mysql+hive单机配置

Hive 依赖于 HDFS 存储数据，Hive 将 HQL 转换成 MapReduce 执行，所以说 Hive 是基于 Hadoop 的一个数据仓库工具，实质就是一款基于 HDFS 的 MapReduce 计算框架，对存储在 HDFS 中的数据进行分析和管理。

1. **安装jdk，hadoop，mysql，hive，环境变量配置**

**java，hadoop，hive，mysql、start-yarn.sh、start-dfs.sh、start-all.sh、stop-all.sh命令直接敲就可以。不用再去找文件所在目录。**

[root@kafka Desktop]# **cat ~/.bash\_profile**

JAVA\_HOME=/opt/jdk

JRE\_HOME=/opt/jdk/jre

CLASS\_PATH=.:$JAVA\_HOME/lib/dt.jar:$JAVA\_HOME/lib/tools.jar:$JRE\_HOME/lib

PATH=$PATH:$JAVA\_HOME/bin:$JRE\_HOME/bin

export JAVA\_HOME JRE\_HOME CLASS\_PATH PATH

export HADOOP\_HOME=/u01/hadoop

export PATH=$PATH:$HADOOP\_HOME/bin:/$HADOOP\_HOME/sbin

export HIVE\_HOME=/u01/hive

export HIVE\_CONF\_DIR=$HIVE\_HOME/conf

export PATH=$PATH:$HIVE\_HOME/bin:$HIVE\_CONF\_DIR

export PATH=/u01/mysql/bin:$PATH

1. **安装mysql**
2. **下载**[https://dev.mysql.com/downloads/mysql/5.6.html#downloads](https://dev.mysql.com/downloads/mysql/5.6.html" \l "downloads)

2、创建mysql用户组及用户

**useradd mysql**

3、初始化mysql:

[root@11g Desktop]# **mkdir /u01/mysql/data**

[root@kafka support-files]# **/u01/mysql/bin/mysqld --initialize --user=mysql --datadir=/u01/mysql/data --basedir=/u01/mysql (会生成临时密码)**

4、修改配置文件mysql.server

[root@kafka support-files]# **vim /u01/mysql/support-files/mysql.server**

basedir=/u01/mysql

datadir=/u01/mysql/data

4、把mysql加到系统服务

[root@kafka Desktop]# **cp /u01/mysql/support-files/mysql.server /etc/init.d/mysql**

[root@kafka Desktop]# **service mysql start**

Starting MySQL. SUCCESS!

[root@kafka bin]# **mysql -u root -p**

Enter password:

mysql> **set password='12345';**

Query OK, 0 rows affected (0.00 sec)

mysql> **update user set host="%" where user="root";**

Query OK, 1 row affected (0.11 sec)

Rows matched: 1 Changed: 1 Warnings: 0

mysql> **flush privileges;**

Query OK, 0 rows affected (0.00 sec)

4 rows in set (0.00 sec)

mysql> **create database hive;**

Query OK, 1 row affected (0.01 sec)

**CentOS7安装mysql**

mkdir /u01/mysql/data

useradd mysql

/u01/mysql/bin/mysqld --initialize --user=mysql --datadir=/u01/mysql/data --basedir=/u01/mysql

mv /etc/my.cnf /etc/xxx

./mysql.server start

mysql -u root -p

**Centos7添加到系统服务**

[root@localhost support-files]# cp /u01/mysql/support-files/mysql.server /etc/init.d/mysql

[root@localhost support-files]# vim /usr/lib/systemd/system/mysql.service

[Unit]

Description=mysql

SourcePath=/etc/init.d/mysql

Before=shutdown.target

[Service]

User=mysql

Type=forking

ExecStart=/u01/mysql/support-files/mysql.server start

ExecStop=/u01/mysql/support-files/mysql.server stop

[Install]

WantedBy=multi-user.target

[root@localhost support-files]# systemctl start mysql

[root@localhost support-files]# systemctl stop mysql

**三、安装hadoop**

参考：<https://www.cnblogs.com/xuwujing/p/8017108.html>

1、创建目录

**mkdir /u01/cache**

**mkdir /u01/cache/tmp**

**mkdir /u01/cache/var**

**mkdir /u01/cache/dfs**

**mkdir /u01/cache/dfs/name**

**mkdir /u01/cache/dfs/data**

2、[root@kafka hadoop]# **vim /u01/hadoop/etc/hadoop/core-site.xml**

<configuration>

<property>

<name>hadoop.tmp.dir</name>

<value>/u01/cache/tmp</value>

<description>Abase for other temporary directories.</description>

</property>

<property>

<name>fs.default.name</name>

<value>hdfs://kafka:9000</value>

</property>

</configuration>

3、[root@kafka hadoop]# **vim /u01/hadoop/etc/hadoop/hdfs-site.xml**

<configuration>

<!-- 设置namenode的http通讯地址 -->

<property>

<name>dfs.namenode.http-address</name>

<value>kafka:50070</value>

</property>

<property>

<name>dfs.namenode.name.dir</name>

<value>/u01/cache/dfs/name</value>

</property>

<!-- 设置hdfs副本数量 -->

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

<!-- 设置datanode存放的路径 -->

<property>

<name>dfs.datanode.data.dir</name>

<value>/u01/cache/dfs/data</value>

</property>

</configuration>

4、[root@kafka hadoop]# **vim /u01/hadoop/etc/hadoop/mapred-site.xml**

<configuration>

<!-- 通知框架MR使用YARN -->

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

<property>

<name>mapreduce.application.classpath</name>

<value>

/u01/hadoop/etc/hadoop,

/u01/hadoop/share/hadoop/common/\*,

/u01/hadoop/share/hadoop/common/lib/\*,

/u01/hadoop/share/hadoop/hdfs/\*,

/u01/hadoop/share/hadoop/hdfs/lib/\*,

/u01/hadoop/share/hadoop/mapreduce/\*,

/u01/hadoop/share/hadoop/mapreduce/lib/\*,

/u01/hadoop/share/hadoop/yarn/\*,

/u01/hadoop/share/hadoop/yarn/lib/\*

</value>

</property>

</configuration>

5、[root@kafka Desktop]# **cat /u01/hadoop/etc/hadoop/workers**

192.168.60.1

6、[root@kafka hadoop]# v**im /u01/hadoop/etc/hadoop/hadoop-env.sh**

export JAVA\_HOME=/opt/jdk

7、[root@hive Desktop]# **ssh-keygen -t rsa**

[root@hive Desktop]# **cd /root/.ssh/**

[root@hive .ssh]# **cat id\_rsa.pub >> authorized\_keys**

[root@hive u01]# **vim /etc/ssh/sshd\_config**

PermitRootLogin yes

RSAAuthentication yes

PubkeyAuthentication yes

AuthorizedKeysFile .ssh/authorized\_keys

8、[root@kafka bin]# **hadoop namenode -format**

9、[root@kafka sbin]# **vim /u01/hadoop/sbin/start-dfs.sh/stop-dfs.sh**(修改两个文件，因为用到stop-dfs.sh时还是报相同的错误)

在文件开头加入：

HDFS\_DATANODE\_USER=root

HDFS\_DATANODE\_SECURE\_USER=hdfs

HDFS\_NAMENODE\_USER=root

HDFS\_SECONDARYNAMENODE\_USER=root

[root@kafka sbin]# **vim /u01/hadoop/sbin/start-yarn.sh/stop-yarn.sh**(修改两个文件，因为用到stop-yarn.sh时还是报相同的错误)

在文件开头加入：

YARN\_RESOURCEMANAGER\_USER=root

HADOOP\_SECURE\_DN\_USER=yarn

YARN\_NODEMANAGER\_USER=root

[root@kafka sbin]# **start-dfs.sh**

Starting namenodes on [kafka]

Starting datanodes

Starting secondary namenodes [kafka]

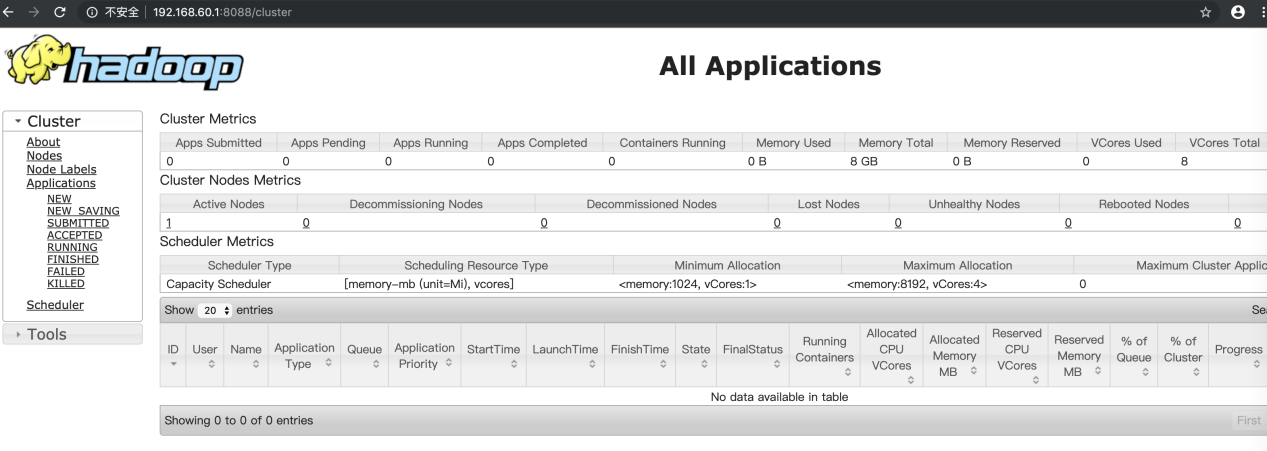
2019-10-14 14:09:40,075 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

[root@kafka sbin]# **start-yarn.sh**

Starting resourcemanager

Starting nodemanagers

hadoop页面



HDFS管理界面

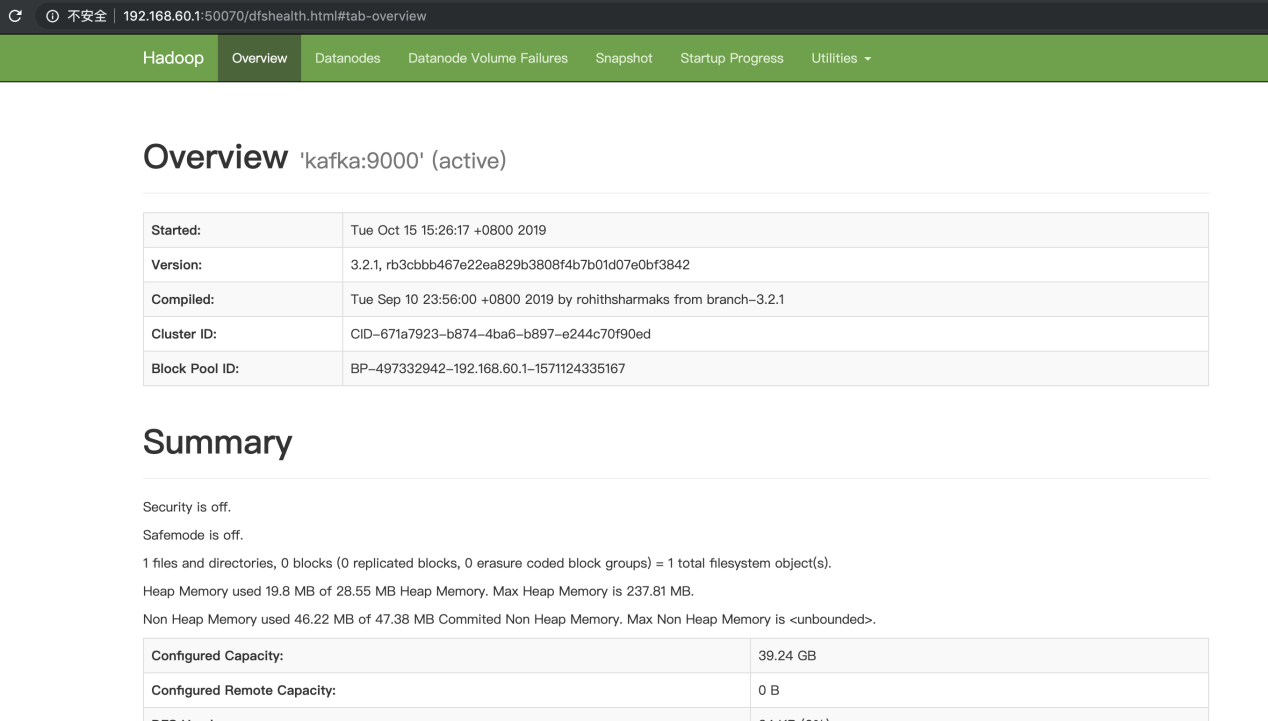
**core-site.xml配置和该页面有关，少了该配置页面就访问不了**

<property>

<name>fs.default.name</name>

<value>hdfs://kafka:9000</value>

</property>



**四、安装Hive**

参考：<https://www.cnblogs.com/dxxblog/p/8193967.html>

1、新建一个hive-site.xml

[root@kafka u01]# vim **/u01/hive/conf/hive-site.xml**

<?xml version="1.0" encoding="UTF-8" standalone="no"?>

<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<configuration>

<property>

<name>javax.jdo.option.ConnectionURL</name>

<value>jdbc:mysql://localhost:3306/hive?&amp;createDatabaseIfNotExist=true&amp;characterEncoding=UTF-8&amp;useSSL=false</value>

</property>

<property>

<name>javax.jdo.option.ConnectionDriverName</name>

<value>com.mysql.jdbc.Driver</value>

</property>

<property>

<name>javax.jdo.option.ConnectionUserName</name>

<value>root</value>

</property>

<property>

<name>javax.jdo.option.ConnectionPassword</name>

<value>12345</value>

</property>

<property>

<name>hive.metastore.schema.verification</name>

<value>false</value>

</property>

</configuration>

1. 下载地址：https://dev.mysql.com/downloads/connector/j/5.1.html

将java连接mysql的jar包拷贝到**/u01/hive/lib/**

**cp mysql-connector-java-5.1.48.jar /u01/hive/lib/**

3、执行**schematool -dbType mysql -initSchema**

4、**敲hive**，进入hive操作

hive> **create database hive\_1;**

OK

Time taken: 0.943 seconds

hive> **use hive\_1;**

OK

Time taken: 0.066 seconds

hive> **CREATE TABLE students (name VARCHAR(64), age INT, gpa DECIMAL(3, 2))**

> **CLUSTERED BY (age) INTO 2 BUCKETS STORED AS ORC;**

OK

Time taken: 0.959 seconds

hive> **INSERT INTO TABLE students**

> **VALUES ('fred flintstone', 35, 1.28), ('barney rubble', 32, 2.32);**

hive> **insert into students values ('huan',1,2.5);**

hive> **select \* from students;**

OK

barney rubble 32 2.32

fred flintstone 35 1.28

huan 1 2.50

Time taken: 0.466 seconds, Fetched: 3 row(s)



登录beeline

[root@kafka Desktop]# **cat /u01/hadoop/etc/hadoop/hdfs-site.xml**

<property>

<name>dfs.permissions</name>

<value>false</value>

</property>

<property>

<name>dfs.safemode.threshold.pct</name>

<value>0</value>

</property>

[root@kafka Desktop]# **cat /u01/hive/conf/hive-site.xml**

<property>

<name>hive.server2.thrift.port</name>

<value>10000</value>

</property>

<property>

<name>hive.server2.thrift.bind.host</name>

<value>192.168.60.1</value>

</property>

[root@kafka Desktop]#nohup hive --service metastore >> log.out 2>&1 &

[root@kafka Desktop]#nohup hive --service hiveserver2 >> log.out 2>&1 &

等一会就可以连接了



beeline -u jdbc:hive2://kafka:10000

beeline登录hive，执行insert，报错

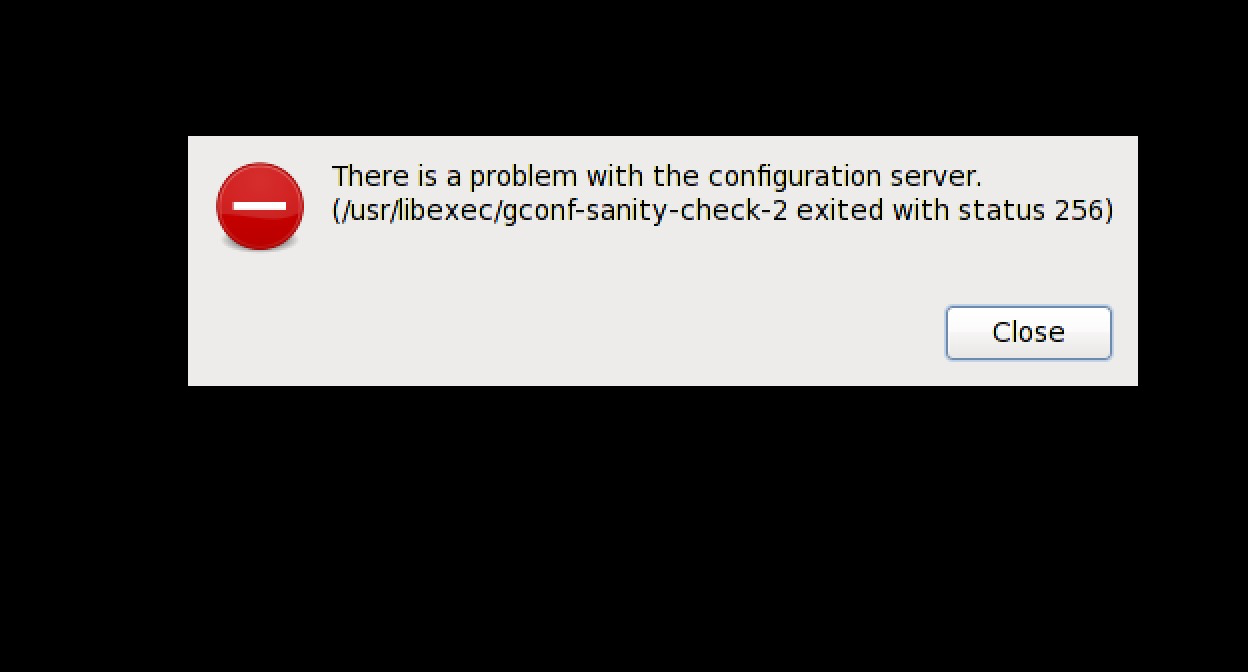
Permission denied. user=anonymous is not the owner of inode=/tmp/hadoop-yarn/staging/anonymous/.staging

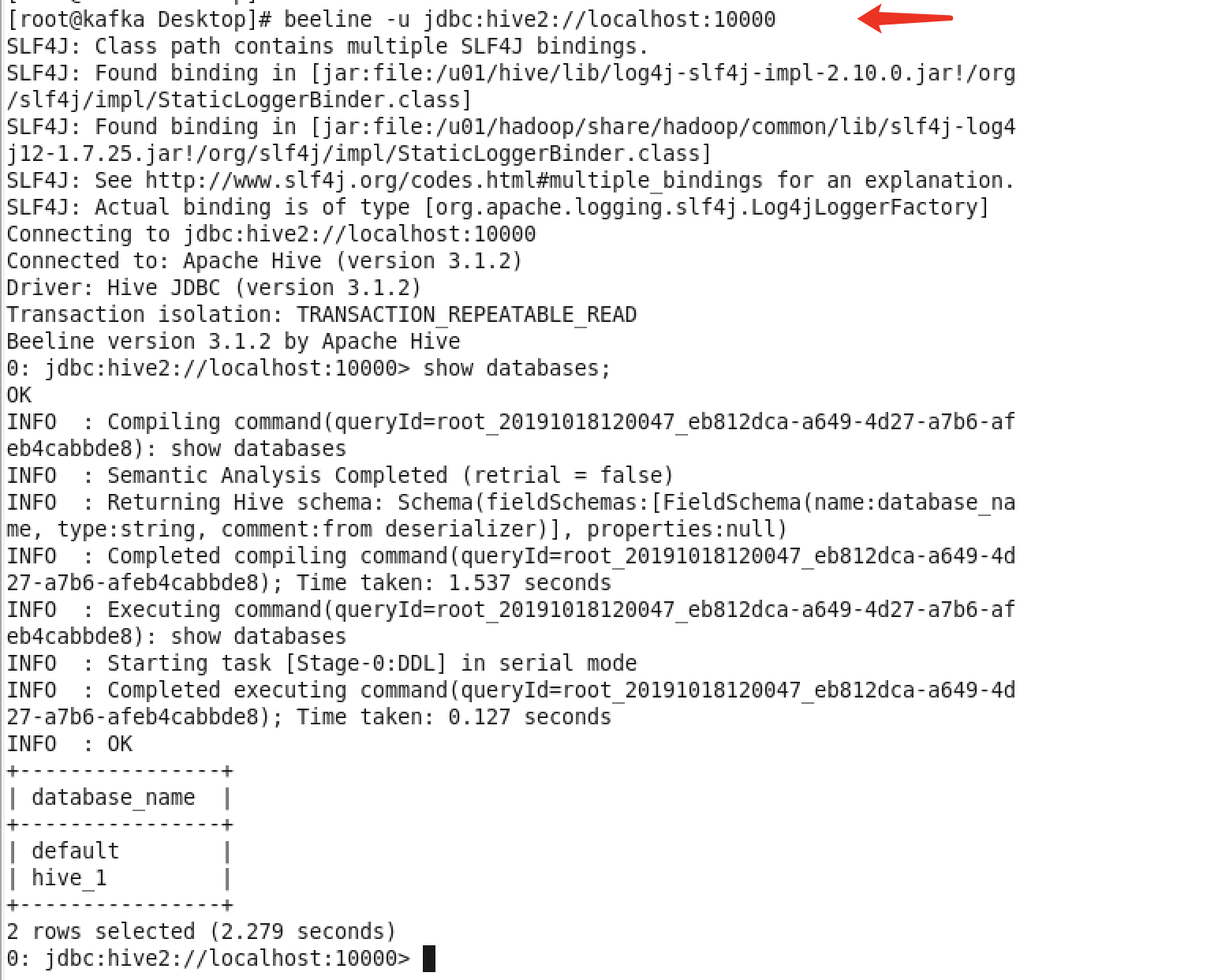
实在解决不了，我就重新格式化hdfs。然后就可以insert了

hadoop namenode -format

机器重启，出现如图，mysql启动报错pid无法更新

解决：chmod 777 /tmp





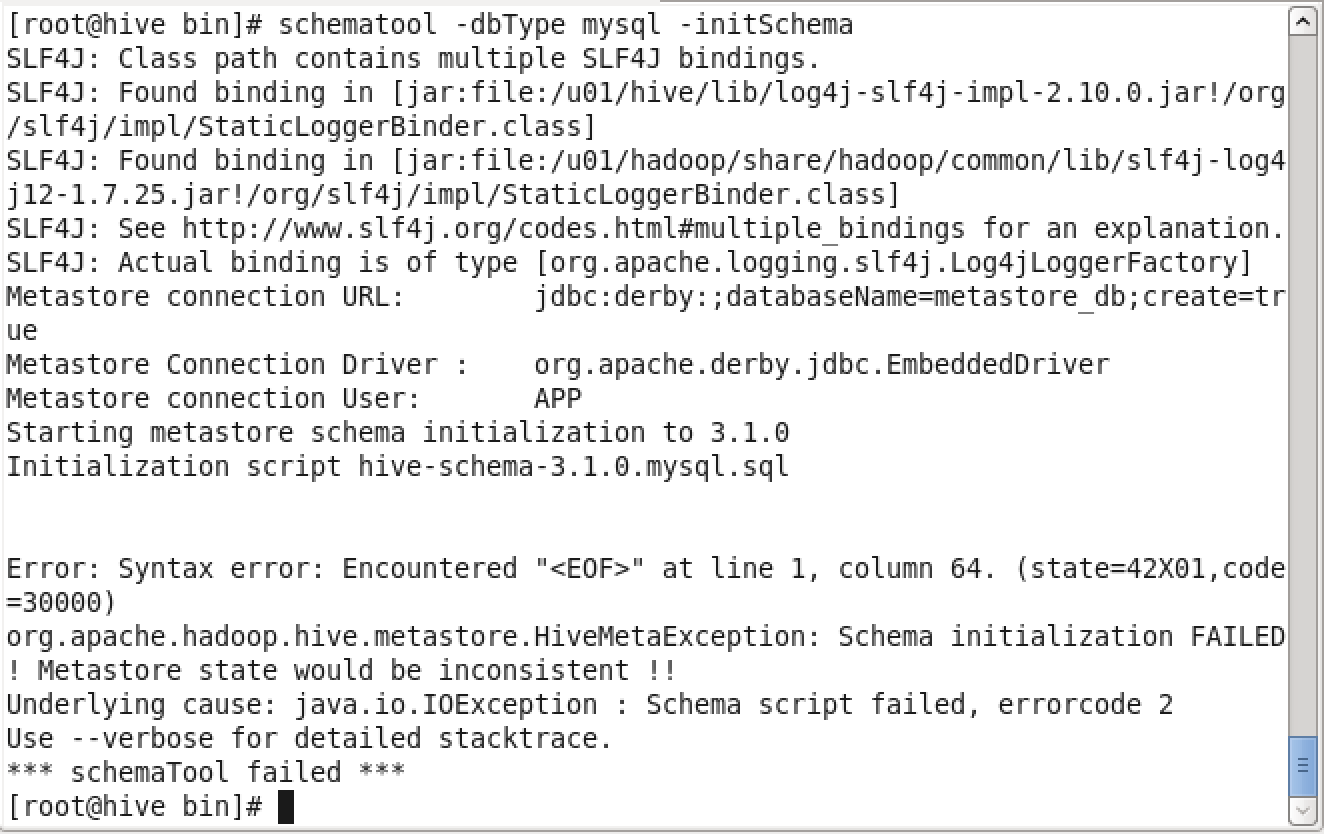
学习资料：

<https://www.cnblogs.com/qingyunzong/p/8707885.html>

**遇到的报错：**



解决：上面的第3213行，第96个字符是非法字符，注释掉就行了



解决：Hadoop，hive全都用最新的版本。之前用的stable-2目录下的包



解决：

[root@kafka Desktop]# find / -name guava\*

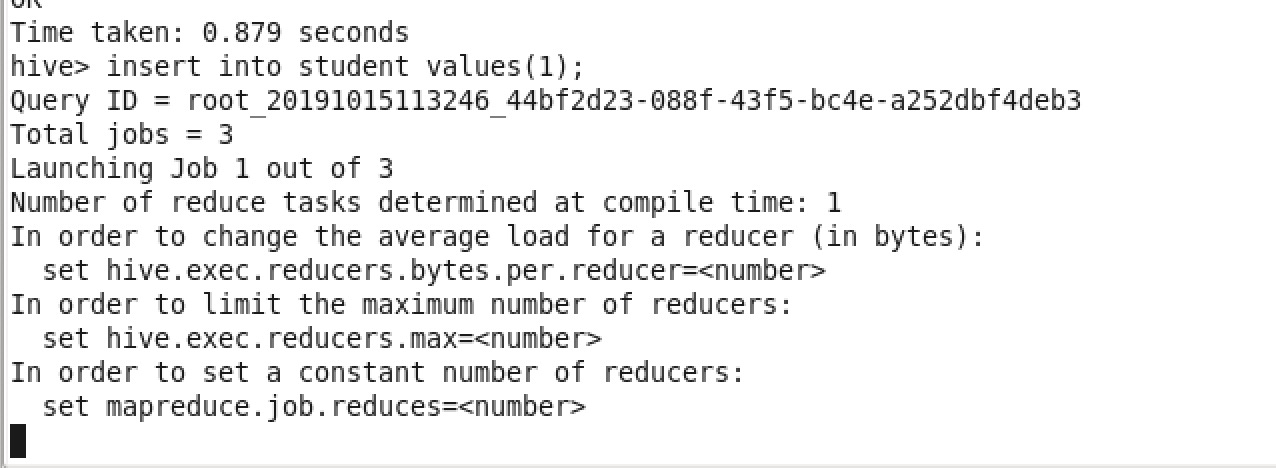
/u01/hive/lib/guava-19.0.jar

/u01/hadoop/share/hadoop/common/lib/guava-27.0-jre.jar

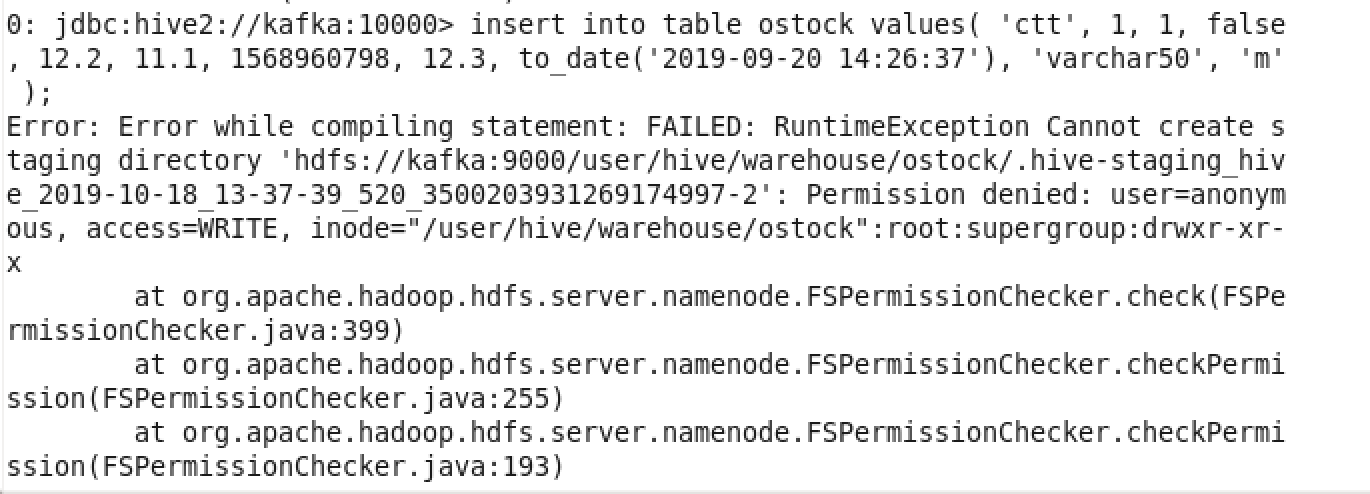
/u01/hadoop/share/hadoop/hdfs/lib/guava-27.0-jre.jar

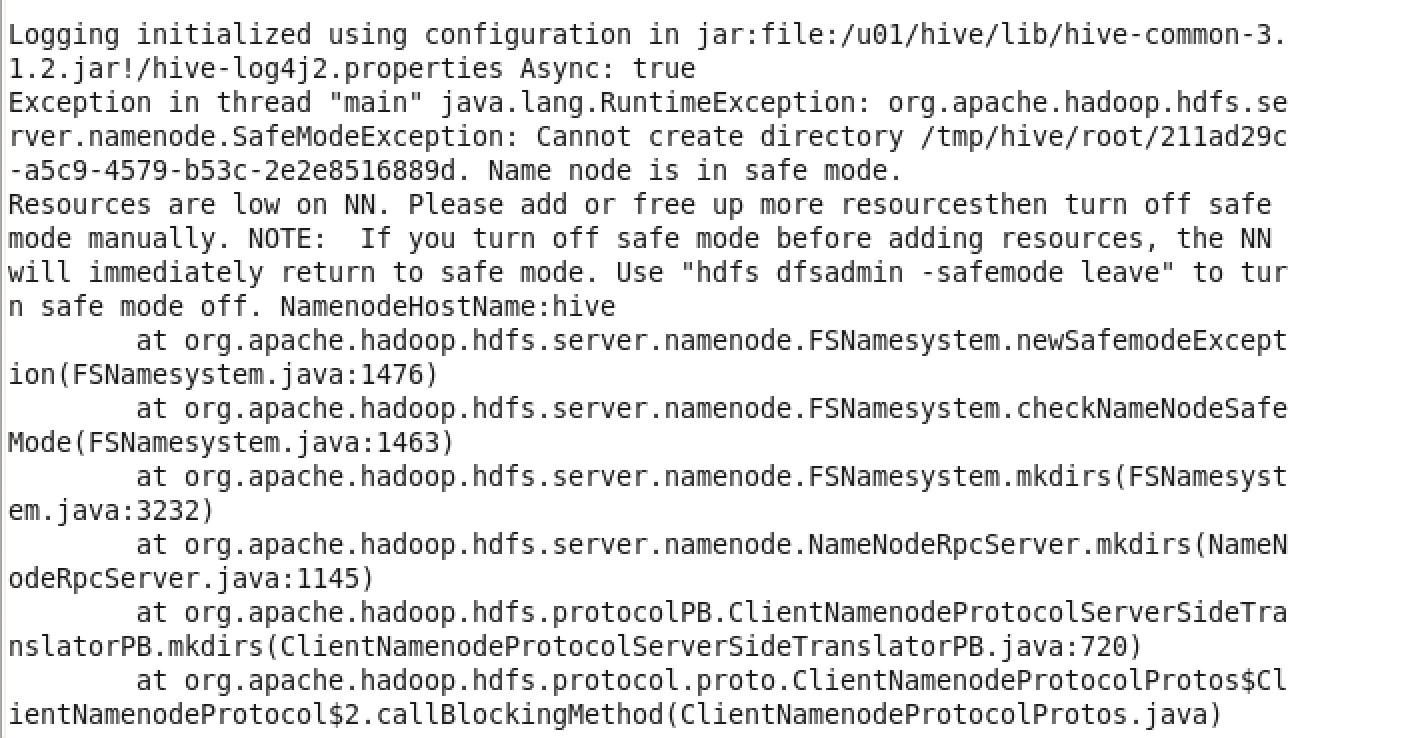
[root@kafka Desktop]# mv /u01/hive/lib/guava-19.0.jar /u01/hive/

[root@kafka Desktop]# cp /u01/hadoop/share/hadoop/hdfs/lib/guava-27.0-jre.jar /u01/hive/lib/



解决：查资料重新配置多次，修改了dfs目录，最后可以insert,**猜测**可能是空间不足。因为之前访问hdfs的页面显示占用是100%。





解决：

[root@hive hadoop]# hdfs dfsadmin -safemode leave



[root@hive Desktop]# vim /u01/hadoop/etc/hadoop/yarn-site.xml

<configuration>

<!-- Site specific YARN configuration properties -->

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

</property>

<property>

<name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>

<value>org.apache.hadoop.mapred.ShuffleHandler</value>

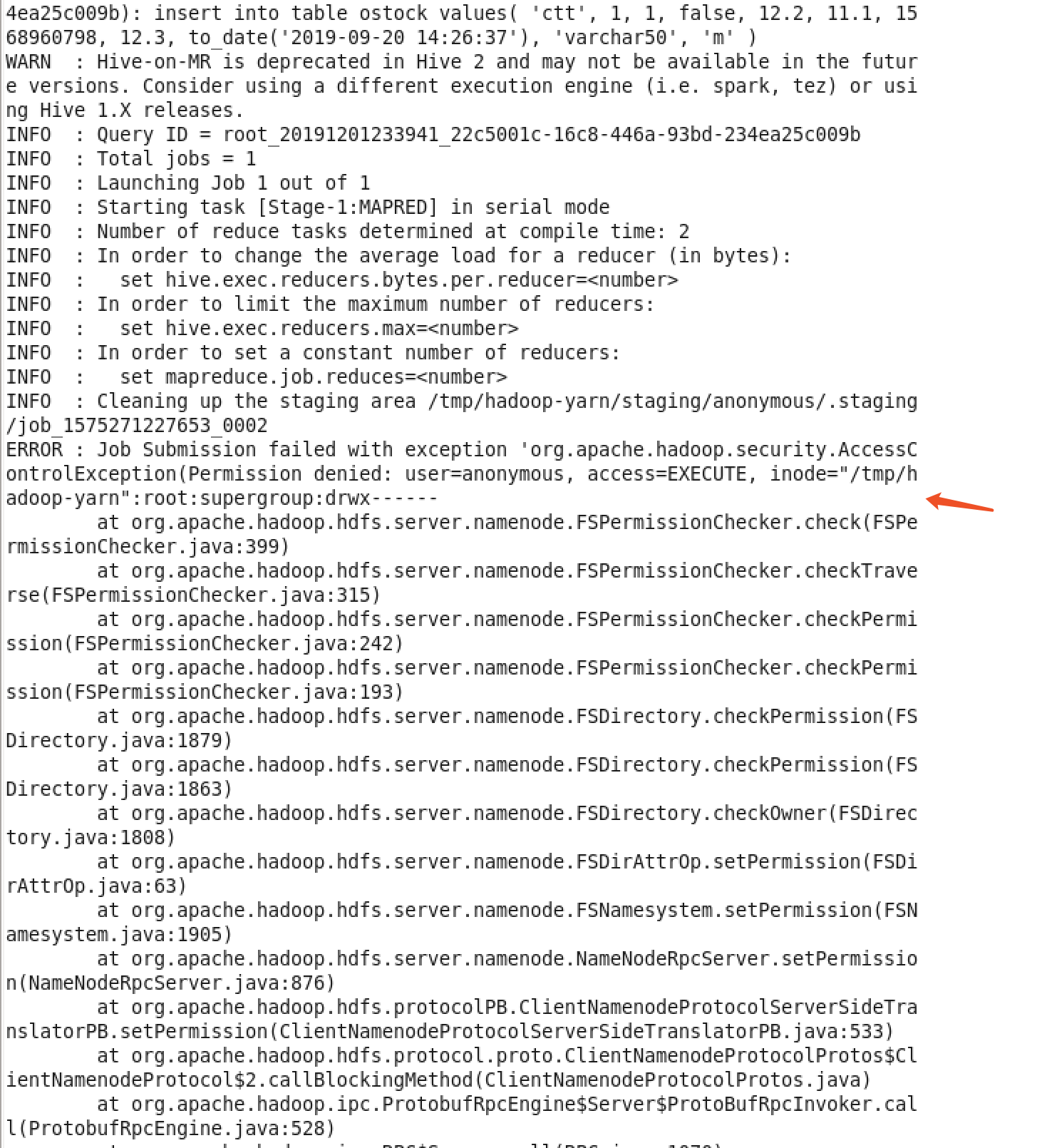
</property>

</configuration>

只是重启hdfs、yarn服务还是报错

重启机器，删除/tmp/\*、/u01/cache，格式化hdfs

还有我的/u01/hadoop/etc/hadoop/workers 文件里写的是主机名,改为IP 192.168.60.1



解决：

[root@hive Desktop]# hadoop fs -chmod -R 777 /tmp

2019-12-01 23:41:41,754 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

[root@hive Desktop]# hadoop fs -ls /tmp

2019-12-01 23:41:48,209 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

Found 2 items

drwxrwxrwx - root supergroup 0 2019-11-29 09:27 /tmp/hadoop-yarn

drwxrwxrwx - root supergroup 0 2019-12-01 22:23 /tmp/hive

建表报ACID不可以在NO ACID模式下

在vim /u01/hive/conf/hive-site.xml添加如下，重启机器

<property>

<name>hive.txn.manager</name>

<value>org.apache.hadoop.hive.ql.lockmgr.DbTxnManager</value>

</property>

<property>

<name>hive.compactor.initiator.on</name>

<value>true</value>

</property>

<property>

<name>hive.compactor.worker.threads</name>

<value>1</value>

</property>

<property>

<name>hive.support.concurrency</name>

<value>true</value>

</property>

<property>

<name>hive.enforce.bucketing</name>

<value>true</value>

</property>

<property>

<name>hive.exec.dynamic.partition.mode</name>

<value>nonstrict</value>

</property>