



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

01

GreatSQL简介

02

GreatSQL特性

GreatSQL简介

GreatSQL是什么?



由万里数据库主导的开源MySQL分支

全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

为什么要做GreatSQL?

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

Steinar H. Gunderson

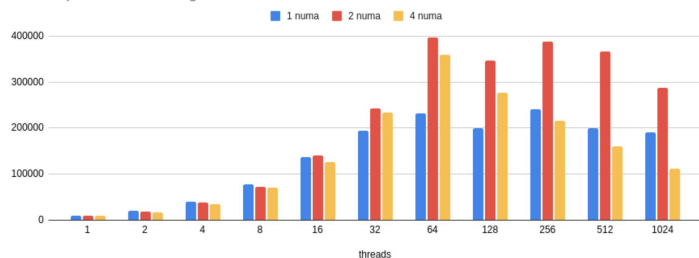
Sun, 05 Dec 2021 - Leaving MySQL

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 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

ARM - tpcc benchmarking 1/2/4 NUMA nodes



About

Products

For Teams

Search...

Home

How do I create a sequence in MySQL?

PUBLIC

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

Questions



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 of Ukraine and for the people of Ukraine, Oracle Corporation has already suspended all operations in the Russian Federation.

全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

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

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	2022-07-30 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 TemtableAggregatedIterator::Init
106629	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
99528	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



02 修复Replication

01 修复NDB Cluster
bug > 100

03 完成NDB Cluster 新功能
(WorkLog) > 10

04 完成Replication 新功能
考手册翻译

05 完成MySQL 5.1 中文参
考手册翻译

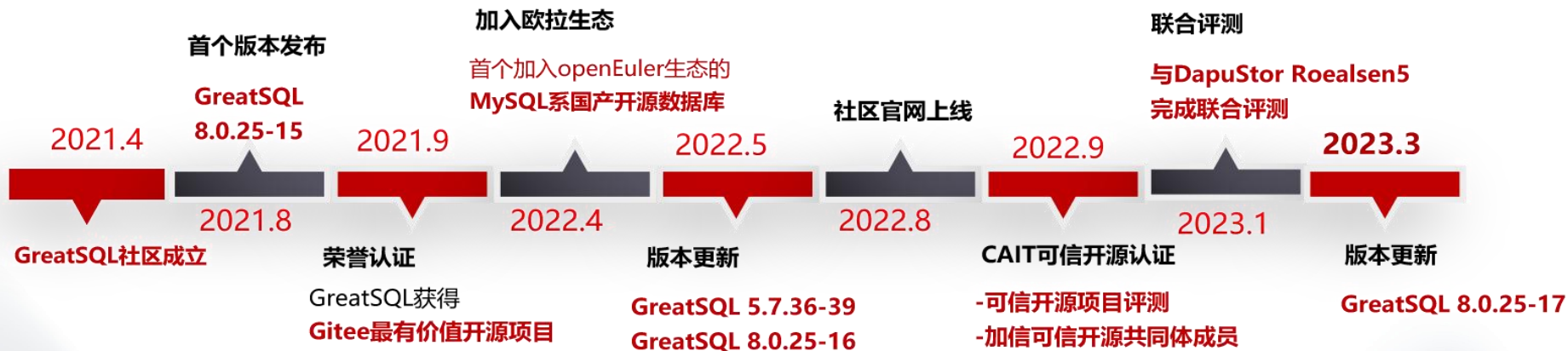
全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

GreatSQL社区

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



全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

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



- 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>

- **某大型股份制银行**：基于GreatSQL源码开发的数据库产品，已上线61个系统，包括：缴费平台、统一支付平台等。
- **某大型国有银行**：下属科技公司基于GreatSQL源码开发的数据库产品，已在该银行上线90个应用系统，包括：综合分账结算服务系统、进出口银行、银保监等外拓业务系统15个。
- **中信建投**：2022年加入GreatSQL社区，2023年2月向社区**首次贡献相关代码**。
- **国家电网**：基于GreatSQL技术开发的思极有容数据库（SG-RDB），累计支撑国网公司100余个业务系统建设，包括国网全业务数据中心、S6000等应用覆盖总部、27家网省公司。
- 恒生芸擎网络
- 福建福富
- 华润网络
- 作业帮

金融行业

能源行业

互联网行业

GreatSQL特性

产品目标



1. 高性能
2. 高可靠
3. 易用性
4. 高安全

全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

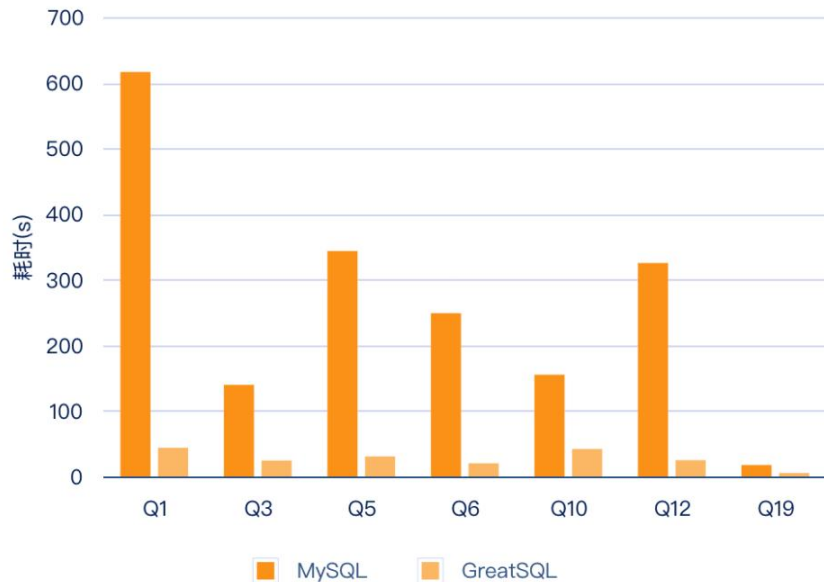
- Parallel Query

- 充分利用多核CPU
- 将单线程计划，转化为多线程并行计算
- TPC-H测试中，SQL执行效率平均提升**15倍**，最高提升**40+倍**；

- Parallel Load

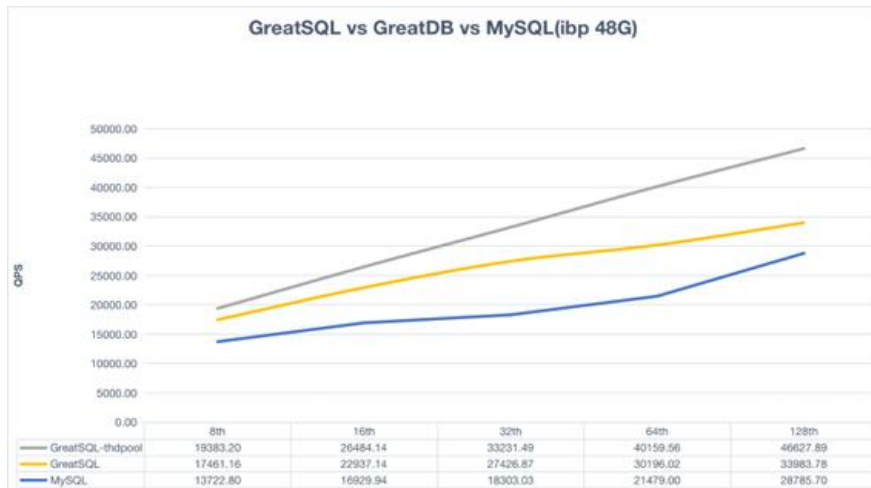
- 充分利用多核CPU
- 主线程解析数据文件
- 多个工作线程并发load数据
- 导入性能和工作线程数正比

TPC-H



- Thread pool
 - 降低了线程创建和销毁的代价
 - 保证高并发下，性能稳定不衰退
- 事务系统优化
 - 锁拆分，将单把大锁拆分为多个粒度更细的锁
 - 无锁化改造，通过实现机制改造，去掉readview获取时的锁保护

TP场景整体性能提升20%以上



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

读写节点绑定VIP



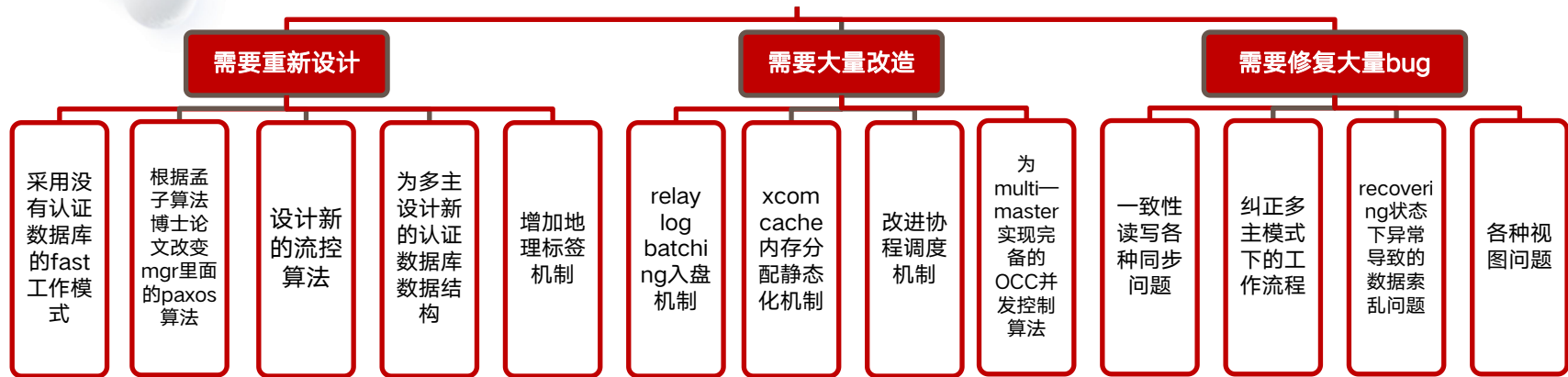
- ☐ 单主模式
- ☐ 读写节点自动绑定VIP
- ☐ 支持在线动态修改
- ☐ 高可用切换更便捷

全新流控算法



- ☐ 原生流控缺陷，导致明显抖动
- ☐ 增加主从延迟时间来计算流控阈值
- ☐ 同时考虑大事务处理因素
- ☐ group_replication_flow_control_rep_lag_lag_behind

MGR完善



更快: 快速探测异常情况; 流控机制更精准;

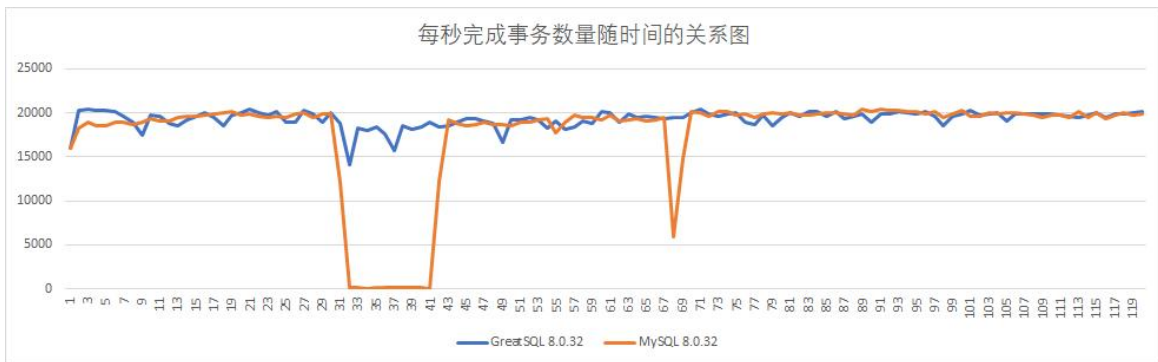
更高: 高并发下, 吞吐持续且稳定;

更强: 更强的鲁棒性、更多的功能性。

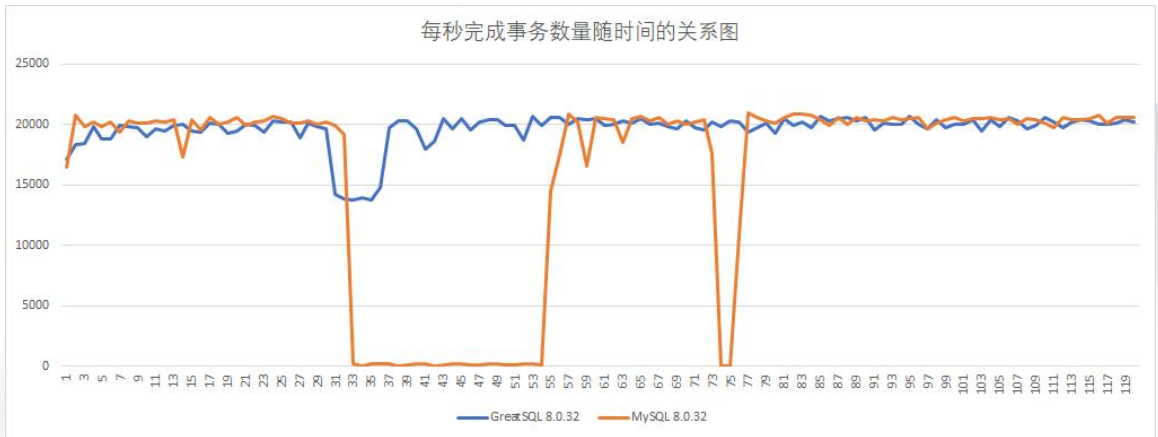
每秒订单数随时间关系图



加入节点



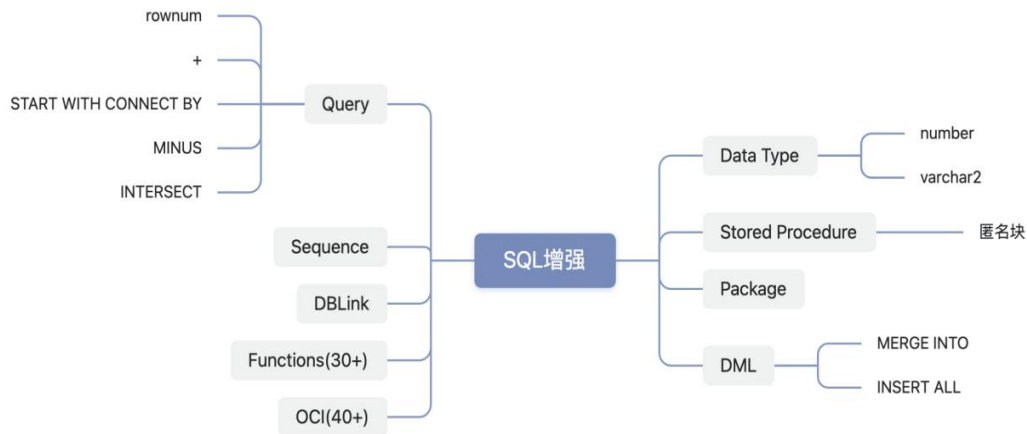
杀节点



- 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

- 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

易用性



安全性

1. 三权分立
2. 强制访问控制
3. 审计增强
4. 国密支持

交易型应用场景

中小型并发、中小型数据量、以联机事务处理为主的交易型应用,如运营商计费/CRM/服务开通、金融缴费/网银/贷款、能源风电监测/农电管理/电费结算等场景。

轻量数据分析场景

中小型数据量的联机分析应用,如周期性数据汇总报表之类的SAP、财务统计等场景。

Oracle替代场景

提供常见业务场景下的Oracle数据库替换。

MySQL替代场景

提供MySQL社区版本的完全覆盖,可无缝进行迁移替换。

1	特性	GreatSQL 8.0.32-24	MySQL 8.0.32
2	开源	✓	✓
3	ACID完整性	✓	✓
4	MVCC特性	✓	✓
5	支持行锁	✓	✓
6	Crash自动修复	✓	✓
7	表分区(Partitioning)	✓	✓
8	视图(Views)	✓	✓
9	子查询(Subqueries)	✓	✓
10	触发器(Triggers)	✓	✓
11	存储过程(Stored Procedures)	✓	✓
12	外键(Foreign Keys)	✓	✓
13	窗口函数(Window Functions)	✓	✓
14	通用表表达式CTE	✓	✓
15	地理信息(GIS)	✓	✓

1	特性	GreatSQL 8.0.32-24	MySQL 8.0.32
16	基于GTID的复制	✓	✓
17	组复制(MGR)	✓	✓
18	MyRocks引擎	✓	✗
19	SQL兼容扩展	1.数据类型扩展 2.SQL语法扩展 共超过20个扩展新特性	✗
20	MGR提升	1.地理标签 2.仲裁节点 3.读写节点绑定VIP 4.快速单主模式 5.智能选主机 6.全新流控算法	✗
21	性能提升	1.InnoDB并行查询 2.并行load data	✗
22	安全提升	1.国密支持 2.备份加密 3.审计日志入库	✗

TPC-C性能

TPC-H性能

同步性能

开发效率

全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE

- 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

GreatSQL, 更流畅



全球开源技术峰会

THE GLOBAL OPENSOURCE TECHNOLOGY CONFERENCE