



在 Cloudera Enterprise 5.x 中安装 Kafka

概述

Kafka 一般是建议单独部署的，因为 Kafka 的数据缓存也是基于磁盘的，所以不建议在生产环境中与 HDFS 共存。

操作步骤

1. 在 httpd 服务器目录下创建 kafka 的安装包目录，并将 kafka 的安装包文件拷贝到新创建的目录下

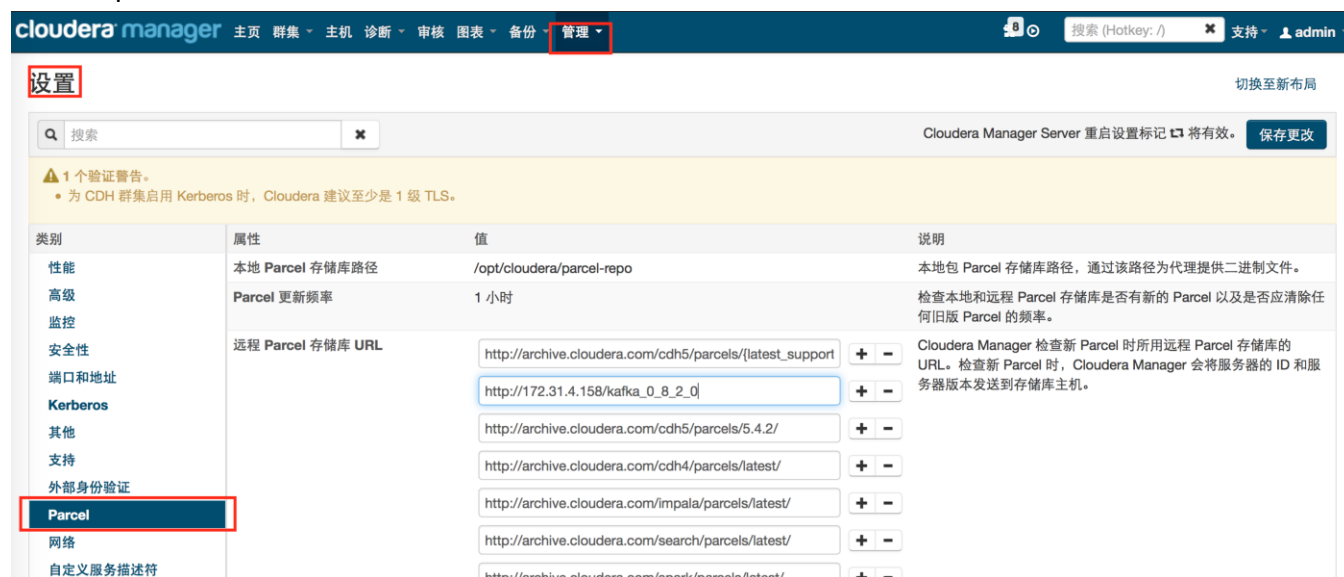
```
cd /var/www/html
```

```
mkdir kafka_0_8_2_0
```

```
cp/path/from/kafka-0.8.2.0-1.kafka1.3.1.p0.9-el6.parcel/var/www/html/kafka_0_8_2_0
```

```
cp/path/from/manifest.json /var/www/html/kafka_0_8_2_0
```

2. 添加 parcel 源



cloudera manager 主页 群集 主机 诊断 审核 图表 备份 管理

搜索 (Hotkey: /) 支持 admin

设置 切换至新布局

1 个验证警告。
• 为 CDH 群集启用 Kerberos 时，Cloudera 建议至少是 1 级 TLS。

类别	属性	值	说明
性能	本地 Parcel 存储库路径	/opt/cloudera/parcel-repo	本地包 Parcel 存储库路径，通过该路径为代理提供二进制文件。
高级	Parcel 更新频率	1 小时	检查本地和远程 Parcel 存储库是否有新的 Parcel 以及是否应清除任何旧版 Parcel 的频率。
监控	远程 Parcel 存储库 URL	http://archive.cloudera.com/cdh5/parcels/latest_support + - http://172.31.4.158/kafka_0_8_2_0 + - http://archive.cloudera.com/cdh5/parcels/5.4.2/ + - http://archive.cloudera.com/cdh4/parcels/latest/ + - http://archive.cloudera.com/impala/parcels/latest/ + - http://archive.cloudera.com/search/parcels/latest/ + - http://archive.cloudera.com/snark/parcels/latest/ + -	Cloudera Manager 检查新 Parcel 时所用远程 Parcel 存储库的 URL。检查新 Parcel 时，Cloudera Manager 会将服务器的 ID 和服务版本发送到存储库主机。

Parcel

3. 将 parcel 源中的安装包下载到 cloudera 的存储库

cloudera manager 主页 群集 主机 诊断 审核 图表 备份 管理 搜索 (Hotkey: /) 支持 admin

主机 状态 配置 模板 磁盘概述 **Parcel**

Parcel

位置

Cluster 1
远程提供

筛选器

PARCEL 名称

全部 8
CDH 4 1
CDH 5 4
IMPALA 1
KAFKA 1
KETTREL 1

状态

全部 1
其他 1
已分配 0

Cluster 1

Parcel 名称	版本	状态	操作
KAFKA	0.8.2.0-1.kafka1.3.1.p0.9	远程提供	下载

Parcel 使用情况 编辑设置 检查新 Parcel

4. 在集群中添加服务 Kafka 服务

cloudera manager 主页 群集 主机 诊断 审核 图表 备份 管理 搜索 (Hotkey: /) 支持 admin

30 分钟 在 2015年9月8日, 下午1点57 UTC 之前 添加群集

主页 状态 所有运行状况问题 配置 9 所有最新命令

Cluster 1 (CDH 5.4.5, Parcel)

主机 添加服务

启动
停止
重启
滚动重启
部署客户端配置
部署 Kerberos 客户端配置
升级群集
刷新群集
刷新动态资源池
启用 Kerberos
Host Inspector (Cluster)

查看客户端配置 URL

重命名群集

进入维护模式

视图维护模式状态

图表

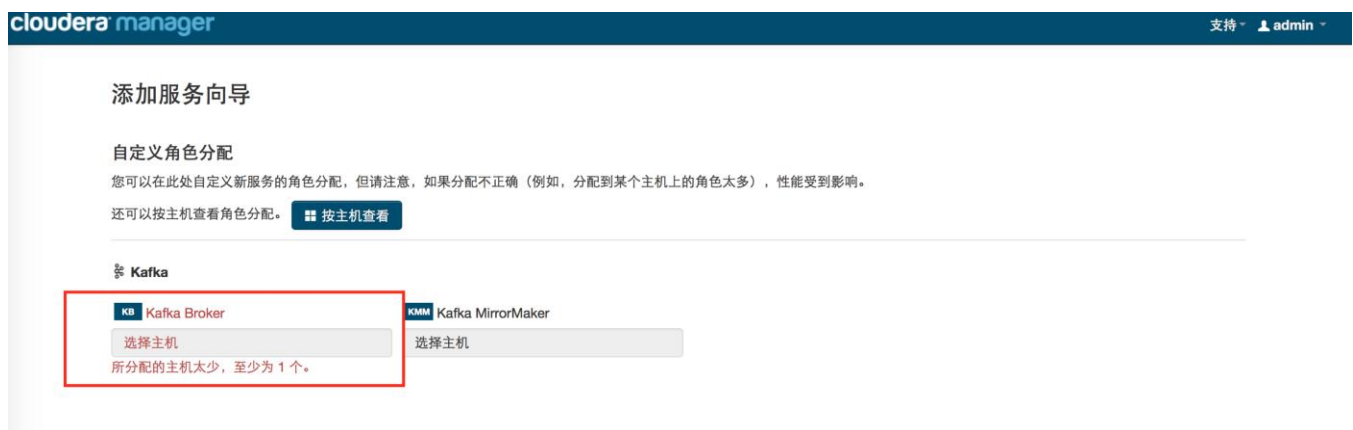
群集网络 IO

群集磁盘 IO

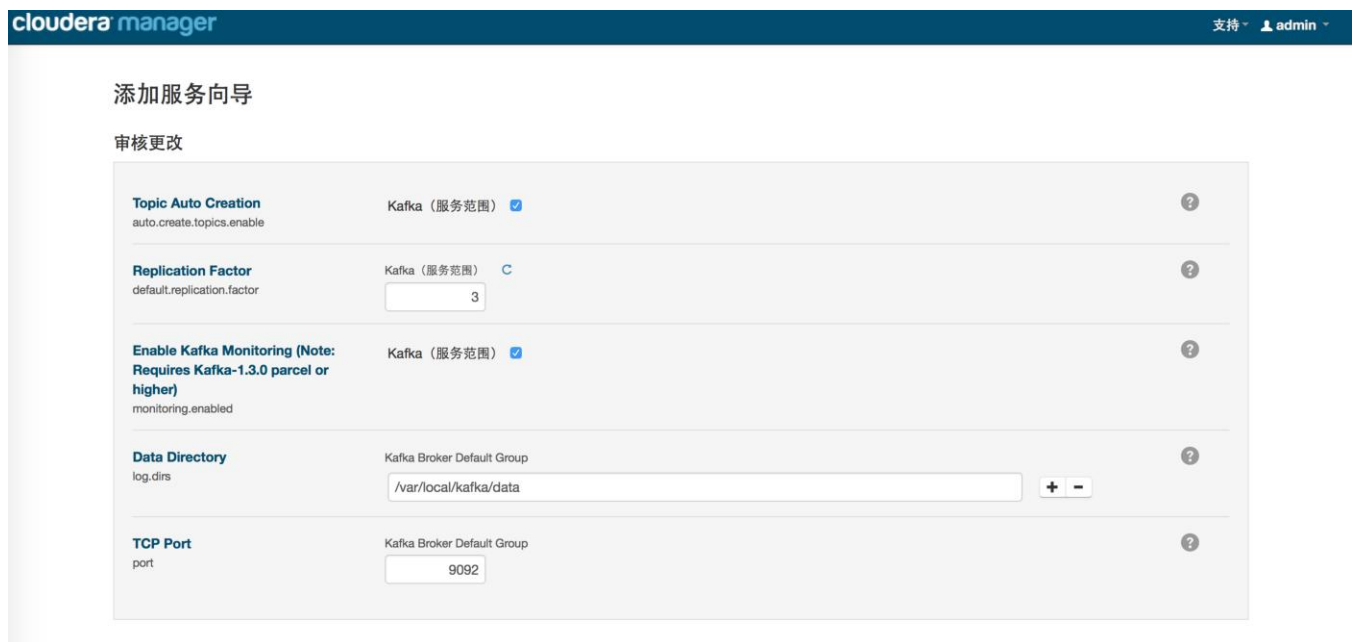
群集 CPU

HDFS IO

5. 选择需要安装的 Kafka Broker 节点；暂时不需要 MirrorMaker（这是一个 Kafka 集群备份的功能）



6. 将 Kafka 的 replication 设置成 3，然后再配置一下它的“Data Directory”；Data Directory 一般需要配置多个。



7. 启动服务即可