**Impala安装与配置**

# Impala组件介绍

1.impala-state-store   -- 协调服务

2.impala-catalog --名称服务

3.impala-server(impalad)  节点服务

4.impala-shell 交互服务

其中 state-store和catalog类似于 namenode一般部署在主节点，而impala-server类似于datenode部署在子节点上

最后impala-shell可以部署在任意几点上。

# 安装技巧

由于使用的是yum在线安装，这就涉及到一个下载软件包的问题。由于国内网络环境的原因，下载超慢的。

所以，先安装一个节点，不要带参数 -y。

下载完成后先不要输入 y 确认安装，因为安装后包会被删除。

在确认安装前，先到 /var/cache/yum/...目录下把安装包拷贝到其它待安装节点，这样其它节点安装时就不用再下载一次了。

# Impala安装

## yum更新

**yum更新每个节点都要操作**

下述操作可概括为：添加yum源、更新yum

### 配置cloudera源

参考《yum使用.docx》文档第三章节“配置cloudera源”

### 更新yum库

执行命令

yum clean all

yum update

yum list |grep impala 确认是否cdh5.13.3版本

yum install impala

应当是如下这种cdh为5.13.3的版本：

impala.x86\_64 2.10.0+cdh5.13.3+0-1.cdh5.13.3.p0.2.el6

impala-server.x86\_64 2.10.0+cdh5.13.3+0-1.cdh5.13.3.p0.2.el6

impala-shell.x86\_64 2.10.0+cdh5.13.3+0-1.cdh5.13.3.p0.2.el6

impala-udf-devel.x86\_64 2.10.0+cdh5.13.3+0-1.cdh5.13.3.p0.2.el6

## 安装 Impala协调服务

执行命令

yum install impala-state-store -y

yum install impala-catalog -y

# yum impala-shell -y 此服务一般安装在计算节点，当然，安装在协调节点也行

# yum mpala-server -y 此服务一般安装在计算节点，当然，安装在协调节点也行

/etc/profile中添加export IMPALA\_HOME=/usr/lib/impala

执行

## 安装 Impala计算服务

执行命令

yum install impala-server -y

yum install impala-shell -y

在计算节点中配置协调节点信息

修改/etc/default/impala，把里面的IMPALA\_CATALOG\_SERVICE\_HOST= 和 IMPALA\_STATE\_STORE\_HOST= 设置为“协调服务”所在的机器IP

/etc/profile中添加export IMPALA\_HOME=/usr/lib/impala

执行 source /etc/profile

## 配置元数据库

拷贝 hadoop配置文件core-site.xml、hdfs-site.xml和hive配置文件hive-site.xml到impala配置文件目录：/etc/impala/conf/

修改hive-site.xml里的连接方式为 通过rpc连接hive metastore的方式----参考《hive安装.docx》第3章节。

## 启动impala

启动各组件要按照如下顺序：

service impala-state-store start

service impala-catalog start

service impala-server start

# 与Kudu集成

修改 /etc/default/impala中IMPALA\_SERVER\_ARGS配置项为如下标红处所示：

IMPALA\_SERVER\_ARGS=" \

-log\_dir=${IMPALA\_LOG\_DIR} \

-catalog\_service\_host=${IMPALA\_CATALOG\_SERVICE\_HOST} \

-state\_store\_port=${IMPALA\_STATE\_STORE\_PORT} \

-use\_statestore \

-state\_store\_host=${IMPALA\_STATE\_STORE\_HOST} \

-be\_port=${IMPALA\_BACKEND\_PORT} \

-kudu\_master\_hosts=kudu1:7051,kudu2:7051,kudu3:7051"

# 对其它组件的影响

由于安装impala的时候，会安装其依赖的hadoop和hive，这就导致这两个组件的命令调用的位置发生了变化。以hadoop为例，原来执行hdfs命令，调用的是$HADOOP\_HOME/sbin下的hdfs脚本；安装impala之后，随之安装的hadoop会在/bin下创建hdfs脚本，且/bin下的命令执行是优先于$HADOOP\_HOME/sbin下的命令。

所以，解决办法有两个：

1. 删除/bin下的hadoop和hive相关的所有命令；
2. 安装impala的时候，跳过hadoop和hive。

# 安装报错 解决

## 计算节点启动报错

报错：Impalad services did not start correctly,exiting.Error:Couldn't open transport for impala2:24000 (connect() failed: Connection refused)

分析：“节点服务”肯定要连接到“协调服务”和“名称服务”

解决：参考 3.3章节 “在计算节点中配置协调节点信息”描述的操作。

## 安装报错

Error:Package:impala-2.10.0+cdh5.13.3+0-1.cdh5.13.3.p0.2.el7.x86\_64(cloudera-cdh)

Requires: /lib/lsb/init-functions

......

解决：

yum install redhat-lsb

## 创建数据库报错

报错：[impala5:21000] > create database impalaTest;

Query: create database impalaTest

ERROR: ImpalaRuntimeException: Error making 'createDatabase' RPC to Hive Metastore:

CAUSED BY: MetaException: Got exception: org.apache.hadoop.security.AccessControlException Permission denied: user=impala, access=WRITE, inode="/user/hive/warehouse":root:supergroup:drwxr-xr-x

解决：hadoop的hdfs里用户权限问题，如果想设置权限，要开启ACL功能：

在hdfs-site.xml里添加：

<property>

<name>dfs.namenode.acls.enabled</name>

<value>true</value>

</property>

然后进行授权：切到 $hadoop/bin下执行

sudo -u root ./hdfs dfs -setfacl -m user:impala:rwx /user/hive/\*

这里的root应替换为/user/hive/\*所属用户名