FK集群升级方案

目录

[1 旧集群版本 3](#_Toc455475469)

[2 目标集群版本 3](#_Toc455475470)

[3 升级方案 3](#_Toc455475471)

[4 新集群安装 4](#_Toc455475472)

[4.1 离线环境 4](#_Toc455475473)

[4.2 安装包 4](#_Toc455475474)

[4.3 安装组件 5](#_Toc455475475)

[5 数据迁移 5](#_Toc455475476)

[5.1 HDFS 数据 5](#_Toc455475477)

[5.2 ORACLE 数据 5](#_Toc455475478)

[6 新集群部署并启动应用系统 5](#_Toc455475479)

[7 FAQ 5](#_Toc455475480)

# 旧集群版本

Cloudera manager 5.2.0

# 目标集群版本

Cloudera manager 5.x.y

# 升级方案

1. 拷贝必要的数据，其中dwd和dm

其生成拷贝命令脚本,

Dwd数据：

全局一表：（拷贝昨天）

DWD\_IMSI\_IMEI\_MAPPING

DWD\_IMSI\_IMEI\_MAPPING\_DAY

DWD\_IMSI\_MSISDN\_MAPPING

DWD\_INBJ\_IMSI\_MSISDN\_MAPPING

DWD\_IN\_OUT\_BEIJING

DWD\_MSISDN\_PREFIX

DWD\_PUBLIC\_IMEI

DWD\_PUBLIC\_IMSI

DWD\_PUBLIC\_MSISDN

DWD\_RELATION\_INFO\_ALL

每日一表：（拷贝90天数据）：

DWD\_FOREIGN\_POSITION\_INFO

DWD\_IMSI\_DAY\_KEEP\_ALIVE

DWD\_IMSI\_KEEP\_ALIVE

DWD\_INBJ\_RELATION\_INFO

DWD\_POSITION\_INFO

DWD\_RELATION\_INFO

DWD\_RELATION\_ZIP\_INFO

DWD\_ROAM\_STAT

DWD\_SILENT\_RELATION\_INFO

Dm数据：

码表：由脚本从oracle同步

|  |  |
| --- | --- |
| 脚本 | HDFS |
| Dm\_distance\_day.pig | 标准化coninfo目录(90) |
| dm\_political\_core\_area.pig | /hdata/fkv2/dm/political\_filter 90天数据 |
| dm\_imei\_imsi\_chang\_new\_oneday.pig | /hdata/fkv2/dm/imei\_imsi\_change |
| dm\_zero\_second\_relation.pig | /hdata/fkv2/dm/zero\_second\_filter[90] |
| sms\_one\_day.pig | /hdata/fkv2/bakdata/$day/coninfo/smdr(90) |
| dm\_relation\_combine.pig | /hdata/fkv2/src/new |
| 935 | /hdata/fkv2/bakdata/$day/express/all (90) |
| dm\_relation\_combine.pig | /hdata/fkv2/dm/code\_case\_manage |

依赖fkv2.0-edw\_keyword流程，生成/hdata/fkv2/srcdata目录数据

2、为了数据安全和现有系统稳定，先安装新集群，将旧集群数据拷贝到新集群，并在新集群搭建现有系统运行，直到新集群能够与旧集群数据保持一致，停掉旧集群。

# 新集群安装

## 操作系统配置

**来自cloudera admin文档，以下操作都是在root用户**

### 配置集群NTP服务

指定集群中的一台为NTPServer

NTPServer配置：只需修改/etc/ntp.conf文件中的如下行即可，

server **193.1.10.225** prefer

其他及其的ntp服务修改其红色的ip到NTPServer IP即可

修改完成后，首次需手动同步时间，执行如下脚本即可(客户机安装即可)：



必须先安装sshpass（解压执行依次执行即可A、./configure B、make && make install）,后执行（./syncDate.sh NtpServerIP root用户密码 hostsfile）



查看ntp同步状态(执行命令：./syncNTP.sh root用户密码 hostsfile)：



### 关闭THP服务

如果存在，不需要重启

echo never > /sys/kernel/mm/redhat\_transparent\_hugepage/enabled

echo never > /sys/kernel/mm/redhat\_transparent\_hugepage/defrag

## 离线环境

Apache 提供一个离线服务器，将要安装的安装包下载后，拷贝到apache服务中。

## 安装包

* cm：cloudera-manager-el6-cm5.x.y\_x86\_64.tar.gz
* cdh：parcels(5.x.y)
* cloudera-cdh.repo
* cloudera-manager.repo
* RPM-GPG-KEY-cloudera

注：将以上文件解压缩放在/var/www/html目录下，目录结构如下：

cdh5/parcels/5.7/

CDH-5.x.y-1.cdh5.x.y.p0.45-el6.parcel

CDH-5.x.y-1.cdh5.x.y.p0.45-el6.parcel.sha

manifest.json

cm5

## 安装组件

* HDFS （HA）
* YARN （HA）
* ZOOKEEPER
* SPARK (on yarn)
* OOZIE
* HUE
* HIVE

# 数据迁移

## HDFS 数据

执行命令：Hadoop distcp –i –update -skipcrccheck hftp://oldnamenode:50070/PATH /PATA

## ORACLE 数据

TIP/DM\_TIP用户下的表

负责人：十所

# 新集群部署并启动应用系统

* 标准化数据服务（十所）
* Fk应用数据服务（东信）

Dwd

Dm

# FAQ