# 修改准备工作

以下操作除MySQL外所有节点都要执行，且要用root权限执行！

## （一）关闭防火墙、SELINUX、IPv6

### 1、修改主机名及hosts

1.1 修改主机名

|  |
| --- |
| sudo hostnamectl set-hostname cdh001 |

1.2 修改hosts

|  |
| --- |
| sudo vim /etc/hosts  增加集群映射  30.23.77.57 cdh001  30.23.77.58 cdh002  30.23.77.59 cdh003  …… |

### 2、关闭防火墙

2.1 临时关闭

|  |
| --- |
| sudo systemctl stop firewalld.service |

2.2 关闭开机启动

|  |
| --- |
| sudo systemctl disable firewalld.service |

### 3、关闭SELINUX

3.1 编辑配置文件

|  |
| --- |
| sudo vim /etc/selinux/config  更改：SELINUX=disable |

3.2 重启生效

### 4、关闭IPv6

4.1 编辑配置文件

|  |
| --- |
| sudo vim /etc/sysvtl.conf  增加如下内容：  net.ipv6.conf.all.disable\_ipv6=1  sudo vim /etc/sysconfig/network  增加如下内容：  NETWORKING\_IPV6=no |

4.2 编辑网卡配置

|  |
| --- |
| sudo vim /etc/sysconfig/network-scripts/ifcfg-eth0  更改或增加：  IPV6INIT=no |

4.3 重启或执行sysctl -p命令生效

## （二）配置集群时间同步

### 1、所有节点安装ntp服务

|  |
| --- |
| sudo yum -y install ntp |

### 2、配置server

2.1 修改ntp配置文件

|  |
| --- |
| sudo vim /etc/ntp.conf  修改：  # Hosts on local network are less restricted.  restrict 30.23.77.0 mask 255.255.255.0 nomodify notrap  注释以下内容：  #server 0.contos.pool.ntp.org iburst  #server 1.contos.pool.ntp.org iburst  #server 2.contos.pool.ntp.org iburst  #server 3.contos.pool.ntp.org iburst  增加：  server 127.127.1.0  fudge 127.127.1.0 stratum 10 |

2.2 修改ntpd配置文件

|  |
| --- |
| sudo vim /etc/sysconfig/ntpd  增加：  SYNC\_HWCLOCK=yes |

### 3、配置其他节点

修改ntp配置文件

|  |
| --- |
| sudo vim /etc/ntp.conf  注释掉以下内容：  #server 0.contos.pool.ntp.org iburst  #server 1.contos.pool.ntp.org iburst  #server 2.contos.pool.ntp.org iburst  #server 3.contos.pool.ntp.org iburst  增加：  server 30.23.77.57 |

### 4、开启ntp服务

|  |
| --- |
| sudo systemctl start ntpd  sudo systemctl enable ntpd |

## （三）配置ssh免密登录

### 1、生成秘钥

|  |
| --- |
| ssh-keygen -t rsa  三次回车 |

### 2、发送秘钥到其他节点

|  |
| --- |
| ssh-copy-id 主机名 |

### 3、测试免密登录

|  |
| --- |
| ssh 主机名  第一次连接需要输入密码 |

### 4、切换root账户重复以上操作

|  |
| --- |
| su root  输入密码： |

## （四）安装java

### 1、卸载原生的open jdk

|  |
| --- |
| 查询已安装的jdk  sudo rpm -qa | grep -i jdk  卸载open jdk  sudo yum remove -y copy-jdk-config… |

### 2、安装java

|  |
| --- |
| tar -zxvf jdk-8u181-linux-x64.tar.gz -C /usr/openv |

### 3、配置环境变量

3.1 修改profile文件

|  |
| --- |
| sudo vim /etc/profile  增加：  # JAVA\_HOME  export JAVA\_HOME=/usr/openv/jdk1.8.0\_181  export PATH=$PATH:$JAVA\_HOME/bin |

3.2 在/usr/java/下创建jdk软连接

CDH不会自动使用系统环境变量中的JAVA\_HOME，而是使用Bigtop进行管理，如果不是在默认位置安装的jdk，就需要创建软连接工CDH使用

|  |
| --- |
| sudo ln -s /usr/openv/jdk1.8.0\_181 jdk1.8 |

## （五）安装MySQL

### 1、卸载已安装的MySQL

1.1 查看已安装的MySQL

|  |
| --- |
| sudo rpm -qa | grep -i mysql |

1.2 卸载已安装的MySQL

|  |
| --- |
| sudo rpm -e postfix-2.10.1-6.el7.x86\_64  sudo rpm -e mariadb-libs-5.5.56-2.el7.x86\_64-nodeps  …… |

### 2、安装MySQL

!!!注意：所有节点都需要安装shared-compat组件

|  |
| --- |
| sudo rpm -ivh MySQL-client-5.6.45-1.el7.x86\_64.rpm  sudo rpm -ivh MySQL-shared-compat-5.6.45-1.el7.x86\_64.rpm  sudo rpm -ivh MySQL-server-5.6.45-1.el7.x86\_64.rpm |

### 3、配置MySQL

3.1 启动MySQL服务

|  |
| --- |
| sudo systemctl start mysql.service |

3.2 登录MySQL进行配置

|  |
| --- |
| mysql -uroot -p临时密码 |

3.2.1 修改密码

|  |
| --- |
| set password=password(‘123456’); |

3.2.2 配置远程连接

|  |
| --- |
| 修改连接方式为%：  use mysql;  update user set host=‘%’ where host=‘localhost’;  刷新权限：  flush privileges; |

3.2.3 创建CM使用的数据库

|  |
| --- |
| create database hive DEFAULT utf8 COLLATE utf8\_general\_ci;  create database amon DEFAULT utf8 COLLATE utf8\_general\_ci;  create database hue DEFAULT utf8 COLLATE utf8\_general\_ci;  create database oozie DEFAULT utf8 COLLATE utf8\_general\_ci;  create database sentry DEFAULT utf8 COLLATE utf8\_general\_ci; |

# 二、安装Cloudera Manager

CM下载地址：

<https://archive.cloudera.com/cm5/cm/5/>

parcel下载地址：

<https://archive.cloudera.com/cdh5/parcels/5.15.2.3/>

## 1、解压CM的tar包

|  |
| --- |
| 创建安装路径：  mkdir /usr/openv/cdh-5.15.2  解压CM：  tar -zxvf cloudera-manager-centos7-cm5.15.2\_x86\_64.tar.gz -C \  /usr/openv/cdh-5.15.2 |

## 2、检查安装第三方依赖

|  |
| --- |
| sudo yum -y install chkconfig python bind-utils psmisc libxslt zlib sqlite cyrus-sasl-plain cyrus-sasl-gssapi fuse fuse-libs redhat-lsb httpd mod\_ssl |

## 3、配置MySQL的JDBC驱动

3.1 复制驱动包到/usr/share/java

|  |
| --- |
| sudo cp mysql-connector-java-5.1.14-bin.jar \  /usr/share/java/mysql-connector-java.jar |

3.2 复制驱动包到cm-5.15.2/share/cmf/lib

|  |
| --- |
| cp mysql-connector-java-5.1.14-bin.jar \  /usr/share/java/mysql-connector-java.jar |

## 4、修改Agent配置

|  |
| --- |
| vim cm-5.15.2/ect/cloudera-scm-agent/config.ini  修改：  server\_host=cdh001 |

## 5、分发文件到其他节点

5.1 分发脚本：

|  |
| --- |
|  |

5.2 分发文件

|  |
| --- |
| xsync cdh-5.15.2 |

## 6、所有节点创建CM用户

|  |
| --- |
| sudo useradd --system --home=cm-5.15.2/run/cloudera-scm-server/ \  --no-create-home --shell=/bin/false --comment “Cloudera SCM User” \  cloudera-scm |

## 7、初始化CM数据库

|  |
| --- |
| sudo cm-5.15.2/share/cmf/schema/scm\_prepare\_database.sh \  mysql cm -h cdh001 -uroot -p123456 –-scm-host cdh001 scm scm scm |

## 8、上传CDH安装文件

上传CDH的三个安装文件到cdh-5.15.2/cloudera/parcel-repo路径下

## 9、所有节点创建parcels文件夹

9.1 创建文件夹

|  |
| --- |
| mkdir cdh-5.15.2/cloudera/parcels |

9.2 修改parcel-repo和parcels拥有者为cloudera-scm

|  |
| --- |
| sudo chown cloudera-scm:cloudera-scm parcel-repo parcels |

## 10、启动CM服务

10.1 主节点

|  |
| --- |
| sudo cm-5.15.2/etc/init.d/cloudera-scm-server start  sudo cm-5.15.2/etc/init.d/cloudera-scm-agent start |

10.2 其他节点

|  |
| --- |
| sudo cm-5.15.2/etc/init.d/cloudera-scm-agent start |

## 11、查看端口

|  |
| --- |
| sudo netstat -anp | grep 7180 |

端口被占用则服务启动成功，后续使用Web页面安装CDH各组件

# 安装CDH

## （一）Web页面安装CDH组件

### 1、登录Web页面

用户名：admin

密码：admin

### 2、选择CM版本

### 3、选择主机

### 4、选择parcel方式安装

### 5、将CDH部署到所有节点并激活

### 6、主机检查

### 7、选择需要安装的组件

### 8、配置主机角色

### 9、配置数据库

所有节点都要安装MySQL的shared-compat组件，否则此处会连接不到数据库

### 10、使用默认配置至安装结束

## （二）Parcel方式安装Spark2

### 1、下载地址

1.1 csd：archive:cloudera.com/spark/csd

1.2 parcel：archive:cloudera.com/spark2/parcels/2.3.0.cloudera4

parcel的版本必须与csd相匹配

### 2、上传文件

2.1 将在csd页面下载的jar包上传到……cloudera/csd文件下

2.2 将剩余的3个文件上传到主机的……cloudera/parcel-repo文件夹下，并将…….sha1文件改名为…….sha，修改所有者为cloudera-scm

### 3、集群部署

3.1 CM主页：管理->设置，搜索本地，修改csd、parcel-repo、parcels路径，重启server和agent服务

3.2 CM主页：主机->parcel，选择Spark2，分配->激活

3.3 CM主页：添加服务，选择Spark2

3.4 选择依赖关系

3.5 分配主机，Gateway所有主机都要安装

3.6 默认配置至安装结束

## （三）Parcel方式安装LZO

### 1、下载地址

<http://archive.cloudera.com/gplextras5/parcels/5.15.2.3>

### 2、上传文件

将下载的3个文件上传到主机的……cloudera/parcel-repo文件夹下，并将…….sha1文件改名为…….sha，修改所有者为cloudera-scm

### 3、集群部署

3.1 CM主页：主机->parcel，检查新parcel

3.2 选择GPLEXTRAS，分配->激活

3.3 修改HDFS配置

3.3.1 压缩格式添加：

|  |
| --- |
| com.hadoop.compression.lzo.LzoCodec  com.hadoop.compression.lzo.LzopCodec |

3.4 修改YARN配置

3.4.1 MR程序classpath添加：

|  |
| --- |
| ……/CDH/cloudera/parcels/GPLEXTRAS/lib/hadoop/lib |

3.4.2 MR程序运行环境添加：

|  |
| --- |
| ;……/ CDH/cloudera/parcels/GPLEXTRAS/lib/hadoop/lib/native |

# 四、安装过程中遇到的问题及解决办法

## （一）Server启动问题

1、错误信息：

|  |
| --- |
| ERROR main:org.hibernate.engine.jdbc.spi.SqlExceptionHelper: Table ‘cm.CM\_VERSION’ doesn’t exist |

处理办法：不影响使用，可忽略

2、错误信息：

|  |
| --- |
| ERROR WebServerImpl:com.cloudera.server.web.cmf.search.components.SearchRepositoryManager: The server storage directory [/var/lib/cloudera-scm-server] doesn’t exist |

处理办法：手动创建/var/lib/cloudera-scm-server

3、错误信息：

|  |
| --- |
| ERROR WebServerImpl:com.cloudera.server.web.cmf.search.components.SearchRepositoryManager: No read/write permission to the server storage directory [/var/lib/cloudera-scm-server] |

处理办法：手动修改[/var/lib/cloudera-scm-server]文件夹所有者

4、错误信息：

|  |
| --- |
| ERROR ParcelUpdateService:com.cloudera.parcel.components.ParcelDownloaderImpl: Unable to retrieve remote parcel repository manifest |

处理办法：服务器未联网，无法从cm官网获取manifest，不影响使用，可忽略

## （二）Agent启动问题

1、错误信息：

|  |
| --- |
| MainThread agent ERROR Failed to connect to previous supervisor. |

处理办法：

1. 确保集群NTP同步服务有效后重启agent
2. 服务器重启后在启动agent有时也会报此错误，此时杀掉supervisor进程，重启agent

2、错误信息：

|  |
| --- |
| MainThread parcel ERROR Failed to activate alternatives for parcel ……  MainThread parcel ERROR Failed to deactivate alternatives for parcel …… |

处理办法：

1. 查看/etc/alternatives/、/usr/bin/路径下是否缺少对应的软连接，并手动创建
2. 脚本命令连接至：parcels/CDH/bin
3. conf路径连接至：parcels/CDH/etc/组件名/conf.

3、错误信息：

|  |
| --- |
| MonitorDaemom-Reporter throttling\_logger ERROR sending messages to firehose: mgmt-HOSTMONITOR-…… |

处理办法：错误2未处理agent进程也不会dead，但在Web安装组件时会报此错误，错误2处理完即可解决

4、错误信息：

|  |
| --- |
| Monitor-GenericMonitor throttling\_logger ERROR fetching metrics at ‘http://host:60030/jmx’ |

处理办法：检查主机名、hosts、agent的server配置是否一致

## （三）HDFS权限报错

CDH默认用户hdfs:supergroup为HDFS的最高管理用户，普通用户启动spark、kylin等对HDFS有写需求的服务时会报对HDFS没有写权限的错，此时可通过将用户添加到supergroup组来解决权限问题：

1. 添加supergroup组

|  |
| --- |
| sudo groupadd supergroup  grep supergroup /etc/group |

1. 将用户添加到supergroup组中

|  |
| --- |
| sudo usermod -a -G supergroup 用户名  id 用户名 |

1. 将上述信息同步到HDFS

|  |
| --- |
| sudo -u hdfs hdfs dfsadmin -refreshUserToGroupsMappings |

## （四）Hue相关问题

### 1、Hue的Load Balancer无法启动

1.1 安装httpd和mod\_ssl依赖

|  |
| --- |
| sudo yum -y install httpd mod\_ssl |

1.2 重启Load Balancer

### 2、配置Hue查看HBase

2.1 开启HBase Thrift服务

2.2 修改HBase配置（core-site.xml）

|  |
| --- |
| <property>  <name>hadoop.proxyuser.hue.hosts</name>  <value>\*</value>  </property>  <property>  <name>hadoop.proxyuser.hue.groups</name>  <value>\*</value>  </property>  <property>  <name>hadoop.proxyuser.hbase.hosts</name>  <value>\*</value>  </property>  <property>  <name>hadoop.proxyuser.hbase.groups</name>  <value>\*</value>  </property> |

2.3 修改Hue配置（hue\_safety\_value）

|  |
| --- |
| [hbase]  hbase\_conf\_dir={{HBASE\_CONF\_DIR}}  thrift\_transport=buffered |