

CDH 6.3.1 安装

一.官网下载地址

1.官网

https://docs.cloudera.com/documentation/enterprise/6/release-notes/topics/rg_cm_6_version_download.html

CDH6.3.3 于 2020 年 2 月 4 日发布, 从该版本开始, Cloudera 将不再提供 CDH 免费版, 用户只能通过试用或者购买正式许可证才能够下载 CDH 安装包及使用

因此选择 6.3.1 版本进行安装

2.安装包概览

```
CDH-6.3.1-1.cdh6.3.1.p0.1470567-el7.parcel
CDH-6.3.1-1.cdh6.3.1.p0.1470567-el7.parcel.sha1
cm6.3.1-redhat7.tar.gz
jdk-8u45-linux-x64.gz
manifest.json
mysql-5.7.11-linux-glibc2.5-x86_64.tar.gz
mysql-connector-java-5.1.47.jar
```

以上 jdk 及 mysql 安装包可为其他版本, 具体要求参考 cdh6 环境准备文档

二.环境配置:root 用户操作

1.基本配置

1.1 系统确认

```
[root@localhost ~]# cat /etc/redhat-release
CentOS Linux release 7.5.1804 (Core)
```

1.2 修改时区

```
[root@localhost ~]# ll /etc/localtime
lrwxrwxrwx. 1 root root 35 Jul  9 2019 /etc/localtime
-> ../usr/share/zoneinfo/Asia/Shanghai
```

是上海时区不用改,为其他时区改为上海时区能避免很多后续麻烦

```
[root@localhost ~]# ln -sf /usr/share/zoneinfo/Asia/Shanghai /etc/localtime
```

或者

```
timedatectl set-timezone Asia/Shanghai
```

1.3 修改主机名

```
hostnamectl set-hostname cdh1
```

```
hostnamectl set-hostname cdh2
```

```
hostnamectl set-hostname cdh3
```

也可以

```
vi /etc/hostname
```

```
cdh1
```

1.4 修改/etc/hosts:注修改此项不要更改行首为主机名称,否则识别不出名称和 ip

```
[root@localhost ~]# vi /etc/hosts
```

```
[root@localhost ~]# cat /etc/hosts
```

```
127.0.0.1    localhost localhost.localdomain localhost4 localhost4.localdomain4
::1         localhost localhost.localdomain localhost6 localhost6.localdomain6
```

```
172.16.8.134 cdh1
```

```
172.16.8.137 cdh2
```

```
172.16.8.156 cdh3
```

1.5 修改 /etc/sysconfig/network

```
[root@localhost ~]# vi /etc/sysconfig/network
```

```
[root@localhost ~]# cat /etc/sysconfig/network
```

```
# Created by anaconda
```

```
NETWORKING=yes
```

```
HOSTNAME=cdh1
```

```
[root@localhost ~]#
```

1.6 关闭防火墙 #centos 7 默认使用的是 firewall, 不是 iptables

```
systemctl stop firewalld.service
```

```
systemctl mask firewalld.service
```

#注销服务,不会开机启动

1.7 关闭 SELinux

```
vi /etc/selinux/config  
SELINUX=disabled
```

#设置 SELINUX=disabled

1.8 设置 swap 空间

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

1.9 关闭大页面压缩

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

2.配置 ntp 服务(需设置所有节点)

2.1 安装 ntp

```
yum install ntp -y
```

#安装

```
chkconfig ntpd on
```

#设置开机启动

如为内网环境且未配置 yum 源可下载 rpm 包使用 rpm 安装,需先安装依赖

```
rpm -ivh autogen-libopts-5.18-5.el7.x86_64.rpm
```

```
rpm -ivh openssl-libs-1.0.2k-16.el7.x86_64.rpm
```

```
rpm -ivh ntp-4.2.6p5-28.el7.centos.x86_64.rpm
```

2.2 主节点配置为时间服务器(CM 安装节点)

```
vi /etc/ntp.conf
```

#按以下内容修改

```
server 127.127.1.0
```

#无外网,配置为自身

```
fudge 127.127.1.0 stratum 8
```

#无外网,配置为自身无法生效就改为 Fudge

```
server ntp1.aliyun.com iburst
```

#有外网,阿里云时钟服务器

```
server ntp2.aliyun.com iburst
server ntp3.aliyun.com iburst
server ntp4.aliyun.com iburst
server ntp5.aliyun.com iburst
```

```
service ntpd start          #启动服务
ntpstat                     #查看状态,如下则成功
```

```
[root@master9101 ~]# ntpstat
synchronised to local net at stratum 6
  time correct to within 12 ms
  polling server every 64 s
[root@master9101 ~]#
```

```
vi /etc/ntp.conf            #设置其他节点的时间服务器为主节点
server cdh1 prefer
```

2.3 安装 ntpdate 手动同步时间工具

```
yum install ntpdate          #yum 安装或使用如下 rpm 离线安装
#rpm -ivh ntpdate-4.2.6p5-28.el7.centos.x86_64.rpm
```

3.SSH 免密登录(配置集群内节点间互相免密登录)

3.1 创建公钥(需在所有节点内创建)

```
cd /root/.ssh                #没有目录.ssh 的话直接 ssh 登录一次其他节点就有
了
ssh-keygen -t rsa             #直接回车
```

3.2 合并所有节点 id_rsa.pub 内容为 authorized_keys 文件(就是复制所有节点 id_rsa.pub 文件内容,粘贴到文件 authorized_keys 中)

```
[root@cdh-dev1 ~]# cd /root/.ssh/
[root@cdh-dev1 .ssh]# ll
total 16
-rw-r--r--. 1 root root 1185 Jun 11 09:44 authorized_keys
-rw-r--r--. 1 root root 1675 Jun 11 09:42 id_rsa
-rw-r--r--. 1 root root 395 Jun 11 09:42 id_rsa.pub
-rw-r--r--. 1 root root 1198 Jul 19 06:24 known_hosts
[root@cdh-dev1 .ssh]# cat id_rsa.pub
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCNpVd+sYqD0YYq0Plz784kgL156YtpoUg61cg90z/6Gr/ScjA2DKf74G/MuPWwGqvnXbA19U59wRkBu+jGhYDvCBq6QaN/3Uk4vT5oQ2RMmPi11StFLiylpANiv3mh5JA2J59PdcD07GzdIEEVANvXMch9X7njEbKnoVpdGkGEhjXjPS2wkgEtpijCWP2VFb08ab60qDD0HYtOnewhemGP4DsicpU+2M1mjJC3Zyl/p6npXKN38IxiZy8zac5hRsg3TgTxlG4AE1+sFxA0UiiEyc8y8T0y1nr0I0BpVL root@cdh-dev1
[root@cdh-dev1 .ssh]# █

[root@cdh-dev1 ~]# cat /root/.ssh/authorized_keys
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCNpVd+sYqD0YYq0Plz784kgL156YtpoUg61cg90z/6Gr/ScjA2DKf74G/MuPWwGqvnXbA19U59wRkBu+jGhYDvCBq6QaN/3Uk4vT5oQ2RMmPi11StFLiylpANiv3mh5JA2J59PdcD07GzdIEEVANvXMch9X7njEbKnoVpdGkGEhjXjPS2wkgEtpijCWP2VFb08ab60qDD0HYtOnewhemGP4DsicpU+2M1mjJC3Zyl/p6npXKN38IxiZy8zac5hRsg3TgTxlG4AE1+sFxA0UiiEyc8y8T0y1nr0I0BpVL root@cdh-dev1
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCNpVd+sYqD0YYq0Plz784kgL156YtpoUg61cg90z/6Gr/ScjA2DKf74G/MuPWwGqvnXbA19U59wRkBu+jGhYDvCBq6QaN/3Uk4vT5oQ2RMmPi11StFLiylpANiv3mh5JA2J59PdcD07GzdIEEVANvXMch9X7njEbKnoVpdGkGEhjXjPS2wkgEtpijCWP2VFb08ab60qDD0HYtOnewhemGP4DsicpU+2M1mjJC3Zyl/p6npXKN38IxiZy8zac5hRsg3TgTxlG4AE1+sFxA0UiiEyc8y8T0y1nr0I0BpVL root@cdh-dev2
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCNpVd+sYqD0YYq0Plz784kgL156YtpoUg61cg90z/6Gr/ScjA2DKf74G/MuPWwGqvnXbA19U59wRkBu+jGhYDvCBq6QaN/3Uk4vT5oQ2RMmPi11StFLiylpANiv3mh5JA2J59PdcD07GzdIEEVANvXMch9X7njEbKnoVpdGkGEhjXjPS2wkgEtpijCWP2VFb08ab60qDD0HYtOnewhemGP4DsicpU+2M1mjJC3Zyl/p6npXKN38IxiZy8zac5hRsg3TgTxlG4AE1+sFxA0UiiEyc8y8T0y1nr0I0BpVL root@cdh-dev3
[root@cdh-dev1 ~]# █
```

```
[root@cdh1 .ssh]# cat id_rsa.pub
ssh-rsa
AAAAB3NzaC1yc2EAAAADAQABAAQDg4+y8WwcKoVAungfThogHKG2eL2NY3CG2JI/03j2nKk7ZRP
x6RcNKUC+k1aTe1RqIeojnrocIgGmX7iXQWTsdBuoGDeYnaeRtEQo7dJit07IsJuQSwdOyMCXKMfHkT
WvNPo7IF2i/nUoSvVQXCd7LXzUnFjOKiHeDQbHyb0/KWYgYa7EhDSsy6Bx/vT3iha1yaRaJ63aWZik+
V9rDoDjQKUKYa6uL1Xyab0FrZebdtJtDofpNGDMteZUyUo995NFaWrVwGhXiwfnAnvbFCvPiT34D70d
fa2jKAsxFU0/v5TvD0w0Rqr8VWXJG0hXEz1FglavDuC2qGYDk74hK1tL root@cdh1
[root@cdh1 .ssh]# vi authorized_keys
```

3.3 更改权限并分发到全部节点

```
chmod 600 authorized_keys
scp authorized_keys root@cdh-dev2:~/.ssh
scp authorized_keys root@cdh-dev3:~/.ssh
```

4.Java 环境(需修改所有节点:具体jdk 支持版本请参考 cdh6 环境准备文档)

4.1 查看 java 版本,满足需求即可,不满足需重新安装

```
[root@cdh1 parcel-repo]# java -version
openjdk version "1.8.0_212"
OpenJDK Runtime Environment (build 1.8.0_212-b04)
OpenJDK 64-Bit Server VM (build 25.212-b04, mixed mode)
```

4.2 若为系统自带 jdk,卸载自带 jdk

```
rpm -qa | grep -i java          #查看安装信息
rpm -e java* --nodeps          #根据查找到的软件包信息卸载
```

4.3 安装 jdk,配置环境变量

```
mkdir /usr/java
tar xzvf jdk-8u202-linux-x64.tar.gz -C /usr/java
vi /etc/profile                  #文件末尾添加下面三行

export JAVA_HOME=/usr/java/jdk1.8.0_202
export PATH=$JAVA_HOME/bin:$PATH
export CLASSPATH=.:$JAVA_HOME/lib/dt.jar:$JAVA_HOME/lib/tools.jar

source /etc/profile             #使文件立即生效
java -version                   #查看版本是否为安装版本
```

以上为安装 oracle jdk8 的步骤 ,下面介绍 安装 openjdk 的步骤

```
rpm -ivh libICE-1.0.9-9.el7.x86_64.rpm
rpm -ivh libSM-1.2.2-2.el7.x86_64.rpm
rpm -ivh libX11-1.6.7-2.el7.x86_64.rpm
rpm -ivh libXext-1.3.3-3.el7.x86_64.rpm
rpm -ivh libXi-1.7.9-1.el7.x86_64.rpm
rpm -ivh giflib-4.1.6-9.el7.x86_64.rpm
yum install -y copy-jdk-configs
yum install -y cups-libs
yum install -y jpackage-utils
yum install -y fontconfig
yum install -y gtk2
rpm -ivh libXtst-1.2.2-2.1.el7.x86_64.rpm
yum install -y lksctp-tools
yum install -y nss
yum install -y pcsc-lite-libs
yum install -y tzdata-java
rpm -ivh java-1.8.0-openjdk-headless-1.8.0.212.b04-0.el7_6.x86_64.rpm
yum install -y xorg-x11-fonts-Type1
yum install -y libgif.so.4
rpm -ivh java-1.8.0-openjdk-1.8.0.212.b04-0.el7_6.x86_64.rpm
```

以上为环境实测,部分环境可能无法 yum 安装,请下载相关 rpm 包安装依赖

```
rpm -ivh java-1.8.0-openjdk-1.8.0.212.b04-0.el7_6.x86_64.rpm
```

根据上述命令提示的缺失依赖进行下载安装,具体可参照一下网页

http://rpm.pbone.net/index.php3/stat/4/idpl/54995513/dir/scientific_1linux_7/com/java-1.8.0-openjdk-1.8.0.212.b04-0.el7_6.x86_64.rpm.html

5.安装 MySQL(主节点安装即可,用于存储 CDH 相关信息)

5.1 卸载 mariadb(centos7 默认安装 mariadb,不卸载会报错 mariadb-libs is obsoleted by)

```
rpm -qa | grep mariadb          #查看 mariadb 安装信息,未安装可跳过一步
rpm -e mariadb-libs* --nodeps  #根据查找到的软件包信息卸载
```

5.2 解压安装 MySQL(严格按照顺序安装)

```
tar zxvf mysql-5.7.24-rpm.tar.gz && cd mysql-5.7.24-rpm
rpm -ivh mysql-community-common-5.7.24-1.el7.x86_64.rpm --nodeps
rpm -ivh mysql-community-libs-5.7.24-1.el7.x86_64.rpm --nodeps
rpm -ivh mysql-community-client-5.7.24-1.el7.x86_64.rpm --nodeps
rpm -ivh mysql-community-server-5.7.24-1.el7.x86_64.rpm --nodeps
```

上一步不强制执行可能报错 缺失依赖 libaio.so.1 在安装完成起不来的情况下,可以不强制执行

```
wget
http://mirror.centos.org/centos/6/os/x86_64/Packages/libaio-0.3.107-10.el6.x86_64.rpm
rpm -ivh libaio-0.3.107-10.el6.x86_64.rpm
```

5.3MySQL 的基本设置

```
service mysqld start          #启动服务
cat /var/log/mysqld.log |grep password      #查看初始密码
mysql -uroot -p                #登录,密码为刚查看的初始密码
mysql> set password = password('Upchina#3#2#1');    #设置密码
                #授权用户 root 使用密码从任意主机连接到 mysql 服务器
mysql> grant all privileges on *.* to 'root'@'%' identified by 'Upchina#3#2#1' with
grant option;
```

```
mysql> flush privileges;
```

5.4 相关数据库的创建(用于存储 CDH 相关组件的信息,这一块参照之前的即可,需修改)

```
create database cmf DEFAULT CHARACTER SET utf8;
create database amon DEFAULT CHARACTER SET utf8;
grant all on cmf.* TO 'cmf'@'%' IDENTIFIED BY 'Upchina#3#2#1';
grant all on amon.* TO 'amon'@'%' IDENTIFIED BY 'Upchina#3#2#1';
flush privileges;

create database hive DEFAULT CHARSET utf8 COLLATE utf8_general_ci;
create database oozie DEFAULT CHARSET utf8 COLLATE utf8_general_ci;
create database hue DEFAULT CHARSET utf8 COLLATE utf8_general_ci;

[root@cdh1 mysql]# mkdir -p /usr/share/java/
重命名不能带版本号
cp mysql-connector-java-5.1.47.jar /usr/share/java/mysql-connector-java.jar
```

三.CDH 的安装

1.准备工作

1.1 所有节点创建目录及解压

```
mkdir /opt/cloudera-manager
tar -xzf cm6.3.1-redhat7.tar.gz -C /opt/cloudera-manager/
tar -xzf cm6.3.1-redhat7.tar.gz -C /opt/cloudera-manager/
```


这里有一点需要注意,不同于如 5.13.0 等低版本 cdh,解压出来的 cm 直接是 rpm 包而不是免安装版本的 cm

1.2 选择 cdh1 为 cm server, (可以尝试不下载依赖包直接部署 --nodeps --force

,但不建议,因为缺失很多依赖可能出问题,建议根据缺失依赖逐步安装)

```
cd /opt/cloudera-manager/cm6.3.1/RPMS/x86_64/  
rpm -ivh cloudera-manager-daemons-6.3.1-1466458.el7.x86_64.rpm  
rpm -ivh cloudera-manager-server-6.3.1-1466458.el7.x86_64.rpm  
rpm -ivh cloudera-manager-agent-6.3.1-1466458.el7.x86_64.rpm
```

1.3 选择 cdh2-3 为 cm agent, (可以尝试不下载依赖包直接部署 --nodeps --force

,但不建议,因为缺失很多依赖可能出问题,建议根据缺失依赖逐步安装)

```
cd /opt/cloudera-manager/cm6.3.1/RPMS/x86_64/  
rpm -ivh cloudera-manager-daemons-6.3.1-1466458.el7.x86_64.rpm  
rpm -ivh cloudera-manager-agent-6.3.1-1466458.el7.x86_64.rpm
```

以下为可能缺失的依赖的 yum 安装方式

```
yum install -y bind-utils  
yum install -y psmisc  
yum install -y cyrus-sasl-plain  
yum install -y cyrus-sasl-gssapi  
yum install -y portmap  
yum install -y lsb  
yum install -y httpd  
yum install -y mod_ssl  
yum install -y openssl-devel  
yum install -y python-psycopg2  
yum install -y MySQL-python
```

1.4 主节点执行以下建表语句

```
/opt/cloudera/cm/schema/scm_prepare_database.sh mysql cm -h 172.16.8.134 -uroot  
-pUpchina#3#2#1 --scm-host 172.16.8.134 scm Scm#3#2#1
```

1.5 配置 Agent

在所有节点上执行修改 agent 的配置，

```
vim /etc/cloudera-scm-agent/config.ini
server_host=cdh1
```

1.6 准备 parcels，将 CDH 相关文件拷贝到主节点/opt/cloudera/parcel-repo/

```
# 相关文件
CDH-6.3.1-1.cdh6.3.1.p0.1470567-el7.parcel
CDH-6.3.1-1.cdh6.3.1.p0.1470567-el7.parcel.sha1
manifest.json
需要将 CDH-6.3.1-1.cdh6.3.1.p0.1470567-el7.parcel.sha1 改成
CDH-6.3.1-1.cdh6.3.1.p0.1470567-el7.parcel.sha，否则安装时会重新下载，很重要！！

mv CDH-6.3.1-1.cdh6.3.1.p0.1470567-el7.parcel.sha1
CDH-6.3.1-1.cdh6.3.1.p0.1470567-el7.parcel.sha
```

1.7 启动

```
#主节点
systemctl start cloudera-scm-server

#主节点及其他节点
systemctl start cloudera-scm-agent
```

启动需要点时间，可以查看数据库 scm 中的 hosts 表，如果里面出现节点信息表示 server 启动完成

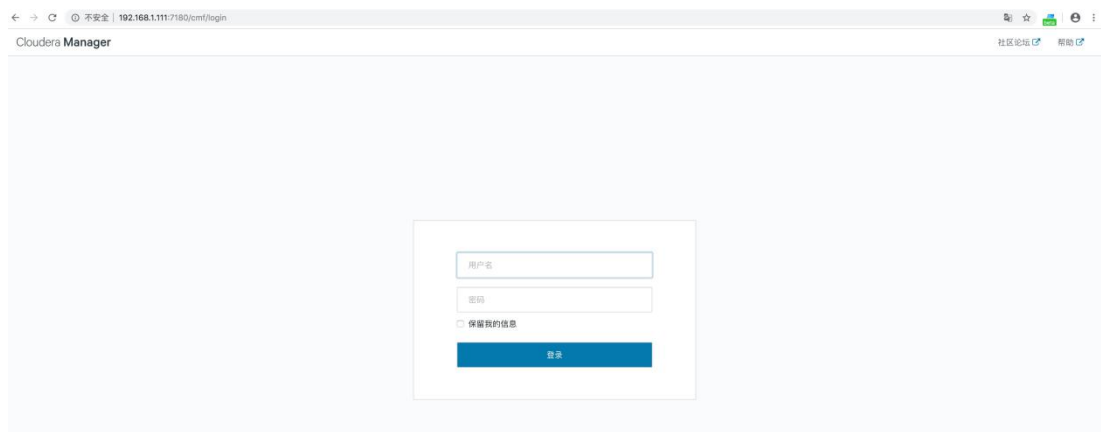
如果启动不成功，查看日志解决问题

```
/var/log/cloudera-scm-server
```

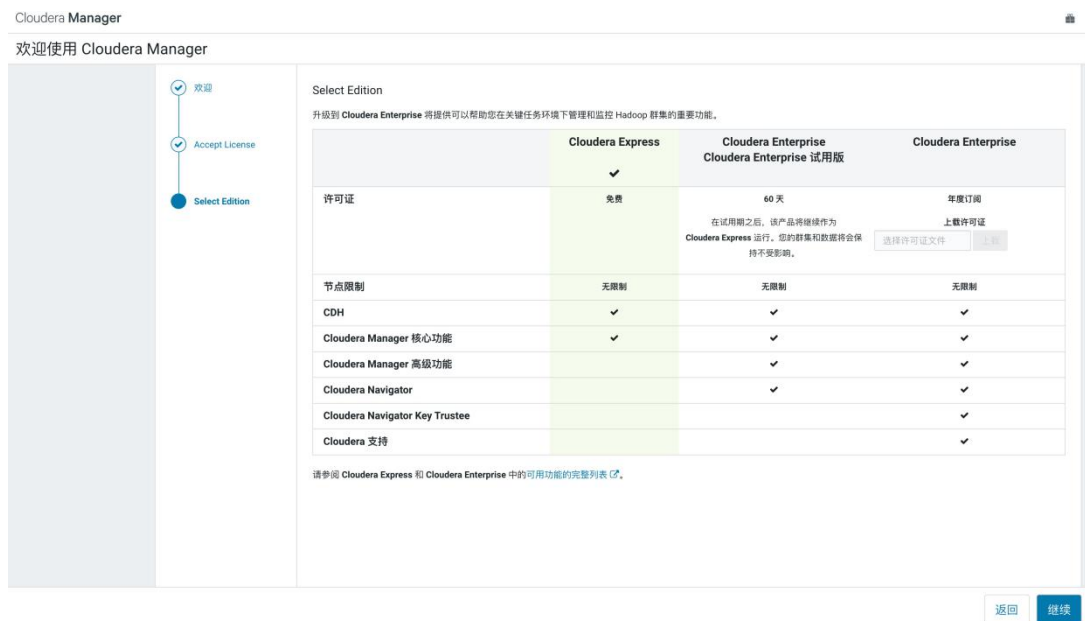
2. 页面操作

2.1 登录 <http://172.16.8.134:7180/cmfd/login>

登录界面,默认登录用户名和密码都是 admin



2.2 选择版本(这里我们选择免费版,根据具体需求选择)



2.3 修改集群名称



2.4 为集群安装指定主机

此处是在当前管理的主机里选择机器，不要在新主机里查找，另外这个地方显示的主机信息一定要对，因为主机名配置出问题，会导致无法识别主机名称和 IP，虽然 agent 的 uuid 是分配到了机器上，但是后边的安装无法进行。

欢迎

Cluster Basics

Specify Hosts

选择存储库

Install Parcels

Specify Hosts

In Cloudera Express, you can install CDH 6.0 or higher on up to 100 hosts.

新主机 当前管理的主机 (3)

这些主机不属于任何群集。请选择组成群集的主机。

单击第一个复选框，按住 Shift 键并单击最后一个复选框以选择范围。

<input checked="" type="checkbox"/> 主机名称 *	IP 地址	机架	CDH 版本	内核
<input checked="" type="checkbox"/> elink-1	192.168.1.111	/default	无	4
<input checked="" type="checkbox"/> elink-2	192.168.1.112	/default	无	4
<input checked="" type="checkbox"/> elink-3	192.168.1.113	/default	无	4

Displaying 1 - 3 of 3

2.5 选择 CDH 版本安装

这里我们的 parcel 会有 CDH 6.3.2 和 6.3.1 两个版本,我选择了 6.3.1 版本进行安装

欢迎

Cluster Basics

Specify Hosts

选择存储库

Install Parcels

Inspect Cluster

选择存储库

CDH and other software

Cloudera 建议使用 parcel 来代替软件包进行安装。因为 parcel 可以使服务二进制文件的部署和升级自动化，让 Cloudera Manager 轻松地管理群集上的软件。如果选择不使用 parcel，当有软件更新可用时，将需要您手动升级群集中所有主机上的包，并会阻止您使用 Cloudera Manager 的滚动升级功能。

选择方法 ☐ 使用数据包 ☒ 使用 Parcel (建议) [更多选项](#) [代理设置](#)

CDH 版本 对于此 Cloudera Manager 版本 (6.3.1) 太新的 CDH 版本不会显示。
☒ CDH-6.3.1-1.cdh6.3.1.p0.1470567
☐ CDH-5.16.2-1.cdh5.16.2.p0.8

其他 Parcel
☐ ACCUMULO-1.9.2-1.ACCUMULO6.1.0.p0.908695
☐ ACCUMULO-1.7.2-5.5.0.ACCUMULO5.5.0.p0.8
☒ 无
☐ KAFKA-4.1.0-1.4.1.0.p0.4
☒ 无
☐ SPARK-0.9.0-1.cdh4.6.0.p0.98
☒ 无
☐ SQOOP_NETEZZA_CONNECTOR-1.5.1c6
☐ SQOOP_NETEZZA_CONNECTOR-1.5.1c5
☒ 无
☐ SQOOP_TERADATA_CONNECTOR-1.7c5
☒ 无
☐ mkl-2019.5.281
☒ 无

返回

继续

2.6 检查集群环境

分别点击 [Inspect Network Performance](#) 与 [Inspect Hosts](#)

验证

✔ 检查器在所有 3 个主机上运行。
✔ 个别主机正确地解析了自己的主机名称。
✔ 查询存在冲突的初始脚本时未发现错误。
✔ 检查 /etc/hosts 时未发现错误。
✔ 所有主机均将 localhost 解析为 127.0.0.1。
✔ 检查过的主机均正确且及时地解析了彼此的主机名称。
✔ 主机时钟几乎同步（10 分钟内）。
✔ 整个群集中的主机时区一致。
✔ 无用户或组缺失。
✔ 软件包和 parcel 之间未检测到冲突。
✔ 没有存在已知错误的内核版本在运行。
❗ Cloudera 建议将 /proc/sys/vm/swappiness 设置为最大值 10。当前设置为 60。使用 sysctl 命令在运行时更改该设置并编辑 /etc/sysctl.conf，以在重启后保存该设置。您可以继续进行安装，但 Cloudera Manager 可能会报告您的主机由于交换而运行状况不良。以下主机将受到影响： ➤ 查看详细信息
❗ 已启用透明大页面压缩，可能会导致重大性能问题。请运行"echo never > /sys/kernel/mm/transparent_hugepage/defrag"和"echo never > /sys/kernel/mm/transparent_hugepage/enabled"以禁用此设置，然后将同一命令添加到 /etc/rc.local 等初始化脚本中，以便在系统重启时予以设置。以下主机将受到影响： ➤ 查看详细信息
✔ 已满足 CDH 5 Hue Python 版本依赖关系。
✔ Hue Psycopg2 version for PostgreSQL is satisfied for both CDH 5 and CDH 6.
✔ 0 hosts are running CDH 5 and 3 hosts are running CDH 6.
✔ 每个群集中检查过的主机均在运行相同版本的组件。
✔ 所有托管的主机都拥有不一致的 Java 版本。
✔ 所检查的所有 Cloudera Management Daemon 版本与服务器一致。
✔ 所检查的所有 Cloudera 管理代理版本与服务器一致。

点击查看信息会显示具体检测出问题的节点,这里可选择在全部节点执行以下命令或者选择检测出问题的节点

解决方案

第一个警告直接在所有节点上执行

```
echo 10 > /proc/sys/vm/swappiness
```

第二个警告按照上面的说明，在所有节点上执行这两条命令

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

然后将命令写到 /etc/rc.local，保证每次重启时能执行

配置完成后。重新运行监测程序

欢迎

Cluster Basics

Specify Hosts

选择存储库

Install Parcels

Inspect Cluster

Inspect Cluster

You have created a new empty cluster. Cloudera recommends that you run the following inspections. For accurate measurements, Cloudera recommends that they are performed sequentially.

Inspect Network Performance

高级选项

You can use this tool to evaluate the network performance between hosts, such as ping latency.

Ping Timeout

10

秒

Amount of time after which the inspector reports a failure.

Ping Count

10

Number of times the inspector pings each host.

Ping Packet Size

56

字节

Size of the test packet sent when pinging the hosts.

状态

上次运行

几秒前

持续时间

19.9s

显示检查结果

重新运行

更多

Inspect Hosts

No issues were detected, review the inspector results to see what checks were performed.

状态

上次运行

几秒前

持续时间

11.77s

显示检查结果

重新运行

更多

2.7 群集设置--选择安装服务--按需选择

Select Services

自定义角色分配

数据库设置

审核更改

命令详细信息

汇总

Select Services

选择要安装的服务组合。

Essentials

Management and support for Cloudera's distribution including Hadoop.
服务: HDFS、YARN (含 MapReduce 2) 、ZooKeeper、Oozie、Hive 和 Hue

Data Engineering

Process, develop, and serve predictive models.
服务: HDFS、YARN (含 MapReduce 2) 、ZooKeeper、Oozie、Hive、Hue 和 Spark

Data Warehouse

The modern data warehouse for today, tomorrow, and beyond.
服务: HDFS、YARN (含 MapReduce 2) 、ZooKeeper、Oozie、Hive、Hue 和 Impala

Operational Database

Real-time insights for modern data-driven business.
服务: HDFS、YARN (含 MapReduce 2) 、ZooKeeper、Oozie、Hive、Hue 和 HBase

所有服务

Everything you need to become information-driven, with complete use of the platform.
服务: HDFS、YARN (含 MapReduce 2) 、ZooKeeper、Oozie、Hive、Hue、HBase、Impala、Solr、Spark 和 Key-Value Store Indexer

自定义服务

选择您自己的服务。将自动包含所选服务需要的服务。只有在设置了初始群集之后才能添加 Flume。

本向导还将安装 Cloudera Management Service。有一系列组件可启用监控、报告、事件和警报；这些组件需要数据库存储信息，这将在下一页面上配置。

可选择自定义服务进行执行

2.8 集群设置--自定义角色

默认即可，如有特殊需求可自定义配置（比如 DataNode 建议最少三个机器，但是默认 master 不含 DataNode 角色，可以更改）

Select Services

自定义角色分配

数据库设置

审核更改

命令详细信息

汇总

自定义角色分配

您可在自定义新群集的角色分配。但如果分配不正确（例如，分配到某个主机上的角色太多）会影响服务性能。除非您有特殊需求，如已为特定角色预先选择特定主机，否则 Cloudera 建议改变分配情况。

还可以按主机查看角色分配。[按主机查看](#)

HBase

Master × 1 新建

elink-1

HBase REST Server

选择主机

HBase Thrift Server

选择主机

RegionServer × 2 新建

与 DataNode 相同

HDFS

NameNode × 1 新建

elink-1

SecondaryNameNode × 1 新建

elink-1

Balancer × 1 新建

elink-1

HttpFS

选择主机

NFS Gateway

选择主机

DataNode × 2 新建

elink-2-3

Cloudera Management Service

Service Monitor × 1 新建

elink-1

Activity Monitor

选择一个主机

Host Monitor × 1 新建

elink-1

Event Server × 1 新建

elink-1

Alert Publisher × 1 新建

elink-1

YARN (MR2 Included)

ResourceManager × 1 新建

elink-1

JobHistory Server × 1 新建

elink-1

NodeManager × 2 新建

与 DataNode 相同

ZooKeeper

Server × 1 新建

elink-1

2.9 开始执行安装命令

Select Services

自定义角色分配

审核更改

命令详细信息

汇总

首次运行 命令

状态 正在运行 10月 16, 10:31:21 上午 [中止](#)

✓ 已完成 0 个步骤（共 1 个）。

Show All Steps

Show Only Failed Steps

Show Only Running Steps

✓ Run a set of services for the first time

已完成 0/1 个步骤。

10月 16, 10:31:21 上午

➤ 依次运行 6 步骤

已完成 0/1 个步骤。

10月 16, 10:31:21 上午

首次运行 命令

状态 已完成 10月 16, 10:31:21 上午 2.5m

Finished First Run of the following services successfully: ZooKeeper, HDFS, HBase, YARN (MR2 Included), Cloudera Management Service.

✓ 已完成 1 个步骤（共 1 个）。

Show All Steps	Show Only Failed Steps	Show Only Running Steps
✓ Run a set of services for the first time	已成功完成 5 个步骤。	10月 16, 10:31:21 上午 2.5m
✓ 依次运行 6 步骤	已成功完成 5 个步骤。	10月 16, 10:31:21 上午 2.5m
➤ Ensuring that the expected software releases are installed on hosts.		10月 16, 10:31:21 上午 5.02s
➤ 并行运行 4 步骤		10月 16, 10:31:26 上午 15.71s
➤ 并行运行 2 步骤		10月 16, 10:31:42 上午 19.07s
➤ 并行运行 3 步骤		10月 16, 10:32:01 上午 89.84s
➤ 并行运行 2 步骤		10月 16, 10:33:31 上午 6.82s
➤ Verifying successful startup of services		10月 16, 10:33:38 上午 13.42s

2.10 安装完成



汇总

✔ 服务已安装、配置并在群集中运行。



四.部分参考

1. 安装 kudu 可参考配置

将 Kudu 服务添加到 CDH

审核更改

Kudu Master WAL Directory
Master Default Group
fs_wal_dir: /data1/kudu/master
Missing required value: Kudu Master WAL Directory

Kudu Master Data Directories
Master Default Group
fs_data_dirs: /data1/kudu/master, /data2/kudu/master, /data3/kudu/master
Missing required value: Kudu Master Data Directories

Kudu Tablet Server WAL Directory
Tablet Server Default Group
fs_wal_dir: /data1/kudu/tserver
Missing required value: Kudu Tablet Server WAL Directory

Kudu Tablet Server Data Directories
Tablet Server Default Group
fs_data_dirs: /data1/kudu/tserver, /data2/kudu/tserver, /data3/kudu/tserver
Missing required value: Kudu Tablet Server Data Directories

激活 Windows

2. Impala 安装时未添加 kudu 服务范围,后面添加



3. su hdfs 报错问题处理

```
[root@cdh1 mysql-5.7.24-rpm]# su hdfs
This account is currently not available.
cat /etc/passwd | grep hdfs
hdfs:x:995:992:Hadoop HDFS:/var/lib/hadoop-hdfs:/sbin/nologin
cat /etc/passwd
hdfs:x:995:992:Hadoop HDFS:/var/lib/hadoop-hdfs:/bin/bash
```

4. root 无法访问 mysql 问题处理

参考 <https://blog.csdn.net/u010316188/article/details/82717527>

五. 相关下载

1. CM

<https://archive.cloudera.com/cm6/6.3.1/repo-as-tarball/cm6.3.1-redhat7.tar.gz>

2. Parcel

<https://archive.cloudera.com/cdh6/6.3.1/parcels/CDH-6.3.1-1.cdh6.3.1.p0.1470567-el7.parcels>

<https://archive.cloudera.com/cdh6/6.3.1/parcels/CDH-6.3.1-1.cdh6.3.1.p0.1470567-el7.parcels.sha1>

<https://archive.cloudera.com/cdh6/6.3.1/parcels/manifest.json>

3. Jdk

<https://www.oracle.com/technetwork/java/javase/downloads/java-archive-javase8-2177648.html>

4. Mysql

<https://dev.mysql.com/downloads/mysql/5.7.html#downloads>

5. Mysql Jdbc

<http://central.maven.org/maven2/mysql/mysql-connector-java/5.1.47/mysql-connector-java-5.1.47.jar>

以上软件可在百度云下载

链接：<https://pan.baidu.com/s/1UH50Uweyi7yg6bV7dI02mQ>

提取码：nx7p

链接：<https://pan.baidu.com/s/1AqhC4mhv029NvaaSVT1kuQ>

提取码：mt1z

