**Cloudera Manager分析**

**目录**

[1. 相关目录](https://github.com/zouhc/MyHadoop/blob/master/doc/Cloudera%20Manager%E5%88%86%E6%9E%90.md#1)

[2. 配置](https://github.com/zouhc/MyHadoop/blob/master/doc/Cloudera%20Manager%E5%88%86%E6%9E%90.md#2)

[3. 数据库](https://github.com/zouhc/MyHadoop/blob/master/doc/Cloudera%20Manager%E5%88%86%E6%9E%90.md#3)

[4. CM结构](https://github.com/zouhc/MyHadoop/blob/master/doc/Cloudera%20Manager%E5%88%86%E6%9E%90.md#4)

[5. 升级](https://github.com/zouhc/MyHadoop/blob/master/doc/Cloudera%20Manager%E5%88%86%E6%9E%90.md#5)

[6. 卸载](https://github.com/zouhc/MyHadoop/blob/master/doc/Cloudera%20Manager%E5%88%86%E6%9E%90.md#6)

[7. 开启postgresql远程访问](https://github.com/zouhc/MyHadoop/blob/master/doc/Cloudera%20Manager%E5%88%86%E6%9E%90.md#7)

**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**](http://lib.csdn.net/base/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**](http://lib.csdn.net/base/mysql)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**](http://lib.csdn.net/base/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）。它主要做三件事件：

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**](http://lib.csdn.net/base/python)调用相应的服务shell脚本。Server端为[**Java**](http://lib.csdn.net/base/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。