北京三期数据迁移使用说明

目录

[北京三期数据迁移使用说明 1](#_Toc496522924)

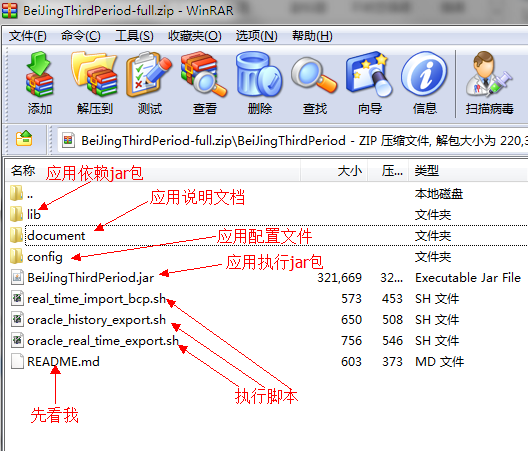
[一、 压缩包内容说明： 2](#_Toc496522925)

[二、修改配置文件 2](#_Toc496522926)

[二、 部署说明 6](#_Toc496522927)

## 压缩包内容说明：

项目压缩包名：BeiJingThirdPeriod-full.zip



可执行脚本说明：

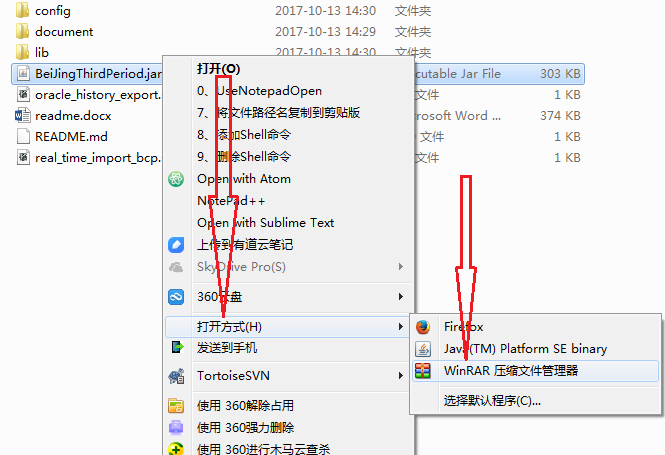
oracle\_history\_export.sh Oracle历史数据导入脚本

oracle\_real\_time\_export.sh Oracle数据实时导入脚本

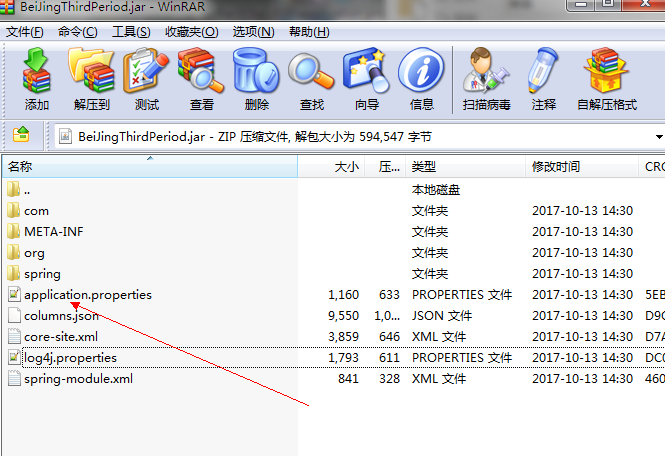
real\_time\_import\_bcp.sh Bcp文件实时导入脚本

## 二、修改配置文件

1. 用解压缩软件打开Jar包



1. 双击修改application.properties配置文件



根据情况主要修改以下几个配置项:

**Solr单机版配置信息:**

solr.url: Solr单机版连接地址

**Solr集群版配置信息:**

zkHost:大数据集群Zookeeper服务所在的节点名

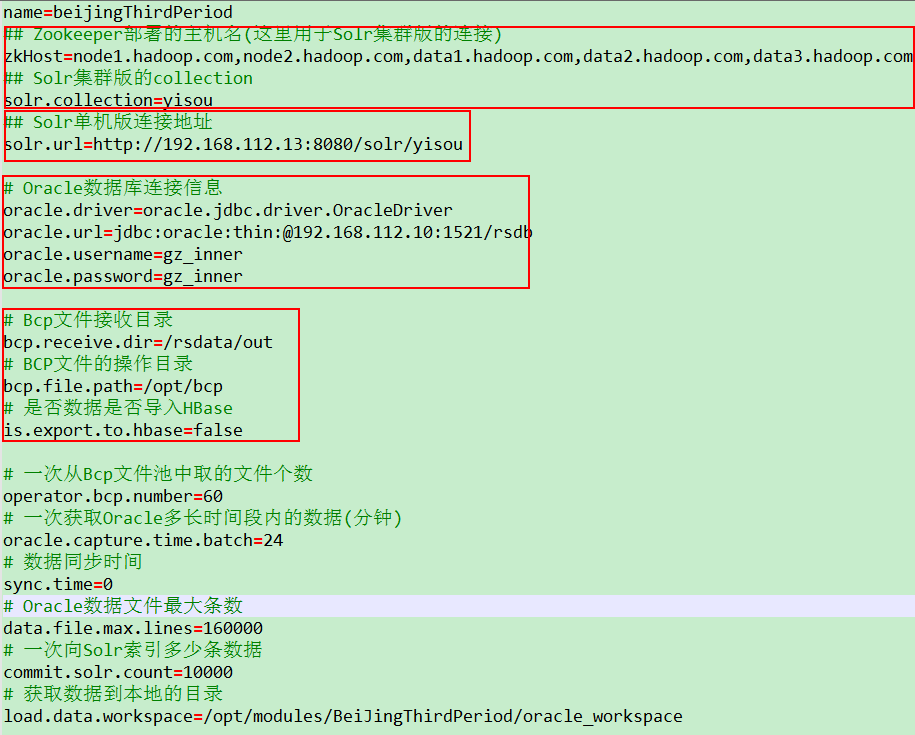
solr.collection:Solr当前使用的collection(一般不变)

**Oracle连接信息:**

oracle.url:Oracle连接地址

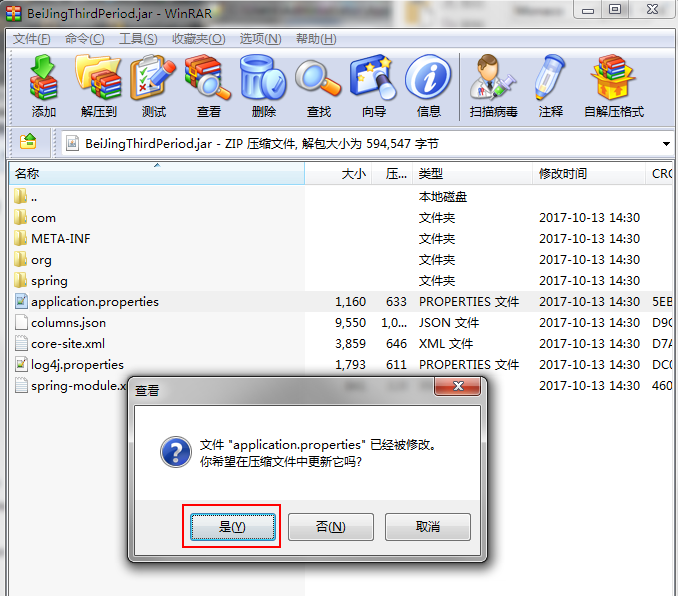
oracle.username:Oracle用户名

oracle.password:Oracle密码



其它的配置项使用默认的就行

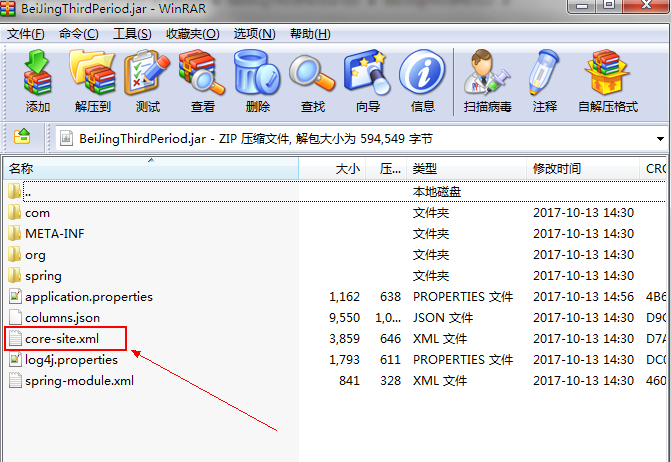
1. Ctrl + s保存配置文件解压缩软件会提示你是否希望在压缩文件中更新它，点是



1. 修改配置文件：可以双击打开修改过直接Ctrl + S 保存，也可以把它拖出来，修改，修改完后再拖进行.
2. 替换jar包中core-site.xml将集群的core-site.xml文件配置jar包内的此文件

集群上此文件的位置一般为:

/opt/cloudera/parcels/CDH/lib/hadoop/etc/hadoop/core-site.xml



## 部署说明

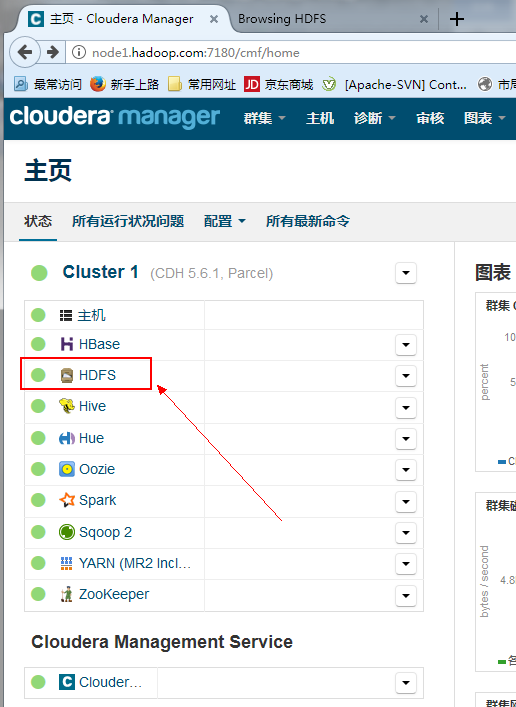
1. 将BeiJingThirdPeriod-full.zip上传到服务器(必须)
2. 创建目录(必须)

**mkdir** -p **/**opt**/**modules

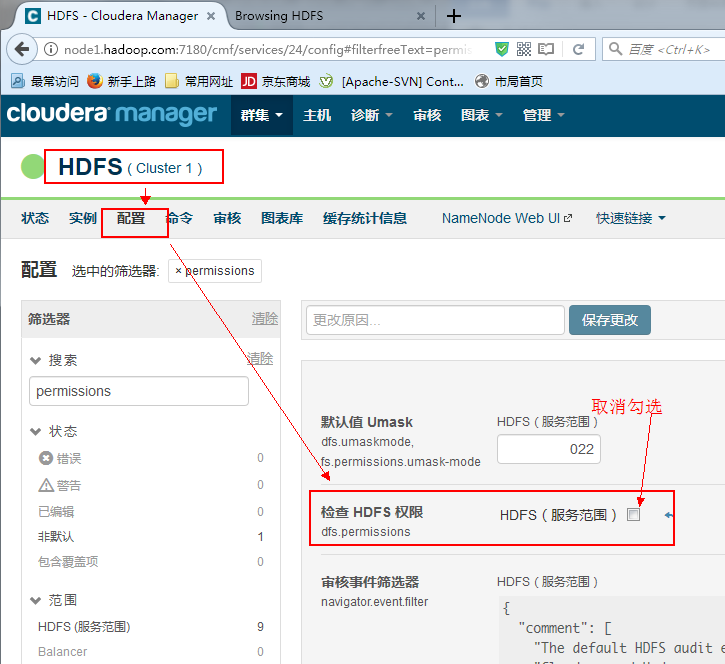
1. 解压压缩包(必须)

**unzip** BeiJingThirdPeriod-full.zip -d **/**opt**/**modules**/**

1. 将你在Windows平台下修改过配置文件的jar包替换刚刚解压的jar包(必须)
2. ClouderaManager集群关闭HDFS服务的权限检查(集群版)
3. 登陆到ClouderaManager的管理界面



1. 关闭HDFS的权限检查



1. 重启集群让配置生效



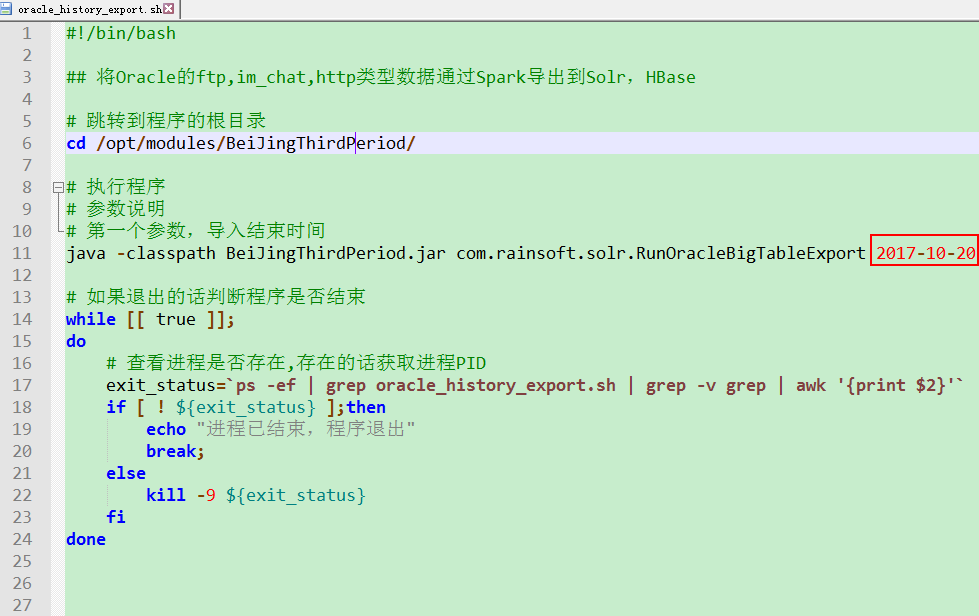
1. 跑Bcp实时任务(根据情况确定执行)

**nohup** sh **/**opt**/**modules**/**BeiJingThirdPeriod**/**real\_time\_import\_bcp.sh **&**

1. 从Oracle取数据跑历史任务(根据情况执行)
   1. 修改历史任务脚本:

**vim** oracle\_history\_export.sh

* 1. 修改oracle\_history\_export.sh历史任务的结束时间



参数:导入结束时间

* 1. 执行历史任务脚本

**nohup** sh **/**opt**/**modules**/**BeiJingThirdPeriod**/**oracle\_history\_export.sh **&**

1. 从Oracle取数据跑实时任务
   1. 看之前是否导出过数据,如果没有导出过的话此步跳过，如果导出过的话执行以下步骤:
      1. 根据情况修改index-record.txt



根据前面的Key将日期修改成你之前最后导入的日期

Key的前半部分是表名

Key和Value以制表符分隔,千万不要变成空格

* + 1. 将修改后的文件放在以下目录:

/opt/modules/BeiJingThirdPeriod/createIndexRecord/index-record.txt

* 1. 执行Oracle实时任务脚本:

**nohup** sh **/**opt**/**modules**/**BeiJingThirdPeriod**/**oracle\_real\_time\_export.sh **&**

1. 查看程序是否正常执行:可以查看执行日志:

**vim** **/**opt**/**modules**/**BeiJingThirdPeriod**/**logs**/**log.log