目录

[1 禁止 5](#_Toc486433301)

[1.1 禁止配置DNS 5](#_Toc486433302)

[1.2 禁止开启防火墙 5](#_Toc486433303)

[1.3 文件数超过1.5亿，禁止执行hdfs balance 5](#_Toc486433304)

[1.4 禁止勾选或者取消 Cluster Admin Operations 5](#_Toc486433305)

[1.5 禁止使用jmap命令 6](#_Toc486433306)

[1.6 C30和C60禁止任何业务在保存修改参数时，勾选如下项”重启服务” 6](#_Toc486433307)

[1.7 C30和C60禁止任何业务在重启业务时，勾选如下项“Restart all the associated services” 6](#_Toc486433308)

[1.8 C30禁止调整hive的日志级别 6](#_Toc486433309)

[1.9 禁止打开hdfs如下日志级别 7](#_Toc486433310)

[1.10 C30禁止只重启hiveserver ，或者只重启metastore，必须一起重启 7](#_Toc486433311)

[1.11 禁止修改hdfs namenode如下参数 8](#_Toc486433312)

[2 必须 8](#_Toc486433313)

[2.1 C30新增Yarn ResourceManager 必须运行如下脚本清理 cg垃圾 8](#_Toc486433314)

[2.2 FI参数配置界面，除LOG的级别修改可以不用重启立即生效外，**不管参数是否打星号，都必须重启才生效**。 10](#_Toc486433315)

[2.3 新增Yarnresource node，必须修改excel中设置vcore为80 10](#_Toc486433316)

[2.4 C30安装lzo后，必须重启yarn 10](#_Toc486433317)

[2.5 Hdfs namenode 机器名小的，必须处于主用状态 10](#_Toc486433318)

[2.6 必须配置Hiveserver连接zookeeper时间= 内存回收时间\*5 11](#_Toc486433319)

[2.7 必须关闭C60 kafka的自动创建 11](#_Toc486433320)

[2.8 调整C60 kafka日志保存时间 12](#_Toc486433321)

[2.9 定期清除hdfs /tmp/hadoop-yarn/staging下面的文件 12](#_Toc486433322)

[3 调优 12](#_Toc486433323)

[3.1 Yarn vcore调整 12](#_Toc486433324)

[3.2 打印dbservice执行超长的sql 12](#_Toc486433325)

[3.3 Gaussdb中对分区表加索引 13](#_Toc486433326)

[3.4 单节点hdfs balance 13](#_Toc486433327)

[3.5 优化hdfs namenode 13](#_Toc486433328)

[3.6 什么时候调大AM的内存/AM OOM 14](#_Toc486433329)

[3.7 Loader任务内存优化 14](#_Toc486433330)

[3.8 join操作触发的数据偏移 14](#_Toc486433331)

[3.9 Kafka中的topic partition配置为多少? 15](#_Toc486433332)

[3.10 ftp loader配置原则 15](#_Toc486433333)

[3.11 AM报OOM错误 15](#_Toc486433334)

[3.12 hive on spark 合并文件 15](#_Toc486433335)

[4 日常经验 16](#_Toc486433336)

[4.1 metadata数据多过hdfs 16](#_Toc486433337)

[4.2 添加hdfs文件后，需要关联metadata 16](#_Toc486433338)

[4.3 查看当前yarn任务资源使用情况 16](#_Toc486433339)

[4.4 yarn中默认的SparkSQL被删除怎么办？ 16](#_Toc486433340)

[4.5 大面积，node agent停止失败，重启2个controller。管理各个节点，各节点上报状态。 17](#_Toc486433341)

[4.6 重启业务，从这里启动，避免列表中有太多主机，忘记选择 17](#_Toc486433342)

[4.7 Ommdb过期 17](#_Toc486433343)

[4.8 分区是乱码无法删除 18](#_Toc486433344)

[4.9 Gaussdb的账号account配置为不锁定 19](#_Toc486433345)

[4.10 hive 发生问题时，如何收集日志信息 20](#_Toc486433346)

[4.11 服务无法启动/健康状态不正常/Noteagent 停止失败: 20](#_Toc486433347)

[4.11.1 **现象**：服务重启，无法停止。健康状态为红色。 20](#_Toc486433348)

[**4.11.2** **原因：** 21](#_Toc486433349)

[**4.11.3** **恢复步骤：** 21](#_Toc486433350)

[4.12 c60 reflect方法不可用，怎么办？ 22](#_Toc486433351)

[4.13 到底是hive慢还是hdfs慢？ 23](#_Toc486433352)

[4.14 如何通过命令切换hiveserver和metastore? 23](#_Toc486433353)

[4.15 查看hdfs哪些盘坏了 24](#_Toc486433354)

[4.16 更换web界面图标 24](#_Toc486433355)

[4.17 当Yarn无法查看日志是，如何看应用日志 25](#_Toc486433356)

[4.18 GaussDB中 The account has been locked 25](#_Toc486433357)

[4.18.1 现象 25](#_Toc486433358)

[4.18.2 解锁账号的方法 27](#_Toc486433359)

[4.19 机器失联如何找回 27](#_Toc486433360)

[4.19.1 操作步骤 27](#_Toc486433361)

[4.19.2 打lzo补丁(C30特有) 32](#_Toc486433362)

[4.20 Web ui停止，怎么启动? 33](#_Toc486433363)

[4.21 无法添加FTP Server服务 33](#_Toc486433364)

[4.22 FI web触发错误的命令，如何终止 35](#_Toc486433365)

[4.23 hdfs如何关闭对坏盘的检测？ 35](#_Toc486433366)

[4.24 为什么yarn application –list会 Failing over to ? 35](#_Toc486433367)

[4.25 C60 Hive DB改名字的实施方案 35](#_Toc486433368)

[4.26 hdfs/hadoop 命令报内存溢出 36](#_Toc486433369)

[4.27 total number of created files now is \*\*\*\*\* 36](#_Toc486433370)

[4.28 为什么spark任务提交时写明占用16G内存，实际会超过一些？ 37](#_Toc486433371)

[4.29 unable to close file becuase the last block does not have enough number of replicas 37](#_Toc486433372)

[4.30 还没解决，先记录 37](#_Toc486433373)

[4.31 变更表的存储格式后，数据无法正常查询 37](#_Toc486433374)

[4.32 sql优化方法 38](#_Toc486433375)

[4.33 坏盘替换方法 38](#_Toc486433376)

[4.34 C30版本增大高斯数据连接数方法 44](#_Toc486433377)

[4.35 告警12001处理方法 45](#_Toc486433378)

[4.36 node fault告警无法恢复 45](#_Toc486433379)

[4.37 表变更属组 46](#_Toc486433380)

[4.38 慢盘检测方法 47](#_Toc486433381)

[4.39 admin重做密码 47](#_Toc486433382)

[4.40 应用起不来，内存不够 48](#_Toc486433383)

[4.41 hdfs文件不可读 48](#_Toc486433384)

[4.42 Kafka topic在同一节点 不同磁盘之间分区迁移的方法 48](#_Toc486433385)

[4.43 Kafka topic在不同节点之间分区迁移的方法 49](#_Toc486433386)

[4.44 对特定 Kafka topic 设定日志保存时间 50](#_Toc486433387)

[4.45 查询表类型和路径 50](#_Toc486433388)

[4.46 管理节点ssh服务被停止导致集群告警 50](#_Toc486433389)

[4.47 FI界面添加hive的角色失败 51](#_Toc486433390)

[4.48 浮动IP和下面绑定的服务器IP要在同一个网段 52](#_Toc486433391)

[4.49 安装FI节点时，报磁盘容量不匹配 52](#_Toc486433392)

[4.50 C30报Exceed the reconnecting time limits 53](#_Toc486433393)

[4.50.1 现象 53](#_Toc486433394)

[4.50.2 原因： 53](#_Toc486433395)

[4.50.3 解决方法： 53](#_Toc486433396)

[4.51 C60任务提交到已经删除的队列 54](#_Toc486433397)

[4.52 C30从开源平台迁移数据的方法 55](#_Toc486433398)

[4.53 如何查询有问题的udf？ 57](#_Toc486433399)

[4.54 kafka启动失败，报java.lang.OutOfMemoryError: Map failed 58](#_Toc486433400)

[4.55 kafka进程存在，但是健康状态为bad 58](#_Toc486433401)

[4.56 FI portal监控数据在哪里? 59](#_Toc486433402)

[4.57 如何获取 服务的主备关系？ 59](#_Toc486433403)

[4.58 分析jhistory的日志。 59](#_Toc486433404)

# 禁止

## 禁止配置DNS

cat /etc/resolv.conf 内容要为空，否则整个集群健康状态时好时坏。

## 禁止开启防火墙

开启防火墙，集群状态时好时坏

cd /etc/sysconfig

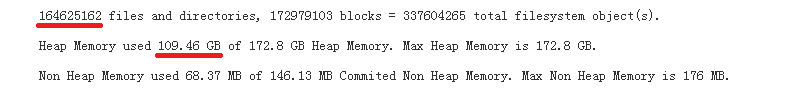
rcSuSEfirewall2 status

如果运行，则停止

rcSuSEfirewall2 stop

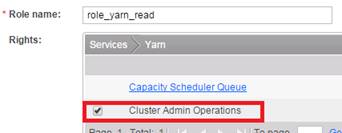
## 文件数超过1.5亿，禁止执行hdfs balance

当hdfs文件数在1.6亿时，在namenode内存约100G。执行全hdfs datanode的 hdfs balance，namenode会被撑爆。然后瘫痪。



## 禁止勾选或者取消 Cluster Admin Operations

创建用户角色时，选中和取消这个权限，都会导致yarn延迟重启。在C30中， yarn超过2个小时才重启；在C60中yarn在2小时内重启



禁止在

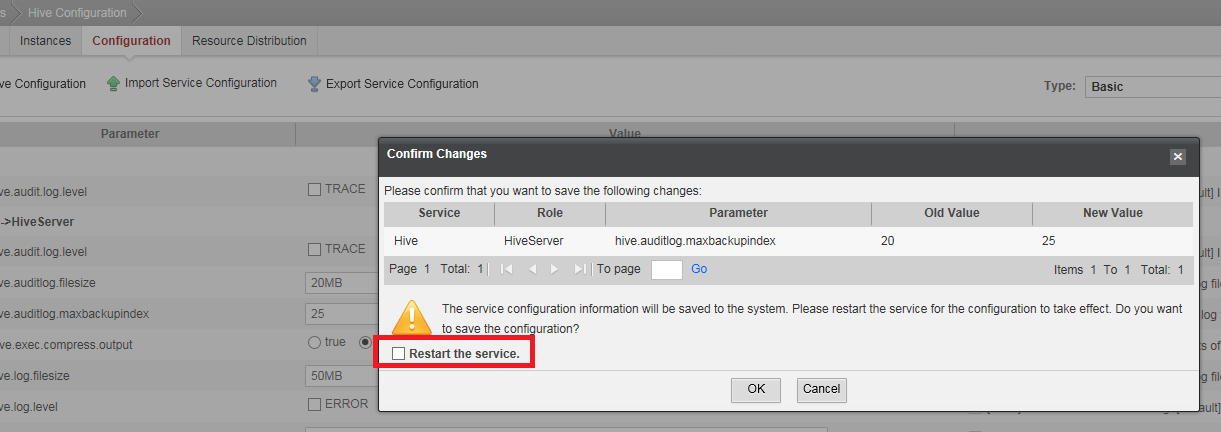
## 禁止使用jmap命令

执行jmap高概率导致进程崩溃。 无特别要求，禁止执行jmap命令

## C30和C60禁止任何业务在保存修改参数时，勾选如下项”重启服务”

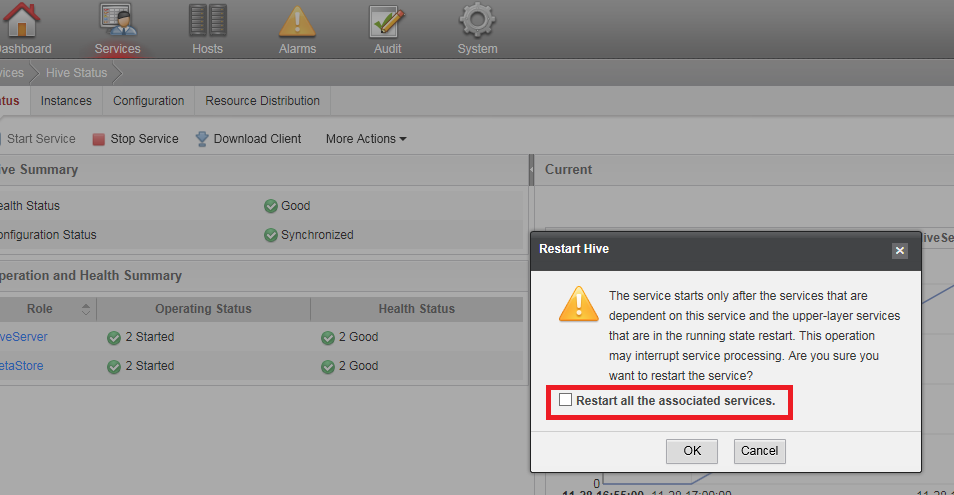
勾选此项，重启服务范围不可控。

关联的业务要重启，与之无关但是配置失效的业务也要重启。



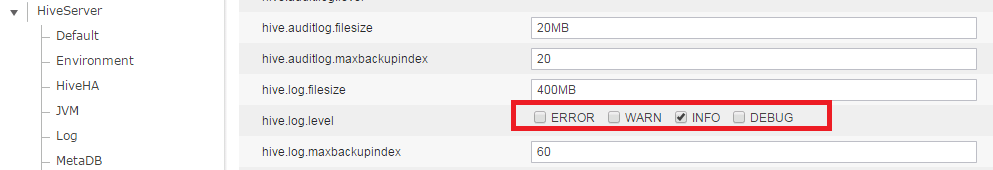
## C30和C60禁止任何业务在重启业务时，勾选如下项“Restart all the associated services”

关联范围不可控。

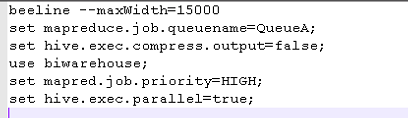


## C30禁止调整hive的日志级别

调整hive.log.level此日志级别，将导致beeline客户端无法显示任务进度



同时



执行set hive.exec.parallel=true 命令，它会导致无法显示进度消息

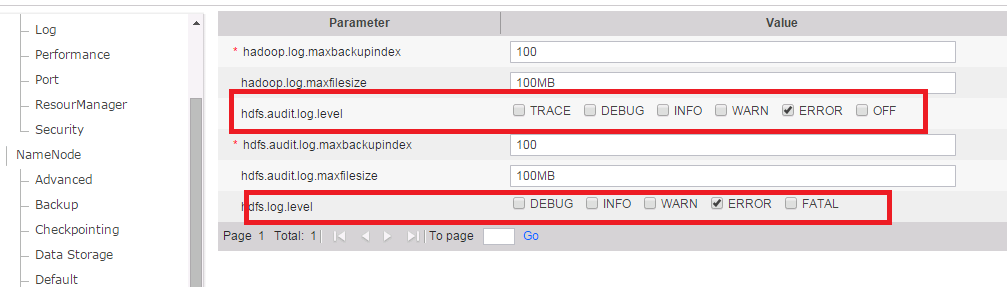
再， hiveserver返回给客户端的日志，都保存在本地如下目录中，检查目录 /tmp/omm/operation\_logs目录是否存在，omm用户是否能访问此目录。



## 禁止打开hdfs如下日志级别

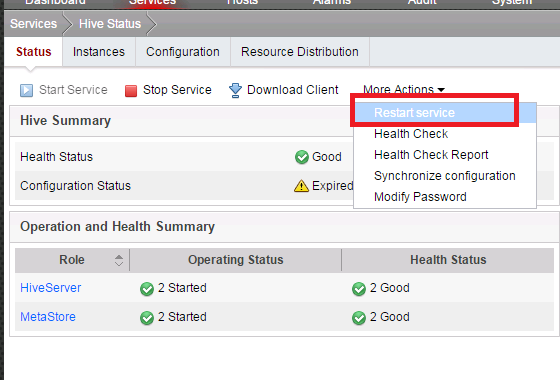
Hdfs写日志，会降低10%的性能。

写了日志，数据太多，你也看不懂。如非专家要求，日志级别只准保持在ERROR。

c

## C30禁止只重启hiveserver ，或者只重启metastore，必须一起重启

只重启一个业务，会导致lzo的入库缓慢。多次教训表明，必须在下图界面一起重启



## 禁止修改hdfs namenode如下参数

Namenode在172G内存在，配置dfs.namenode,handler.count=100，数值越大，需要的内存越多，能提供hdfs查询的线程也越多。

dfs.namenode.service,handler.count 与健康检查相关，如果配置过小，会认为服务挂了，会被切换。

以上2个参数，如无专家指导，禁止调整。



# 必须

## C30新增Yarn ResourceManager 必须运行如下脚本清理 cg垃圾

YARN nodemanage 需要配置定时任务，删除/dev/cgroup/cpu/hadoop-yarn下的container文件。具体原因见附件docx。



复制后，每台机器/home/omm如下

cid:image003.jpg@01D2328A.3324C890

**需统一：**

1. 以root用户登录，

vi /etc/cron.allow

添加用户名omm， 目的是允许该用户使用crontab

  2.以omm用户登录，执行

       crontab –e

       添加一行：0 22 \* \* \* sh /home/omm/cgroup-clean.sh &

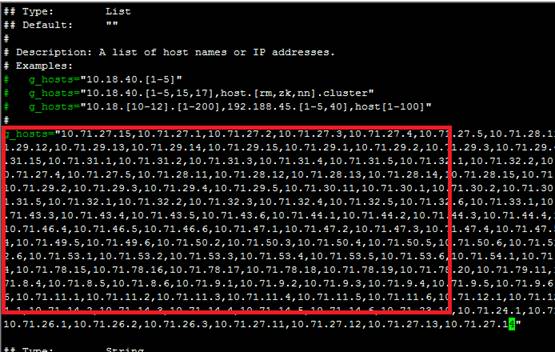
      每天22点清理垃圾目录

以后扩容任何机器时，都加上这个功能(见cluster.zip)，并配置好crontab。

批量安装的命令在cluster.zip中是，在cluster.ini中配置了要安装的主机。运维估计有神器，不一定用户它的批量安装。



./clusterscp.sh put /home/omm/cgroup-clean.sh /home/omm/



## FI参数配置界面，除LOG的级别修改可以不用重启立即生效外，**不管参数是否打星号，都必须重启才生效**。

## 新增Yarnresource node，必须修改excel中设置vcore为80

一般机器是48个逻辑CPU，查看命令是

cat /proc/cpuinfo | grep "processor" | wc -l

vcore配置逻辑核数的1.5-2倍

48 - 8（留8个核给操作系统）=40 \*2 = 80

新增加yarn resource manange节点时，不能按excel取默认值，必须设定为80

## C30安装lzo后，必须重启yarn

必须重启yarn，否则lzo不能正常工作。

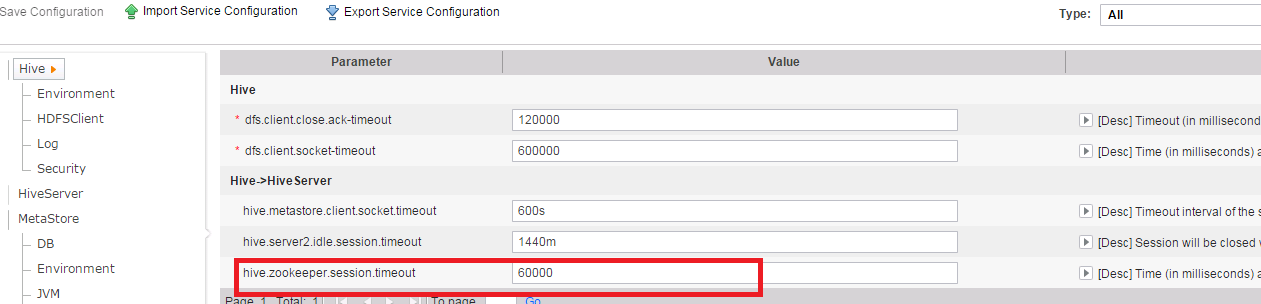
## Hdfs namenode 机器名小的，必须处于主用状态

Hive发起hdfs请求时，默认按照主机名，先去访问字母顺序小的主机，如果此机器处于备机，则返回鉴权失败。Hive根据配置文件中的hdfs信息，再访问第2个hdfs name的。

让hdfs namenode机器名小的机器，处于主用状态，有利于提高性能。

## 必须配置Hiveserver连接zookeeper时间= 内存回收时间\*5

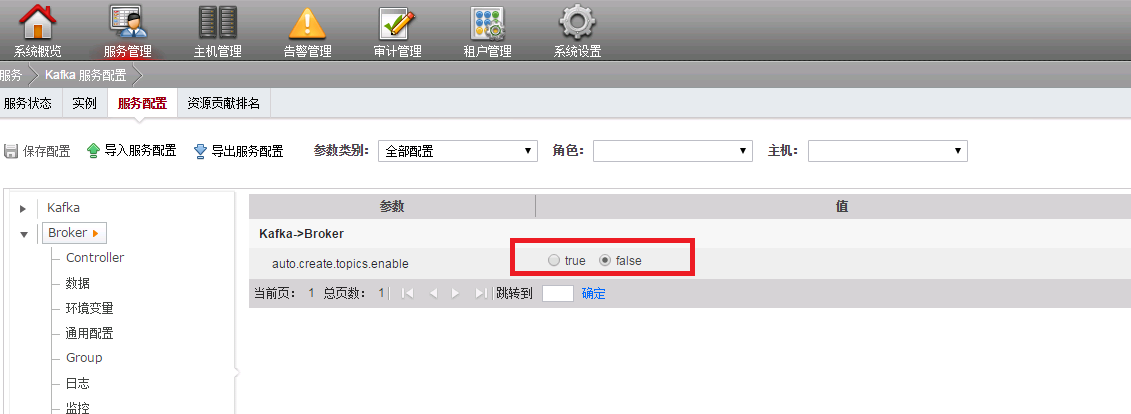
Hiveserver配置为40G， 运行一段时间，hiveserver会回收内存占用约50秒，接近与zookeeper连接超时时间。容易导致zookeeper认为服务不可用，发生主备切换。



设置hive.zookeeper.session.timeout=内存回收时间\*5

## 必须关闭C60 kafka的自动创建

改参数默认开启，如果不存在要访问的topic，则自动创建1个，且partition为2



自动创建有如下问题：

* + - 1. 大概率删不掉，后面维护难
      2. Partitioin为2，没考虑实际kafka的节点数，不合理。

必须关掉.

## 调整C60 kafka日志保存时间

7天变3天



## 定期清除hdfs /tmp/hadoop-yarn/staging下面的文件

下面放任务的临时文件，任务执行完毕后，都会自动删除。 但是，当AM失败时，目录无法清理。日积月累无法删除的目录越来越多。截止 2016.11.24日，该目录有110万文件和目录，1T临时数据。

# 调优

## Yarn vcore调整

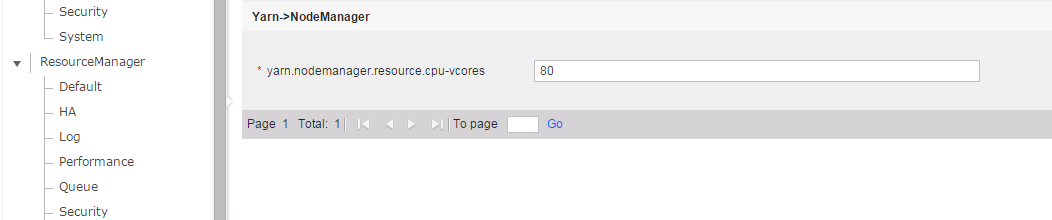
Vcore不对，yarn中资源偏少。

一般机器是48个逻辑CPU，查看命令是

cat /proc/cpuinfo | grep "processor" | wc -l

vcore配置逻辑核数的1.5-2倍

48 - 8（留8个核给操作系统）=40 \*2 = 80



修改完后要重启。

## 打印dbservice执行超长的sql

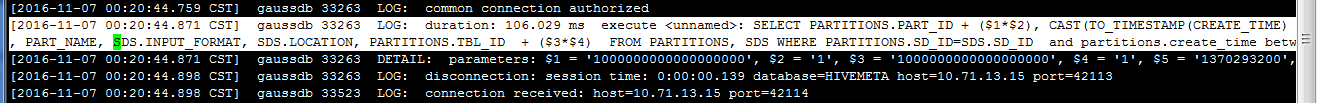
建议改成大于100毫秒都打印

source /opt/huawei/Bigdata/dbservice/.dbservice\_profile  
gs\_guc reload -c log\_min\_duration\_statement=100 -D /srv/BigData/dbdata\_service/data

查看如下，文件名等于日期。11月3日，则看\_03的日志



Grep “duration” logfile



最后用文本和excel处理，放在excel里面排序，得到最慢的SQL

## Gaussdb中对分区表加索引

分区表超过600万后，查询分区耗时会超过4秒，需要加分区，加快查询速度

执行source：  
source /opt/huawei/Bigdata/dbservice/.dbservice\_profile  
登录高斯数据库：  
gsql -p 20051 -d hivemeta -U hive -W HiveUser@  
执行以下命令创建索引

CREATE INDEX "SDS\_N50" ON "SDS" USING btree ("CD\_ID");

CREATE INDEX “PARTITIONS\_N51” on “PARTITIONS”using btree(“CREATE\_TIME”)

## 单节点hdfs balance

由FI部门提供



## 优化hdfs namenode

1) HDFS中，fsimage 和editlog 均是元数据信息。NameNode会周期性对元数据进行合并归档， Fsimage是合并归档前的全量元数据信息，editlog周期内增量元数据信息。

2) 查看当前集群约25分钟新生成一个 fsimage文件，16G 文件按照100MB/S的速度，写入磁盘会耗时150S。

3) 非读操作都需要记录editlog 才会反馈操作结果，如果此时正在刷editlog，会影响到HDFS性能。

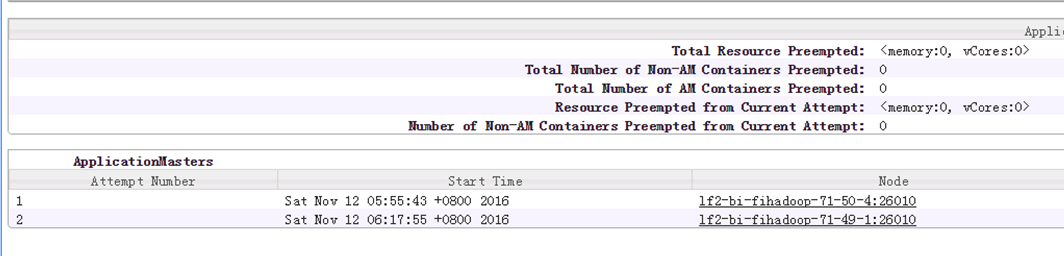
因此将editlog 和fsimage 分磁盘存放，会提升大集群HDFS的NameNode 稳定性。

操作步骤原型验证流程，由 彭瑞林、李亿、秦信伟、蔡小峰 验证并输出：《EMUI NameNode性能调优.docx》



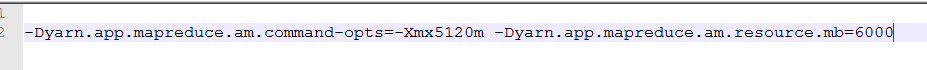
## 什么时候调大AM的内存/AM OOM

点击job详情，ApplicationMasters有多条记录，多半表示AM内存不够，实际要按4.18查看日志。



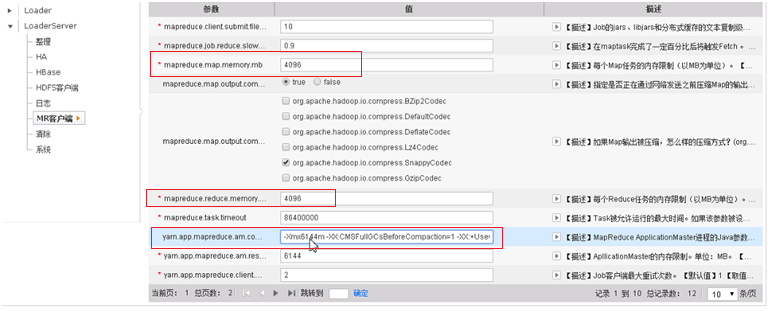
单AM内存不够时，设置如下2个参数

这是hadoop命令中设置的参数



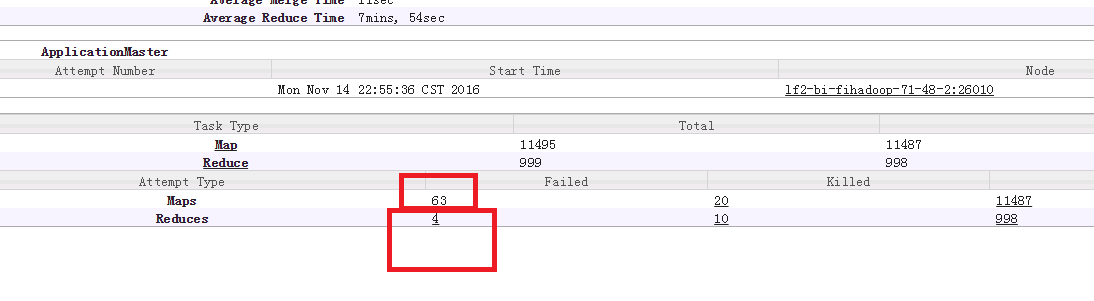
## Loader任务内存优化

下面3个框对应map内存，reduce内存，am内存，哪个报错，就增加哪个。



## join操作触发的数据偏移

SQL正确，reduce执行失败，且map也有失败。SQL含join操作



点开Failed的reduce能看到oom错误。

脚本加 set hive.optimize.skewjoin=true

## Kafka中的topic partition配置为多少?

创建topic时，最优的 partition = kafka节点数量 \*2

## ftp loader配置原则

loader导数据用  每个map都需要连接ftp，保证所有loader任务起的map数小于ftp-server限制的链接线程数即可

ftp-server的总连接数大于并发的loader起的map数即可

## AM报OOM错误

加大AM内存:

set yarn.app.mapreduce.am.command-opts=-Xmx4096m;

set yarn.app.mapreduce.am.resource.mb=6144

## hive on spark 合并文件

hive.merge.sparkfiles=true

# 日常经验

## metadata数据多过hdfs

元数据还存在NN里，但是没有对应的blcok了，报block missing，那么用hdfs fsck <patch> -delete命令删除掉元数据就行了

## 添加hdfs文件后，需要关联metadata

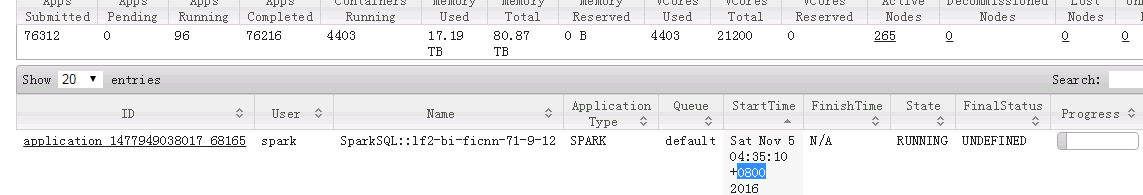
msck repair table

## 查看当前yarn任务资源使用情况

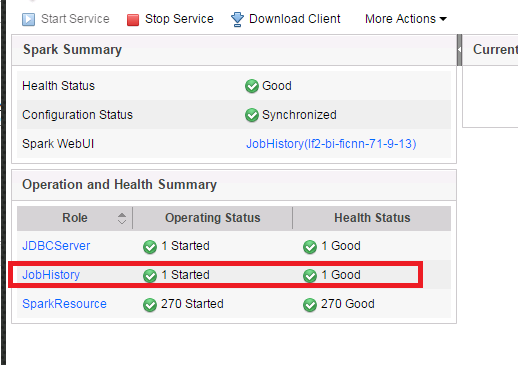
mapred job –list

## yarn中默认的SparkSQL被删除怎么办？

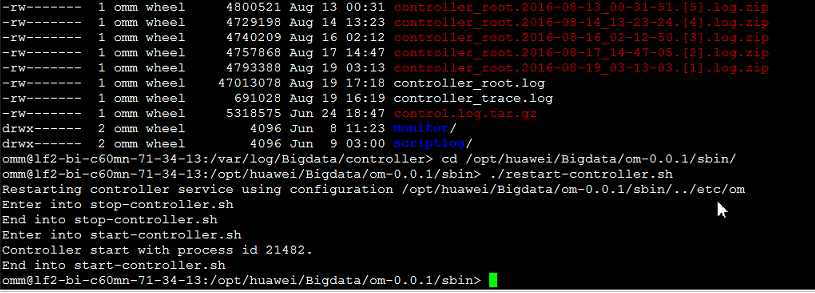
如下任务在yarn中应该一直存在，负责统计集群的信息



如被杀掉，yarn中看不到任务列表。 请重启spark 🡪JobHistory



## 大面积，node agent停止失败，重启2个controller。管理各个节点，各节点上报状态。





## 重启业务，从这里启动，避免列表中有太多主机，忘记选择

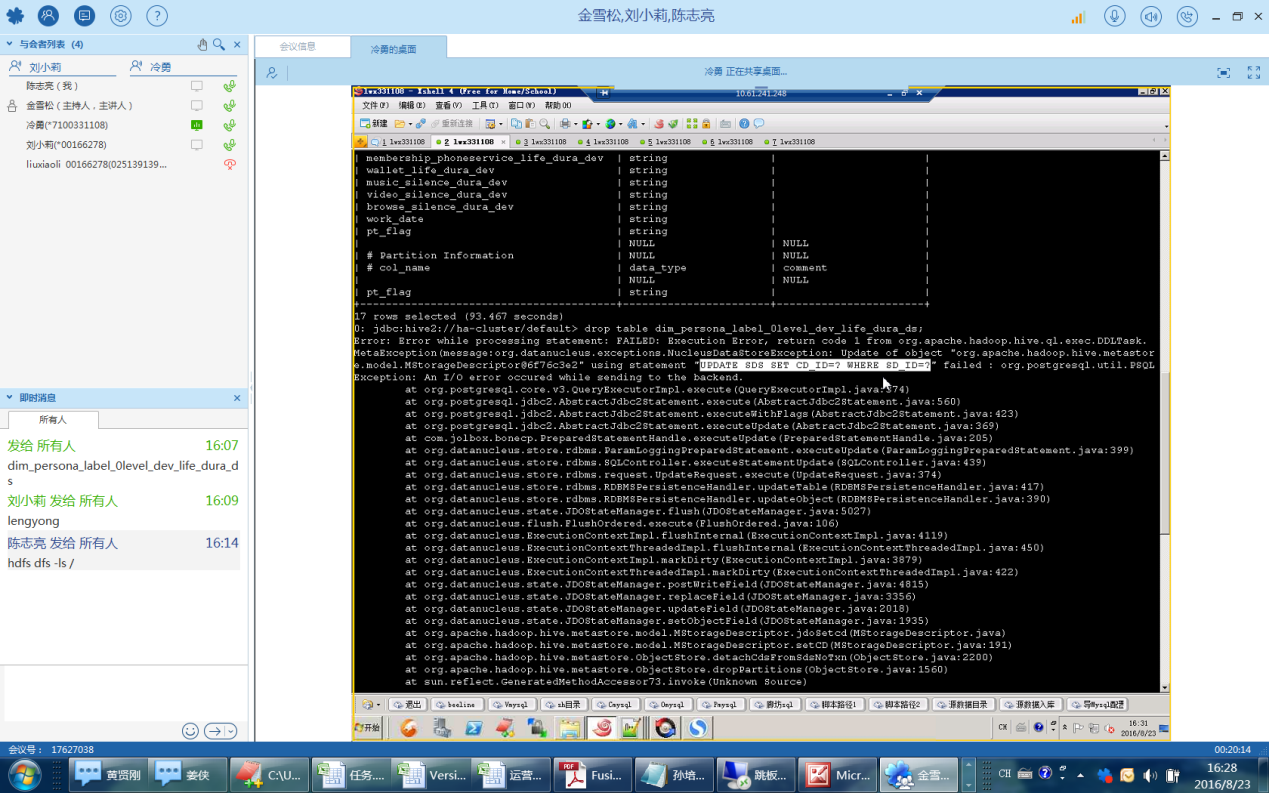


## Ommdb过期

这是Linux用户的管控机制



## 分区是乱码无法删除



从元数据删除无效分区方法 ：

Step 1) 登录OM页面-->DBService-->Configuration-->All-->HA   找到dbservice.floatip:160.161.0.145

Step 2) 用OM用户执行source /opt/huawei/Bigdata/dbservice/.dbservice\_profile

Step 3)连接到DB: gsql -p 20051 -h 160.161.0.145 -U hive -W HiveUser@ -d hivemeta

Step 4)执行以下语句：

select \* from tbls where tbl\_name='XXX';

select \* from partitions where tbl\_id=XXX;

delete from partitions where tbl\_id=XXX and part\_id=XXX;

select \* from PARTITION\_KEY\_VALS where part\_id=XXX;

delete from PARTITION\_KEY\_VALS where part\_id=XXX;

select \* from PARTITION\_PARAMS where part\_id=XXX;

delete from PARTITION\_PARAMS where part\_id=XXX;

## Gaussdb的账号account配置为不锁定

访问gaussdb的账号hive，如果重试多次，密码不对，会锁定。导致服务中断

去掉锁定设置数据库安装用户调用命令 gs\_guc reload –c password\_lock\_time=0 另外看看产品是不是有备份高斯配置文件 那个也要改。 DBService的安装目录下有一份高斯配置文件，需要手动修改下：/opt/huawei/Bigdata/dbservice/setup/conf/postgresql.conf，主备都改

omm@189-39-151-16:/opt/huawei/Bigdata/FusionInsight/dbservice> source .dbservice\_profile omm@189-39-151-16:/opt/huawei/Bigdata/FusionInsight/dbservice> vi /opt/huawei/Bigdata/dbservice/setup/conf/postgresql.conf



## hive 发生问题时，如何收集日志信息

1. 问题发生时的HiveServer运行日志，主备都收集一下：  
目录：/var/log/Bigdata/hive/hiveserver/  
2. 问题发生时的metastore日志，两个节点都收集一下：  
目录： /var/log/Bigdata/hive/metastore/  
3. 审计日志，两个节点的都收集一下：  
目录：/var/log/Bigdata/audit/hive/  
4. 当前主 HiveServer的jstack日志  
切换到omm用户,   
执行ps -ef | grep HiveServer获取HiveServer的pid  
每隔3秒执行一次 jstack pid > 1.out/2.out....，共执行5次  
5. MetaStore的jstack日志，两个节点都执行一下  
ps -ef | grep MetaStore  
每隔3秒执行一次jstack pid > 1.out/2.out....，共执行5次

多谢了！

说这么多，改改这个脚本



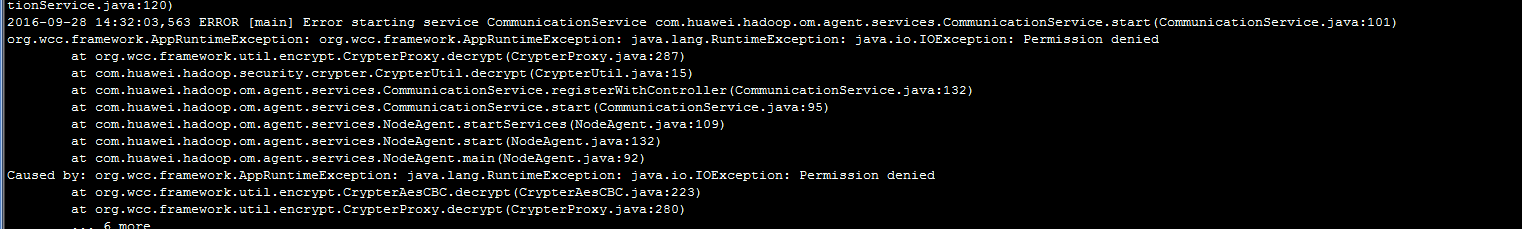
## 服务无法启动/健康状态不正常/Noteagent 停止失败:

### **现象**：服务重启，无法停止。健康状态为红色。

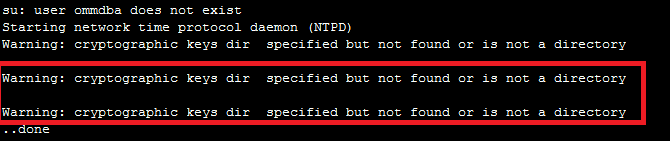
/var/log/Bigdata/nodeagent/agentlog/heart\_beat.log报错误

Java.io.IOException:Permission denied

At org.wcc.framwork.util.encrypt.CrypterProxy.decrypt



/var/log/Bigdata/nodeagent/scriptlog/start-agent.log报



### **原因：**

/tmp权限不对，/tmp目录下nodeagent的文件被删除。导致nodeagent启动读取不到必须的\_wcc\_file\_locker\_文件。然后业务无法启动。

/tmp目录的权限被磁盘软件distmgt修改了

### **恢复步骤：**

ls –l /

tmp目录的正确权限应该如下，否则root登录，执行

chmod 777 /tmp

chmod o+t /tmp

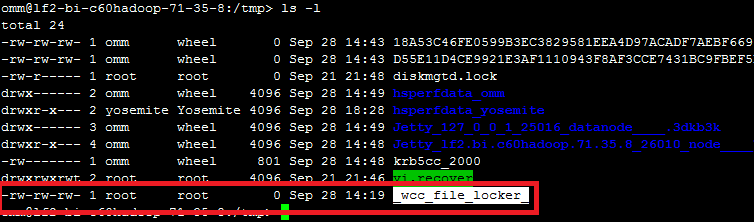
最终变成



如不存在\_wcc\_file\_locker\_， 则root用户 执行：

touch \_wcc\_file\_locker\_

chmod 666 \_wcc\_file\_locker\_



------------------以下非必须---------------------------------

重启nodeagent（非必要，实在没招了，再搞）:





观察

/var/log/Bigdata/nodeagent/agentlog/heart\_beat.log日志

其它有用信息在

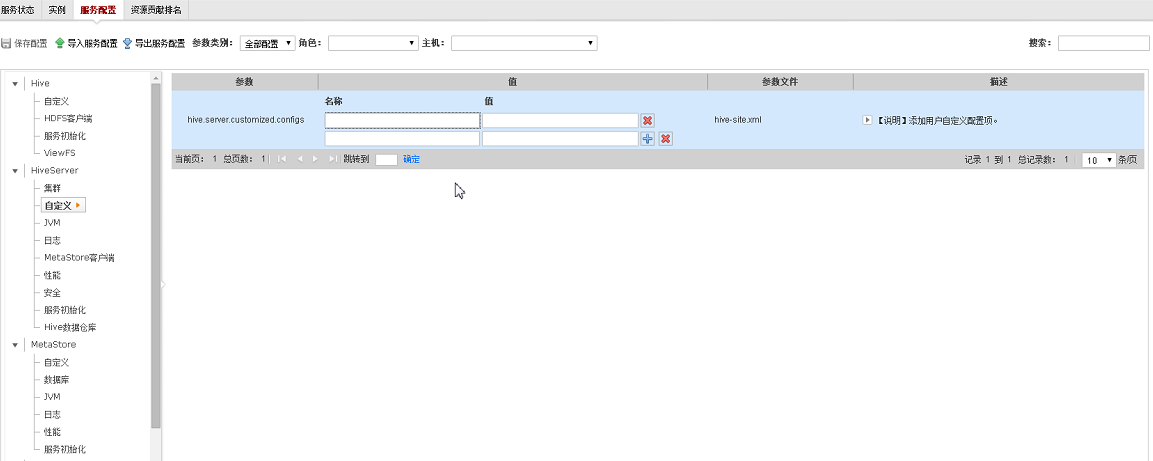
/var/log/Bigdata/nodeagent/scriptlog/nodeagent\_ctl.log

/var/log/Bigdata/nodeagent/scriptlog/start-agent.log (nodeagent启动日志)

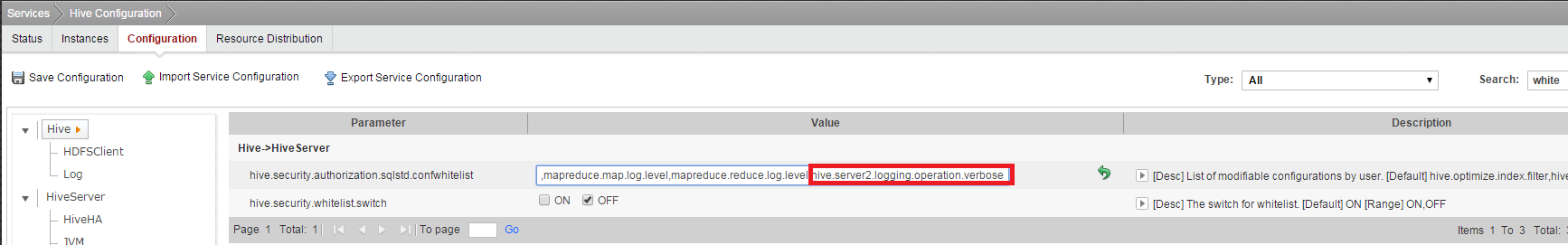
## c60 reflect方法不可用，怎么办？

安全原因，c60共reflect,reflect2,java\_method不可用

如果想使用reflect，可以在manager界面增加参数hive.server2.builtin.udf.blacklist  并设置为reflect2,java\_method



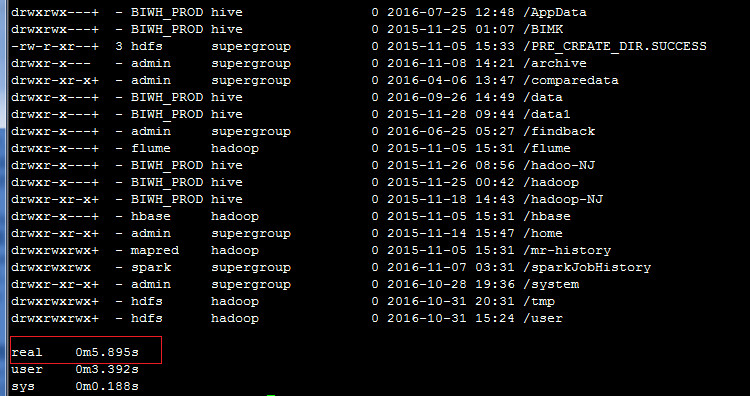
修改白名单, 加入hive.server2.logging.operation.verbose



## 到底是hive慢还是hdfs慢？

执行time hdfs dfs –ls 某个目录，为避免缓存，每次执行不同的目录，看平均的时长。

不超过10秒，表明hdfs正常；时间忽高忽低，hdfs不正常



## 如何通过命令切换hiveserver和metastore?

先软杀，等一会，再硬杀。

#/bin/sh

t=`date +"%Y%m%d%H%M%S"`

p1=`ps -ef| grep org.apache.hadoop.hive.om.service.HiveService | grep -v grep |awk -F ' ' '{print $2}'`

p2=`ps -ef| grep metastore | grep -v grep |awk -F ' ' '{print $2}' `

kill $p1

kill $p2

sleep 10

if [ $p1 != "" ]; then

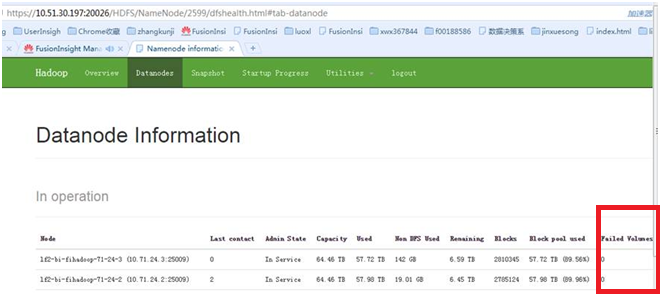
kill -9 $p1

kill -9 $p2

fi

## 查看hdfs哪些盘坏了

从FI界面，点击HDFS namenode的web界面 Datanodes, Failed Volumes列表示每个机器有多少块坏盘。



## 更换web界面图标

安装太多集群后，搞不清这是干什么的集群。在web界面加标识，方便识别



进入omsservice节点，一共2个。

cd /opt/huawei/Bigdata/apache-tomcat-7.0.68/webapps/web/resources/css

chmod -R a+w ./images/

cd /opt/huawei/Bigdata/apache-tomcat-7.0.68/webapps/web/resources/css/images

cp indlogo.png indlogo.png2

将indlogo.png传出来，并修改图片，再传回去覆盖。不用重启。

登录首页的图片在

C30 /opt/huawei/Bigdata/apache-tomcat-7.0.68/webapps/cas/images/login\_top\_bg.gif

C60 /opt/huawei/Bigdata/apache-tomcat-7.0.68/webapps/cas/images/login\_top\_bg.png

## 当Yarn无法查看日志是，如何看应用日志

Hdfs dfs –ls /tmp/logs/提交任务的FI用户名/logs/应用ID

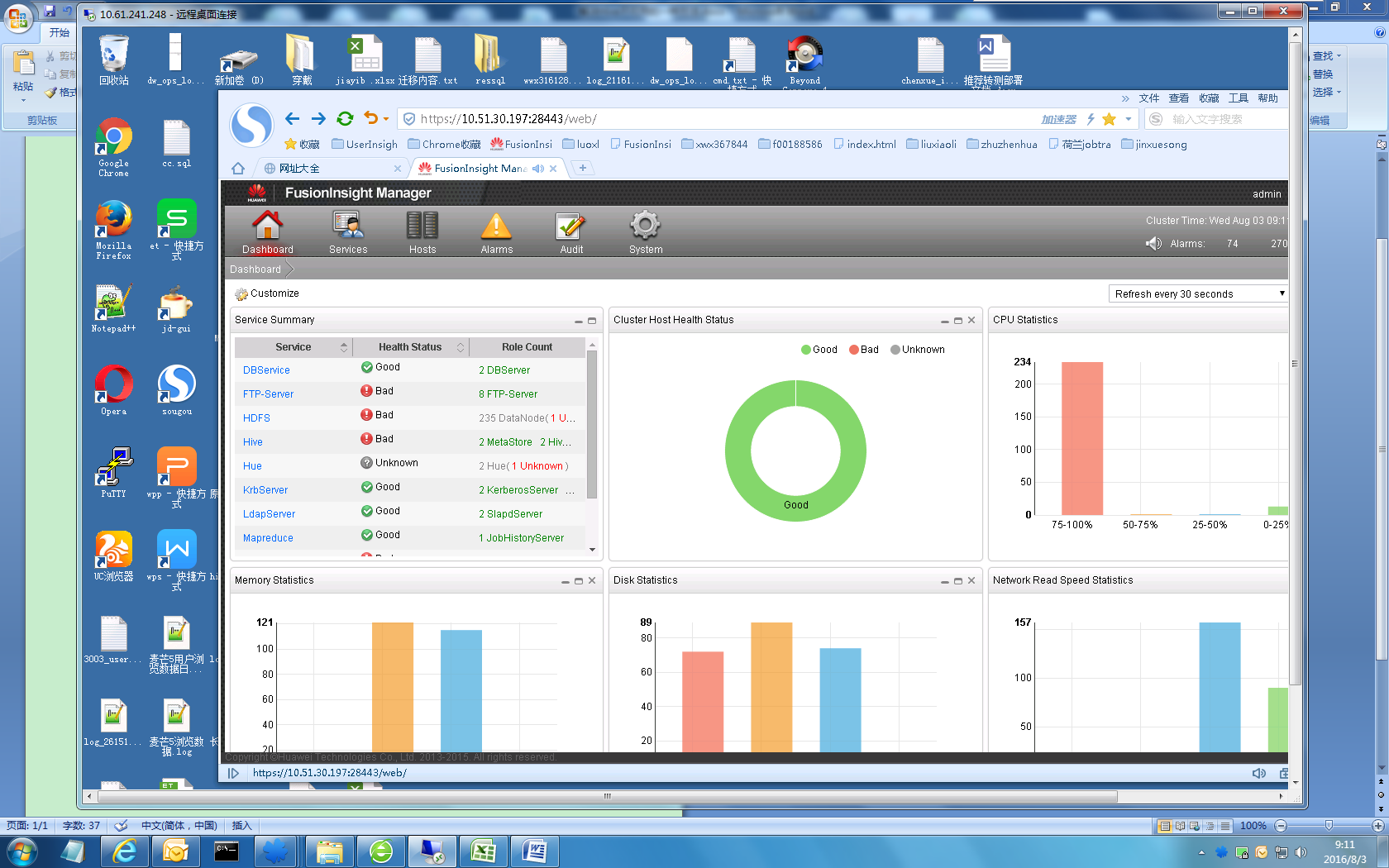
然后找大一点的日志看。

## GaussDB中 The account has been locked

### 现象

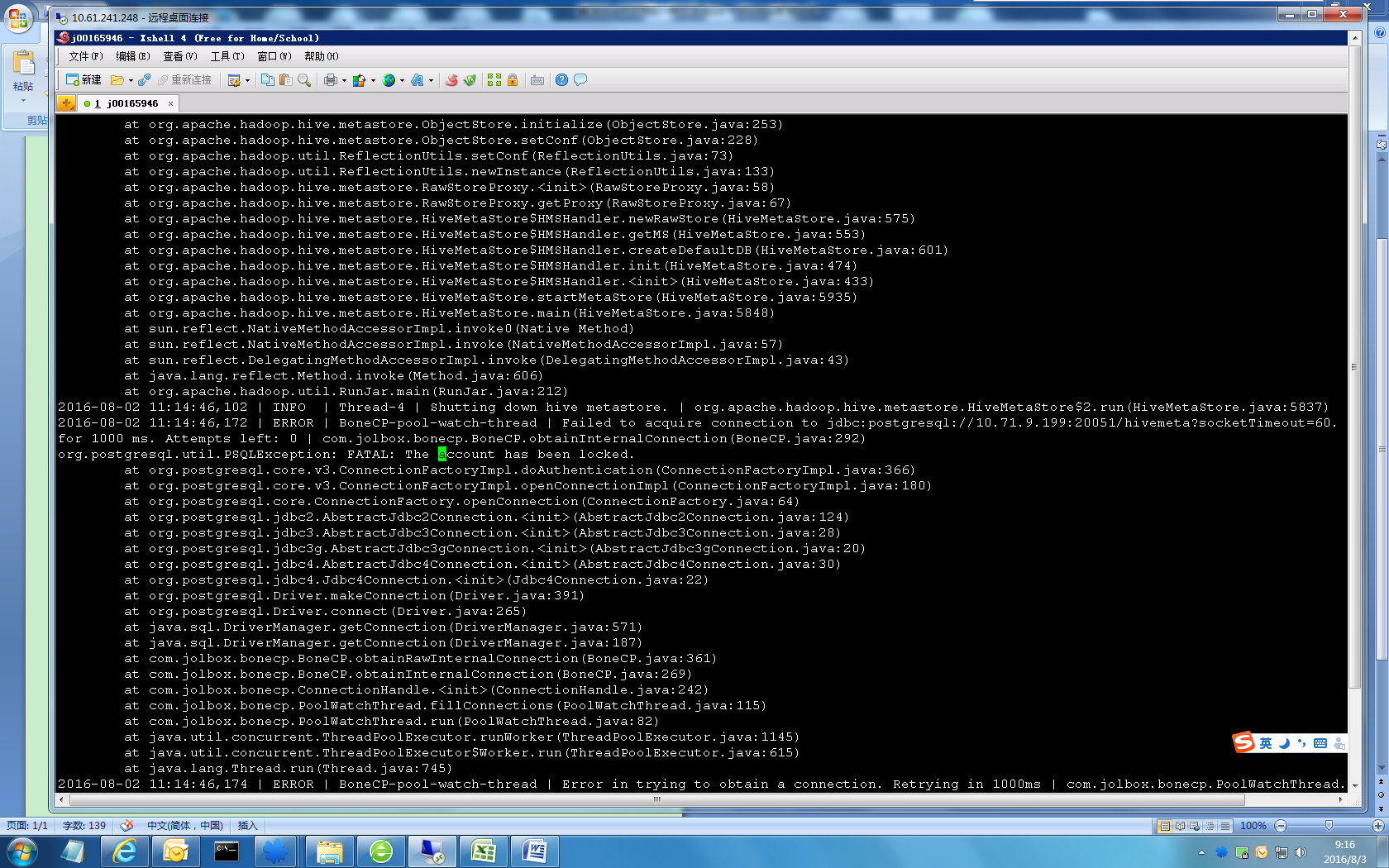
2016年8月2日10:00左右，业务反馈FIC30集群的beeline无法使用，HUE也无法正常向外提供服务。

查看集群管理页面，发现hive的health状态处于bad状态。



然后想到通过重启解决问题，于是重启hiveserver及metastore服务，重启后，问题依然存在。于是再重启一遍，发现无法重启，提示任务队列正在执行，无法重启。

然后进入metastore查看日志，发现日志中提示：The account has been locked。



但是日志没有提示是哪个用户，于是再3ms上查找相关资料。没有重复的案例，找到一个关于spark锁定的案例。在确定解锁账号不会对业务及集群有任何影响后。准备解锁相关账号尝试一下，看是否能够解决问题。

进入DBservice，解锁Spark。然后重启hive集群，问题依然存在。

于是尝试解锁metastore，提示没有该用户。

继续尝试解锁Hive。提示解锁成功。然后重启集群的hive和metastore服务，重启OK。故障消失。

在11:20左右问题解决。

### 解锁账号的方法

解锁方法：

1、登陆dbservice主节点，切换到omm用户

2、导入环境变量

source /opt/huawei/Bigdata/dbservice/.dbservice\_profile;

3、连接DBservice

gsql -p 20051 -d hivemeta -U omm -W ommHuawei@123

4、解锁Hive用户

SET TRANSACTION READ WRITE;   
alter role hive account unlock;

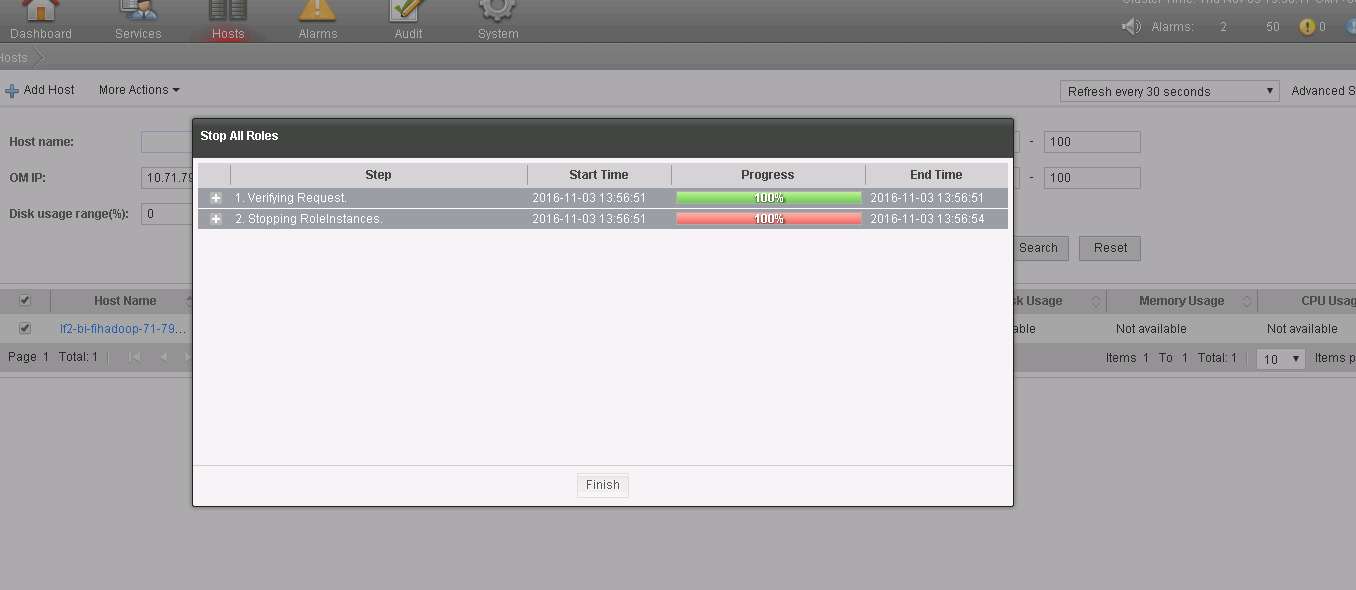
## 机器失联如何找回

### 操作步骤

1. 将机器退服

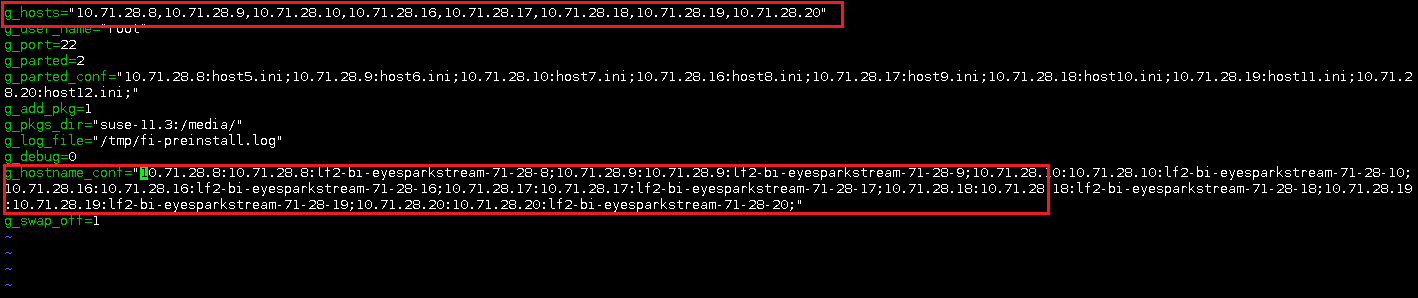
到对应的集群中，主机管理中，找到相应的服务器，删除该服务器





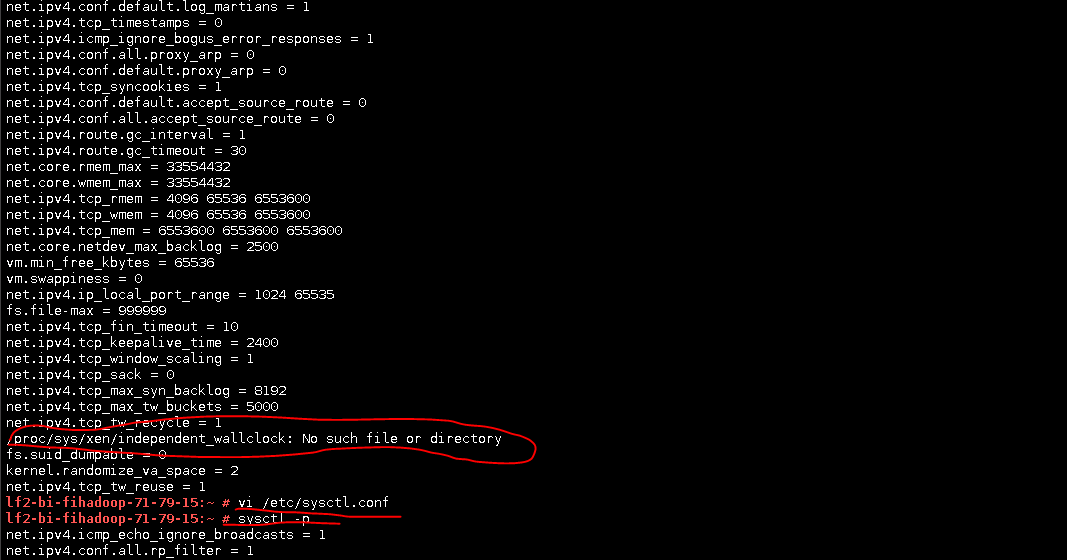
1. 修改preinstall.ini文件

到对应集群管理节点下，进入preinstall.ini,将失联的机器IP，填到g\_hosts、g\_hostname\_count,中



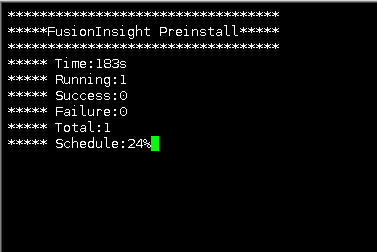
在管理节点下输入命令：df –h，查看挂盘情况，如果有挂盘，要将盘卸掉，输入命令unmount 路径+盘名

进入主机节点执行 sysctl -p,检查/etc/sysctl.conf是否有问题，如果有问题，先解决问题。执行命令 echo ‘omm’ >> /etc/cron.allow ,将omm用户加到/etc/cron.allow中。或直接添加，vi /etc/cron.allow,将omm加到最后一行



1. 执行脚本preinstall.sh

./preinstall.sh





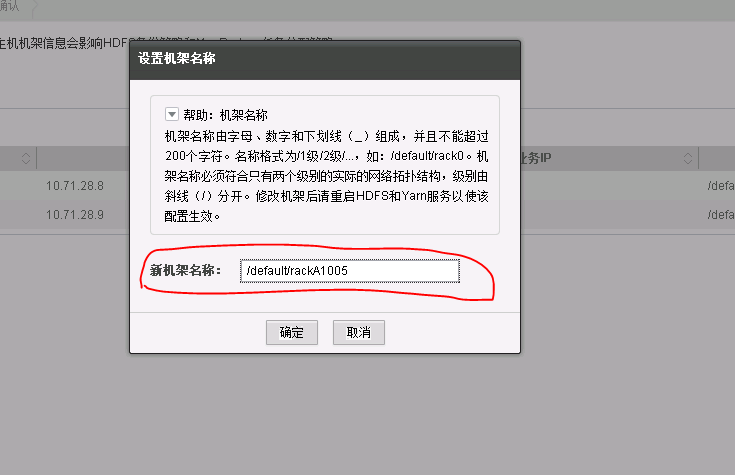
脚本执行完成，到对应的节点下，df –h，查看是否所有盘都加载成功。

1. 重新添加主机

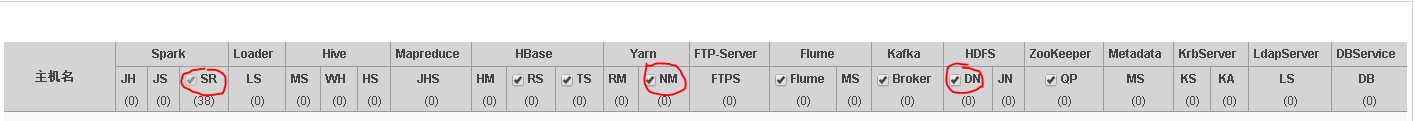
4.1登录对应的集群>主机管理>添加主机，进入第一个页面（查找主机页面），将所有的IP输入‘匹配表达式’中。点击查找主机，下方会出现所添加的所有主机，勾选所有主机，点下一步。



4.2进入第二个页面（设置机架名称页面），勾选全部主机，点击‘设置选中主机的机架名称’，给机架取个名称，点下一步。

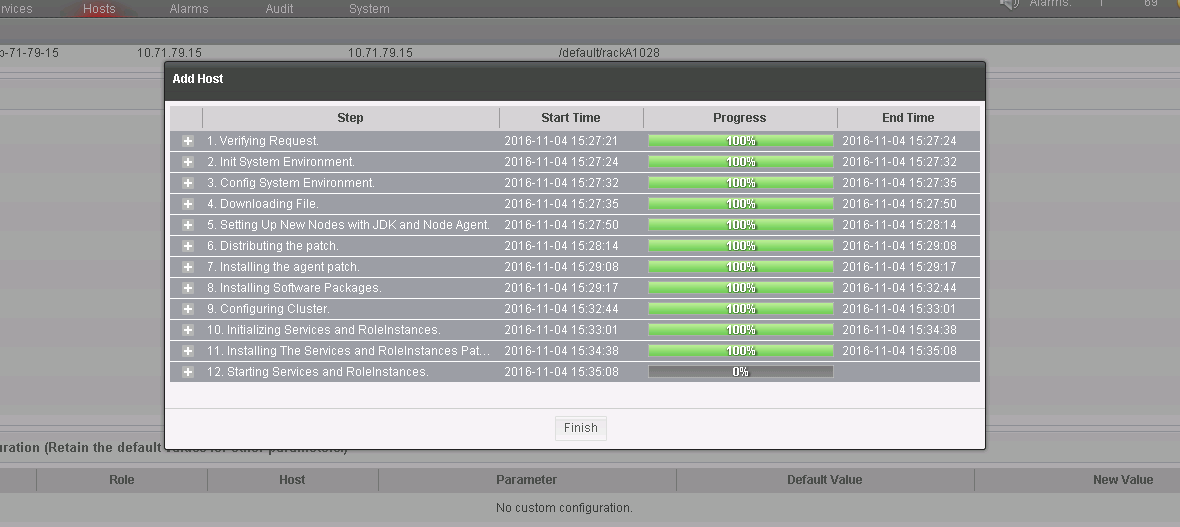


4.3进入第三个页面（定义拓扑页面），给主机选择对应的DataNode、NodeManager和SparkResource，点‘下一步’。



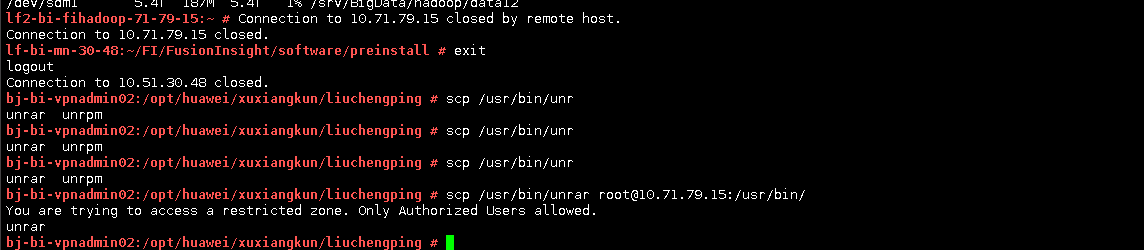
4.4进入第四个页面（配置页面），直接点击下一步。

4.5进入最后一个页面（确认页面），核对相关信息，确认无误后，点击‘提交按钮’。



### 打lzo补丁(C30特有)

在主机节点上下载lzo.zip压缩文件，解压文件中的文档





输入执行命令

cp lib64/liblzo\* /usr/lib64;

chmod og+r /usr/lib64/liblzo\*;

cp lzop /usr/local/bin;

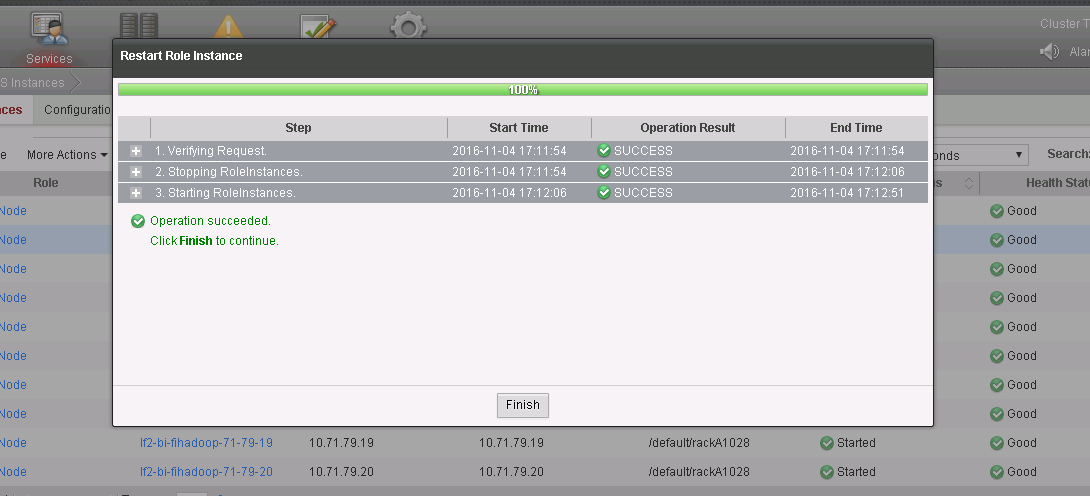
chmod o+x /usr/local/bin/lzop;

ln -s /usr/lib64/liblzo.so.1.0.0 /usr/lib64/liblzo.so;

ln -s /usr/lib64/liblzo.so.1.0.0 /usr/lib64/liblzo.so.1

1. 重新启动机器

登陆集群，进入主机管理，找到对应的机器，重新启动。



## Web ui停止，怎么启动?

C30当前是root用户启动的web ui。必须用root启动(一般是用omm用户启动)，因为文件目录数组不知道怎么改成root了。

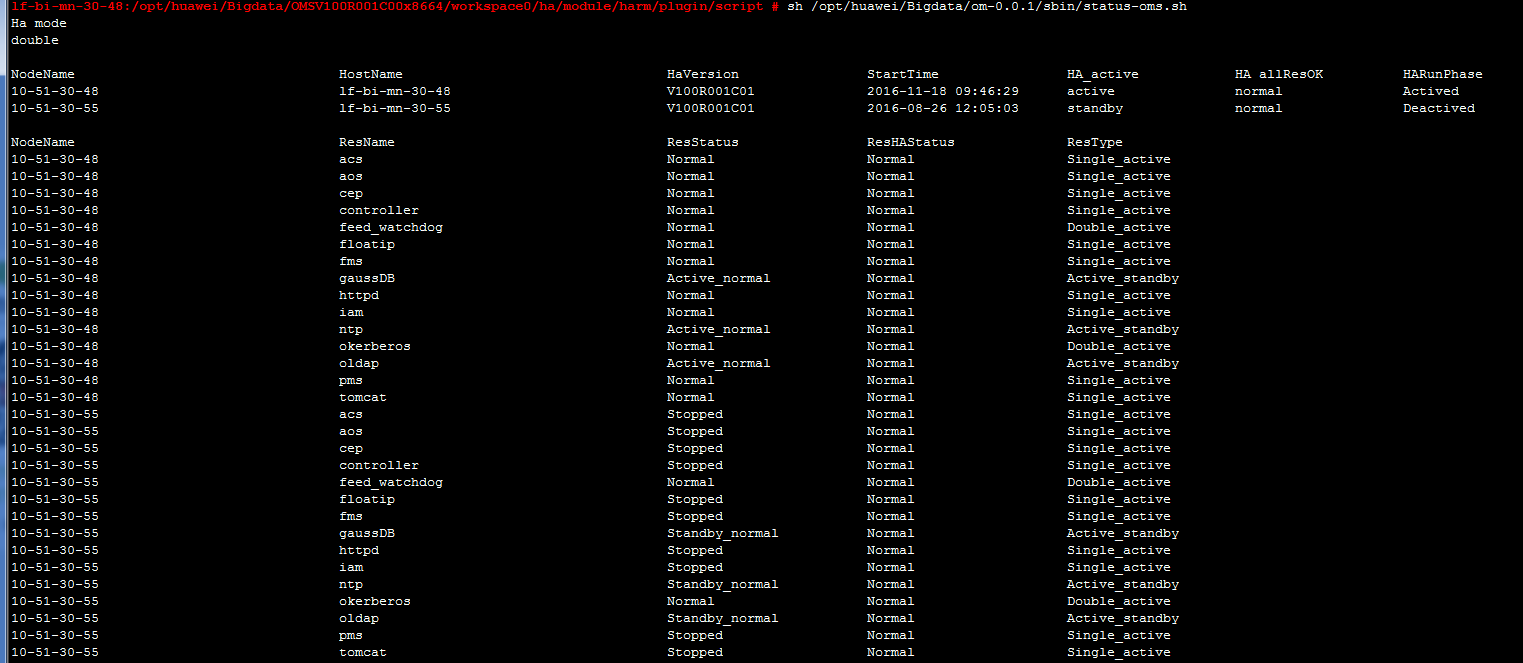
cd /opt/Huawei/Bigdata/om-0.0.1/sbin

sh ./stop-oms.sh

sh ./restart-oms.sh

查看业务状态

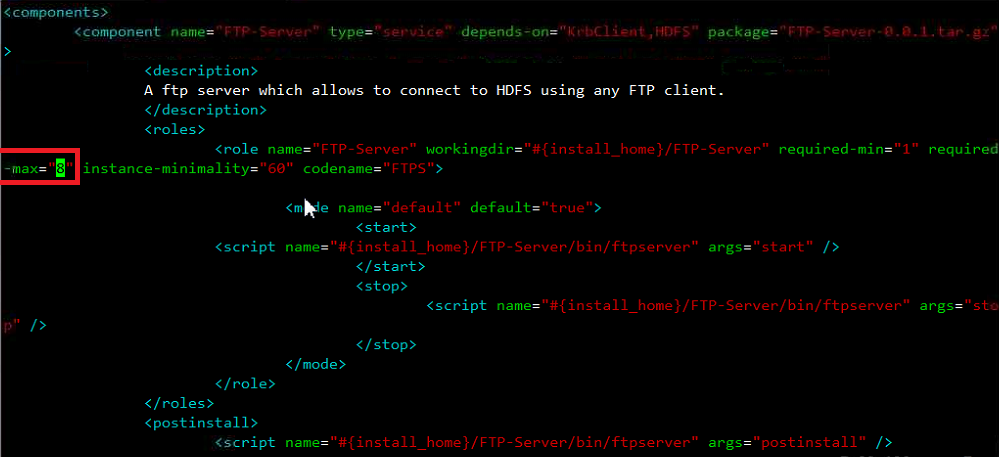
sh ./status-oms.sh



## 无法添加FTP Server服务

C30 ftpserver 最大只有8个，配置在这里

vi /opt/huawei/Bigdata/om-0.0.1/etc/components/FusionInsight\_HD\_V100R002C30SPC603/FTP-Server/components.xml



修改后，重启controller

cd /opt/huawei/Bigdata/om-0.0.1/sbin

如果节点太多，则增加controller内存

vi controller.sh

修改配置项“JAVA\_HEAP\_MAX”的参数值。例如，集群中包含了400个以上的节点，建议修改如下，表示Controller最大可使用8GB内存：

JAVA\_HEAP\_MAX=-Xmx8192m

再到备机修改controller.sh的内存

sh ./restart-controller.sh

执行sh /opt/huawei/Bigdata/om-0.0.1/sbin/status-oms.sh，查看Controller的“ResHAStatus”是否为“Normal”，并可以重新登录FusionInsight Manager表示重启成功

主备会自动同步，无需再更改备机

## FI web触发错误的命令，如何终止

找到主管理节点，执行



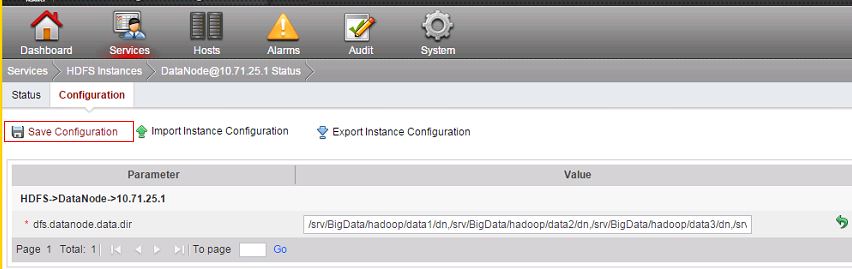
终止FI界面上的任务

## hdfs如何关闭对坏盘的检测？

当有坏盘时，会触发hdfs datanode重启。

登录坏盘节点，执行mount ，看到盘与挂载目录的关系。

点击有坏盘的节点



修改此字符串，去掉有坏盘对应的挂载点。

然后保存配置(不可勾选重启服务)， 单独重启此节点的datanode服务。

## 为什么yarn application –list会 Failing over to ?

主备机制，每个点分配一个id，客户端连接时，先连接最小的id，失败，再连接大的id。 如果恰巧，主机在小id上，则不会报这个错误；主机在大id上，就报这个错误。



## C60 Hive DB改名字的实施方案

**1,DB 改名字**

update DBS set NAME='nemonlou' where name='nemon';

update  PARTITION\_EVENTS set NAME='nemonlou' where name='nemon';

update TAB\_COL\_STATS set NAME='nemonlou' where name='nemon';

update PART\_COL\_STATS set NAME='nemonlou' where name='nemon';

**2，DB改location**

select char\_length(DB\_LOCATION\_URI) from dbs where name='nemonlou'; ---拿到的长度为45，另外31是db的id

update sds set location = overlay(location placing 'hdfs://hacluster/user/hive/warehouse/nemonlou.db' from 1 for 45) where sd\_id in (select sd\_id from tbls where db\_id = 31) and char\_length(location) >45 and  substring(location from 1  for 45)='hdfs://hacluster/user/hive/warehouse/nemon.db';

update SKEWED\_COL\_VALUE\_LOC\_MAP set location = overlay(location placing 'hdfs://hacluster/user/hive/warehouse/nemonlou.db' from 1 for 45) where sd\_id in (select sd\_id from tbls where db\_id = 31) and char\_length(location) >45 and  substring(location from 1  for 45)='hdfs://hacluster/user/hive/warehouse/nemon.db';

update dbs set DB\_LOCATION\_URI = 'hdfs://hacluster/user/hive/warehouse/nemonlou.db' where db\_id = 31;

**3,修改HDFS 路径：**

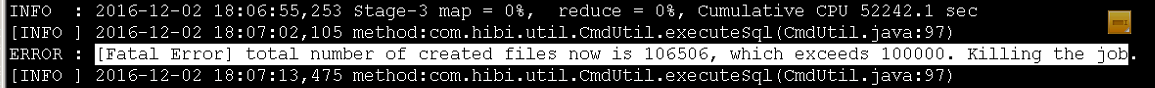
hdfs dfs -mv /user/hive/warehouse/nemon.db /user/hive/warehouse/nemonlou.db

## hdfs/hadoop 命令报内存溢出

dfs命令默认128M内存，加红色参数增大内存

cid:image003.jpg@01D24C63.89260810

## total number of created files now is \*\*\*\*\*



hive.exec.max.created.files 控制select过程中出现的文件个数给10万个。该参数在FI界面不可见。

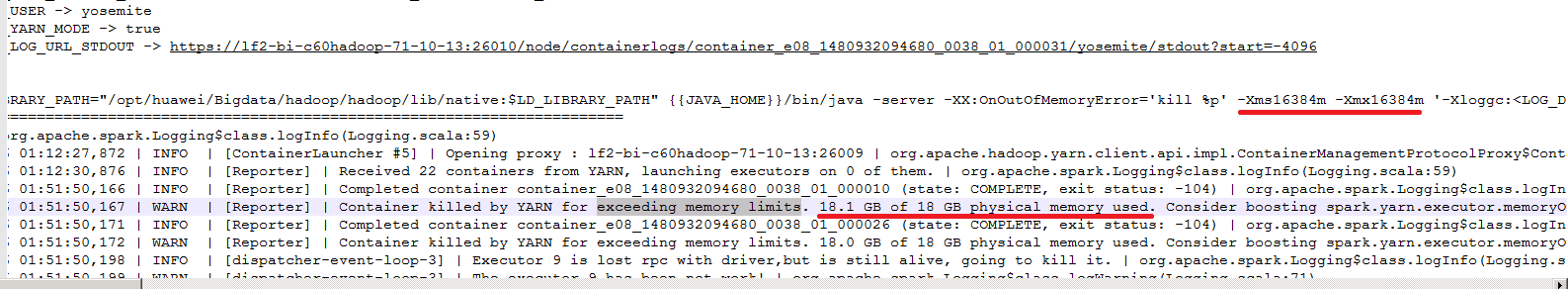
不该修改此参数，应该减少文件数。 修改脚本加distribute by

## 为什么spark任务提交时写明占用16G内存，实际会超过一些？

提交指名占用内存16G，实际超过18G才报错。

因为系统会多占一点内存，多占的内存配置在这里

spark.yarn.executor.memoryOverhead  
spark.yarn.driver.memoryOverhead

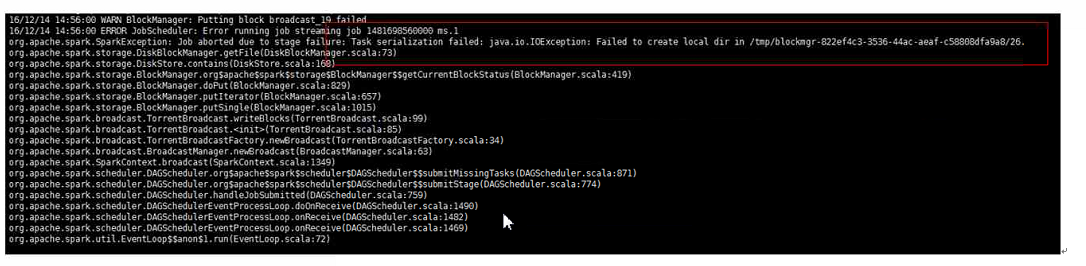


## unable to close file becuase the last block does not have enough number of replicas

* + - 1. 磁盘剩余空间不够。个别磁盘要被写
      2. 写入磁盘后，hdfs没回应，beeline认为写入失败。sql加set dfs.client.block.write.locateFollowingBlock.retries=6；

## 还没解决，先记录

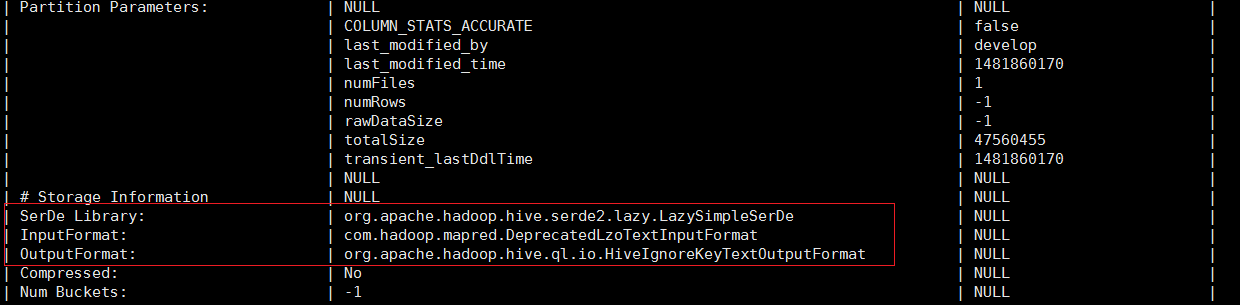
spark streaming报错



## 变更表的存储格式后，数据无法正常查询

变更表的存储格式，修改InputFormat和OutputFormat, 但是SerDe保持老格式没变。所以要显示的修改SerDe

desc formatted ods\_dev\_cloud\_folder\_app\_view\_dm partition(pt\_d='20161214')



alter table ods\_dev\_cloud\_folder\_app\_view\_dm set SerDe 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe';

alter table ods\_dev\_cloud\_folder\_app\_view\_dm partition(pt\_d=20161214) set SerDe 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe';

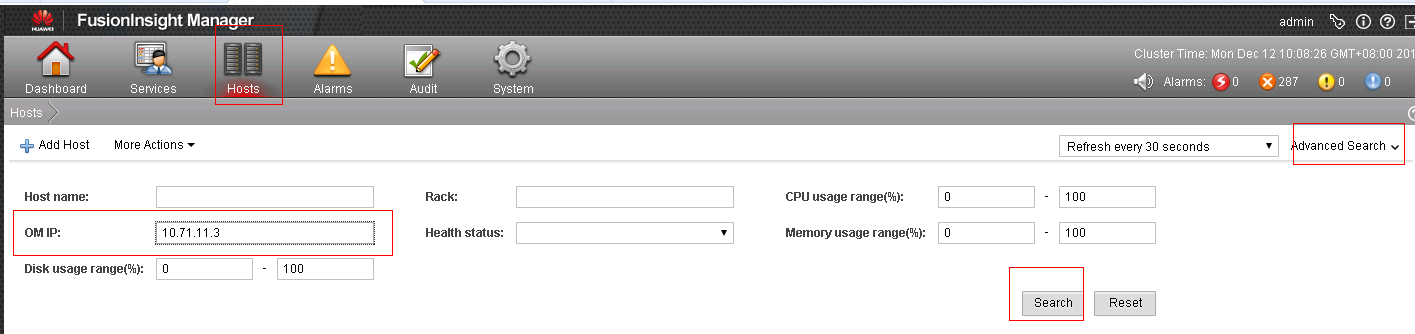
alter table ods\_dev\_cloud\_folder\_app\_view\_dm partition(pt\_d=20161215) set SerDe 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe';

## sql优化方法

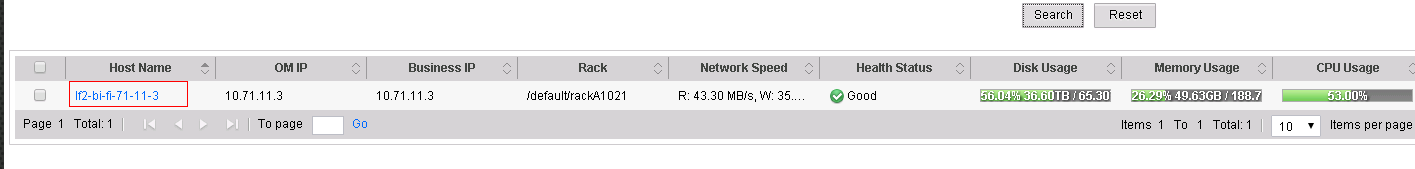
优化前：  
SELECT COUNT( DISTINCT id ) FROM TABLE\_NAME WHERE ...;  
优化后：  
SELECT COUNT(\*) FROM (SELECT DISTINCT id FROM TABLE\_NAME WHERE … ) t;  
优化后的效果：分两轮处理，先获取不同的ID，然后再去count，并发度获得提升。

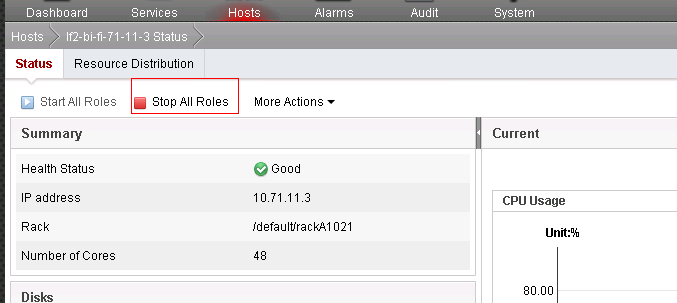
## 坏盘替换方法

1. 停止主机服务
2. 登录相应集群>主机管理>高级搜索>管理IP，输入需要替换磁盘的IP，点击搜索

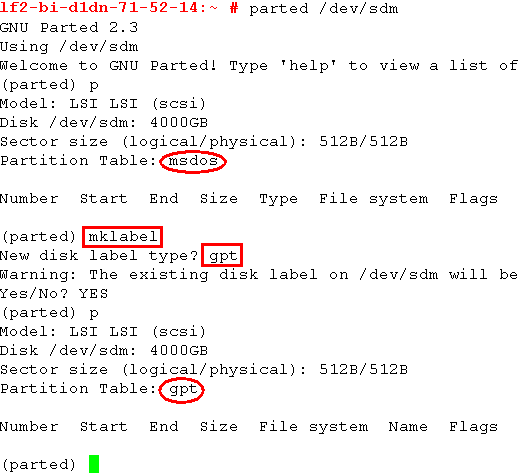


1. 停止所有角色

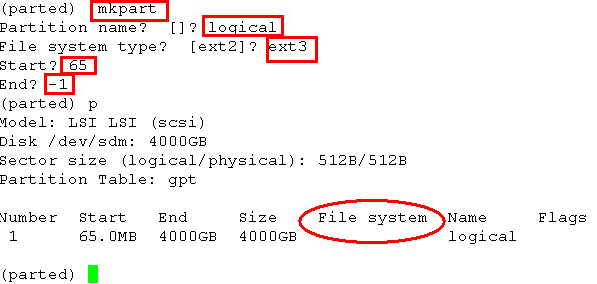




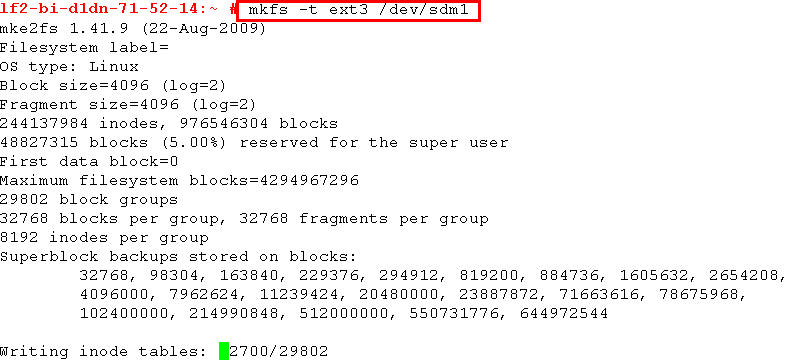
1. 通知钱晓瑞重启机器
2. 格式化磁盘
3. 使用如下命令更改分区格式为GPT



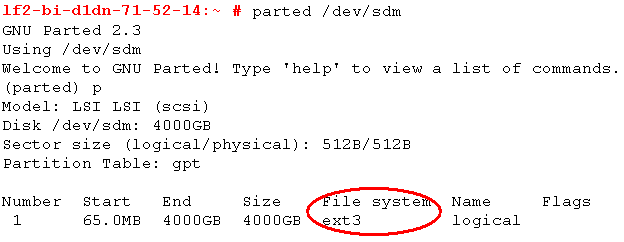
1. 创建分区，注意：parted命令不支持ext3文件系统（分区后图中File system为空），分区之后需要重新格式化



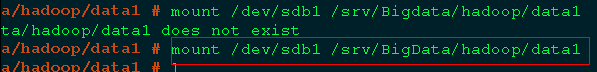
1. 格式化分区为ext3

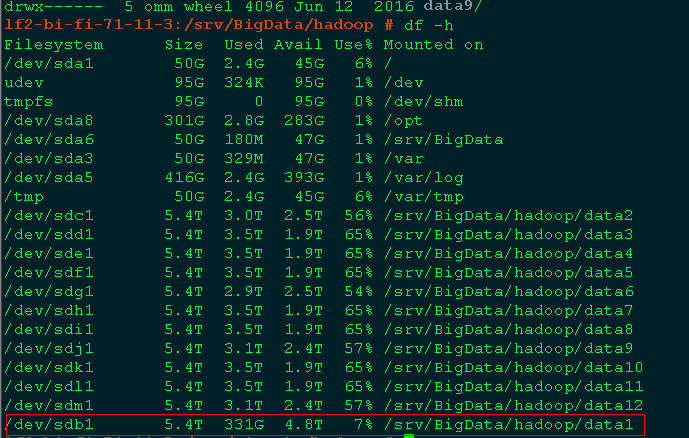


1. 检查格式化结果



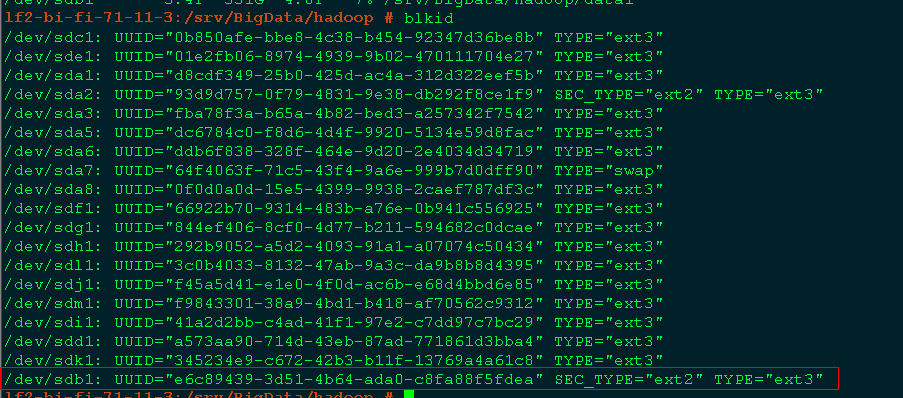
1. 挂载磁盘



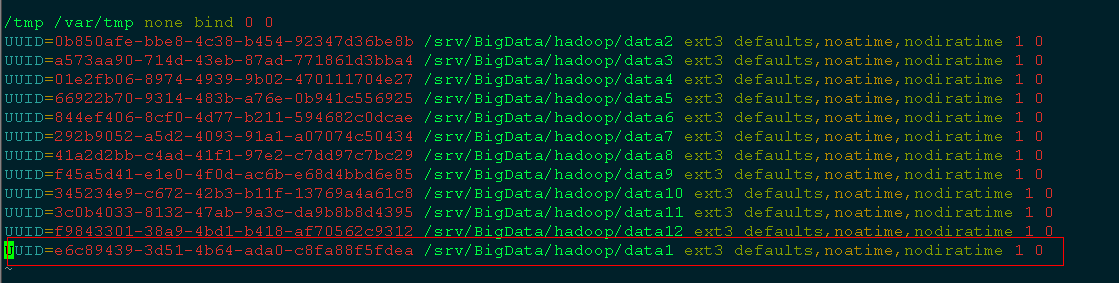


1. 配置/etc/fstab文件

执行命令：blkid，找到新格式化的磁盘对应的UUID

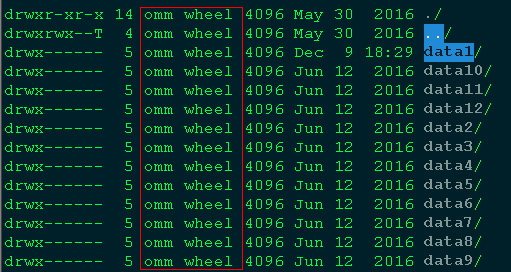


修改/etc/fstab文件，将UUID和对应的挂载目录配置进去



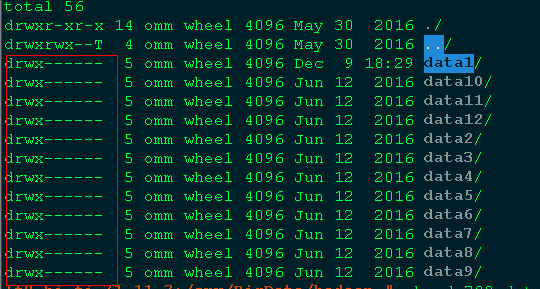
1. 修改属主，与其他盘保持一致



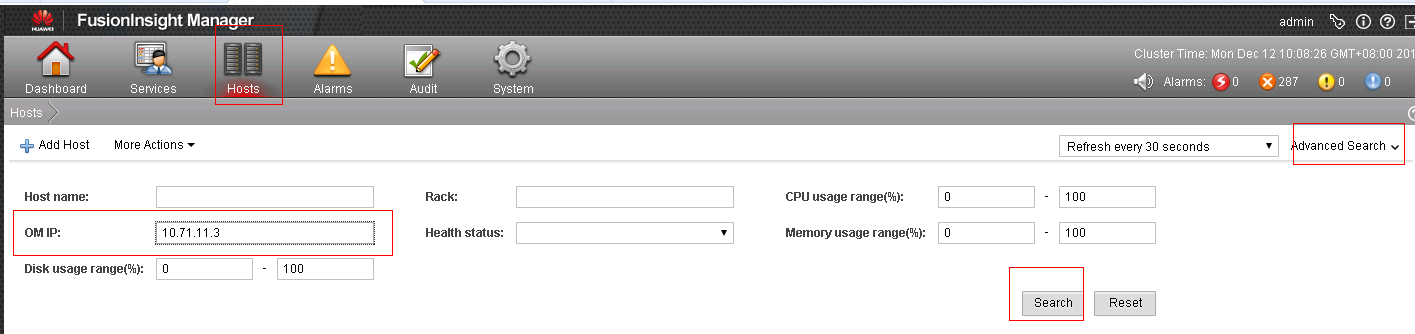


1. 修改目录权限，与其他盘保持一致

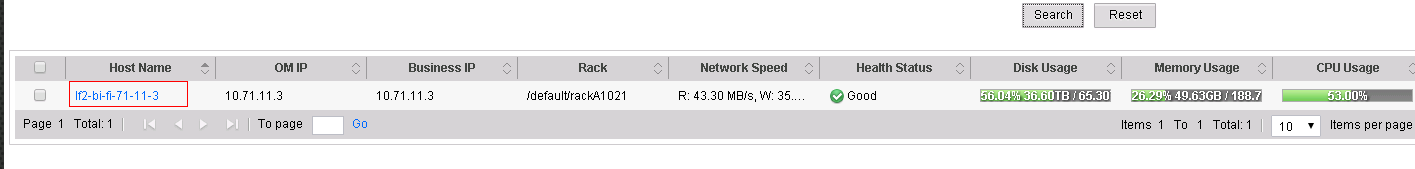


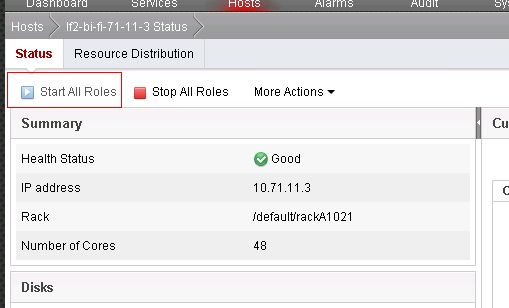


1. 重启服务
2. 登录相应集群>主机管理>高级搜索>管理IP，输入需要替换磁盘的IP，点击搜索



1. 重启服务





## C30版本增大高斯数据连接数方法

1. 登陆主DBServer节点，切换到omm用户：su – omm;

2. 打开配置文件/opt/huawei/Bigdata/dbservice/setup/conf/postgresql.conf

3. 修改配置文件中的max\_connections参数（连接数上限）；

4. 修改完成后，执行重启gaussdb操作，参考命令：

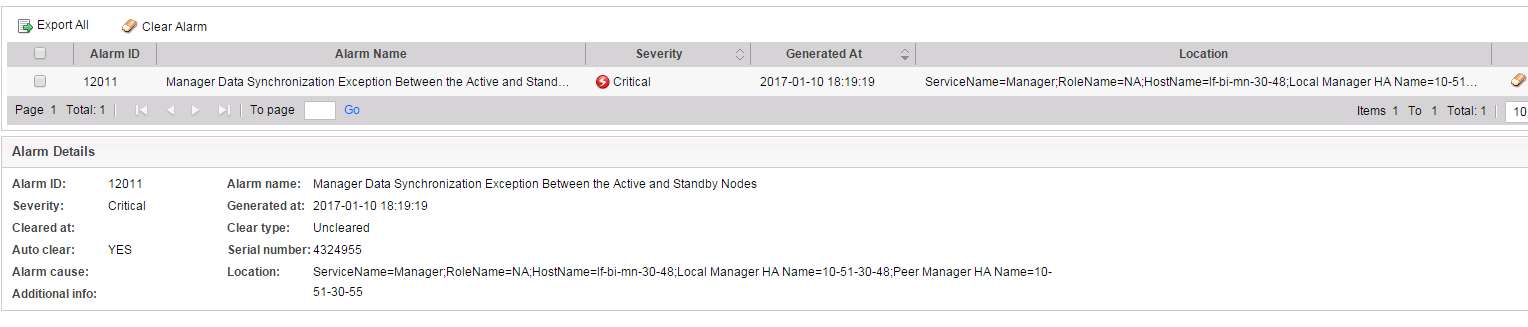
source /opt/huawei/Bigdata/dbservice/.dbservice\_profile

gs\_ctl -D $GAUSSDATA restart

5. 登陆备DBServer，执行1-3步同步修改备DBServer的内容；

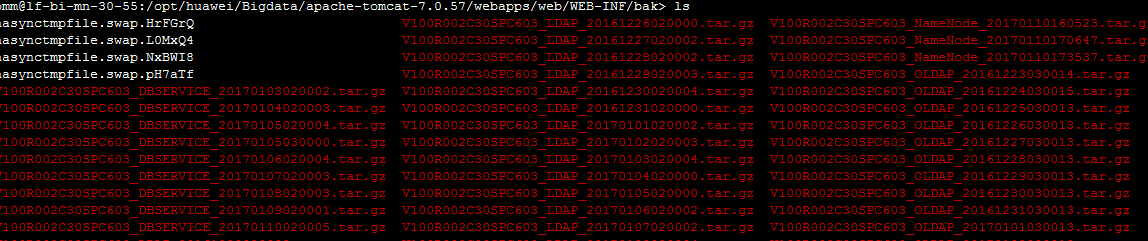
注：max\_connections参数可以调整为500-800，与系统的信号量设置有关系。

## 告警12001处理方法



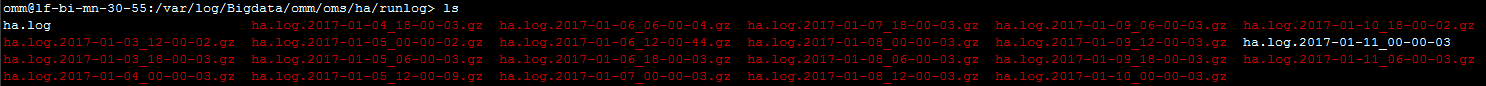
管理节点上报的告警。

各服务的数据，会定时备份到管理节点。备份到此目录。



管理节点主备均要复制此文件。复制超时，会有这个告警。

进入此目录，查看ERROR日志，看具体是哪个业务告警。



## node fault告警无法恢复

查看告警节点

下面的日志会周期性的出现， 如果连接管理节点的日志是正常的，则手动恢复告警。



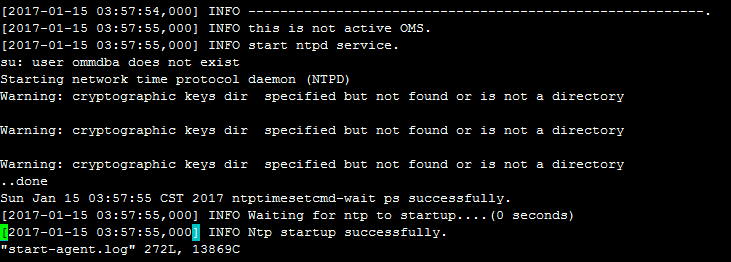
原因：

oms 管理节点重启过，C30管理节点重启，恢复告警会丢失，界面无法自动恢复，实际节点已经正常。

oms管理节点重启时，会告诉所有的节点重启nodeagent，可以看日志



最后一次启动时间和管理节点重启的时间一致。



## 表变更属组

用admin建立的表想让其他用户拥有完全权限。

1.table 属组变更为develop,可以alter， drop

source /opt/huawei/Bigdata/dbservice/.dbservice\_profile;

gsql -p 20051 -h 160.161.0.145 -U hive -W HiveUser@ -d hivemeta

select \* from DBS;

update tbls set OWNER='develop' where OWNER='admin' and DB\_ID='106' and TBL\_NAME='t6' ;

2.补充表权限，develop 可以insert

grant all on table t6 to user develop;

3.database属组变为 develop,默认可查询

alter database db\_c00239107 set owner user develop;

grant all on database db\_c00239107 to user develop WITH GRANT OPTION

不加WITH GRANT OPTION，以后不能执行 show create table \*\*.

4. hdfs 目录要组内可读.模拟push工具，能放文件进去

admin hadoop

/user/hive/warehouse drwxrwx---+

/user/hive/warehouse/db\_c00239107.db drwx------+

/user/hive/warehouse/db\_c00239107.db/t2 drwx------+

模拟push工具，能放文件进去

/user/hive/warehouse/db\_c00239107.db/t2/b.txt -rw-r-----+ (develop hadoop)

还有问题，就检查 hdfs acl权限

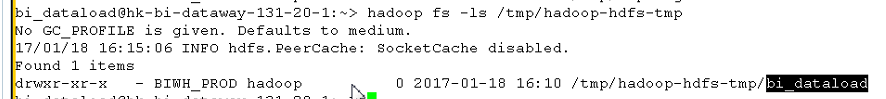
设置：

hdfs dfs -setfacl -m user:hdd1:rwx /user/hive/warehouse/3rdparty.db

清理：

默认的ACL ==>hdfs dfs -setfacl -k -R /user/hive

datapush 目录要变更属组



## 慢盘检测方法

iostat -x -t 1 600

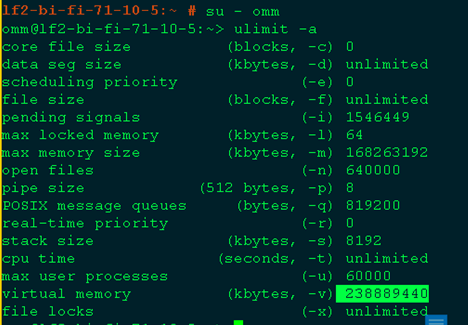
<http://3ms.huawei.com/hi/group/2692269/blog_2093127.html?mapId=3309917&for_statistic_from=mail_share>

## admin重做密码

1：su - omm (如果已经在omm用户下忽略此步骤)  
2：source /opt/huawei/Bigdata/om-0.0.1/meta-0.0.1-SNAPSHOT/kerberos/scripts/component\_env  
3：kadmin -p kadmin/admin  
默认密码是[Admin@123](mailto:Admin@123)，输入完以后需要修改密码,请记住这个密码，下次执行kadmin -p kadmin/admin后输入密码就不是默认密码[Admin@123](mailto:Admin@123)，而是此次改的密码  
4：cpw -pw password admin  
其中password是你要设置的新密码，admin需要修改的用户

## 应用起不来，内存不够

virtual memor 238889440



在 /etc/profile中添加 ulimit –v 238889440，重启nodeagent 。因为nodeagent的从参数还是老的，没有生效，需要重启进程生效。

## hdfs文件不可读

hdfs ckfs /usr/ddd/\*\* 报文件正在写，无法去读。

从hdfs web界面，下载该文件失败。 该文件是flume写入的。

强制去掉写锁:

hdfs debug recoverLease [-path <path>] [-retries <num-retries>]

## Kafka topic在同一节点 不同磁盘之间分区迁移的方法

1. 运营流出现kafka磁盘容量不足的告警。发现该磁盘上分布了5个flumestreaming的topic的分区，每个达1.1T。解决办法是将其中的三个分区迁移到磁盘使用率低的另三个磁盘。步骤如下：

步骤一、登录FusionInsight Manager，单击“服务管理 > Kafka > 实例”，停止该Broker实例。

步骤二、进入“kafka-logs”目录，执行**du -sh \***命令，选择一个待移动的Partition文件夹，其名称命名规则为“Topic名称-Partition标识”，记录Topic及Partition。

步骤三、修改“kafka-logs”目录下的“recovery-point-offset-checkpoint”和“replication-offset-checkpoint”文件（两个文件做同样的修改）。

* 1. 减少文件中第二行的数字（若移出多个目录，则减少的数字为移出的目录个数）。
  2. 删除待移出的Partition所在的行（行结构为“Topic名称 Partition标识 Offset”，删除前先将该行数据保存，后续此内容还要添加到目的目录下的同名文件中）。

步骤四、修改目的数据目录下（例如/“srv/Bigdata/kafka/data2/kafka-logs”）的“recovery-point-offset-checkpoint”和“replication-offset-checkpoint”文件（两个文件做同样的修改）。

* 1. 增加文件中第二行的数字（若移入多个Partition目录，则增加的数字为移入的Partition目录个数）。
  2. 添加待移入的Partition行到文件末尾（行结构为“Topic名称 Partition标识 Offset”，直接复制步骤三中保存的行数据即可）。

步骤五、移动数据，将待移动的Partition文件夹移动到目的目录下，移动完成后执行**chown omm:wheel -R** *Partition目录*命令修改Partition目录属组。

步骤六、登录FusionInsight Manager，单击“服务管理 > Kafka > 实例”，启动停止的Broker实例。

步骤七、等待5-10分钟后查看Broker实例的健康状态是否为“良好”。

## Kafka topic在不同节点之间分区迁移的方法

## 对特定 Kafka topic 设定日志保存时间

**现象**：个别topic数据增长量太大。多数topic 日志增长类似 。需要缩短个别topic的日志保存时间。

**步骤**：

对topic\_source保留数据周期为5分钟

kafka-topics.sh --zookeeper $zk --alter --topic topic\_source --config retention.ms=300000

**观察**：

日志时间删除时间= log.retention.check.interval.ms(在FI kafka配置界面) + retention.ms(topic 单独的配置)

## 查询表类型和路径

Step 1) 登录OM页面-->DBService-->Configuration-->All-->HA   找到dbservice.floatip:160.161.0.145

Step 2) 用OM用户执行source /opt/huawei/Bigdata/dbservice/.dbservice\_profile

Step 3)连接到DB: gsql -p 20051 -h 160.161.0.145 -U hive -W HiveUser@ -d hivemeta

select t1.TBL\_ID, t1.SD\_ID,t1.TBL\_NAME, t1.TBL\_TYPE,t2.location

from TBLS t1

left outer join

SDS t2

on t1.SD\_ID = t2.SD\_ID;

## 管理节点ssh服务被停止导致集群告警

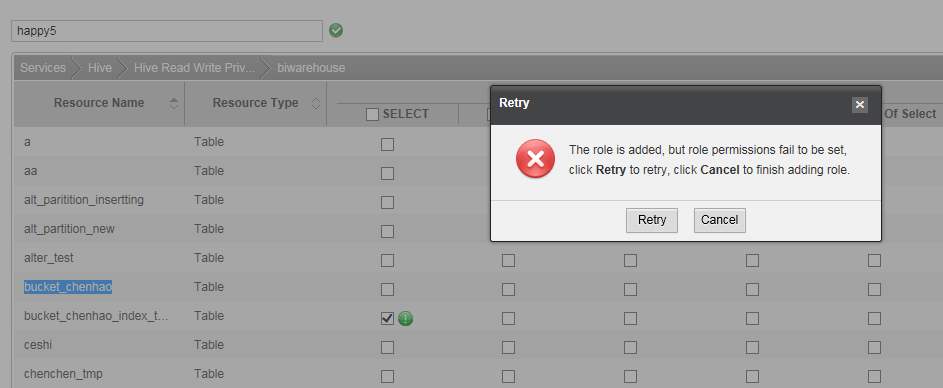
1. dataone集群出现大量致命告警。最早的告警为KrbServer不可用。

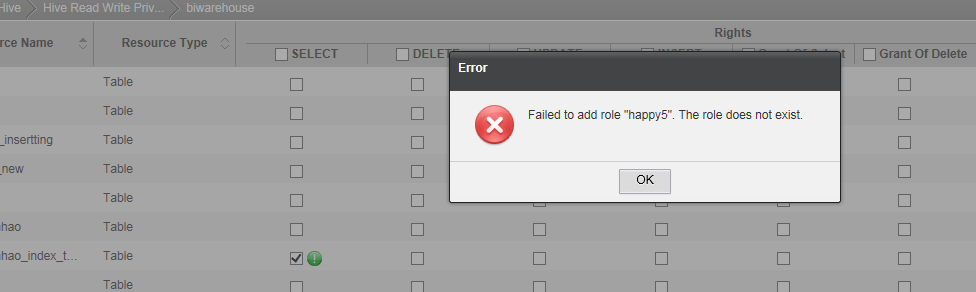
执行kinit admin正常，查看kerboros服务，日志也正常。然后发现ssh不到管理节点浮动IP，但可以ping通浮动IP。 检查发现FI管理节点的主节点的sshd服务没有起来，执行service sshd restart后，告警逐渐消失。



## FI界面添加hive的角色失败

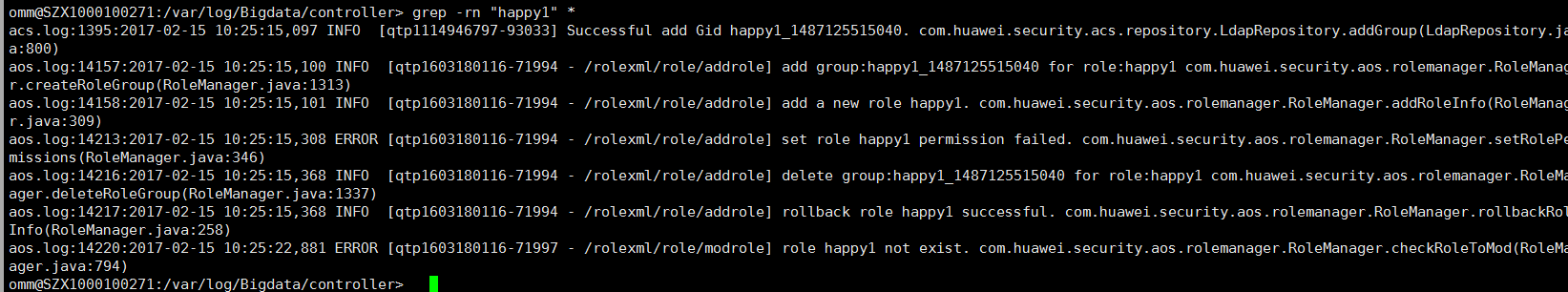
添加hive角色时，对于选中的表报如下错误





该表在hive中存在， 在hdfs的目录已经被删除。 授权本质是给hdfs目录加权限。访问hdfs目录不存在所以报错。

在FI portal所在的服务器搜索新加的角色名，可见添加角色成功，授权时失败。



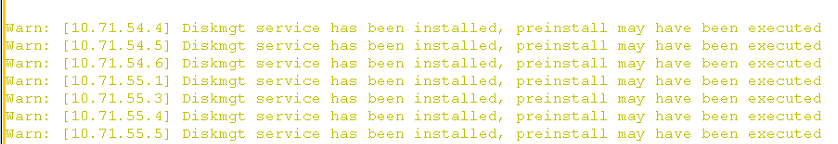
## 浮动IP和下面绑定的服务器IP要在同一个网段

dataone集群退服4T机器时，将Loader所在机器10.71.53.12和10.71.53.13也退服了，导致Loader服务不可用。于是在另外两个数据节点10.71.78.2和10.71.78.4上重新安装了Loader服务，仍然使用旧浮动IP10.71.53.199。但重新安装后，Loader不能正常工作，一直做主备倒换，WebUI打不开。从管理节点10.71.53.8上ping不通浮动IP10.71.53.199，也ping不通浮动IP所在机器的IP，当可以ping通浮动IP不在的另一台Loader服务的机器。将浮动IP改为浮动IP10.71.53.198也不能正常工作。将浮动IP改为浮动IP10.71.78.199（事先已经确认该IP没有使用，ping不通）后，Loader工作正常。原因是浮动IP需和下面绑定的服务器IP要在同一个网段。

## 安装FI节点时，报磁盘容量不匹配

**现象：**

有的节点是以前安装过FI软件的，再次安装时，会报磁盘容量不够。



**原因：**

diskmgt存在， 则安装脚本认为已经格式化过了，不会重新按照Excel配置格式化磁盘。而老的格式化并不能满足当前安装的要求。

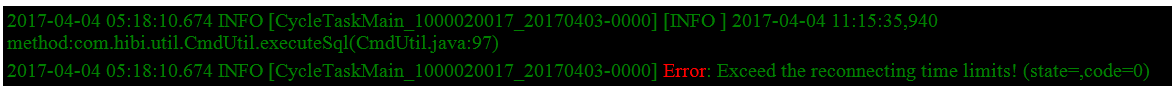
**解决方法：**

root登录已经未格式化的节点，执行 /usr/local/diskmgt/script/uninstall.sh，重新再安装

## C30报Exceed the reconnecting time limits

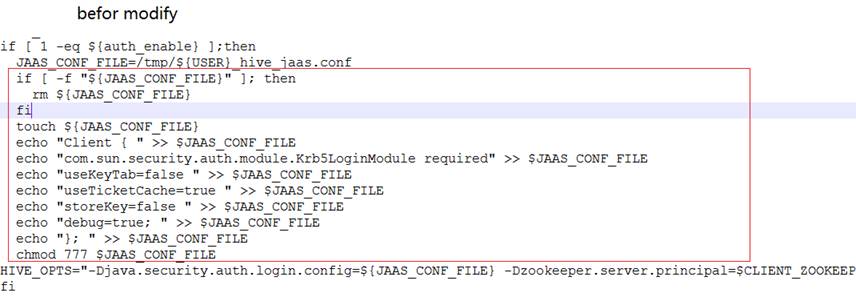
### 现象

TCC执行脚本报如下错误。



### 原因：

$FI\_client/Hive/Beeline/bin/beeline

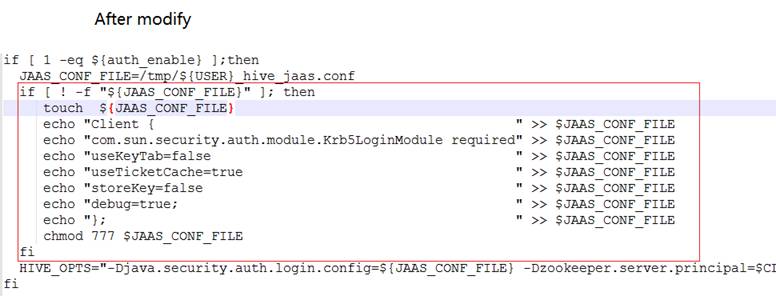


beeline执行过程中会创建 jaas.conf文件，当同一FI 用户同时运行beeline命令时，上图代码可能导致有的命令的jass.conf文件为空，导致鉴权失败。

C30SPC602已经解决了这个问题。

### 解决方法：

改成这样：



## C60任务提交到已经删除的队列

OneData集群的一个角色配置了一个队列，这个队列删除后，RM的配置文件中任然有角色与队列的映射关系。导致某个用户分配新队列后，任务任然提交到这个旧队列而失败。

临时解决办法：删除主备RM节点对应配置文件中的信息，然后在Web页面执行刷新队列后，任务就可以提交到正确的队列了。

**问题原因：**

创建租户tenant时，同步生成tenant\_123456租户角色。

新建角色绑定租户队列tenant，会在queue-mappings中产生该角色与tenant的映射。

而删除租户时，只删除了tenant\_123456与tenant的映射关系。新建角色与tenant的映射信息残留下来，触发了该问题。

**目前规避方案：**

访问om controller数据库，手动清理yarn.scheduler.capacity.queue-mappings\_capacity-scheduler配置中的残留信息，重启controller，并同步配置，重启yarn服务。

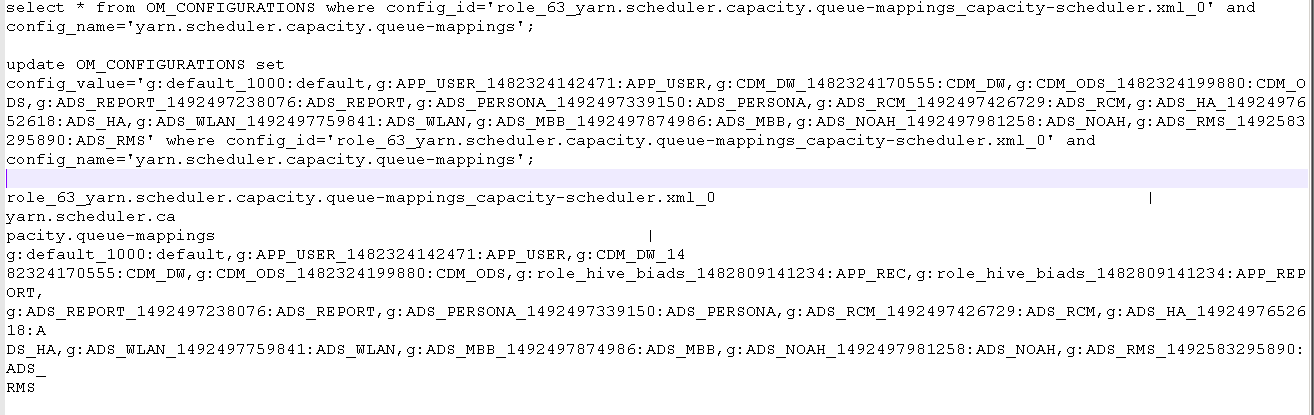
在管理节点执行：

gsql -U omm -W [ommHuawei@123](mailto:ommHuawei@123) -p 20015

select \* from OM\_CONFIGURATIONS;

select \* from OM\_CONFIGURATIONS where config\_id='' and config\_name='';

update OM\_CONFIGURATIONS set config\_value='' where



**建议：**

后续使用yarn队列，直接绑定创建租户时生成的默认租户角色即可，无需再创建角色去绑定该队列。

## C30从开源平台迁移数据的方法

从非安全hadoop集群迁移数据到C30，FI C30版本不支持，需要在搬迁之初替换hadoop-hdfs-V100R001C00.jar,搬迁完毕后，恢复该jar包

需要的文件



**客户端更换方法：**

  1.在FI管理节点新建hadoop用户，在此用户下安装FI客户端，初始化kinit admin账号。

进入fusioninsight客户端安装目录，拷贝补丁包到对应目录：

cp hadoop-hdfs-V100R001C00.jar $FI安装目录/HDFS/hadoop/share/hadoop/hdfs/

cp hadoop-hdfs-V100R001C00.jar $FI安装目录/HDFS/hadoop/share/hadoop/httpfs/webapps/webhdfs/WEB-INF/lib/

cp hadoop-hdfs-V100R001C00.jar $FI安装目录/HDFS/adapter/authorization/controller/lib/

cp hadoop-hdfs-V100R001C00.jar $FI安装目录/HBase/hbase/lib/

2. 在FI管理节点/etc/hosts里面添加开源hadoop平台所有主机名和IP的对应关系

**服务端更换方法**

1. omm用户登录FI管理节点，检查所有节点文件大小

Hdfs 所有实例和yarn所有nodemanager IP写到ips.txt

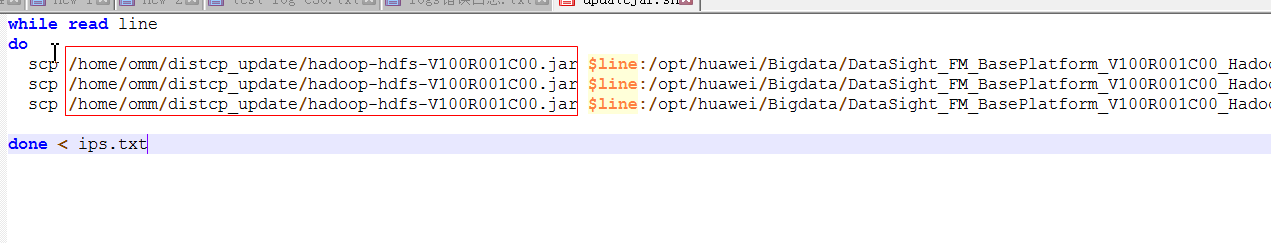
运行check\_file\_size.sh，检查每个节点的hadoop-hdfs-V100R001C00.jar大小是否一致。

如果大小都一样，就备份一个就好了。放在管理节点 /home/omm/distcp\_rollback下

2.omm用户登录FI管理节点，需更新的文件放在/home/omm/distcp\_update

cd /home/omm/

见updateJar.sh，更新时，用红色框中的路径；恢复换成/home/omm/distcp\_rollback/hadoop-hdfs-V100R001C00.jar



1. 在FI WEB界面重启hdfs所有服务(含datanode和namenode)， 重启yarn 所有服务(放resourcemanager和nodemanager).

3. 如果job报not known host错，则需要把开源平台的主机名和ip地址添加到全部datanode和全部nodemanager节点的/etc/hosts文件中

搬迁：

1.在FI 管理节点hadoop用户下，kinit admin，然后执行

**请确保源目录对所有用户可读，如不可读，在开源hadoop平台执行会报http 400错误**

hadoop fs -chmod 755 –R 目录

GC\_PROFILE="custom" GC\_OPTS="-Xmx2000M" hadoop distcp -Ddfs.checksum.type=NULL -Dmapreduce.map.memory.mb=4096 -Dmapreduce.map.java.opts=-Xmx3276M h<ftp://10.161.24.30:50070/AppData/CommonProd/hadoop-NJ/data/DIM/KPI2/dim_cloudservice_kpi2_imei_uuid_device_ds> /AppData/hota/DIM/KPI2/dim\_cloudservice\_kpi2\_imei\_uuid\_device\_ds

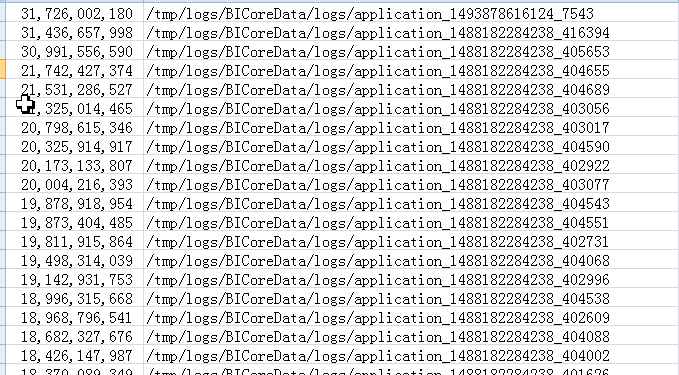
## 如何查询有问题的udf？

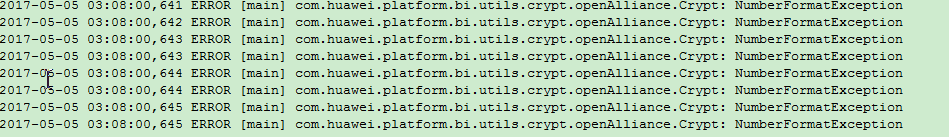
udf功能一般都会正确，不对的地方是拼命打印日志。如何通过打印日志查找有问题的udf

按如下方法查看有问题的UDF。

hdfs dfs -du /tmp/logs/用户名/logs

哪个目录日志大，哪个的udf日志输出就有问题。



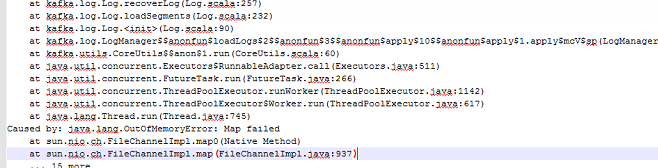
第1个日志，是因为下面的udf拼命打日志  


自我感觉，超过100M的job，在udf使用上一定有问题。

## kafka启动失败，报java.lang.OutOfMemoryError: Map failed

FI界面显示进程未启动， 登录节点 ps-ef| grep kafka ，查不到进程

一般是虚拟机的虚拟内存不够，导致的。



root用户 vi /etc/profile , 最后一行加 ulimit -v unlimited

切换到 omm用户，重启nodeagent

cd /opt/huawei/Bigdata/nodeagent/bin

./stop\_agent.sh

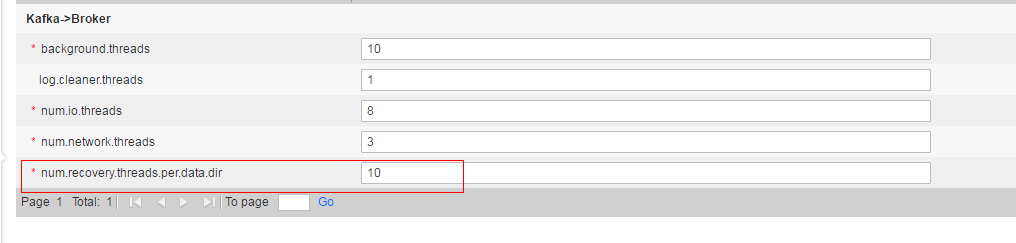
./start\_agent.sh

从FI界面重启kafka

## kafka进程存在，但是健康状态为bad

一般是重启kafka后， 节点上数据太多，加载数据超时(加载超过10分钟表示超时)

修改如下参数，增加加载的线程数量，重启



## FI portal监控数据在哪里?

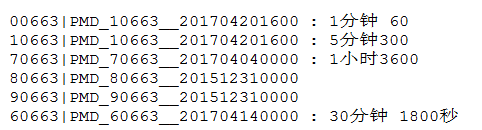
登录FI管理portal

进入/opt/huawei/Bigdata/omm\*\*, 搜索xlsx，这是所有监控指标的定义

gsql -p 20015 -U omm -d omm -W ommHuawei@123 –t

查询 omm\_1下的表，这些都是监控数据。

public是集群组网结构。



## 如何获取 服务的主备关系？

小心点，别在里面乱搞，千万不能执行删除操作。

服务的主备关系是在zookeeper中管理的。 安装FI客户端后，

zkCli.sh -server 10.21.16.237:24002,10.21.17.55:24002



服务管理都在这里。获取hive的主机，命令如下

get /hive/active\_hive

## 分析jhistory的日志。

对job做任何条件限制，分析这个目录下的全部job，转换格式的结果在a.json中

hadoop jar $FI\_Client/HDFS/hadoop/share/hadoop/tools/lib/hadoop-rumen-2.7.2.jar org.apache.hadoop.tools.rumen.TraceBuilder file:///tmp/a.json file:///tmp/b.json hdfs://hacluster/mr-history/done/2017/06/24/000000

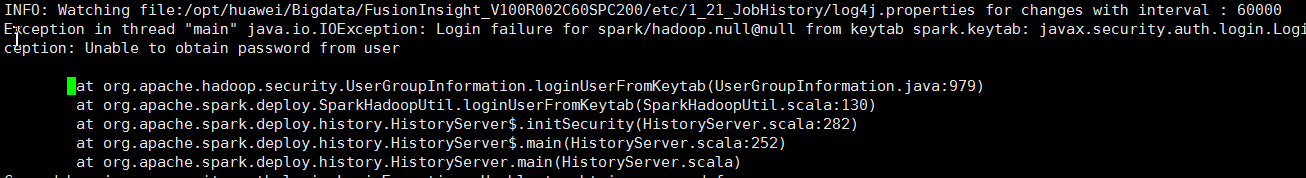
Folder 等比例缩小日志规模

hadoop jar hadoop-rumen-2.7.2.jar org.apache.hadoop.tools.rumen.Folder -output-duration 1h -input-cycle 20m file:///tmp/a.json file:///tmp/job-trace-1hr.json

## sparkresource无法启动



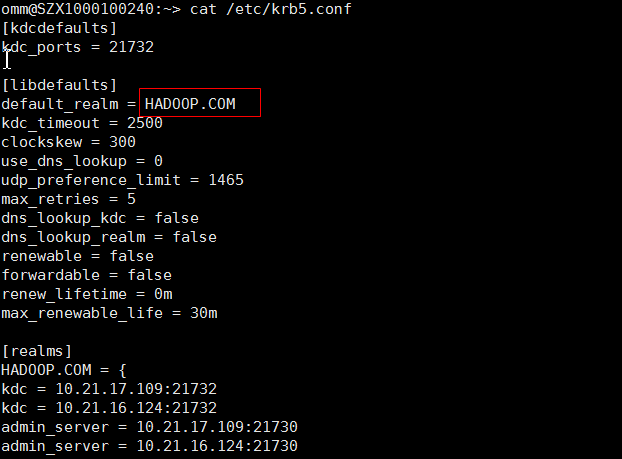
sparkresouce的启动日志

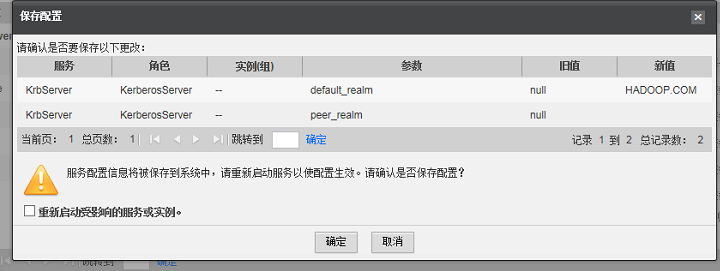


根因：

kerberos的/opt目录曾经被暂满，导致配置文件中域被清空。

解决方法：

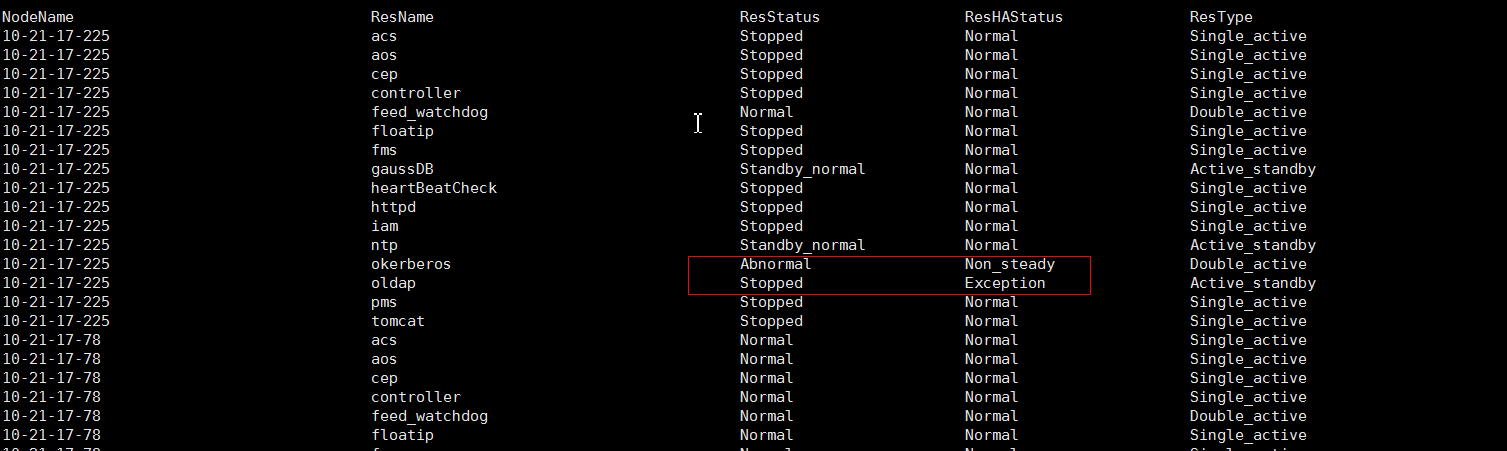
* + - 1. 登录kerberos节点，查看真实的域名
      2. 修改配置。



* + - 1. 整个集群同步配置， 停止集群，启动集群。

## oms的oldap无法启动

管理节点磁盘/opt曾经爆满，导致这2个服务启动异常



根因：

备机不断在同步主机日志，也在刷新备机的环境变量。当磁盘不够时，数据文件和 环境变量均被清空。 导致无法启动了。

解决方法：

1. 停止主节点 oms

/opt/huawei/Bigdata/om-0.0.1/sbin/stop-oms.sh

停止备节点oms

/opt/huawei/Bigdata/om-0.0.1/sbin/stop-oms.sh

1. 删除备节点数据

rm /opt/huawei/Bigdata/om-0.0.1/ldapserver/ldapserver/local/data/\*

复制主节点数据到备节点

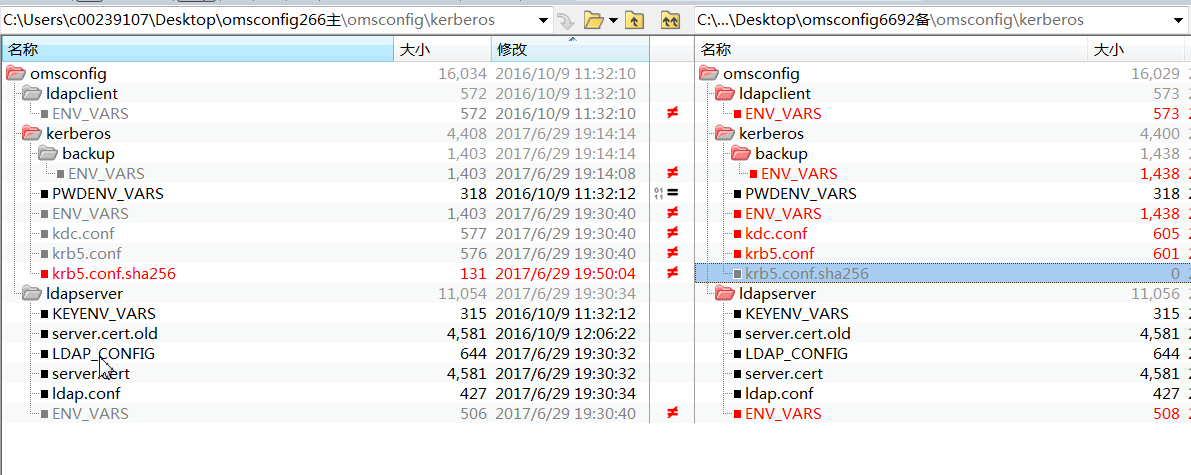
/opt/huawei/Bigdata/om-0.0.1/ldapserver/ldapserver/local/data> scp \* [omm@10.21.17.225:/opt/huawei/Bigdata/om-0.0.1/ldapserver/ldapserver/local/data](mailto:omm@10.21.17.225:/opt/huawei/Bigdata/om-0.0.1/ldapserver/ldapserver/local/data)

1. 对比主备节点环境变量

/opt/huawei/Bigdata/om-0.0.1/etc/om/omsconfig下的所有文件

备节点文件大小为0的都不对。对于为0的文件，将主的复制到备机，并改变IP的顺序，以本机优先

对于 krb5.conf.sha256文件，替换完备机文件后，使用命令 /opt/huawei/Bigdata/om-0.0.1/etc/om/omsconfig/kerberos> sha256sum krb5.conf > krb5.conf.sha256 更新



启动主备节点的oms

## FI web界面添加用户失败



