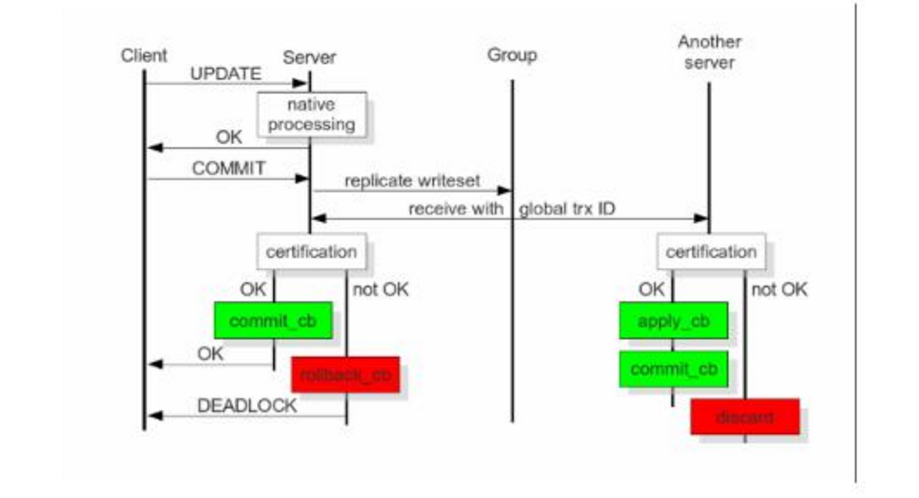
# PXC集群的只读库切换主库(Master:A → Master:B)步骤

说明：在我们的日常生产过程中，有事需要将slave数据库的源数据库切换到PXC集群当中的另一个node节点上。通常认为直接在 slave 数据库上执行change master 命令，这种操作是不可取的。此操作有可能会导致读库的数据与源数据库产生遗漏或者重复的情况。

根据 Galera Cluster Replication 的原理图可以看出



当事务向node1 进行操作请求时，node1将ws广播到整个集群内，集群内每个node分别进行校验，校验成功后分别异步执行。这个时候就有可能 node2执行完了 但是node3还没有开始执行。

但是在整个集群内 Xid（事务id）是一定一致的 。

**当我们要对master 服务器进行切换的时候，我们只要根据slave数据库当前的GTID找到目前的node节点上GTID相对应的Xid，再通过xid找到新node节点上所对应的NEXT\_GTID，即可实现切换master数据库的需求。**

环境描述：

PXC集群:

PXC-node1:192.168.1.30

PXC-node2:192.168.1.31

PXC-node3:192.168.1.20(M)

只读库服务器:

SLAVE\_server：192.168.1.40(S)

需求:目前 只读库 192.168.1.40的master数据库是192.168.1.20,我们要将master数据库改为192.168.1.31

接下来详细说明下整个手动操作的流程。

停止slave数据库的复制

root@localhost : fuscent 16:09:23>stop slave;

记录停止复制的时间  
停止复制的时间为 "2016-11-02 16:09:23"

确认当前GTID中所对应的uuid和目前的偏移量

 show slave status for channel 'db20';

查得结果:

Retrieved\_Gtid\_Set: [b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113322-656146407](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113322-656146407)

Executed\_Gtid\_Set:2289e9db-9e90-11e3-8cd2-2c44fd9276b8:1-62833249,  
51b9a16e-[ca1a-11e4-9cea-40a8f029fabc:1-4](http://ca1a-11e4-9cea-40a8f029fabc:1-4),  
[abb7be49-6564-11e4-8fb2-a0d3c1f2f948:1-288](http://abb7be49-6564-11e4-8fb2-a0d3c1f2f948:1-288),  
[b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:1-656113235](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:1-656113235),  
[c0a13993-9e92-11e3-8ce3-2c44fd929938:1-554331](http://c0a13993-9e92-11e3-8ce3-2c44fd929938:1-554331),  
[c4e47fae-313c-11e6-872a-c4346bb85e64:1-13](http://c4e47fae-313c-11e6-872a-c4346bb85e64:1-13),  
[db43a5d7-9a9d-ee1b-79ad-3d97713cb73b:1-777559907](http://db43a5d7-9a9d-ee1b-79ad-3d97713cb73b:1-777559907)  
根据 Retrieved\_Gtid\_Set 属性部分可以确认 DB20渠道 PXC集群内统一变更的UUID为  [b8e20a8d-35e5-ee1b-7941-b51bcf7abca8](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113322-656146407)

再在 Executed\_Gtid\_Set 属性部分 相对应的UUID和便宜量分别为  [b8e20a8d-35e5-ee1b-7941-b51bcf7abca8](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113322-656146407) 和 [656113235](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:1-656113235)

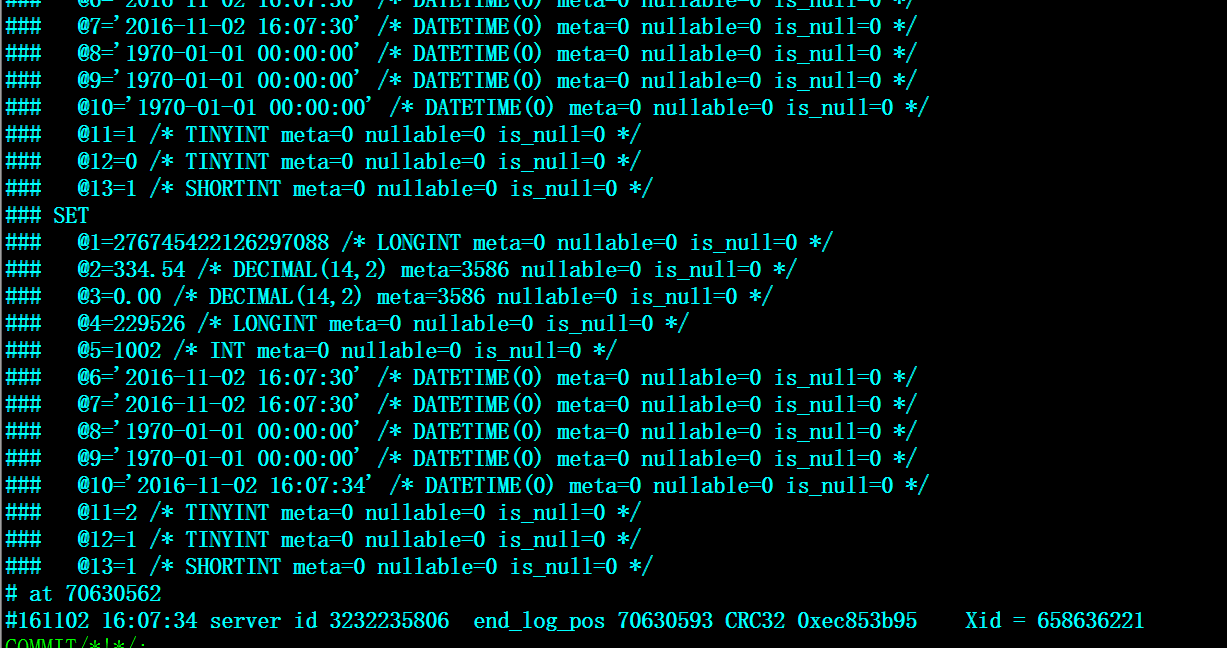
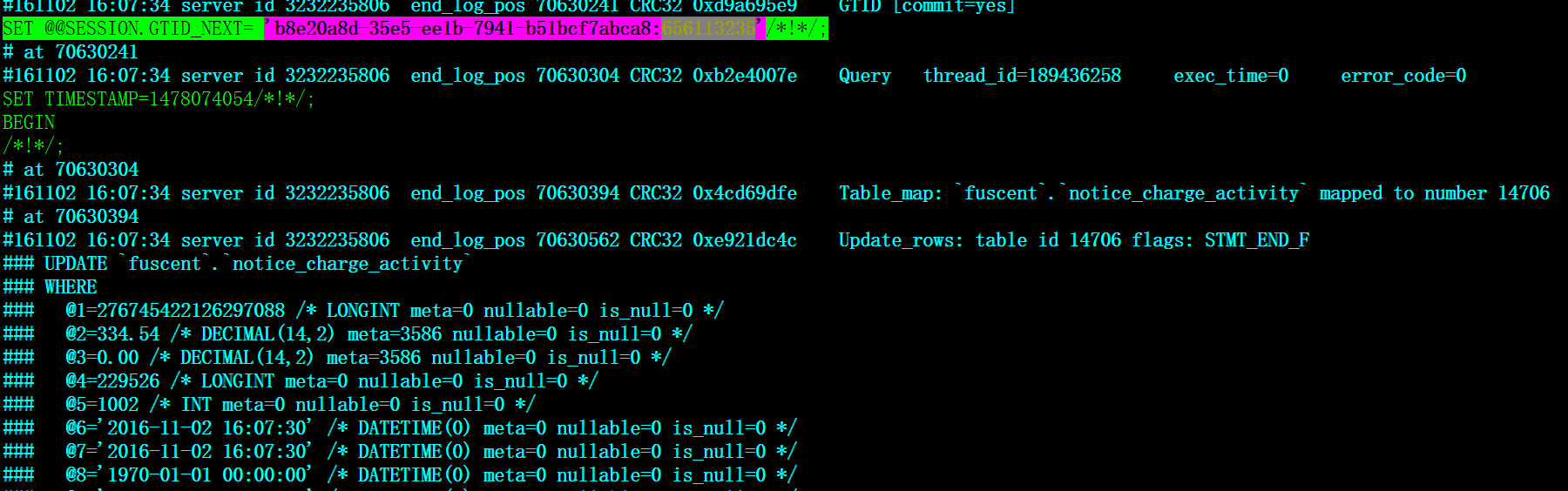
分别导出原master服务器和新master服务器 时间段内的binlog  
 node3:1.20

mysqlbinlog --no-defaults --base64-output=decode-rows -v -v --start-datetime "2016-11-02 16:00:00" --stop-datetime "2016-11-02 16:09:00" /data/mysqldata/mysql-bin.003801 >/tmp/binlog.log

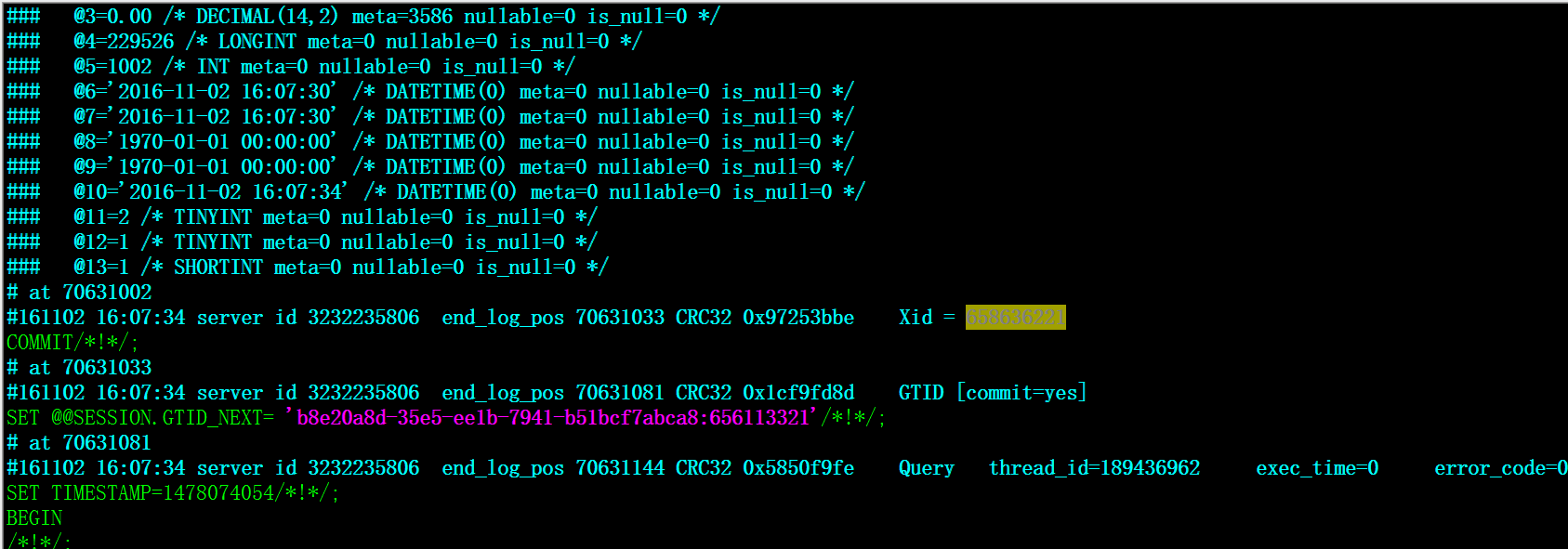
node2:1.31

mysqlbinlog --no-defaults --base64-output=decode-rows -v -v --start-datetime "2016-11-02 16:00:00" --stop-datetime "2016-11-02 16:09:00" /data/emc5600/mysqldata/mysql-bin.003793 >/tmp/binlog.log

1. 在原master服务器的binlog日志内查找GTID对应的XID  
   在1.20服务器上打开导出的 binlog 并查找GTID偏移量 [656113235](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:1-656113235)  
     
     
   我们需要找到 SET @@SESSION.GTID\_NEXT= '[b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113235'/\*!\*/](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113235/); 之后所执行的SQL语句所对应的Xid   
   从图上我们可以看到 之后的SQL语句动作是 UPDATE `fuscent`.`notice\_charge\_activity` 其所对应的 Xid = 658636221 我们将此Xid记录下来

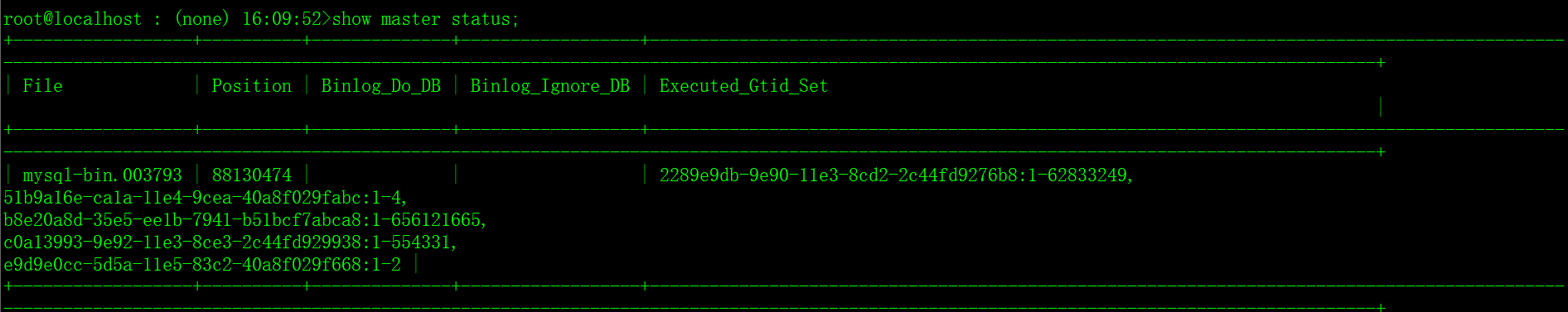


1. 在新master服务器的binlog日志内查找XID对应的NEXT\_GTID  
   在1.31服务器打开导出的binlog 并查找Xid 658636221  
     
   通过查询我们找到 1.31服务器上binlog内对应的Xid 其所执行的SQL语句和1.20服务器是完全一致的 。那么现在我们只需要找到此Xid之后的 SET @@SESSION.GTID\_NEXT 语句所对应的UUID及其偏移量就可以完成了  
   根据上图我们看到 SET @@SESSION.GTID\_NEXT= '[b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113321'/\*!\*/](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113321); 那么 [656113321](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113321/) 就是新的偏移量  
   注意：此处应该为Xid = 658636221所对应的当前的GTID，而非SET @@SESSION.GTID\_NEXT= '[b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113321'/\*!\*/](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:656113321/);这样设置的话，会丢失掉一个事务。



重新修改SLAVE数据库的GITD  
有了新的偏移量我们还需要补全GTID内缺少的 UUID 在1.31服务器上执行

show master status;



比对slave数据库的GTID

**2289e9db-9e90-11e3-8cd2-2c44fd9276b8:1-62833249,**  
**51b9a16e-**[ca1a-11e4-9cea-40a8f029fabc:1-4](http://ca1a-11e4-9cea-40a8f029fabc:1-4)**,**  
[abb7be49-6564-11e4-8fb2-a0d3c1f2f948:1-288](http://abb7be49-6564-11e4-8fb2-a0d3c1f2f948:1-288)**,**  
[b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:1-656113235](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:1-656113235)**,**  
[c0a13993-9e92-11e3-8ce3-2c44fd929938:1-554331](http://c0a13993-9e92-11e3-8ce3-2c44fd929938:1-554331)**,**  
[c4e47fae-313c-11e6-872a-c4346bb85e64:1-13](http://c4e47fae-313c-11e6-872a-c4346bb85e64:1-13)**,**  
[db43a5d7-9a9d-ee1b-79ad-3d97713cb73b:1-777559907](http://db43a5d7-9a9d-ee1b-79ad-3d97713cb73b:1-777559907)

发现缺少 一个UUID和其偏移量 [e9d9e0cc-5d5a-11e5-83c2-40a8f029f668:1-2](http://e9d9e0cc-5d5a-11e5-83c2-40a8f029f668:1-2)

我们将此部分补充到原有的 GTID内并将新的1.31服务器上的GTID偏移量做相应的调整及得到了最新的GTID

2289e9db-9e90-11e3-8cd2-2c44fd9276b8:1-62833249,  
51b9a16e-[ca1a-11e4-9cea-40a8f029fabc:1-4](http://ca1a-11e4-9cea-40a8f029fabc:1-4),  
[abb7be49-6564-11e4-8fb2-a0d3c1f2f948:1-288](http://abb7be49-6564-11e4-8fb2-a0d3c1f2f948:1-288),  
[b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:1-656113321](http://b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:1-656113321),        红色的seqno需修改  
[c0a13993-9e92-11e3-8ce3-2c44fd929938:1-554331](http://c0a13993-9e92-11e3-8ce3-2c44fd929938:1-554331),  
[c4e47fae-313c-11e6-872a-c4346bb85e64:1-13](http://c4e47fae-313c-11e6-872a-c4346bb85e64:1-13),  
[db43a5d7-9a9d-ee1b-79ad-3d97713cb73b:1-777559907](http://db43a5d7-9a9d-ee1b-79ad-3d97713cb73b:1-777559907),  
[e9d9e0cc-5d5a-11e5-83c2-40a8f029f668:1-2](http://e9d9e0cc-5d5a-11e5-83c2-40a8f029f668:1-2)

启动复制

rester master;

SET GLOBAL gtid\_purged="2289e9db-9e90-11e3-8cd2-2c44fd9276b8:1-62833249,

51b9a16e-ca1a-11e4-9cea-40a8f029fabc:1-4,

abb7be49-6564-11e4-8fb2-a0d3c1f2f948:1-288,

b8e20a8d-35e5-ee1b-7941-b51bcf7abca8:1-656113321,

c0a13993-9e92-11e3-8ce3-2c44fd929938:1-554331,

c4e47fae-313c-11e6-872a-c4346bb85e64:1-13,

db43a5d7-9a9d-ee1b-79ad-3d97713cb73b:1-777559907,

e9d9e0cc-5d5a-11e5-83c2-40a8f029f668:1-2";

change master to master\_host="192.168.1.31",master\_user="repl",master\_password="xxxxx",MASTER\_PORT=3306,master\_auto\_position=1 for channel "db20";

start slave;

1. 检查复制状态  
     
   show slave status 后 Master\_host已经更新为192.168.1.31 且 是两个yes 至此 整个切换成功

