

第五下课 Greenplum常见问题的 分析与排查

线上直播 | II月28日 | 14:00 - 15:00

目录



第一节 排查问题的方法

第二节 数据库启停中的问题分析

第三节 数据库状态不正常分析及处理

第四节 扩容问题分析及处理

第五节 运行中几类常见问题分析

排查问题的方法





PID file "/data/master/gpseg-1/postmaster.pid" does not exist

Unexpected internal error

server closed the connection unexpectedly
This probably means the server terminated abnormally
before or while processing the request.



● 管理工具日志

- ➤ 数据库管理工具: gpinitsystem、gpstate、gpstop、gpstart、gprecoverseg、gpcheckcat 等
- ➤ GP集群内所有服务器上: /<User_home_directory>/gpAdminLogs/

● GP数据库日志

- ➤ Master实例日志: \$MASTER_DATA_DIRECTORY/pg_log
- ➤ Segment实例日志: /<segment_datadir>/pg_log
- ➤ CSV文件
- > startup.log

管理工具的日志



- 日志文件是*.log,每天工具第一次被执行就会新生成一个log文件
- 工具在屏幕上显示的信息会全部保存在日志文件中
- 有部分工具日志文件中信息量会比屏幕上丰富,如: gpinitsystem、gpcheckcat、gpbackup等
- 如需要更详细日志,可以使用verbose参数,如:gprecoverseg -v 、gpcheckcat -v -A 、 gpstate -e -v
- 工具内部调用的子命令日志
 - gpstate: 在master上gpstate_YYYYMMDD.log, 在Segment上 gpgetstatususingtransition.py_<hostname>:<username>_YYYYMMDD.log
 - gpstart: 在master上gpstart_YYYYMMDD.log, 在Segment上gpsegstart.py_<hostname>:<username>_YYYYMMDD.log
 - ➤ gpstop: 在master上gpstop_YYYYMMDD.log, 在Segment上 gpsegstop.py_<hostname>:<username>_YYYYMMDD.log

GP数据库日志



startup.log

- ➢ 实例进程启动时写入的日志,一般每次启动写入几行。日志中的一些Warning可以忽略
- → 如果实例启动即失败,可以先从startup.log日志中查找原因

2020-11-06 20:54:04.628382 CST,,,p7929,th-1048955104,,,,0,,,seg-1,,,,,"FATAL","XX000","could not create shared memory segment: Invalid argument (pg_shmem.c:183)","Failed system call was shmget(key=50001001, size=1029073696, 03600).","This error usually means that PostgreSQL's request for a segment exceeded your kernel's SHMMAX parameter. You can either reduce the request size or reconfigure the kernel with larger SHMMAX. To reduce (currently 1029073696 bytes), reduce PostgreSQL's shared_buffers parameter (currently 8192) and/or its max_connections parameter (currently 4000). If the request size is already small, it's possible that it is less than your kernel's SHMMIN parameter, in which case raising the request size or SHMMIN is called for.

● CSV日志

- ➤ 实例进程启动后,操作记录都会写入CSV日志。
- ➤ log_rotation_age、log_rotation_size可以调整日志文件切换的策略
- ➤ log_min_duration_statement设置打印SQL运行时间
- ➤ log_statement可以设置输出DDL、DML或者ALL日志
- ➤ log_connections、 log_disconnections可以打开数据库连接和断链的日志
- ➤ log_min_messages,可设置不同的日志级别



```
1,,dx5277475,x789280442,sx1,"LOG","00000","statement: CREATE TEMP TABLE TEMP Transfer Amt Cnt
  Cust_Id
                                     VARCHAR (50)
                                                          NOT NULL
 ,Sign_Corp_Name
                                     VARCHAR (200)
                                                          NOT NULL
                                                                                     客户端IP地址
                                                          not NULL
 ,Agent_Corp_Distr_Cnt
                                     DECIMAL(18,2)
with(appendonly=true.orientation=row, compresstype=zlib.compresslevel=5)
distributed by (Cust_Id)
;",,,,,"CREATE TEMP TABLE TEMP_Transfer_Amt_Cnt
  Cust_Id
                                     VARCHAR (50)
                                                          NOT NULL
 ,Sign Corp Name
                                     VARCHAR (200)
                                                          NOT NULL
 ,Agent Corp Distr Cnt
                                     DECIMAL(18,2)
                                                          not NULL
with(appendonly=true,orientation=row, compresstype=zlib,compresslevel=5)
distributed by (Cust_Id)
;",0,,"postgres.c",1619,
2019-10-06 12:02:50.122843 CST, "gpusername", "dbname", p30435, th511477536, "10.xxx.xxx.x", "37032", 2019-10-06 12:02:31 CST, 789280321, con1932938, cmd10, seg-
1,,dx5277395,x789280321,sx1,"ERROR","42P16","new partition overlaps existing partition ""p201804""",,,,,"SQL statement ""alter table
dwmart_crm.perf_acct_relasp_results split default partition start('20180501') end('20180601') into (partition p201805, partition extra);""
PL/pgSQL function ""table_add_partition"" line 93 at execute statement"," select
table add partition('dwmart crm','perf acct relasp results',('20191005'::DATE+interval '3 day')::text);",0,,"cdbpartition.c",6560,
2019-10-06 12:03:19.249098 CST, "gpusername", "dbname", p16543, th511477536, "10.xxx.xxx.x", "36791", 2019-10-06 12:02:10 CST, 789280194, con1932907, cmd52, seg-
1,,dx5277332,x789280194,sx1,"ERROR","58M01","Error on receive from seq19 slice1 seqhostXX:40003 pid=31053: server closed the connection unexpectedly","
 This probably means the server terminated abnormally
 before or while processing the request.",,,,"SQL statement ""select count(*)::float4 from (select Ta.arrange_repay_day from pg_temp_1932907.pg_analyze_
 as Ta group by Ta.arrange_repay_day) as Tb""", "analyze dwmart_freeq.rtl_fm_multi_improve_detail;",0,,"cdbdisp.c",1322,
2020-11-04 02:56:56.948450 CST,"gpusername","dbname",p60607,th625108864,"11.xx.xx.xxx","61620",2020-11-04 02:33:59 CST,65897011,con1380376,cmd4,seg-
1,,dx1915760,x65897011,sx1,"ERROR","53000","workfile per segment size limit exceeded (seg179 slice4 seghostXX:40005 pid=37528)",,,,,"INSERT INTO
EXTW 20201104023349 wn19etlsvap0022 6891
```

2019-10-06 12:02:50.023775 CST,"gpusername","bigdata",p23638,th511477536,"10.xxx.xxx.x", 53519",2019-10-06 11:48:13 CST,789280442,con1931767,cmd772,seg-



```
2019-10-06 12:02:50.023775 CST,"gpusername","bigdata",p23638,th511477536,"10.xxx.xxx.x","53519",2019-10-06 11:48:13 CST,789280442,con1931767,cmd772,seg-
1,,dx5277475,x789280442,sx1,"LOG","00000","statement: CREATE TEMP TABLE TEMP Transfer Amt Cnt
  Cust_Id
                                     VARCHAR (50)
                                                          NOT NULL
 ,Sign_Corp_Name
                                     VARCHAR (200)
                                                          NOT NULL
                                                                                                                      会话ID
                                     DECIMAL(18,2)
                                                          not NULL
 ,Agent_Corp_Distr_Cnt
with(appendonly=true.orientation=row, compresstype=zlib.compresslevel=5)
distributed by (Cust Id)
;",,,,,"CREATE TEMP TABLE TEMP_Transfer_Amt_Cnt
  Cust_Id
                                     VARCHAR (50)
                                                          NOT NULL
 ,Sign Corp Name
                                     VARCHAR (200)
                                                          NOT NULL
 ,Agent Corp Distr Cnt
                                    DECIMAL(18,2)
                                                          not NULL
with(appendonly=true,orientation=row, compresstype=zlib,compresslevel=5)
distributed by (Cust_Id)
;",0,,"postgres.c",1619,
2019-10-06 12:02:50.122843 CST,"gpusername","dbname",p30435,th511477536,"10.xxx.xxx.x","37032",2019-10-06 12:02:31 CST,789280321 con1932938,cmd10,seg-
1,,dx5277395,x789280321,sx1,"ERROR","42P16","new partition overlaps existing partition ""p201804""",,,,,"SQL statement ""alter table
dwmart_crm.perf_acct_relasp_results split default partition start('20180501') end('20180601') into (partition p201805, partition extra);""
PL/pgSQL function ""table_add_partition"" line 93 at execute statement"," select
table add partition('dwmart crm','perf acct relasp results',('20191005'::DATE+interval '3 day')::text);",0,,"cdbpartition.c",6560,
2019-10-06 12:03:19.249098 CST,"gpusername","dbname",p16543,th511477536,"10.xxx.xxx.x","36791",2019-10-06 12:02:10 CST,789280194 con1932907,cmd52,seg-
1,,dx5277332,x789280194,sx1,"ERROR","58M01","Error on receive from seg19 slice1 seghostXX:40003 pid=31053: server closed the connection unexpectedly","
 This probably means the server terminated abnormally
 before or while processing the request.",,,,"SQL statement ""select count(*)::float4 from (select Ta.arrange_repay_day from pg_temp_1932907.pg_analyze_
 as Ta group by Ta.arrange_repay_day) as Tb""", "analyze dwmart_freeq.rtl_fm_multi_improve_detail;",0,,"cdbdisp.c",1322,
2020-11-04 02:56:56.948450 CST,"gpusername","dbname",p60607,th625108864,"11.xx.xx.xxx","61620",2020-11-04 02:33:59 CST,65897011 con1380376,cmd4,seg-
1,,dx1915760,x65897011,sx1,"ERROR","53000","workfile per segment size limit exceeded (seg179 slice4 seghostXX:40005 pid=37528).,,,,,"INSERT INTO
EXTW 20201104023349 wn19etlsvap0022 6891
```



```
1,,dx5277475,x789280442,sx1,"LOG","00000","statement: CKLSIE TEMP TABLE TEMP_Transfer_Amt_Cnt
  Cust_Id
                                     VARCHAR (50)
 ,Sign_Corp_Name
                                     VARCHAR (200)
 ,Agent_Corp_Distr_Cnt
                                     DECIMAL(18,2)
with(appendonly=true,orientation=row, compresstype=zlib,compresslevel=5)
distributed by (Cust_Id)
;",,,,,"CREATE TEMP TABLE TEMP_Transfer_Amt_Cnt
  Cust_Id
                                     VARCHAR (50)
                                                          NOT NULL
 ,Sign Corp Name
                                     VARCHAR (200)
                                                          NOT NULL
 ,Agent Corp Distr Cnt
                                     DECIMAL(18,2)
                                                          not NULL
with(appendonly=true,orientation=row, compresstype=zlib,compresslevel=5)
distributed by (Cust_Id)
;",0,,"postgres.c",1619,
2019-10-06 12:02:50.122843 CST,"gpusername","dbname",p30435,th511477536,"10.xxx.xxx.x","37032",2019-10-06 12:02:31 CST,789280321,con1932938,cmd10,seg-
1,,dx5277395,x789280321,sx1,"ERROR","42P16","new partition overlaps existing partition ""p201804""",,,,,"SQL statement ""alter table
dwmart_crm.perf_acct_relasp_results split default partition start('20180501') end('20180601') into (partition p201805, partition extra);""
PL/pgSQL function ""table_add_partition"" line 93 at execute statement"," select
table add partition('dwmart crm','perf acct relasp results',('20191005'::DATE+interval '3 day')::text);",0,,"cdbpartition.c",6560,
2019-10-06 12:03:19.249098 CST,"gpusername","dbname",p16543,th511477536,"10.xxx.xxx.x","36791",2019-10-06 12:02:10 CST,789280194,con1932907,cmd52,seg-
1,,dx5277332,x789280194,sx1,"ERROR","58M01","Error on receive from seq19 slice1 seqhostXX:40003 pid=31053: server closed the connection unexpectedly","
 This probably means the server terminated abnormally
 before or while processing the request.",,,,"SQL statement ""select count(*)::float4 from (select Ta.arrange_repay_day from pg_temp_1932907.pg_analyze_
 as Ta group by Ta.arrange_repay_day) as Tb""", "analyze dwmart_freeq.rtl_fm_multi_improve_detail;",0,,"cdbdisp.c",1322,
2020-11-04 02:56:56.948450 CST,"gpusername","dbname",p60607,th625108864,"11.xx.xx.xxx","61620",2020-11-04 02:33:59 CST,65897011,con1380376,cmd4,seg-
1,,dx1915760,x65897011,sx1,"ERROR","53000","workfile per segment size limit exceeded (seg179 slice4 seghostXX:40005 pid=37528)",,,,,"INSERT INTO
EXTW 20201104023349 wn19etlsvap0022 6891
```

2019-10-06 12:02:50.023775 CST,"gpusername","bigdata",p23638,th511477536,"10.xxx.xxx.x","53519",2019-10-06 11:48:13 CST,789280442,con1931767,cmd772,seg-



EXTW 20201104023349 wn19etlsvap0022 6891

```
2019-10-06 12:02:50.023775 CST,"gpusername","bigdata",p23638,th511477536,"10.xxx.xxx.x","53519",2019-10-06 11:48:13 CST,789280442,con1931767,cmd772,seg-
1,,dx5277475,x789280442,sx1,"LOG","00000","statement: CREATE TEMP TABLE TEMP Transfer Amt Cnt
  Cust_Id
                                     VARCHAR (50)
                                                          NOT NULL
 ,Sign_Corp_Name
                                     VARCHAR (200)
                                                          NOT NULL
                                                          not NULL
 ,Agent_Corp_Distr_Cnt
                                     DECIMAL(18,2)
with(appendonly=true.orientation=row, compresstype=zlib.compresslevel=5)
distributed by (Cust_Id)
;",,,,,"CREATE TEMP TABLE TEMP_Transfer_Amt_Cnt
  Cust_Id
                                     VARCHAR (50)
                                                          NOT NULL
 ,Sign Corp Name
                                     VARCHAR (200)
                                                          NOT NULL
 ,Agent Corp Distr Cnt
                                     DECIMAL(18,2)
                                                          not NULL
with(appendonly=true,orientation=row, compresstype=zlib,compresslevel=5)
                                                                                               报错信息
distributed by (Cust_Id)
;",0,,"postgres.c",1619,
2019-10-06 12:02:50.122843 CST, "gpusername", "dbname".p30435.th511477536."10.xxx.xxx.x"."37 32".2019-10-06 12:02:31 CST, 789280321, con1932938, cmd10, seg-
1,,dx5277395,x789280321,sx1,"ERROR","42P16","new partition overlaps existing partition ""p201804" ",,,,,"SQL statement ""alter table
dwmart_crm.perf_acct_relasp_results split default partition start('20180501') end('20180601') into (partition p201805, partition extra);""
PL/pgSQL function ""table_add_partition"" line 93 at execute statement"," select
table add partition('dwmart crm','perf acct relasp results',('20191005'::DATE+interval '3 day')::text);",0,,"cdbpartition.c",6560,
2019-10-06 12:03:19.249098 CST, "gpusername", "dbname", p16543, th511477536, "10.xxx.xxx.x", "36791", 2019-10-06 12:02:10 CST.789280194.con1932907.cmd52.seg-
1,,dx5277332,x789280194,sx1,"ERROR","58M01","Error on receive from seg19 slice1 seghostXX:40003 pid=31053: server closed the connection unexpectedly ,"
 This probably means the server terminated abnormally
 before or while processing the request.",,,,"SQL statement ""select count(*)::float4 from (select Ta.arrange_repay_day from pg_temp_1932907.pg_analyze_
 as Ta group by Ta.arrange_repay_day) as Tb""", "analyze dwmart_freeq.rtl_fm_multi_improve_detail;",0,,"cdbdisp.c",1322,
2020-11-04 02:56:56.948450 CST, "gpusername", "dbname", p60607, th625108864, "11.xx.xx.xxx", "61620", 2020-11-04 02:33:59 CST, 65897011, con1380376, cmd4, seq-
```

1,,dx1915760,x65897011,sx1,"ERROR","53000", workfile per segment size limit exceeded (seg179 slice4 seghostXX:40005 pid=37528)",,,,,"INSERT INTO



EXTW 20201104023349 wn19etlsvap0022 6891

```
2019-10-06 12:02:50.023775 CST,"gpusername","bigdata",p23638,th511477536,"10.xxx.xxx.x","53519",2019-10-06 11:48:13 CST,789280442,con1931767,cmd772,seg-
1,,dx5277475,x789280442,sx1,"LOG","00000","statement: CREATE TEMP TABLE TEMP Transfer Amt Cnt
  Cust_Id
                                     VARCHAR (50)
                                                           NOT NULL
 ,Sign_Corp_Name
                                     VARCHAR (200)
                                                           NOT NULL
                                                           not NULL
 ,Agent_Corp_Distr_Cnt
                                     DECIMAL(18,2)
with(appendonly=true.orientation=row, compresstype=zlib.compresslevel=5)
distributed by (Cust_Id)
;",,,,,"CREATE TEMP TABLE TEMP_Transfer_Amt_Cnt
  Cust_Id
                                     VARCHAR (50)
                                                           NOT NULL
 ,Sign Corp Name
                                     VARCHAR (200)
                                                           NOT NULL
 ,Agent Corp Distr Cnt
                                     DECIMAL(18,2)
                                                           not NULL
with(appendonly=true,orientation=row, compresstype=zlib,compresslevel=5)
distributed by (Cust_Id)
;",0,,"postgres.c",1619,
2019-10-06 12:02:50.122843 CST, "gpusername", "dbname", p30435, th511477536, "10.xxx.xxx.x", "37032", 2019-10-06 12:02:31 CST, 789280321, con1932938, cmd10, seg-
1,,dx5277395,x789280321,sx1,"ERROR","42P16","new partition overlaps existing partition ""p201984"""
dwmart_crm.perf_acct_relasp_results split default partition start('20180501') end('20180601'
PL/pgSQL function ""table_add_partition"" line 93 at execute statement"," select
table add partition('dwmart crm','perf acct relasp results',('20191005'::DATE+interval '3 da
2019-10-06 12:03:19.249098 CST, "gpusername", "dbname".p16543.th511477536."10.xxx.xxx.x"."36791".2019 20-06 12:02:10 CST, 789280194, con1932907, cmd52, seg-
1,,dx5277332,x789280194,sx1,"ERROR","58M01","Error on receive from seg19 slice1 seghostXX:40003 pid=31053: server closed the connection unexpectedly","
 This probably means the server terminated abnormally
 before or while processing the request.",,,,"SQL statement ""select count(*)::float4 from (select Ta.arrange_repay_day from pg_temp_1932907.pg_analyze
 as Ta group by Ta.arrange_repay_day) as Tb""", "analyze dwmart_freeq.rtl_fm_multi_improve_detail;",0,,"cdbdisp.c",1322,
2020-11-04 02:56:56.948450 CST, "gpusername", "dbname", p60607, th625108864, "11.xx.xx.xxx". "61620". 2020-11-04 02:33:59 CST.65897011.con1380376, cmd4, seg-
```

1,,dx1915760,x65897011,sx1,"ERROR","53000","workfile per segment size limit exceeded (seg179 slice4 seghostXX:40005 pid=37528)",,,,,"INSERT INTO

CSV日志



- CSV日志详细的格式说明参考管理员手册
- CSV日志可通过外部表的方式加载到GP数据库中,便于进一步分析
- 外部表的定义可参考gp_toolkit.__gp_log_master_ext

当遇到一个问题时



- 找对日志
- 凭过往的经验,是否已知问题
- 寻找规律
 - ➤ 经常出现还是偶然出现?
 - ➤ 是否集中在某个时间段?
 - ➤ 是否发生在固定的机器?
 - ➤ 是否发生在固定的实例?
 - ➤ 报错时是否总会有某SQL在运行?
- 可能需要测试和场景复现

数据库启停中的问题分析



gpstop有哪几步操作



- 1. 从master中获取所有Segment的地址,端口,实例状态等信息
- 2. 停止Master实例
- 3. 停止Standby master实例
- 4. 并发停止primary实例
- 5. 并发停止mirror实例
- 6. 清理共享内存

【建议】:gpstop执行完后,使用gpssh连接到所有服务器检查postgres进程数是否为0,如果进程数不为0,要具体分析排查

too many clients already



- 由于连接数太多,gpadmin也无法连接数据库获取segment实例的信息。此时,使用gpstate也会遇到同样的报错。统计master上的postgres进程,可以评估当前客户端连接数
- 首先考虑清理客户端连接,可以让gpadmin连上数据库,就可以操作gpstop
 - ➤ 如果还有可用的客户端使用gpadmin连着数据库,可以查看pg_stat_activity,通过pg_terminate_backend杀会话释放 部分可用连接数。
 - ➤ 查看是否有很多starting up的postgres进程:ps -ef lgrep postgres lgrep start。可以用kill杀这类进程(不能用kill -9)。确认是否能够释放部分可用连接
- 如果无法释放可用连接,使用pg_ctl停止master实例
 - pg_ctl stop -W -m fast -D \$MASTER_DATA_DIRECTORY
 - ➤ 如果-m fast无法把进程停下来,就只能使用-m immediate
 - ➤ 然后gpstart -m, 然后再次操作gpstop -fa
- 参数superuser_reserved_connections
 - ➤ 默认值3,建议调大。保证gpadmin有足够的预留连接数

database is in recovery mode / is starting up



- 是由于master实例发生了PANIC,无法自行恢复正常。此时用正常手段无法连接数据库
- 只能使用pg_ctl停止master实例
 - pg_ctl stop -W -m fast -D \$MASTER_DATA_DIRECTORY
 - ➤ 如果-m fast无法把进程停下来,就只能使用-m immediate
 - ➤ 然后gpstart -m,然后再次操作gpstop -fa

gpstart有哪几步操作



- 1. 启动Master实例,获取所有Segment的地址,端口,实例状态等信息
- 2. 停止Master实例
- 3. 启动Segment实例(开始打点),Primary和mirror的进程都启动,并且启动Primary与mirror之间的同步进程。如果停库之间实例状态为down,启动时会忽略。
- 4. 再次启动Master实例
- 5. 启动Standby master实例



- 报错的原因与该文件不存在无关,该文件用于保存实例主进程号,由程序自行管理不需要人工干预。报错是代表实例的主进程刚启动就失败。
- Master和Segment都有可能遇到这个报错
 - ➤ 如果是Master报错通常在gpstart命令输出信息中可看到
 - ➤ 如果是Segment报错可在gpstart -v(verbose)的输出信息中看到,也可以在启动失败的主机的/ <User_home_directory>/gpAdminLogs/目下, gpsegstart.py_<hostname>:<username>_YYYYMMDD.log日志中可以看到类似信息
- 问题定位方法:
 - > gpssh到所有服务器检查进程启动状况。确定哪些实例未启动成功
 - 查看startup.log,通常在startup.log中就可以找到报错原因。



- 导致这类报错的一些原因及解决方法
 - ➤ 信号量设置偏小,可调整/etc/sysctl.conf中的参数kernel.sem,或者调低max_connections
 - ➢ 共享内存设置不够大,可调整OS参数kernel.shmmax,或者调低使用共享内存的一些参数配置
 - ➤ /tmp目录(No space lefo on device),可能是满了(df -h),或者inode用满了(df -i);解决方法就是删文件



• 几种启动卡住的状况

```
→ 在第一步启动master时、长时间不往下走。这种情况就重点分析master的问题。
gpstart:mdw:gpadmin-[INF0]:-Starting gpstart with args: -a
gpstart:mdw:gpadmin-[INF0]:-Gathering information and validating the environment...
gpstart:mdw:gpadmin-[INF0]:-Greenplum Binary Version: 'postgres (Greenplum Database) 4.3.33.2 build 1'
gpstart:mdw:gpadmin-[INF0]:-Greenplum Catalog Version: '201310150'
gpstart:mdw:gpadmin-[INF0]:-Starting Master instance in admin mode
```

- ➢ 启动时一直在打点,很长时间不结束。这通常是因为启动Segment实例时出现了问题。
- ➤ gpstart命令执行到最后,看上去执行完了,但命令一直不结束
- 问题定位方法:
 - 查看startup.log
 - 查看相应实例的CSV日志
 - ➤ gpssh到所有服务器,检查postgres进程数,检查服务器状态是否正常
 - ➤ gpssh到所有服务器,检查是否有starting up进程,如果有重点检查这些实例

启动过程卡住——配置文件有问题



- pg_hba.conf文件有问题
 - ➤ 在pg_hba.conf文件中有格式错误的规则,会导致对应的实例启动时停住。从CSV日志中可以看到报错信息。

- postgresql.conf中有不规范的信息:
- ➤ 在手工修改postgresql.conf时,误操作遗留的错误。系统启动时读取失败。从startup.log中可以看到报错信息。

2020-06-03 07:46:04.567355 GMT,,,p65137,th737126176,,,,0,,,seg-1,,,,,"FATAL","42601" "syntax error in file ""/data/master/gpseg-1/postgresql.c onf"" line 519, near token ""configuration""",,,,,,,"ParseConfigFile","guc-file.l",369,

- ➢ 如果是个别参数名或者参数值写错,启动是正常。检查startup<u>.log可以看到参数错误的报错信息。</u>
- 2020-06-03 07:54:53.840290 GMT,,,p65508,th221579040,,,,0,,,seg-1,,,,,"LOG","22023","parameter ""gp_enable_gpperfmon"" requires a Boolean value ",,,,,,,"set_config_option","guc.c",10055,
 - 处理方法
 - ➤ gpstart会有超时机制,等待一段时间后,会自动报超时退出
 - ➤ 可以把gpstart命令内部调用的pg_ctl命令进程杀掉。可使用kill,不能用kill -9
 - ➢ 修正配置文件后重新启动

启动过程卡住——Segment启动问题排查



- 首先凭经验、判断本次启动是不是比以往启动过程要慢很多。
- 检查各台服务器是否正常
 - ➤ 用gpssh连接所有服务器,看是否可以连通。如果gpssh也会卡住,首要排查那台机器ssh有问题
 - ➤ gpssh连通所有服务器后,先查看各台启动postgres进程数量 gpssh -f allhosts "ps -ef |grep postgres |grep -v grep |wc -l"
 - ➢ 对启动进程数明显较少的服务器,ssh到服务器上检查
 - □ 排查服务器状态,检查文件系统,内存等是否正常,检查文件系统使用率
 - 到/<User_home_directory>/gpAdminLogs/目录下,查看 gpsegstart.py_<hostname>:<username>_YYYYMMDD.log日志
 - □ 检查服务器上各实例(primary和mirror)的端口号,检查是否缺少某个端口号,也就是某个实例未启动
 - 检查服务器上每个实例的进程数量。是否有实例进程数与别的实例不同。
 - □ 进入怀疑有问题的实例目录下,查看startup.log和CSV文件。

启动过程卡住——分析Segment日志



- 通过分析日志,排除上面所提到的一些问题,如:配置文件,系统信号量,/tmp目录空间,/data目录空间
- 服务端口被占用问题
 - 通常primary和mirror实例端口不会被占用。
 - ➤ GP5及以前版本,还有一个端口replication_port(在gp_segment_configuration中可查)。用于建立primary与mirror之间数据同步(file replication)的端口。这个端口有一定几率被随机端口所占用,导致primary与mirror之间无法建立数据同步关系,结果就是启动后有mirror实例down了
 - ➢ 通过调整系统参数net.ipv4.ip_local_port_range和net.ipv4.ip_local_reserved_ports,调整系统中端口号使用范围。规 避随机端口与GP数据库所使用的端口冲突。
- 遇到系统表的相关报错
 - ➤ 部分系统表的问题,可能会导致实例启动失败。通常的方法是通过设置一些参数,让实例能够启动,然后修复系统表,最后做gprecoverseg

"FATAL","XX000","Number of freeTIDs 191, do not match maximum free order number 258, for 'gp_persistent_relation_node'

veixin.qq.com/s/7bnXJ6fHC1Kba-nAanFshQ

➤ 建议:找原厂的工程师协助解决

gpstart error: Do not have enough valid segments to start the array.



mdw	gpadmin-[INFO]:		
mdw	gpadmin-[INFO]:- Successful segment starts	= 30	
mdw	gpadmin-[WARNING]:-Failed segment starts	= 1 <<<<<<	
mdw	gpadmin-[WARNING]:-Skipped segment starts (segments are m	narked down in configuration) = 1	<<<<<<
mdw	gpadmin-[INFO]:		
mdw	gpadmin-[INFO]:-		
mdw	gpadmin-[INFO]:-Successfully started 30 of 31 segment instance	es, skipped 1 other segments <<<<	<<<<
mdw	gpadmin-[INFO]:		
mdw	gpadmin-[WARNING]:-Segment instance startup failures report	ed	
mdw	gpadmin-[WARNING]:-Failed start 1 of 31 segment instances <	<<<<<	
mdw	gpadmin-[WARNING]:-Review /home/gpadmin/gpAdminLogs/	/gpstart_20160831.log	
	gpadmin-[INFO]:		
mdw	gpadmin-[WARNING]:-*********************	**********	******
mdw	gpadmin-[WARNING]:-There are 1 segment(s) marked down in	the database	
mdw	gpadmin-[WARNING]:-To recover from this current state, review	w usage of the gprecoverseg	
mdw	gpadmin-[WARNING]:-management utility which will recover to	failed segment instance databases.	
mdw	gpadmin-[WARNING]:-********************	**********	******
mdw	gpadmin-[INFO]:-Commencing parallel segment instance shutde	own, please wait	
mdw	gpadmin-[ERROR] -gpstart error: Do not have enough valid seg	gments to start the array.	

• 原因分析

- ➤ 通常是在启动前就有 down的实例;启动时对 应的实例也没有启动成功
- ➤ 也可能是刚好primary和 mirror所在的服务器同时 故障

• 问题定位方法:

- ▶ 如果是服务器故障,则只能等到服务器修复好后再启动
- 排除服务器故障,则重点 关注启动失败的实例,查 看日志。逐一排查上述提 到的一些报错情况。

数据库状态不正常分析及处理





● segment状态正常,会显示"All segments are running normally"

- 数据库状态的重要手段之一,可对接客户监控平台调用
- 如果发现无法获取某些Segment状态,可以反复检查2~3次。检查是否由于网络闪断,服务器短暂 无响应导致的不正常。
- 如果遇到gpstate一直打点,不输出结果。可以用gpssh连接所有服务器,检查连通性,定位是否有服务器ssh不通。



● GP5及以前版本

```
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-local Greenplum Version: 'postgres (Greenplum Database) 5.19.0 build commi
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-master Greenplum Version: 'PostgreSQL 8.3.23 (Greenplum Database 5.19.0 bu
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Obtaining Segment details from master...
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Gathering data from segments...
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:----
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Segment Mirroring Status Report
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:---
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Segments with Primary and Mirror Roles Switched
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                                                   Mirror
                                          Current Primary
                                                            Port
                                                                                     Port
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                          ZZD170000DWHDB1
                                                            7100
                                                                   ZZD170000DWHDB8
                                                                                     6100
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                                            7101
                                                                   ZZD170000DWHDB8
                                                                                     6101
                                          ZZD170000DWHDB1
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                          ZZD170000DWHDB2
                                                            7102
                                                                   ZZD170000DWHDB8
                                                                                     6102
                                          ZZD170000DWHDB2
                                                            7103
                                                                   ZZD170000DWHDB8
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                                                                     6103
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:--
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Primaries in Change Tracking
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                                                   Change tracking size
                                          Current Primary
                                                            Port
                                                                                         Mirror
                                                                                                            Port
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                          ZZD170000DWHDB6
                                                            6102
                                                                   60.3 MB
                                                                                          ZZD170000DWHDB8
                                                                                                           7102
                                                                   60.3 MB
qpstate:ZZD170000DWHGL1:qpadmin-[INF0]:-
                                                            6103
                                                                                                           7103
                                          ZZD170000DWHDB6
                                                                                          ZZD170000DWHDB8
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                          ZZD170000DWHDB7
                                                            6100
                                                                   60.2 MB
                                                                                          ZZD170000DWHDB8
                                                                                                           7100
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                                                   60.3 MB
                                                                                                           7101
                                          ZZD170000DWHDB7
                                                            6101
                                                                                          ZZD170000DWHDB8
qpstate:ZZD170000DWHGL1:qpadmin-[INF0]:-
                                          ZZD170000DWHDB1
                                                            7100
                                                                   60.3 MB
                                                                                          ZZD170000DWHDB8
                                                                                                           6100
                                                                   60.3 MB
qpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                                            7101
                                                                                                           6101
                                          ZZD170000DWHDB1
                                                                                          ZZD170000DWHDB8
gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-
                                                            7102
                                                                   60.3 MB
                                                                                          ZZD170000DWHDB8
                                                                                                           6102
                                          ZZD170000DWHDB2
qpstate:ZZD170000DWHGL1:qpadmin-[INF0]:-
                                          ZZD170000DWHDB2
                                                            7103
                                                                   60.3 MB
                                                                                          ZZD170000DWHDB8
                                                                                                           6103
```



● GP5及以前版本

```
20200519:18:58:21:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Starting gpstate with args: -e
20200519:18:58:22:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INFO]:-local Greenplum Version: 'postgres (Greenplum Database) 5.19.0 build
commit:7f5d6a22614522e47fa1020933e0a231a122b00a'
20200519:18:58:22:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-master Greenplum Version: 'PostgreSQL 8.3.23 (Greenplum Database 5.19.0 build
commit:7f5d6a22614522e47fa1020933e0a231a122b00a) on x86_64-pc-linux-gnu, compiled by GCC gcc (GCC) 6.2.0, 64-bit compiled on May 15 2019 16:16:39'
20200519:18:58:23:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Obtaining Segment details from master...
20200519:18:58:23:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Gathering data from segments...
20200519:18:58:50:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:---
20200519:18:58:50:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Segment Mirroring Status Report
20200519:18:58:50:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Segments with Primary and Mirror Roles Switched
20200519:18:58:50:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:- Current Primary
                                                                  Port
                                                                                      Port
                                                                        Mirror
ZZD170000DWHDB1
                                                                  7100
                                                                        ZZD170000DWHDB8
                                                                                      6100
ZZD170000DWHDB1
                                                                  7101
                                                                        ZZD170000DWHDB8
                                                                                      6101
ZZD170000DWHDB2
                                                                  7102
                                                                        ZZD170000DWHDB8
                                                                                      6102
ZZD170000DWHDB2
                                                                  7103
                                                                        ZZD170000DWHDB8
                                                                                      6103
20200519:18:58:50:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:-Segment Pairs in Resynchronization
<u> 20200519:18:58:50:045523 g</u>pstate:ZZD170000DWHGL1:gpadmin-[INFO]:- Current Frimary Fort Resync ساملو
                                                                                   Est. resync progress
                                                                                                    Total resync objects
                                                                                                                      Objects to
resync Data synced Est. total to sync Est. resync end time
                                                   Change tracking size Mirror
                                                                                   Port
ZZD170000DWHDB6
                                                                                  99.23%
                                                                                                    83
                                                                  6102
                                                                      Incremental
                                                                                  7102
     1.47 GB
                1.48 GB
                                2020-05-19 18:58:45
                                                  352 kB
                                                                    ZZD170000DWHDB8
99.22%
                                                                                                    83
                                                                  6103
                                                                        Incremental
                1.48 GB
                                2020-05-19 18:58:48
                                                  352 kB
     1.47 GB
                                                                    ZZD170000DWHDB8
                                                                                  7103
Incremental
                                                                                  99.25%
                                                                                                    83
                1.50 GB
                                2020-05-19 18:58:39
     1.49 GB
                                                  352 kB
                                                                    ZZD170000DWHDB8
                                                                                  7100
99.25%
                                                                                                    83
                                                                  6101 Incremental
                                2020-05-19 18:58:42
     1.47 GB
                1.48 GB
                                                  352 kB
                                                                    ZZD170000DWHDB8
                                                                                  7101
                                                                                                    83
20200519:18:58:50:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:- ZZD170000DWHDB1
                                                                  7100
                                                                      Incremental
                                                                                  99.21%
     1.49 GB
                1.50 GB
                                2020-05-19 18:58:46
                                                  352 kB
                                                                    ZZD170000DWHDB8
                                                                                  6100
20200519:18:58:50:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:- ZZD170000DWHDB1
                                                                  7101 Incremental
                                                                                  99.23%
                                                                                                     83
     1.47 GB
                1.48 GB
                                2020-05-19 18:58:49
                                                  352 kB
                                                                    ZZD170000DWHDB8
                                                                                  6101
20200519:18:58:50:045523 gpstate:ZZD170000DWHGL1:gpadmin-[INF0]:- ZZD170000DWHDB2
                                                                  7102 Incremental
                                                                                  99.23%
                                                                                                    83
     1.47 GB
                1.48 GB
                                2020-05-19 18:58:46
                                                  352 kB
                                                                    ZZD170000DWHDB8
                                                                                  6102
ZZD170000DWHDB2
                                                                                  99.32%
                                                                                                    83
                                                                                                                      0
                                                                        Incremental
             1.48 GB
                                2020-05-19 18:58:49
     1.47 GB
                                                                    ZZD170000DWHDB8
                                                                                 6103
```



GP6

```
gpstate:mdw:gpadmin-[INF0]:-Starting gpstate with args: -e
gpstate:mdw:gpadmin-[INF0]:-local Greenplum Version: 'postgres (Greenplum Database) 6.6.0 build commit:
gpstate:mdw:gpadmin-[INF0]:-master Greenplum Version: 'PostgreSQL 9.4.24 (Greenplum Database 6.6.0 build
i4-unknown-linux-gnu, compiled by gcc (GCC) 6.4.0, 64-bit compiled on Apr 7 2020 02:46:13'
gpstate:mdw:gpadmin-[INF0]:-Obtaining Segment details from master...
gpstate:mdw:gpadmin-[INF0]:-Gathering data from segments...
gpstate:mdw:gpadmin-[WARNING]:-pg stat replication shows no standby connections
gpstate:mdw:gpadmin-[INF0]:-----
qpstate:mdw:qpadmin-[INF0]:-Segment Mirroring Status Report
gpstate:mdw:gpadmin-[INF0]:-Segments with Primary and Mirror Roles Switched
gpstate:mdw:gpadmin-[INF0]:-
                             Current Primary
                                                     Mirror
                                              Port
                                                              Port
gpstate:mdw:gpadmin-[INF0]:-
                             mdw
                                              7001
                                                     mdw
                                                              6001
gpstate:mdw:gpadmin-[INF0]:
gpstate:mdw:gpadmin-[INF0]:-Unsynchronized Segment Pairs
gpstate:mdw:gpadmin-[INF0]:-
                             Current Primary
                                              Port
                                                     Mirror
                                                              Port
gpstate:mdw:gpadmin-[INF0]:-
                             mdw
                                              7001
                                                     mdw
                                                              6001
gpstate:mdw:gpadmin-[INF0];
gpstate:mdw:gpadmin-[INF0]:-Downed Segments (may include segments where status could hot be retrieved)
gpstate:mdw:gpadmin-[INF0]:-
                             Segment
                                       Port
                                             Config status
                                                             Status
qpstate:mdw:gpadmin-[INF0]:-
                                                             Down in configuration
                                       6001
                             mdw
                                              Down
```

发现Segment down后分析和收集日志方法



- 1. 使用gpstate -e确定down状态实例的情况,确定涉及哪些主机和端口
- 2. 从系统表gp_segment_configuration中检查状态为down的实例,记录dbid、content、端口号。
 - select * from gp_segment_configuration order by 2,1;
 - > select * from gp_segment_configuration where status='d' order by 1;
- 3. 从系统表gp_configuration_history中检查实例状态变更的具体时间点
 - select * from gp_configuration_history order by 1 desc;
- 4. 先检查master的日志
 - ➤ 依据具体的实例变更的时间点,查找相应的CSV日志
 - ➤ 通过FTS关键字查找日志,如:grep "FTS:"gpdb-YYYY-MM-DD_000000.csv
 - ➤ 通过FTS的信息可以判断,是否为master的ftsprobe进程主动探测到的。默认情况下FTS探测实例状态异常,会重试5次
 - ➢ 对于连接超时,服务器宕机,网络中断等情况,可以在FTS的日志信息中看到明确的报错信息
 - ➢ primary与mirror之间断连(如:gp_segment_connect_timeout超时,mirror空间满等),primary实例会向FTS进程汇 报。FTS日志中会显示相关信息
- 5. 进一步去检查Segment的日志,依据前两步中定位到的实例ID,登录对应服务器中检查
 - → 一对primary和mirror实例的日志都应该关注
 - ➢ ps -ef lgrep green检查对应实例目录(关注端口号和contentID),进入实例目录查看CSV日志。
 - ▶ 检查出问题时间点前后的日志,查找可疑的,异常的信息。

Segment状态异常常见问题分析



• 文件系统空间用完

- ➤ 通常mirror实例会被置为down,就是mirror实例在写入数据时,遇到空间不足。
- ➤ 在mirror实例的CSV日志中能够看到相关信息"No space left on device"
- → 如果是某个SQL在写入数据时,在primary上遇到空间不足问题。SQL会报错。此时primary实例不会down
- → 如果primary实例在写入XLOG时,碰到空间不足。primary实例也会down。

Segment状态异常常见问题分析



- gp_segment_connect_timeout相关
 - "threshold '75' percent of 'gp_segment_connect_timeout=600' is reached, mirror may not be able to keep up with primary, primary may transition to change tracking"
 - ➤ primary日志中看到上述信息,代表primary会主动与mirror断开连接,mirror会被置为down
 - ➤ 日志中会提示,调大gp_segment_connect_timeout的设置值。可适当调大参数值,但不能一味通过调大该参数 来规避同类问题。必须要排查是什么原因导致的服务器资源消耗过大
 - ➤ 分析报错规律:是否集中在某些或某台服务器上?出现时间点是否有规律?出现时是否有相同的SQL在运行?出 现时系统并发数高不高?
 - ➤ 常见集中原因:
 - □ 因为mirror所在的服务器资源消耗过高,Workload过高,甚至服务器hang住无响应。导致primary与mirror之间的 通讯大量超时。
 - □ 故障期间服务器网络有较长时间大流量,把带宽基本用满,导致网络传输大量超时
 - □ IO性能是否存在问题,故障期间硬盘故障并伴随有大量的IO读写操作,也可能导致mirror响应慢导致超时



查看mirror日志:

```
2020-08-12 04:02:12.523545 CST,,,p15751,th1506146176,,,2000-01-01 08:00:00 CST,0,,,seg-1,,,, "WARNING","XX000","receive close on connection: Success (cdbfilerepconnserver.c:337)",,,,,,,0,,"cdbfilerepconnserver.c",337, 2020-08-12 04:08:43.735834 CST,,,p15751,th1506146176,,,0,,,seg-1,,,,"Log","00000","'set segment state', mirroring role 'mirror role' mirroring state 'sync' segment state 'up and running' process name(pid) 'mirror receiver process(15751)' 'cdbfilerep.c' 'L2480' 'FileR
ep_SetSegmentState'",,,,,,0,,"cdbfilerep.c",1850,"Stack trace:
      0xb0d20e postgres errstart (elog.c:502)
       0xbf9004 postgres <symbol not found> (cdbfilerep.c:1832)
       0xc06c0e postgres FileRepMirror_StartReceiver (cdbfilerepmirror.c:169)
       0xc002ed postgres FileRepSubProcess_Main (cdbfilerepservice.c:815)
      Oxbf6b6a postgres <symbol not found> (cdbfilerep.c:2703)
Oxbfe9d4 postgres FileRep_Main (cdbfilerep.c:3636)
       0x5f5f97 postgres AuxiliaryProcessMain (bootstrap.c:499)
       0x8f0f54 postgres <symbol not found> (postmaster.c:7603)
       0x8f113c postgres StartFilerepProcesses (postmaster.c:1631)
      0x9048af postgres doReguestedPrimaryMirrorModeTransitions (primary_mirror_mode.c:1752)
      0x8fd3af postgres <symbol not found> (postmaster.c:2484)
      0x8fec90 postgres PostmasterMain (postmaster.c:7603)
      0x7feb6f postgres main (main.c:206)
0x7f0455417555 libc.so.6 __libc_start_main + 0xf5
      0x4bee79 postgres <symbol not found> + 0x4bee79
2020-08-12 04:08:43.735948 CST,,,p15765,th1506146176,,,,0,,,seg-1,,,,,"LOG","00000","'set segment state', mirroring role 'mirror role' mirroring state 'sync' segment state 'in fault' filerep state 'fault' process name(pid) 'mirror consumer append only process(15765)' 'cdbfilerep.c' 'L2480' 'FileRep_
SetSegmentState'",,,,,,0,,"cdbfilerep.c",1850,"Stack trace:
      0xb0d20e postgres errstart (elog.c:502)
       0xbf9004 postgres <symbol not found> (cdbfilerep.c:1832)
       0xbff5e2 postgres FileRepSubProcess_ProcessSignals (cdbfilerepservice.c:550)
       0xc06e2a postgres <symbol not found> (cdbfilerepmirror.c:664)
       0xc0bf55 postgres FileRepMirror_StartConsumer (cdbfilerepmirror.c:547)
       0xc0023d postgres FileRepSubProcess_Main (cdbfilerepservice.c:825)
      Oxbf6b6a postgres <symbol not found> (cdbfilerep.c:2703)
Oxbfea0a postgres FileRep_Main (cdbfilerep.c:3651)
       0x5f5f97 postgres AuxiliaryProcessMain (bootstrap.c:499)
      0x8f0f54 postgres <symbol not found> (postmaster.c:7603)
       0x8f113c postgres StartFilerepProcesses (postmaster.c:1631)
      0x9048af postgres doRequestedPrimaryMirrorModeTransitions (primary_mirror_mode.c:1752)
       0x8fd3af postgres <symbol not found> (postmaster_c:2484)
       0x8fec90 postgres PostmasterMain (postmaster.c:7603)
      0x7feb6f postgres main (main.c:206)
       0x7f0455417555 libc.so.6 __libc_start_main + 0xf5
      0x4bee79 postgres <symbol not found> + 0x4bee79
```



• 查看primary日志:

```
2020-08-12 03:59:30.115695 CST,,,p44215,th1249773440,,,,0,,,seg-1,,,,,"WARNING","01000","threshold '75' percent of 'gp_segment_connect_timeout=600' is reached, mirror may not be able to keep up with primary, primary may transition to change tracking",,"increase guc 'gp_segment_connect_timeout' by 'g
pconfig' and 'gpstop -u'",,,,,0,,"cdbfilerepprimaryack.c",838,
pconfig and gpstop -u ,,,,,o,, confifereprimaryack.c ,oso,
2020-08-12 04:02:05.207753 CSI,,,p44215,th1249773440,,,,0,,,seg-1,,,,"warning","01000", mirror failure, could not complete mirrored request identifie
r 'heartBeat' ack state 'waiting for ack', failover requested",,"run gprecoverseg to re-establish mirror connectivity",,,"mirroring role 'primary role
' mirroring state 'sync' segment state 'up and running' process name(pid) 'primary recovery process(44215)' filerep state 'up and running'
position ack begin '0x7f6a32e05040' position ack end '0x7f6a32e85040' position ack insert '0x7f6a32e25bf0' position ack consume '0x7f6a32e25bf0' position ack wraparound '0x7f6a32e85040' insert count ack '1549958' consume count ack '1549958' ",,0,,"cdbfilerepprimaryack.c",876,
2020-08-12 04:02:05.226125 CST,,p44215,th1249773440,,,,0,,,seg-1,,,,,"WARNING","01000","mirror failure, could not complete operation on mirror, failo
ver requested","identifier 'heartBeat' operation 'heart beat' relation type 'control message' message count '-1'","run gprecoverseg to re-establish mi
rror connectivity",,,"mirroring role 'primary role' mirroring state 'sync' segment state 'up and running' process name(pid) 'primary recovery process(
44215)' fileren state 'up and running'
44215)' filerep staté 'up and running'
position ack begin '0x7f6a32e05040' position ack end '0x7f6a32e85040' position ack insert '0x7f6a32e25bf0' position ack consume '0x7f6a32e25bf0' position ack wraparound '0x7f6a32e85040' insert count ack '1549958' consume count ack '1549958'
position begin '0x7f6a32604040' position end '0x7f6a32e04100' position insert '0x7f6a32be2dd0' position consume '0x7f6a32be2dd0' position wraparound '
0x7f6a32e04100' insert count '174734534' consume count '174734534' ",,0,,"cdbfilerepprimary.c",1394,
2020-08-12 04:02:12.457847 CST,,,p44211,th1249773440,,,,0,,,seg-1,,,,,"LOG","00000","'set segment state', mirroring role 'primary role' mirroring state 'sync' segment state 'in fault' filerep state 'fault' process name(pid) 'primary consumer ack process(44211)' 'cdbfilerep.c' 'L2480' 'FileRep_SetSeg
mentState'",,,,,,,0,,"cdbfilerep.c",1850,"Stack trace:
          0xb0d20e postgres errstart (elog.c:502)
          Oxbf9004 postgres <symbol not found> (cdbfilerep.c:1832)
          Oxbff5e2 postgres FileRepSubProcess_ProcessSignals (cdbfilerepservice.c:550)
          OxcOd223 postgres <symbol not found> (cdbfilerepprimaryack.c:995)
          0xc0dbd8 postgres FileRepAckPrimary_StartConsumer (cdbfilerepprimaryack.c:903)
          0xc002ad postgres FileRepSubProcess_Main (cdbfilerepservice.c:837)
          Oxbf6b6a postgres <symbol not found> (cdbfilerep.c:2703)
          Oxbfdeba postgres FileRep_Main (cdbfilerep.c:3585)
          0x5f5f97 postgres AuxiliaryProcessMain (bootstrap.c:499)
         0x8f0f54 postgres <symbol not found> (postmaster.c:7603)
         0x8f113c postgres StartFilerepProcesses (postmaster.c:1631)
```

0x90415a postgres doRequestedPrimaryMirrorModeTransitions (primary_mirror_mode.c:1752)

0x8fd3af postgres <symbol not found> (postmaster.c:2484) 0x8fec90 postgres PostmasterMain (postmaster.c:7603)

0x7f6a45f98555 libc.so.6 __libc_start_main + 0xf5 0x4bee79 postgres <symbol not found> + 0x4bee79

0x7feb6f postgres main (main.c:206)

Segment状态异常常见问题分析



Cannot allocate memory

- ➤ 如果是SQL执行过程遇到内存不足,SQL报错结束。报错信息是Out of memory,不会影响实例状态。
- → 如果是primary的服务进程,如: Checkpoint等进程遇到内存不足时, primary状态会down。master会切换mirror
 实例
- ► 检查当时的并发度
- ➤ 检查gp_vmem_protect_limit和系统参数vm.overcommit_ratio等配置是否合理

I/O error

- → 如果只是某个表的数据文件发生损坏,涉及该表的SQL会报错。这不会影响segment实例状态。
- ➤ 如果由于阵列卡等问题,导致整个盘无法访问,IO error报错。则会导致盘上的Segment实例down

gprecoverseg过程中无法启动被恢复的实例



- 在gprecoverseg命令执行过程中,启动被恢复的实例时就报错。
- 参考上面gpstart失败的分析方法,分析启动不成功的实例的startup.log和CSV日志。
- 几类常见问题:
 - ➤ 与gpstart类似的问题:/tmp/空间满了,配置文件中有错误,OS参数问题等
 - ➤ 如果一次操作恢复的实例比较多时,在primary与mirror之间建立数据复制连接时,有可能会超时
 - ➤ 如果系统表元数据有问题,也有可能导致实例启动失败。

gprecoverseg过程服务器再次宕机



- 如果在gprecoverseg过程中,服务器反复宕机,建议认真检修故障服务器,不要没搞清楚问题就反复尝试恢复。
- 如果无法修复,或者无法找到问题根源,可考虑更换服务器,或重装操作系统
- 如果遇到阵列卡故障,需由硬件厂家确认文件系统不会损坏,数据不会损坏。如果相关方面无法保证,可考虑采用重新格盘,做全量恢复。

gprecoverseg过程实例状态再变为down



- gprecoverseg命令正常结束,已经开始恢复数据,隔一阵发现被恢复的实例又变成down。
- 建议先排查实例down的原因,不要盲目重试gprecoverseg
- 排查方法,跟上面gpstate发现Segment状态不正常的检查方法相同。一对primary和mirror实例的CSV日志都需要 分析
- gprecoverseg命令结束,实例已正常启动,primary与mirror之间已开始resync。此时primary和mirror的CSV日志中会打印很多resync的LOG信息,可重点查找WARNING信息

• 几种错误分析:

- primary实例上数据文件不存在,情况一:对应表确实已经不存在,可能是当时系统繁忙,正准备恢复的表刚好被drop掉了。重做 gprecoverseg一般不会碰到同样的问题。但建议在系统压力较小时再做gprecoverseg
- ▶ primary实例上数据文件不存在,情况二:对应表是存在的,但该实例上数据文件不存在。这时表应该已不可用。推荐方案是重建表,从其他系统或者备份中恢复表数据,问题表drop掉。然后再做gprecoverseg。系统闲时还可以做一次gpcheckcat,查看是否还有其他系统表问题
- ➤ mirror实例上数据文件不存在,如果primary实例上数据文件是正常的,表是可用的。可以考虑两个方案:
 - alter table tablename set with (reorganize=true);
 - □ 原表rename,新建表并导入数据,drop原表
- ➤ 如果有人建议遇缺文件就touch一个空文件,建议咨询原厂工程师后再操作



● Standby master状态正常时,显示信息如下:

```
gpstate:mdw:gpadmin-[INF0]:-Starting gpstate with args: -f
gpstate:mdw:gpadmin-[INF0]:-local Greenplum Version: 'postgres (Greenplum Database) 5.19.0
gpstate:mdw:gpadmin-[INF0]:-master Greenplum Version: 'PostgreSQL 8.3.23 (Greenplum Databas
gpstate:mdw:gpadmin-[INF0]:-Obtaining Segment details from master...
gpstate:mdw:gpadmin-[INF0]:-Standby master details
gpstate:mdw:gpadmin-[INF0]:---
gpstate:mdw:gpadmin-[INF0]:-
                              Standby address
                                                        = smdw
gpstate:mdw:gpadmin-[INF0]:-
                              Standby data directory
                                                       = /data/master/gpseg-1
gpstate:mdw:gpadmin-[INF0]:-
                              Standby port
                                                       = 5432
gpstate:mdw:gpadmin-[INF0]:-
                              Standby PID
                                                        = 14344
gpstate:mdw:gpadmin-[INF0]:-
                              Standby status
                                                       = Standby host passive
gpstate:mdw:gpadmin-[INF0]:---
gpstate:mdw:gpadmin-[INF0]:--pg_stat_replication
gpstate:mdw:gpadmin-[INF0]:----
gpstate:mdw:gpadmin-[INF0]:--WAL Sender State: streaming
gpstate:mdw:gpadmin-[INF0]:--Sync state: sync
gpstate:mdw:gpadmin-[INF0]:--Sent Location: C6/8C01B0C0
gpstate:mdw:gpadmin-[INF0]:--Flush Location: C6/8C01B0C0
gpstate:mdw:gpadmin-[INF0]:--Replay Location: C6/8C01AB58
apstate:mdw:apadmin-[TNF0]:-----
```



● Standby master的状态未启动时,显示如下:

```
gpstate:dwtestcn001:gpadmin-[INFO]:-Starting gpstate with args: -f
gpstate:dwtestcn001:gpadmin-[INF0]:-local Greenplum Version: 'postgres (Greenplum Database) 5.19.0 build
gpstate:dwtestcn001:gpadmin-[INF0]:-master Greenplum Version: 'PostgreSQL 8.3.23 (Greenplum Database 5.19
gpstate:dwtestcn001:gpadmin-[INF0]:-Obtaining Segment details from master...
gpstate:dwtestcn001:gpadmin-[INF0]:-Standby master details
gpstate:dwtestcn001:gpadmin-[INF0]:----
gpstate:dwtestcn001:gpadmin-[INF0]:- Standby address
                                                               = dwtestdn009
gpstate:dwtestcn001:gpadmin-[INF0]:-
                                      Standby data directory
                                                               = /data/master/gpseg-1
gpstate:dwtestcn001:gpadmin-[INF0]:- Standby port
                                                               = 5432
gpstate:dwtestcn001:gpadmin-[WARNING]:-Standby PID
                                                               = 0
                                                                                               <<<<<<
gpstate:dwtestcn001:gpadmin-[WARNING]:-Standby status
                                                               = Standby process not running
gpstate:dwtestcn001:gpadmin-[INF0]:------
gpstate:dwtestcn001:gpadmin-[INF0]:--pg stat replication
gpstate:dwtestcn001:gpadmin-[INF0]:----
gpstate:dwtestcn001:gpadmin-[INF0]:--WAL Sender State: streaming
gpstate:dwtestcn001:gpadmin-[INF0]:--Sync state: sync
gpstate:dwtestcn001:gpadmin-[INF0]:--Sent Location: C2/11C08000
gpstate:dwtestcn001:gpadmin-[INF0]:--Flush Location: C2/11C08000
gpstate:dwtestcn001:gpadmin-[INF0]:--Replay Location: C2/11C00A40
gpstate:dwtestcn001:gpadmin-[INF0]:------
```

● 原因分析:

- ➤ 服务器被重启后,需要在master上重新触发同步: gpinitstandby –n
- Standby master启动失败、按照上面介绍的检查实例启动失败的方法进行分析。



Standby master与master之间不同步

```
gpstate:DWHGL1:gpadmin-[INF0]:-Starting gpstate with args: -f
gpstate:DWHGL1:gpadmin-[INF0]:-local Greenplum Version: 'postgres (Greenplum Database) 5.19.0
gpstate:DWHGL1:gpadmin-[INF0]:-master Greenplum Version: 'PostgreSQL 8.3.23 (Greenplum Databas
gpstate:DWHGL1:gpadmin-[INF0]:-Obtaining Segment details from master...
gpstate:DWHGL1:gpadmin-[INF0]:-Standby master details
gpstate:DWHGL1:gpadmin-[INF0]:-
gpstate:DWHGL1:gpadmin-[INF0]:-
                                  Standby address
                                                           = DWHGL2
                                  Standby data directory
gpstate:DWHGL1:gpadmin-[INF0]:-
                                                           = /data/master/gpseg-1
gpstate:DWHGL1:gpadmin-[INF0]:-
                                  Standby port
                                                           = 5432
gpstate:DWHGL1:gpadmin-[INF0]:-
                                  Standby PID
                                                           = 456105
gpstate:DWHGL1:gpadmin-[INF0]:-
                                  Standby status
                                                           = Standby host passive
gpstate:DWHGL1:gpadmin-[INF0]:----
gpstate:DWHGL1:gpadmin-[INF0]:--pg_stat_replication
gpstate:DWHGL1:gpadmin-[INF0]:-
gpstate:DWHGL1:gpadmin-[INF0]:-No entries found.
gpstate:DWHGL1:gpadmin-{INFO}:
```

● 重点查看Standby master的CSV日志

- ➤ 主机名,网络等问题
- ➤ Master的pg_hba.conf配置有问题,在master上,针对Standby主机IP,replication的配置行,必须是trust
- ➤ Standby master可能长时间未通信,master上较老的WAL已经丢弃。此时就要删除Standby master然后重建

扩容问题分析及处理



gpexpand——集群新增Segment的操作



- 1. 检查GP集群中各数据库的表
- 2. 检查新服务器配置一致性
- 3. 在master上打包元数据
- 4. 将元数据包分发到所有新服务器上
- 5. 配置新的Segment实例(包括priamry和mirror)
- 6. 清理临时文件
- 7. 配置新的Segment实例的filespace/tablespace
- 8. 清理新的Segment上多余的元数据
- 9. 重启数据库
- 10. 创建gpexpand模式,修改所有数据库下所有表(除根分区)的分布键为random(GP5及之前版本),把表信息插 入gpexpand.status_detail表中
- 11. 重启数据库
- 12. 执行gprecoverseg -F,对新服务器上的mirror实例进行全量恢复

gpexpand新增Segment时一些常见问题处理



- 目录权限问题
 - ➤ 需要确认扩容服务器上各个数据目录的权限,如:/data1/primary、/data1/filespace等
- 目录空间问题
 - ➤ master空间是否足够
 - 提前统计Master实例上元数据的空间,包括\$MASTER_DATA_DIRECTORY及各个Filespace(GP6是Tablespace)的空间容量。
 - 程序在Master数据目录(\$MASTER_DATA_DIRECTORY)下创建临时目录暂存所有扩容所需的元数据,需确保空间足够
 - □ 程序会在当前目录下打包元数据(创建gpexpand_schema.tar),文件不做压缩,需确保当目前目录空间足够
 - ➤ 新扩容服务器上空间是否足够
 - 建议使用 -t 参数指定在新服务器上接收gpexpand_schema.tar文件的目录
 - 确保接收gpexpand_schema.tar文件的目录空间充足
- 关注Master实例目录大小
 - ➤ 元数据打包时会忽略\$MASTER_DATA_DIRECTORY/pg_log和\$MASTER_DATA_DIRECTORY/gpperfmon/data两个目录,建议提前清理CSV日志让gpexpand更高效
 - ➤ 扩容前需确认,\$MASTER_DATA_DIRECTORY是否有其他不合理的大文件或者目录。如:coredump文件,备份的文件等。提前做好清理工作

gpexpand新增Segment时一些常见问题处理



- 扩容前建议进行gpcheckcat检查
 - ➤ 建议在gpexpand之前,安排专门的停机窗口,做系统表一致性检查(gpcheckcat)
 - ➤ 可与系统表vacuum full安排在一次停机窗口内操作
 - ➤ 系统表一致性检查需覆盖所有Database,包括gpperfmon库
 - ➤ 如gpcheckcat发现问题,请及时修复。需修复好之后才可进行gpexpand
- 检查是否安装perl和rsync
 - > which perl; which rsync
 - ➢ 可通过安装盘或者yum源可安装
- 参数-S | --simple_progress
 - ➤ gpexpand默认统计每张表的size,如果库内表数量很多,会导致gpexpand耗时很长。
 - ➤ 可使用-S参数, 忽略计算表size的过程, 效率可大幅提升
- 部分用户操作窗口(如: CRT)会超时退出甚至挂死,有可能会导致gpexpand命令被退出。如果类似情况,需要分析日志(~/gpAdminLogs/gpexpand_YYYYMMDD.log)看操作停在哪里,一般需要回滚重做,同时建议用nohup提交gpexpand命令



操作步骤:

- 1. 再次启动gpexpand命令,指定运行时长,指定处理并发,自动执行表重分布操作
- 2. 依据gpexpand.status_detail表中的信息,按照优先级排序(rank),逐个表执行重分布操作
- 3. 使用nohup 把 gpexpand放后台操作。重分布操作随时可以杀掉重调。

常见问题:

- ➤ 部分表被删除,重分布(alter table)的SQL报错。该报错忽略即可
- ▶ 如果在重分布过程中遇到宕机等硬件故障,导致重分布(alter table)的SQL报错。后续重新启动gpexpand重分布即可
- → 如果一张超大表并且严重倾斜,表的重分布操作会十分慢。建议直接alter table修改分布键,修改为合理的数据可均匀分布的分布键字段。

其他建议:

- ▶ 扩容前先做一次全面系统巡检,把严重倾斜的表修正过来。
- ▶ 扩容前务必做好新服务器的IO、网络性能测试;如果集群较大,新老机器之间也需要做网络压测。

运行中几类常见问题分析





- 内存相关问题
- PANIC
- Interconnect encountered a network error

out of shared memory



- 通常是由于某个参数设置值无法满足实际需求
- 查看相关的报错日志,日志中通常会提示需要增加那个参数值

20160918:03:23:47|gp_dump_agent-[ERROR]:-The command was: LOCK TABLE dw_dc.dw_oflogasset_1_prt_m20160520 IN ACCESS SHARE MODE 20160918:03:23:47|gp_dump_agent-[ERROR]:-*** aborted because of error: ERROR: out of shared memory HINT: You might need to increase max_locks_per_transaction.

2017-06-09 02:27:43.297676 EDT,"xxxxxxxxxxx","xxx",p633215,th65394464,"xx.xxx.xx.xx","55283",2017-06-09 02:27:43 EDT,98317350,con9254060,cmd2,seg-1,dx22330340,x98317350,sx1,"ERROR","53200","out of shared memory adding portal increments",,"You may need to increase max_resource_portals_per_transaction.",,,,"select nt.nspname as table_schema,

• 参数也不是调的越大越好。参数调的太大也可能会导致共享内存不足,会导致启动失败。

insufficient memory



- 报错发生在生成执行计划阶段。一般调大statement_mem可解决
- 什么情况下容易出现"insufficient memory"
 - ▶ 使用了一张子分区个数超多的分区表,对全表进行操作
 - ➤ 生成一个很复杂的执行计划,其中切片(slice)很多
- 优化建议
 - ➤ 合理规划和设计分区表,控制子分区的粒度和个数
 - ➤ 分析执行计划是否可以优化, SQL是否可以拆分

Out of memory



- 常见报错信息
 - Cannot allocate memory
 - Out of memory
 - Canceling query because of high VMEM usage
- 分析日志
 - Master和Segment都可能发生OOM,所有实例的日志都可能需要关注

- ▶ 检查报错时刻系统的并发度
- ▶ 内存严重不足的情况,服务进程可能都无法获取内存,会引起down实例
- 分析报错的SQL
 - ➢ 对于简单的SQL,维护的SQL,则可排除嫌疑
 - ➢ 对于日常能正常运行的SQL,如果出现OOM,可分析统计信息和数据是否有问题。如果重新提交SQL可以正常执行则可排除嫌疑

内存相关配置



- OS
 - vm.overcommit_memory = 2
 - vm.overcommit_ratio = 95
 - ➤ OS总可用内存计算方法: RAM × (vm.overcommit_ratio ÷ 100) + SWAP
- Resource Queue
 - ➤ gp_vmem_protect_limit -- 限制每个Instance上所有语句可以使用的内存总量的上限值。
 - ➤ resource queue中的memory_limit,建议如果加多大SQL在运行的队列,都建议设置memory limit,所有多些的 memory limit总和不要超过gp_vmem_protect_limit
 - ➤ 为了提升效率减少workfile的输出,可调大Statement_mem。但不能滥用,以免造成浪费
 - ➤ 内存配置计算指引: https://greenplum.org/calc/

Resource Group

- ➤ gp_resource_group_memory_limit设置了每个GP Host主机上可以分配给资源组的系统内存的最大百分比。缺省值为0.7(70%)
- ➤ 每个Instance能获得的内存算法: rg_perseg_mem = ((RAM × (vm.overcommit_ratio ÷ 100) + SWAP) × gp_resource_group_memory_limit) ÷ num_active_primary_segments
- ➤ 每个资源组可设置专用的内存MEMORY_LIMIT,其中在设置每个组内共享内存MEMORY_SHARED_QUOTA。当资源组内共享内存用尽,全局共享内存也用尽,SQL还需要内存则会报错。

应用优化



• 并发控制

- ➤ 系统并发没有控制好,或者不做控制,OOM和系统不稳定的风险就会增加
- ➤ 通过Resource Queue或Resource Group做好并发度控制
- ➤ 24小时内,可以在不同时间段采用不同的控制策略

• 分区表设计和使用

- ➢ 要控制好单表子分区的数目。
- ➢ 分区表的维护策略,如:添加分区策略,删除旧分区策略,合并旧分区策略
- ➢ 及时收集统计信息

• SQL优化

- ➤ 分区表全表关联delete或update,保证分布键一致及用分布键关联
- ➤ 注意执行计划中分区裁剪是否合理,分区字段过滤条件是否不生效
- ▶ 减少超复杂超长的SQL,合理的简化和拆分SQL,效率可大幅提升

• 可调整参数

gp_max_slices



- 内存相关问题
- PANIC
- Interconnect encountered a network error

了解一下PANIC



• 什么是PANIC

- ➤ 有postgres进程异常退出或被kill
- ➤ Master或者Segment实例都可能发生PANIC
- ➤ 通常会查到"PANIC"这个关键字(注意不是ERROR/FATAL)
- ➤ 如果没找到"PANIC",还可以查找"process (PID XXXXX) was terminated by signal X"类似的关键字
- 发生PANIC后,数据库实例会做什么
 - ➤ 终止所有活动链接(包括所有在运行的SQL),在运行的SQL都会报错,报错信息是"server closed the connection unexpectedly"或者"Unexpected internal error"
 - ➢ 服务进程都被停止,清理临时文件等操作。此时新发起的SQL会收到报错"the database system is in recovery mode"
 - ➤ 服务进程自动重启,启动期间新发起的SQL会收到报错"the database system is starting up"
 - ➤ 服务进程启动完成后,重新恢复可接受连接,从日志中会看到信息"database system is ready to accept connections"
- PANIC可能是BUG或人为kill进程或其他外部因素引起的。
 - ➤ 如果从日志中发现was terminated by signal 9,则可以确定是被kill或者其他外因导致的。
- 发生PANIC后如何应对

如何定位问题



- 可以怀疑发生了PANIC的报错信息
 - server closed the connection unexpectedly
 - Unexpected internal error
 - the database system is in recovery mode
 - the database system is starting up
- 分析日志技巧
 - ➤ Master和Segment实例日志都需要检查

➤ 如果找不到PANIC关键字,按照"process (PID XXXXX) was terminated by signal X"查找,定位到那个进程号被终

2018-09-07 17:21:40.908165 CST,,,p152741,th-2105387232,,,,0,,,seg-1,,,,,"LOG","00000","server process (PID <mark>152752</mark>) was terminated by signal 6: Aborted",,,,,,,0,,"postmaster.c",5620,

- ➤ 如果SQL在master上发生PANIC,可能是生成执行计划阶段产生的PANIC
- ➤ 如果SQL在segment上发生的PANIC,可能是SQL执行过程中产生的PANIC

辅助分析PANIC的方法



- 开case或ticket,需要提供core文件和数据库日志帮助分析
- 操作系统允许输出进程的core文件
 - ➤ 在/etc/sysctl.conf中设置core位置,如:kernel.core_pattern = /data2/coredump/core.%e.%p.%t.%s.%u.%g
 - ➤ 在/etc/security/limits.conf中设置gpadmin不限制core文件大小"gpadmin core unlimited";或者在环境变量中设置 "ulimit -c unlimited"
- 打包core文件的工具
 - packcore
 - ➤ gpmt
- SQL问题重现工具minirepro
 - ▶ 收集问题SQL涉及表的元数据以及统计信息,方便在其他环境上分析和复现问题
 - ➤ 参考社区的文章《Greenplum SQL问题重现利器Minirepro》https://mp.weixin.qq.com/s/a505azzg7Qduqlhl-kPmBA



- 内存相关问题
- PANIC
- Interconnect encountered a network error



- Interconnect encountered a network error是一种发生在SQL执行过程中的实例之间通讯的报错
- 从某个实例到另一个实例的通讯链路上的阻塞,跟普通的网络丢包无关

```
2020-05-03 14:49:35.088777 CST,"gpmon","gpperfmon",p450729,th-56326272,"10.192.68.203","14676",2020-05-03 12:49:32
CST,37526683,con6914329,cmd16,seg11,slice1,dx53101981,x37526683,sx1,"ERROR","58M01","Interconnect encountered a network error, please check your network","Failed to
send packet (seg 1) to 10.192.68.195:16147 (pid 136323 cid 0) after 3574 retries in 3600 seconds",,,,,"INSERT INTO gpmetrics.gpcc plannode history SELECT * FROM
gpmetrics. gpcc plannode history",0,,"ic udpifc.c",4920,
2020-05-04 01:39:26.021365 CST,"hdm usr","zzgpdb",p214194,th-56326272,"10.192.68.203","8847",2020-05-04 00:39:25
CST,0,con6939534,cmd14,seg11,slice1,dx53245527,,sx1,"ERROR","58M01","Interconnect encountered a network error, please check your network","Failed to send packet (seq

    to 10.192.68.195:4791 (pid 309803 cid 1) after 3581 retries in 3600 seconds",,,,,"INSERT INTO hdm.CBV7_KCEB_PZYJCZ (

2020-05-04 02:44:30.450438 CST,"hdm_usr","zzqpdb",p292526,th-56326272,"10.192.68.203","61266",2020-05-04 01:44:30
CST,0,con6946659,cmd14,seg11,slice1,dx53267620,,sx1,"ERROR","58M01","Interconnect encountered a network error, please check your network","Failed to send packet (seq

    to 10.192.68.199:4791 (pid 65862 cid 18) after 3581 retries in 3600 seconds",,,,,"INSERT INTO hdm.SEAI_NWBK_UKEY (

2020-05-04 03:10:30.965715 CST,"fdm_usr","zzgpdb",p326196,th-56326272,"10.192.68.203","30931",2020-05-04 02:10:30
CST,0,con6949860,cmd60,seg11,slice12,dx53277148,,sx1,"ERROR","58M01","Interconnect encountered a network error, please check your network","Failed to send packet
(seq 1) to 10.192.68.195:4791 (pid 388523 cid 3) after 3580 retries in 3600 seconds",,,,,"INSERT INTO fdm.dcev_zq_zcgl_j01_tm(
2020-05-04 03:10:31.655621 CST, "fdm_usr", "zzgpdb", p326129, th-56326272, "10.192.68.203", "30573", 2020-05-04 02:10:30
CST,0,con6949860,cmd60,seg11,slice11,dx53277148,,sx1,"ERROR","58M01","Interconnect encountered a network error, please check your network","Failed to send packet
(seq 1) to 10.192.68.195:4791 (pid 388523 cid 3) after 3581 retries in 3600 seconds",,,,,"INSERT INTO fdm.dcev_zg_zcgl_j01_tm(
```

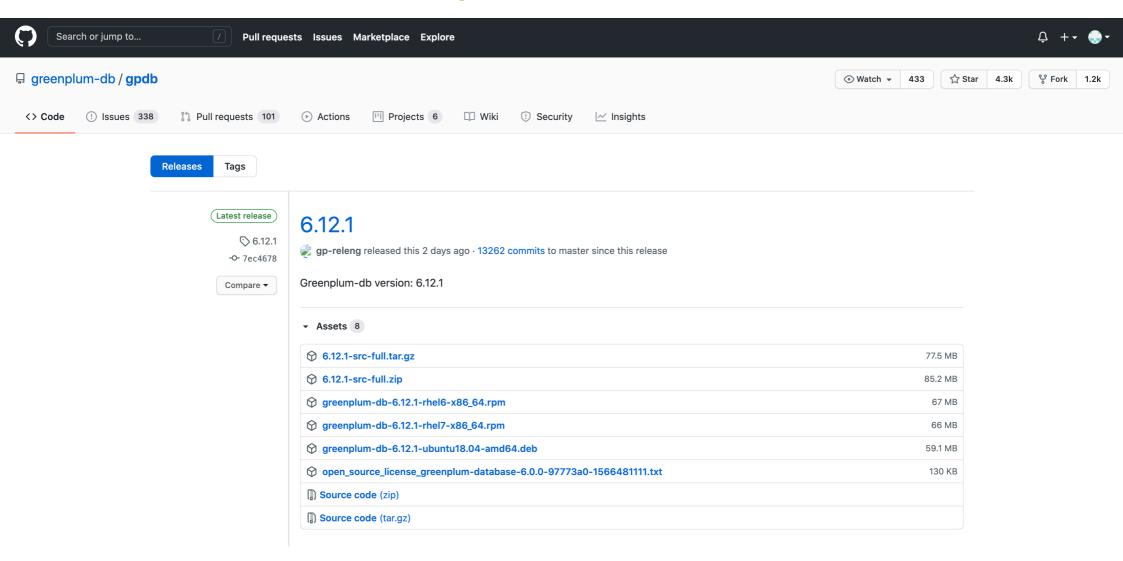
• 几类遇到过的问题

- ➤ 交换机故障
- ➤ 某些特殊的端口不可用,但在主机上看不到端口侦听
- ➤ 一些复杂的SQL或者写法待优化的SQL、导致服务器负荷过高
- 应对方法
 - ➤ 提前做好网络压力测试,提前发现问题,建议使用能够测试UDP的工具,如: iperf3, netperf
 - ➢ 通过net.ipv4.ip_local_reserved_ports跳过某些特殊端口,通过net.ipv4.ip_local_port_range限制端口范围

试试看吧! GREENPLUM DATABASE° ፟多腾讯云大学

欢迎下载试用Greenplum





一些小工具



https://github.com/davidch3/gp_simpletools

感谢观看



Greenplum中文社区公众号



Greenplum微信技术讨论群



