

# GreatSQL构建金融级场景高可用架构

叶金荣  
2022.06.29



1. About GreatSQL
2. 金融级场景特点
3. GreatSQL优势特性
4. GreatSQL高可用架构

# About GreatSQL

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。

# About GreatSQL



加入欧拉生态

首个版本发布

GreatSQL

8.0.25-15

首个加入openEuler生态的  
MySQL系国产开源数据库

版本更新

GreatSQL

8.0.25-16

更多...

2021.4

2021.9

2022.4

2021.8

2022.4

2022.5

GreatSQL社区成立

荣誉认证

GreatSQL获得  
Gitee最有价值开源项目

大版本发布

发布5.7系列大版本  
GreatSQL  
5.7.36-39

- 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

- 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>

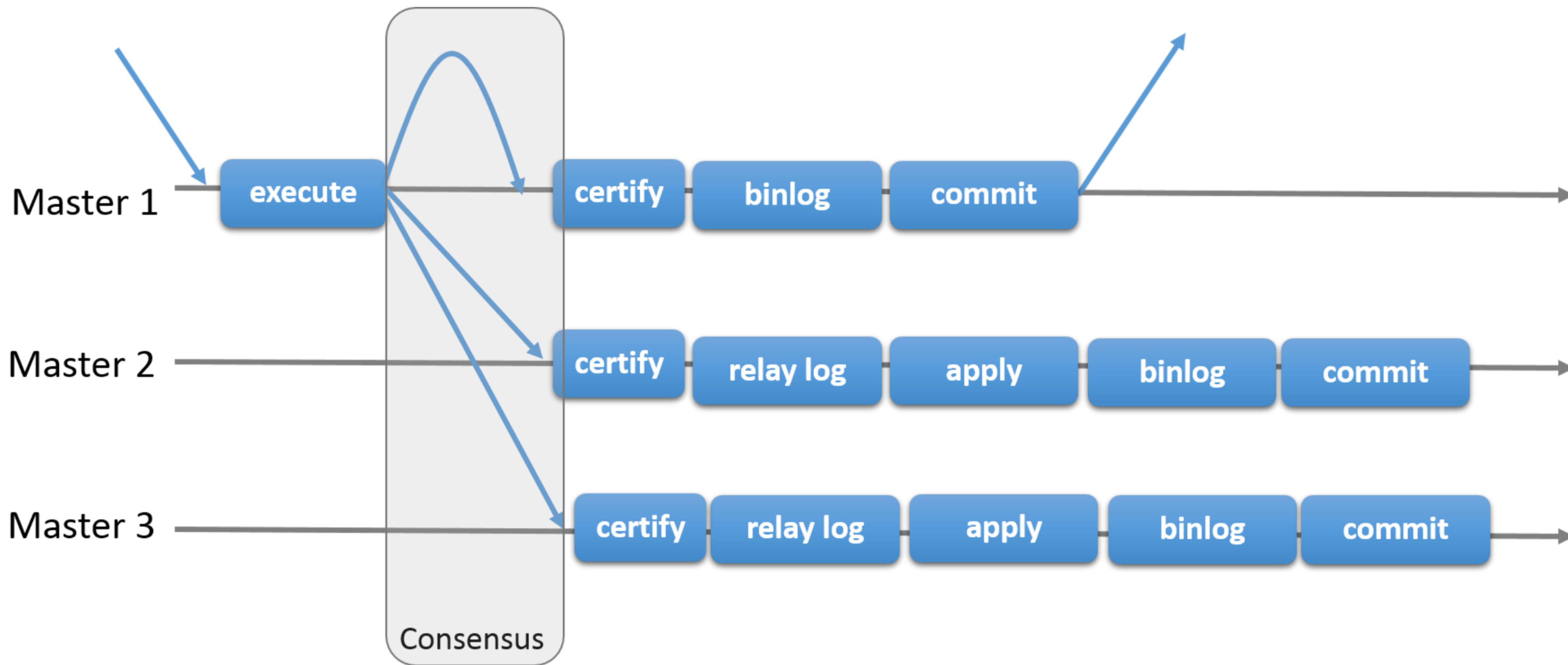
# 金融级场景特点

- 服务稳定性
- 数据一致性
- 业务连续性
- 尽可能实现
  - RTO=0
  - RPO=0



- MySQL Group Replication, 简称MGR/GR, 组复制
- shared-nothing
- 数据一致性及高可用集群方案
- 自带高可用, 自动故障检测、自动切换
- 支持单主、多主模式
- 可弹性扩展, 快速新增和移除节点

# About MGR



# 为什么说传统主从复制不适合金融场景



- 传统异步复制非常容易产生延迟
- 增强半同步虽有所改进，但从节点响应延迟较大时，主节点也会跟着受到影响
- 无论异步复制还是增强半同步复制，都存在众多因素会导致数据不一致
- 在MGR之前，并没有官方提供的配套高可用方案
- 主从复制在切换过程中，通常要进行辅助的数据校验及数据补差等工作
- 要持续不断校验数据，不适合数据高一致性业务场景

# 为什么说PXC不适用于金融场景



- Group Communication System
- Binlogs & Gcache
- Node Provisioning
- Partition
- Flow Control
- Cross platform
- DDL



# 为什么说MGR更适用于金融场景



- MGR采用多数派投票方案，个别节点出现异常时，不影响整体服务稳定性
- 更灵活的流控机制，避免木桶短板效应
- 全局事务冲突校验保障事务数据一致性
- 可根据业务特征自定义一致性级别
- 自动选主&切换，故障自动检测保障服务的持续性
- 在线添加、删除节点更便捷
- 原生自带MySQL Shell、Router等组件，应用、运维都更灵活方便

# GreatSQL优势特性

# 目录

1. 地理标签
2. 仲裁节点
3. 快速单主
4. 智能选主
5. 并行查询

- 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



- 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
  - 0
  - 1

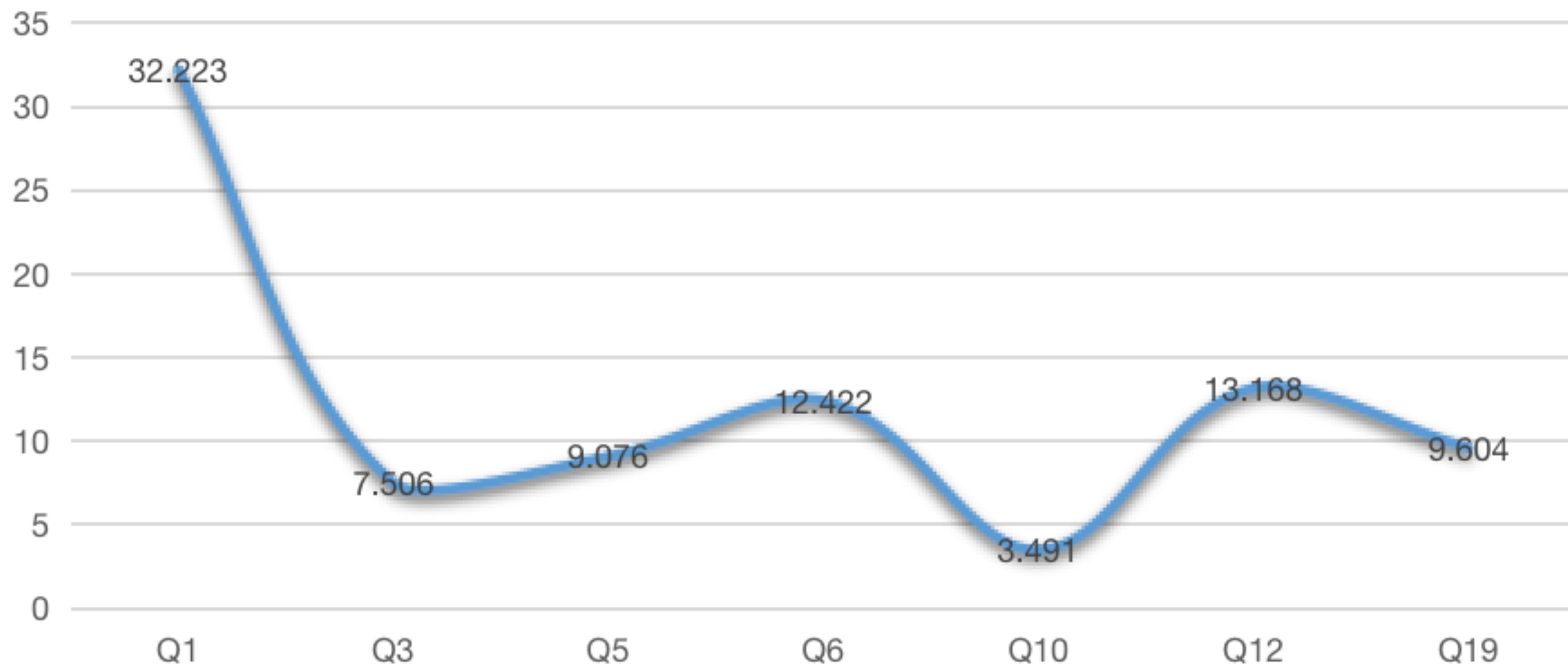
- 不使用原来的事务认证模式，只需在本地认证
- 降低内存消耗，提升高并发时的MGR性能
- 特别适合单主模式且跨IDC部署场景
- `group_replication_single_primary_fast_mode`
  - 0
  - 1

- 原来的选主策略中没有判断各节点最新事务状态
- 可能会导致丢失部分事务数据
- group\_replication\_primary\_election\_mode
  - WEIGHT\_ONLY
  - GTID\_FIRST
  - WEIGHT\_FIRST

- 对B+树多个子树并行扫描后再聚合，提升查询效率
- TPC-H测试中，最高可提升30倍，平均提升15倍
- 特别适合汇总报表之类的SAP、财务统计等业务
- force\_parallel\_execute
  - ON
  - OFF



## InnoDB并行查询性能提升



# Features Improved & Bugs fixed

- 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



# CPU、内存消耗更低



```
# top
top - 08:50:11 up 192 days, 18:23, 1 user, load average: 0.04, 0.07, 0.12
Tasks: 133 total, 1 running, 131 sleeping, 1 stopped, 0 zombie
%Cpu(s): 1.8 us, 0.9 sy, 0.0 ni, 97.0 id, 0.2 wa, 0.0 hi, 0.1 si, 0.0 st
KiB Mem : 16266176 total, 167328 free, 7059720 used, 9039128 buff/cache
KiB Swap: 0 total, 0 free, 0 used. 8858872 avail Mem
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
24102	mysql	20	0	6610980	3.3g	10036	S	5.3	21.3	406:03.23	mysqld
9139	mysql	20	0	6796472	2.7g	11192	S	1.3	17.5	1138:50	mysqld
4819	root	20	0	1070316	82216	13792	S	1.0	0.5	140:50.06	YDService
10068	root	20	0	1219652	27720	5072	S	0.0	0.2	145:05.22	dockerd-current
398	root	20	0	55828	17852	17372	S	0.0	0.1	17:06.42	systemd-journal

... 同样是在空跑状态下，MySQL 相比 GreatSQL 消耗的 CPU 和内存都更高

```
# ps -eo pid,command | grep -v grep | grep mysqld
9139 /usr/local/GreatSQL-8.0.25-15-Linux-glibc2.17-x86_64/bin/mysqld
24102 /usr/sbin/mysqld
```

- 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

# GreatSQL优势特性



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	✓	✗



# GreatSQL优势特性

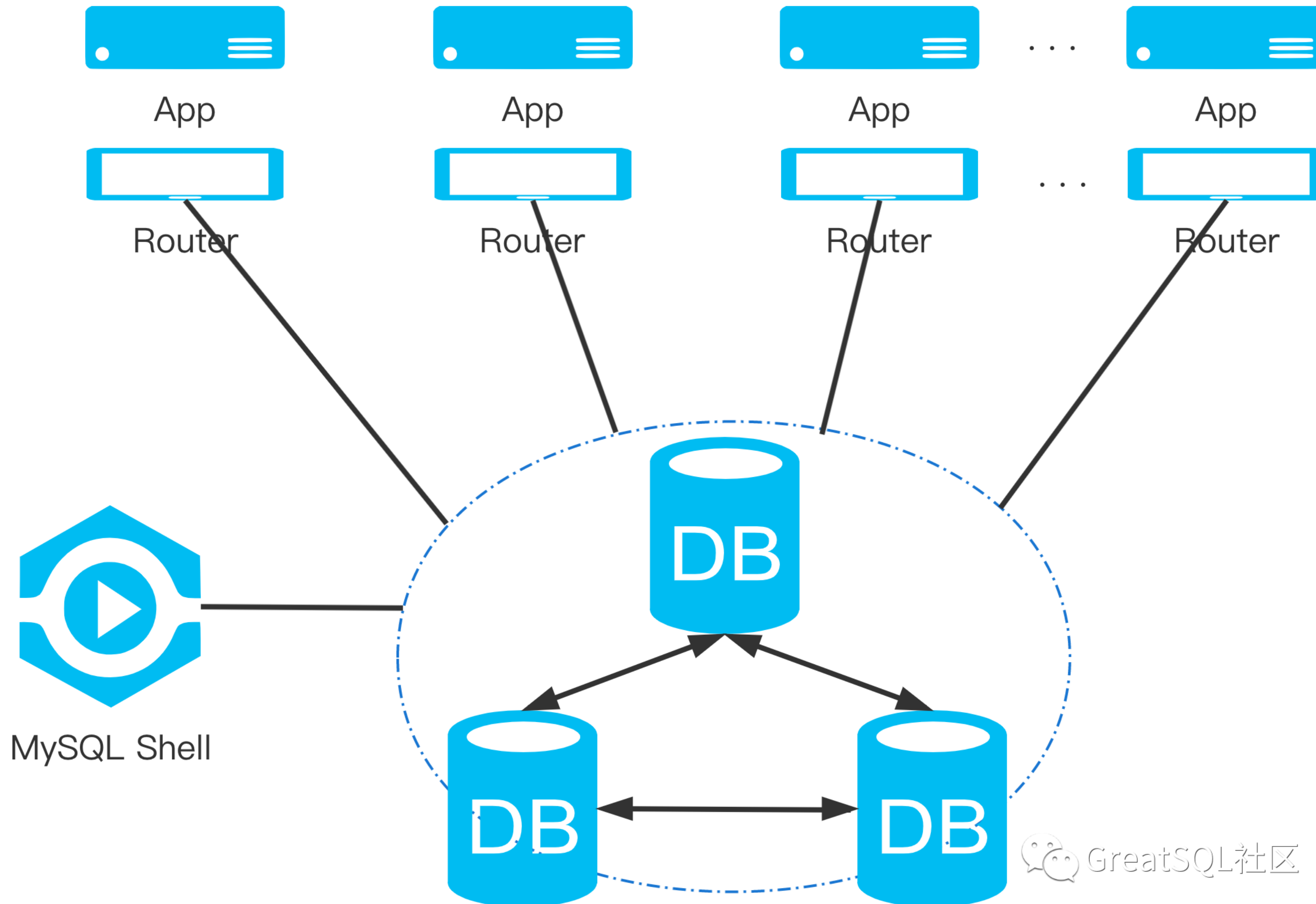


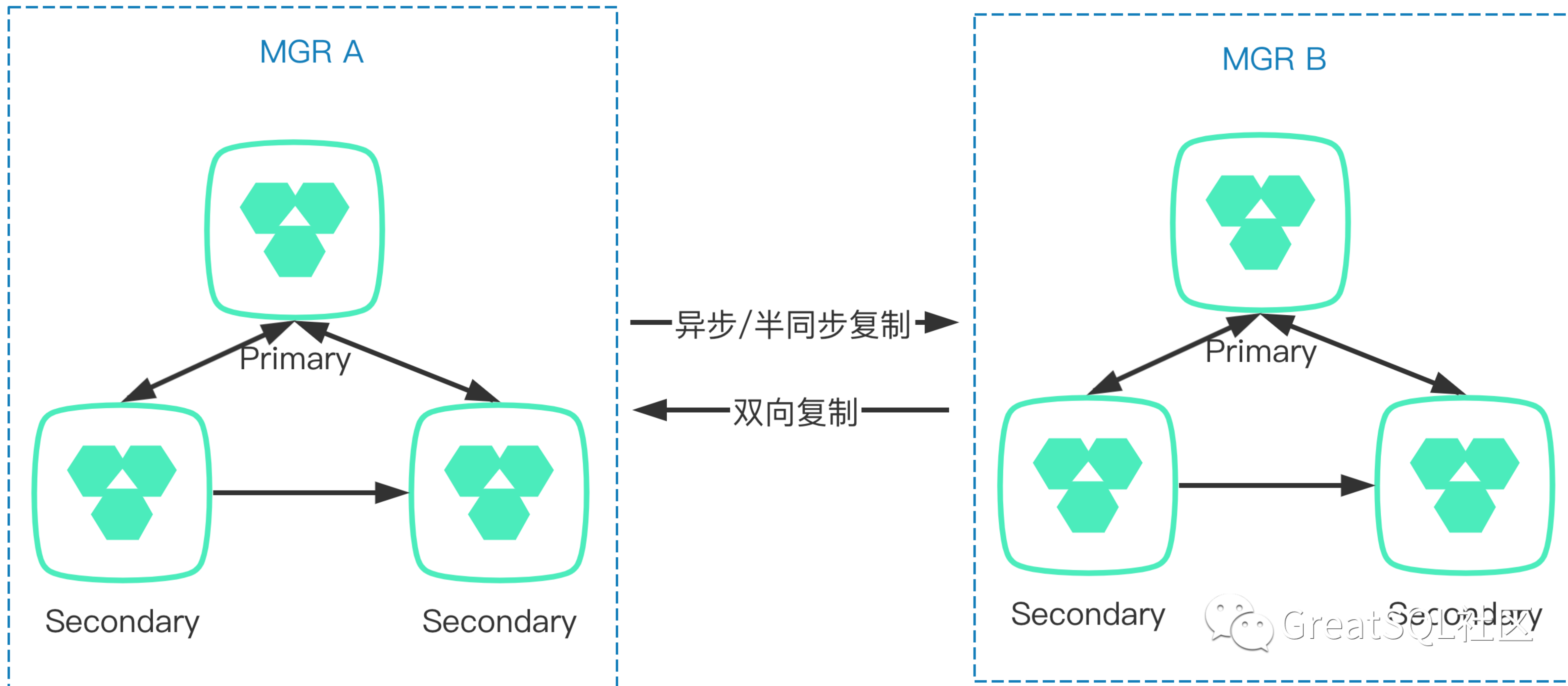
特性	GreatSQL 8.0.25-16	MySQL 8.0.25 社区版
InnoDB I/O性能提升	★★★★★	★★
网络分区异常应对	★★★★★	★
完善节点异常退出处理	★★★★★	★
一致性读性能	★★★★★	★
提升MGR吞吐量	★★★★★	★
统计信息增强	★★★★★	★
slow log增强	★★★★★	★
大事务处理	★★★★	★
修复多写模式下可能丢数据风险	★★★★★	/
修复单主模式下切主丢数据风险	★★★★★	/
MGR集群启动效率提升	★★★★★	/
集群节点磁盘满处理	★★★★★	/
修复TCP self-connect问题	★★★★★	/
PROCESSLIST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

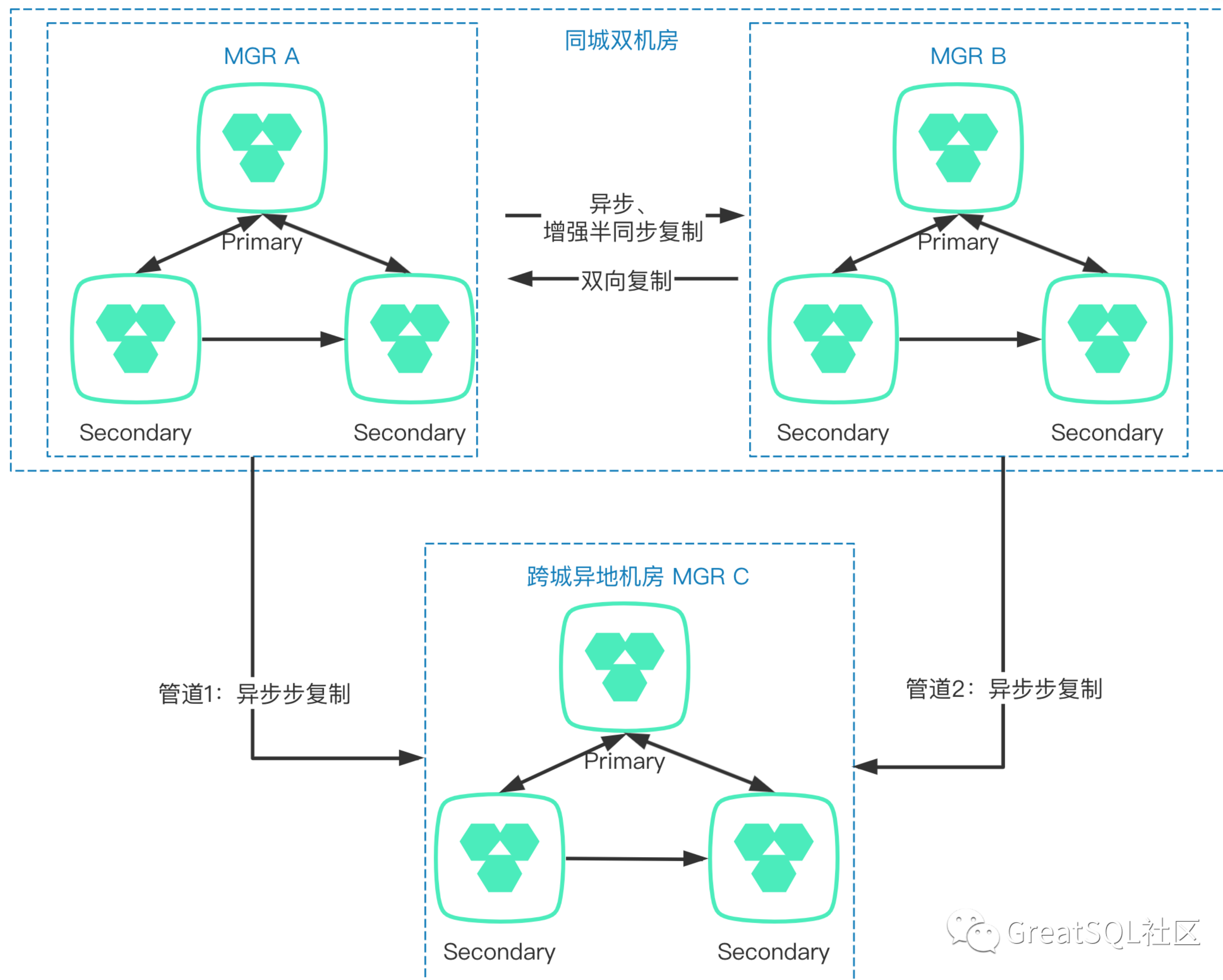
# GreatSQL高可用架构



# GreatSQL高可用架构







# GreatSQL展望

- 更多企业级特性，更适用于金融级应用场景
- 持续优化Paxos机制，提升稳定性及性能
- 优化MGR中的relay log，缓解本地节点事务应用的瓶颈
- XCom cache静态化，提升MGR通信性能
- 大事务支持，应对更多业务场景
- 解决多主模式下DDL & DML之间的冲突问题



- GreatSQL是由万里数据库维护的MySQL分支
- 专注于提升MGR可靠性及性能，支持InnoDB并行查询特性
- 是适用于金融级应用的MySQL分支版本
- 地理标签，提升多机房架构数据可靠性
- 仲裁节点，用更低的服务器成本实现更高可用
- 单主模式下更快，选主机制更完善
- InnoDB表也支持并行查询，让CPU资源不再浪费
- 全新流控机制，让MGR运行更流畅不频繁抖动
- 相对官方社区版，MGR运行更稳定、可靠
- 其他...

不忘DB初心，牢记万里使命

# GreatSQL，更流畅

