**Hadoop安装与配置**

# 安装方式概述

hbase和hadoop一样也分为单机版、伪分布式版和完全分布式集群版本，这篇文件介绍如何搭建完全分布式集群环境搭建。

hbase依赖于hadoop环境，搭建habase之前首先需要搭建好hadoop的完全集群环境。**本文中没有按照独立的zookeeper，使用了hbase自带的zookeeper**。

# 环境准备

hbase软件包: http://mirror.bit.edu.cn/apache/hbase/1.3.1/hbase-1.3.1-bin.tar.gz

Hadoop集群（安装方式参考《hadoop集群搭建.docx》文档）

# 安装步骤

## 下载解压

首先在hadoop-master节点安装配置好HBase之后，在复制到从节点

wget <http://mirror.bit.edu.cn/apache/hbase/1.3.1/hbase-1.3.1-bin.tar.gz>

解压

tar -xzvf hbase-1.3.1-bin.tar.gz -C /usr/local/

重命名

mv hbase-1.3.1 hbase

## 配置环境变量

vi /etc/profile

内容如下：

export HBASE\_HOME=/usr/local/hbase

export PATH=$HBASE\_HOME/bin:$PATH

使立即生效

source /etc/profile

修改系统变量ulimit

ulimit -n 10240

## 修改配置文件

hbase 相关的配置主要包括hbase-env.sh、hbase-site.xml、regionservers三个文件，都在 /usr/local/hbase/conf目录下面：

配置hbase-env.sh

vi hbase-env.sh

内容如下

export JAVA\_HOME=/usr/lib/jvm/jre-1.7.0-openjdk.x86\_64

export HBASE\_CLASSPATH=/usr/local/hbase/conf

# 此配置信息，设置由hbase自己管理zookeeper，不需要单独的zookeeper。

export HBASE\_MANAGES\_ZK=true

export HBASE\_HOME=/usr/local/hbase

export HADOOP\_HOME=/usr/local/hadoop

#Hbase日志目录

export HBASE\_LOG\_DIR=/usr/local/hbase/logs

配置 hbase-site.xml

<configuration>

<property>

<name>hbase.rootdir</name>

<value>hdfs://hadoop-master:9000/hbase</value>

</property>

<property>

<name>hbase.cluster.distributed</name>

<value>true</value>

</property>

<property>

<name>hbase.master</name>

<value>hadoop-master:60000</value>

</property>

<property>

<name>hbase.zookeeper.quorum</name>

<value>hadoop-master,hadoop-slave1,hadoop-slave2,hadoop-slave3</value>

</property>

</configuration>

配置regionservers

vi /usr/local/hbase/conf/regionservers

内容如下

hadoop-master

hadoop-slave1

hadoop-slave2

hadoop-slave3

复制hbase到从节点中

scp -r /usr/local/hbase hadoop-slave1:/usr/local/

scp -r /usr/local/hbase hadoop-slave2:/usr/local/

scp -r /usr/local/hbase hadoop-slave3:/usr/local/

## 启动hbase

启动仅在master节点上执行即可

~/hbase/bin/start-hbase.sh

启动后，master上进程和slave进程列表

master中执行jps

[hadoop@master ~]$ jps

6225 Jps

2897 SecondaryNameNode # hadoop进程

2710 NameNode # hadoop master进程

3035 ResourceManager # hadoop进程

5471 HMaster # hbase master进程

2543 HQuorumPeer # zookeeper进程

salve中执行jps

[hadoop@slave1 ~]$ jps

4689 Jps

2533 HQuorumPeer # zookeeper进程

2589 DataNode # hadoop slave进程

4143 HRegionServer # hbase slave进程

因为hbase依赖于hadoop，因此启动和停止都是需要按照顺序进行

如果安装了独立的zookeeper

启动顺序: hadoop-> zookeeper-> hbase  
停止顺序：hbase-> zookeeper-> hadoop

使用自带的zookeeper

启动顺序: hadoop-> hbase  
停止顺序：hbase-> hadoop

重启hbase

stop-hbase.sh

stop-all.sh

start-all.sh

start-hbase.sh

# 错误处理

在搭建的过程中，报了这么一个错误，错误信息如下：

Unhandled: org.apache.hadoop.hbase.ClockOutOfSyncException: Server hadoop-slave3,16020,1500526355333

Caused by: org.apache.hadoop.hbase.ipc.RemoteWithExtrasException(org.apache.hadoop.hbase.ClockOutOfSyncException):

org.apache.hadoop.hbase.ClockOutOfSyncException:

Server hadoop-slave3,16020,1500526355333 has been rejected; Reported time is too far out of sync with

master. Time difference of 77348ms > max allowed of 30000ms

看大概的意思是主节点连接从节点超时了。可能有两方面的原因，第一、linux服务器时间不一致导致，第二、由于网络其它原因导致连接的时间超长。

解决方案：

1、修改各服务器时间保持一致。最终的解决方案是：设置一个定时使用ntp从某个服务器定时同步时间。

参考《ntp时钟同步.docx》文档。

2、可以修改hbase默认的最大链接时间长一些。

HBase配置文件hbase-siter.xml中添加连接时长的属性：

<property>

<name>hbase.master.maxclockskew</name>

<value>120000</value>

</property>

参考:http://www.ixirong.com/2015/05/25/how-to-install-hbase-cluster/