

# CDH5.16.1集群企业真正离线部署(全网最细，配套视频，生产可实践)

视频:<https://www.bilibili.com/video/av52167219>

PS:建议先看课程视频1-2篇，再根据视频或文档部署，如有问题，及时与@若泽数据J哥联系。

## 一.准备工作

### 1.离线部署主要分为三块:

- a.MySQL离线部署
- b.CM离线部署
- c.Parcel文件离线源部署

### 2.规划:

节点	MySQL部署组件	Parcel文件离线源	CM服务进程	大数据组件
hadoop001	MySQL	Parcel	Activity Monitor	NN RM DN NM
hadoop002			Alert Publisher Event Server	DN NM
hadoop003			Host Monitor Service Monitor	DN NM

### 3.下载源:

- CM  
[cloudera-manager-centos7-cm5.16.1x8664.tar.gz](#)
- Parcel  
[CDH-5.16.1-1.cdh5.16.1.p0.3-el7.parcel](#)  
[CDH-5.16.1-1.cdh5.16.1.p0.3-el7.parcel.sha1](#)  
[manifest.json](#)
- JDK

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

下载jdk-8u202-linux-x64.tar.gz

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

下载mysql-5.7.26-el7-x86\_64.tar.gz

- MySQL jdbc jar

[mysql-connector-java-5.1.47.jar](#)

下载完成后要重命名去掉版本号,

mv mysql-connector-java-5.1.47.jar mysql-connector-java.jar

准备好百度云,下载安装包:

链接:<https://pan.baidu.com/s/10s-NaFLfztKuWImZTiBMjA> 密码:viqp

## 二.集群节点初始化

### 1.阿里云上海区购买3台, 按量付费虚拟机

CentOS7.2操作系统, 2核8G最低配置

### 2.当前笔记本或台式机配置hosts文件

- MAC: /etc/hosts
- Window: C:\windows\system32\drivers\etc\hosts

公网地址:

106.15.234.222 hadoop001

106.15.235.200 hadoop002

106.15.234.239 hadoop003

### 3.设置所有节点的hosts文件

私有地址、内网地址:

```
echo "172.19.7.96 hadoop001">> /etc/hosts
```

```
echo "172.19.7.98 hadoop002">> /etc/hosts
```

```
echo "172.19.7.97 hadoop003">> /etc/hosts
```

### 4.关闭所有节点的防火墙及清空规则

```
systemctl stop firewalld  
systemctl disable firewalld  
iptables -F
```

## 5.关闭所有节点的selinux

```
vi /etc/selinux/config  
将SELINUX=enforcing改为SELINUX=disabled  
设置后需要重启才能生效
```

## 6.设置所有节点的时区一致及时钟同步

## 6.1.时区

```
[root@hadoop001 ~]# date
Sat May 11 10:07:53 CST 2019
[root@hadoop001 ~]# timedatectl
    Local time: Sat 2019-05-11 10:10:31 CST
    Universal time: Sat 2019-05-11 02:10:31 UTC
    RTC time: Sat 2019-05-11 10:10:29
    Time zone: Asia/Shanghai (CST, +0800)
    NTP enabled: yes
NTP synchronized: yes
    RTC in local TZ: yes
    DST active: n/a
```

#查看命令帮助, 学习至关重要, 无需百度, 太👍

```
[root@hadoop001 ~]# timedatectl --help
timedatectl [OPTIONS...] COMMAND ...
```

Query or change system time and date settings.

-h --help	Show this help message
--version	Show package version
--no-pager	Do not pipe output into a pager
--no-ask-password	Do not prompt for password
-H --host=[USER@]HOST	Operate on remote host
-M --machine=CONTAINER	Operate on local container
--adjust-system-clock	Adjust system clock when changing local RTC mode

### Commands:

status	Show current time settings
set-time TIME	Set system time
set-timezone ZONE	Set system time zone
list-timezones	Show known time zones
set-local-rtc BOOL	Control whether RTC is in local time
set-ntp BOOL	Control whether NTP is enabled

### #查看哪些时区

```
[root@hadoop001 ~]# timedatectl list-timezones
Africa/Abidjan
Africa/Accra
Africa/Addis_Ababa
Africa/Algiers
Africa/Asmara
Africa/Bamako
```

### #所有节点设置亚洲上海时区

```
[root@hadoop001 ~]# timedatectl set-timezone Asia/Shanghai
[root@hadoop002 ~]# timedatectl set-timezone Asia/Shanghai
[root@hadoop003 ~]# timedatectl set-timezone Asia/Shanghai
```

## 6.2.时间

#所有节点安装ntp

```
[root@hadoop001 ~]# yum install -y ntp
```

#选取hadoop001为ntp的主节点

```
[root@hadoop001 ~]# vi /etc/ntp.conf
```

```
#time
```

```
server 0.asia.pool.ntp.org
```

```
server 1.asia.pool.ntp.org
```

```
server 2.asia.pool.ntp.org
```

```
server 3.asia.pool.ntp.org
```

#当外部时间不可用时, 可使用本地硬件时间

```
server 127.127.1.0 iburst local clock
```

#允许哪些网段的机器来同步时间

```
restrict 172.19.7.0 mask 255.255.255.0 nomodify notrap
```

#开启ntpd及查看状态

```
[root@hadoop001 ~]# systemctl start ntpd
```

```
[root@hadoop001 ~]# systemctl status ntpd
```

• ntpd.service - Network Time Service

```
Loaded: loaded (/usr/lib/systemd/system/ntpd.service; enabled; vendor preset: disabled)
```

```
Active: active (running) since Sat 2019-05-11 10:15:00 CST; 11min ago
```

```
Main PID: 18518 (ntpd)
```

```
CGroup: /system.slice/ntpd.service
```

```
└─18518 /usr/sbin/ntpd -u ntp:ntp -g
```

```
May 11 10:15:00 hadoop001 systemd[1]: Starting Network Time Service...
```

```
May 11 10:15:00 hadoop001 ntpd[18518]: proto: precision = 0.088 usec
```

```
May 11 10:15:00 hadoop001 ntpd[18518]: 0.0.0.0 c01d 0d kern kernel time sync enabled
```

```
May 11 10:15:00 hadoop001 systemd[1]: Started Network Time Service.
```

#验证

```
[root@hadoop001 ~]# ntpq -p
```

remote	refid	st	t	when	poll	reach	delay	offset	jitter
=====									
LOCAL(0)	.LOCL.	10	1	726	64	0	0.000	0.000	0.000

#其他从节点停止禁用ntpd服务

```
[root@hadoop002 ~]# systemctl stop ntpd
```

```
[root@hadoop002 ~]# systemctl disable ntpd
```

```
Removed symlink /etc/systemd/system/multi-user.target.wants/ntpd.service.
```

```
[root@hadoop002 ~]# /usr/sbin/ntpdate hadoop001
```

```
11 May 10:29:22 ntpdate[9370]: adjust time server 172.19.7.96 offset 0.000867 sec
```

#每天凌晨同步hadoop001节点时间

```
[root@hadoop002 ~]# crontab -e
00 00 * * * /usr/sbin/ntpdate hadoop001

[root@hadoop003 ~]# systemctl stop ntpd
[root@hadoop004 ~]# systemctl disable ntpd
Removed symlink /etc/systemd/system/multi-user.target.wants/ntpd.service.
[root@hadoop005 ~]# /usr/sbin/ntpdate hadoop001
11 May 10:29:22 ntpdate[9370]: adjust time server 172.19.7.96 offset 0.000867 sec
#每天凌晨同步hadoop001节点时间
[root@hadoop003 ~]# crontab -e
00 00 * * * /usr/sbin/ntpdate hadoop001
```

## 7.部署集群的JDK

```
mkdir /usr/java
tar -xzvf jdk-8u45-linux-x64.tar.gz -C /usr/java/
#切记必须修正所属用户及用户组
chown -R root:root /usr/java/jdk1.8.0_45

echo "export JAVA_HOME=/usr/java/jdk1.8.0_45" >> /etc/profile
echo "export PATH=${JAVA_HOME}/bin:${PATH}" >> /etc/profile
source /etc/profile
which java
```

## 8.hadoop001节点离线部署MySQL5.7(假如觉得困难哟,就自行百度RPM部署,因为该部署文档是我司生产文档)

- 文档链接:<https://github.com/Hackeruncle/MySQL>
- 视频链接:<https://pan.baidu.com/s/1jdM8Welg8syU0evL1-tDOQ> 密码:whic

## 9.创建CDH的元数据库和用户、amon服务的数据库及用户

```
create database cmf DEFAULT CHARACTER SET utf8;
create database amon DEFAULT CHARACTER SET utf8;
grant all on cmf.* TO 'cmf'@'%' IDENTIFIED BY 'Ruozedata123456!';
grant all on amon.* TO 'amon'@'%' IDENTIFIED BY 'Ruozedata123456!';
flush privileges;
```

## 10.hadoop001节点部署mysql jdbc jar

```
mkdir -p /usr/share/java/
cp mysql-connector-java.jar /usr/share/java/
```

## 三.CDH部署

### 1.离线部署cm server及agent

#### 1.1.所有节点创建目录及解压

```
mkdir /opt/cloudera-manager
tar -zxvf cloudera-manager-centos7-cm5.16.1_x86_64.tar.gz -C /opt/cloudera-manager
/
```

#### 1.2.所有节点修改agent的配置，指向server的节点hadoop001

```
sed -i "s/server_host=localhost/server_host=hadoop001/g" /opt/cloudera-manager/cm-5.16.1/etc/cloudera-scm-agent/config.ini
```

#### 1.3.主节点修改server的配置：

```
vi /opt/cloudera-manager/cm-5.16.1/etc/cloudera-scm-server/db.properties
com.cloudera.cmf.db.type=mysql
com.cloudera.cmf.db.host=hadoop001
com.cloudera.cmf.db.name=cmf
com.cloudera.cmf.db.user=cmf
com.cloudera.cmf.db.password=Ruozedata123456!
com.cloudera.cmf.db.setupType=EXTERNAL
```

#### 1.4.所有节点创建用户

```
useradd --system --home=/opt/cloudera-manager/cm-5.16.1/run/cloudera-scm-server/ -
-no-create-home --shell=/bin/false --comment "Cloudera SCM User" cloudera-scm
```

#### 1.5.目录修改用户及用户组

```
chown -R cloudera-scm:cloudera-scm /opt/cloudera-manager
```

### 2.hadoop001节点部署离线parcel源

### 2.1.部署离线parcel源

```
$ mkdir -p /opt/cloudera/parcel-repo
$ ll
total 3081664
-rw-r--r-- 1 root root 2127506677 May  9 18:04 CDH-5.16.1-1.cdh5.16.1.p0.3-el7.parcel
-rw-r--r-- 1 root root          41 May  9 18:03 CDH-5.16.1-1.cdh5.16.1.p0.3-el7.parcel.sha1
-rw-r--r-- 1 root root  841524318 May  9 18:03 cloudera-manager-centos7-cm5.16.1_x86_64.tar.gz
-rw-r--r-- 1 root root  185515842 Aug 10 2017 jdk-8u144-linux-x64.tar.gz
-rw-r--r-- 1 root root    66538 May  9 18:03 manifest.json
-rw-r--r-- 1 root root   989495 May 25 2017 mysql-connector-java.jar
$ cp CDH-5.16.1-1.cdh5.16.1.p0.3-el7.parcel /opt/cloudera/parcel-repo/
```

#切记cp时，重命名去掉1，不然在部署过程cm认为如上文件下载未完整，会持续下载

```
$ cp CDH-5.16.1-1.cdh5.16.1.p0.3-el7.parcel.sha1 /opt/cloudera/parcel-repo/CDH-5.16.1-1.cdh5.16.1.p0.3-el7.parcel.sha
$ cp manifest.json /opt/cloudera/parcel-repo/
```

### 2.2.目录修改用户及用户组

```
$ chown -R cloudera-scm:cloudera-scm /opt/cloudera/
```

## 3.所有节点创建软件安装目录、用户及用户组权限

```
mkdir -p /opt/cloudera/parcels chown -R cloudera-scm:cloudera-scm /opt/cloudera/
```

## 4.hadoop001节点启动Server

### 4.1.启动server

```
/opt/cloudera-manager/cm-5.16.1/etc/init.d/cloudera-scm-server start
```

4.2.阿里云web界面，设置该hadoop001节点防火墙放开7180端口

4.3.等待1min，打开 <http://hadoop001:7180> 账号密码:admin/admin

4.4.假如打不开，去看server的log，根据错误仔细排查错误

## 5.所有节点启动Agent

```
/opt/cloudera-manager/cm-5.16.1/etc/init.d/cloudera-scm-agent start
```

## 6.接下来，全部Web界面操作

<http://hadoop001:7180/>

账号密码:admin/admin



## 7.欢迎使用Cloudera Manager--最终用户许可条款与条件。勾选

clouderaMANAGER

Support

Welcome to Cloudera Manager

End User License Terms and Conditions

Cloudera Standard License

Version 2016-05-26

END USER LICENSE TERMS AND CONDITIONS

THESE TERMS AND CONDITIONS (THESE "TERMS") APPLY TO YOUR USE OF THE PRODUCTS (AS DEFINED BELOW) PROVIDED BY CLOUDERA, INC. ("CLOUDERA").

PLEASE READ THESE TERMS CAREFULLY.

IF YOU ("YOU" OR "CUSTOMER") PLAN TO USE ANY OF THE PRODUCTS ON BEHALF OF A COMPANY OR OTHER ENTITY, YOU REPRESENT THAT YOU ARE THE EMPLOYEE OR AGENT OF SUCH COMPANY (OR OTHER ENTITY) AND YOU HAVE THE AUTHORITY TO ACCEPT ALL OF THE TERMS AND CONDITIONS SET FORTH IN AN ACCEPTED REQUEST (AS DEFINED BELOW) AND THESE TERMS (COLLECTIVELY, THE "AGREEMENT") ON BEHALF OF SUCH COMPANY (OR OTHER ENTITY).

BY USING ANY OF THE PRODUCTS, YOU ACKNOWLEDGE AND AGREE THAT:

(A) YOU HAVE READ ALL OF THE TERMS AND CONDITIONS OF THIS AGREEMENT;

(B) YOU UNDERSTAND ALL OF THE TERMS AND CONDITIONS OF THIS AGREEMENT;

(C) YOU AGREE TO BE LEGALLY BOUND BY ALL OF THE TERMS AND CONDITIONS SET FORTH IN THIS AGREEMENT

IF YOU DO NOT AGREE WITH ANY OF THE TERMS OR CONDITIONS OF THESE TERMS, YOU MAY NOT USE ANY PORTION OF THE PRODUCTS.

☒ Yes, I accept the End User License Terms and Conditions.

If your download and use of Cloudera Manager are on behalf of a company that has an existing agreement with Cloudera for the use of the software, your action does not modify that existing agreement.


Back

12

Continue

## 8.欢迎使用Cloudera Manager--您想要部署哪个版本？ 选择Cloudera Express免费版

Upgrading to **Cloudera Enterprise** provides important features that help you manage and monitor your Hadoop clusters in mission-critical environments.

	Cloudera Express	Cloudera Enterprise Cloudera Enterprise Trial	Cloudera Enterprise
License	Free	60 Days  After the trial period, the product will continue to function as <b>Cloudera Express</b> . Your cluster and your data will remain unaffected.	Annual Subscription  Upload License Key <div>Select License FileUpload</div> Cloudera Enterprise is available in three editions: <ul style="list-style-type: none"><li>Basic Edition</li><li>Flex Edition</li><li>Cloudera Enterprise</li></ul>
Node Limit	Unlimited	Unlimited	Unlimited
CDH	✓	✓	✓
Core Cloudera Manager Features	✓	✓	✓
Advanced Cloudera Manager Features		✓	✓
Cloudera Navigator		✓	✓
Cloudera Navigator Key Trustee			✓
Cloudera  Support			✓

See [full list of features available](#) in Cloudera Express and Cloudera Enterprise.

Back

12

Continue

9.感谢您选择Cloudera Manager和CDH

Thank you for choosing Cloudera Manager and CDH.

This installer will install **Cloudera Express 5.16.1** and enable you to later choose packages for the services below (there may be some license implications).

- Apache Hadoop (Common, HDFS, MapReduce, YARN)
- Apache HBase
- Apache ZooKeeper
- Apache Oozie
- Apache Hive
- Hue (Apache licensed)
- Apache Flume
- Apache Impala
- Apache Sentry
- Apache Sqoop
- Cloudera Search (Apache licensed)
- Apache Spark

You are using Cloudera Manager to install and configure your system. You can learn more about Cloudera Manager by clicking on the **Support** menu above.

Before you proceed, be sure to checkout the [CDH and Cloudera Manager Requirements and Supported Versions](#)

- [Supported Operating Systems](#)
- [Supported Databases](#)
- [Supported JDK Versions](#)

Continue

## 10.为CDH集群安装指导主机。选择[当前管理的主机]，全部勾选

clouderaMANAGER

Supportadmin

Specify hosts for your CDH cluster installation.

New Hosts

Currently Managed Hosts (3)

These hosts do not belong to any clusters. Select some to form your cluster.

<input checked="" type="checkbox"/>	Name	IP	Rack	CDH Version	Status	Last Heartbeat
	<input type="text" value="Any Name"/>	<input type="text" value="Any IP"/>	<input type="text" value="Any Rack"/>	<div>All</div>	<div>All</div>	<div>All</div>
<input checked="" type="checkbox"/>	hadoop001	172.19.7.96	/default	None	Unknown Health	13.22s ago
<input checked="" type="checkbox"/>	hadoop002	172.19.7.98	/default	None	Unknown Health	10.9s ago
<input checked="" type="checkbox"/>	hadoop003	172.19.7.97	/default	None	Unknown Health	9.45s ago

Back

Continue

## 11.选择存储库

Select Repository

Cloudera recommends the use of parcels for installation over packages, because parcels enable Cloudera Manager to easily manage the software on your cluster, automating the deployment and upgrade of service binaries. Electing not to use parcels will require you to manually upgrade packages on all hosts in your cluster when software updates are available, and will prevent you from using Cloudera Manager's rolling upgrade capabilities.

Choose Method

☐ Use Packages

☒ Use Parcels (Recommended)

More Options

Proxy Settings

CDH Version

☒ CDH-5.16.1-1.cdh5.16.1.p0.3

☐ CDH-4.7.1-1.cdh4.7.1.p0.47

Versions of CDH that are too new for this version of Cloudera Manager (5.16.1) will not be shown.

Additional Parcels

☐ ACCUMULO-1.7.2-5.5.0.ACCUMULO5.5.0.p0.8

☐ ACCUMULO-1.4.4-1.cdh4.5.0.p0.65

☒ None

☐ IMPALA-2.1.0-1.impala2.0.0.p0.1995

☒ None

☐ KAFKA-4.0.0-1.4.0.0.p0.1

Back

123

Continue

## 12.集群安装--正在安装选定Parcel假如

本地parcel离线源配置正确，则"下载"阶段瞬间完成，其余阶段视节点数与内部网络情况决定。

## Cluster Installation

### Install Parcels

The selected parcels are being downloaded and installed on all the hosts in the cluster.



## 13.检查主机正确性

## Cluster Installation

Inspect hosts for correctnessRun Again

### Validations

Inspector ran on all 3 hosts.
Individual hosts resolved their own hostnames correctly.
No errors were found while looking for conflicting init scripts.
No errors were found while checking /etc/hosts.
All hosts resolved localhost to 127.0.0.1.
All hosts checked resolved each other's hostnames correctly and in a timely manner.
Host clocks are approximately in sync (within ten minutes).
Host time zones are consistent across the cluster.
No users or groups are missing.
No conflicts detected between packages and parcels.
No kernel versions that are known to be bad are running.
No problems were found with /proc/sys/vm/swappiness on any of the hosts.

13.1.建议将/proc/sys/vm/swappiness设置为最大值10。  
swappiness值控制操作系统尝试交换内存的积极；  
swappiness=0：表示最大限度使用物理内存，之后才是swap空间；  
swappiness=100：表示积极使用swap分区，并且把内存上的数据及时搬迁到swap空间；  
如果是混合服务器，不建议完全禁用swap，可以尝试降低swappiness。

临时调整：

```
sysctl vm.swappiness=10
```

永久调整：

```
cat << EOF >> /etc/sysctl.conf
# Adjust swappiness value
vm.swappiness=10
EOF
```

13.2.已启用透明大页面压缩，可能会导致重大性能问题，建议禁用此设置。

临时调整：

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

永久调整：

```
cat << EOF >> /etc/rc.d/rc.local
# Disable transparent_hugepage
echo never > /sys/kernel/mm/transparent_hugepage/defrag
echo never > /sys/kernel/mm/transparent_hugepage/enabled
EOF
```

```
# centos7.x系统，需要为"/etc/rc.d/rc.local"文件赋予执行权限
chmod +x /etc/rc.d/rc.local
```

## 14.自定义服务，选择部署Zookeeper、HDFS、Yarn服务

Core Hadoop

HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, and Hue

Core with HBase

HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, and HBase

Core with Impala

HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, and Impala

Core with Search

HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, and Solr

Core with Spark

HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, and Spark

All Services

HDFS, YARN (MapReduce 2 Included), ZooKeeper, Oozie, Hive, Hue, HBase, Impala, Solr, Spark, and Key-Value Store Indexer

Custom Services

Choose your own services. Services required by chosen services will automatically be included. Flume can be added after your initial cluster has been set up.

Service Type	Description
<input type="checkbox"/> HBase	Apache HBase provides random, real-time, read/write access to large data sets (requires HDFS and ZooKeeper).
<input checked="" type="checkbox"/> HDFS	Apache Hadoop Distributed File System (HDFS) is the primary storage system used by Hadoop applications. HDFS creates multiple replicas of data blocks and distributes them on compute hosts throughout a cluster to enable reliable, extremely rapid computations.

Back

123456

Continue

## 15.自定义角色分配

NameNode × 1 New

hadoop001

SecondaryNameNode × 1 New

hadoop001

Balancer × 1 New

hadoop001

HttpFS

Select hosts

NFS Gateway

Select hosts

DataNode × 3 New

All Hosts ▾

Cloudera Management Service

Service Monitor × 1 New

hadoop001

Activity Monitor × 1 New

hadoop001 ▾

Host Monitor × 1 New

hadoop001

Event Server × 1 New

hadoop001

Alert Publisher × 1 New

hadoop001

YARN (MR2 Included)

ResourceManager × 1 New

hadoop001

JobHistory Server × 1 New

hadoop002 ▾

NodeManager × 3 New

Same As DataNode ▾

ZooKeeper

Server × 3 New

hadoop[001-003] ▾

Back

123456

Continue

## 16.数据库设置

# Cluster Setup

## Setup Database

Configure and test database connections. Create the databases first according to the **Installing and Configuring an External Database** section of the [Installation Guide](#).

Activity Monitor ✓ Successful

Currently assigned to run on **hadoop001**.

Database Host Name: \*

Database Type:

Database Name : \*

Username: \*

Password:

☒ Show Password

Notes:

- The value in the **Database Host Name** field must match the value you used for the hostname when creating the database. [Learn more](#)
- If the database is not running on its default port, specify the port number using **host:port** in the **Database Host Name** field.
- It is highly recommended that each database is on the same host as the corresponding role instance.

123456

## 17. 审改设置，默认即可

**DataNode Failed Volumes Tolerated**  
dfs.datanode.failed.volumes.tolerated  
[Edit Individual Values](#)

Cluster 1 > DataNode Default Group ...and 2 others

?

**DataNode Data Directory**  
dfs.data.dir,  
dfs.datanode.data.dir  
[Edit Individual Values](#)

Cluster 1 > DataNode Default Group ...and 2 others

?

**NameNode Data Directories**  
dfs.name.dir,  
dfs.namenode.name.dir  
[Edit Individual Values](#)

Cluster 1 > NameNode Default Group

?

**HDFS Checkpoint Directories**  
fs.checkpoint.dir,  
dfs.namenode.checkpoint.dir  
[Edit Individual Values](#)

Cluster 1 > SecondaryNameNode Default Group

?

**Alerts: Mail Server Hostname**

Alert Publisher Default Group

?



123456

## 18. 首次运行

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

✔ Completed 5 of 5 step(s).

☒ Show All Steps    ☐ Show Only Failed Steps    ☐ Show Running Steps

> ✔ Ensuring that the expected software releases are installed on hosts.	May 11, 1:04:22 PM	59ms
✔ Deploying Client Configuration Successfully deployed all client configurations.	 Cluster 1 	May 11, 1:04:22 PM 15.32s
> ✔ Execute DeployClusterClientConfig for {hdfs,yarn} in parallel.	May 11, 1:04:22 PM	15.32s
✔ Start Cloudera Management Service, ZooKeeper Successfully completed 2 steps.	May 11, 1:04:37 PM	26.36s
> ✔ Execute 1 steps in sequence	May 11, 1:04:37 PM	22.88s
> ✔ Execute 2 steps in sequence	May 11, 1:04:38 PM	25.46s
✔ Start HDFS Successfully completed 1 steps.	May 11, 1:05:04 PM	67.33s
> ✔ Execute 3 steps in sequence	May 11, 1:05:04 PM	67.33s
> ✔ Start YARN (MR2 Included)	May 11, 1:06:11 PM	28.32s

19.恭喜您!

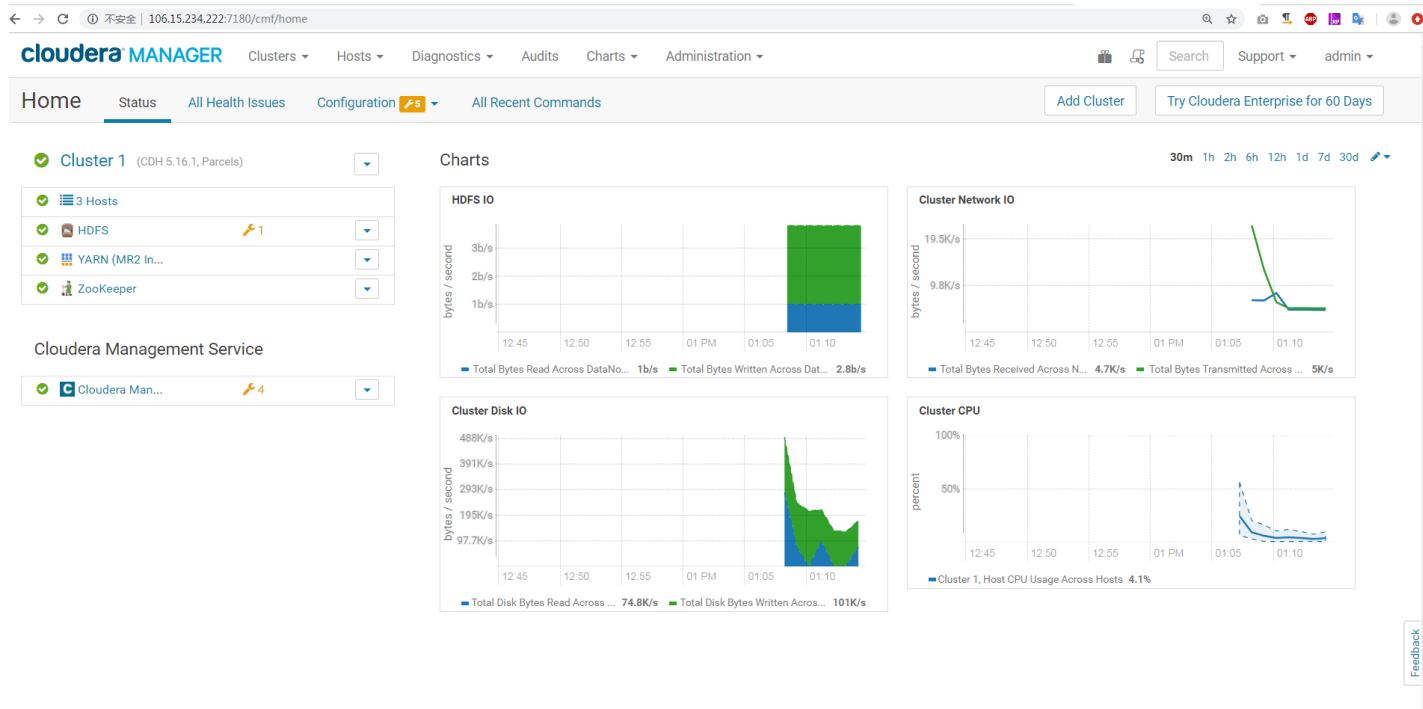
Cluster Setup

Congratulations!

 The services are installed, configured, and running on your cluster.

20.主页





## CDH全套课程目录，如有buy，加微信(ruoze\_star)

### 0. 青云环境介绍和使用

#### 1. Preparation

- 谈谈怎样入门大数据
- 谈谈怎样做好一个大数据平台的运营工作
- Linux机器,各软件版本介绍及安装(录播)

#### 2. Introduction

- Cloudera、CM及CDH介绍
- CDH版本选择
- CDH安装几种方式解读

#### 3. Install&UnInstall

- 集群节点规划,环境准备(NTP,Jdk and etc)
- MySQL编译安装及常用命令
- 推荐:CDH离线安装(踩坑心得,全面剖析)
- 解读暴力卸载脚本

#### 4. CDH Management

- CDH体系架构剖析
- CDH配置文件深度解析
- CM的常用命令
- CDH集群正确启动和停止顺序
- CDH Tsquery Language
- CDH常规管理(监控/预警/配置/资源/日志/安全)

#### 5. Maintenance Experiment

- HDFS HA 配置 及hadoop/hdfs常规命令
- Yarn HA 配置 及yarn常规命令

Other CDH Components HA 配置  
CDH动态添加删除服务(hive/spark/hbase)  
CDH动态添加删除机器  
CDH动态添加删除及迁移DataNode进程等  
CDH升级(5.10.0-->5.12.0)

## 6.Resource Management

Linux Cgroups  
静态资源池  
动态资源池  
多租户案例

## 7.Performance Tunning

Memory/CPU/Network/Disk及集群规划  
Linux参数  
HDFS参数  
MapReduce及Yarn参数  
其他服务参数

## 8.Cases Share

CDH4&5之Alternatives命令 的研究  
CDH5.8.2安装之Hash verification failed  
记录一次CDH4.8.6 配置HDFS HA 坑  
CDH5.0集群IP更改  
CDH的active namenode exit(GC)和彩蛋分享

## 9. Kerberos

Kerberos简介  
Kerberos体系结构  
Kerberos工作机制  
Kerberos安装部署  
CDH启用kerberos  
Kerberos开发使用(真实代码)

## 10.Summary

总结

**Join us if you have a dream.**

若泽数据官网: <http://ruozedata.com>

腾讯课堂, 搜若泽数据: <http://ruoze.ke.qq.com>

Bilibili网站,搜若泽数据: <https://space.bilibili.com/356836323>

[若泽大数据--官方博客](#)

[若泽大数据--博客一览](#)

[若泽大数据--内部学员面试题](#)

扫一扫，学一学：

# 学大数据就来若泽数据



扫一扫关注小程序



扫一扫关注公众号



扫一扫联系客服星星