
GOTC 2023
全球开源技术峰会
THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE
OPEN SOURCE, INTO THE FUTURE
「数据与数据库技术」专场
做中国广受欢迎的开源数据库
叶金荣
2023.05.28

目录

CONTENTS

GreatSQL简介

GreatSQL是什么?

由万里数据库主导的开源MySQL分支

为什么要做GreatSQL?

- MySQL is The world's most popular open source database
- But

Steinar H. Gunderson

Sun, 05 Dec 2021 - Leaving MySQL

Today was my last day at Oracle, and thus also in the MySQL team.

When a decision comes to switch workplaces, there's always the question of "why", but that question always has multiple answers, and perhaps the simplest one is that I found another opportunity, and and as a whole, it was obvious it was time to move on when that arrived.

But it doesn't really explain why I did go looking for that somewhere else in the first place. The reasons for that are again complex, and it's not possible to reduce to a single thing. But nevertheless, let me point out something that I've been saying both internally and externally for the last five years (although never on a stage—which explains why I've been staying away from stages talking about MySQL): *MySQL is a pretty poor database, and you should strongly consider using Postgres instead.*[1]

stackoverflow About Products For Teams Search...

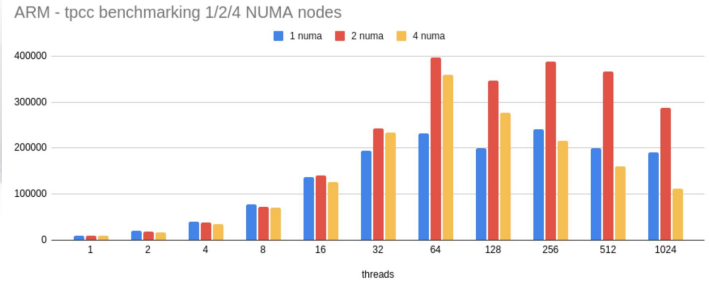
Home

PUBLIC

Questions

How do I create a sequence in MySQL?

Asked 8 years ago Modified 2 months ago Viewed 201k times



Oracle
@Oracle

Replying to @FedorovMykhailo and @SAP

On behalf of Oracle's 150,000 employees around the world and in support of both the elected government o Ukraine and for the people of Ukraine, Oracle Corporation has already suspended all operations in the Russian Federation.

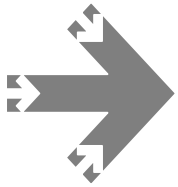
为什么是万里数据库?

01 中国第一个MySQL认证的金牌合作伙伴

02 中国第一个MySQL研发中心

03 中国第一个MySQL教育中心

04 中国第一个MySQL商业案例



01 修复NDB Cluster bug > 100

02 修复Replication bug > 300

03 完成NDB Cluster 新功能 (WorkLog) > 10

04 完成Replication 新功能 (WorkLog) > 20

05 完成MySQL 5.1 中文参考手册翻译

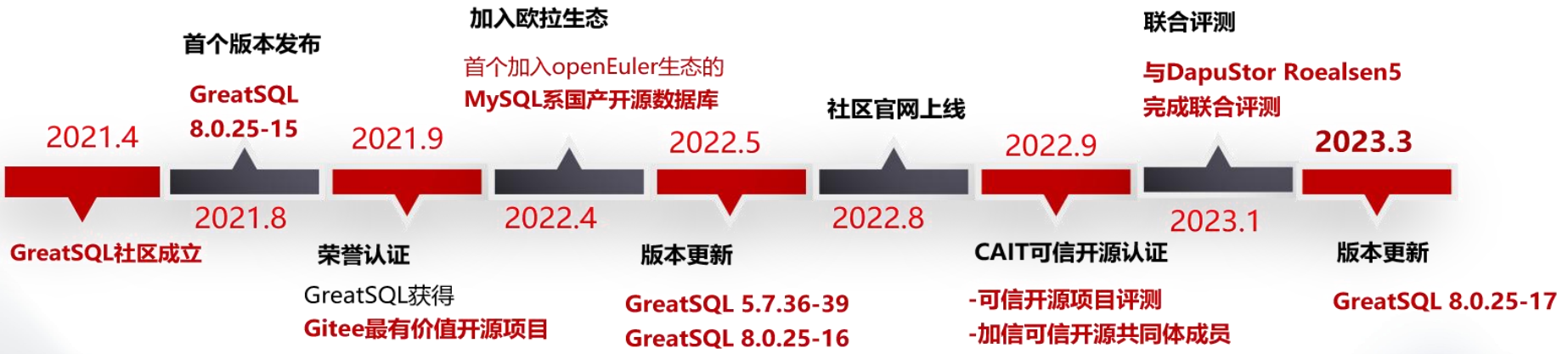
ID#	Date	Updated	Type	Status	Sev	Version	OS	CPU	Summary
108190	2022-08-18 13:02	2022-09-05 2:52	Connector / ODBC	Analyzing (52 days)	S1	8.0.30	Any	x86	ODBC SQLGetData return wrong value
107969	2022-07-26 3:48	2022-07-26 5:29	MySQL Server: Optimizer	Verified (93 days)	S2	8.0.29	Any	Any	wrong result when search binary columns
107635	2022-06-22 14:47	2022-07-20 15:43	MySQL Server: Group Replication	Closed (104 days)	S1	8.0.*	Any	Any	event scheduler cause error on group replication
107559	2022-06-14 8:45	2022-06-14 12:23	MySQL Server: Optimizer	Not a Bug (135 days)	S3		Any	Any	Why Switch_ref_item_slice in TemptableAggregateIterator::Init
104629	2021-08-16 3:08	2021-08-16 7:09	MySQL Server: Optimizer	Verified (437 days)	S1	8.0.25,5.7.35, 8.0.26	Any	Any	wrong result when outer join prune partition tables with is null predicate
103040	2021-03-18 14:50	2021-03-19 9:46	MySQL Server: Group Replication	Verified (587 days)	S3	8.0.*	Any	Any	minor fix for DEBUG message in XCOM
100800	2020-09-10 12:51	2021-12-02 11:20	MySQL Server: Optimizer	Closed (469 days)	S2	8.0.21, 8.0.11	Any	Any	wrong result when select int column with range
100783	2020-09-09 12:29	2020-09-10 4:50	MySQL Server: Optimizer	Can't repeat (777 days)	S2	8.0.19	Any	Any	wrong result with hash join
99647	2020-05-20 13:11	2020-05-22 12:02	MySQL Server: DML	Not a Bug (888 days)	S5	8.0.*	Any	Any	call file->position when necessary in sql_delete.cc
99628	2020-05-19 9:16	2020-05-28 14:59	MySQL Server: Replication	Verified (889 days)	S2	8.0.*	Any	Any	semi sync master not handle ack packet correctly when recv packet timeout
99094	2020-03-27 4:52	2020-03-27 10:38	MySQL Server: Information schema	Verified (944 days)	S1	8.0, 8.0.19	Any	Any	coredump when install information schema plugin

GreatSQL社区

社区简介

分类	介绍
关于GreatSQL社区	GreatSQL社区成立于2021年，是一个MySQL开源数据库社区，由万里数据库发起，致力于通过开放的社区合作，构建国内自主MySQL版本及开源数据库技术，推动中国开源数据库及应用生态繁荣发展。
愿景	成为中国广受欢迎的开源数据库社区
关于GreatSQL开源数据库	GreatSQL开源数据库是适用于金融级应用的国内自主MySQL版本，专注于提升MGR可靠性及性能，支持InnoDB并行查询等特性，可以作为MySQL或Percona Server的可选替换，用于线上生产环境，且完全免费并兼容MySQL或Percona Server。

GreatSQL社区发展历程



GreatSQL社区现状

01

国内活跃的MySQL开源社区

- 活跃的社区微信群、QQ群 & 微信公众号
- 活跃参与者超2000人

02

获得中国信通院 CAIT可信开源认证

- 可信开源项目认证
- 入选可信开源共同体成员

03

加入openEuler生态

- 首个加入openEuler生态的MySQL系国产开源数据库，openEuler22.09版本正式合入GreatSQL

GreatSQL社区技术支持与服务

- 技术支持与服务
  - 免费技术支持
  - 在线技术交流群
  - 提供Docker镜像
  - 提供Ansible一键安装包
  - 相关文档、视频

- 相关资源
  - 官网：https://greatsql.cn
  - 代码：https://gitee.com/GreatSQL
  - 文档：https://gitee.com/GreatSQL/GreatSQL-Doc
  - 社区：微信群、QQ群、微信公众号
  - openEuler生态 https://gitee.com/src-openeuler/greatsql

GreatSQL当前主要用户

GreatSQL社区从上线初期就收获了社区用户的高度关注，短时间内就聚集了上千位专业DBA的粉丝群体，以及有几十位社区企业用户：

- 恒生电子旗下的芸擎网络科技
- 深圳华润
- 靠谱云
- 中信建投
- 福建福富
- 作业帮
- 建信金科
- 直真科技
- 通达信
- 力维智联

GreatSQL特性

产品目标

1. 性能
2. 稳定性
3. 易用性

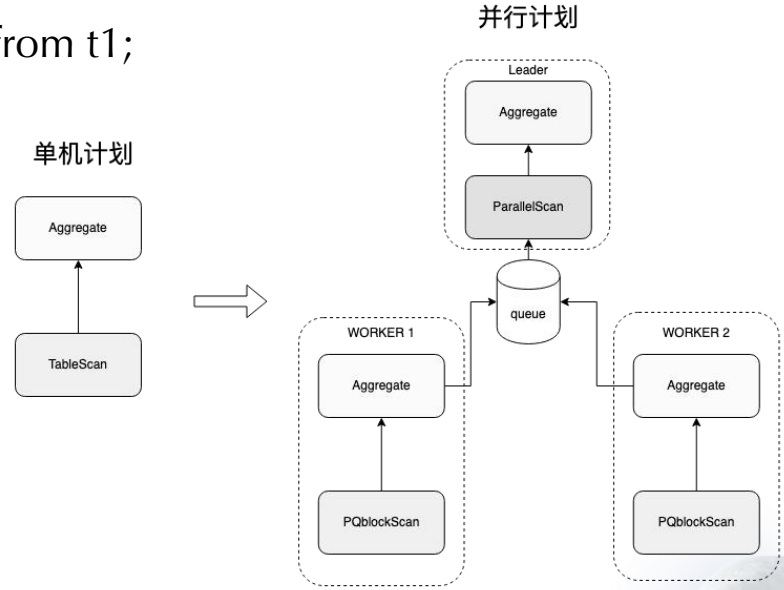
性能提升

- InnoDB Parallel Query
- Parallel Load data

性能提升：InnoDB Parallel Query

SELECT AVG(a) from t1;

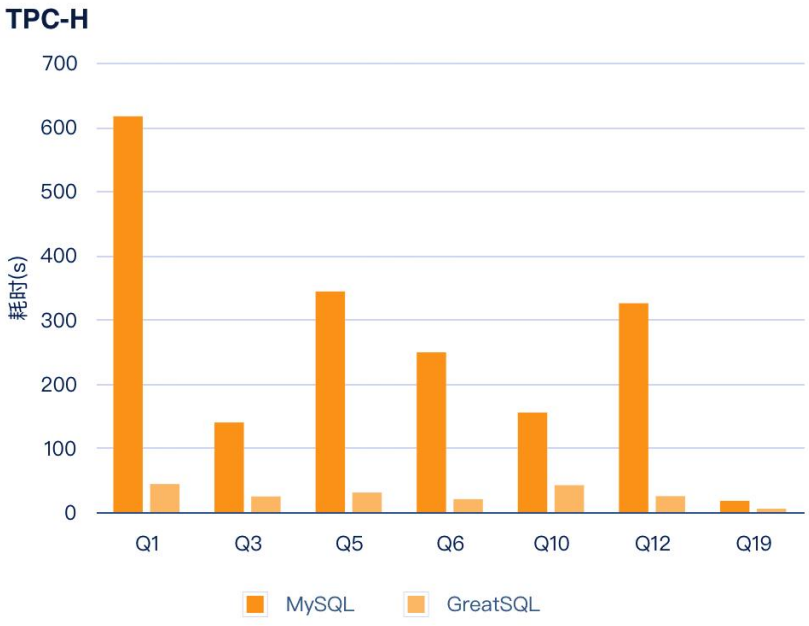
- 单线程计划 => 多线程并行计划
  - Leader & Worker Threads
    - worker线程：并发抽取数据，执行计划
    - Leader线程：汇聚worker线程的计算结果
  - 各个线程拥有各自的计划
    - TableScan=> ParallelScan/PQBblockScan
    - 数据传递通过leader的上的队列




```
-> Aggregate: avg(`avg(a)`)
  -> Parallel scan on <temporary>
      -> Aggregate:
          -> PQblock scan on t1 using PRIMARY  (cost=4346.95 rows=43067)
```

性能提升：InnoDB Parallel Query

1. 充分利用多核优势，加快语句执行效率
2. TPC-H测试中，平均提升15倍+，最高42倍
3. 子查询等部分特性暂不支持



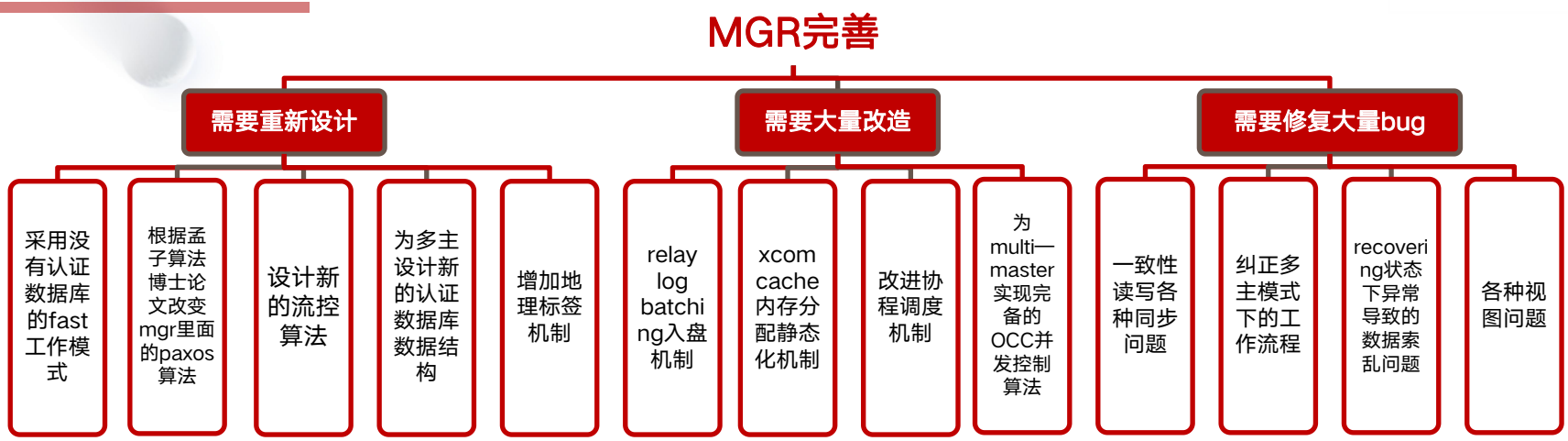
性能提升：Parallel Load data

- 原生load data存在缺陷
  - 单线程读取本地文件（或收取client传来的网络数据包），逐行获取内容并插入数据
  - 单个文件很大时，单线程处理模式无法充分利用数据库的资源，导致执行时间很长
  - 导入的数据在一个事务内，当binlog事务超过2G时，无法使用binlog在MGR集群间同步
- GreatSQL Parallel Load data
  - 自动将导入的文件切分文件成多个小块
  - 然后启动多个worker线程导入文件块
  - 与存储引擎无关，理论上可以支持所有的储引擎

性能提升：Parallel Load data

- 两种方式启用Parallel Load data
  - 设置session级变量启用
  - load data语句中增加hint启用
- 使用限制
  - 非原子性
  - 不分session变量支持受限，例如connection_id()
  - 不支持replace into
- 受限于master session的文件分割速度，并行导入速度可能区别较大。经过测试，在磁盘IO和CPU核心资源都充足的前提下启动32个worker，最大的加速比大概为20倍

稳定性提升——MGR



更快：快速探测异常情况；流控机制更精准；

更高：高并发下，吞吐持续且稳定；

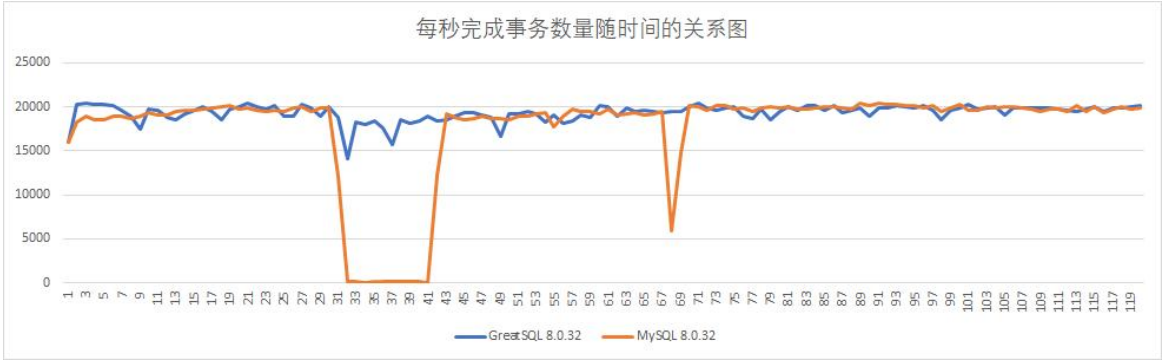
更强：更强的鲁棒性、更多的功能性。

稳定性提升——MGR

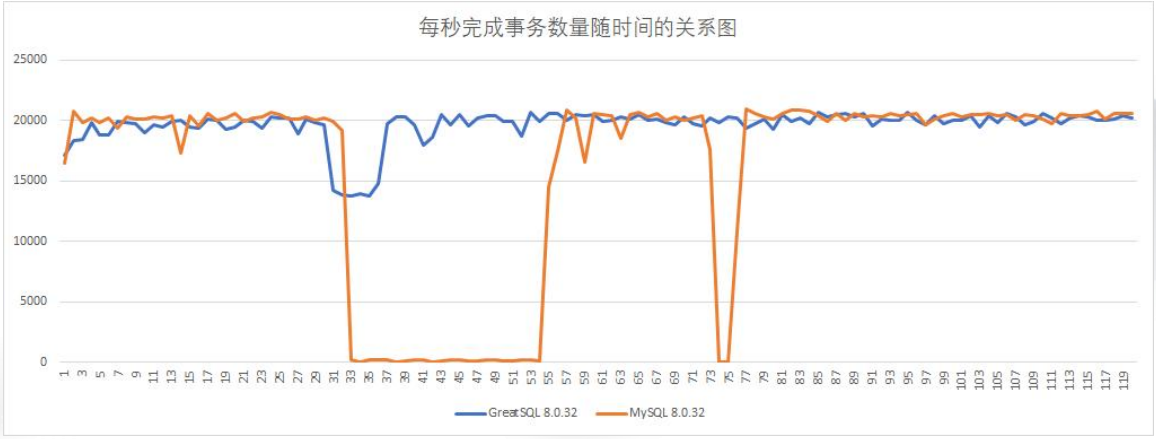

每秒订单数随时间关系图
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 64 67 70 73 76 79 82 85 88 91 94 97 100 103 106 109 112 115 118 121
流控
流控
认证数据库在垃圾回收
认证数据库在垃圾回收
改进后的版本
官方版本

稳定性提升——MGR

加入节点



杀节点



稳定性提升——MGR

地理标签

- 解决多IDC数据同步的问题
- 事务提交时，每个IDC中至少一个节点确认事务
- 每个IDC中至少一个节点有最新事务
- group_replication_zone_id
  - 0 ~ 8

快速单主模式

- 不使用原来的事务认证模式，只需在本地认证
- 降低内存消耗，提升高并发时的MGR性能
- 特别适合单主模式且跨IDC部署场景
- group_replication_single_primary_fast_mode
  - 0
  - 1

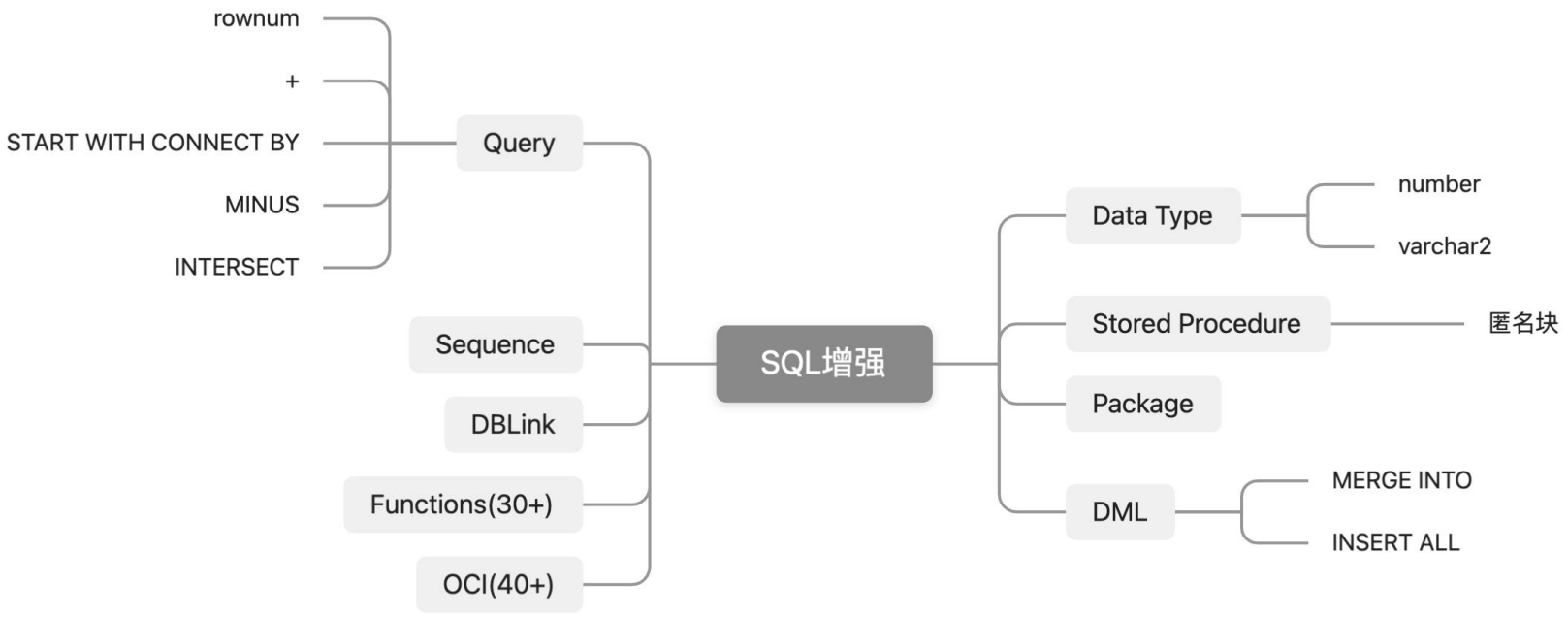
仲裁节点

- 不存储用户数据
- 没有binlog，也不需要回放relay log
- 只参与MGR状态投票/仲裁
- 系统负载非常低，可以在一个服务器上部署多实例
- group_replication_arbitrator

智能选主

- 优化选主策略
- 避免可能丢失数据风险
- group_replication_primary_election_mode
  - WEIGHT_ONLY
  - GTID_FIRST
  - WEIGHT_FIRST

易用性提升


rownum
+
START WITH CONNECT BY
MINUS
INTERSECT
Query
Sequence
DBLink
Functions(30+)
OCI(40+)
SQL增强
Data Type
number
varchar2
Stored Procedure
匿名块
Package
DML
MERGE INTO
INSERT ALL

Features Improved

&

Bugs fixed

Features Improved

- **优化了加入节点时可能导致性能剧烈抖动的问题**
- **解决节点异常退出集群时导致性能抖动的问题**
  - MySQL 中，paxos通信机制较为粗糙，当节点异常退出时，会造成较长时间（约20~30秒）的性能抖动，最差时TPS可能有好几秒都降为0
  - GreatSQL中优化后只会产生约1~3秒的性能小抖动，最差时TPS可能只损失约20% ~ 30%
- **解决磁盘空间爆满时导致MGR集群阻塞的问题**
- **解决了长事务造成无法选主的问题**
- 完善MGR中的外键约束机制，降低或避免从节点报错退出MGR的风险
- 解决多主模式下或切主时可能导致丢数据的问题

Features Improved

- 优化事务认证队列清理算法，规避每60s抖动问题
- 修复了recover过程中长时间等待的问题
- 修复了传输大数据可能导致逻辑判断死循环问题
- 节点异常状态判断更完善

Bug Fixed

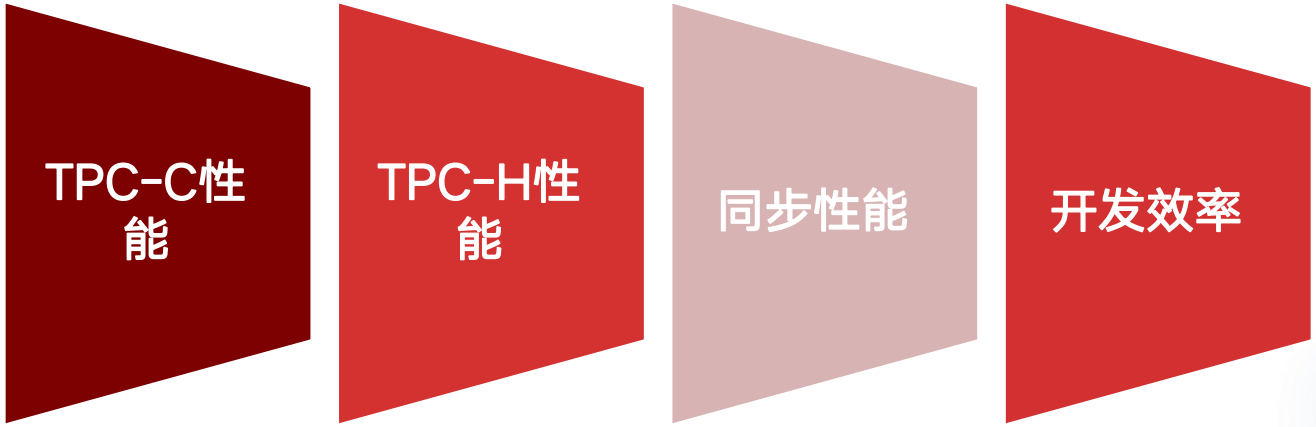
- 修复了InnoDB并行查询crash的问题
- 修复了协程调度不合理可能会造成在大事务时系统错误判断为网络错误的问题
- 修复了新加入节点在追数据时，由于超时导致连接提前关闭的问题
- 修复了recovering节点被中途停止导致的数据异常问题
- 修复了将传统主从环境下产生的binlog导入MGR可能引起死循环的问题
- 修复了多个可能导致MGR视图异常的问题
- 修复了多个可能导致MGR异常崩溃的问题

Others

特性	GreatSQL 8.0.25-16	MySQL 8.0.25 社区版
投票节点/仲裁节点	✅	❎
快速单主模式	✅	❎
地理标签	✅	❎
全新流控算法	✅	❎
InnoDB并行查询优化	✅	❎
线程池（Thread Pool）	✅	❎
审计	✅	❎
InnoDB事务锁优化	✅	❎
SEQUENCE_TABLE(N)函数	✅	❎
InnoDB表损坏异常处理	✅	❎
强制只能使用InnoDB引擎表	✅	❎
杀掉空闲事务，避免长时间锁等待	✅	❎
Data Masking（数据脱敏/打码）	✅	❎
InnoDB碎片页统计增强	✅	❎
支持MyRocks引擎	✅	❎

特性	GreatSQL 8.0.25-16	MySQL 8.0.25 社区版
InnoDB I/O性能提升	★★★★★	★★
网络分区异常应对	★★★★★	★
完善节点异常退出处理	★★★★★	★
一致性读性能	★★★★★	★
提升MGR吞吐量	★★★★★	★
统计信息增强	★★★★★	★
slow log增强	★★★★★	★
大事务处理	★★★★	★
修复多写模式下可能丢数据风险	★★★★★	/
修复单主模式下切主丢数据风险	★★★★★	/
MGR集群启动效率提升	★★★★★	/
集群节点磁盘满处理	★★★★★	/
修复TCP self-connect问题	★★★★★	/
PROCESSLIST增强	★★★★★	/

GreatSQL展望



GreatSQL展望

- 性能提升
  - AP性能提升
  - TP性能提升
  - 物理备份工具
- 稳定性提升
  - 双机高可用方案
  - 三副本高可用方案

- 其他增强
  - InnoDB表空间加密支持国密
  - 审计日志入库
  - 逻辑备份加密

GreatSQL，更流畅

全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE