mysql-community-client-5.7.19-1.el7.x86\_64.rpm
* rpm -ivh mysql-community-server-5.7.19-1.el7.x86\_64.rpm
* rpm -ivh mysql-community-libs-compat-5.7.19-1.el7.x86\_64.rpm

MYSQL配置如下:

* mysqld --initialize --user=mysql # 初始化mysql使mysql目录的拥有者为mysql用户
* cat /var/log/mysqld.log # 最后一行将会有随机生成的密码
* systemctl start mysqld.service # 设置mysql服务自启
* mysql -uroot –p 如果不能登陆
* systemctl restart mysqld
* #登录并修改mysql的管理者密码

　　　　$> mysql -u root  
　　　　mysql> use mysql;  
　　　　mysql> set password = PASSWORD('root');  
　　　　mysql> exit;

* 创建库(后续安装服务等使用)

create database cmserver default charset utf8 collate utf8\_general\_ci;

grant all on cmserver.\* to 'cmserveruser'@'%' identified by 'root';

create database metastore default charset utf8 collate utf8\_general\_ci;

grant all on metastore.\* to 'hiveuser'@'%' identified by 'root';

create database amon default charset utf8 collate utf8\_general\_ci;

grant all on amon.\* to 'amonuser'@'%' identified by 'root';

create database rman default charset utf8 collate utf8\_general\_ci;

grant all on rman.\* to 'rmanuser'@'%' identified by 'root';

create database oozie default charset utf8 collate utf8\_general\_ci;

grant all on oozie.\* to 'oozieuser'@'%' identified by 'root';

create database hue default charset utf8 collate utf8\_general\_ci;

grant all on hue.\* to 'hueuser'@'%' identified by 'root';

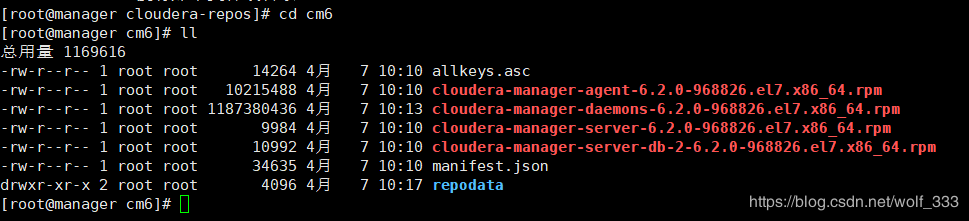
### 2.1.13. 安装Httpd服务（manager）

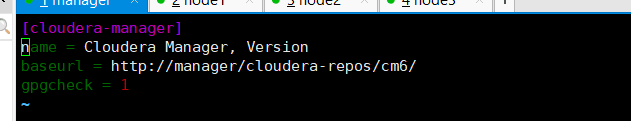
小知识：这个东西，是用来做自己源的，离线安装必备

真实线上案例：为了脱离收费墙，只能离线，幸亏之前服务器有原始CDH包版本的备份，要不全网找CDH固定版本，是非常非常非常非常非常麻烦的。

* yum install httpd
* systemctl start httpd
* systemctl enable httpd.service 设置httpd服务开机自启

### 2.1.14. 配置Cloudera Manager包yum源（manager节点）

* mkdir -p /var/www/html/cloudera-repos/
* 将下载的cm包文件移到此目录下:
* mv cm6 /var/www/html/cloudera-repos/
* cd /var/www/html/cloudera-repos/cm6/
* 创建repodata： createrepo **.**
* vim /etc/yum.repos.d/cloudera-manager.repo

　注意路径：http://manager/cloudera-repos/cm6/

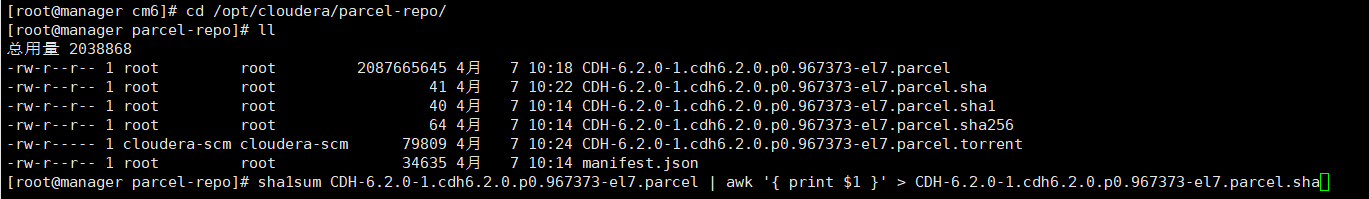
* yum clean all
* yum makecache

### 2.1.15.导入GPG key（如果没有这步操作，很可能cloudera服务安装失败）manager节点

* rpm --import http://manager/cloudera-repos/cm6/RPM-GPG-KEY-cloudera

### 2.1.16. 安装 Cloudera Manager（manager节点）

* sudo yum install cloudera-manager-daemons cloudera-manager-agent cloudera-manager-server
* 安装完CM后/opt/ 下会出现cloudera目录
* mv /opt/parcels/\* /opt/cloudera/parcel-repo # 将parcel包移动到指定位置
* 在/opt/cloudera/parcel-repo执行以下命令：
* sha1sum CDH-6.2.0-1.cdh6.2.0.p0.967373-el7.parcel | awk ‘{ print $1 }’ > CDH-6.2.0-1.cdh6.2.0.p0.967373-el7.parcel.sha

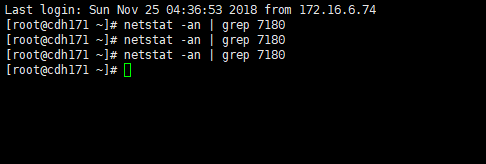


* 执行初始化脚本:
* /opt/cloudera/cm/schema/scm\_prepare\_database.sh mysql cmserver cmserveruser password
* 打开server服务:
* service cloudera-scm-server start
* 静候几分钟，打开http://manager:7180

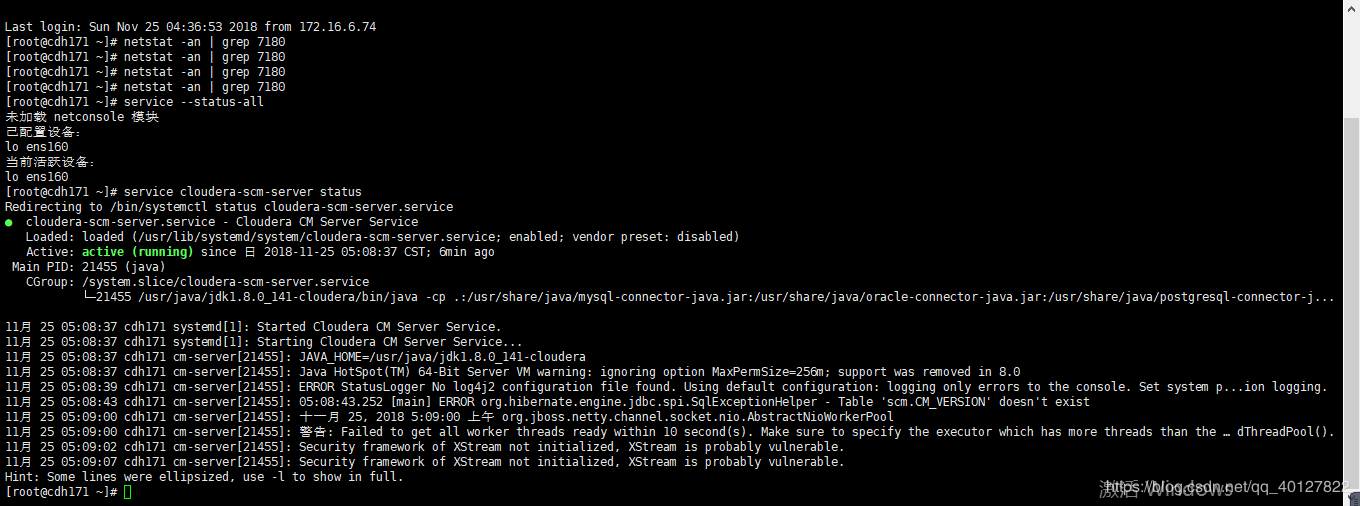
## 2.2 其他服务安装

### 2.2.1 登录cm WEB界面

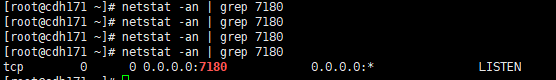
[http://主机ip:7180/cmf/login](http://xn--ip-wz2cm89g:7180/cmf/login) 访问CM  
用户名admin  
密码admin  
遇到问题：7180服务没有启动  
解决方法：



查看端口服务，未启动



cm服务启动显示正常。  
我在刚启动服务后7180没有启动，没找到什么原因，后来 晾了它一晚上，第二天一查端口，居然启动了，可能是cm服务要启动的东西太多，主机一时没启动。

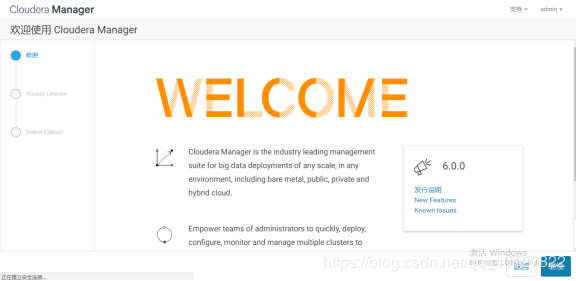


启动成功！

备注：  
linux查看端口：[https://www.cnblogs.com/Archmage/p/7570716.html](https://www.cnblogs.com/Archmage/p/7570716.html" \t "_blank)

### 2.2.2 具体安装步骤

#### WELCOME



#### Accept License

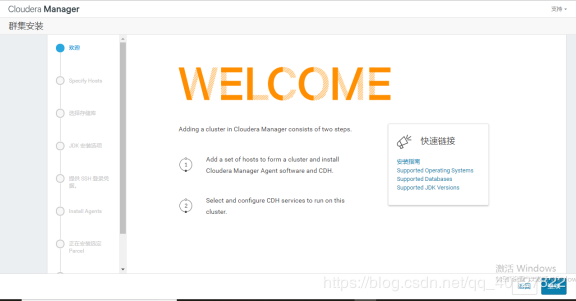


#### Select Edition

版本选择免费版，已经够用。



#### Welcome (Add Cluster - Installation)



#### Specify Hosts

主机是自己规划安装agent的主机

#### C:\Users\liuliyu\Desktop\1217276-20190604112025480-101815659.pngSelect Repository

小知识：使用自己配置的源地址

线上真实案例：不用自己的源地址，默认会去官方源，但是官方加收费墙了啊，这辈子都下载不下来，这是一个巨坑



#### JDK 安装选项



#### Enter Login Credentials

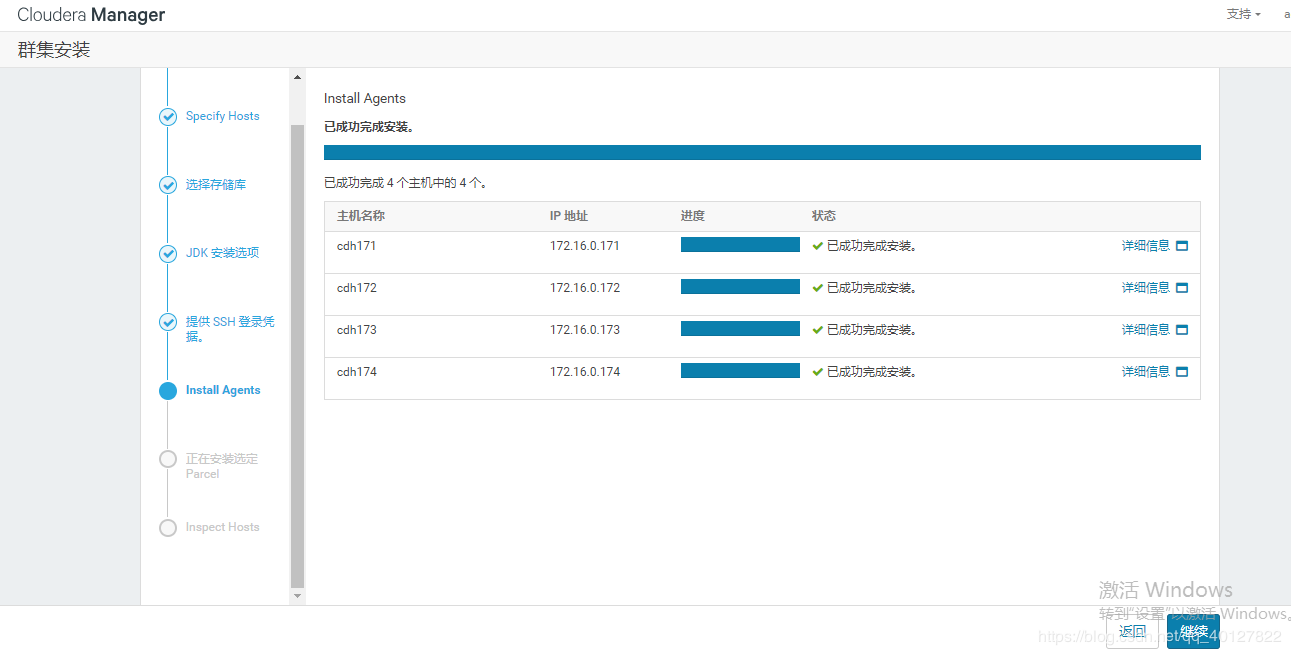


#### Install Agents

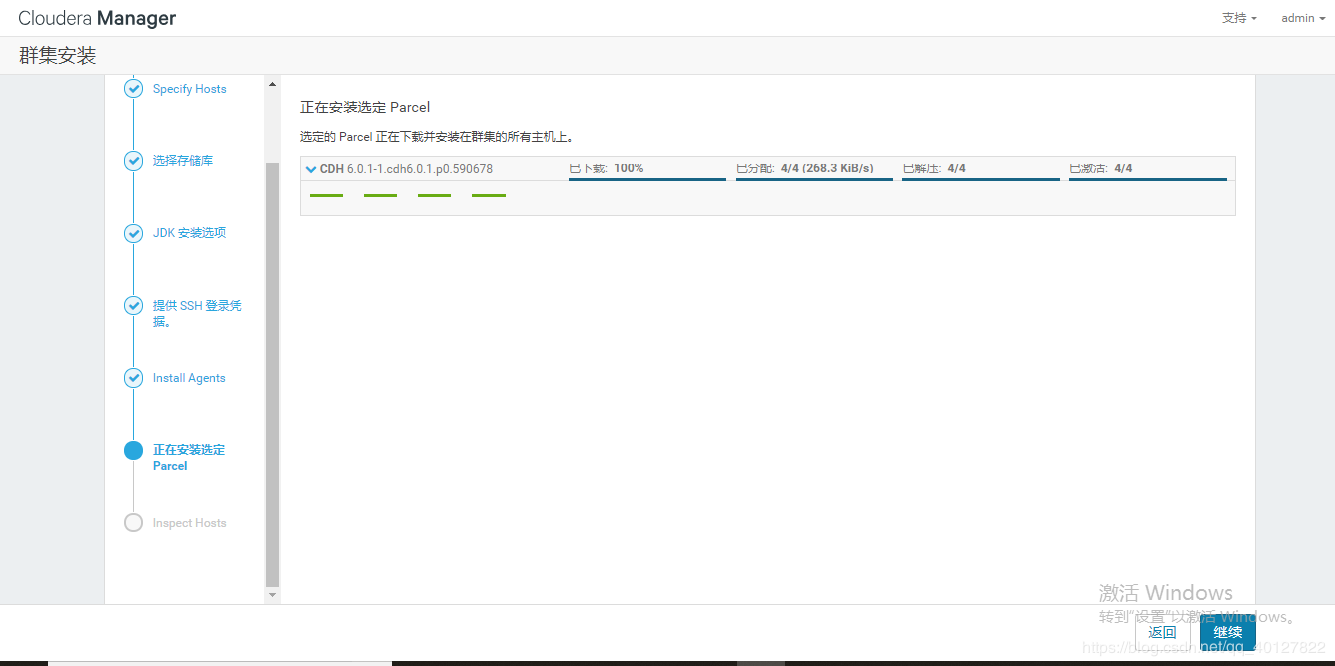
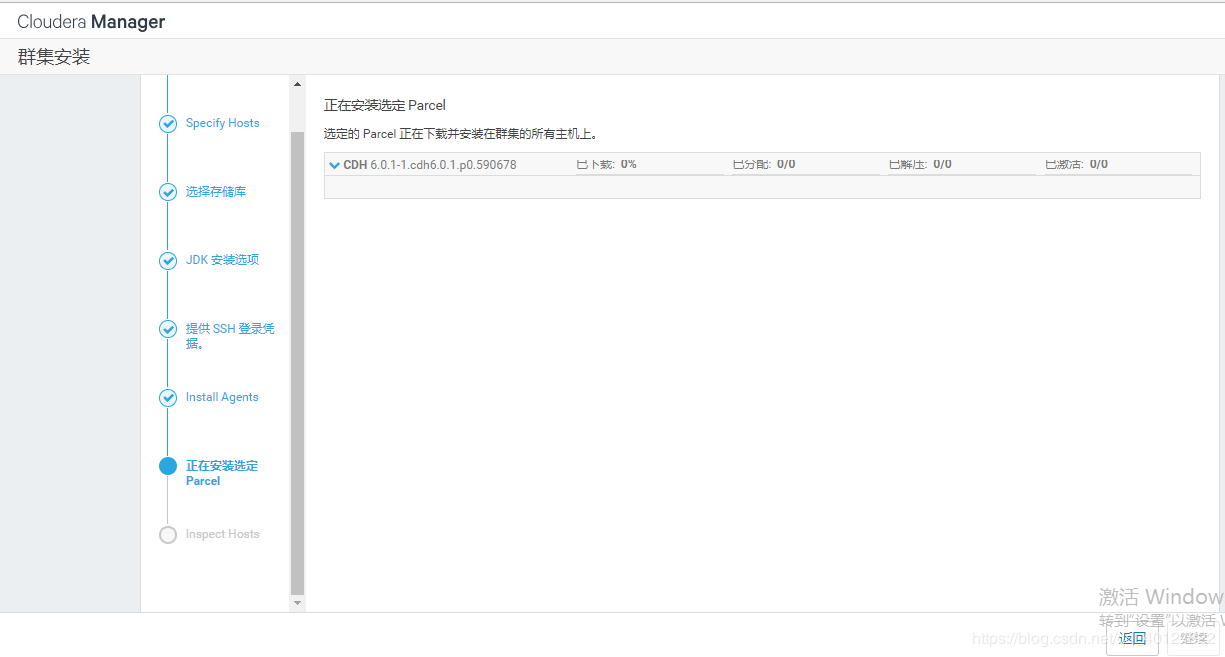
最到考验网速的时候了，该页面使用js进行刷新，千万别手动刷新，手动刷新的话安装列表中之前已经功成的会消失，未成功的显示，未成功即使安装成功了，cm会管理不到之前已经成功但刷新后未显示的主机，在安装集群时只能选择本次显示的（原因未知）。网速过慢的话安装会失败，一定要耐心等待，别做无关操作。



#### 失败重试直到成功，再次说明，耐心等待。



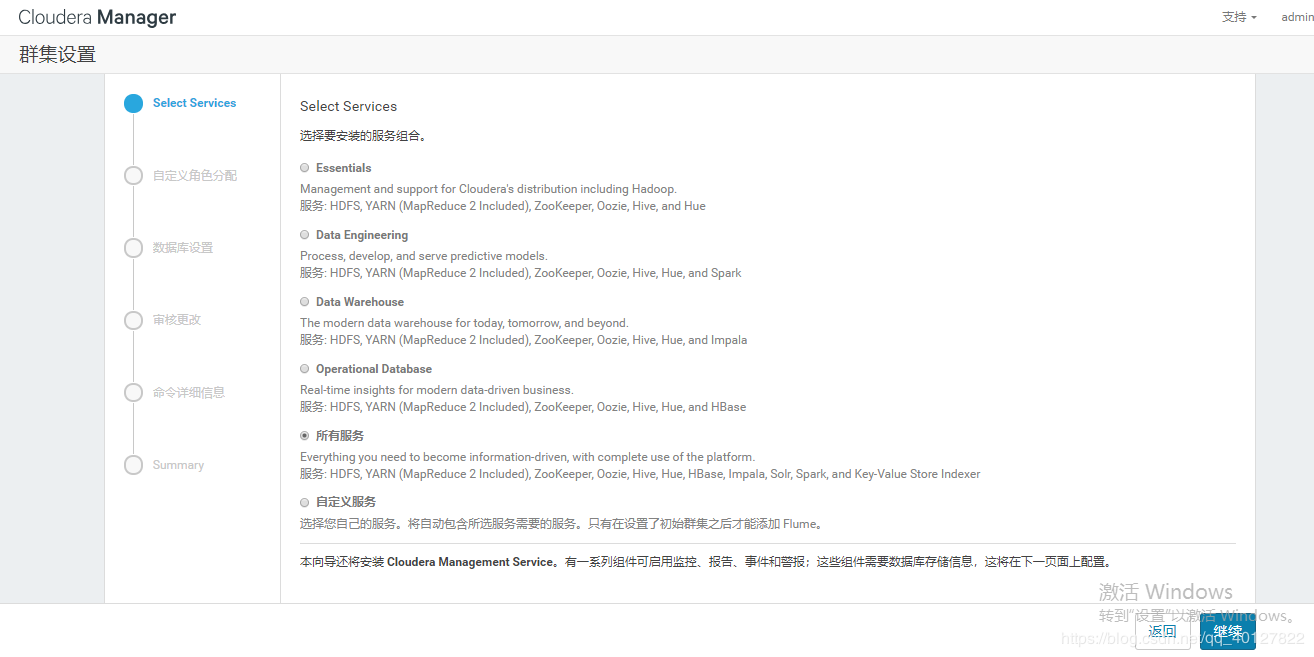
n次失败之后终于安装成功！

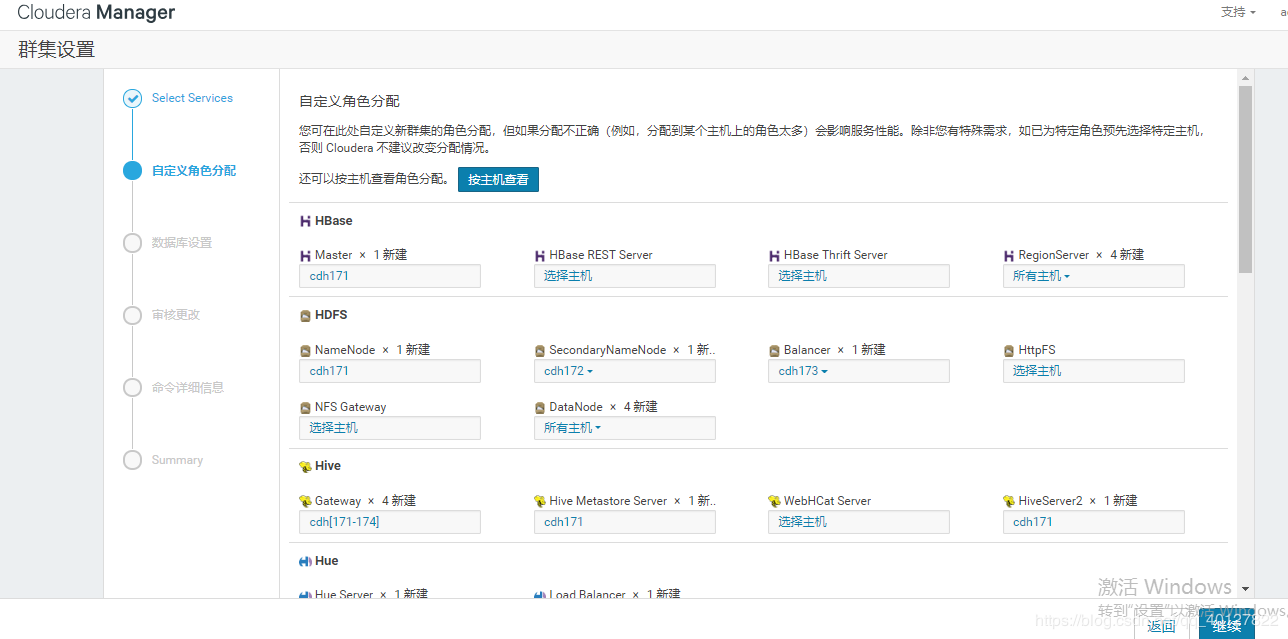


以上基础CDH环境搭建完成，开始集成自己的大数据组件，剩余大数据组件安装步骤跳过，每个安装按照实际使用场景规划制定

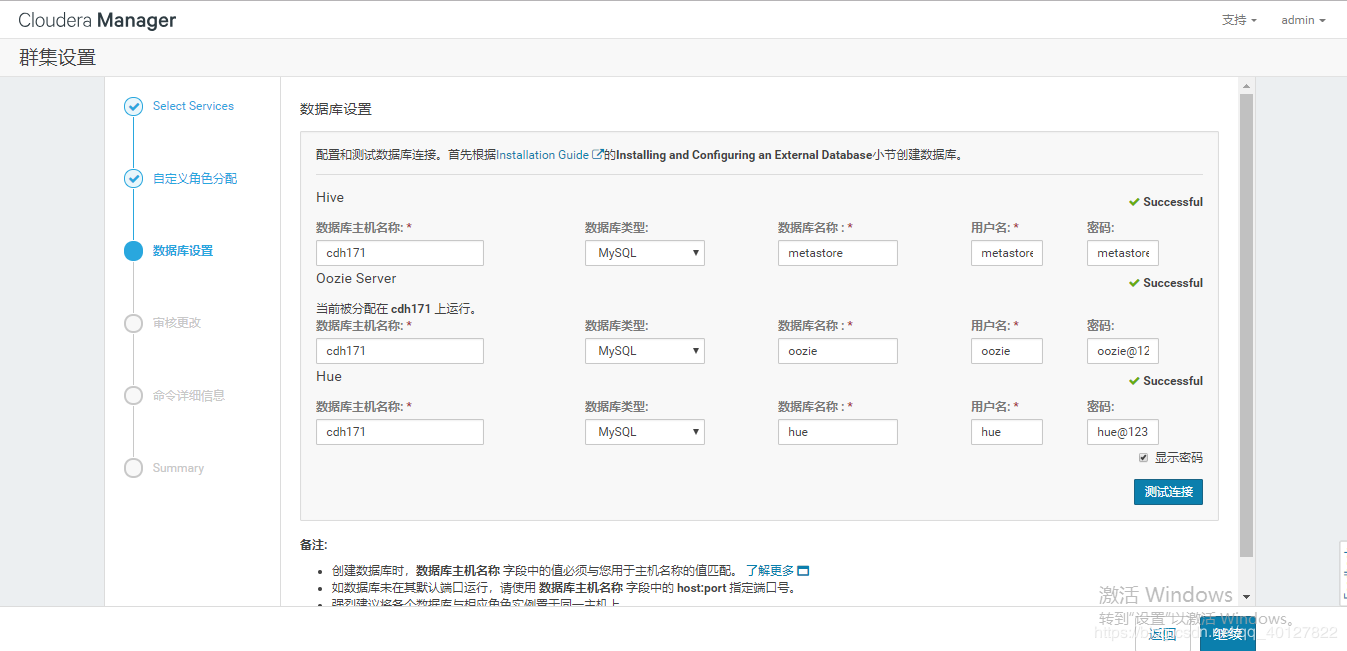
## 2.3 大数据组件安装

小知识：正式安装前拍个快照，可以避免很多很多问题

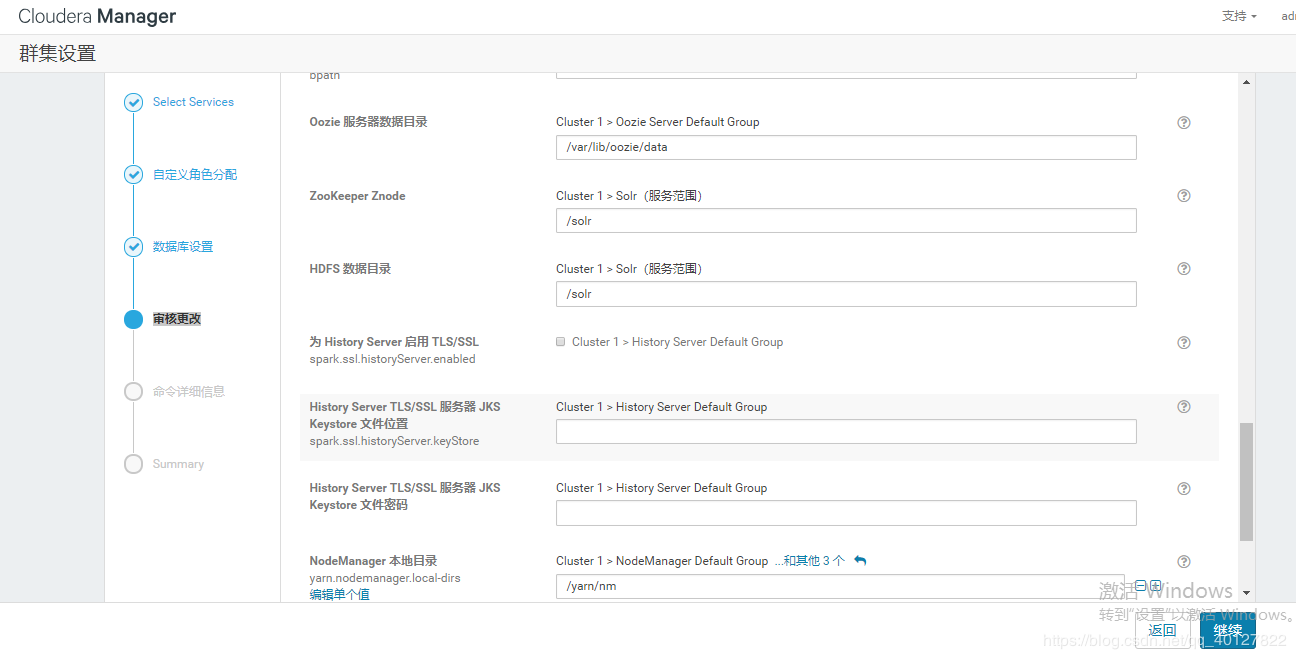




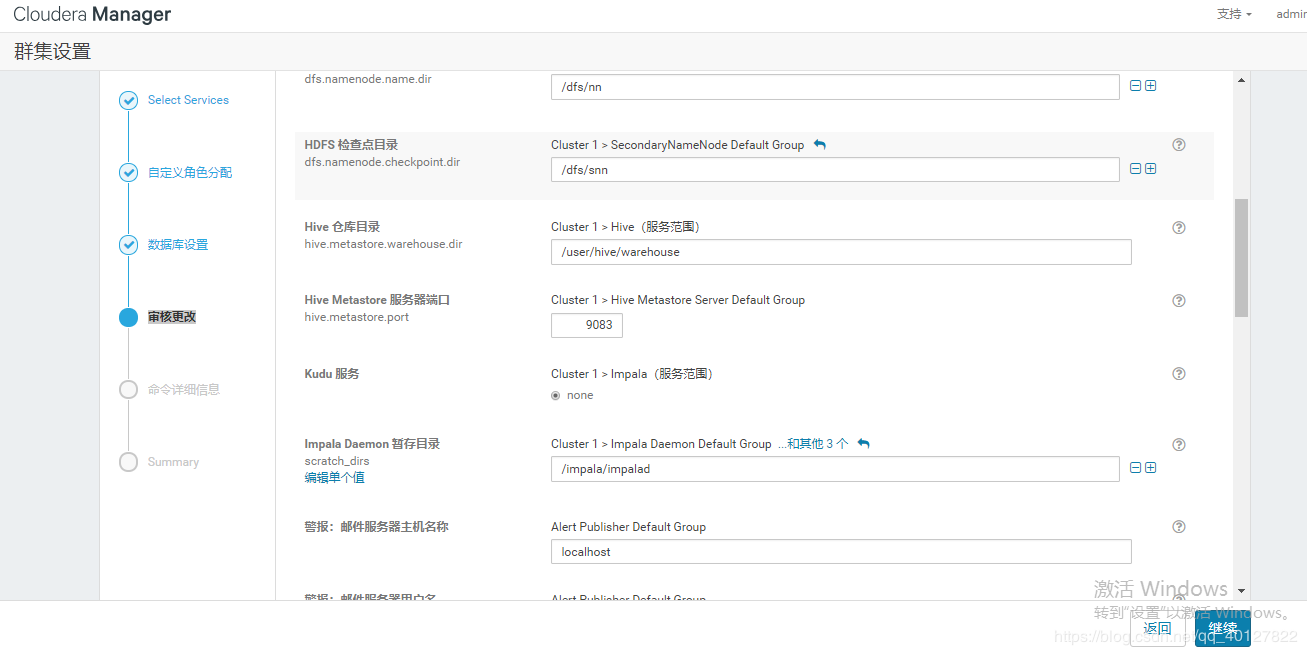
小知识：这个就是之前配置的mysql地址信息，这步也容易让人怀疑人生

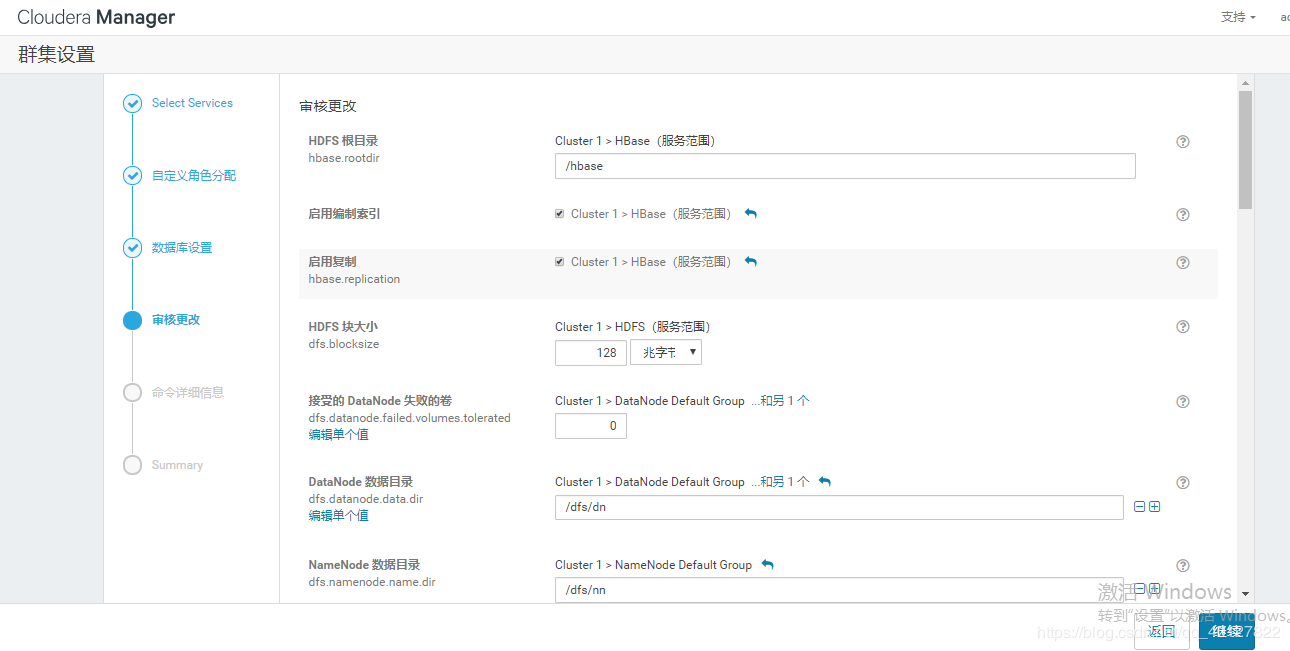


这些全部默认



这些全部默认





【参考资料】

https://blog.csdn.net/wolf\_333/article/details/89071203

http://www.cnblogs.com/mylovelulu/p/10384732.html

https://blog.csdn.net/qq\_40127822/article/details/84441869

https://www.cnblogs.com/raphael5200/p/5293960.html

https://www.waitig.com/%E5%AF%B9cloudera-hadoop%E5%A4%9A%E4%B8%80%E4%BA%9B%E4%BA%86%E8%A7%A3.html cdh切换日志目录

https://www.520mwx.com/view/46525 cdh服务器磁盘划分

https://www.cnblogs.com/swordfall/p/10816797.html