

CDH大数据平台

安装测试文档



2017-12-21

华融科技

目录

[1. 环境准备 2](#_Toc500920743)

[1.1. 网络配置(所有节点) 2](#_Toc500920744)

[1.2. 关闭防火墙： 2](#_Toc500920745)

[1.3. 关闭SELINUX 2](#_Toc500920746)

[1.4. yum源配置 2](#_Toc500920747)

[2. JDK安装（已有就无需安装） 5](#_Toc500920748)

[3. 修改设置参数 6](#_Toc500920749)

[4. 修改机器名和hosts文件 7](#_Toc500920750)

[4.1. 修改机器名 7](#_Toc500920751)

[4.2. 启动ntp时间同步服务 7](#_Toc500920752)

[5. 安装和配置Mysql【只有主节点需要安装】 7](#_Toc500920753)

[5.1. 检查是否已经安装mysql 7](#_Toc500920754)

[5.2. 安装Mysql 7](#_Toc500920755)

[5.3. 配置mysql数据库[下面的语句全部执行了] 8](#_Toc500920756)

[6. 配置SSH无密码登录 9](#_Toc500920757)

[重启 9](#_Toc500920758)

[7. 正式开工 9](#_Toc500920759)

[7.1. 完全离线安装 9](#_Toc500920760)

[7.2. 以前5.5.2的一次失败重来经历 15](#_Toc500920761)

[8. 半离线安装：RPM方式安装[要网络] 15](#_Toc500920762)

[8.1. 当前主机安装rpm 16](#_Toc500920763)

[8.2. 其他节点安装rpm 16](#_Toc500920764)

[8.3. 配置文件修改 16](#_Toc500920765)

[8.4. Mysql驱动和数据库初始化【主机上操作】 17](#_Toc500920766)

[8.5. Parcel包 17](#_Toc500920767)

[8.6. 正在安装选定 Parcel ：在这个界面中需要等待一定的时间。 18](#_Toc500920768)

[8.7. CDH5的安装配置 19](#_Toc500920769)

[9. 重启 19](#_Toc500920770)

[9.1. 主机 19](#_Toc500920771)

[9.2. 其他节点 19](#_Toc500920772)

# 环境准备

以下操作均需要(root)管理员权限

平台安装在默认文件夹为系统/opt/目录下

机器最好分配100G左右系统盘存储空间 16G以上内存

设备信息：

四台虚拟机

服务器信息：Red Hat Enterprise Linux Server 6.5

内核：Linux 2.6.32-431.el6.x86\_64 内存：16G

CPU：Intel Xeon E5-4640 v2 @ 2.20GHz

## 网络配置(所有节点)

修改主机名：vi /etc/sysconfig/network

NETWORKING=yes

HOSTNAME=dev-bigdata1

NETWORKING\_IPV6=no

PEERNTP=no

通过service network restart重启网络服务（通常需要重启机器才管用）

华融集团设备 不要随便重启 会导致网络连接不通

修改ip与主机名的对应关系：vi /etc/hosts（所有节点操作）

10.1.8.193 dev-bigdata1

10.1.8.194 dev-bigdata2

10.1.8.195 dev-bigdata3

10.1.8.196 dev-bigdata4

## 关闭防火墙

service iptables stop （临时关闭）

chkconfig iptables off（重启后生效）

## 关闭SELINUX

setenforce 0 （临时生效）

修改 /etc/selinux/config 下的 SELINUX=disabled （重启后永久生效）

查看selinux状态：/usr/sbin/sestatus -v

注意：每台机器都需要进行相同操作

## yum源配置

centos6.5-Yum服务器搭建：

1.服务端挂载iso系统镜像

确保系统挂载iso安装镜像文件

 [root@localhost ~]# mount /dev/cdrom  /mnt //挂载到mnt下

mount: block device /dev/sr0 is write-protected, mounting read-only

 (/dev/cdrom如果没有就用cdrom1)

2.服务端安装ftp

[root@localhost ~]# cd /mnt/

[root@localhost mnt]# cd Packages/

[root@localhost Packages]# ll vsftp\* //检查rpm是否存在

[root@localhost Packages]# rpm -ivh vsftpd-2.2.2-11.el6\_4.1.x86\_64.rpm

3.服务端ftp下建立rpm文件存放路径(该路径后面用于yum.repo文件配置

例如设置路径为：/var/ftp/pub/centos6.5那么下面设置repo文件的路径应用/pub/centos6.5/)

[root@localhost Packages]# ll /var/ftp/pub/

total 0

[root@localhost Packages]# cd /var/ftp/pub/

[root@localhost pub]# mkdir centos6.5

root@localhost pub]# cd centos6.5/

[root@localhost centos6.5]# pwd

/var/ftp/pub/centos6.5

4.服务端拷贝RPM文件

[root@localhost centos6.5]# cp -rav  /mnt/\* ./

5.服务端启动ftp、ftp服务

[root@localhost centos6.5]# chkconfig vssftpd onf

chkconfig version 1.3.49.3 - Copyright (C) 1997-2000 Red Hat, Inc.

This may be freely redistributed under the terms of the GNU Public License.

usage:   chkconfig [--list] [--type <type>] [name]

         chkconfig --add <name>

         chkconfig --del <name>

         chkconfig --override <name>

         chkconfig [--level <levels>] [--type <type>] <name> <on|off|reset|resetpriorities>

[root@localhost centos6.5]# chkconfig vsftpd on

[root@localhost centos6.5]# chkconfig vs

[root@localhost centos6.5]# chkconfig --list  | grep ftp

vsftpd          0:off 1:off 2:on 3:on 4:on 5:on 6:off

[root@localhost centos6.5]# service  vsftpd status

vsftpd is stopped

[root@localhost centos6.5]# service  vsftpd start

Starting vsftpd for vsftpd: [  OK  ]

[root@localhost centos6.5]# service  vsftpd status

vsftpd (pid 2630) is running...

6.服务端机器-安装ftp客户端（为了测试yum是否可用）

[root@localhost centos6.5]# cd Packages/

[root@localhost Packages]# ll  \*ftp\*

[root@localhost Packages]# rpm  -ivh ftp-0.17-54.el6.x86\_64.rpm

7.服务端机器-测试ftp客户端

ftp 10.1.8.193

name：ftp

password：

Login successful连接成功即可

8.服务端修改机器配置

设置enforce为permissive：

[root@localhost Packages]# setenforce 0

[root@localhost Packages]# getenforce

Permissive

修改selinux：

[root@localhost Packages]# vi /etc/sysconfig/selinux

修改为：selinux=disabled

9.服务端关防火墙

chkconfig iptables off（重启后生效）

service iptables stop（即时）

service iptables status（检查状态)

10.服务端替换repo.d文件

路径如下：

[root@localhost yum.repos.d]# pwd

/etc/yum.repos.d

创建原有文件保存路径：

[root@localhost yum.repos.d]# mkdir bk

把当前.repo文件移动至新建的路径下：

[root@localhost yum.repos.d]# mv \*.repo bk

在/etc/yum.repos.d路径放入yum.repo，打开yum.repo 修改baseurl中的ip

例如：baseurl=ftp://10.1.8.193/pub/centos6.5/ （参考后面附录）

10.1.8.193为服务器机器地址，/pub/centos6.5/目录为服务器ftp下建立的rpm存放路径（参考第3步路径）

11.服务端测试当前yum

[root@localhost yum.repos.d]# yum list

[root@localhost yum.repos.d]# yum list | grep gcc

12.服务端前清理缓存

yum clean all

13.客户端使用yum

在[root@localhost yum.repos.d]# pwd

/etc/yum.repos.d

放入yum.repo

14.客户端关防火墙

chkconfig iptables off（重启后生效）

service iptables stop（即时）

15.客户端前清理缓存

yum clean all

16.客户端yum测试

yum list|grep gcc

有对应包列表即可

附录：

repo文件内容：

[base]

name=RedHat-$releasever -$basearch -Source

#修改为yum服务器对应路径

baseurl=ftp://10.1.8.193/pub/centos6.5/

gpgcheck=1

enabled=1

gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-redhat-release

# JDK安装（最好1.7.0\_79版本）

上传jdk-7u79-linux-x64.tar.gz

解压tar zvxf jdk-7u79-linux-x64.tar.gz

移动mv jdk1.7.0\_79 /usr/lib/

vi /etc/profile

export JAVA\_HOME=/usr/lib/jdk1.7.0\_79

export PATH=$JAVA\_HOME/bin:$PATH

export CLASSPATH=.:$JAVA\_HOME/lib/dt.jar:$JAVA\_HOME/lib/tools.jar

source /etc/profile

java -version

# 修改设置参数

## 3.1 优化虚拟内存

vim /etc/security/limits.conf

\* soft noproc 11000

\* hard noproc 11000

\* soft nofile 65535

\* hard nofile 65535

虚拟内存修改

1)检查虚拟内存需求率

cat /proc/sys/vm/swappiness

显示如下：

1

2)临时降低虚拟内存需求率

sysctl vm.swappiness=0

3)永久降低虚拟内存需求率

vim /etc/sysctl.conf

kernel.shmall = 4294967296

vm.swappiness = 0

并运行如下命令使生效

sysctl -p

## 3.2 配置权限永久生效

临时生效

echo 0 > /proc/sys/vm/swappiness

echo never > /sys/kernel/mm/transparent\_hugepage/defrag

永久初始化

vim /etc/rc.local

echo 0 > /proc/sys/vm/swappiness

echo never > /sys/kernel/mm/transparent\_hugepage/defrag

配置权限

vim /etc/sudoers

在Defaults requiretty下面增加一行

cloudera-scm ALL=(ALL) NOPASSWD:ALL //cloudera-scm用户后面会创建

:wq!

保存退出

# 修改机器名和hosts文件

## 修改机器名

vi /etc/sysconfig/network

vi /etc/hosts

修改后需要重启

## 启动ntp时间同步服务

Node02和node03进行下面的修改：

vi /etc/ntp.conf注释一行，添加一行；以master为参照服务器，下面的144

#restrict default kod nomodify notrap nopeer noquery

restrict default nomodify notrap server 192.168.64.144 preferincludefile /etc/ntp/crypto/pwkeys /etc/ntp/keys

service ntpd start

chkconfig ntpd on

ntpdate -u ip

用date检查发现时间基本同步，认为成功

# 安装MySql(只需主节点需要安装)

mount -o loop /dev/sr0 /mnt/cdrom

## 检查是否已经安装mysql

rpm -qa | grep -i mysql 检查mysql是否安装

rpm -e --nodeps mysql 彻底卸载mysql包括依赖关系

rpm -qa | grep mysql 查看当前安装的mysql包

rpm -e --nodeps mysql-server-5.1.71-1.el6.x86\_64 删除安装包

## 安装Mysql

yum install -y mysql-server

yum install -y mysql-devel

## 配置mysql数据库

启动mysql数据库

service mysqld start

修改初始密码：

修改MySQL数据库root用户的初始密码,登陆命令为：

mysql

执行如下命令修改数据库root用户密码：

set password=password('root');

刷新权限表：

flush privileges;

exit;

将mysql设置成开机启动，测试使用新密码进行登陆。

chkconfig mysqld on

mysql -u root -p

本次安装需要创建如下数据库 --hive数据库

create database amon DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

create database hive DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

create database hue DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

create database nav DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

create database nava DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

create database navm DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

create database oozie DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

create database rman DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

create database reports DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

create database sentry DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

create database metastore DEFAULT CHARSET utf8 COLLATE utf8\_general\_ci;

mysql -u root -p

insert into mysql.user(Host,User,Password) values('dev-bigdata1','cdh',password('cdh'));

insert into mysql.user(Host,User,Password) values('localhost','cdh',password('cdh'));

insert into mysql.user(Host,User,Password) values('127.0.0.1','cdh',password('cdh'));

flush privileges;

exit

mysql -u root -p

GRANT ALL PRIVILEGES ON \*.\* TO cdh@localhost IDENTIFIED BY 'cdh' WITH GRANT OPTION;

GRANT ALL PRIVILEGES ON \*.\* TO cdh@dev-bigdata1 IDENTIFIED BY 'cdh' WITH GRANT OPTION;

GRANT ALL PRIVILEGES ON \*.\* TO cdh@127.0.0.1 IDENTIFIED BY 'cdh' WITH GRANT OPTION; //这里注意写的时候要和界面上的主机名一致，否则登录不了

//刷新系统权限表

flush privileges;

exit

mysql -h dev-bigdata1 -u cdh -p这样就可以用cdh登录mysql

use hive

show tables;

# 配置SSH无密码登录

SSH配置无密钥登录：

在各个节点上分别执行：

#cd /root/.ssh/

#ssh-keygen -t rsa

主节点执行：

#cat id\_rsa.pub >> authorized\_keys

将所有从节点的id\_rsa.pub文件内容都添加到主节点的authorized\_keys中

再将生成的包含所有节点的密钥从主节点scp到各个从节点：

#scp /root/.ssh/authorized\_keys dev-bigdata2:/root/.ssh/

#scp /root/.ssh/authorized\_keys dev-bigdata3:/root/.ssh/

#scp /root/.ssh/authorized\_keys dev-bigdata4:/root/.ssh/

其他机器同理，但是主机名适当调整一下，别乱了就好了

重启服务：

mount -o loop /dev/sr0 /mnt/cdrom

service mysqld start

# 正式开工

## 完全离线安装

上传文件/opt

CDH-5.9.0-1.cdh5.9.0.p0.7-el6.parcel

CDH-5.9.0-1.cdh5.9.0.p0.7-el6.parcel.sha

cloudera-manager-el6-cm5.9.0\_x86\_64.tar.gz

manifest.json

mysql-connector-java-5.1.34-bin.jar

jdk-7u79-linux-x64.tar.gz

相关安装文件在【D:\工作笔记\学习积累\Hadoop\cdh5.9.0】

cloudera manager的目录默认位置在/opt下，

1、解压：tar xzvf cloudera-manager\*.tar.gz 将解压后的cm-5.9.0和cloudera目录放到/opt目录下【我把文件放在opt下解压，所以直接就在这个目录】。

2、mysql-connector-java-5.1.30-bin.jar，放到/opt/cm-5.9.0/share/cmf/lib中。

cp /opt/mysql-connector-java-5.1.34-bin.jar /opt/cloudera/parcels/CDH-5.9.0-1.cdh5.9.0.p0.23/lib/hive/lib (稍后执行)

cp /opt/mysql-connector-java-5.1.34-bin.jar /usr/share/java/

mv mysql-connector-java-5.1.34-bin.jar mysql-connector-java.jar

在主节点初始化CM5的数据库：

3、cd /opt/cm-5.9.0/share/cmf/schema

./scm\_prepare\_database.sh mysql cm -hlocalhost -uroot -proot --scm-host localhost scm scm scm（可以先执行有错误再赋权限GRANT ALL PRIVILEGES ON \*.\* TO scm @dev-bigdata1 IDENTIFIED BY ' scm ' WITH GRANT OPTION;

GRANT ALL PRIVILEGES ON \*.\* TO scm @127.0.0.1 IDENTIFIED BY 'cdh' WITH GRANT OPTION;

）

此步可以完全手动安装配置

insert into mysql.user(Host,User,Password) values('dev-bigdata1',' scm ',password('cdh'));

insert into mysql.user(Host,User,Password) values('127.0.0.1',' scm ',password('cdh'));

GRANT ALL PRIVILEGES ON \*.\* TO scm @’dev-bigdata1’ IDENTIFIED BY ' scm ' WITH GRANT OPTION;

GRANT ALL PRIVILEGES ON \*.\* TO scm @127.0.0.1 IDENTIFIED BY ' scm ' WITH GRANT OPTION;

GRANT ALL PRIVILEGES ON \*.\* TO scm @localhost IDENTIFIED BY ' scm ' WITH GRANT OPTION;

vim /opt/cm-5.9.0/etc/cloudera-scm-agent/config.ini

修改其中的server\_host为主节点的主机名。

server\_host=dev-bigdata1原值为localhost

scp -r cm-5.9.0/ root@dev-bigdata2:/opt/

scp -r cm-5.9.0/ root@dev-bigdata3:/opt/

scp -r cm-5.9.0/ root@dev-bigdata4:/opt/

在所有节点创建cloudera-scm用户

useradd --system --home=/opt/cm-5.9.0/run/cloudera-scm-server/ --no-create-home --shell=/bin/false --comment "Cloudera SCM User" cloudera-scm

将CHD5相关的Parcel包放到主节点的/opt/cloudera/parcel-repo目录中（parcel-repo需要手动创建）。

相关的文件如下：

·CDH-5.9.0-1.cdh5.9.0.p0.7-el6.parcel

·CDH-5.9.0-1.cdh5.9.0.p0.7-el6.parcel.sha

· manifest.json

在Master节点上通过/opt/cm-5.9.0/etc/init.d/cloudera-scm-server start 启动服务端。

在所有节点上通过/opt/cm-5.9.0/etc/init.d/cloudera-scm-agent start 启动Agent服务。

我们启动的其实是个service脚本，需要停止服务将以上的start参数改为stop就可以了，重启是restart。

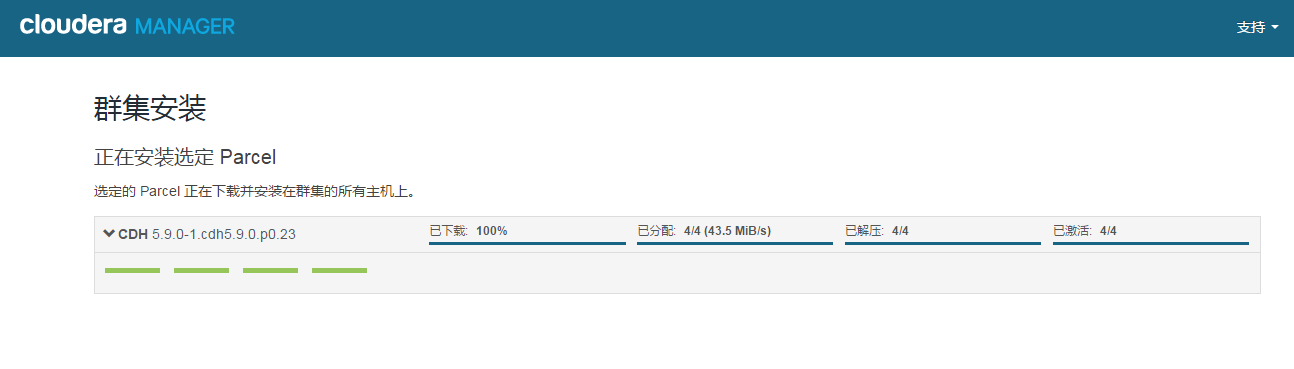
Cloudera Manager Server和Agent都启动以后，就可以进行CDH5的安装配置了。<http://10.1.8.193:7180>这个需要等待一会儿的，没有那么快启动

admin/admin





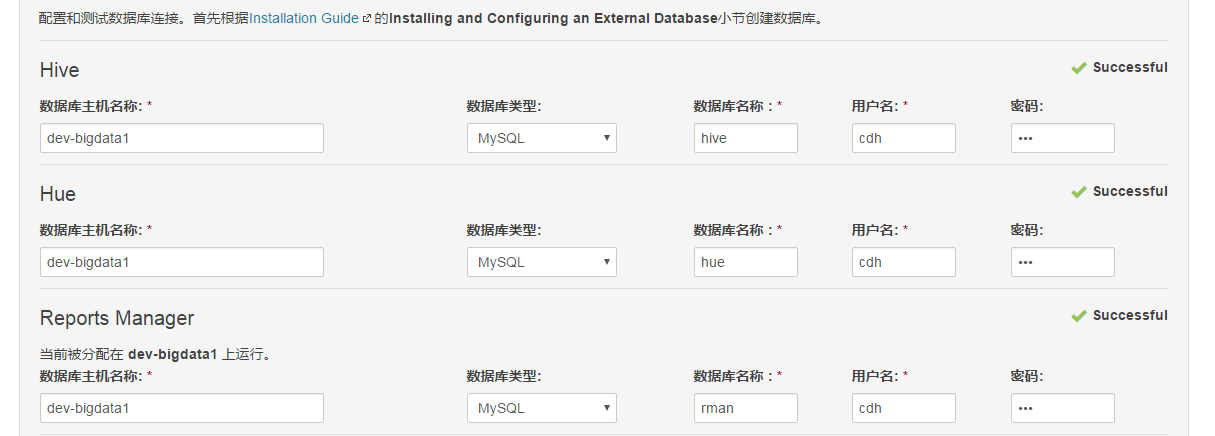
前一个界面千万要选对版本号啊，选对了的话，下载那个地方很快就要100%；否则是从网络上下载，整个过程很慢很慢的：





vi /etc/sysctl.conf

vm.swappiness=10







注：在当前管理的主机里面可以选择；

在安装过程中出现hive和oozie找不到mysql驱动程序，用下面的方式去解决。查看日志即可自行解决问题。需要注意的是stdout里面的命令是需要关注的，oozie就是检查该命令后发现如何解决的问题

上传驱动包到

/opt/cloudera/parcels/CDH-5.9.0-1.cdh5.9.0.p0.23/lib/hive/lib/mysql-connector-java-5.1.34-bin.jar

cd /opt/cloudera/parcels/CDH-5.9.0-1.cdh5.9.0.p0.23/lib/hive/lib

scp mysql-connector-java-5.1.34-bin.jar /opt/cloudera/parcels/CDH-5.9.0-1.cdh5.9.0.p0.23/lib/hive/lib/mysql-connector-java-5.1.34-bin.jar

mkdir -p /usr/share/java

cd /opt/cloudera/parcels/CDH-5.9.0-1.cdh5.9.0.p0.23/lib/oozie/lib

上传mysql-connector-java-5.1.34-bin.jar

cp mysql-connector-java-5.1.34-bin.jar /usr/share/java/mysql-connector-java.jar

cp mysql-connector-java-5.1.30-bin.jar

到Master的这个目录/opt/cloudera/parcels/CDH-5.5.2-1.cdh5.5.2.p0.4/lib/hive/lib

Parcels界面要是等待很久就说明是从网站上下载parcels文件，这个时候把parcels放入指定目录，然后重启主机的server和agent，其他节点只需要重启agent

/opt/cm-5.9.0/etc/init.d/cloudera-scm-server restart

/opt/cm-5.9.0/etc/init.d/cloudera-scm-agent restart

重启server后需要等待好一会儿才能登录网页，否则一直是无法访问

前面创建的数据库，在安装界面需要使用的，全部数据库的用户名和密码都用cdh（因为上面只创建了cdh用户）

# 重启集群

## 主节机

mount -o loop /dev/sr0 /mnt/cdrom

service mysqld start

/etc/init.d/cloudera-scm-server start

/etc/init.d/cloudera-scm-agent start

## 从节点

mount -o loop /dev/sr0 /mnt/cdrom

/etc/init.d/cloudera-scm-agent start

/etc/init.d/cloudera-scm-server stop

/etc/init.d/cloudera-scm-agent stop

/etc/init.d/cloudera-scm-agent restart

# 安装完成

浏览器输入<http://10.1.8.193:7180>

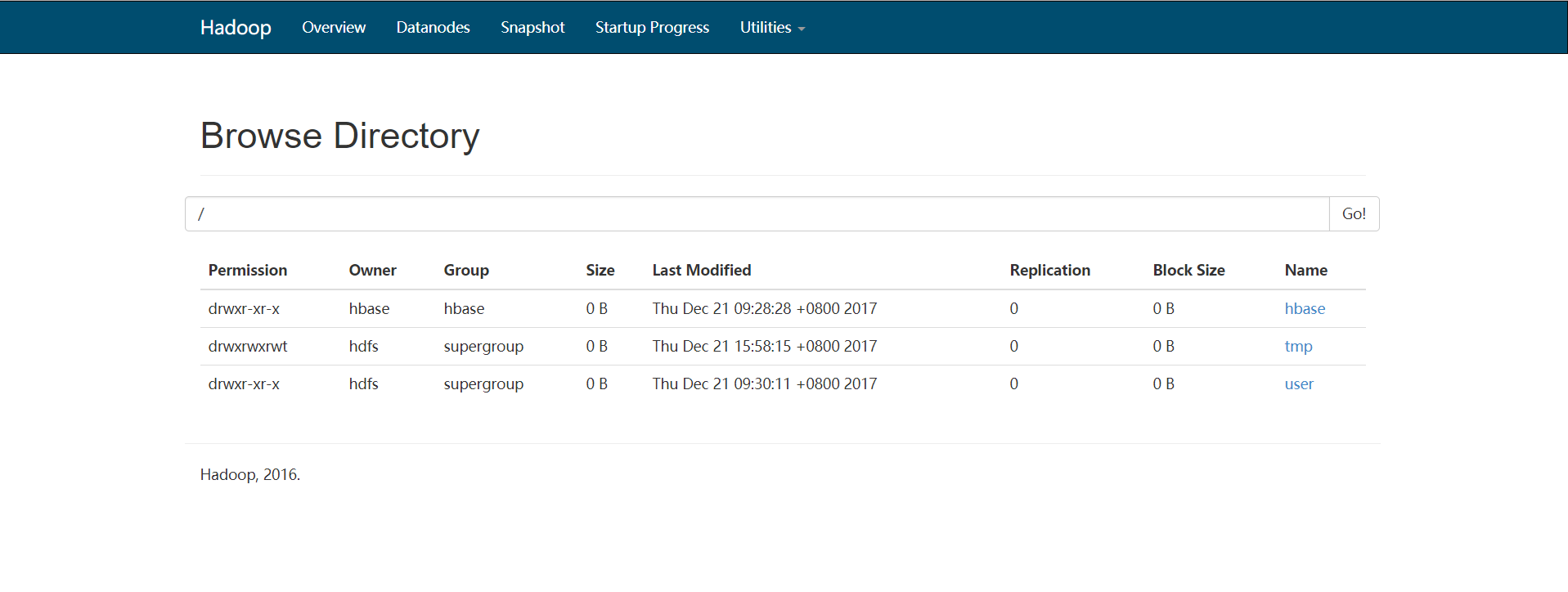
使用用户名：admin和密码：admin登录cloudera manager （请勿随意操作）

就会进入如下cloudera manager操作界面：



浏览器输入<http://10.1.8.193:50070>

登录Hadoop Web UI界面如下图 查看分布式文件存储系统



# 数据引入

# 测试