Statement 模式 ：一些函数不能被复制，例如 insert into tb（col1） values （uuid()）,类似这样的的函数无法被复制，导致主从数据不一致，后来推出了row模式。

Zh 巧 00a0 年 0 
statement 格 式 复 制 不 足 及 解 决 方 法 
· 理 R?binlog 
． 根 据 日 志 格 式 复 制 分 ： 
． 根 据 日 志 定 义 的 格 式 不 一 样 ， 可 以 分 为 ： Statement(SBIR) 格 式 ， Row(RBR) 格 式 或 
是 M 工 格 式 
' 记 录 最 小 的 单 位 是 一 lEvent. 日 志 前 4 个 字 节 是 一 个 mag number, 接 下 19 个 字 
节 讠 己 录 Format des c event ： FDE 
个 事 务 由 多 个 E “ nt 组 成 
相 关 的 包 含 ： binary log 和 binary log index 文 件 
•Bin 10g 
•MYSQL 5 、 6 增 加 了 G “ D 复 制 

(root@dzst143 trysq133071# /usr/local/mysql/bin/mysqld 
16498 
Croot@dzst143 mysql -S /tmp/mysq133Ø7 sock -p 
Enter password; 
Welcome +0 the MySQL monitor, Cmnand5 end with ; or 
Your VySQL connecticn id is 10 
Server version: 5.7.24-10g "YSQL Catm,lnity Server (CPL) 
Copyright (c) 2øøø, 2018, Oracle ard/or its affiliates. All rights reserved- 
Oracle is a registered trademark of Oracle Corporation and/or its 
affiliates, Other nanes my be traderarks of their respective 
Type • help;' or • for help, Type to clear the current input 5taternent- 
rysql> shc*/ global variables like 

为了查看statement格式有什么不足，在session级别把binlog改位statement模式。

bi nl og_fo rmat 
1 n _ sync _ e ay 
bi nl I ay_count 
bi nl mpl ecovery 
bi nl ry_l og_eve nts 
bi nl og_transact epend ency_hi story_si ze 
bi nl og_tra nsact i on—d epend ency_tracki ng 
max—b i z e 
max—b z e 
sync_bi nl og 
22 in set sec) 
set statement' ; 
rysq 
rows a ec e 
sec 
SERVER 
ROW 
ON 
ON 
I FULL 
ON 
1 32768 
1 25øøø 
I COWIT_ORDER 
OFF 
OFF 
ON 
1 18446744737054752ø 
1073741824 
1 18446744737054752ø 
Il 

面试可能遇到装逼的问题：你的binlog里面结构是什么样子？Binlog event有那些？

binlog前四个字节是magic number

æøøøøø 62fe 
wøømø 7bØØ øøøø 
wøøøsø Ød38 Ø8øø 12øø 
elab 
wøøæø øøøø øøøø 
42f 
wøøøæ lau øøe8 øøøø 0808 0208 øøøø 
æøøø7ø 2a2a 12ØØ Øø34 701 b67F 2f9f 
øøøø 
5bfb 
wøøæø 6f6c øø67 øøøø øøøø øøøø øøøø øøøø øøøø 
æøøø30 øøøø øøøø øøøø øøøø øøøø øøøø øøøø øøøø 
wøøuø øøøø øøøø øøøø øøøø øøøø øøøø øøøø 13øø 
Ø4.øø 
alel 
Croot@dzst143 10 s # hexdUmp mysqI-bin.ØØØØØ6 Il head —n 10 
ø4Ø4 12Ø4 øøøø øø5f 
?feg ?8f1 e811 2f82 
abØf 
øøøl øø47 øøøø øøc2 øøøø øøgø øøøl 
6e69 
ø4øø 35øø 372e 322e 2d34 
77ØØ ØøøØ 
Øaøø 
fbf4 
235b 

Croot@dzst143 logsJ# echo "bin" I hexdump 
æøøøøø 6962 ea6e 
Croot@dzst143 logsl# echo 
Croot@dzst143 logsJ# echo I hexdump 
æøøøøø 
æøøøø2 
Croot@dzst143 logsJ# echo 
æøøøøø øa62 
2 
Croot@dzst143 echo 
æøøøøø øa6e 
"b" I hexdunp 
"n" I hexdump 

69是I 62是b 6e是n，那么magic number怎么解释？magic number就是在写某个文件之前大家约定俗成固定的写某个数，作为这个文件的开头。

[root@dzst143 mysqlbirilog -v --base64-output=decode-rows mysq1-bin.ØØØØØ9 >9.sq1 
[root@dzst143 

然后vim 9.sql，可以看到

# 123 
#181126 22: server id 1233Ø7 end- 
log—pos 
# 1-18 
# 194 
#181128 server id 1233Ø7 end- 
log—pos 
rbr-only=no 
SET 
194 CRC32 exfb4e6453 
CRC32 exe6bf4473 
Previous-GTIDs 
GTID 
Query 
last—conumi tted=Ø 
# 259 
#181128 server id 1233Ø7 end-log-pos 365 CRC32 ex3bØ8d1cc 
se 
SET 
SET Qsession. æsession. @@session.unique—checks=l, 
SET 
SET 
utf8 
SET @@session. 
SET 
q:reate table tb3(c1 varchar(32)) 
sequence—number—I 
error—cod 
15901270211 A. 
# at 365 
#181128 ze:46 
at 430 
server id 123307 end-log-pos 43e CRC32 ex483bO%b 
GTID 
5765809 
last—commi teed—I 

ddl：sbr，rbr 全部记成statement模式。

BEGIN 
# 862 
= 627 
#181128 server id 1233Ø7 
SET 
#181128 server id 123307 
end—log—pos 
end—log—pos 
end—log—pos 
966 CRC32 
862 CRC32 exe6b66f7S 
ex4ec1a8ca 
997 CRC32 ex4d61eefØ 
Query 
Query 
Xid 
error—cod 
error—cod 
insert into tb3 
# at 966 
#181128 server id 1233Ø7 
57.1 

我们知道复制是通过binlog重放来运行的，在statement模式下，binlog记录了上面的insert 语句，因为uuid()每次执行的结果都不一样，可能会造成主从数据不一致。所以类似uuid()每次执行结果不一样的函数在statement模式下都会出现主从数据不一致的问题。

知 数 堂 
statement 格 式 复 制 不 足 及 解 决 方 法 
． 基 于 语 旬 级 的 制 
早 in100 了 or 们 at 一 St 
． 优 点 ； 
` 0 “ 00 文 件 较 小 
` 日 志 是 包 含 用 户 执 行 的 原 始 50L 、 方 便 计 和 市 计 
出 现 最 早 可 的 兼 容 较 好 
． ] 明 方 使 阅 读 ， 方 使 故 嶂 10 复 
． 矇 点 ， 
` 存 安 全 喼 患 ． 可 能 导 致 主 从 不 一 致 
` 对 一 些 系 毓 函 数 不 准 复 制 或 是 不 能 复 制 
一 的 aci i10 0 
《 00r 0 
卜 ' out 〕 ci rows 的 
[mysqld] 
binlog_format— 、 statement 

在row模式下进行同样的创建表，insert数据使用uuid()这样的函数时binlog时什么样的？

shay birilog events in 'mysql-bin.ØØØØ1e• 

上面的语句也可以查看日志。

mysql -bin 
mysqL -bin , 
mysql -bin . 
mysql -bin . 
mysqL -bin 
mysql -bin . 
504 1 BEGIN 
mysql -bin 
mysqL -bin , 
mysql -bin . 
mysql -bin . 
mysql -bin 
mysql -bin . 
øøøetø 
oøøø10 
aaaølø 
øøømø 
oaøø10 
øaømø 
øøøølø 
oaøø10 
øøømø 
oaaø10 
4 
123 
194 
259 
367 
432 
504 
558 
606 
711 
776 
p rev ious_gtid5 
I Gtid 
Query 
I Gtid 
Query 
ry 
Wri 
Xid 
I Gtid 
Query 
123307 
123307 
123307 
123307 
123307 
123307 
123307 
123307 
123307 
123307 
123307 
123307 
123 
Server vert S. 7.24-log, Binlog ver: 4 
194 
I ela1e97f-ftrg-11e8-822f-Ø242ac12ec:1-24 
SET GTID-NEXT- 'ela1e97F-f178-11e8-822F-0242 
259 
367 
use 'wubx•; create table tb4 (cl varchar(128)) 
432 1 
SET —SESSION. GTID-NEXT- 'ela1e97f-f178-11e8-822f-Ø242 
SSE 
* 'nsert into tb4 
I table—id: 158 Gubx.±b4) 
6Ø6 
158 Flags: STMT-END-F 
680 
711 
776 
856 
cowrT xid-636 
SET &SESSION. GTID-NEXT- ie1a1e97f-f178-11e8-822f-Ø242 
BEGIN 

解析binlog10，使用vim来查看。主要看insert部分。

# 558 
*181178 ze. 
INSERT 
SET 
@1-'9e607a13-f3ec-11e8-822f-Ø2420c120tOc' 
= 06 
•53:35 
INTO 
server 
- tb'l 
12.3307 
end—log-pos 
686 CRC32 Øx5e97c4c7 
RC32 
Øx4b6c4Sac 
RC32 Øx1d2672e7 
T dile—nnp : 
Wri te_ rows : 
'wubx• . ' napped 
table id 158 flags: 
at 68e 
server 
en 
og_pos 

可以看到insert语句中在statement模式下是uuid()的部分改为了具体的值，而且在insert的时指明了那个数据库。

statement 
． 解 决 办 法 、 使 用 Ro “ 格 式 的 b og 
5 165809 
15901270211 
、 优 点 ： 
． 相 比 “ TE 旺 下 T 更 加 安 全 的 复 制 格 式 
制 不 足 及 解 决 
[mysqld] 
binlog format— 、 row' 
、 在 某 些 情 况 下 复 制 速 度 更 快 00 ， 复 杂 ， 表 有 主 建 ） 
． 系 统 的 特 殊 函 数 也 以 复 制 
、 更 少 的 锁 
比 较 大 〖 支 持 0n1 四 row_image) 
•Binary 10q 
． 单 语 句 更 新 [ 删 除 ] 表 的 行 数 过 多 ， 会 形 成 大 量 b 工 n 上 “ 
、 无 法 从 凿 。 № g 看 见 用 户 执 行 的 ， (Event ． 
， 记 录 用 户 的 q 过 “ 司 
rO query 10q events 

row模式：

ZhiShutang.corn 
row 格 式 binlog 执 行 流 程 
新 ： 地 数 
0 0 苤 
Slave IO_TH READ 

主键不建议使用uuid，因为uuid没有顺序，而且uuid的值时前面几位不一样，如果使用uuid会造成主键离散度比较大。前缀儿索引无法使用。

mysql> select Wid(); 
I uuid() 
0826423 f3Ød-11e8-822f-Ø242ac12WØc 
ow in t (O.øø sec) 
sel ct Wid(); 
uidO 
Øfa859d f3Ød-11e8-822f-Ø242ac12WØc 
ow in t (O.øø sec) 
sel ct Wid(); 
1 
1 
1 
uidO 
f16fudc 
f3Ød-11e8-822f-Ø242ac12WØc 
row in set (0.01 sec) 

在实际工作中可以时int或者bigint来做主键。

Mixed 模式：过度版本。

新特点： GTID，给每个事物（写入动作）做一个唯一的编号。这要主库执行了那些日志从库执行了那些日志就可以看清了。

gtid指明了事物由那个机器产生的，一共产生了多少事物，那些事物是复制过的那些事物是没有复制过的。

编号：server-uuid:1-n

Gtid\_pured

dh•shutang.errrr 
rootVlocaLhost sock global variables like ; 
I Variable—name 
I gtid_executed 
41eScfdd-7aØO-11e8-bd24-0242ac11ØØØ2:1-85Ø148 1 
I gtid_mode 
I gtid-owned 
I gtid_purged 
Value 
I ON 
1 123bgS38-SfSf-11e8-a6d6-Ø242ac11ØØØ2:1, 
loøø 
I ON 
123b9538- Ile8-a6d6-Ø242ac11W02 1, 
41eScfdd-7aØO-11e8-bd24-0242ac11ØØØ2:1-85Ø133 
I OFF 
8 rows in set CO. W sec) 

I gizid—mode 
I gtid—owned 
gtid—purged 
ON 
I ela1e97f-fyg-11eg-822f-e242ac12ØØØc:1-13 
OFF 
8 
session_t rack—gtå d5 
rows in set 0.01 sec) 
rysql> purge binary logs to 'nysq1-bin.WØØ11' • 
Q'ery 
rows a 
sec 
ysql> shmv global variables like 
Variabl e—nane 
bi e_recove ry 
gkid—execueed 
gtid-execuked—cotnprcssion-period 
I gtid—mode 
gEid_owned 
_ purged 
rows in set (O.øø sec) 
8 
Value 
ON 
elcle97i 
OFF 

GTID的魅力之一：

sı 
Sil 
S22 
Sil 
S22 

如果S1节点挂了，那么S11可以很快指向S2,S22,M三者之一。

具体做法：利用搭建从库的命令，其中master\_auto\_position=1的作用很明显。可以很快的同步数据。具体为什么看day08笔记中的master\_auto\_position=1是什么意思的解答。

change ester to 
•aster_passvord—• rep 14sL&ve•, 
te it ion=l ; 

====背下来===

MySQL 复 制 原 理 
℃ thread 
。 关 联 复 制 分 类 
· 多 源 复 制 
· 基 *Filename 十 position 
。 基 于 GTID 自 动 找 同 步 位 置 
•replication crash recovery 
有 处 
复 制 中 断 ， 一般 是 SQL-thread 出 错 了 。 
SQL thread 
· 早 期 单 进 程 
· 5 “ 6 基 于 database9i 别 的 并 行 
· 5 “ 7 基 于 事 务 级 别 的 并 行 
•binlog group commit 
明 ， 0 基 于 行 级 别 的 并 行 

每个版本复制的时候每个线程再干什么？

疑问？为什么是sql\_thread出错？

复 制 应 用 场 景 
主 库 珏 力 较 大 抗 不 住 
· 把 读 压 力 分 担 一 部 分 到 从 库 上 
· 业 务 高 崆 数 哐 隆 连 接 数 太 多 
· 例 如 高 悭 的 认 证 ， 只 需 要 把 这 个 业 务 分 拆 到 从 库 上 
· 有 一 些 大 的 SOL 出 现 不 多 ， 不 好 优 化 ． > 专 用 从 
． 后 面 的 缔 计 分 析 奕 的 SQ 〔 
钶 利 用 库 高 可 用 切 换 
5 化 09 

认证的

读压力比较大，分到从库上去读取。

主从库的索引不一致没有问题。

从库没有复制完可以切换为主库么？

复 制 应 用 场 景 
· 利 用 主 从 复 制 ： 实 现 多 IDC 架 构 

row+gtid 知数堂的复制。

双活不是再mysql实现的，都是再业务层控制的。