



零基础(Cloudera CDH)

入门线路指导

(about 云网络书籍)

2015 年 03 月 21 日

整理自如下内容：

[cloudera\(CDH\)官网介绍：安装包、离线包该如何下载、官方文档等介绍](#)

[Cloudera Manager、CDH 零基础入门、线路指导](#)

[Cloudera Manager5 及 CDH5 在线（cloudera-manager-installer.bin）安装详细文档](#)

[Cloudera Manager Server5 及 Cloudera Manager Agents5 命令整理（about 云）](#)

[Cloudera Manager 5 和 CDH5 本地（离线）安装指南](#)

错误处理：

[Cloudera Manager Server5 启动后又挂掉的原因总结](#)

[卸载 Cloudera Manager 5.1.x.和 相关软件【官网翻译：高可用】](#)

[解析 Cloudera Manager 内部结构、功能包括配置文件、目录位置等](#)

cloudera(CDH)官网介绍：安装包、离线包该如何下载、官方文档等介绍

问题导读：

- 1.如何进入 **cloudera** 官网？
- 2.如何下载 **CM**、**CDH** 离线包？
- 3.如何找到 **CM** 的在线安装包？
- 4.**cloudera** 在哪可以找到？

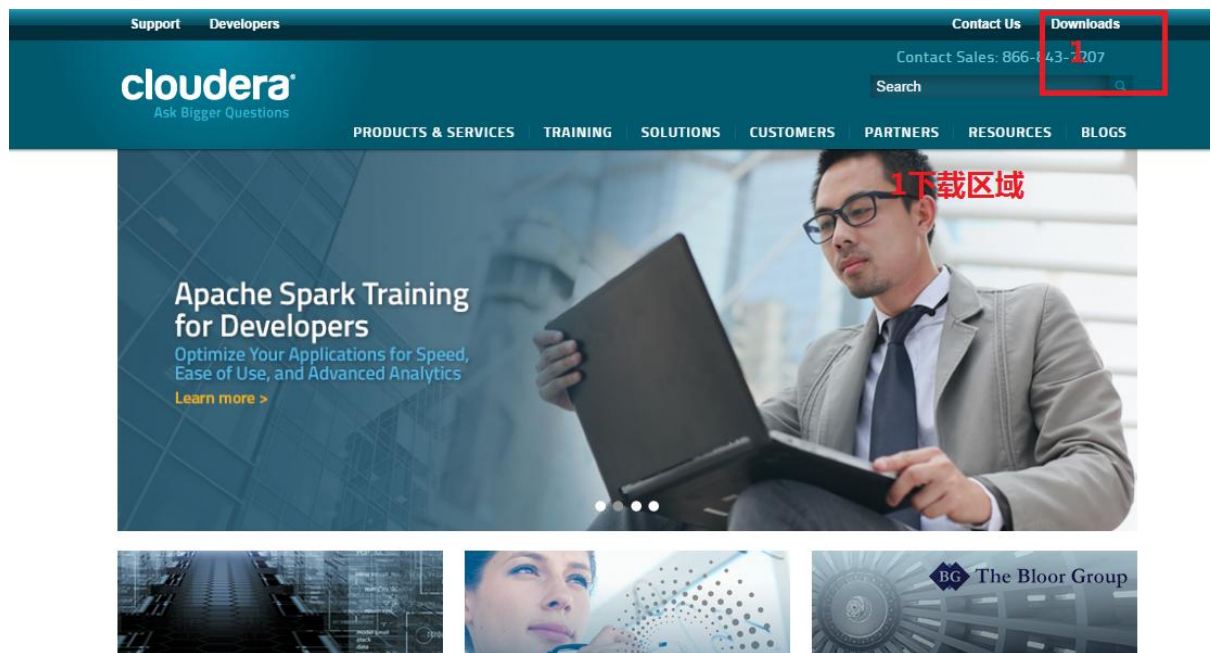


进入官网：

主页：

<http://www.cloudera.com/content/cloudera/en/home.html>

进入下载区域： 进入主页，我们会看到 download，点击 download



我们看到下面页面：

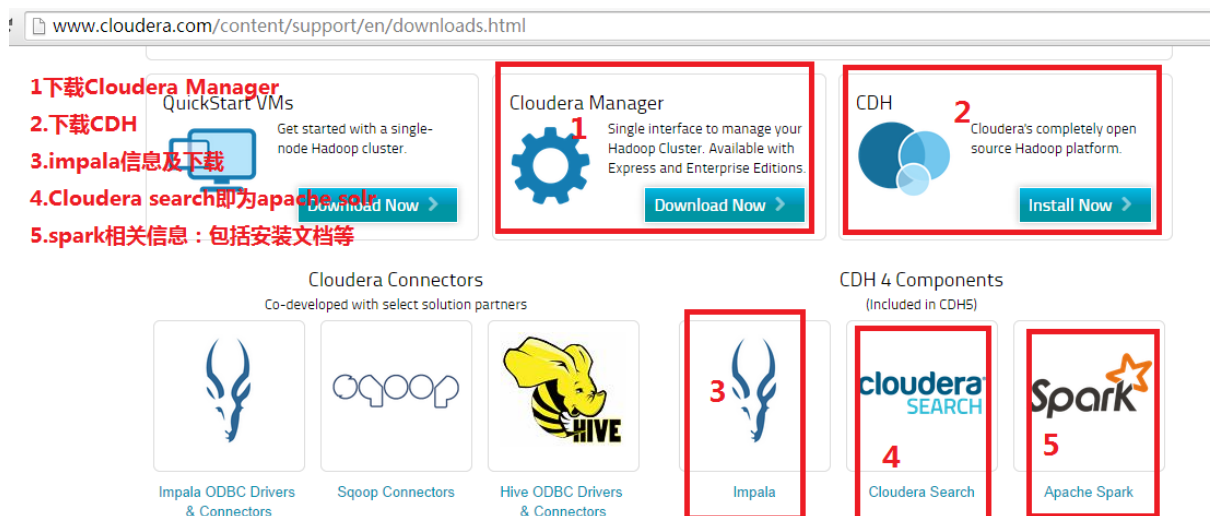


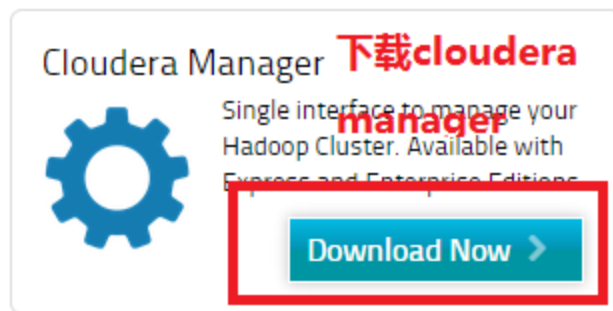
图 1

上面重点介绍：

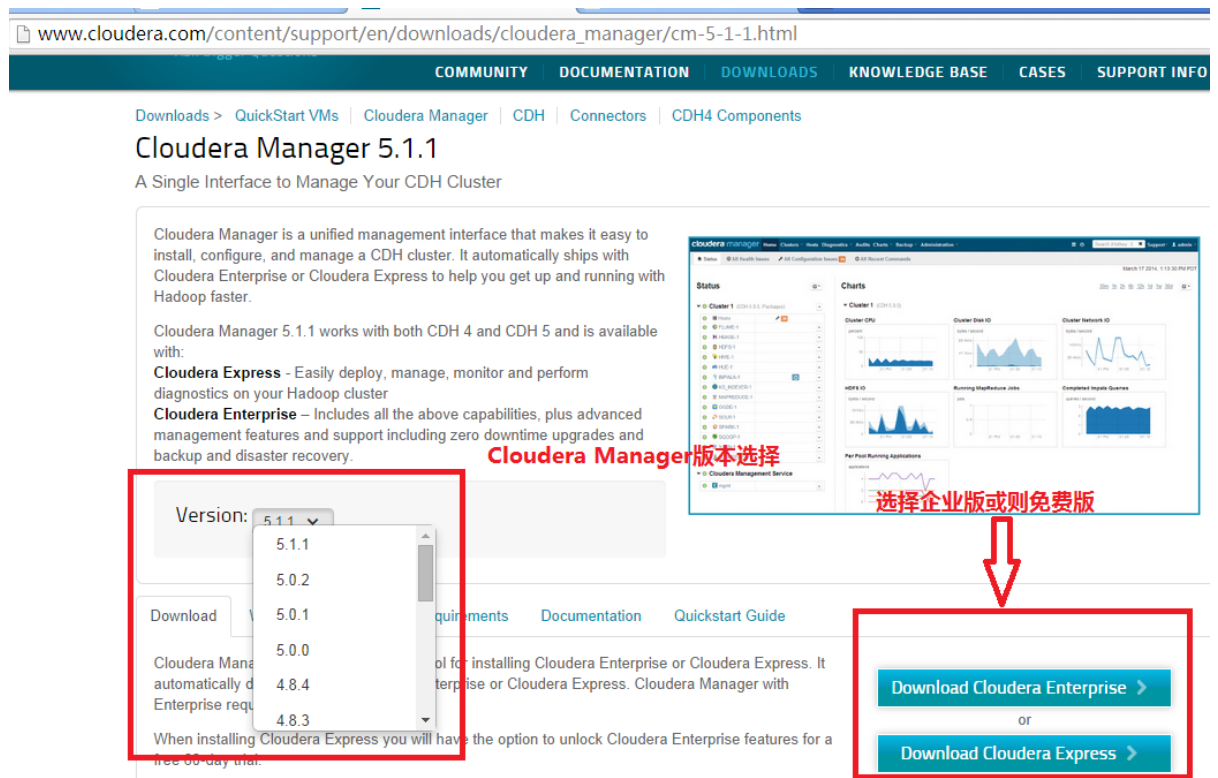
Cloudera Manager 与 CDH 如何识别安装包

一.Cloudera Manager 在线安装包识别

点击上图中



进入下面页面



我们选择版本之后，单击下载。

上面

1.Cloudera Enterprise 与 Cloudera Express 的区别是什么？

引用官网的内容：

Cloudera Manager is the recommended tool for installing Cloudera Enterprise or Cloudera Express. It automatically downloads with Cloudera Enterprise or Cloudera Express. Cloudera Manager with Enterprise requires a license.

When installing Cloudera Express you will have the option to unlock Cloudera Enterprise features for a free 60-day trial.

Once the trial has concluded, the Cloudera Enterprise features will be disabled until you obtain and upload a license.

大意应该是：

Cloudera Enterprise 是需要注册码的

Cloudera Expresss 中你是可以选择 Cloudera Enterprise 的，但是这个 60 天的试用期，试用期到了，你就不能用了。

除了选择 Cloudera Enterprise，当然还有 Cloudera Expresss，这个是长期使用没有问题的。

如下图所示说明：

欢迎使用 **Cloudera Manager**。您想要部署哪一个版本？

升级到 **Cloudera Enterprise Data Hub Edition** 将提供可以帮助您在关键任务环境下管理和监控 Hadoop 群集的重要功能。

| | | | |
|--|---|--|---|
| 1.免费版（长期使用） 2.试用版（试用60天） 3.企业版（上传许可证） | Cloudera Express 1 免费 节点限制：无限制 CDH：✓ Cloudera Manager 核心功能：✓ Cloudera Manager 高级功能：✓ Cloudera Navigator 审核组件：✓ Cloudera 支持：✓ | Cloudera Enterprise 数据集线器试用版 2 60 天 在试用期之后，该产品将继续作为 Cloudera Express 运行。您的群集和数据将会保持不受影响。 节点限制：无限制 CDH：✓ Cloudera Manager 核心功能：✓ Cloudera Manager 高级功能：✓ Cloudera Navigator 审核组件：✓ Cloudera 支持：✓ | Cloudera Enterprise 3 年度订阅 上传许可证 Cloudera Enterprise 在三个版本中可用： • Basic Edition • Flex Edition • Data Hub Edition 节点限制：无限制 CDH：✓ Cloudera Manager 核心功能：✓ Cloudera Manager 高级功能：✓ Cloudera Navigator 审核组件：✓ Cloudera 支持：✓ |
|--|---|--|---|

如需获得 Cloudera Express 和 Cloudera Enterprise 中提供的功能的完整列表，单击[此处](#)。

继续

2. 下载文件还是比较小的

文件名称为：cloudera-manager-installer.bin

附上这个在线包：

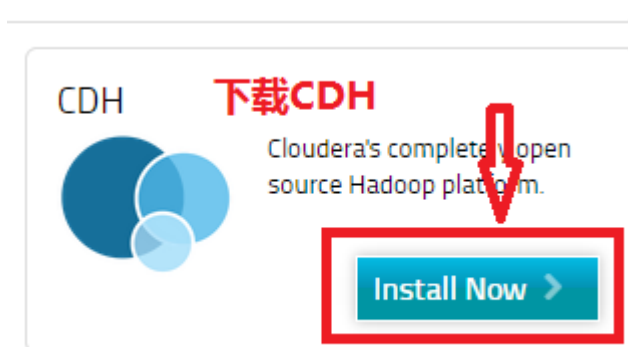


[cloudera-manager-installer.rar](#) （253.69 KB，下载次数：29）

说明：这个文件是我们在线安装所必须的文件，后面会讲到该如何在线安装

二.CDH 安装包识别

点击图 1 中，下面 CDH



CDH 下载图示 1: 点击可以查看以前版本的下载

2.是属于离线包

3、4 为版本发布说明及改变

| CDH 5.1.0 Packaging and Tarballs | | | | |
|---|---------------------------------|---------|---------------|--------------|
| <div>1</div> <div>Note: This section only contains packaging information for the current release. To see packaging and tarball information for older releases, refer CDH Packaging Information for Previous Releases.</div> | | | | |
| To view the overall release notes for CDH 5.x.x, CDH 5 Release Notes . | | | | |
| Component | Package Version | Tarball | Release Notes | Changes File |
| Apache Avro | avro-1.7.5+cdh5.1.0+30 | Tarball | Release notes | Changes |
| Apache Crunch | crunch-0.10.0+cdh5.1.0+14 | Tarball | Release notes | Changes |
| DataFu | pig-udf-datafu-1.1.0+cdh5.1.0+8 | Tarball | Release notes | Changes |
| Apache Flume | flume-ng-1.5.0+cdh5.1.0+10 | Tarball | Release notes | Changes |
| Apache Hadoop | hadoop-2.3.0+cdh5.1.0+795 | Tarball | Release notes | Changes |
| Apache HBase | hbase-0.98.1+cdh5.1.0+64 | Tarball | Release notes | Changes |
| HBase-Solr | hbase-solr-1.5+cdh5.1.0+12 | Tarball | Release notes | Changes |
| Apache Hive | hive-0.12.0+cdh5.1.0+369 | Tarball | Release notes | Changes |
| Hue | hue-3.6.0+cdh5.1.0+86 | Tarball | Release notes | Changes |
| Cloudera Impala | impala-1.4+cdh5.1.0+0 | (none) | Release notes | Changes |
| Kite SDK | kite-0.10.0+cdh5.1.0+120 | Tarball | Release notes | Changes |

1.可以下载以前版本比如CDH4、CDH3

2.是tar包下载

3.版本发布说明

4.做了哪些改变

| | | | | |
|----------------------------|---------------------------------|---------|---------------|---------|
| Llama | llama-1.0.0+cdh5.1.0+0 | Tarball | Release notes | Changes |
| Apache Mahout | mahout-0.9+cdh5.1.0+11 | Tarball | Release notes | Changes |
| Apache Oozie | oozie-4.0.0+cdh5.1.0+249 | Tarball | Release notes | Changes |
| Parquet | parquet-1.2.5+cdh5.1.0+130 | Tarball | Release notes | Changes |
| Parquet-format | parquet-format-1.0.0+cdh5.1.0+6 | Tarball | Release notes | Changes |
| Apache Pig | pig-0.12.0+cdh5.1.0+33 | Tarball | Release notes | Changes |
| Cloudera Search | search-1.0.0+cdh5.1.0+0 | Tarball | Release notes | Changes |
| Apache Sentry (incubating) | sentry-1.3.0+cdh5.1.0+155 | Tarball | Release notes | Changes |
| Apache Solr | solr-4.4.0+cdh5.1.0+231 | Tarball | Release notes | Changes |
| Apache Spark | spark-1.0.0+cdh5.1.0+41 | Tarball | Release notes | Changes |
| Apache Sqoop | sqoop-1.4.4+cdh5.1.0+55 | Tarball | Release notes | Changes |
| Apache Sqoop2 | sqoop2-1.99.3+cdh5.1.0+26 | Tarball | Release notes | Changes |
| Apache Whirr | whirr-0.9.0+cdh5.1.0+9 | Tarball | Release notes | Changes |
| Apache ZooKeeper | zookeeper-3.4.5+cdh5.1.0+29 | Tarball | Release notes | Changes |



Cloudera Manager 与 CDH 直接下载地址:

一、Cloudera Manager

Cloudera Manager5 下载

<http://archive.cloudera.com/cm5/>

| | | |
|-------------------------|-------------------|---|
| Parent Directory | 28-Oct-2013 04:49 | - |
| cm/ 1 | 28-Oct-2013 04:49 | - |
| debian/ | 09-Feb-2014 18:05 | - |
| installer/ | 28-Jul-2014 22:21 | - |
| redhat/ | 28-Oct-2013 04:46 | - |
| repo-2s-tarball/ | 28-Jul-2014 22:19 | - |
| sles/ | 28-Oct-2013 04:46 | - |
| ubuntu/ | 28-Oct-2013 04:46 | - |

看到下图:

1.离线包下载:

下面不同 Linux 对应不同的版本:

如下图

cloudera-manager-el6-cm5.1.2_x86_64.tar.gz

对应的为 **Centos el6_X64**

| | | |
|---|--|------------------------|
|  | Parent Directory | - |
|  | cloudera-manager-el5-cm5.0.0-beta-1_x86_64.tar.gz | 28-Oct-2013 04:53 184M |
|  | cloudera-manager-el5-cm5.0.0-beta-2_x86_64.tar.gz | 09-Feb-2014 11:40 335M |
|  | cloudera-manager-el5-cm5.0.0_x86_64.tar.gz | 01-Apr-2014 16:08 324M |
|  | cloudera-manager-el5-cm5.0.1_x86_64.tar.gz | 13-May-2014 14:36 328M |
|  | cloudera-manager-el5-cm5.0.2_x86_64.tar.gz | 07-Jun-2014 05:53 328M |
|  | cloudera-manager-el5-cm5.1.0_x86_64.tar.gz | 17-Jul-2014 04:19 385M |
|  | cloudera-manager-el5-cm5.1.1_x86_64.tar.gz | 25-Jul-2014 23:44 385M |
|  | cloudera-manager-el6-cm5.0.0-beta-1_x86_64.tar.gz | 28-Oct-2013 04:53 184M |
|  | cloudera-manager-el6-cm5.0.0-beta-2_x86_64.tar.gz | 09-Feb-2014 11:40 335M |
|  | cloudera-manager-el6-cm5.0.0_x86_64.tar.gz | 01-Apr-2014 16:08 324M |
|  | cloudera-manager-el6-cm5.0.1_x86_64.tar.gz | 13-May-2014 14:36 328M |
|  | cloudera-manager-el6-cm5.0.2_x86_64.tar.gz | 07-Jun-2014 05:53 328M |
|  | cloudera-manager-el6-cm5.1.0_x86_64.tar.gz | 17-Jul-2014 04:20 384M |
|  | cloudera-manager-el6-cm5.1.1_x86_64.tar.gz | 25-Jul-2014 23:44 384M |
|  | cloudera-manager-lucid-cm5.0.0-beta-1_amd64.tar.gz | 28-Oct-2013 04:53 183M |
|  | cloudera-manager-lucid-cm5.0.0-beta-2_amd64.tar.gz | 09-Feb-2014 11:40 335M |
|  | cloudera-manager-lucid-cm5.0.0_amd64.tar.gz | 01-Apr-2014 16:08 325M |
|  | cloudera-manager-lucid-cm5.0.1_amd64.tar.gz | 13-May-2014 14:36 328M |
|  | cloudera-manager-lucid-cm5.0.2_amd64.tar.gz | 07-Jun-2014 05:53 328M |

更多版本对应:

Linux ubuntu 对应

lucid(debian)

precise (ubuntu)

Centos 对应

el*

如 el5、el6

debian 对应

squeeze



wheezy

redhat 对应

sles*

2.针对不同操作的安装的安装如 **ubuntu**

3.**cloudera-manager-installer.bin**

| | | |
|--|--|------------------------|
|  | Parent Directory | - |
|  | cloudera-manager-installer.bin | 18-Jul-2014 22:21 499K |

这里也是我们在在线安装的一个文件,执行下面命令即可安装: 但是这里需要网络比较好

```
1. ./cloudera-manager-installer.bin
```

二、CDH 下载

我们通过下面地址:

<http://archive.cloudera.com/cdh5/>

进入下图, 图示中, 我们可以使用 CDH 离线包下载, 也可以使用 one-click-install.

| <u>Name</u> | <u>Last modified</u> | <u>Size</u> | <u>Description</u> |
|--|----------------------|-------------|--------------------|
|  Parent Directory | | - | |
|  cdh/ | 28-Oct-2013 03:09 | - | |
|  debian/ | 09-Feb-2014 18:02 | - | |
|  one-click-install/ | 01-Apr-2014 15:44 | - | |
|  parcels/ | 15-Jul-2014 20:25 | - | |
|  redhat/ | 28-Oct-2013 02:42 | - | |
|  repo-as-tarball/ | 15-Jul-2014 19:33 | - | |
|  sles/ | 28-Oct-2013 02:42 | - | |
|  ubuntu/ | 28-Oct-2013 02:42 | - | |

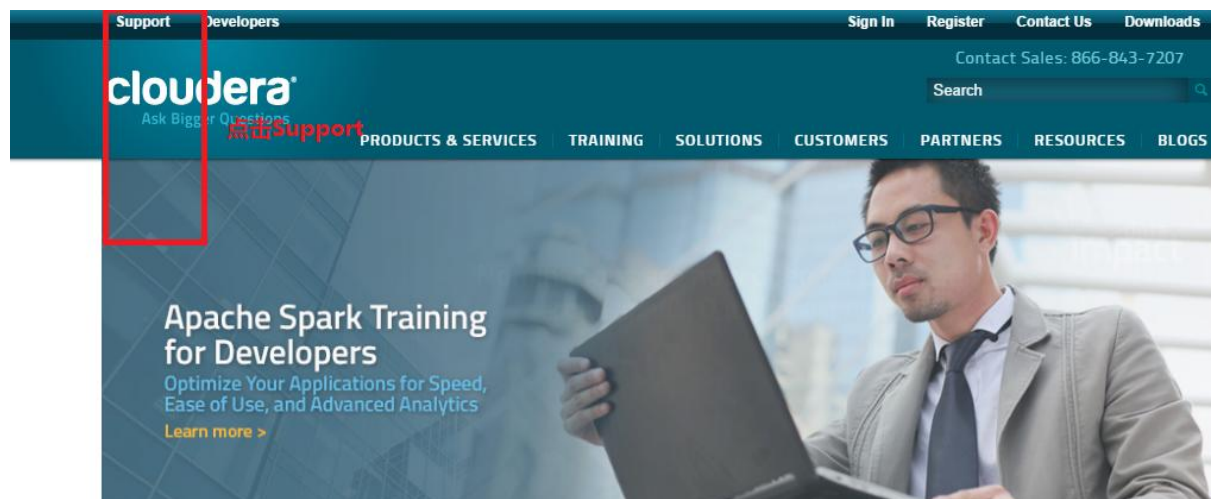


官方文档：

我们进入官方：

<http://www.cloudera.com/content/cloudera/en/home.html>

单击 Support



我们看到下面页面：

看到 document

<http://www.cloudera.com/content/support/en/home.html>



单击上图中 documentation

我们看到下面页面：

文档汇总区域：

这样我们就找到了官方的汇总文档区域。

包括：

Cloudera Manager 文档

CDH 文档等

<http://www.cloudera.com/content/support/en/documentation.html>

Cloudera Documentation

Welcome to the Cloudera Documentation page. Use the links below to find documentation that describes how to install, configure, secure, and use Cloudera Manager, CDH, Cloudera Impala, Cloudera Search, Cloudera Navigator, and other products.

Installation and User Documentation

Cloudera Manager

[Cloudera Manager 5 Documentation](#)

[Cloudera Manager 4 Documentation](#)

1

CDH

[CDH 5 Documentation](#)

[CDH 4 Documentation](#)

2

Cloudera Impala

[Cloudera Impala Documentation for CDH 5](#)

[Cloudera Impala Documentation for CDH 4](#)

3

Cloudera Search

[Cloudera Search Documentation for CDH 5](#)

[Cloudera Search Documentation for CDH 4](#)

Cloudera Navigator

[Cloudera Navigator Documentation](#)

Cloudera Connectors

[Cloudera Connectors Documentation](#)

1. Cloudera Manager 文档

2. CDH 文档

3. Cloudera Impala 文档

Other Documentation and Information

[Product Compatibility Matrix](#)

[Cloudera 5.x and 4.x Product Compatibility Matrix](#) 

For New Users

[Cloudera Manager 5 Quick Start Guide](#)

[Cloudera QuickStart VM](#)

[Cloudera Glossary](#)

[Hadoop Tutorial](#)

[CDH Package Versions, Tarballs, Repositories, and Repo File Locations](#)

[Cloudera Packaging and Versions Documentation](#)

[CDH Tarball Documentation](#)

Cloudera Security Bulletins

[Cloudera Security Bulletins](#)

Cloudera Third-Party License Information

[Cloudera Third-party Licenses](#)



汇总：

Cloudera 文档汇总

<http://www.cloudera.com/content/support/en/documentation.html>

CDH5 包汇总

<http://archive.cloudera.com/cdh5/>

CM5 包汇总

<http://archive.cloudera.com/cm5/>

官网 CDH5 下载

<http://www.cloudera.com/content/.../cdh/cdh-5-1-0.html>



Cloudera Manager 5.1.x.安装向导官网文档汇总

包括安装 Cloudera Manager 和 CDH，Cloudera Manager 5 支持管理 CDH4 和 CDH5

[Cloudera Manager 安装向导官网文档汇总](#)

- [Introduction to Cloudera Manager Installation](#)
- [Cloudera Manager Requirements](#)
- [Cloudera Manager and Managed Service Databases](#)
- [Installing Cloudera Manager, CDH, and Managed Services](#)
- [Managing Software Distribution](#)
- [Understanding Custom Installation Solutions](#)
- [Deploying Clients](#)
- [Testing the Installation](#)
- [Uninstalling Cloudera Manager and Managed Software](#)
- [Troubleshooting Installation and Upgrade Problems](#)
- [Configuring Ports for Cloudera Manager](#)

Cloudera Manager、CDH 零基础入门、线路指导

问题导读：

- 1.什么是 cloudera CM 、CDH?
- 2.CDH、CM 有哪些版本?
- 3.CDH、CM 有哪些安装方式?
- 4.CDH 如何开发?





我们知道 cloudera CDH 是为简化 hadoop 的安装,也对对 hadoop 做了一些封装。那么我们就像尝试学习 cloudera。cloudera 本质 hadoop 的封装,那么学起来,应该不难。不就是下载下来,然后自动安装,最后管理 hadoop 生态系统的一些服务。

的确是这样的,但是也没有我们想象的那么简单。

基础知识

那么我们就开始第一步吧,第一步之前,那就是你的基础,这个是很关键的,如果不会使用 Linux,那么你需要看一下另外一篇文章:[零基础学习 hadoop 到上手工作线路指导\(初级篇\)](#)。可以看看 Linux 的基础知识模块。需要掌握的内容还是不少的:

[集群搭建: 主机宽带拨号上网,虚拟机使用桥接模式,该如何 ping 通外网](#)

[集群搭建必备: 虚拟机之一实现 Host-only 方式上网](#)

[集群搭建必备: nat 模式设置静态 ip, 达到上网与主机相互通信](#)

[云技术基础: 学习 hadoop 使用零基础 linux\(Ubuntu\)笔记](#)

[搭建集群必知: Linux 常用命令及修改文件总结\(不断更新\)](#)

[Linux 网络接口 ifconfig 命令及认识网络接口 lo](#)

[linux 入门详细介绍](#)

[虚拟机安装 linux 网络配置资料大全](#)

[解决遇到 Linux 网络配置,从熟悉网络配置文件入手\(大体了解即可\)](#)

[linux 入门大全: 包括零基础入门, Linux 详细介绍](#)

[Linux 重启方式 init 0 init 1 init 3 init 5 init 6 这几个启动级别都代表什么意思?](#)

[Ubuntu 常用命令总结及修改 DNS 的多种方法总结](#)

[Linux 关机各种关机命令总结](#)

[Linux 基础必懂: eth0,eth1,eth2,lo 是什么意思?](#)

Linux 掌握了，那么我们是否需要掌握虚拟机的基础知识，在搭建 hadoop 集群中，可以这些是必须掌握的，那么 cloudera 是否需要，也是需要的，这里就不在罗列了，详细参考[零基础学习 hadoop 到上手工作线路指导（初级篇）](#)。但是 cloudera CM 的安装比起 hadoop 集群的安装对硬件的要求更高。内存至少 10G，为什么会这么多，如果少于 10G 是否可以，答案是可以的，但是后面你会遇到各种问题，或许都找不到答案。

对于 cloudera-scm-server 就需要至少 4G 的内存，cloudera-scm-agent 的内存至少也需要 1.5G 以上。那么如果你的机器是 8G 的，还是很吃力的。

在[安装](#)的过程中，后面有很多的服务安装遇到问题的可能性是很大的。

什么是 CDH

hadoop 是一个开源项目，所以很多公司在这个基础进行商业化，Cloudera 对 hadoop 做了相应的改变。

Cloudera 公司的发行版，我们将该版本称为 CDH。

很多新手问的最多的问题是，哪个是收费的，那个是免费的。

Cloudera Express 版本是免费的

Cloudera Enterprise 是需要购买注册码的

更多内容：

[Cloudera Hadoop 什么是 CDH 及 CDH 版本介绍](#)

[CDH（Cloudera）与 hadoop（apache）对比](#)

[大数据架构师基础：hadoop 家族，Cloudera 产品系列等各种技术](#)



官网介绍



当我们学习的时候，我们该如何学习，有两种方式：第一种：先理论，后实践
第二种：先实践，后理论

上面两种方式各有各的场合，如果你要去面试、考试、搞理论思想类，第一种方式是没有问题的。因为这些都是停留在思想层面的。

对于一个编程技术人员来讲，本人觉得第二种方式更加的短平快。

上面两种方式都有些绝对，最佳的方式就是实践与理论相结合。

这里撰文一下，引用孔子的一句话

学而不思则罔 思而不学则殆

只看书，不实践，或则只倒腾部署、停留在某些问题得不到解决，但是又不去看书、百度、谷歌查找解决问题，这样也是不行的。

好了，开始我们的官网：

官网是我们学习标准，所以我们首先要记住官网地址：

主页：

<http://www.cloudera.com/content/cloudera/en/home.html>

进入官网我们可以做些什么事情：

- 1.提供 API
- 2.查看部署文档
- 3.下载安装包

那么我们如何查找 API，如何查找部署文档，如何下载安装包？

详细可以查看

[cloudera\(CDH\)官网介绍：安装包、离线包该如何下载、官方文档等介绍](#)



安装包下载：

同时附上各个版本包的地址：

Cloudera 文档汇总

<http://www.cloudera.com/content/support/en/documentation.html>

CDH4、CDH5 包汇总

<http://archive.cloudera.com/cdh4/>

<http://archive.cloudera.com/cdh5/>

CM4、CM5 包汇总

<http://archive.cloudera.com/cm4/>

<http://archive.cloudera.com/cm5/>

官网 CDH5 下载

[http://www.cloudera.com/content/ ... /cdh/cdh-5-1-0.html](http://www.cloudera.com/content/.../cdh/cdh-5-1-0.html)

以前版本地址：

CDH1~CDH3

<http://archive-primary.cloudera.com/cdh/>

当我们看到安装的时候，同时也困扰着我们，我们该选择哪个安装包。

首先需要了解下 CM(Cloudera Manager)及 CDH 的安装方式：

CM(Cloudera Manager)有三种安装方式：

1.第一种使用 cloudera-manager-installer.bin 安装

这种安装方式，只要从官网下载 cloudera-manager-installer.bin，然后执行这个 bin 文件，剩下的就是等待下载和安装。但是这个时间不是一般的长，最好吃个饭，睡个觉，最后看到还在安装过程中。此帖安装步骤及遇到问题记录很详细，可参考

[Cloudera Manager5 及 CDH5 在线（cloudera-manager-installer.bin）安装详细文档](#)

[Cloudera Manager5 及 CDH5 安装指导（终极在线安装）](#)

遇到的问题：

当我们安装完毕 CM，那么我们就要添加主机，主机添加完毕，我们会看到安装进度：

cloudera manager

支持 admin

群集安装

正在安装。

已成功完成 3 个主机中的 0 个。

中止安装

| 主机名称 | IP 地址 | 进度 | 状态 | |
|--------------|--------------|-------------|---------------------------------|----------------------|
| 172.16.77.3 | 172.16.77.3 | <div></div> | 正在检测 Cloudera Manager Server... | 详细信息 |
| 172.16.77.30 | 172.16.77.30 | <div></div> | 正在刷新包元数据... | 详细信息 |
| 172.16.77.31 | 172.16.77.31 | <div></div> | 正在检测 Cloudera Manager Server... | 详细信息 |

返回

1 2 3 4 5 6

继续

但是并不是每个人都能看到上面进度，而是看到下面错误：

```
Detecting Cloudera Manager Server...
Detecting Cloudera Manager Server...
BEGIN host -t PTR 192.168.1.198
198.1.168.192.in-addr.arpa domain name pointer localhost.
END (0)
using localhost as scm server hostname
BEGIN which python
/usr/bin/python
END (0)
BEGIN python -c 'import socket; import sys; s = socket.socket(socket.AF_INET); s.settimeout(5.0);
s.connect((sys.argv[1], int(sys.argv[2]))); s.close();' localhost 7182
Traceback (most recent call last):
File "<string>", line 1, in <module>
File "<string>", line 1, in connect
socket.error: [Errno 111] Connection refused
END (1)
could not contact scm server at localhost:7182, giving up
waiting for rollback request
```

上面错误该如何解决：

可以使用下面方法解决：

```
1. mv /usr/bin/host /usr/bin/host.bak
```

2.第二种使用 rpm、yum、apt-get 方式在线安装

这种安装方式，在网上看的有点眼花缭乱，可能是因为我们对于 yum、apt、rpm 了解的不多造成的，对于 Linux 基础缺乏的，可以参考：[yum 与 rpm、apt 的区别：rpm 的缺陷及 yum 的优势](#)，总的来说：由于 rpm 依赖关系不好，所以产生了 yum，而 yum 和 apt 则都是 Linux 的包管理工具，并且解决了包与包的依赖关系。

这里列出一些安装文档

[ubuntu 12.04 安装 Cloudera Manager5 及 CDH5\(Mysql\)【添加 yum 源方式安装集群】](#)

遇到问题 1：内存过小

使用 yum 安装因为 mysql 的配置至少需要对 Cloudera server 4G 的内存，否则 mysql 的重启过程中，就会遇到 unknown instance.

遇到问题 2：界面无法访问

界面无法访问，分为两种情况

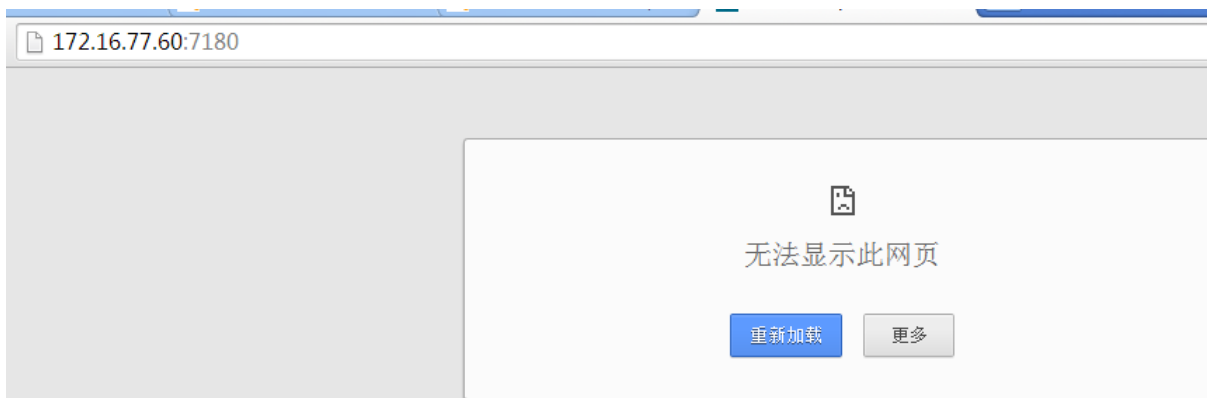
一种是 Cloudera-manager-server 没有完全起来

当我们运行下面命令：

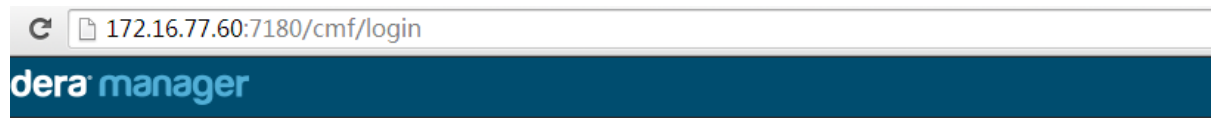
```
aboutyun@CDH:~$ sudo service cloudera-scm-server restart
```

如果现在，输入 url，访问 web，可能会是无法访问

host/ip:7180,我这里是 172.16.77.60: 7180



大概需要等 10 分钟左右，界面就可以访问了。



登录

👤 用户名:

🔒 密码:

☐ 在此计算机上保留我的信息。

登录

一种则是权限问题

所以我们初次安装尽量使用 root.还有我们需要会看日志。

Linux 日志的位置:

下面为 server 日志

```
root@CDH: /var/log/cloudera-scm-server# ls
cloudera-scm-server.log  cloudera-scm-server.out
```

下面为 agent 日志

```
root@CDH: /var/log/cloudera-scm-agent# ls
cloudera-scm-agent.log  cmf_listener.log  supervisord.out
cloudera-scm-agent.out  supervisord.log
```

同时在我们安装服务的过程中，我们还需要会查看角色日志，这样从 web 界面就是可以看到，因为在我们安装过程中，服务的安装没有那么顺利。

这里暂时没有截图，我们只要记住角色日志，在界面中找到即可。

遇到问题 3：无法找到 cloudera agent 客户端

我们使用下面命令安装了 agent

```
1. sudo apt-get install cloudera-manager-agent cloudera-manager-daemons
```

当然还有很多其它需要安装，当我们的 agent 安装成功后，为什么我们的的界面中不能发现 agent.

本来我们有三台，但是却只有一台，而且是 cloudera server 本地的 agent

这是因为我们的在安装的过程中，agent 都指向了本地 localhost.

我们打开配置文件

```
1. sudo nano /etc/cloudera-scm-agent/config.ini
```

修改

server_host 为 server ip 地址

server_port 为 7180

下图 CDH 即为 cloudera server 的 hostname

```
GNU nano 2.2.6      File: /etc/cloudera-scm-agent/config.ini

[General]
# Hostname of the CM server.
server_host=CDH
# Port that the CM server is listening on.
server_port=7182
## It should not normally be necessary to modify these.
# Port that the CM agent should listen on.
# listening_port=9000

# IP Address that the CM agent should listen on.
# listening_ip=
```

修改完毕：

首先重启 Agent

`sudo service cloudera-scm-agent restart`

然后重启 server

`sudo service cloudera-scm-server restart`

这时候会发现已经生效。

更详细内容可以参考

[记录 cloudera Manager 安装 Cloudera-Scm-Agent 如何指向 Cloudera-Scm-Server](#)

更多文档可以参考：

[Cloudera Manager 和 CDH 4 终极安装（一）](#)

[ClouderaManager 以及 CHD5.1.0 集群部署安装](#)

[Cloudera Manager （centos）离线安装详细介绍](#)

[CM5、CDH5 安装（CDH5 半自动离线安装）](#)

[Cloudera Manager 5 和 CDH5 本地（离线）安装指南](#)

[Cloudera Manager 和 CDH5 本地（离线）安装指导](#)

3.第三种使用是 Tarballs 的方式

这种方式网上资料很少

安装的命令如下：

```
1. $ tar xzf cloudera-manager*.tar.gz -C /opt/cloudera-manager
```

这个跟 **hadoop** 的安装方式是一样的

安装的方法详细可参考：

- [Installation Path C - Manual Installation Using Cloudera Manager Tarballs](#)

这里在列出官网给出的三种安装方式：

- [Installation Path A - Automated Installation by Cloudera Manager](#)
- [Installation Path B - Manual Installation Using Cloudera Manager Packages](#)
- [Installation Path C - Manual Installation Using Cloudera Manager Tarballs](#)

如果想安装 CM，这次再次强调一定要有足够的内存。并且我们安装完毕 CM，一般会

4. 安装失败处理

上面如果我们安装失败该如何处理：[使用 cloudera-manager-installer.bin\(Parcles\)安装失败后卸载 cloudera](#)

[卸载 Cloudera Manager 5.1.x.和 相关软件【官网翻译：高可用】](#)

5. 目录结构：

由于我们大多采用在线安装的方式，所以出血，我们根本不知道 **cloudera** 安装在在了什么位置，详细可以参考下面内容

[解析 Cloudera Manager 内部结构、功能包括配置文件、目录位置等](#)

亦可参考下面帖子：

[卸载 Cloudera Manager 5.1.x.和 相关软件【官网翻译：高可用】](#)



CDH 安装方式：

Yum/Apt 包，Tar 包，RPM 包，CM 安装

这些包的下载，可以在上文安装包汇总中可以找到。

CDH1~CDH3

<http://archive-primary.cloudera.com/cdh/>

CDH4~CDH5

<http://archive.cloudera.com/cdh4/>

<http://archive.cloudera.com/cdh5/>

安装文档，下面可以参考

[Linux（ubuntu12.04）单节点伪分布安装 CDH5.1.X 及提交 wordcount 到 yarn 高可靠文档](#)
[各个版本 Linux 单节点伪分布安装 CDH5.1.X 及提交 wordcount 到 yarn 高可靠文档](#)

[Hadoop CDH5 手动安装伪分布式模式](#)

通过 CM 方式[安装](#)，大多在安装 CM 的时候，已经安装了 CDH。



CDH 开发

安装完毕，我们还想开发提到开发，我们就想到 eclipse 插件，其实我们在开发过程中，插件作用就是帮助我们能够方便的看到在 Linux 的文件。

所以开发方式也有两种，

一种插件开发

我们如何找到 eclipse 插件，可以参考

[cloudera CDH（5）开发方式及 CDH eclipse 插件编译总结](#)

一种是无插件开发

无插件开发，也就是直接添加开发包

可以参考：[hadoop 开发方式总结及操作指导](#)

先总结到此，希望大家有所收获。



有的同学反映写的不够细，这里在补充一些内容：

Cloudera Manager Server 启动后又挂掉的原因总结

我们安装 CDH 的时候，会碰到 cloudera server 启动后又挂掉了，这是什么原因？

查看日志：包下面问题

可是明明已经安装 jdk，并且已经设置了 JAVA_HOME

```
1. java -version
```

也能返回版本

这是怎么回事，猜测 cloudera server（5）认 **oracle jdk**，对于 sun jdk 也需要是在线安装（至于是不识别 sun jdk 还未验证）。同时还必须是 jdk1.7.

```
root@CDH:~# sudo service cloudera-scm-server restart
cloudera-scm-server is already stopped
Starting cloudera-scm-server:  * Couldn't start cloudera-scm-server

root@CDH:~# cd /var/log/cloudera-scm-server/
root@CDH:/var/log/cloudera-scm-server# ls
cloudera-scm-server.out
root@CDH:/var/log/cloudera-scm-server# cat cloudera-scm-server.out
+=====+
|      Error: JAVA_HOME is not set and Java could not be found      |
+-----+
| Please download the latest Sun JDK from the Sun Java web site    |
|       > http://java.sun.com/javase/downloads/ <                 |
|                                                                     |
| Cloudera SCM requires Java 1.6 or later.                           |
| NOTE: This script will find Sun Java whether you install using the |
|       binary or the RPM based installer.                           |
+=====+
```

```
root@CDH:/var/log/cloudera-scm-server# sudo /usr/share/cmf/schema/scm_prepare_database.sh mysql -uroot -p --scm-host localhost scm scm scm_password
Enter database password:
=====+
|      Error: JAVA_HOME is not set and Java could not be found      |
|-----+
| Please download the latest Oracle JDK from the Oracle Java web site |
| > http://www.oracle.com/technetwork/java/javase/index.html <      |
|-----+
| Cloudera Manager requires Java 1.6 or later.                        |
| NOTE: This script will find Oracle Java whether you install using   |
| the binary or the RPM based installer.                              |
|-----+
=====+
```

也就是需要执行下面命令：

```
1. sudo apt-get -o Dpkg::Options::=--force-confdef -o Dpkg::Options::=--force-confold -y install
    oracle-j2sdk1.7
```

原因是在后面执行 **scm-server** 数据库配合的时候，如果找不到 **Java_home** 可能会初始化失败。

Java_home 默认安装路径如下。

```
1. /usr/lib/jvm/java-7-oracle-cloudera
```

从上面我们得出，如果想安装顺序，我们最好使用下面命令来安装 **jdk1.7**

```
sudo apt-get -o Dpkg::Options::=--force-confdef -o Dpkg::Options::=--force-confold -y install oracle-j2sdk1.7
```

执行上面命令，那么还需要做下面工作：

```
JAVA_HOME=/usr/lib/jvm/java-7-oracle-cloudera
CLASSPATH=.:/usr/java/jdk1.7.0_51/jre/lib
PATH="$JAVA_HOME/bin:/usr/local/sbin:/usr/local/bin:/usr/sbin:/
```

然后我们接着执行命令：

```
1. service cloudera-scm-server restart
```

认为终于好了，但是不幸的是，这次坚持了不到两分钟又挂掉了。

这次为什么会挂掉，不得不查看日志了：

```
Caused by: org.springframework.beans.factory.BeanCreationException: Error creating bean with name 'entityManagerFactoryBean': FactoryBean threw exception on object creation; nested exception is javax.persistence.PersistenceException: org.hibernate.exception.GenericJDBCException: Could not open connection
    at org.springframework.beans.factory.support.FactoryBeanRegistrySupport.doGetObjectFromFactoryBean(FactoryBeanRegistrySupport.java:149)
    at org.springframework.beans.factory.support.FactoryBeanRegistrySupport.getObjectFromFactoryBean(FactoryBeanRegistrySupport.java:102)
    at org.springframework.beans.factory.support.AbstractBeanFactory.getObjectForBeanInstance(AbstractBeanFactory.java:1440)
    at org.springframework.beans.factory.support.AbstractBeanFactory.doGetBean(AbstractBeanFactory.java:247)
    at org.springframework.beans.factory.support.AbstractBeanFactory.getBean(AbstractBeanFactory.java:192)
    at org.springframework.beans.factory.support.BeanDefinitionValueResolver.resolveReference(BeanDefinitionValueResolver.java:322)
    ... 17 more
```

```
Caused by: javax.persistence.PersistenceException: org.hibernate.exception.GenericJDBCException: Could not open connection
    at org.hibernate.ejb.AbstractEntityManagerImpl.convert(AbstractEntityManagerImpl.java:1387)
    at org.hibernate.ejb.AbstractEntityManagerImpl.convert(AbstractEntityManagerImpl.java:1310)
    at org.hibernate.ejb.AbstractEntityManagerImpl.throwPersistenceException(AbstractEntityManagerImpl.java:1397)
    at org.hibernate.ejb.TransactionImpl.begin(TransactionImpl.java:62)
    at com.cloudera.enterprise.AbstractWrappedEntityManager.beginForRollbackAndReadOnly(AbstractWrappedEntityManager.java:85)
    at com.cloudera.enterprise.dbutil.DbUtil.isInnoDBEnabled(DbUtil.java:472)
)
```

2014-09-22 17:51:27,328 ERROR [main:spi.SqlExceptionHandler@147] An attempt by a client to checkout a Connection has timed out.

2014-09-22 17:51:27,339 INFO [main:support.DefaultListableBeanFactory@422] Destroying singletons in org.springframework.beans.factory.support.DefaultListableBeanFactory@715dcdeb: defining beans [commandLineConfigurationBean,entityManagerFactoryBean,com.cloudera.server.cmf.TrialState,com.cloudera.server.cmf.TrialManager,com.cloudera.cmf.crypto.LicenseLoader]; root of factory hierarchy

2014-09-22 17:51:27,341 ERROR [main:cmf.Main@202] Server failed.

```
org.springframework.beans.factory.BeanCreationException: Error creating bean with name 'com.cloudera.server.cmf.TrialState': Cannot resolve reference to bean 'entityManagerFactoryBean' while setting constructor argument; nested exception is org.springframework.beans.factory.BeanCreationException: Error creating bean with name 'entityManagerFactoryBean': FactoryBean threw exception on object creation; nested exception is javax.persistence.PersistenceException: org.hibernate.exception.GenericJDBCException: Could not open connection
    at org.springframework.beans.factory.support.BeanDefinitionValueResolver.resolveReference(BeanDefinitionValueResolver.java:328)
    at org.springframework.beans.factory.support.BeanDefinitionValueResolver.resolveValueIfNecessary(BeanDefinitionValueResolver.java:106)
    at org.springframework.beans.factory.support.ConstructorResolver.resolveConstructorArguments(ConstructorResolver.java:616)
    at org.springframework.beans.factory.support.ConstructorResolver.autowireConstructor(ConstructorResolver.java:148)
    at org.springframework.beans.factory.support.AbstractAutowireCapableBeanFactory.autowireConstructor(AbstractAutowireCapableBeanFactory.java:1003)
    at org.springframework.beans.factory.support.AbstractAutowireCapableBeanFactory.createBeanInstance(AbstractAutowireCapableBeanFactory.java:907)
    at org.springframework.beans.factory.support.AbstractAutowireCapableBeanFactory.doCreateBean(AbstractAutowireCapableBeanFactory.java:485)
    at org.springframework.beans.factory.support.AbstractAutowireCapableBeanFactory.createBean(AbstractAutowireCapableBeanFactory.java:456)
    at org.springframework.beans.factory.support.AbstractBeanFactory$1.getObject(AbstractBeanFactory.java:293)
    at org.springframework.beans.factory.support.DefaultSingletonBeanRegistry.getSingleton(DefaultSingletonBeanRegistry.java:222)
```

看完日志，你是否还是云里雾里，这不得不说一下，我们该如何利用日志。

日志的作用是什么？日志的作用是提示，它大部分会给我们明确说遇到了什么问题，比如权限拒绝，等，但是有些你是看不明白的，那我们该如何处理？只能靠你的灵感了，**没有灵感怎么办？那就需要自己捋顺序，自己是怎么安装、怎么操作的，然后从中能够发现问题。渐渐的你就能培养出灵感了。**

上面其实我只知道报的是一个 **Java** 异常，而且还跟链接有关系。

灵感来了，就像到了，是因为没有执行下面命令：

配置 cloudera-manager-server 数据库

```
1.  sudo /usr/share/cmf/schema/scm_prepare_database.sh mysql -uroot -p --scm-host localhost scm
    scm scm_password
```

但是又出问题了，如下：

```
root@CDH:/var/log/cloudera-scm-server# sudo /usr/share/cmf/schema/scm_prepare_database.sh mysql -uroot -p --scm-host localhost scm scm scm_password
Enter database password:
+=====+
|      Error: JAVA_HOME is not set and Java could not be found      |
+-----+
| Please download the latest Oracle JDK from the Oracle Java web site |
|   > http://www.oracle.com/technetwork/java/javase/index.html   <  |
|                                                                     |
| Cloudera Manager requires Java 1.6 or later.                        |
| NOTE: This script will find Oracle Java whether you install using   |
|       the binary or the RPM based installer.                        |
+=====+
```

其实这个问题已经解决了，安装我们可以进入上面步骤。

当我们安装完毕，ok 了，终于好了：

当我们看到下面内容的时候，all done，说明我们配置成功了。

```
root@CDH:/var/log/cloudera-scm-server# sudo /usr/share/cmf/schema/scm_prepare_database.sh mysql -uroot -p --scm-host localhost scm scm scm_password
Enter database password:
Verifying that we can write to /etc/cloudera-scm-server
log4j:ERROR Could not find value for key log4j.appender.A
log4j:ERROR Could not instantiate appender named "A".
Creating SCM configuration file in /etc/cloudera-scm-server
Executing: /usr/lib/jvm/java-7-oracle-cloudera/bin/java -cp /usr/share/java/mysql-connector-java.jar:/usr/share/java/oracle-connector-java.jar:/usr/share/cmf/schema/./lib/* com.cloudera.enterprise.dbutil.DbCommandExecutor /etc/cloudera-scm-server/db.properties com.cloudera.cmf.db.
log4j:ERROR Could not find value for key log4j.appender.A
log4j:ERROR Could not instantiate appender named "A".
[2014-09-22 19:19:15,388] INFO      0[main] - com.cloudera.enterprise.dbutil.DbCommandExecutor.testDbConnection(DbCommandExecutor.java:238) - Successfully connected to database.
All done, your SCM database is configured correctly!
```

然后我们再次重启

```
1. service cloudera-scm-server restart
```

ok 始终在坚挺着，没有挂机，问题得到解决

Cloudera Manager5 及 CDH5 在线 (cloudera-manager-installer.bin) 安装详细 文档

问题导读:

1. Cloudera Manager5 如何使用 cloudera-manager-installer.bin 安装?
2. Cloudera Manager5 安装被中断该如何继续安装? 还是重新安装?
3. 安装过程 cloudera server 与 agent 不能通信该如何解决?
4. Cloudera Manager5 安装是否有要求?
5. 安装服务的过程中, 遇到错误, 该如何解决?



Cloudera Manager5 安装遇到很多的问题, 本来 Cloudera Manager 是为了简化部署, 但是由于条件有限, 安装 Cloudera Manager 如果不注意, 会成为一部血泪史。为了避免、多次、重复安装不成功, 所以你需要具备以下条件。

1. 较好的网速

至少 2m, 否则现在慢的要死

2.网络要稳定

一旦网络中断，可能你花费了半天或则一天的时间，就有可能会浪费，然后重装。

3.内存要足够大

我的电脑是 8G:

使用虚拟机安装

cloudera server: 4G

cloudera agent: 分别为 2G

电脑最后很卡，而且内存爆满。

所以官网给出的 10G 不是没有道理的。

安装系统为 ubuntu12.04



准备工作

规划三台:

cloudera scm server 内存分配 4G

cloudera scm agent 内存分配 2G

cloudera scm agent 内存分配 2G

同时三台都需要配置 hostname 及 hosts 如下

(这里一定要配置正确否则，会面会出通信问题)

```
root@CDH:~# cat /etc/hosts
127.0.0.1      localhost
#127.0.1.1    ubuntu
172.16.77.60  CDH.aboutyun.com  CDH
172.16.77.61  CDH1.aboutyun.com CDH1
172.16.77.62  CDG2.aboutyun.com CDH2
```

1. #127.0.1.1 ubuntu
2. 172.16.77.60 CDH.aboutyun.com CDH
3. 172.16.77.61 CDH1.aboutyun.com CDH1
4. 172.16.77.62 CDG2.aboutyun.com CDH2
- 5.

不知道如何配置，参考 [ubuntu 修改 hostname](#)

注意：本来想写一个完美的文档，上面有一个小失误，CDH2 可以写成 CDG2，这里不要照抄，而是失误，不过这个不影响，只要所有的 **hosts**，都配置成这样，通信是没有问题的。

同时三台都安装 ssh

```
1. sudo apt-get install ssh

user@CDH1:~$ sudo apt-get install ssh
Reading package lists... Done
Building dependency tree
Reading state information... Done
ssh is already the newest version.
0 upgraded, 0 newly installed, 0 to remove and 674 not upgraded.
```

在每台机器上执行下面操作：

```
1. sudo mv /usr/bin/host /usr/bin/host.bak

1. echo 0 > /proc/sys/vm/swappiness
```

官网介绍及 cloudera-manager-installer.bin 下载及[安装](#)视频

下载和运行 cloudera-manager-installer.bin

1. 下载 cloudera-manager-installer.bin

附件下载：



[cloudera-manager-installer.rar](#) (253.69 KB, 下载次数：66)

[官网下载](#)

如果使用 windows, 可以使用 scp, 复制到 Linux, 如果不会使用 scp, 可以参考

[新手指导：使用 WinSCP（下载）](#) [上文件到 Linux 图文教程](#)

[ssh、scp 不能远程登录与复制都失败解决方案](#)

2. 改变 cloudera-manager-installer.bin 的执行权限

```
1.  chmod u+x cloudera-manager-installer.bin  
2.
```

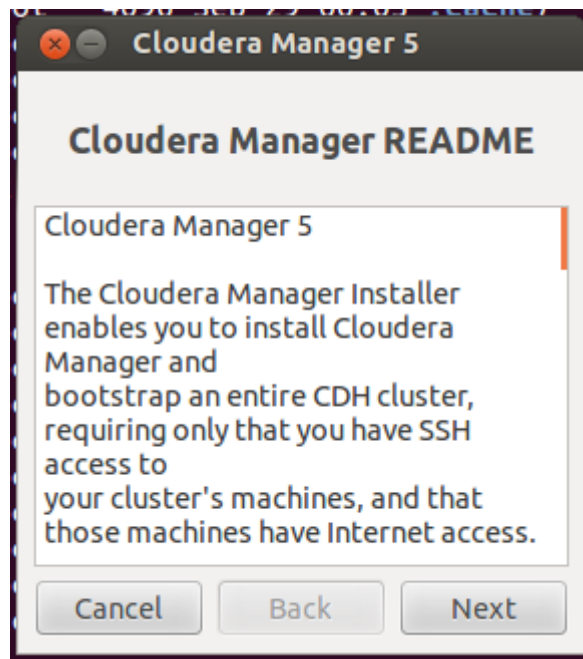
3. 执行 cloudera-manager-installer.bin

```
1.  sudo ./cloudera-manager-installer.bin
```

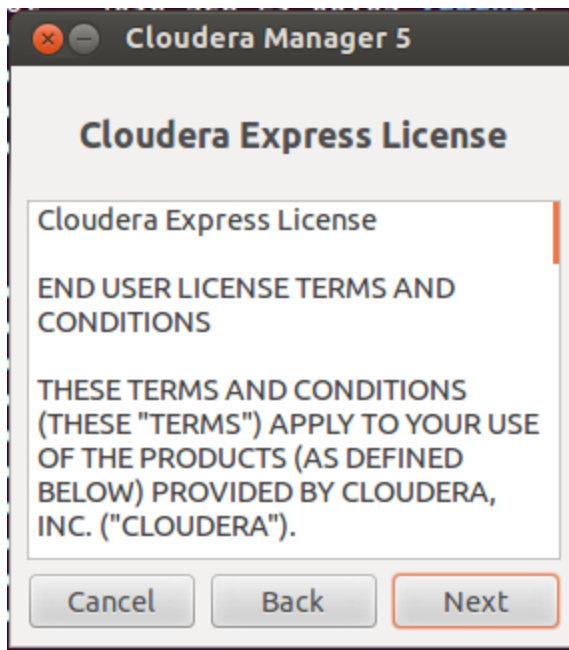
会弹出框，你只需要按下 **next** 和 **yes** 就可以了。

详细步骤如下：

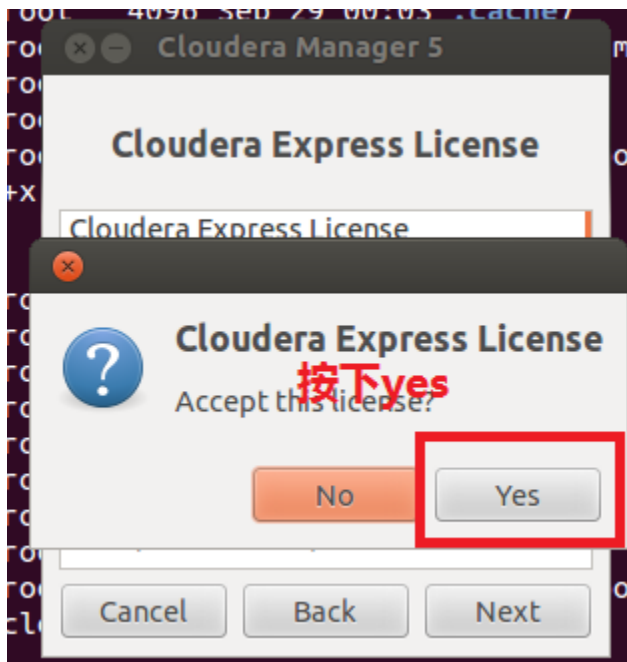
弹出对话框



单击 **NEXT**



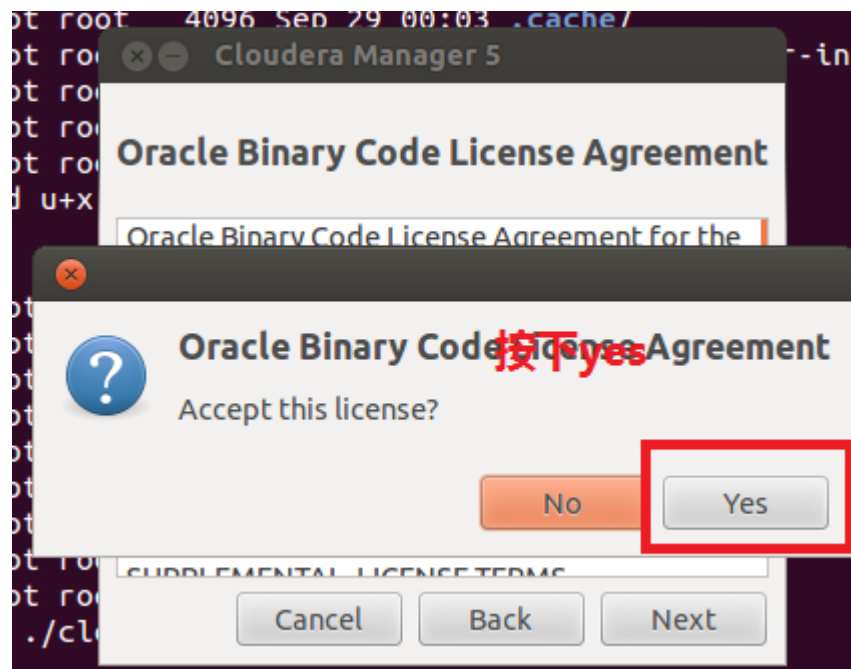
单击 **NEXT**



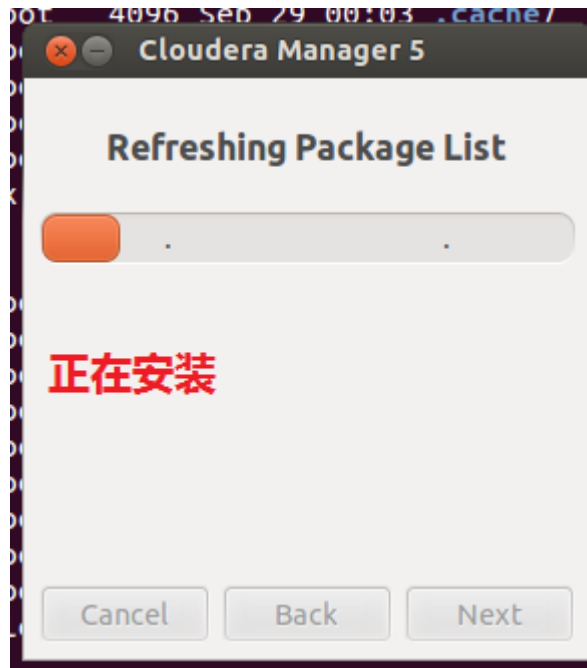
按下 **Yes**



单击 NEXT

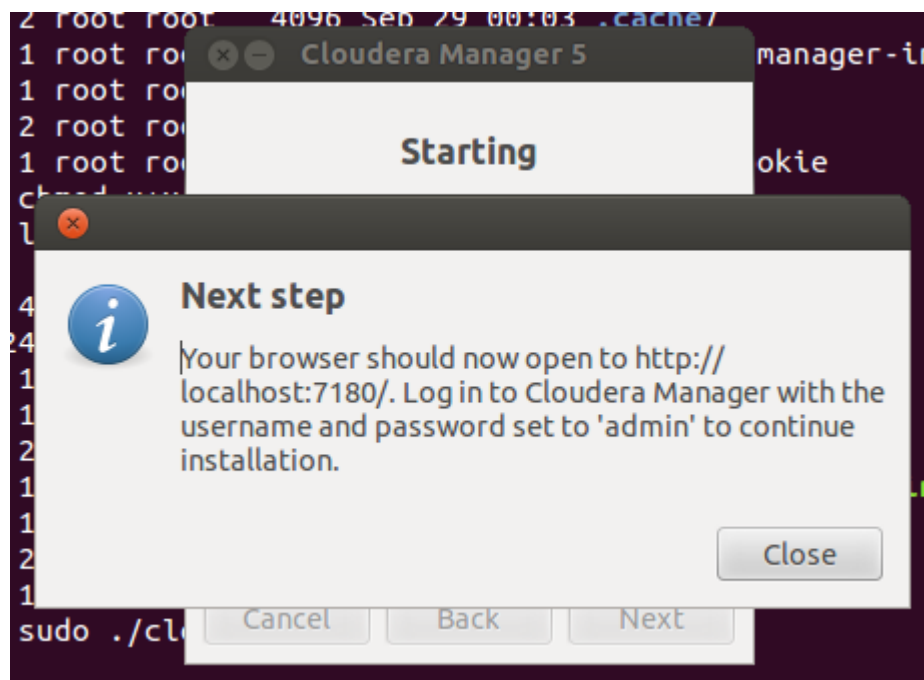


按下 Yes



最后是等待。

最后安装



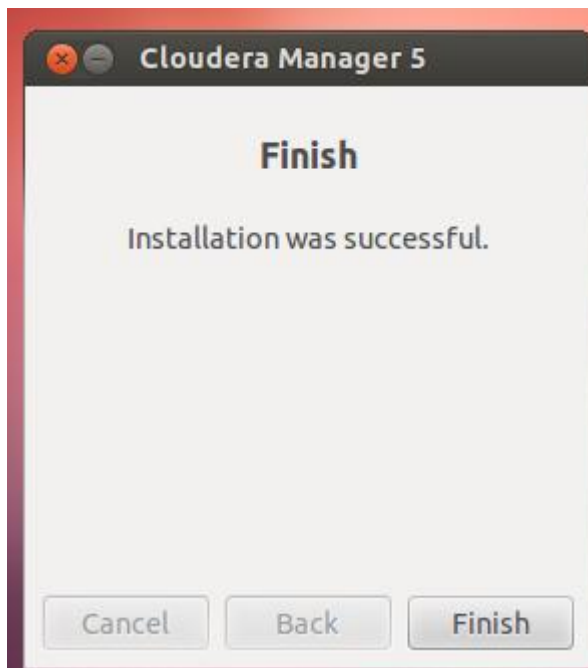
www.aboutyun.com

```
1. Your browser should now open to http://localhost:7180/. Log in to Cloudera Manager with the
   username and password set to 'admin' to continue installation.
```

在安装 cloudera-server 的 pc 机下，输入下面网址即可访问

```
1. http://localhost:7180
```

上面点击 close，会自动访问。关闭浏览器后，会看到下图



远程访问的输入 ip 即可下面形式

```
1. http://ip:7180
```

例如：

```
1. http://172.16.77.60: 7180
```

4.登录

输入用户名和密码

用户名： admin

密码： admin

安装完毕，单击 ok，系统会默认让你访问 localhost:7180。我们可以通过 ip 访问
如下图所示：<http://172.16.77.60>: 7180



登录

用户名:

密码:

☐ 在此计算机上保留我的信息。

登录

z 这样 cloudera-manager-server 就已经安装成功了。



我们接着进入下面：

登录成功，进入下面界面

选择 Cloudera Manager 版本 和指定主机

欢迎使用 Cloudera Manager。您想要部署哪一个版本？

升级到 Cloudera Enterprise Data Hub Edition 将提供可以帮助您在关键任务环境下管理和监控 Hadoop 集群的重要功能。

| | Cloudera Express | Cloudera Enterprise 数据集线器试用版 | Cloudera Enterprise |
|-------------------------|------------------|--|--|
| 许可证 | 免费 | 60 天 在试用期之后，该产品将继续作为 Cloudera Express 运行。您的群集和数据将会保持不受影响。 | 年度订阅 上载许可证 Cloudera Enterprise 在三个版本中可用： <ul style="list-style-type: none">Basic EditionFlex EditionData Hub Edition |
| 节点限制 | 无限制 | 无限制 | 无限制 |
| CDH | ✓ | ✓ | ✓ |
| Cloudera Manager 核心功能 | ✓ | ✓ | ✓ |
| Cloudera Manager 高级功能 | | ✓ | ✓ |
| Cloudera Navigator 审核组件 | | ✓ | ✓ |
| Cloudera 支持 | | | ✓ |

如需获得 Cloudera Express 和 Cloudera Enterprise 中提供的功能的完整列表，单击[此处](#)。

继续

我们选择 Clouder Express。

当然你选择 Clouder Enterprise 试用版也可以。单击继续

感谢您选择 Cloudera Manager 和 CDH。

将安装此安装程序Cloudera Express5.1.1，您可以稍后通过此安装程序选择以下服务的软件包（可能会涉及到许可证）。

- Apache Hadoop（Common、HDFS、MapReduce、YARN）
- Apache HBase
- Apache ZooKeeper
- Apache Oozie
- Apache Hive
- Hue（已获 Apache 许可）
- Apache Flume
- Cloudera Impala（许可的 Apache）
- Apache Sqoop
- Cloudera Search（许可的 Apache）
- Apache Spark

您正在使用 Cloudera Manager 安装和配置您的系统。您可以通过单击上面的支持菜单了解更多有关 Cloudera Manager 的信息。

继续

为 CDH 集群安装指定主机

为 CDH 群集安装指定主机。

应使用主机用于标识自身的同一主机名称 (FQDN) 来指定主机。

Cloudera 建议包括 Cloudera Manager Server 的主机。这样还将对该主机进行运行状况监控。

提示: 使用[模式](#) 搜索主机名称和或 IP 地址。

172.16.77.60,172.16.77.61,172.16.77.62

单击搜索

SSH 端口: 22

搜索

返回

继续

为 CDH 群集安装指定主机。

应使用主机用于标识自身的同一主机名称 (FQDN) 来指定主机。

Cloudera 建议包括 Cloudera Manager Server 的主机。这样还将对该主机进行运行状况监控。

提示: 使用[模式](#) 搜索主机名称和或 IP 地址。

已扫描 3 个主机，其中 3 个正在运行 SSH。 新搜索

| <input checked="" type="checkbox"/> | 已扩展查询 | 主机名称 (FQDN) | IP 地址 | 当前受管 | 结果 |
|-------------------------------------|--------------|-------------------|--------------|------|----------------------|
| <input checked="" type="checkbox"/> | 172.16.77.60 | CDH.aboutyun.com | 172.16.77.60 | 否 | ✓ 主机准备就绪：294 毫秒响应时间。 |
| <input checked="" type="checkbox"/> | 172.16.77.61 | CDH1.aboutyun.com | 172.16.77.61 | 否 | ✓ 主机准备就绪：8 毫秒响应时间。 |
| <input checked="" type="checkbox"/> | 172.16.77.62 | CDG2.aboutyun.com | 172.16.77.62 | 否 | ✓ 主机准备就绪：96 毫秒响应时间。 |

返回

单击继续

继续

都通信完好，我们单击继续

选择存储库

都是采用的默认设置

172.16.77.60:7180/cm/express-wizard/wizard

选择存储库

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

选择 CDH 的版本

☒ CDH-5.13-1.cdh5.13.p0.12

☐ CDH-4.7.0-1.cdh4.7.0.p0.40

其他 Parcel

☐ ACCUMULO-1.6.0-1.cdh5.1.0.p0.51

☐ ACCUMULO-1.4.4-1.cdh4.5.0.p0.65

☒ 无

☐ SQOOP_NETEZZA_CONNECTOR-1.2c5

☒ 无

☐ SQOOP_TERADATA_CONNECTOR-1.2c5

☒ 无

选择您要安装在主机上的 Cloudera Manager Agent 特定发行版。

返回

123456

继续

配置 Java 加密

172.16.77.60:7180/cm/express-wizard/wizard#step=javaOptionsStep

Cloudera Manager

支持

群集安装

配置 Java 加密

☒ Install Oracle Java SE Development Kit (JDK)
Uncheck this box to use a currently installed JDK

☐ 安装 Java 无限制强度加密政策文件
如果当地法律允许您部署无限强度加密，并且您正在运行安全的群集，则选中该复选框以安装无限强度加密政策文件。

返回

123456

继续

采用的是默认形式

提供 ssh 登录凭据

172.16.77.60:7180/cmf/express-wizard/wizard#step=hostCredentialsStep

群集安装

提供 SSH 登录凭据。

安装 Cloudera 包需要有主机的 root 访问权限。此安装程序将通过 SSH 连接到您的主机，然后直接以 root 用户身份登录，或者以另一个具有变为 root 用户的无密码 sudo/pbrun 权限的用户身份登录。

登录到所有主机，作为：

root

其他用户

对以上选定的用户，您可通过密码或公钥身份验证连接。

身份验证方法：

所有主机接受相同密码

所有主机接受相同私钥

输入密码：

...

确认密码：

...|

SSH 端口：

22

同时安装的数量：

10

（同时运行多个安装时将耗费大量的网络带宽和其他系统资源）

返回

123456

继续

这里选用的是 root，并且接受相同的密码，密码都为 123

注释：

指定主机的 SSH 登录方式

一种通过 root 用户，密码需要一致，刚开始安装建议使用 root。

也可以使用非 root 用户，但是需要保证 sudo 无密码访问

实现方法如下

1. 给 aboutyun 用户设置无密码 sudo 权限：
2. `chmod u+w /etc/sudoers`
3. `aboutyun ALL=(root)NOPASSWD:ALL`
4. `chmod u-w /etc/sudoers`
5. 测试：`sudo ifconfig`

来自[创建 sudo 无密码登陆](#)

开始安装

群集安装

正在安装。

已成功完成 3 个主机中的 0 个。

[中止安装](#)

| 主机名称 | IP 地址 | 进度 | 状态 | |
|-------------------|--------------|-------------|-------------|----------------------|
| CDG2.aboutyun.com | 172.16.77.62 | <div></div> | 正在复制安装文件... | 详细信息 |
| CDH.aboutyun.com | 172.16.77.60 | <div></div> | 正在进行身份验证... | 详细信息 |
| CDH1.aboutyun.com | 172.16.77.61 | <div></div> | 正在创建临时目录... | 详细信息 |

[返回](#)

1 2 3 4 5

[继续](#)

如遇到问题 无法检测到 Agent 发出的检测信号，可参考：[Cloudera Manager5 在线 bin 安装遇到 无法检测到 Agent 发出的检测信号 总结](#)

172.16.77.60:7180/cmf/express-wizard/resume

Cloudera manager

支持

群集安装

已成功完成安装。

已成功完成 3 个主机中的 3 个。

| 主机名称 | IP 地址 | 进度 | 状态 | |
|-------------------|--------------|-------------|----------|----------------------|
| CDG2.aboutyun.com | 172.16.77.62 | <div></div> | 已成功完成安装。 | 详细信息 |
| CDH.aboutyun.com | 172.16.77.60 | <div></div> | 已成功完成安装。 | 详细信息 |
| CDH1.aboutyun.com | 172.16.77.61 | <div></div> | 已成功完成安装。 | 详细信息 |

[返回](#)

1 2 3 4 5

[继续](#)

这样就通信成功了，我们单击继续

下载 CDH

图 1

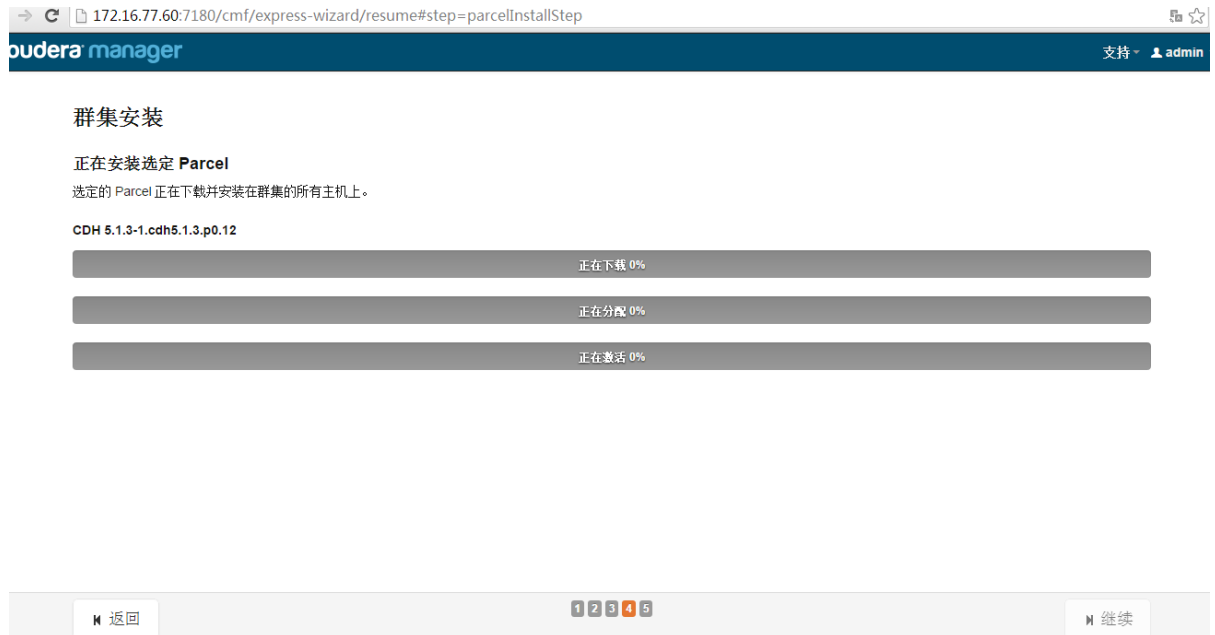


图 2





主机正确性检测

检查主机正确性

重新运行

验证

✓ 检查器在所有 3 个主机上运行。

✓ 检查主机名称时发现以下失败结果...

✓ 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。

✓ 检查过的所有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。

✓ 所有主机都将 /proc/sys/vm/swappiness 设置为 0。

✓ 没有任何性能与“透明大页面”设置有关。

✓ 已满足 CDH 5 Hue Python 版本依赖关系。

✓ 0 台主机正在运行 CDH 4，3 台主机正在运行 CDH5。

✓ 每个群集中检查过的所有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 Services 版本与服务器一致。

返回

1 2 3

完成

这里可能会遇到如下问题：

Cloudera 建议将 `/proc/sys/vm/swappiness` 设置为 0。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](#)，但可能会遇到问题，Cloudera Manager 报告您的[主机](#)由于交换运行状况不佳。以下主机受到影响：

解决办法：执行下面即可

```
1. echo 0 > /proc/sys/vm/swappiness
```

选择服务

由于内存紧张，所以这里安装了最少的服务

群集设置

选择您要在群集上安装的 CDH 5 服务。

选择要安装的服务组合。

☒ 核心 Hadoop

HDFS、YARN（包括 MapReduce 2）、ZooKeeper、Oozie、Hive、Hue 和 Sqoop

☐ 含 HBase 的内核

HDFS、YARN（包括 MapReduce 2）、ZooKeeper、Oozie、Hive、Hue、Sqoop 和 HBase

☐ 含 Impala 的内核

HDFS、YARN（包括 MapReduce 2）、ZooKeeper、Oozie、Hive、Hue、Sqoop 和 Impala

☐ 含 Search 的内核

HDFS、YARN（包括 MapReduce 2）、ZooKeeper、Oozie、Hive、Hue、Sqoop 和 Solr

☐ 含 Spark 的内核

HDFS、YARN（包括 MapReduce 2）、ZooKeeper、Oozie、Hive、Hue、Sqoop 和 Spark

☐ 所有服务

HDFS、YARN（包括 MapReduce 2）、ZooKeeper、Oozie、Hive、Hue、Sqoop、HBase、Impala、Solr、Spark 和键/值 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 Fume。

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。☐ 包括 Cloudera Navigator

单击继续即可



数据库设置：

这里使用默认设置，然后单击测试连接即可

群集设置

数据库设置

配置和测试数据库连接。如果使用自定义数据库，请先依照[Installation Guide](#) 的Installing and Configuring an External Database小节创建数据库。

使用自定义数据库

使用嵌入式数据库

当使用嵌入式数据库时，将会自动生成密码。请将它们复制下来。

Hive

数据库主机名称:

CDH.aboutyun.com:7432

数据库类型:

PostgreSQL

数据库名称:

hive

用户名:

hive

密码:

HjBM3G3rlq

Activity Monitor

当前被分配在 CDH.aboutyun.com 上运行。

数据库主机名称:

CDH.aboutyun.com:7432

数据库类型:

PostgreSQL

数据库名称:

amon

用户名:

amon

密码:

90yvjtmmnh

Reports Manager

当前被分配在 CDH.aboutyun.com 上运行。

数据库主机名称:

CDH.aboutyun.com:7432

数据库类型:

PostgreSQL

数据库名称:

rman

用户名:

rman

密码:

Vc2fONW7r9

测试连接



审核更改

群集设置

审核更改

| 为新角色设置下列配置值。必填值以 * 标出。 | | | |
|--|---|--|---|
| 参数 | 组 | 值 | 说明 |
| 服务 Cloudera Management Service | | | |
| Service Monitor 存储目录* firehose.storage.base.directory | Service Monitor Default Group 显示成员 | /var/lib/cloudera-service-monitor 默认值 | Service Monitor 数据的存储目录。Service Monitor 存储度量时间序列和运行状况信息，以及 Impala 查询和 YARN 应用程序元数据（如果已配置 Impala 和或 YARN）。 |
| Host Monitor 存储目录* firehose.storage.base.directory | Host Monitor Default Group 显示成员 | /var/lib/cloudera-host-monitor 默认值 | Host Monitor 数据的存储目录。Host Monitor 存储度量时间序列和运行状况信息。 |
| 警报：邮件服务器主机名称* | Alert Publisher Default Group 显示成员 | localhost 默认值 | 接收警报的邮件服务器的 IP 地址或主机名称 |
| 警报：邮件服务器用户名 | Alert Publisher Default Group 显示成员 | 默认为空。单击以编辑。 | 用于登录邮件服务器的用户名 |
| 警报：邮件服务器密码 | Alert Publisher Default Group 显示成员 | 默认为空。单击以编辑。 | 用于登录邮件服务器的密码。警告：此密码将通过网络以明文形式发送至 Alert Publisher 主机。另外，密码将存储在 Alert Publisher 主机的纯文本文件中，具有受限文件系统权限。 |
| 警报：邮件收件人* | Alert Publisher Default Group 显示成员 | root@localhost 默认值 | 用于接收警报的电子邮件地址的逗号分隔列表 |
| 服务 HDFS | | | |

同样这里采用的是默认设置



创建服务

这里是一个关键点，很多人会遇到错误。
并且有的服务创建成功，有的创建失败，于是我们点击重试，那么对于成功的服务，他就不会再次出现，这就是为什么，加入有 19 个服务，在此重试的时候，你可能会看到 2 个服务。

在这里遇到的错误是内存不够用，所以我们 server 最好是 5 个 G，另外两个节点是 3G，至少 2G。我这里的内存分布是：有一个 master,三个 datanode.
server 节点有 namenode 和 datanode 两个节点。
因为内存少，所以就遇到了下面问题：

1.

Command aborted because of exception: Command timed-out after 90 seconds
2.

程序: hdfs/hdfs.sh ["mkdir","/tmp","hdfs","supergroup","1777"]

Cluster 1 » HDFS » 命令 »

^ 父命令: 首次运行 ()

命令详细信息: 创建 /tmp 目录

| 命令 | 上下文 | 状态 | 开始日期 |
|--|--|-----|-----------------------|
|  创建 /tmp 目录 |  NameNode_cdh | 已中止 | 2014-9-25 2:13:29 PDT |

Command aborted because of exception: Command timed-out after 90 seconds

程序: hdfs/hdfs.sh ["mkdir", "/tmp", "hdfs", "supergroup", "1777"]

» 最近的日志条目

完整日志的链接: [Stderr](#) [Stdout](#) [角色日志详细信息](#)

上面是命令超时。

解决办法：

再次重试，相信能过的。

假如这里你遇到了其它的问题，那么可以通过上图中的角色日志查看问题原因。

由于上面两个服务安装失败，所以在此点击重试的时候，剩下两个服务，创建成功

群集设置

进度

| 命令 | 上下文 | 状态 | 开始日期 | 结束于 |
|--------|-----|-----|-----------------------|-----------------------|
| ✓ 首次运行 | | 已完成 | 2014-9-25 2:58:43 PDT | 2014-9-25 3:00:04 PDT |

Finished First Run of all services successfully.

命令进度

已完成 2 个步骤（共 2 个）。

✓ 正在创建 HDFS /tmp 目录

HDFS directory /tmp already exists.

[详细信息](#)

✓ 正在部署客户端配置

Successfully deployed all client configurations.

[详细信息](#)



最后集群创建成功

群集设置

恭喜您！

服务已安装、配置并在群集中运行。

◀ 返回

1 2 3 4 5 6

▶ 完成



指定主机开始可能会碰到的问题

注意：问题 1

这里安装的时候，相信你会碰到下面问题

```
Detecting Cloudera Manager Server...
Detecting Cloudera Manager Server...
BEGIN host -t PTR 192.168.1.198
198.1.168.192.in-addr.arpa domain name pointer localhost.
END (0)
using localhost as scm server hostname
BEGIN which python
/usr/bin/python
```

```
END (0)
BEGIN python -c 'import socket; import sys; s = socket.socket(socket.AF_INET); s.settimeout(5.0);
s.connect((sys.argv[1], int(sys.argv[2]))); s.close();' localhost 7182
Traceback (most recent call last):
File "<string>", line 1, in <module>
File "<string>", line 1, in connect
socket.error: [Errno 111] Connection refused
END (1)
could not contact scm server at localhost:7182, giving up
waiting for rollback request
```

解决办法:

```
1. mv /usr/bin/host /usr/bin/host.bak
```

问题 2:

在这个过程中还发生些问题, 就是安装完毕和 cloudera-scm-agent 通信异常, 不能检测到心跳。

Cluster Installation

Installation in progress.

0 of 3 host(s) completed successfully. [Abort Installation](#)

Uninstalled on 1 host(s) after installation failure. [Retry Failed Hosts](#)

| Hostname | IP Address | Progress | Status |
|--------------|--------------|-------------|--|
| 172.16.77.60 | 172.16.77.60 | <div></div> | <div>✗ Installation failed. Failed to receive heartbeat from agent. Retry Details</div> <ul style="list-style-type: none">• Ensure that the host's hostname is configured properly.• Ensure that port 7182 is accessible on the Cloudera Manager Server (check firewall rules).• Ensure that ports 9000 and 9001 are free on the host being added.• Check agent logs in /var/log/cloudera-scm-agent/ on the host being added (some of the logs can be found in the installation details). |
| 172.16.77.61 | 172.16.77.61 | <div></div> | <div>⚙ Installing cloudera-manager-daemons package... Details</div> |
| 172.16.77.62 | 172.16.77.62 | <div></div> | <div>⚙ Installing oracle-j2sdk1.7 package... Details</div> |

[Back](#) 1 2 3 4 5 6 [Continue](#)

具体解决办法, 参考: [Cloudera Manager5 在线 bin 安装遇到 No socket could be created on\[Errno 99\] 总结](#)

问题 3: 集群安装意外终止, 该如何恢复

图 1

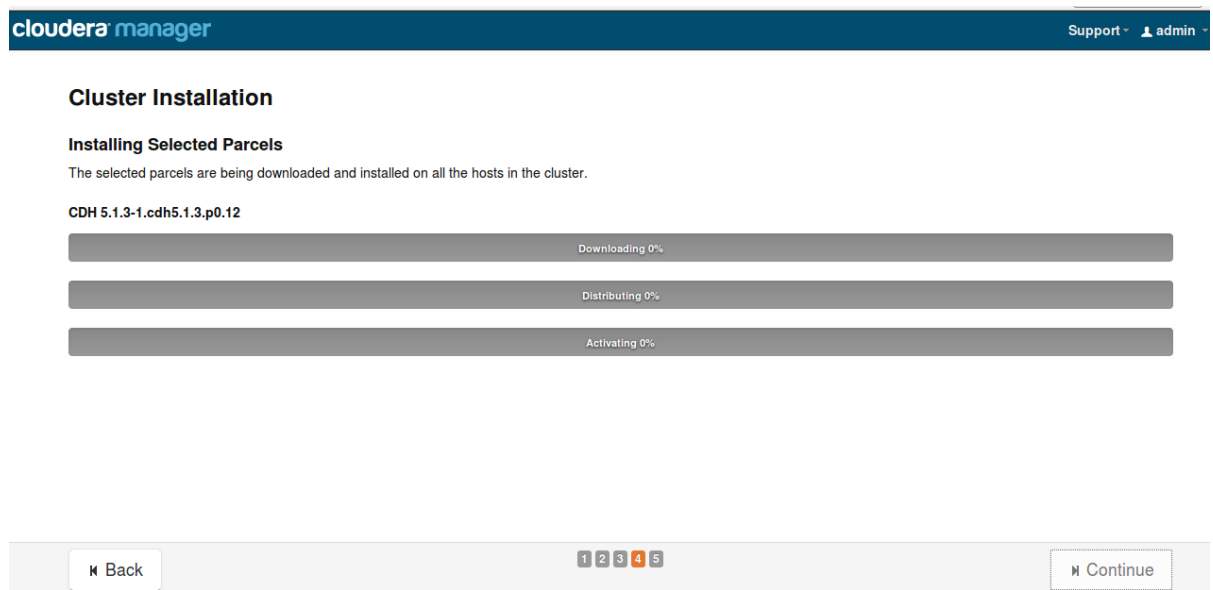


图 2



这是一个很漫长的过程，不幸的是在下载完毕，执行分发的时候，这个过程被打断，安装被终止。导致回不到图 2，这个该如何解决？

难道真的要重装吗？花费了大半天功夫，又不得不重来。

这里面的问题是找不到 cloudera manager5 所管理的节点了？

about 云有这篇文章[卸载 Cloudera Manager 5.1.x.和 相关软件【官网翻译：高可用】](#)

但是这篇文章是卸载 Cloudera Manager 以及 CDH 的，由于 cloudera-scm-server 和 cloudera-scm-agent 查看都

是运行正常的，如何查看状态，可以参考：

[Cloudera Manager Server5 及 Cloudera Manager Agents5 命令整理（about 云）](#)，所以 cloudera manager 不需要卸载，卸载的是 CDH 的相关内容。

于是执行下面命令：

```
1. sudo apt-get remove avro-tools crunch flume-ng hadoop-hdfs-fuse hadoop-hdfs-nfs3 hadoop-httpfs
hbase-solr hive-hbase hive-webhcat hue-beeswax hue-hbase hue-impala hue-pig hue-plugins
hue-rdbms hue-search hue-spark hue-sqoop hue-zookeeper impala impala-shell kite llama mahout
oozie pig pig-udf-datafu search sentry solr-mapreduce spark-python sqoop sqoop2 whirr

1. sudo apt-get clean

1. sudo rm -Rf /var/lib/flume-ng /var/lib/hadoop* /var/lib/hue /var/lib/navigator /var/lib/oozie
/var/lib/solr /var/lib/sqoop* /var/lib/zookeeper

1. sudo rm -Rf /dfs /mapred /yarn
```

进入目录：

删掉/var/lib/cloudera-scm-agent 里面的

response.avro uuid

```
root@CDH:/var/lib/cloudera-scm-agent# ls
response.avro  uuid
root@CDH:/var/lib/cloudera-scm-agent# rm *
root@CDH:/var/lib/cloudera-scm-agent# service cloudera-scm-agent restart
Stopping cloudera-scm-agent: * cloudera-scm-agent stopped
Starting cloudera-scm-agent: * cloudera-scm-agent started
```

最后终于找回

为 CDH 群集安装指定主机。

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。

| <input type="checkbox"/> | 名称 | IP | 机架 | CDH 版本 | 状态 | 上一检测信号 |
|--------------------------|-----------------------------------|------------------------------------|-----------------------------------|---------------|---------------|---------------|
| | <input type="text" value="任何名称"/> | <input type="text" value="任何 IP"/> | <input type="text" value="任何机架"/> | <div>全部</div> | <div>全部</div> | <div>全部</div> |
| <input type="checkbox"/> | CDG2.aboutyun.com | 172.16.77.62 | /default | 无 | ⑦ 未知运行状况 | 12.35s ago |
| <input type="checkbox"/> | CDH.aboutyun.com | 172.16.77.60 | /default | 无 | ⑦ 未知运行状况 | 5.61s ago |
| <input type="checkbox"/> | CDH1.aboutyun.com | 172.16.77.61 | /default | 无 | ⑦ 未知运行状况 | 11.87s ago |

返回

继续

通过上面终于找到所管理的三个节点。然后从新登录，选择三个主机，然后继续继续，最后终于进入了这个界面。
总算没有重新安装

如果还未找回，则进入下面目录：删掉/var/lib/cloudera-scm-agent 里面的
response.avro uuid

```
1. /var/lib/cloudera-scm-agent# rm *
```

找回

如果出现了下面情况，说明 parcel 错误，如果遇到这样的情况不能解决，就只能重装了。

```
1. CDH 5.1.3-1.cdh5.1.3.p0.12 1 个错误

2. parcel CDH-5.1.3-1.cdh5.1.3.p0.12-unknown 的本地 parcel 错误：一个或多个主机没有报告其操作系统分

配。这可能是由于运行已过期的 Cloudera Manager Agent 版本造成的。请运行 Host Inspector 检查 Agent 版

本。。
```


Cloudera Manager Server5 及 Cloudera Manager Agents5 命令整理（about 云）

问题导读

- 1.使用什么命令可以开启、停止、重启 **Cloudera Manager Server**
- 2.如何强制重启 **Agents**
- 3.什么情况下适合强制重启 **Agents**
- 4.如何查看 **Agents** 状态



1.Cloudera Manager Server 命令

开启, 停止, 和 重启 Cloudera Manager Server

1.1 开启

```
1. $ sudo service cloudera-scm-server start
```

1.2 停止

```
1. $ sudo service cloudera-scm-server stop
```

1.3 重启

```
1. $ sudo service cloudera-scm-server restart
```



2.Cloudera Manager Agents 命令

2.1 开启

```
1. $ sudo service cloudera-scm-agent start
```

2.2 清理开启

```
1. $ sudo service cloudera-scm-agent clean_start
```

彻底清除目录 `/var/run/cloudera-scm-agent`；删除所有的文件和子目录，然后执行命令

`/var/run/cloudera-scm-agent` 包含磁盘上运行的 Agent 的状态。

一些 Agent 的状态留在 `/var/lib/cloudera-scm-agent`，不应删除

更多信息，查看 [Server and Client Configuration](#) 和 [Process Management](#)。

2.3 停止

```
1. $ sudo service cloudera-scm-agent stop
```

2.4 重启

```
1. $ sudo service cloudera-scm-agent restart
```

3.强制停止和重启 Agents 命令

注意： `hard_stop`, `clean_restart`, or `hard_restart` 命令杀掉主机上运行的服务进程

3.1 强制停止

```
1. $ sudo service cloudera-scm-agent hard_stop
```

3.2 强制重启

```
1. $ sudo service cloudera-scm-agent hard_restart
```

强制重启适用于下面情况：

- 1.Cloudera Manager 升级，supervisord 代码在当前版本与新版本之间发生变化，为了保证的升级，你也需要重启 supervisor
- 2.supervisord 挂起，需要重新启动
- 3.清除与 Cloudera Manager 有关的运行状态及服务

3.3 清理重启

```
1. $ sudo service cloudera-scm-agent clean_restart
```

先执行 clean_start, 然后执行 hard_stop



4.检查 Agents 状态

检查 Agents 进程

```
1. $ sudo service cloudera-scm-agent status
```

出自 [about 云](#)

本文链接 <http://www.aboutyun.com/thread-9096-1-1.html>

错误处理

卸载 Cloudera Manager 5.1.x.和 相关软件

【官网翻译：高可用

问题导读：

1.不同的安装方式，卸载方法存在什么区别？

2.不同的操作系统，卸载 Cloudera Manager Server and 数据库有什么区别？



重新安装不完整

如果你来到这里，因为你的安装没有完成（例如，如果它是由一个虚拟机超时中断），和你想继续安装，然后重新安装，做到以下几点：

删除文件和目录：

个人建议：

这一步虽然是官网这么做，但是下面命令还是不要执行，因为我们后面还会用到下面命令

```
1. /usr/share/cmf/uninstall-cloudera-manager.sh
```

当然如果你找不到这个 `uninstall-cloudera-manager.sh` 文件，后面也有详细的说明，教你如何做，所以这一步（**1.**

移除文件和目录：）个人认为可以省略。

1.移除文件和目录：

```
1. $ sudo rm -Rf /usr/share/cmf /var/lib/cloudera* /var/cache/yum/cloudera*
```

2.卸载 Cloudera Manager 和管理软件

用户数据路径如下：

Cloudera Manager 数据

`/usr/share/cmf /var/lib/cloudera* /var/cache/yum/cloudera* /var/log/cloudera* /var/run/cloudera*`

数据库存放路径

/var/lib/cloudera-scm-server-db

Cloudera Manager Lock 文件

/tmp/.scm_prepare_node.lock

用户数据

/var/lib/flume-ng /var/lib/hadoop* /var/lib/hue /var/lib/navigator /var/lib/oozie /var/lib/solr /var/lib/sqoop*
/var/lib/zookeeper

/dfs /mapred /yarn

上面的文件都是 Cloudera Manager 的默认安装路径，如果你改变了路径，需要查看每个服务的配置文件

3.停止所有服务


1. 通过 Cloudera Manager 停止集群

- 在主页中，单击  按钮，停止集群
- 单击确认。然后看到服务停止的进度，当停止完毕，任务完成，可以关闭命令停止窗口。
- 进入主页，单击 停止 Cloudera Management Service，关闭命令停止窗口。

2. 停止 Cloudera Manager 服务

角色要求: **Administrator**

1.按照下面步骤

- 选择 **Clusters > Cloudera Management Service > Cloudera Management Service.**
- 选择  Actions > Stop.
- 在主页单击按钮，停止 Cloudera Management Service

2.确定停止，然后会显示停止进度

- 当指令完成 **N / N**，出现成功的命令，任务完成。单击“关闭”。

4.停用和删除的 Parcels

如果你是通过 packages 安装，跳过这步，然后到第五步（卸载 Cloudera Manager Server），然后移除 packages 在第六步（卸载 Cloudera Manager Agent and 相关软件）。如果安装使用 parcels，那么按照下面步骤：

- 单击导航栏 parcel  图标

2. 每一个激活的 parcel，选择 Actions > Deactivate。操作完成，parcel 按钮变成 Activate
3. 每一个激活的 parcel，选择 Actions > Remove from Hosts 操作完成，parcel 按钮变成 **Distribute**
4. 每一个激活的 parcel，选择 **Actions > Delete** 本地库将移除 parcel

可能 parcels 包下载和分布在多处，但是没有被激活，如果这样的话，应该从本地库及其它客户端移除这些 parcels

5. 卸载 Cloudera Manager Server

卸载 Cloudera Manager Server 的方法，取决于你的安装方式，参照下面来选择你的卸载方式

如果你是使用的 **cloudera-manager-installer.bin** 文件，在 Cloudera Manager Server 主机上使用下面命令

```
1. $ sudo /usr/share/cmf/uninstall-cloudera-manager.sh
```

需要注意的是：

如果 uninstall-cloudera-manager.sh 没有安装到集群上，用下面方法卸载 Cloudera Manager Server .

如果没有使用 **cloudera-manager-installer.bin** 文件，如果你安装使用使用其它方法，比如 Puppet, 在 Cloudera Manager Server 主机上，使用下面命令卸载

1. 停止 Cloudera Manager Server and 它的数据库

```
1. sudo service cloudera-scm-server stop
2. sudo service cloudera-scm-server-db stop
```

2. 卸载 Cloudera Manager Server and 数据库，这部描述了移除嵌入式数据库 PostgreSQL，如果你没有使用嵌入式数据库，可以忽略

```
1. cloudera-manager-server-db
```

这步。

Red Hat 系统:

```
1. sudo yum remove cloudera-manager-server
2. sudo yum remove cloudera-manager-server-db-2
```

SLES 系统

1. `sudo zypper -n rm --force-resolution cloudera-manager-server`
2. `sudo zypper -n rm --force-resolution cloudera-manager-server-db-2`

Debian/Ubuntu 系统:

1. `sudo apt-get remove cloudera-manager-server`
2. `sudo apt-get remove cloudera-manager-server-db-2`

6.卸载 Cloudera Manager Agent and 相关软件

在所有的 Agent hosts 上按照下面步骤:

1.停止 Cloudera Manager Agent.

Red Hat/SLES 系统:

1. `$ sudo service cloudera-scm-agent hard_stop`

Debian/Ubuntu 系统:

1. `$ sudo /usr/sbin/service cloudera-scm-agent hard_stop`

2.卸载软件

2.1Red Hat 系统

Parcel 安装:

```
1. $ sudo yum remove 'cloudera-manager-*
```

Package 安装

CDH 4

```
1. $ sudo yum remove 'cloudera-manager-*' bigtop-utils bigtop-jsvc bigtop-tomcat hadoop hadoop-hdfs  
hadoop-httpfs hadoop-mapreduce hadoop-yarn hadoop-client hadoop-0.20-mapreduce hue-plugins  
hbase hive oozie oozie-client pig zookeeper hue impala impala-shell solr-server
```

CDH 5

```
1. $ sudo yum remove 'cloudera-manager-*' avro-tools crunch flume-ng hadoop-hdfs-fuse  
hadoop-hdfs-nfs3 hadoop-httpfs hbase-solr hive-hbase hive-webhcat hue-beeswax hue-hbase  
hue-impala hue-pig hue-plugins hue-rdbms hue-search hue-spark hue-sqoop hue-zookeeper impala  
impala-shell kite llama mahout oozie pig pig-udf-datafu search sentry solr-mapreduce spark-python  
sqoop sqoop2 whirr
```

2.2SLES 系统

Parcel 安装

```
1. $ sudo zypper remove 'cloudera-manager-*
```

Package 安装

CDH 4

```
1. $ sudo zypper remove 'cloudera-manager-*' bigtop-utils bigtop-jsvc bigtop-tomcat hadoop  
hadoop-hdfs hadoop-httpfs hadoop-mapreduce hadoop-yarn hadoop-client hadoop-0.20-mapreduce  
hue-plugins hbase hive oozie oozie-client pig zookeeper hue impala impala-shell solr-server
```

CDH 5

```
1. $ sudo zypper remove 'cloudera-manager-*' avro-tools crunch flume-ng hadoop-hdfs-fuse  
hadoop-hdfs-nfs3 hadoop-httpfs hbase-solr hive-hbase hive-webhcat hue-beeswax hue-hbase  
hue-impala hue-pig hue-plugins hue-rdbms hue-search hue-spark hue-sqoop hue-zookeeper impala  
impala-shell kite llama mahout oozie pig pig-udf-datafu search sentry solr-mapreduce spark-python  
sqoop sqoop2 whirr
```

2.3Debian/Ubuntu 系统

Parcel 安装

```
1. $ sudo apt-get purge 'cloudera-manager-*'
```

Package 安装

CDH 4

```
1. $ sudo apt-get purge 'cloudera-manager-*' bigtop-utils bigtop-jsvc bigtop-tomcat hadoop  
hadoop-hdfs hadoop-httpfs hadoop-mapreduce hadoop-yarn hadoop-client hadoop-0.20-mapreduce  
hue-plugins hbase hive oozie oozie-client pig zookeeper hue impala impala-shell solr-server
```

CDH 5

```
1. $ sudo apt-get purge 'cloudera-manager-*' avro-tools crunch flume-ng hadoop-hdfs-fuse  
hadoop-hdfs-nfs3 hadoop-httpfs hbase-solr hive-hbase hive-webhcat hue-beeswax hue-hbase  
hue-impala hue-pig hue-plugins hue-rdbms hue-search hue-spark hue-sqoop hue-zookeeper impala  
impala-shell kite llama mahout oozie pig pig-udf-datafu search sentry solr-mapreduce spark-python  
sqoop sqoop2 whirr
```

3. 卸载软件

Red Hat

```
1. $ sudo yum clean all
```

SLES

```
1. $ sudo zypper clean
```

Debian/Ubuntu

```
1. $ sudo apt-get clean
```

移除 Cloudera Manager 和用户数据

kill Cloudera Manager and Managed 进程

在所有 Agent hosts，杀掉运行着的 Cloudera Manager and managed 进程：

```
1. $ for u in cloudera-scm flume hadoop hdfs hbase hive httpfs hue impala llama mapred oozie solr  
    spark sqoop sqoop2 yarn zookeeper; do sudo kill $(ps -u $u -o pid=); done
```

注意：这一步不是必须的，如果你已经正确停止了所有的服务和 Cloudera Manager Agent

移除 Cloudera Manager 数据

这一步是永久的移除数据，如果你以后还使用的的话，可以先备份。如果你使用了嵌入式数据库，数据存储在

```
1. /var/lib/cloudera-scm-server-db
```

在所有的 Agent hosts 运行下面命令：

```
1. $ sudo rm -Rf /usr/share/cmf /var/lib/cloudera* /var/cache/yum/cloudera* /var/log/cloudera*  
    /var/run/cloudera*
```

这里官网忘记了，对于 **ubuntu**，是在 **apt** 文件夹中，还需要执行下面命令：

```
1. $ sudo rm -Rf /var/cache/apt/archives/cloudera*
```

移除 Cloudera Manager Lock 文件

在 Agent hosts 上，运行下面命令移除 Cloudera Manager 锁文件：

```
1. $ sudo rm /tmp/.scm_prepare_node.lock
```

移除用户数据

这一步是永久的移除数据.为了保护数据,可以使用 `distcp` 命令复制到另外一个集群。在所有 Agent hosts 上运行下面命令。

```
1. $ sudo rm -Rf /var/lib/flume-ng /var/lib/hadoop* /var/lib/hue /var/lib/navigator /var/lib/oozie  
/var/lib/solr /var/lib/sqoop* /var/lib/zookeeper
```

```
1. $ sudo rm -Rf /dfs /mapred /yarn
```

说明: 关于卸载 CDH, 包括清理 CDH 文件,可以查看 [CDH4 Installation Guide](#) or [CDH 5 Installation Guide](#).

本文链接:

<http://www.aboutyun.com/thread-8992-1-1.html>

相关文章推荐: [卸载 Cloudera CDH 5 beta2 总结](#)

[使用 cloudera-manager-installer.bin\(Parcles\)安装失败后卸载 cloudera](#)

翻译:

http://www.cloudera.com/content/cloudera-content/cloudera-docs/CM5/latest/Cloudera-Manager-Installation-Guide/cm5iq_uninstall_cm.html

解析 Cloudera Manager 内部结构、功能包括配置文件、目录位置等

问题导读

1.CM 的安装目录在什么位置?

2.hadoop 配置文件在什么位置?

3.Cloudera manager 运行所需要的信息存在什么位置?

4.CM 结构和功能是什么?



1. 相关目录

- /var/log/cloudera-scm-installer : 安装日志目录。
- /var/log/* : 相关日志文件 (相关服务的及 CM 的)。
- /usr/share/cmf/ : 程序安装目录。
- /usr/lib64/cmf/ : Agent 程序代码。
- /var/lib/cloudera-scm-server-db/data : 内嵌数据库目录。
- /usr/bin/postgres : 内嵌数据库程序。
- /etc/cloudera-scm-agent/ : agent 的配置目录。
- /etc/cloudera-scm-server/ : server 的配置目录。
- /opt/cloudera/parcels/ : Hadoop 相关服务安装目录。
- /opt/cloudera/parcel-repo/ : 下载的服务软件包数据, 数据格式为 parcels。
- /opt/cloudera/parcel-cache/ : 下载的服务软件包缓存数据。
- /etc/hadoop/* : 客户端配置文件目录。

2. 配置

- Hadoop 配置文件

配置文件放置于 /var/run/cloudera-scm-agent/process/ 目录下。如:

/var/run/cloudera-scm-agent/process/193-hdfs-NAMENODE/core-site.xml。这些配置文件是通过 Cloudera Manager 启动相应服务 (如 HDFS) 时生成的, 内容从数据库中获得 (即通过界面配置的参数)。

在 CM 界面上更改配置是不会立即反映到配置文件中, 这些信息会存储于数据库中, 等下次重启服务时才会生成配置文件。且每次启动时都会产生新的配置文件。

CM Server 主要数据库为 scm 基中放置配置的数据表为 configs。里面包含了服务的配置信息, 每一次配置的更改会把当前页面的所有配置内容添加到数据库中, 以此保存配置修改历史。

scm 数据库被配置成只能从 localhost 访问, 如果需要从外部连接此数据库, 修改 vim /var/lib/cloudera-scm-server-db/data/pg_hba.conf 文件, 之后重启数据库。运行数据库的用户为 cloudera-scm。

- 查看配置内容

1. 直接查询 scm 数据库的 configs 数据表的内容。
2. 访问 REST API: <http://hostname:7180/api/v4/cm/deployment>, 返回 JSON 格式部署配置信息。

- 配置生成方式

CM 为每个服务进程生成独立的配置目录（文件）。所有配置统一在服务端查询数据库生成（因为 scm 数据库只能在 localhost 下访问）生成配置文件，再由 agent 通过网络下载包含配置文件的 zip 包到本地解压到指定的目录。

- 配置修改

CM 对于需要修改的配置预先定义，对于没有预先定义的配置,则通过在高级配置项中使用 xml 配置片段的方式进行配置。而对于/etc/hadoop/下的配置文件是客户端的配置，可以在 CM 通过部署客户端生成客户端配置。

3. 数据库

Cloudera manager 主要的数据库为 scm,存储 Cloudera manager 运行所需要的信息：配置，主机，用户等。

4. CM 结构

CM 分为 Server 与 Agent 两部分及数据库（自带更改过的嵌入 Postgresql）。它主要做三件事：

1. 管理监控集群主机。
2. 统一管理配置。
3. 管理维护 Hadoop 平台系统。

实现采用 C/S 结构，Agent 为客户端负责执行服务端发来的命令，执行方式一般为使用 python 调用相应的服务 shell 脚本。Server 端为 Java REST 服务，提供 REST API，Web 管理端通过 REST API 调用 Server 端功能，Web 界面使用富客户端技术（Knockout）。

1. Server 端主体使用 Java 实现。
2. Agent 端主体使用 Python, 服务的启动通过调用相应的 shell 脚本进行启动，如果启动失败会重复 4 次调用启动脚本。
3. Agent 与 Server 保持心跳，使用 Thrift RPC 框架。

5. 升级

在 CM 中可以通过界面向导升级相关服务。升级过程为三步：

1. 下载服务软件包。

2. 把所下载的服务软件包分发到集群中受管的机器上。
3. 安装服务软件包，使用软链接的方式把服务程序目录链接到新安装的软件包目录上。

6. 卸载

sudo /usr/share/cmf/uninstall-scm-express.sh, 然后删除/var/lib/cloudera-scm-server-db/目录, 不然下次安装可能不成功。

7. 开启 postgresql 远程访问

CM 内嵌数据库被配置成只能从 localhost 访问, 如果需要从外部查看数据, 数据修改 vim /var/lib/cloudera-scm-server-db/data/pg_hba.conf 文件, 之后重启数据库。运行数据库的用户为 cloudera-scm。

文章会随时更新, 如有错误, 欢迎批评指正, 在此基础上会对其扩展录制视频, 如觉得文档有帮助, 亦可捐助 [about 云](#)

如果想关注 about 云, 可以通过下面方式:

1. 加入 about 云群 39327136、371358502、322273151、90371779, 云计算大数据爱好者群, 如群满, 请搜 **about 云**



2.关注腾讯认证空间

[about 云腾讯认证空间](#)

3.惯用手机的用户可关注 **about** 云微信地址:

搜索:

wwwaboutyuncom



4.关注微博:

[新浪微博](#)

5.邮件订阅

[邮件订阅](#)

2014 过去, 2015 年来临, 2014 年, about 云也有一年多, 云技术、大数据迅速火了起来。

过去的一年中, 我们都尽其所能的为大家提供技术性实践资料、文章、视频。但是总的来说,

还是不够。我们将一如既往的, 出一些技术性文章。引导新手入门及学习, 同时希望大数据、

云技术爱好者多多支持 about 云，我们将会出一些视频，也欢迎捐助 about 云。

在 IT 行业，技术不断更新换代，不断抛陈出新，5 年不学习新知识，就会 out 了。about 云

创立之初，面对资料的缺乏，与大家共同学习，相互分享学习经验，帮助了很多初学者入门，

about 云在 2015 年将会更努力，成为更多人的大数据、云技术学习分享基地。

同时希望接触 about 云，知道 about 云，了解 about 云的人，about 云能传播给大家一种学

习精神----活到老，学到老。

2015 年，about 云将会不断的努力、坚持，帮助更多的人，更多的大数据、云技术爱好者。