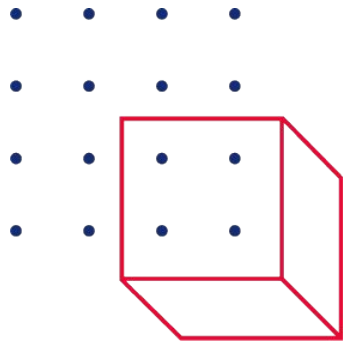
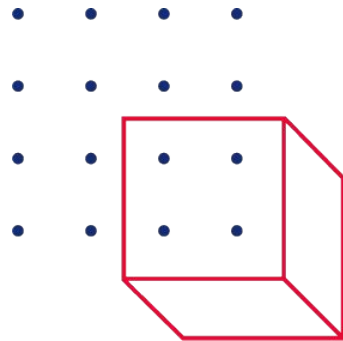


TiDB 5.0 预览

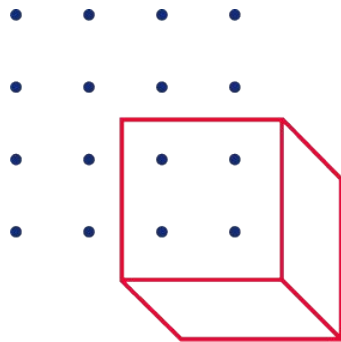


TiDB 5.0 总览



TiDB 5.0 概况

- 预计 2021 年 3 月 GA 发布
- TiDB Cloud 公测
- 更强的 OLTP 金融级核心场景能力
 - 更快
 - 更稳定
- 更完整的 HTAP 架构
 - MPP 架构
 - 更好的大数据生态集成
- 更简易的运维



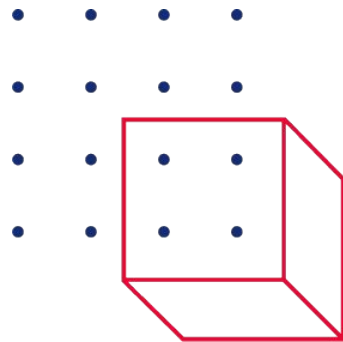
新的火车发版模型

- 火车发版模型
 - 向小步迭代渐进发版靠拢
 - 大型特性也将拆散在小版本逐步 发布
 - 更可控的发版周期
- 为何我需要关心？
 - 持续发版信息
 - 没有发现属于你的惊喜，不代表没有

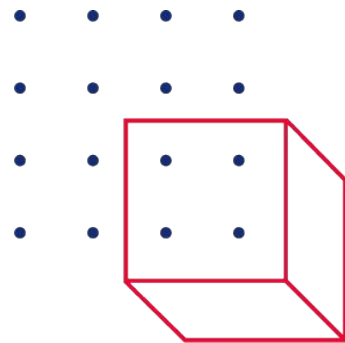
这不是总结，这只是一个开始



OLTP 核心场景



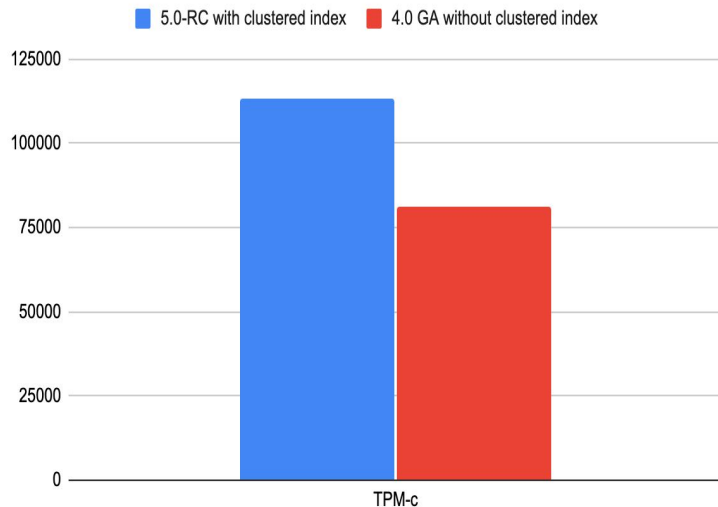
更快



聚簇索引

- 5.0 之前, 表数据以表数据 + 主键索引分离的形式存储
 - 行写入需要同时写入数据和索引两部分
 - 根据主键读取行需要先读取索引再查找数据本身
- 5.0 开始, 表的行数据可以完全以主键序存储
 - 写入省去一次索引插入
 - 读取省去一次数据查找
 - **读写均可更快完成**

5.0-RC vs 4.0 GA

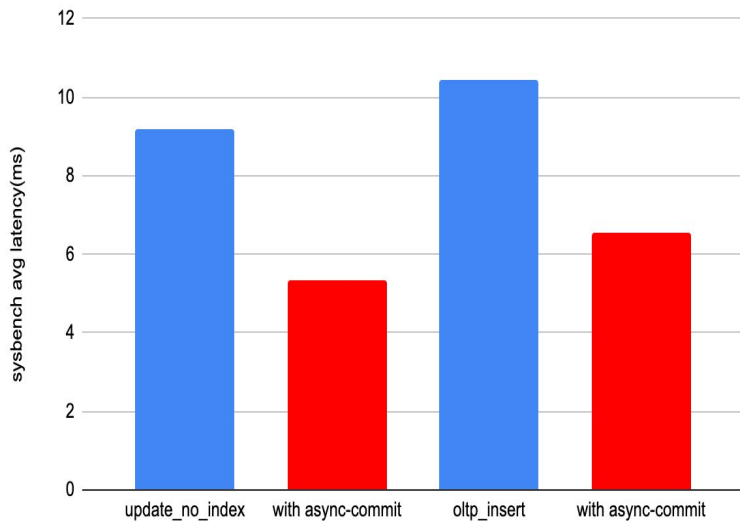


TPC-C tpmC 提升 39.3%

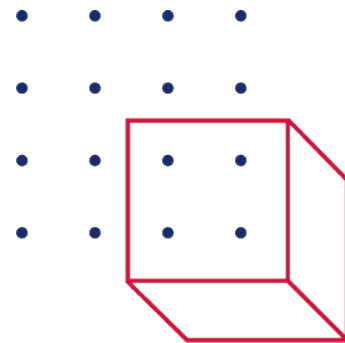
异步提交

- 5.0 之前, 事务需要经历完整二阶段提交才能返回
 - 2 阶段提交 = Prewrite + Commit
 - 两阶段叠加带来高延迟
- 5.0 开始, 事务的提交阶段可被隐藏
 - 锁信息合并到主锁, 维持原有的原子性
 - Prewrite 完成即可返回
 - Commit 阶段以异步方式进行
 - **降低写入延迟近一倍**

sysbench latency



更稳定

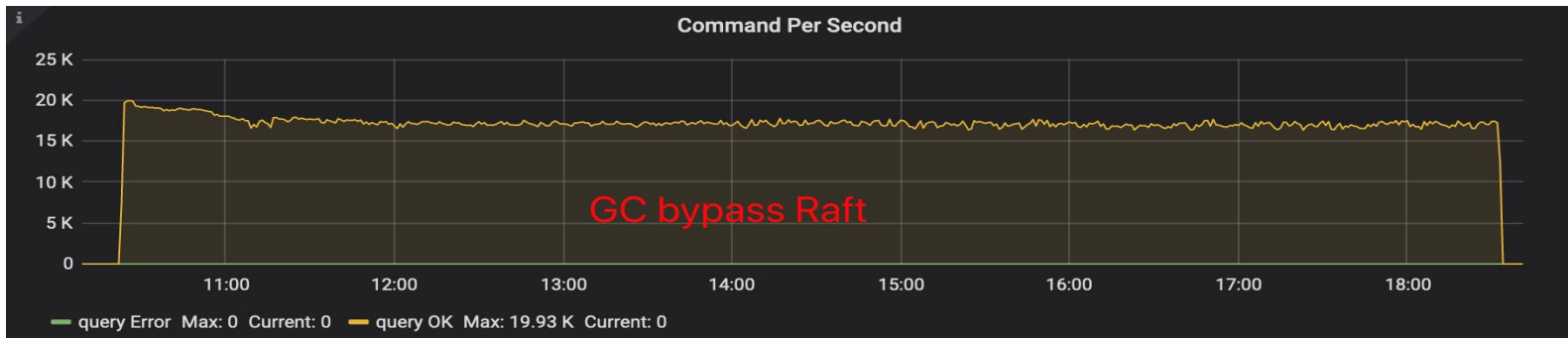
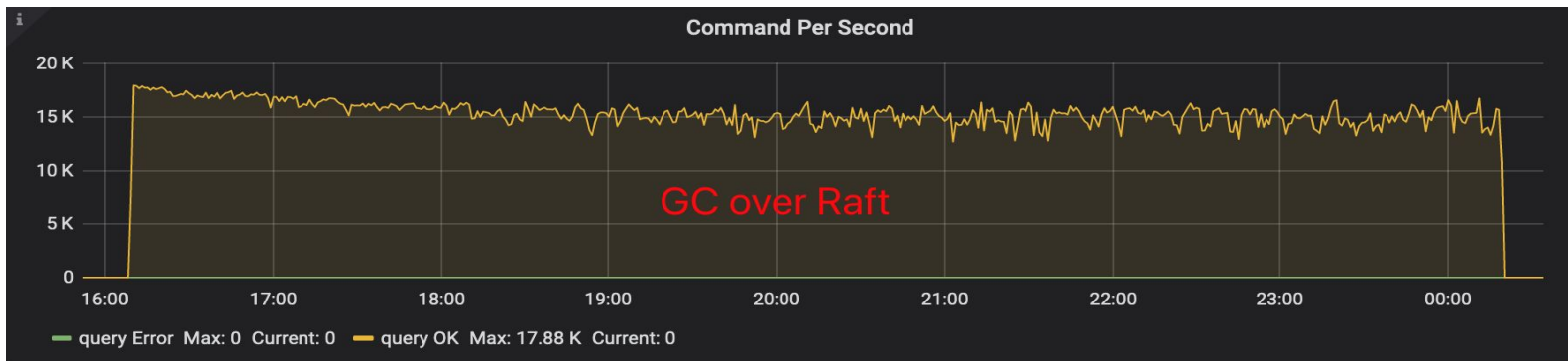


更稳定的 GC / Compaction 过程

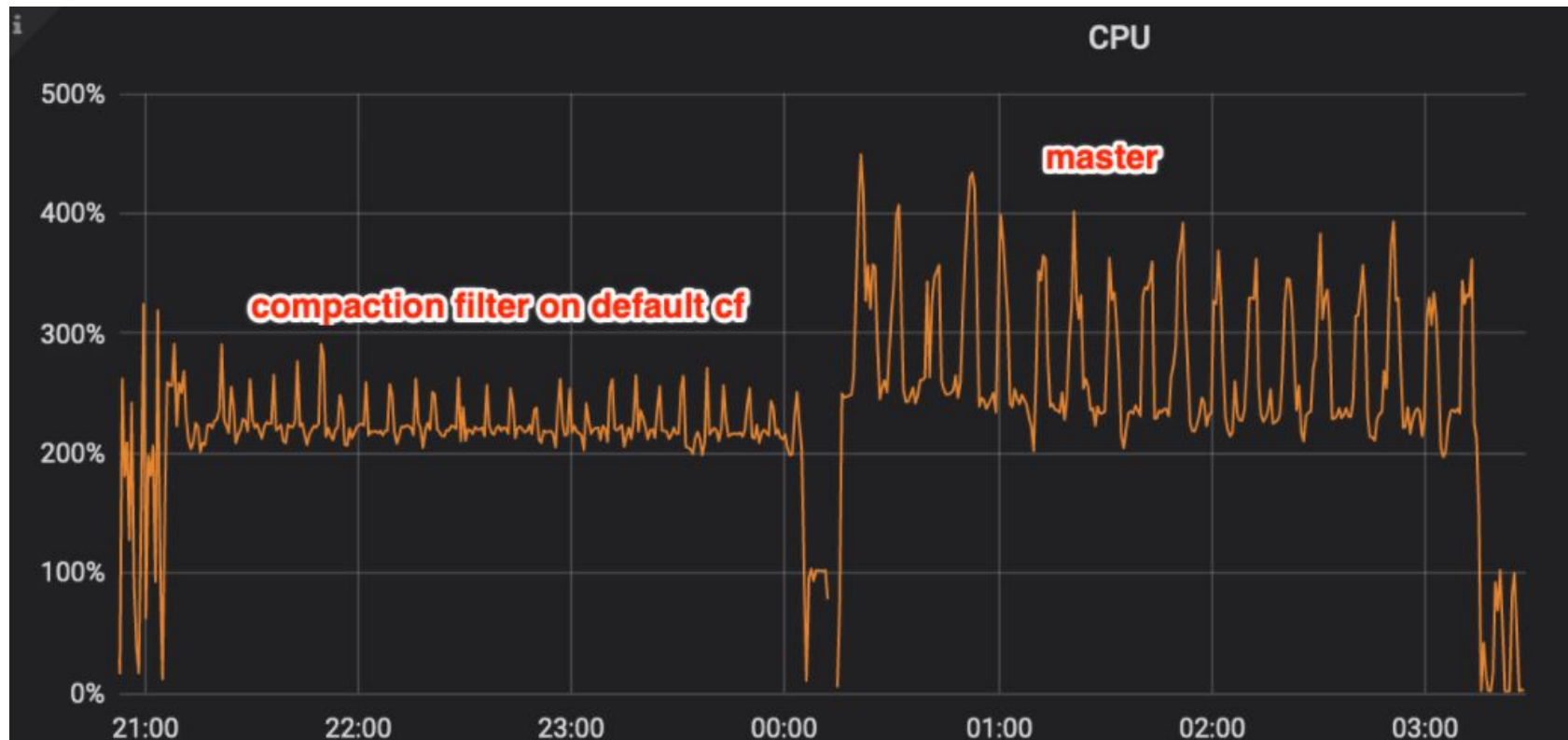


- 5.0 之前, GC / Compaction 容易引入性能抖动
 - 都会大量读写数据, 引入抖动
- 5.0 开始, GC 和 Compaction 被全面优化
 - GC 可与 Compaction 合并完成
 - LSM Compaction 同时过滤过期数据
 - 直接忽略过期数据不写入目标 SST 文件
 - GC 可绕过 Raft 复制
 - Compaction 个多环节引入自动限流机制
 - **更低性能抖动**

GC 算法优化 - Bypass Raft



GC 算法优化 - Compaction 合并 GC



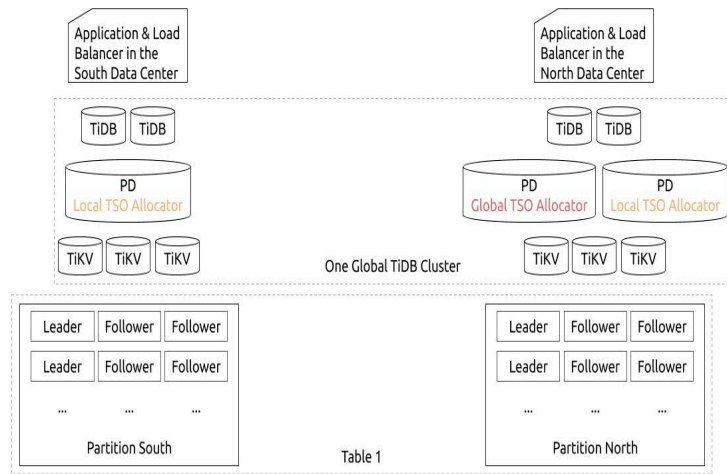
更稳定的执行计划和调度

- 增强的执行计划选择
 - 扩展的统计信息: multi-column NDV, multi-column order dependency, multi-column function dependency 等
 - 更多维度考虑数据相关性
 - SPM: 为 SQL Binding 增加 DML 语句支持
- 增强的 PD 调度算法
 - 更优的调度评分算法
 - 消除空 Region 带来的冗余调度



更稳定的跨区部署能力

- 增强跨域部署
 - 允许将分区数据向不同机房 调度
- 局部事务
 - 允许事务读取本地 PD 授时, **大幅降低跨机房延迟**
- 从副本提供 Stale Read 能力
 - 如果读取非最新数据可无视跨机房一致性校对
 - 允许根据新鲜度换取更短的读取延迟
 - **提供跨域从集群低延迟读取能力**
- 加强的跨中心调度可用性
 - Joint Consensus: 原子化 Peer 增减操作
 - 防止 Peer 变更瞬间中心不可用带来的故障



更好的内存管理

- 加强的 Spill 算法
 - 减少 OOM 可能性
 - Spill 数据 Checksum 和加密支持
- 降低统计信息占用内存
- TiFlash 加强内存控制
 - 引入 Delta 数据索引淘汰

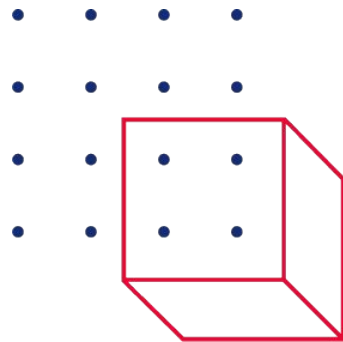


其他引擎特性与改进



- Coprocessor cache GA
- 改进分区表性能
- 支持 collation 'utf8mb4_unicode_ci' and 'utf8_unicode_ci'
- 集合算子 INTERSECT / EXCEPT 支持
- 支持 Invisible Index
- 表达式索引 GA

HTAP 能力

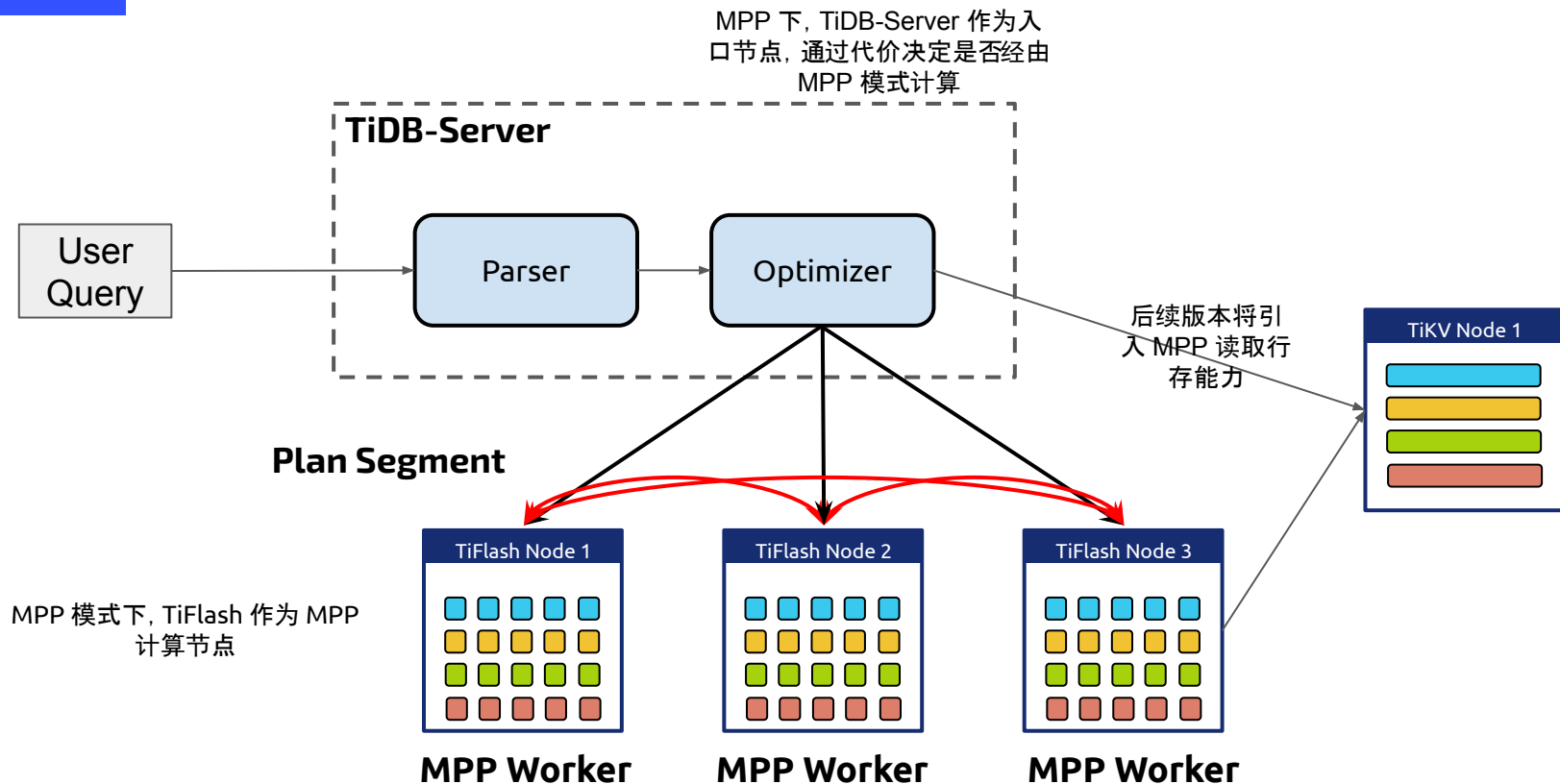


MPP 引擎



- 5.0 引入 MPP 架构, 提供与存储匹配的分布式计算引擎
 - 解决大数据量分析场景下 TiDB 无法扩展问题
- 与 TiDB-Server 共享 SQL 前端
 - 共享的 Parser 和优化器
 - 一体化的 TP 和 AP 入口
 - 根据优化器自动选择单机执行或 MPP 模式
- 高性能的交互式查询体验
 - 数倍于 Apache Spark 的执行速度

MPP 架构介绍

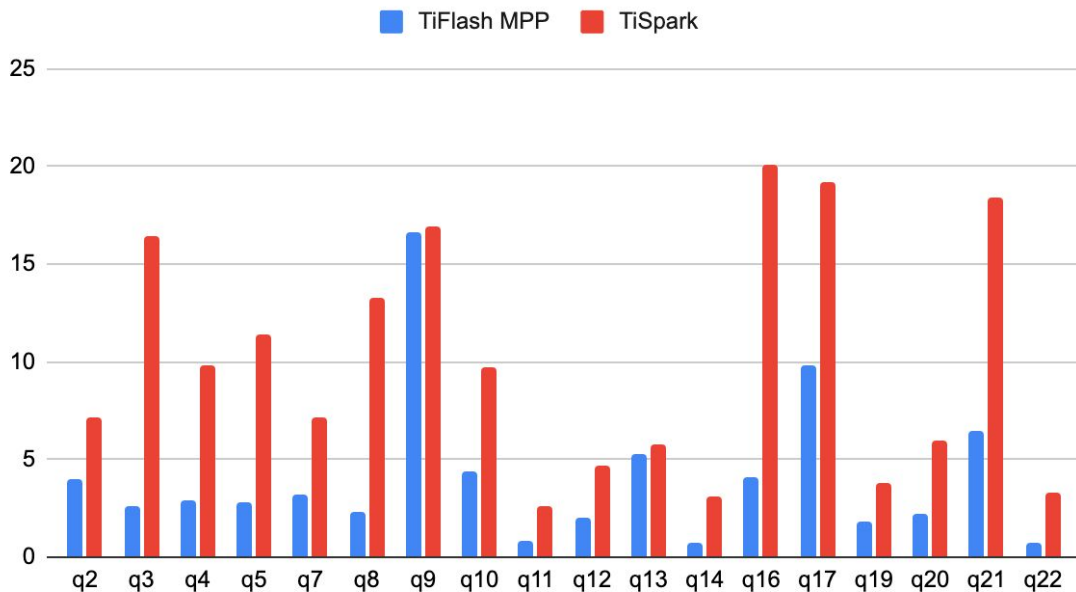


MPP 性能

这只是调优前的 Benchmark...

	TiFlash MPP	TiSpark
q2	3.97	7.151
q3	2.59	16.463
q4	2.86	9.8
q5	2.79	11.395
q7	3.25	7.128
q8	2.32	13.243
q9	16.62	16.949
q10	4.38	9.728
q11	0.83	2.655
q12	2.05	4.706
q13	5.23	5.76
q14	0.73	3.077
q16	4.05	20.08
q17	9.82	19.215
q19	1.81	3.814
q20	2.17	5.958
q21	6.5	18.424
q22	0.72	3.28

TPC-H 20 3 Nodes

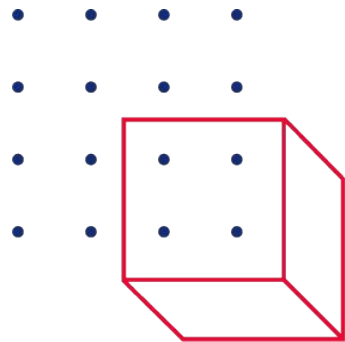


大数据生态集成

- 社区主导的 Apache Flink 集成
 - 由**知乎**主导的 Apache Flink 深度集成
 - 提供原生的 Apache Flink 读写能力
 - 向批流一体迈进
- TiBigdata Repo 提供除 TiSpark 外的大数据集成
 - 更多生态建设逐步完善中



更简易的运维



数据迁移



- 数据导入 / 导出
 - Dumping 支持向 S3 导出数据
 - Lightning 支持从 S3 导入 AWS Aurora 快照
 - Lightning 去除 tikv-importer 组件, 提供 all-in-one binary 简化使用
 - Lightning 数据导入速度提升 40%
- Backup & Restore
 - BR 支持基于 S3 / GCS 备份与恢复数据
 - BR 支持持续日志备份与恢复 (experimental)
- Change Data Capture
 - TiDB 成为 Confluent 认证 Kafka 合作伙伴
 - TiCDC 支持 TiDB 集群之间的环形复制 (experimental)

集群部署



- 增强的 TiUP Mirror 功能
 - 支持合并镜像
 - 支持向本地镜像发布组件
 - 为镜像管理增加身份体系
- 增强的标准环境部署
 - 新增提供标准参数模板功能
 - 支持 server 配置全局参数
 - 支持更全面的环境标准检查及一键参数配置
 - 支持将所有配置合并于 meta.yaml 以便修改

集群升级



- 增强集群升级
 - 新增 PD leader 状态的主动识别, 优化 PD leader 滚动升级带来的性能抖动
 - 支持断点重起功能
 - 新增快速离线升级方式
 - 新增升级前环境自动检查, 增加升级的成功率。

加强的诊断功能



- EXPLAIN ANALYZE 加强
 - 支持 UPDATE / INSERT
 - 支持 TiKV RPCs, Network 延迟, RocksDB tombstones, 锁信息等
 - 支持慢查询日志自动打印分析
- 支持观察运行中 SQL 状态

更详细的诊断信息

4.0 > explain analyze select * from t where a in (1,2,3);

id	execution info
Batch_Point_Get_1	time:861.607μs, loops:2

5.0 > explain analyze select * from t where a in (1,2,3);

id	execution info
Batch_Point_Get_1	time:3.61s, loops:2, BatchGet:{num_rpc:6, total_time:3.42ms}, ResolveLock:{num_rpc:1, total_time:3.61s}

自动诊断信息日志

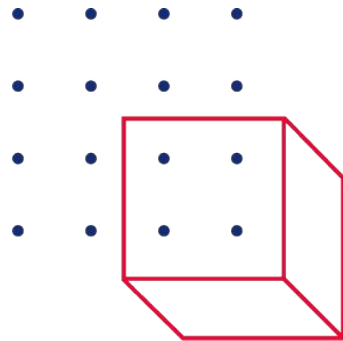
4.0 Slow Log

+-----+				
	plan			
+-----+				
	TableReader_7	root	196608	data:TableFullScan_6
	└TableScan_6	cop	196608	table:t, keep order:false
+-----+				

5.0 Slow Log

+-----+			
	plan		
+-----+			
	id	execution info
	Insert_1	time:393.3ms, loops:1, prepare:267ms, insert:126.3ms, ...
	└TableReader_7	time:216.2ms, loops:66, cop_task: {num: 1, max: 215.7ms, ...
	└└TableFullScan_6	tikv_task:{time:213ms, loops:69}, total_keys: 851969, ...
+-----+			

TiDB Cloud



TiDB Cloud 开放公测



- TiDB Cloud 将在 2021 上半年启动 Public Preview Program
 - AWS / GCP
 - 开放公开账号注册
 - 免费 2 天 Free Trial
 - 公测期间单价 -30% off
 - 抢先体验 TiDB 5.0 (GA 发布后)

这只是一个开始。

