

# CDH5.11.2 安装

## 目录

---

提示.....	2
Redhat 下载和安装.....	2
Redhat7.2 下载.....	2
redhat7.2 的安装步骤.....	4
配置 Linux.....	20
CM、CDH、KAFKA 下载.....	27
下载 CM.....	27
下载 CDH.....	28
下载 KAFKA.....	29
安装 CM、CDH、KAFKA.....	31
配置 SSH 免密码登陆.....	31
配置 NTP.....	32
配置 CM 的 yum 安装源.....	33
安装 Oracle JDK.....	34
安装和配置 mariadb.....	35

安装 CM 和 CDH .....	41
添加 KAFKA 服务 .....	56

## 提示

---

以下安装步骤均在一台机器上安装，仅作为演示步骤所用。在多台机器安装时，部分操作只需要在主节点上完成，无需在所有节点上操作，而一部分操作需要在所有机器上进行操作，所以请格外留意图文中的红色字体以及红色方框标注的内容，实际安装时 IP 地址以及其他环境因素会有变化，请勿照搬照抄文中的部分变量。

所有安装步骤均在 root 用户下完成

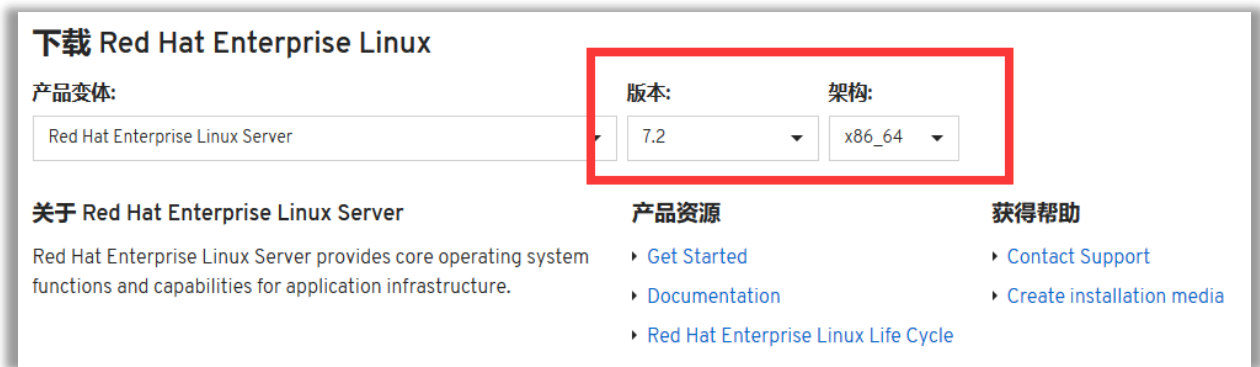
## Redhat 下载和安装

---

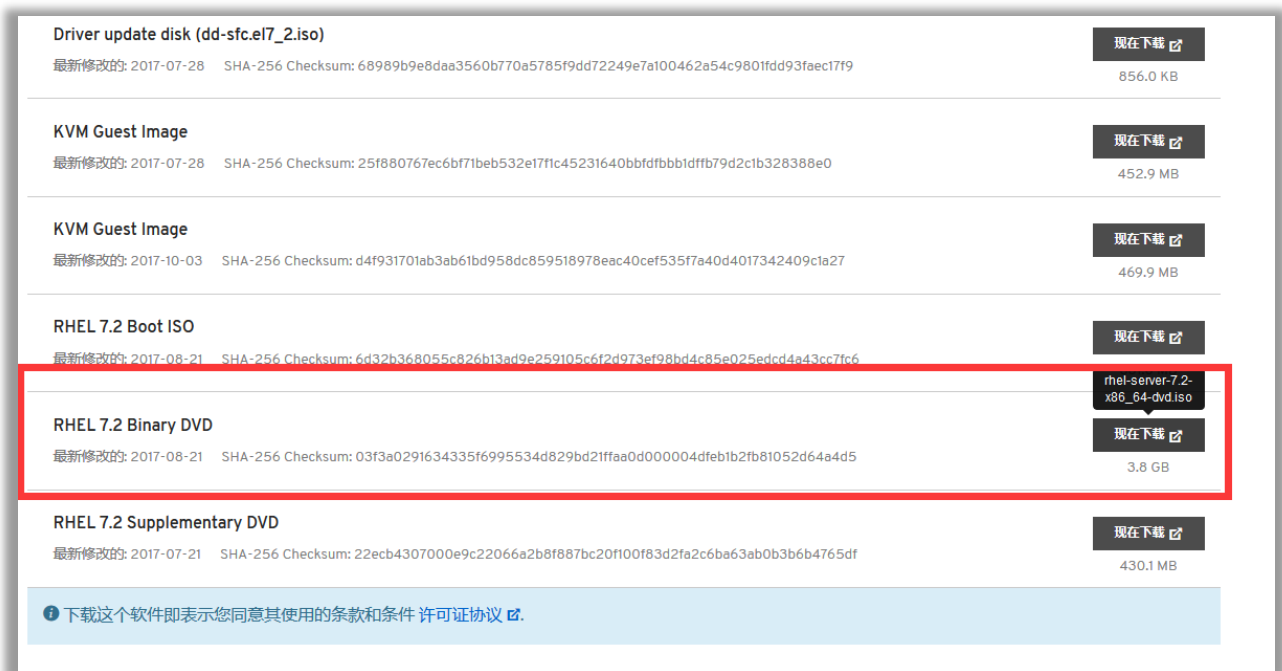
### Redhat7.2 下载

版本：redhat7.2

下载页面：[https://access.redhat.com/downloads/content/69/ver=/rhel---7/7.2/x86\\_64/product-software](https://access.redhat.com/downloads/content/69/ver=/rhel---7/7.2/x86_64/product-software) (打开页面需要登录 redhat 账号) 如图



在下载页面中选择镜像文件下载地址（如图）：



如果上述地址无法下载，可尝试以下方法：

地址：<https://developers.redhat.com/products/rhel/download/>

此方法也需要红帽账号。

Overview

**Download**

Hello World!

Docs and APIs

Help

Learn

Buzz

## Download for Development Use

**TRY IT**

To download this product, [you must have an account](#) and accept the terms and conditions of the [Red Hat Developer Program](#), which provides no-cost subscriptions for development use only.

[DOWNLOAD](#)

Red Hat Enterprise Linux 7.4.0

**ALL DOWNLOADS**

Version	Release Date	Description	Download
7.4.0	2017-08-02	Boot iso	<a href="#">Boot iso (456 MB)</a>
		DVD iso	<a href="#">DVD iso (4 GB)</a>
		RHEL	<a href="#">RHEL-Server-7.4</a>

[View Older Downloads](#)

**Other Developer Subscription Options:**

- Supported versions of Red Hat Enterprise Linux Developer Subscriptions are also available. [See this complete list](#) to choose from.
- If you're a Red Hat technology partner (e.g. an ISV), no-cost (Not for Resale - NFR) subscriptions are available by joining [Red Hat Connect for Technology Partners](#). Once there, register your company and join the "Zone" for Red Hat Enterprise Linux or Containers.

[View Older Downloads](#)

Version	Release Date	Description	Download
7.3.0	2016-11-03	Boot iso	<a href="#">Boot iso (408 MB)</a>
		DVD iso	<a href="#">DVD iso (4 GB)</a>
		RHEL	<a href="#">RHEL-Server-7.3</a>
7.2.0	2016-03-04	Boot iso	<a href="#">Boot iso (361 MB)</a>
		DVD iso	<a href="#">DVD Iso (4 GB)</a>
		RHEL	<a href="#">RHEL-Server-7.2</a>

下载完成后通过 iso 文件制作 U 盘或者光盘的安装介质。

[redhat7.2 的安装步骤](#)

安装步骤中机器的硬件配置仅供参考，以实际配置为准。

Red Hat Enterprise Linux 7.2

Install Red Hat Enterprise Linux 7.2  
Test this media & install Red Hat Enterprise Linux 7.2

Troubleshooting >

Press Tab for full configuration options on menu items.



## 欢迎使用 RED HAT ENTERPRISE LINUX 7.2。

您在安装过程中想使用哪种语言？

తెలుగు	Telugu	简体中文(中国)
Тоҷикӣ	Tajik	繁體中文(台灣)
ไทย	Thai	繁體中文(中華人民共和國香港特別行政區)
Türkçe	Turkish	简体中文(新加坡)
Українська	Ukrainian	
اردو	Urdu	
Tiếng Việt	Vietnamese	
中文	Chinese	
IsiZulu	Zulu	

在这里进行搜索。

退出(Q)

继续(C)

安装信息摘要

RED HAT ENTERPRISE LINUX 7.2 安装

cn Help!

**SECURITY POLICY**  
No profile selected

软件

**安装源(I)**  
本地介质

**软件选择(S)**  
最小安装

系统

**安装位置(D)**  
已选择自动分区

**KDUMP**  
已启用 Kdump

**网络和主机名(N)**  
未连接

退出(Q) 开始安装(B)

在点击“开始安装”按钮前我们并不会操作您的磁盘。

请先完成带有此图标标记的内容再进行下一步。

软件选择

RED HAT ENTERPRISE LINUX 7.2 安装

完成(D) cn Help!

基本环境

- 最小安装  
基本功能。
- 基础设施服务器  
用于操作网络基础设施服务的服务器。
- 文件及打印服务器  
用于企业的文件、打印及存储服务器。
- 基本网页服务器  
提供静态及动态互联网内容的服务器。
- 虚拟化主机  
最小虚拟化主机。
- 带 GUI 的服务器  
带有用于操作网络基础设施服务 GUI 的服务器。

已选环境的附加选项

- 备份服务器  
集中管理基础设施备份的软件。
- DNS 名称服务器  
该软件包组允许您在系统上运行 DNS 名称服务器 (BIND)。
- 电子邮件服务器  
允许将系统作为 SMTP 和 (或者) IMAP 电子邮件服务器使用。
- FTP 服务器  
允许将系统作为 FTP 服务器使用。
- 文件及存储服务器  
CIFS, SMB, NFS, iSCSI, iSER 及 iSNS 网络存储服务器。
- 硬件监控工具  
一组用来监控服务器硬件的工具。
- 身份管理服务器  
用户、服务器和认证策略的集中管理。
- Infiniband 支持  
用来支持集群和使用 RDMA InfiniBand 和 iWARP 光纤的网状连接性的软件。
- Java 平台  
红帽企业版 Linux 服务器和桌面平台的 Java 支持。





完成(D)

cn

Help!

▼ 新 Red Hat Enterprise Linux 7.2 安装

您还没有为 Red Hat Enterprise Linux 7.2 的安装创建任何挂载点。您可以：

- [点这里自动创建他们 \(C\)。](#)
- 通过点击“+”按钮创建新挂载点。

新挂载点将使用以下分区方案 (N)：

LVM

在您为 Red Hat Enterprise Linux 7.2 安装创建挂载点后，您可在这里浏览它们的详细信息。

+ - ↻

可用空间

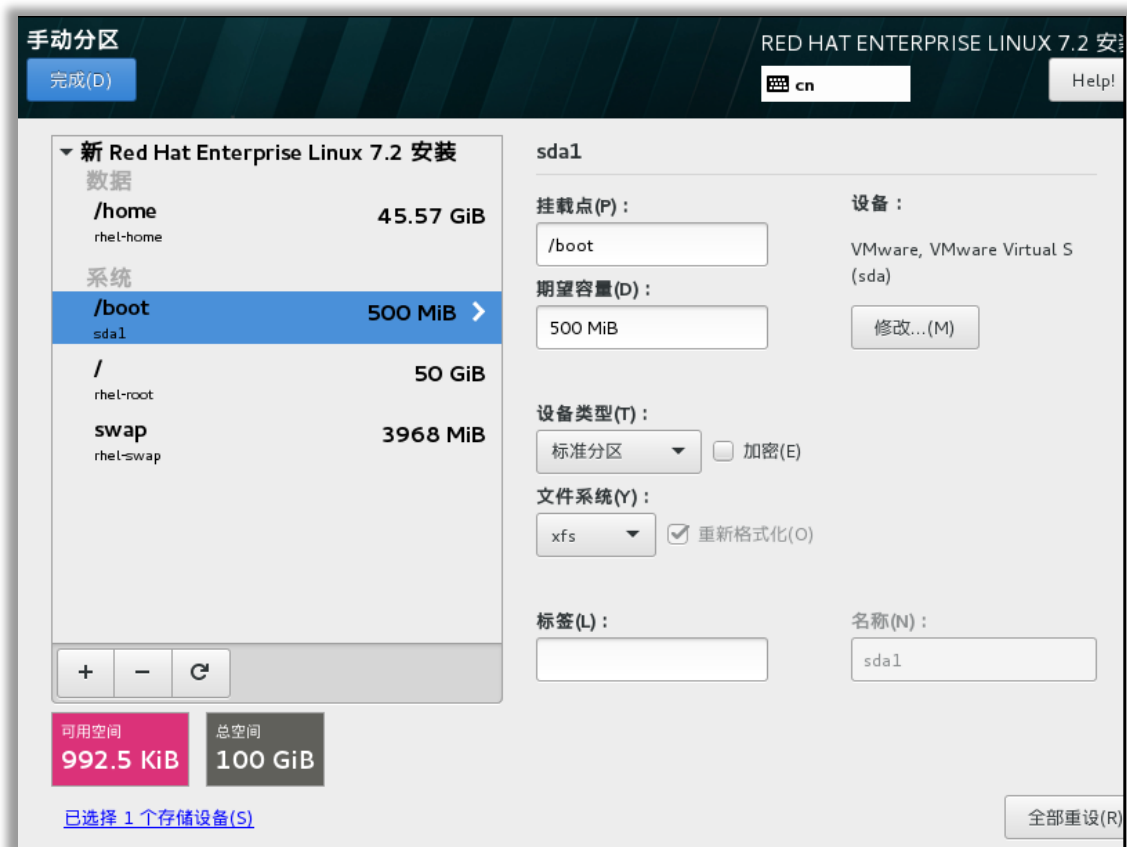
100 GiB

总空间

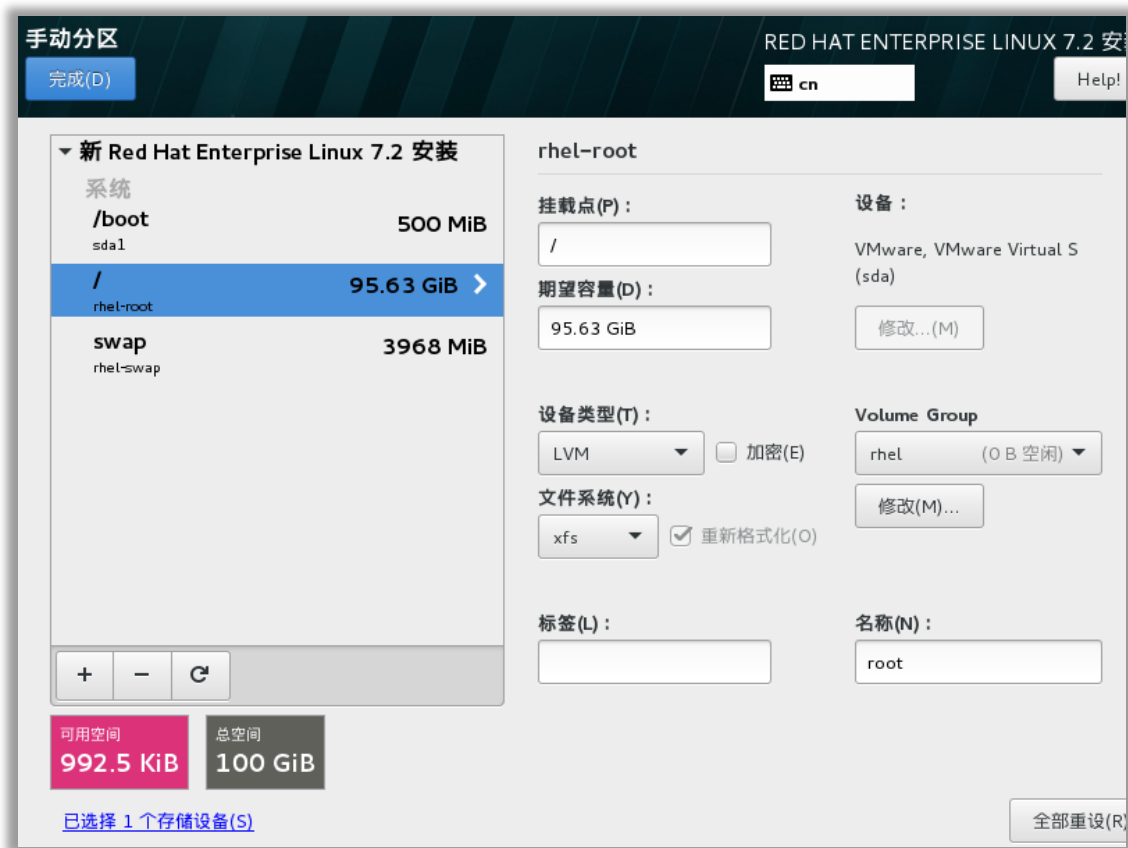
100 GiB

[已选择 1 个存储设备\(S\)](#)

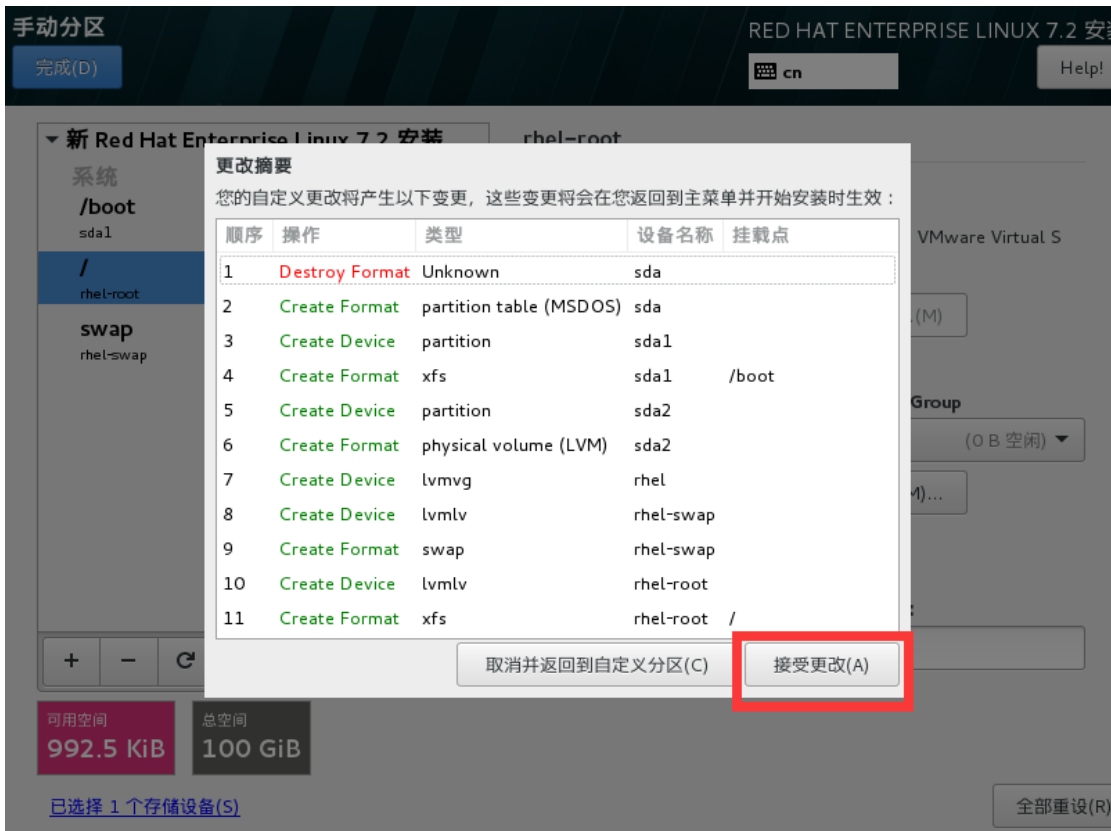
全部重设(R)

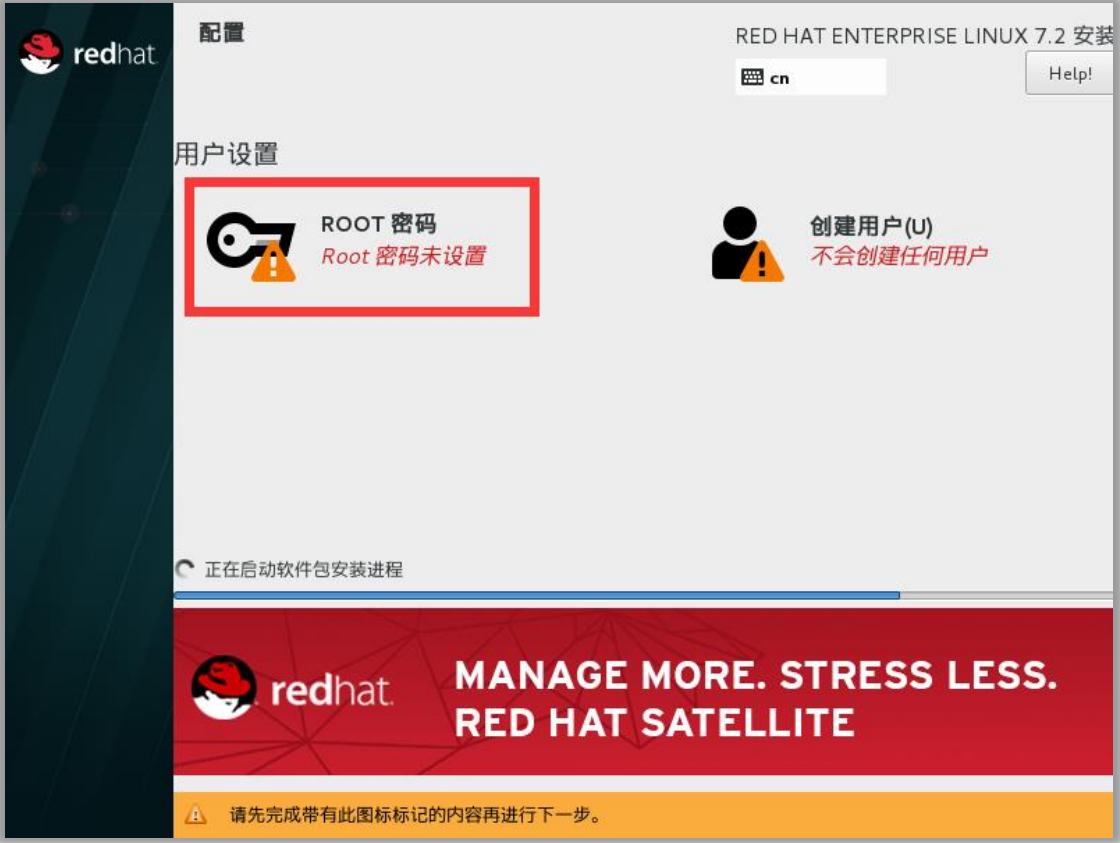


上图中分区方案既可以默认，也可以根据需求进行更改。由于 CDH 组件实在 root 用户下安装，home 分区可以删除，将硬盘全部划分到根分区，如下图。



分区完毕后点击“完成”按钮，继续安装





完成(D)

cn

Help!

root 帐户用于管理系统。为 root 用户输入密码。

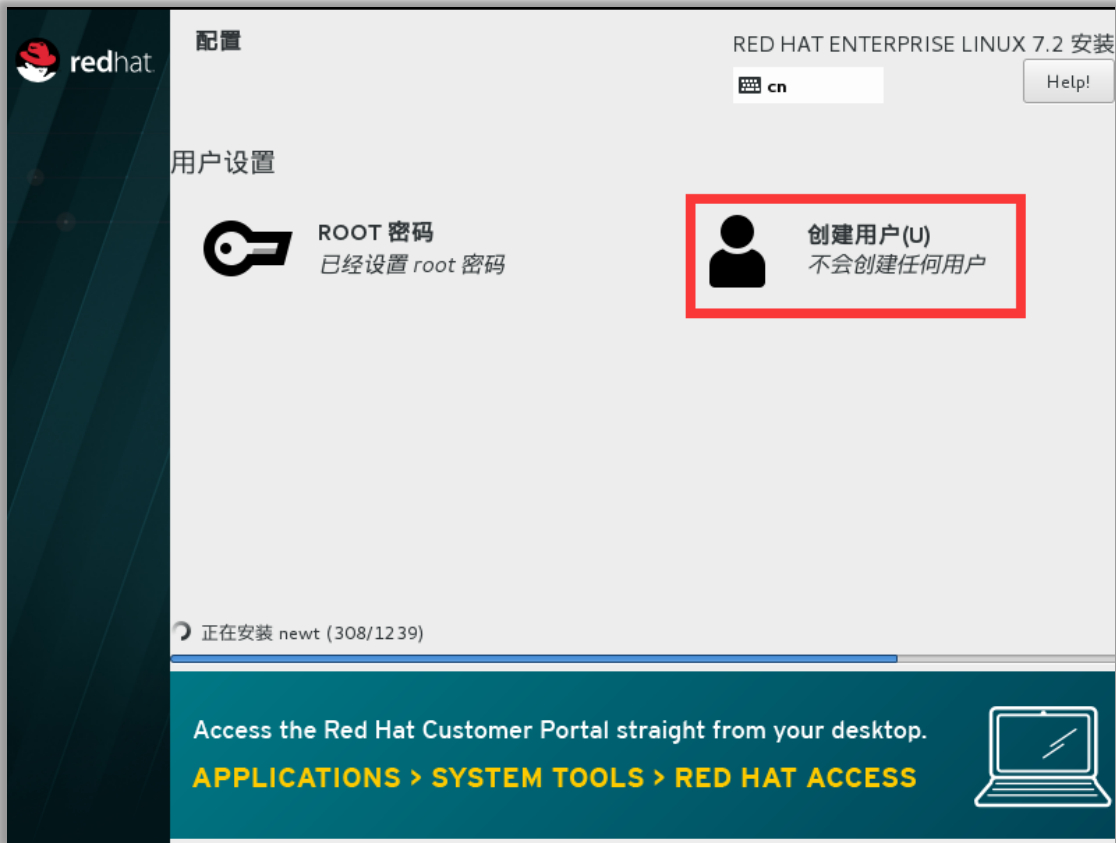
Root 密码 (R) :

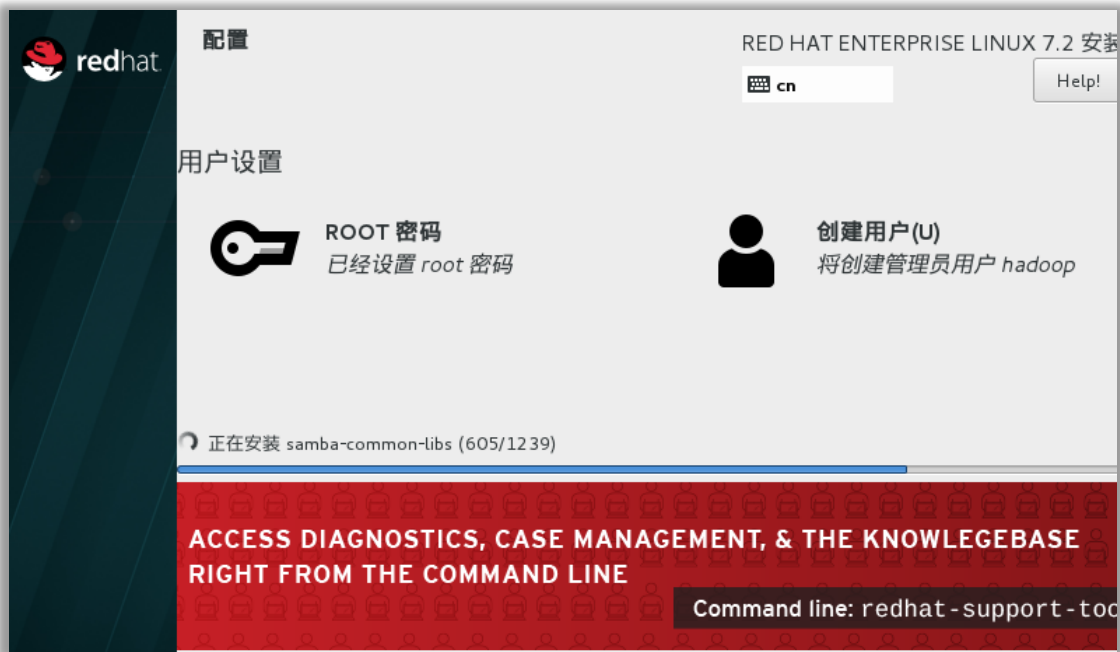
空白

确认(C) :

此处设置好root用户密码点击完成返回，若密码简单，则需要点击两次完成返回

 密码为空。





重启后按照系统提示进行确认进图操作系统，若碰到命令行界面的认证按照下图进行认证：



```
[ 9.244330] sd 2:0:0:0: [sd] Assuming drive cache: write through
[ OK ] Started Avahi mDNS/DNS-SD Stack.
[ OK ] Started RealtimeKit Scheduling Policy Service.
      Starting D-Bus System Message Bus...
[ OK ] Started Service for virtual machines hosted on VMware.
      Starting Service for virtual machines hosted on VMware...
[ OK ] Started System Logging Service.
=====
Initial setup of Red Hat Enterprise Linux Server 7.2 (Maipo)

1) [!] License information
   (License not accepted)
   Please make your choice from [ '1' to enter the License information spoke | 'q' to quit |
   'c' to continue | 'r' to refresh] 1
=====
License information

   1) Read the License Agreement

[ ] 2) I accept the license agreement.

   Please make your choice from above [ 'q' to quit | 'c' to continue |
   'r' to refresh] 2
=====
License information

   1) Read the License Agreement

[x] 2) I accept the license agreement.

   Please make your choice from above [ 'q' to quit | 'c' to continue |
   'r' to refresh] q
=====
Question

Are you sure you want to quit the configuration process?
You might end up with an unusable system if you do. Unless the License agreement
is accepted, the system will be rebooted.

Please respond 'yes' or 'no': yes
```



未列出？





hadoop

密码：

取消



登录





以上步骤在所有主机上执行

配置 Linux

系统安装好后，先不要将 U 盘或光盘介质取出

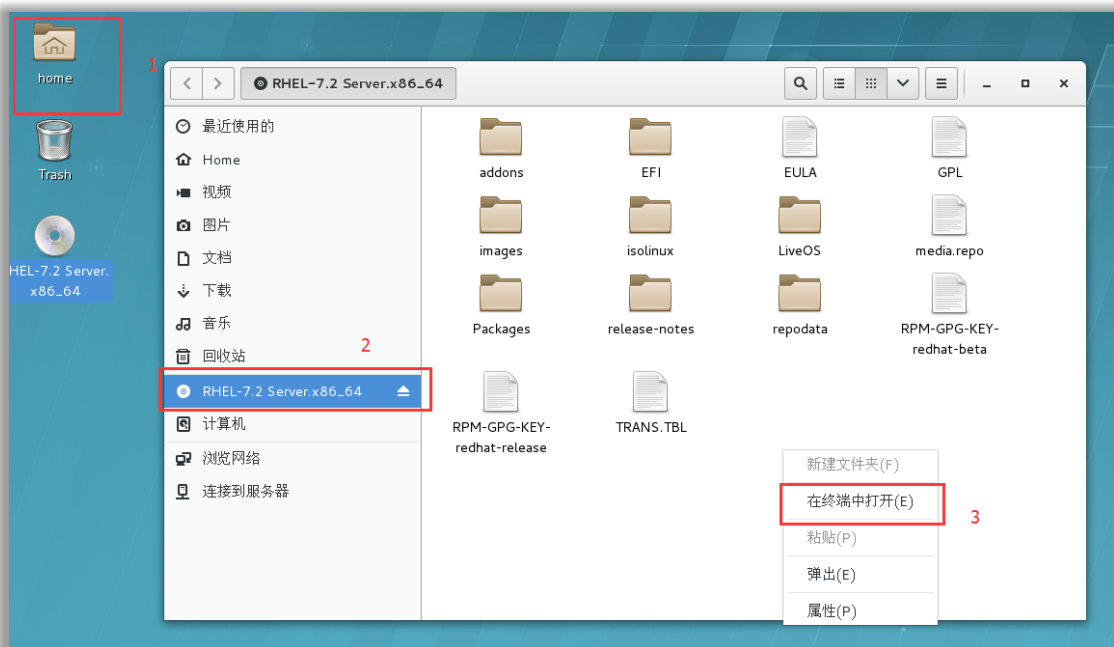
所有主机关闭并禁用防火墙

```
systemctl stop firewalld.service  
systemctl disable firewalld.service
```

编辑文件 `/etc/selinux/config` 将属性 SELINUX 改为 disabled ，如下：

```
# This file controls the state of SELinux on the system.
# SELINUX= can take one of these three values:
#   enforcing - SELinux security policy is enforced.
#   permissive - SELinux prints warnings instead of enforcing.
#   disabled - No SELinux policy is loaded.
SELINUX=disabled
# SELINUXTYPE= can take one of three two values:
#   targeted - Targeted processes are protected,
#   minimum - Modification of targeted policy. Only selected processes are p
rotected.
#   mls - Multi Level Security protection.
SELINUXTYPE=targeted
```

在主节点机器上打开文件管理器进入安装介质目录。并在此处打开命令行，切换至 root 用户。



在 /root 目录创建文件夹 rhel，并将安装介质中所有的文件及目录拷贝到 /root/rhel 之下

```
mkdir ~/rhel
```

```
cp -rfv ./ * ~/rhel/
```

将文件 /root/rhel/media.repo 拷贝到目录 /etc/yum.repos.d/

```
cp /root/rhel/media.repo /etc/yum.repos.d/
```

编辑文件 `/etc/yum.repos.d/media.repo` , 在文件中添加 `baseurl` 参数, 如下

```
[root@localhost rhel]# cat /etc/yum.repos.d/media.repo
[InstallMedia]
name=Red Hat Enterprise Linux 7.2
baseurl=file:///root/rhel
mediaid=1446216863.790260
metadata_expire=-1
gpgcheck=0
cost=500
[root@localhost rhel]#
```

`baseurl` 属性是当前镜像文件所在的位置。

运行命令 `yum makecache` 验证:

```
[root@localhost ~]# yum makecache
已加载插件: langpacks, product-id, search-disabled-repos, subscription-manager
This system is not registered to Red Hat Subscription Management. You can use
subscription-manager to register.
InstallMedia | 4.1 kB 00:00
(1/4): InstallMedia/group_gz | 136 kB 00:00
(2/4): InstallMedia/primary_db | 3.6 MB 00:00
(3/4): InstallMedia/filelists_db | 3.2 MB 00:00
(4/4): InstallMedia/other db | 1.4 MB 00:00
元数据缓存已建立
[root@localhost ~]#
```

仅在主节点上安装 `httpd`

```
yum install -y httpd
```

将文件夹 `/root/rhel/` 拷贝至目录 `/var/www/html/` 下

```
cp -rfv /root/rhel/ /var/www/html/
```

```
[root@localhost ~]# ls /var/www/html/
rhel
[root@localhost ~]# ls /var/www/html/rhel/
addons  images  Packages  RPM-GPG-KEY-redhat-release
EFI     isolinux release-notes  TRANS.TBL
EULA    LiveOS  repodata
GPL     media.repo  RPM-GPG-KEY-redhat-beta
[root@localhost ~]#
```

启动 httpd 并设置 httpd 服务开机自动运行

```
systemctl start httpd
```

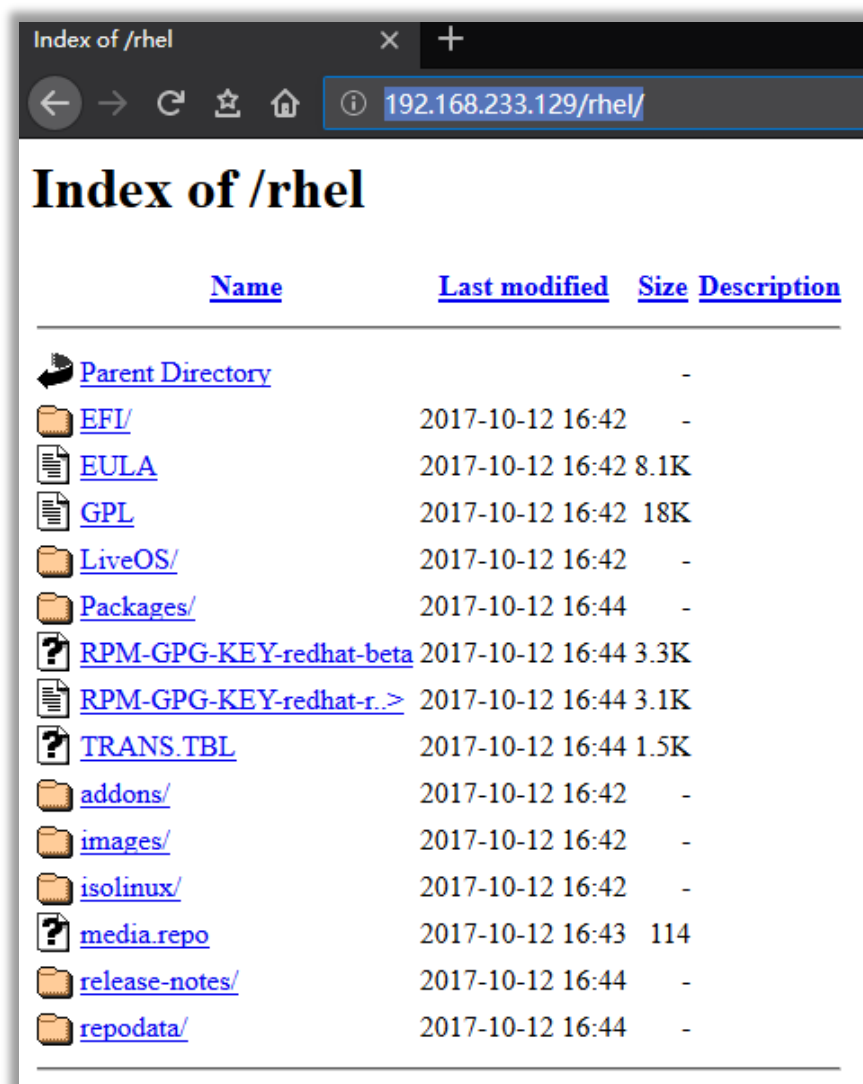
```
systemctl enable httpd
```

编辑文件 /etc/yum.repos.d/media.repo 将 baseurl 参数修改成 http 地址

```
baseurl= http://192.168.233.129/rhel/
```

192.168.233.129 是当前主机 IP , 需要根据实际情况进行更改。

打开浏览器可看到配置好的 yum 源



运行一下命令验证 yum 源是否安装成功：

```
yum clean all  
yum makecache
```

```
[root@localhost ~]# yum clean all  
已加载插件: langpacks, product-id, search-disabled-repos, subscription-manager  
This system is not registered to Red Hat Subscription Management. You can use  
subscription-manager to register.  
正在清理软件源: InstallMedia  
Cleaning up everything  
[root@localhost ~]# yum makecache  
已加载插件: langpacks, product-id, search-disabled-repos, subscription-manager  
This system is not registered to Red Hat Subscription Management. You can use  
subscription-manager to register.  
InstallMedia | 4.1 kB 00:00  
(1/4): InstallMedia/group_gz | 136 kB 00:00  
(2/4): InstallMedia/filelists_db | 3.2 MB 00:00  
(3/4): InstallMedia/primary_db | 3.6 MB 00:00  
(4/4): InstallMedia/other_db | 1.4 MB 00:00  
元数据缓存已建立  
[root@localhost ~]#
```

将文件 /etc/yum.repos.d/media.repo 通过 SSH 传送到所有主机的

/etc/yum.repos.d/ 目录，此时，可以卸载安装介质，所有主机可以安装镜像中的所有软件。

以上步骤只需在主节点上执行

接着对 linux 系统进行进一步的配置（此步骤需要在所有主机上执行）：

执行命令

```
systemctl start tuned  
tuned-adm off  
systemctl stop tuned  
systemctl disable tuned  
echo never > /sys/kernel/mm/transparent_hugepage/enabled  
echo never > /sys/kernel/mm/transparent_hugepage/defrag
```

编辑文件 /etc/rc.d/rc.local 将上面最后两行命令添加至文件末尾，如图：



```
#!/bin/bash
# THIS FILE IS ADDED FOR COMPATIBILITY PURPOSES
#
# It is highly advisable to create own systemd services or udev rules
# to run scripts during boot instead of using this file.
#
# In contrast to previous versions due to parallel execution during boot
# this script will NOT be run after all other services.
#
# Please note that you must run 'chmod +x /etc/rc.d/rc.local' to ensure
# that this script will be executed during boot.

touch /var/lock/subsys/local

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

给文件/etc/rc.d/rc.local 添加可执行的权限

```
chmod +x /etc/rc.d/rc.local
```

编辑文件/etc/default/grub 将 transparent\_hugepage=never 添加至如图所示的位置处

```
GRUB_TIMEOUT=5
GRUB_DISTRIBUTOR="$(sed 's, release .*$,,g' /etc/system-release)"
GRUB_DEFAULT=saved
GRUB_DISABLE_SUBMENU=true
GRUB_TERMINAL_OUTPUT="console"
GRUB_CMDLINE_LINUX="crashkernel=auto rd.lvm.lv=cl/root rd.lvm.lv=cl/swap rhgb quiet transparent_hugepage=never"
GRUB_DISABLE_RECOVERY="true"
~
```

运行命令 grub2-mkconfig -o /boot/grub2/grub.cfg

```
[root@centos ~]# grub2-mkconfig -o /boot/grub2/grub.cfg
Generating grub configuration file ...
Found linux image: /boot/vmlinuz-3.10.0-514.el7.x86_64
Found initrd image: /boot/initramfs-3.10.0-514.el7.x86_64.img
Found linux image: /boot/vmlinuz-0-rescue-f9e1d042b87c4e8c8327a27ff0d8e76c
Found initrd image: /boot/initramfs-0-rescue-f9e1d042b87c4e8c8327a27ff0d8e76c.img
done
[root@centos ~]#
```

编辑文件 `/etc/sysctl.conf` 在文件末尾添加 `vm.swappiness=1` 如图

```
# sysctl settings are defined through files in
# /usr/lib/sysctl.d/, /run/sysctl.d/, and /etc/sysctl.d/.
#
# Vendors settings live in /usr/lib/sysctl.d/.
# To override a whole file, create a new file with the same in
# /etc/sysctl.d/ and put new settings there. To override
# only specific settings, add a file with a lexically later
# name in /etc/sysctl.d/ and put new settings there.
#
# For more information, see sysctl.conf(5) and sysctl.d(5).

vm.swappiness=1
```

设置所有主机的 `hostname`，在全部的主机上执行以下命令：

```
hostnamectl set-hostname <hostname>
```

<hostname> 更换为对应主机需要设置的 `hostname`

```
[root@localhost ~]# hostnamectl set-hostname redhat_vm
[root@localhost ~]# hostname
redhat_vm
```

在所有主机的 `/etc/hosts` 文件中写入所有主机的 IP 地址以及 `hostname`，中间用空格分开，例如：

```
193.100.100.175 hostname1
193.100.100.176 hostname2
193.100.100.177 hostname3
...
...
...
```

卸载系统自带的 `openJDK`：

```
yum remove -y java*
```

接着重启所有节点主机

```
reboot
```

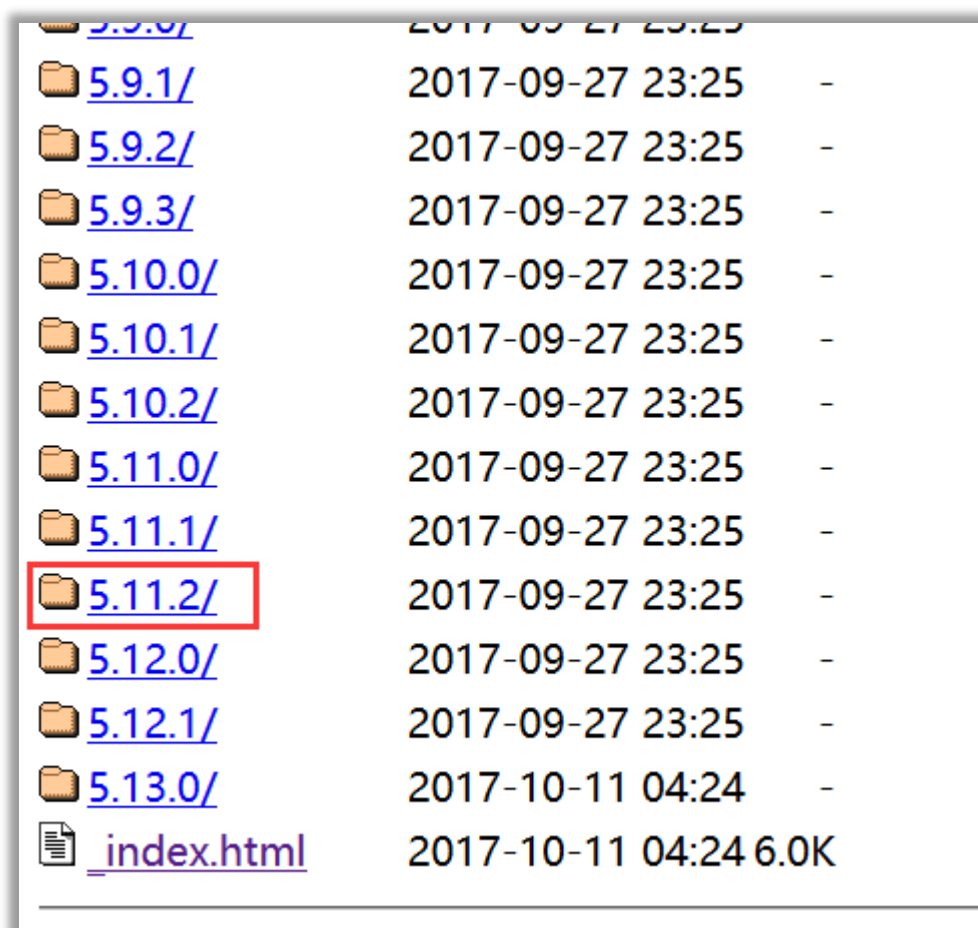
CM、CDH、KAFKA 下载

---

下载 CM

下载地址：<http://archive.cloudera.com/cm5/repo-as-tarball/>

页面打开后选择需要的版本进入相应的目录下，此处以 5.11.2 为例：



<a href="#">5.9.0/</a>	2017-09-27 23:25	-
<a href="#">5.9.1/</a>	2017-09-27 23:25	-
<a href="#">5.9.2/</a>	2017-09-27 23:25	-
<a href="#">5.9.3/</a>	2017-09-27 23:25	-
<a href="#">5.10.0/</a>	2017-09-27 23:25	-
<a href="#">5.10.1/</a>	2017-09-27 23:25	-
<a href="#">5.10.2/</a>	2017-09-27 23:25	-
<a href="#">5.11.0/</a>	2017-09-27 23:25	-
<a href="#">5.11.1/</a>	2017-09-27 23:25	-
<a href="#">5.11.2/</a>	2017-09-27 23:25	-
<a href="#">5.12.0/</a>	2017-09-27 23:25	-
<a href="#">5.12.1/</a>	2017-09-27 23:25	-
<a href="#">5.13.0/</a>	2017-10-11 04:24	-
<a href="#">_index.html</a>	2017-10-11 04:24	6.0K














进入目录后选择为 redhat 和 centos 适配的 CM 版本进行下载：

 <a href="#">cm5.11.2-centos6.tar.gz.sha1</a>	2017-08-22 20:35	114
 <a href="#">cm5.11.2-centos7.tar.gz</a>	2017-08-22 20:36	891M
 <a href="#">cm5.11.2-centos7.tar.gz.md5</a>	2017-08-22 20:36	106
 <a href="#">cm5.11.2-centos7.tar.gz.sha1</a>	2017-08-22 20:36	114
 <a href="#">cm5.11.2-debian-jessie.tar.gz</a>	2017-08-22 20:32	923M

下载 CDH

下载地址：<http://archive.cloudera.com/cdh5/parcels/>

CDH 选择与当前 CM 版本号一致的版本下载

 <a href="#">5.11.1/</a>	2017-09-29 23:40	-
 <a href="#">5.11.2.4/</a>	2017-09-29 23:40	-
 <a href="#">5.11.2/</a>	2017-09-29 23:40	-
 <a href="#">5.11/</a>	2017-09-29 23:40	-
 <a href="#">5.12.0.29/</a>	2017-09-29 23:40	-
 <a href="#">5.12.0/</a>	2017-09-29 23:40	-
 <a href="#">5.12.1.3/</a>	2017-09-29 23:40	-
 <a href="#">5.12.1/</a>	2017-09-29 23:40	-
 <a href="#">5.12/</a>	2017-09-29 23:40	-
 <a href="#">5.13.0.29/</a>	2017-10-10 01:21	-
 <a href="#">5.13.0/</a>	2017-10-10 01:21	-
 <a href="#">5.13/</a>	2017-10-10 01:21	-
 <a href="#">5/</a>	2017-10-10 01:21	-

同样选择为 redhat 和 centos 适配的版本进行下载，此处下载 .parcel 和 .parcel.sha1 格式的两个文件。

Parent Directory			
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-el5.parcel</a>	2017-08-22 19:57	1.4G	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-el5.parcel.sha1</a>	2017-08-22 19:57	41	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-el6.parcel</a>	2017-08-22 19:55	1.4G	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-el6.parcel.sha1</a>	2017-08-22 19:55	41	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-el7.parcel</a>	2017-08-22 19:56	1.5G	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-el7.parcel.sha1</a>	2017-08-22 19:56	41	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-jessie.parcel</a>	2017-08-22 19:56	1.5G	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-jessie.parcel.sha1</a>	2017-08-22 19:56	41	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-precise.parcel</a>	2017-08-22 19:56	1.5G	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-precise.parcel.sha1</a>	2017-08-22 19:56	41	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-sles11.parcel</a>	2017-08-22 19:57	1.4G	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-sles11.parcel.sha1</a>	2017-08-22 19:57	41	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-sles12.parcel</a>	2017-08-22 19:54	1.4G	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-sles12.parcel.sha1</a>	2017-08-22 19:54	41	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-trusty.parcel</a>	2017-08-22 19:56	1.5G	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-trusty.parcel.sha1</a>	2017-08-22 19:56	41	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-wheezy.parcel</a>	2017-08-22 19:57	1.5G	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-wheezy.parcel.sha1</a>	2017-08-22 19:57	41	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-xenial.parcel</a>	2017-08-22 19:55	1.5G	
<a href="#">CDH-5.11.2-1.cdh5.11.2.p0.4-xenial.parcel.sha1</a>	2017-08-22 19:55	41	
<a href="#">manifest.json</a>	2017-08-22 19:57	71K	

下载 KAFKA

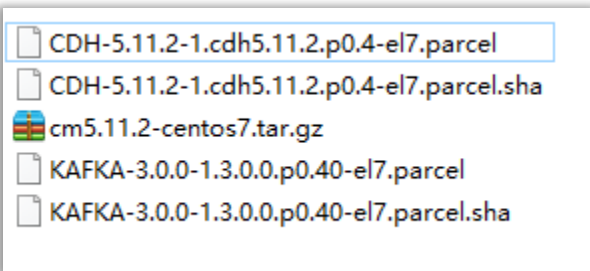
下载地址 : <http://archive.cloudera.com/kafka/parcels/>

KAFKA 选择最新的版本即可

<a href="#">3.0/</a>	2017-10-05 22:19	-
<a href="#">3/</a>	2017-10-05 22:19	-
<a href="#">index.html</a>	2017-10-02 15:59	3.1K
<a href="#">latest/</a>	2017-10-05 22:19	-

<u>Name</u>	<u>Last modified</u>	<u>Size</u>
Parent Directory		-
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-el5.parcel</a>	2017-10-05 22:19	63M
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-el5.parcel.sha1</a>	2017-10-05 22:19	41
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-el6.parcel</a>	2017-10-05 22:19	63M
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-el6.parcel.sha1</a>	2017-10-05 22:19	41
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-el7.parcel</a>	2017-10-05 22:19	63M
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-el7.parcel.sha1</a>	2017-10-05 22:19	41
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-jessie.parcel</a>	2017-10-05 22:19	63M
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-jessie.parcel.sha1</a>	2017-10-05 22:19	41
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-precise.parcel</a>	2017-10-05 22:19	63M
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-precise.parcel.sha1</a>	2017-10-05 22:19	41
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-sles11.parcel</a>	2017-10-05 22:19	63M
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-sles11.parcel.sha1</a>	2017-10-05 22:19	41
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-sles12.parcel</a>	2017-10-05 22:19	63M
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-sles12.parcel.sha1</a>	2017-10-05 22:19	41
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-trusty.parcel</a>	2017-10-05 22:19	63M
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-trusty.parcel.sha1</a>	2017-10-05 22:19	41
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-wheezy.parcel</a>	2017-10-05 22:19	63M
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-wheezy.parcel.sha1</a>	2017-10-05 22:19	41
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-xenial.parcel</a>	2017-10-05 22:19	63M
<a href="#">KAFKA-3.0.0-1.3.0.0.p0.40-xenial.parcel.sha1</a>	2017-10-05 22:19	41
<a href="#">manifest.json</a>	2017-10-05 22:19	5.1K

下载完成之后将所有 “.sha1” 的扩产名改为 “.sha” 。最终得到以下一组文件：



## 配置 SSH 免密码登陆

所有主机编辑文件 /etc/ssh/sshd\_config

将以下两行前面的#号去掉，保存

```
RSAAuthentication yes
PubkeyAuthentication yes
```

在所有主机上执行以下命令

```
ssh-keygen -t rsa
```

按四下回车得到 public key，如图：

```
[root@redhat_vm ~]# ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa):
Created directory '/root/.ssh'.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
5e:63:7c:6f:8f:5f:59:f6:16:03:d4:92:57:1c:d5:c8 root@redhat_vm
The key's randomart image is:
+--[ RSA 2048 ]-----+
|           oo+B|
|           .oEoo|
|            .o |
|            .  |
|           S = . oo|
|            . o o . .*|
|             .  o.+|
|              . +.|
|               ..o|
+-----+
[root@redhat_vm ~]#
```

在**所有主机**上执行以下命令，将所有主机的 public key 拷贝到主节点上

```
ssh-copy-id root@<hostname>
```

<hostname> 更换为**主节点主机**的主机名

```
[root@redhat_vm ~]# ssh-copy-id root@redhat_vm
The authenticity of host 'redhat_vm (192.168.233.129)' can't be established.
ECDSA key fingerprint is eb:49:60:96:5d:c6:46:e8:68:f8:eb:ae:0f:70:89:73.
Are you sure you want to continue connecting (yes/no)? yes
/usr/bin/ssh-copy-id: INFO: attempting to log in with the new key(s), to filter out
/usr/bin/ssh-copy-id: INFO: 1 key(s) remain to be installed -- if you are prompted
root@redhat_vm's password:  主节点root登录密码
Number of key(s) added: 1

Now try logging into the machine, with:  "ssh 'root@redhat_vm'"
and check to make sure that only the key(s) you wanted were added.
```

在**主节点机器**上通过 scp 命令将文件 /root/.ssh/authorized\_keys 发送到**所有其他**节点主机的 /root/.ssh/ 目录下：

```
scp /root/.ssh/authorized_keys root@<hostname>:/root/.ssh/
```

重启**所有主机**的 sshd 服务

```
systemctl restart sshd.service
```

```
systemctl enable sshd.service
```

验证免密码登陆：

```
ssh <hostname>
```

<hostname> 更换为主机名。

配置 NTP

在**所有主机**上安装 ntp



```
yum install -y ntp
```

在主节点上配置 ntp

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

添加如下内容

```
server 127.127.1.0
fudge 127.127.1.0 stratum 10
```

保存退出，然后启动 ntp 服务并设为开机启动：

```
systemctl start ntpd
```

```
systemctl enable ntpd
```

在集群其他节点安装 ntp 服务，并将 ntp 服务 server 指向主节点

```
echo "server <ip> iburst" > /etc/ntp.conf
```

<ip> 替换为主节点 ip 地址

启动所有节点的 ntp 服务

配置 CM 的 yum 安装源

将文件 CM 压缩包 cm5.11.2-centos7.tar.gz 上传至主节点机器，并将其解压到目录

/var/www/html/ 下：

```
tar zxvf cm5.11.2-centos7.tar.gz -C /var/www/html/
```

```
[root@redhat_vm ~]# ls /var/www/html/
cm  rhel
[root@redhat_vm ~]# ls /var/www/html/cm
5  5.11  5.11.2  cloudera-cm.repo  generated_index.html  RPM-GPG-KEY-cloudera
[root@redhat_vm ~]#
```

编辑文件 /var/www/html/cm/cloudera-cm.repo 更改属性 baseurl 和 gpgkey

```
baseurl=http://<ip>/cm/5/
```

```
gpgkey = http://<ip>/cm/RPM-GPG-KEY-cloudera
```

<ip> 替换为主节点 IP 地址，也就是运行有 httpd 服务的主节点地址。

```
[root@redhat_vm ~]# cat /var/www/html/cm/cloudera-cm.repo
[cloudera-cm]
# Packages for Cloudera's Distribution for cm, Version 5, on RedHat or CentOS 7 x86_64
name=Cloudera's Distribution for cm, Version 5
baseurl=http://192.168.233.129/cm/5/
gpgkey = http://192.168.233.129/cm/RPM-GPG-KEY-cloudera
gpgcheck = 1
```

再将文件 /var/www/html/cm/cloudera-cm.repo 拷贝到主节点以及所有其他节点机器的 /etc/yum.repos.d/ 目录下

验证 yum，在主节点以及任意一个其他节点机器运行命令

```
yum list|grep cloudera
```

可看到如下结果

```
[root@redhat_vm ~]# yum list|grep cloudera
cloudera-manager-agent.x86_64          5.11.2-1.cm5112.p0.6.e17 cloudera-cm
cloudera-manager-daemons.x86_64      5.11.2-1.cm5112.p0.6.e17 cloudera-cm
cloudera-manager-server.x86_64        5.11.2-1.cm5112.p0.6.e17 cloudera-cm
cloudera-manager-server-db-2.x86_64   5.11.2-1.cm5112.p0.6.e17 cloudera-cm
enterprise-debuginfo.x86_64           5.11.2-1.cm5112.p0.6.e17 cloudera-cm
jdk.x86_64                             2000:1.6.0_31-fcs        cloudera-cm
oracle-j2sdk1.7.x86_64                 1.7.0+update67-1        cloudera-cm
[root@redhat_vm ~]#
```

## 安装 Oracle JDK

所有主机均要先卸载 linux 自带的 openJDK，再安装 Oracle JDK

上个步骤中可以看出 CM 的安装源中自带一个 JDK7，如果想安装 7 版本的 JDK，可以直接执行命令安装。

```
yum install -y oracle-j2sdk1.7
```

如果想安装其他版本的 JDK 请到 oracle 官网下载安装：

地址：

JDK7：<http://www.oracle.com/technetwork/java/javase/downloads/java-archive-downloads-javase7-521261.html>

JDK8：<http://www.oracle.com/technetwork/java/javase/downloads/java-archive-javase8-2177648.html>

下载 .rpm 版本的安装包进行安装即可，譬如：

```
yum localinstall jdk-8u141-linux-x64.rpm
```

接着编辑文件 /etc/profile 将如下内容添加至文件末尾：

```
JAVA_HOME=/usr/java/jdk1.7.0_67-cloudera/  
JRE_HOME=/usr/java/jdk1.7.0_67-cloudera/jre  
PATH=$PATH:$JAVA_HOME/bin:$JRE_HOME/bin  
CLASSPATH=.:$JAVA_HOME/lib/dt.jar:$JAVA_HOME/lib/tools.jar:$JRE_  
HOME/lib  
export JAVA_HOME JRE_HOME PATH CLASSPATH
```

上述红色字体部分是 JDK 安装目录。请根据实际情况修改。rpm 包安装 JDK 的默认路径在 /usr/java/ 目录下。

安装和配置 mariadb

mariadb 仅需要在主节点安装即可

执行命令：

```
yum install -y mariadb-server
```

```
systemctl enable mariadb
```

启动 MariaDB

```
service mariadb start
```

## 初始化数据库

```
/usr/bin/mysql_secure_installation
```

## 在输出的提示中按照以下方法选择

```
In order to log into MariaDB to secure it, we'll need the current password for the root user. If you've just installed MariaDB, and you haven't set the root password yet, the password will be blank, so you should just press enter here.
```

```
Enter current password for root (enter for none): 此处回车
```

```
OK, successfully used password, moving on...
```

```
Setting the root password ensures that nobody can log into the MariaDB
```

```
root user without the proper authorisation.
```

```
Set root password? [Y/n] y
```

```
New password: 设置密码
```

```
Re-enter new password: 再次输入密码
```

```
Password updated successfully!
```

```
Reloading privilege tables..
```

```
... Success!
```

```
By default, a MariaDB installation has an anonymous user, allowing anyone
```

to log into MariaDB without having to have a user account created for them. This is intended only for testing, and to make the installation go a bit smoother. You should remove them before moving into a production environment.

Remove anonymous users? [Y/n] **y**

... Success!

Normally, root should only be allowed to connect from 'localhost'. This ensures that someone cannot guess at the root password from the network.

Disallow root login remotely? [Y/n] **n**

... skipping.

By default, MariaDB comes with a database named 'test' that anyone can access. This is also intended only for testing, and should be removed before moving into a production environment.

Remove test database and access to it? [Y/n] **y**

- Dropping test database...

... Success!

- Removing privileges on test database...

... Success!

```
Reloading the privilege tables will ensure that all changes made
so far
will take effect immediately.
```

```
Reload privilege tables now? [Y/n] y
```

```
... Success!
```

```
Cleaning up...
```

```
All done! If you've completed all of the above steps, your
MariaDB
installation should now be secure.
```

```
Thanks for using MariaDB!
```

## 登录数据库

```
mysql -uroot -p
```

## 在数据库终端中使用以下命令创建多个数据库

```
create database cm default character set utf8;
CREATE USER 'cm'@'%' IDENTIFIED BY 'password';
GRANT ALL PRIVILEGES ON cm. * TO 'cm'@'%';
FLUSH PRIVILEGES;

create database amon default character set utf8;
CREATE USER 'amon'@'%' IDENTIFIED BY 'password';
```

```
GRANT ALL PRIVILEGES ON amon. * TO 'amon'@'%';
FLUSH PRIVILEGES;

create database rman default character set utf8;
CREATE USER 'rman'@'% ' IDENTIFIED BY 'password';
GRANT ALL PRIVILEGES ON rman. * TO 'rman'@'%';
FLUSH PRIVILEGES;

create database metastore default character set utf8;
CREATE USER 'hive'@'% ' IDENTIFIED BY 'password';
GRANT ALL PRIVILEGES ON metastore. * TO 'hive'@'%';
FLUSH PRIVILEGES;

create database sentry default character set utf8;
CREATE USER 'sentry'@'% ' IDENTIFIED BY 'password';
GRANT ALL PRIVILEGES ON sentry. * TO 'sentry'@'%';
FLUSH PRIVILEGES;

create database nav default character set utf8;
CREATE USER 'nav'@'% ' IDENTIFIED BY 'password';
GRANT ALL PRIVILEGES ON nav. * TO 'nav'@'%';
FLUSH PRIVILEGES;

create database navms default character set utf8;
CREATE USER 'navms'@'% ' IDENTIFIED BY 'password';
GRANT ALL PRIVILEGES ON navms. * TO 'navms'@'%';
FLUSH PRIVILEGES;

create database oozie default character set utf8;
```

```
CREATE USER 'oozie'@'%' IDENTIFIED BY 'password';
GRANT ALL PRIVILEGES ON oozie. * TO 'oozie'@'>';
FLUSH PRIVILEGES;

create database hue default character set utf8;
CREATE USER 'hue'@'%' IDENTIFIED BY 'password';
GRANT ALL PRIVILEGES ON hue. * TO 'hue'@'>';
FLUSH PRIVILEGES;
```

```
MariaDB [(none)]> show databases;
+-----+
| Database |
+-----+
| information_schema |
| amon |
| cm |
| hue |
| metastore |
| mysql |
| nav |
| navms |
| oozie |
| performance_schema |
| rman |
| sentry |
+-----+
12 rows in set (0.00 sec)
```

下载 JDBC 驱动

地址：<https://dev.mysql.com/downloads/connector/j/>

在主节点上新建目录 /usr/share/java/

```
mkdir -p /usr/share/java/
```



将 jdbc 驱动包上传至主节点机器的目录 /usr/share/java/下

## 运行命令

```
cd /usr/share/java  
ln -s mysql-connector-java-5.1.44-bin.jar mysql-connector-java.jar  
ls -al
```

```
[root@redhat_vm java]# ls -al  
总用量 992  
drwxr-xr-x  2 root root   79 10月 13 09:58 .  
drwxr-xr-x 218 root root 8192 10月 13 09:56 ..  
-rw-r--r--  1 root root 999635 8月 29 10:13 mysql-connector-java-5.1.44-bin.jar  
lrwxrwxrwx  1 root root   35 10月 13 09:58 mysql-connector-java.jar -> mysql-connector-java-5.1.44-bin.jar
```


## 安装 CM 和 CDH

### 主节点机器执行命令

```
yum install -y cloudera-manager-server.x86_64  
  
/usr/share/cmf/schema/scm_prepare_database.sh mysql cm cm password
```

```
[root@redhat_vm java]# /usr/share/cmf/schema/scm_prepare_database.sh mysql cm cm password  
JAVA_HOME=/usr/java/jdk1.7.0_67-cloudera  
Verifying that we can write to /etc/cloudera-scm-server  
Creating SCM configuration file in /etc/cloudera-scm-server  
Executing: /usr/java/jdk1.7.0_67-cloudera/bin/java -cp /usr/share/java/mysql-connector-java.jar:/usr/  
le-connector-java.jar:/usr/share/cmf/schema/./lib/* com.cloudera.enterprise.dbutil.DbCommandExecutor  
cm-server/db.properties com.cloudera.cmf.db.  
[ main] DbCommandExecutor INFO Successfully connected to databa  
All done, your SCM database is configured correctly!  
[root@redhat_vm java]#
```

将下图所示的文件上传至主节点的目录 /opt/cloudera/parcel-repo 下



```
CDH-5.11.2-1.cd5.11.2.p0.4-el7.parcel  
CDH-5.11.2-1.cd5.11.2.p0.4-el7.parcel.sha  
cm5.11.2-centos7.tar.gz  
KAFKA-3.0.0-1.3.0.0.p0.40-el7.parcel  
KAFKA-3.0.0-1.3.0.0.p0.40-el7.parcel.sha
```

## 启动 CM 服务

```
service cloudera-scm-server start
```

在浏览器中访问地址：<http://<ip>:7180>

<ip> 为主节点 ip

登陆用户名和密码均为 admin



The screenshot displays a login interface for the Cloudera SCM Server. It features a white background with a light gray border. At the top, there is a text input field containing the username 'admin'. Below it is a password input field with six black dots for masking. A prominent blue button with the white Chinese characters '登录' (Login) is centered below the password field. At the bottom left, there is a checked checkbox followed by the text '保留我的信息' (Remember my information).

## 欢迎使用 Cloudera Manager

### 最终用户许可条款和条件

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.

是的，我接受最终用户许可条款和条件。

如果您代表与 Cloudera 已签订软件使用协议的公司下载并使用 Cloudera Manager，您的操作不会修改该现有协议。

1

2

返回

1 2

继续

## 欢迎使用 Cloudera Manager

### 您想要部署哪个版本？

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。

	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 上传许可证 选择许可证文件 上传 Cloudera Enterprise 在三个版本中可用： • Basic Edition • Flex Edition • Data 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		✓	✓
Cloudera Navigator Key Trustee			✓

返回

1 2

继续

## 感谢您选择 Cloudera Manager 和 CDH。

将安装此安装程序 Cloudera Express 5.11.2，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 Sentry
- 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 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。这还将启用该主机的运行状况监控。

提示: 使用模式 `ip` 搜索主机名和 IP 地址。

192.168.233.129

SSH 端口:

输入所有节点的IP或者主机名，用逗号分隔

返回

继续

搜索框中填入所有主机名或者 IP，用逗号分隔。

## 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。这还将启用该主机的运行状况监控。

提示: 使用 [高级模式](#) 搜索主机名和 IP 地址。

已扫描 1 个主机, 其中 1 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"/>	192.168.233.129	redhat_vm	192.168.233.129	否	✔ 主机准备就绪: 1 毫秒响应时间。

返回

继续

## 群集安装

### 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](#) [代理设置](#)

选择 CDH 的版本

CDH-5.11.2-1.cdh5.11.2.p0.4

对于此 Cloudera Manager 版本 (5.11.2) 太新的 CDH 版本不会显示。

其他 Parcel

KAFKA-3.0.0-1.3.0.0.p0.40

无

KUDU-1.4.0-1.cdh5.12.1.p0.10

无

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

此 Cloudera Manager Server 的匹配发行版

自定义存储库

返回

1 2 3 4 5 6 7

继续

此处点击“更多选项”按钮进行下图设置

Parcel 存储库设置

Parcel 目录  主件 "p  
需要重启代理

本地 Parcel 存储库路径  本

远程 Parcel 存储库 URL

- +
- +
- +
- +
- +
- +
- +
- +
- +

Cl  
U  
服  
接

本地 Parcel 存储库路径

远程 Parcel 存储库 URL

- +
- +
- +
- +
- +
- +
- +
- +
- +

红框中的 IP 地址更换为主节点的 IP，然后删除其他远程地址

Parcel 目录

 需要重启代理

/opt/cloudera/parcels

本地 Parcel 存储库路径

/opt/cloudera/parcel-repo

远程 Parcel 存储库 URL

http://192.168.233.129/cm/5/



## 群集安装

### 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](#) [代理设置](#)

选择 CDH 的版本  CDH-5.11.2-1.cdh5.11.2.p0.4

对于此 Cloudera Manager 版本 (5.11.2) 太新的 CDH 版本不会显示。

其他 Parcel  KAFKA-3.0.0-1.3.0.0.p0.40

无

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

此 Cloudera Manager Server 的匹配发行版

自定义存储库

[返回](#)

1 2 3 4 5 6 7

[继续](#)

## 群集安装

### JDK 安装选项

Oracle Binary Code License Agreement for the Java SE Platform Products and JavaFX

ORACLE AMERICA, INC. ("ORACLE"), FOR AND ON BEHALF OF ITSELF AND ITS SUBSIDIARIES AND AFFILIATES UNDER COMMON CONTROL, IS WILLING TO LICENSE THE SOFTWARE TO YOU ONLY UPON THE CONDITION THAT YOU ACCEPT ALL OF THE TERMS CONTAINED IN THIS BINARY CODE LICENSE AGREEMENT AND SUPPLEMENTAL LICENSE TERMS (COLLECTIVELY "AGREEMENT"). PLEASE READ THE AGREEMENT CAREFULLY. BY SELECTING THE "ACCEPT LICENSE AGREEMENT" (OR THE EQUIVALENT) BUTTON AND/OR BY USING THE SOFTWARE YOU ACKNOWLEDGE THAT YOU HAVE READ THE TERMS AND AGREE TO THEM. IF YOU ARE AGREEING TO THESE TERMS ON BEHALF OF A COMPANY OR OTHER LEGAL ENTITY, YOU REPRESENT THAT YOU HAVE THE LEGAL AUTHORITY TO BIND THE LEGAL ENTITY TO THESE TERMS. IF YOU DO NOT HAVE SUCH AUTHORITY, OR IF YOU DO NOT WISH TO BE BOUND BY THE TERMS, THEN SELECT THE "DECLINE LICENSE AGREEMENT" (OR THE EQUIVALENT) BUTTON AND YOU MUST NOT USE THE SOFTWARE ON THIS SITE OR ANY OTHER MEDIA ON WHICH THE SOFTWARE IS CONTAINED.

1. DEFINITIONS. "Software" means the software identified above in binary form that you selected for download, install or use (in the version You selected for download, install or use) from Oracle or its authorized licensees, any other machine readable materials (including, but not limited to, libraries, source files, header files, and data files), any updates or error corrections provided by Oracle, and any user manuals, programming guides and other documentation provided to you by Oracle under this Agreement. "General Purpose Desktop Computers and Servers" means computers, including desktop and laptop computers, or servers, used for general computing functions under end user control (such as but not specifically limited to email, general purpose Internet browsing, and office suite productivity tools). The use of Software in systems and solutions that provide dedicated functionality (other than as mentioned above) or designed for use in embedded or function-specific software applications, for example but not limited to: Software embedded in or bundled with industrial control systems, wireless mobile telephones, wireless handheld devices, kiosks, TV/STB, Blu-ray Disc devices, telematics and network control switching equipment, printers and storage management systems, and other related systems are excluded from this definition and not licensed under this Agreement. "Programs" means (a) Java technology applets and applications intended to run on the Java Platform, Standard Edition platform on Java-enabled General Purpose Desktop Computers and Servers; and (b) JavaFX technology applications intended to run on the JavaFX Runtime on JavaFX-enabled General Purpose Desktop Computers and Servers. "Commercial Features" means those features identified in Table 1-1 (Commercial Features in Java SE Product Editions) of the Java SE documentation accessible at <http://www.oracle.com/technetwork/java/javase/documentation/index.html>.

安装 Oracle Java SE 开发工具包 (JDK)

选中此复选框以接受“Oracle 二进制代码许可协议”并安装 JDK。取消选中以使用当前安装的 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。

返回

1 2 3 4 5 6 7

继续

此处如果是手动安装了其他版本的 JDK，则不需要打勾，直接继续

## 群集安装

### 启用单用户模式

**仅受 CDH 5.2 及更高版本支持。**

默认情况下，服务进程以不同用户身份在系统中运行。例如，HDFS DataNode 以用户“hdfs”的身份运行，HBase RegionServer 以用户“hbase”的身份运行。启用“单个用户模式”将 Cloudera Manager 配置为以单个用户的身份运行服务进程，默认为“cloudera-scm”，从而使受管服务与系统其他服务之间的分离优先于受管服务之间的分离。

此选项的主要益处是 Agent 不作为根运行。但是，此模式将使安装变得复杂，这将在文档中作完整介绍。最明显的是，在常规模式下由 Agent 自动创建的目录必须通过相应权限在每个主机上手动创建，并且必须为配置的用户设置 sudo（或同等）访问权限。

在单个用户模式与常规模式之间来回切换不受支持。

单用户模式



返回

1 2 3 4 5 6 7

继续



## 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：  
•

root用户登录密码

SSH 端口：

同时安装的数量：  
 (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](#)

1 2 3 4 5 6 7

[继续](#)

## 群集安装

已成功完成安装。



已成功完成 1 个主机中的 1 个。

主机名称	IP 地址	进度	状态
linux176	193.100.100.176	<div style="width: 100%; height: 10px; background-color: green;"></div>	已成功完成安装。 <a href="#">详细信息</a>

[返回](#)

1 2 3 4 5 6 7

[继续](#)

## 群集安装

正在安装选定 Parcel

选定的 Parcel 正在下载并安装在群集的所有主机上。

▶ CDH 5.11.2-1.cdh5.11.2.p0.4	已下载: 100%	已分配: 1/1 (50.3 MiB/s)	已解压: 1/1	已激活: 1/1
▶ KAFKA 3.0.0-1.3.0.0.p0.40	已下载: 100%	已分配: 1/1 (1.2 MiB/s)	已解压: 1/1	已激活: 1/1

返回

1 2 3 4 5 6 7

继续

### 群集安装

检查主机正确性 [C 重新运行](#)

#### 验证

✓	检查器在所有 1 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。
✓	主机时转几乎同步 (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。
✓	所有主机上的 /proc/sys/vm/swappiness 都未发现有问题。
✓	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, 1 台主机正在运行 CDH 5。
✓	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。

[返回](#)

1 2 3 4 5 6 7

[完成](#)

Feedback

#### 版本汇总

Cluster 1 - CDH 5			
主机			
linux176			
组件	版本	发行版	CDH 版本
Bigtop-Tomcat (仅限 CDH 5)	0.7.0+cdh5.11.2+0	1.cdh5.11.2.p0.7	CDH 5
Crunch (仅限 CDH 5)	0.11.0+cdh5.11.2+99	1.cdh5.11.2.p0.7	CDH 5
Flume NG	1.6.0+cdh5.11.2+165	1.cdh5.11.2.p0.7	CDH 5
MapReduce 1	2.6.0+cdh5.11.2+2430	1.cdh5.11.2.p0.7	CDH 5
Hadoop	2.6.0+cdh5.11.2+2430	1.cdh5.11.2.p0.7	CDH 5
HDFS	2.6.0+cdh5.11.2+2430	1.cdh5.11.2.p0.7	CDH 5
HTTPFS	2.6.0+cdh5.11.2+2430	1.cdh5.11.2.p0.7	CDH 5
hadoop-kms	2.6.0+cdh5.11.2+2430	1.cdh5.11.2.p0.7	CDH 5
MapReduce 2	2.6.0+cdh5.11.2+2430	1.cdh5.11.2.p0.7	CDH 5
YARN	2.6.0+cdh5.11.2+2430	1.cdh5.11.2.p0.7	CDH 5
HBase	1.2.0+cdh5.11.2+340	1.cdh5.11.2.p0.7	CDH 5
Lily HBase Indexer	1.5+cdh5.11.2+71	1.cdh5.11.2.p0.7	CDH 5
Hive	1.1.0+cdh5.11.2+1082	1.cdh5.11.2.p0.7	CDH 5
HCatalog	1.1.0+cdh5.11.2+1082	1.cdh5.11.2.p0.7	CDH 5
Hue	3.9.0+cdh5.11.2+5098	1.cdh5.11.2.p0.7	CDH 5
Impala	2.8.0+cdh5.11.2+0	1.cdh5.11.2.p0.7	CDH 5
Kite (仅限 CDH 5)	1.0.0+cdh5.11.2+144	1.cdh5.11.2.p0.7	CDH 5
Llama (仅限 CDH 5)	1.0.0+cdh5.11.2+0	1.cdh5.11.2.p0.7	CDH 5
Mahout	0.9+cdh5.11.2+34	1.cdh5.11.2.p0.7	CDH 5
Oozie	4.1.0+cdh5.11.2+441	1.cdh5.11.2.p0.7	CDH 5
Parquet	1.5.0+cdh5.11.2+187	1.cdh5.11.2.p0.7	CDH 5
Pig	0.12.0+cdh5.11.2+108	1.cdh5.11.2.p0.7	CDH 5
sentry	1.5.1+cdh5.11.2+918	1.cdh5.11.2.p0.7	CDH 5
Solr	4.10.3+cdh5.11.2+514	1.cdh5.11.2.p0.7	CDH 5
spark	1.6.0+cdh5.11.2+519	1.cdh5.11.2.p0.7	CDH 5
Sqoop	1.99.5+cdh5.11.2+46	1.cdh5.11.2.p0.7	CDH 5
Sqoop	1.4.6+cdh5.11.2+112	1.cdh5.11.2.p0.7	CDH 5
Whirr	0.9.0+cdh5.11.2+23	1.cdh5.11.2.p0.7	CDH 5
ZooKeeper	3.4.5+cdh5.11.2+117	1.cdh5.11.2.p0.7	CDH 5
Kafka	0.11.0+kaflka3.0.0	1.3.0.0.p0.50	不适用
Cloudera Manager Management Daemon	5.11.2	1.cm5112.p0.6	不适用
Supervisor	3.0-cm5.11.2	不可用	不适用
Java 7	JAVA_HOME=/usr/java/jdk1.7.0_67-cloudera java version "1.7.0_67" Java(TM) SE Runtime Environment (build 1.7.0_67-b01) Java HotSpot(TM) 64-Bit Server VM (build 24.65-b04, mixed mode)	不可用	不适用
Java 6	JAVA_HOME=/usr/java/jdk1.6.0_31 java version "1.6.0_31" Java(TM) SE Runtime Environment (build 1.6.0_31-b04) Java HotSpot(TM) 64-Bit Server VM (build 20.6-b01, mixed mode)	不可用	不适用
Cloudera Manager Agent	5.11.2	1.cm5112.p0.6.e17	不适用

此处安装哪些服务可根据需要进行选择

### 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 (含 MapReduce 2)、ZooKeeper、Oozie、Hive 和 Hue
- 含 HBase 的内核  
HDFS、YARN (含 MapReduce 2)、ZooKeeper、Oozie、Hive、Hue 和 HBase
- 含 Impala 的内核  
HDFS、YARN (含 MapReduce 2)、ZooKeeper、Oozie、Hive、Hue 和 Impala
- 含 Search 的内核  
HDFS、YARN (含 MapReduce 2)、ZooKeeper、Oozie、Hive、Hue 和 Solr
- 含 Spark 的内核  
HDFS、YARN (含 MapReduce 2)、ZooKeeper、Oozie、Hive、Hue 和 Spark
- 所有服务  
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。

[返回](#) 1 2 3 4 5 6 [继续](#)

角色分配暂且可以默认，待引导结束后可再进行进一步的更改

### 群集设置

#### 自定义角色分配

您可在此处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 不建议改变分配情况。

还可以按主机查看角色分配。 [按主机查看](#)

#### HBase

- M** Master × 1 新建
- HBRS** HBase REST Server
- HHTS** HBase Thrift Server
- RS** RegionServer × 1 新建

#### HDFS

- NN** NameNode × 1 新建
- SNN** SecondaryNameNode × 1 新建
- B** Balancer × 1 新建
- HFS** HttpFS
- NFSG** NFS Gateway
- DN** DataNode × 1 新建

#### Hive

- G** Gateway × 1 新建
- HMS** Hive Metastore Server × 1 新建
- WHCCS** WebHCat Server
- HS2** HiveServer2 × 1 新建

#### Hue

[返回](#) 1 2 3 4 5 6 [继续](#)

下图填写在安装 mariadb 时所用的数据库名称和密码，对应关系如下表所描述

服务名	数据库名称	用户名	密码
Activity Monitor	amon	amon	password
Reports Manager	rman	rman	password
Hive Metastore Server	metastore	hive	password
Sentry Server	sentry	sentry	password
Cloudera Navigator Audit Server	nav	nav	password
Cloudera Navigator Metadata Server	navms	navms	password
Oozie Server	oozie	oozie	password
Hue	hue	hue	password

**群集设置**  
数据库设置

配置和测试数据库连接。首先根据 [Installation Guide](#) 的 [Installing and Configuring an External Database](#) 小节创建数据库。

**Hive** ✔ Successful

数据库主机名称: \*  数据库类型:  数据库名称: \*  用户名: \*  密码:

**Hue** ✔ Successful

数据库主机名称: \*  数据库类型:  数据库名称: \*  用户名: \*  密码:

**Oozie Server** ✔ Successful

当前被分配在 linux176 上运行。

数据库主机名称: \*  数据库类型:  数据库名称: \*  用户名: \*  密码:

显示密码

1

备注:

- 创建数据库时，数据库主机名称 字段中的值必须与您用于主机名称的值匹配。 [了解更多](#)
- 如数据库未在其默认端口运行，请使用 数据库主机名称 字段中的 host:port 指定端口号。
- 强烈建议将各个数据库与相应角色实例置于同一主机上。

1 2 3 4 5 6 2

## 群集设置

### 审核更改

<b>HDFS 根目录</b> hbase.rootdir	Cluster 1 > HBase (服务范围)	<input type="text" value="/hbase"/>	?
<b>启用编制索引</b>	<input checked="" type="checkbox"/> Cluster 1 > HBase (服务范围) ←		?
<b>启用复制</b> hbase.replication	<input checked="" type="checkbox"/> Cluster 1 > HBase (服务范围) ←		?
<b>HDFS 块大小</b> dfs.block.size, dfs.blocksize	Cluster 1 > HDFS (服务范围)	<input type="text" value="128"/> 兆字节 ▾	?
<b>接受的 DataNode 失败的卷</b> dfs.datanode.failed.volumes.tolerated	Cluster 1 > DataNode Default Group	<input type="text" value="0"/>	?
<b>DataNode 数据目录</b> dfs.data.dir, dfs.datanode.data.dir	Cluster 1 > DataNode Default Group ←	<input type="text" value="/dfs/dn"/> - +	?
<b>NameNode 数据目录</b> dfs.name.dir, dfs.namenode.name.dir	Cluster 1 > NameNode Default Group ←	<input type="text" value="/dfs/nn"/> - +	?

[返回](#) 1 2 3 4 5 6 [继续](#)

## 群集设置

### ✓ 首次运行 命令

状态: 已完成 开始时间: 10月 13, 2:15:14 下午 持续时间: 6.4m

Finished First Run of the following services successfully: ZooKeeper, HDFS, HBase, Solr, YARN (MR2 Included), Key-Value Store Indexer, Spark, Hive, Impala, Oozie, Hue, Cloudera Management Service.

全部  只会失败  仅限运行状态

详细信息 [已完成 10 个步骤 \(共 10 个\)](#)

步骤	上下文	开始时间	持续时间	操作
▶ ✓ 并行运行 1 步骤 已成功完成 1 个步骤。		10月 13, 2:15:14 下午	38ms	
▶ ✓ 正在部署客户端配置 Successfully deployed all client configurations.	Cluster 1 ↗	10月 13, 2:15:14 下午	16.05s	
▶ ✓ 启动 Cloudera Management Service, ZooKeeper 已成功完成 2 个步骤。		10月 13, 2:15:30 下午	25.06s	
▶ ✓ 启动 HDFS 已成功完成 1 个步骤。		10月 13, 2:15:55 下午	41.08s	
▶ ✓ 启动 HBase, Solr 已成功完成 2 个步骤。		10月 13, 2:16:36 下午	66.1s	
▶ ✓ 启动 Key-Value Store Indexer, YARN (MR2 Included) 已成功完成 2 个步骤。		10月 13, 2:17:42 下午	27.7s	
▶ ✓ 启动 Oozie		10月 13, 2:18:10 下午	21.01s	

[返回](#)

1 2 3 4 5 6

[继续](#)

默认配置与实际的硬件环境不符，首次启动会有很多的警告

# 主页

状态 所有运行状况问题 **1** 配置 **8** 所有最新命令

- Cluster 1** (CDH 5.11.2, Parcel) [dropdown]
- 主机 [dropdown]
- HBase **1** [dropdown]
- HDFS **1** **2** [dropdown]
- Hive [dropdown]
- Hue **1** [dropdown]
- Impala [dropdown]
- Key-Value ... [dropdown]
- Oozie [dropdown]
- Solr [dropdown]
- Spark [dropdown]
- YARN (MR...) **1** [dropdown]
- ZooKeeper **1** [dropdown]

## Cloudera Management Service

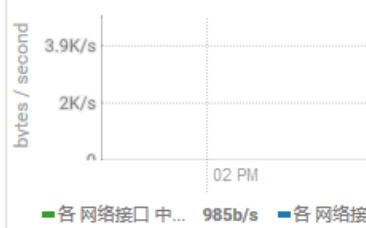
- Cloudera ... **4** [dropdown]

## 图表

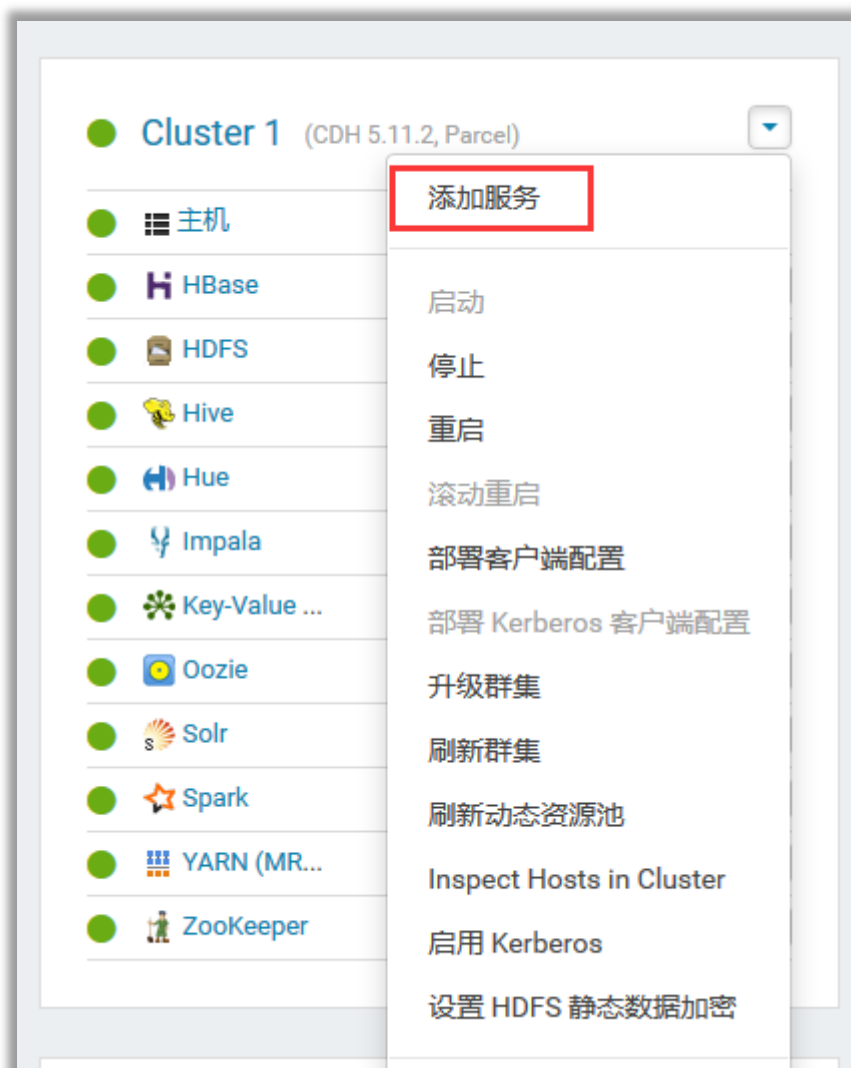
群集 CPU



群集网络 IO



## 添加 KAFKA 服务



<input type="radio"/>	Java KeyStore KMS	The Hadoop Key Management Service with file-based Java KeyStore. Maintains a single copy of keys, using simple password-based protection. Requires CDH 5.3+. <b>Not recommended for production use.</b>
<input checked="" type="radio"/>	Kafka	Apache Kafka is publish-subscribe messaging rethought as a distributed commit log. <b>Before adding this service, ensure that either the Kafka parcel is activated or the Kafka package is installed.</b>
<input type="radio"/>	Key-Value Store Indexer	键/值 Store Indexer 侦听 HBase 中所含表内的数据变化，并使用 Solr 为其创建索引。



## 将 Kafka 服务添加到 Cluster 1

### 自定义 Kafka 的角色分配

您可以在此处自定义新服务的角色分配，但请注意，如果分配不正确（例如，分配到某个主机上的角色太多），性能受到影响。

还可以按主机查看角色分配。 [按主机查看](#)

<b>KB</b> Kafka Broker x 1 新建	<b>KMM</b> Kafka MirrorMaker	<b>G</b> Gateway
linux176 ▾	选择主机	选择主机

返回

1 2 3 4 5 6

继续

## 将 Kafka 服务添加到 Cluster 1

### 审核更改

<b>ZooKeeper Root</b> zookeeper.chroot	<b>Kafka (服务范围)</b> C	?
	<input type="text" value="/kafka"/>	
<b>Enable Kerberos Authentication</b> kerberos.auth.enable	<input type="checkbox"/> Kafka (服务范围)	?
<b>Topic Auto Creation</b> auto.create.topics.enable	<input checked="" type="checkbox"/> Kafka (服务范围)	?
<b>Default Replication Factor</b> default.replication.factor	<input type="text" value="1"/>	?
<b>Offset Commit Topic Number of Partitions</b> offsets.topic.num.partitions	<input type="text" value="50"/>	?
<b>Offset Commit Topic Replication Factor</b> offsets.topic.replication.factor	<input type="text" value="3"/>	?
<b>Enable Kafka Monitoring (Note: Requires Kafka-1.3.0 parcel or higher)</b>	<input checked="" type="checkbox"/> Kafka (服务范围)	?

返回

1 2 3 4 5

继续

启动 Kafka Kafka 10月 13, 2:38:15 下午 17.69s  
 Failed to start service.

正在启动服务上的 1 角色 10月 13, 2:38:15 下午 17.69s  
 服务未成功启动; 所需角色未全部启动: 仅启动了 1 个角色中的 0 个。原因: Service has only 0 Kafka Broker roles running instead of minimum required 1.

对角色 Kafka Broker (linux176) 执行命令 启动此 Kafka Broker Kafka Broker (linux176) 10月 13, 2:38:15 下午 17.68s  
 Failed to start role.

启动角色 Kafka Broker (linux176) 10月 13, 2:38:15 下午 17.68s  
 Supervisor 返回 FATAL。请参阅角色日志文件、Stderr 或 Stdout.

\$> csd/csd.sh ["start"] stdout stderr 角色日志

第一次启动如果出错请暂时回到集群主页

**cloudera MANAGER**

## 将 Kafka 服务添加到 Cluster 1

首次运行 命令

状态: **失败** 开始时间: 10月 13, 2:37:58 下午 持续时间: 34.01s 重试

Failed to perform First Run of services.

详细信息 已完成 3 个步骤 (共 3 个)。

步骤	上下文
> <span style="color: green;">✔</span> 并行运行 1 步骤 已成功完成 1 个步骤。	
> <span style="color: green;">✔</span> 正在部署客户端配置 Successfully deployed all client configurations.	<a href="#">Cluster 1</a>
v <span style="color: red;">✘</span> 启动 Kafka Failed to start service.	<a href="#">Kafka</a>
v <span style="color: red;">✘</span> 正在启动服务上的 1 角色 服务未成功启动; 所需角色未全部启动: 仅启动了 1 个角色中的 0 个。原因: Service has only 0 Kafka Broker roles running instead of minimum required 1.	
v <span style="color: red;">✘</span> 对角色 Kafka Broker (linux176) 执行命令 启动此 Kafka Broker Failed to start role.	<a href="#">Kafka Broker (linux176)</a>

进入 kafka , 点击 “配置” , 点击警告



A screenshot of a Kafka configuration status table. The table has a header '状态' (Status) and a sub-header '警告' (Warning). The '警告' row is highlighted with a red box and shows a count of 1. Other rows include '错误' (Error) with 0, '已编辑' (Edited) with 0, '非默认' (Non-default) with 5, and '包含覆盖项' (Contains overrides) with 1.

状态	数量
错误	0
警告	1
已编辑	0
非默认	5
包含覆盖项	1



A screenshot of the Kafka configuration page for the 'broker\_max\_heap\_size' parameter. The parameter is set to 50兆字节 (50 MB). A red arrow points to the 'Kafka Broker Default Group' dropdown menu, which is highlighted with a red box. Below the configuration, there is a note: '50 小于 256 建议的最小值。 抑制...' (50 is less than the recommended minimum of 256. Suppress...).

然后保存更改，最后重启 kafka 服务即可正常开启

## ✔ 重启 命令

状态: 已完成    上下文: [Kafka](#)    开始时间: 10月 13, 2:50:45 下午    持续时间: 22.21s

Successfully restarted service.

详细信息 已完成 1 个步骤 (共 1 个)。

### 步骤

- ✔ 执行服务 Kafka 上的命令 启动  
Successfully started service.
- ✔ 正在启动服务上的 1 角色  
已成功启动服务上的 1 角色。
  - ✔ 对角色 Kafka Broker (linux176) 执行命令 启动此 Kafka Broker  
Successfully started process.

# 主页

状态 所有运行状况问题 配置 ▾ 所有最新命令

- Cluster 1 (CDH 5.11.2, Parcel) ▾
- 主机
- HBase ▾
- HDFS ▾
- Hive ▾
- Hue ▾
- Impala ▾
- Kafka ▾
- Key-Value ... ▾
- Oozie ▾
- Solr ▾
- Spark ▾
- YARN (MR... ▾
- ZooKeeper ▾