
aws
INNOVATE
ONLINE CONFERENCE

aws INNOVATE
ONLINE CONFERENCE

分会场二：数据库

深入探索兼容 MySQL 的 Amazon Aurora

汪允璋，AWS 解决方案架构师

愿景: Amazon Relational Database Service

提供新的开源与商业数据库选项

云原生引擎

Amazon Aurora

开源引擎

PostgreSQL

MySQL

MariaDB

商业引擎

Microsoft SQL Server

ORACLE

关系数据库系统平台

- 自动故障转移
- 备份与恢复
- 多区域复制
- 隔离与安全
- 行业合规
- 自动补丁修复
- 高级监控
- 日常维护
- 简易规模伸缩

Amazon Aurora…

以开源级成本交付的企业级数据库

- ☑ 媲美高端商业数据库的**速度**与**可用性**
- ☑ 媲美开源数据库的**简单性**与**成本效益**
- ☑ 与 MySQL 及 PostgreSQL 全面 **兼容**
- ☑ 按**使用量计费**的简单定价模式

以**托管**服务形式交付

Amazon Aurora 客户采用情况

AWS 有史以来增长速度最快的服务

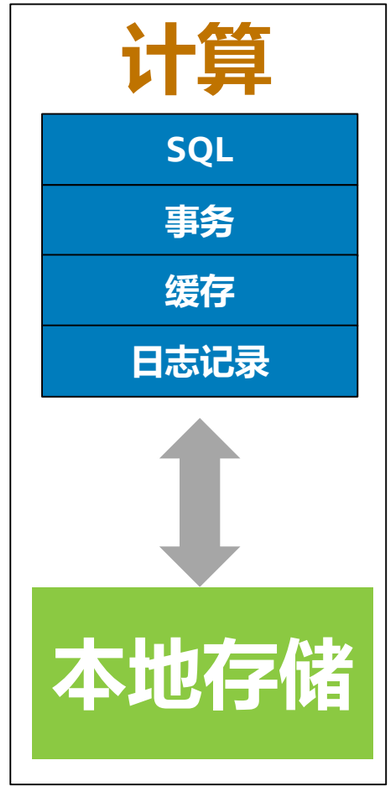
在前百大 AWS 客户当中，有四分之三选择使用 Amazon Aurora

性能

aws INNOVATE
ONLINE CONFERENCE

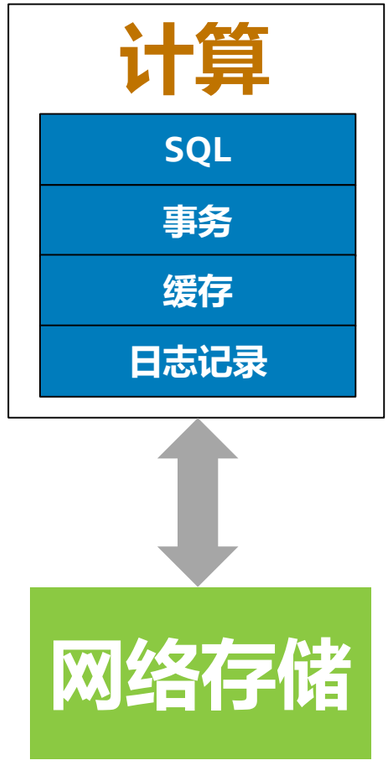
传统数据库架构

单个"box"中的整体式堆栈



传统数据库架构

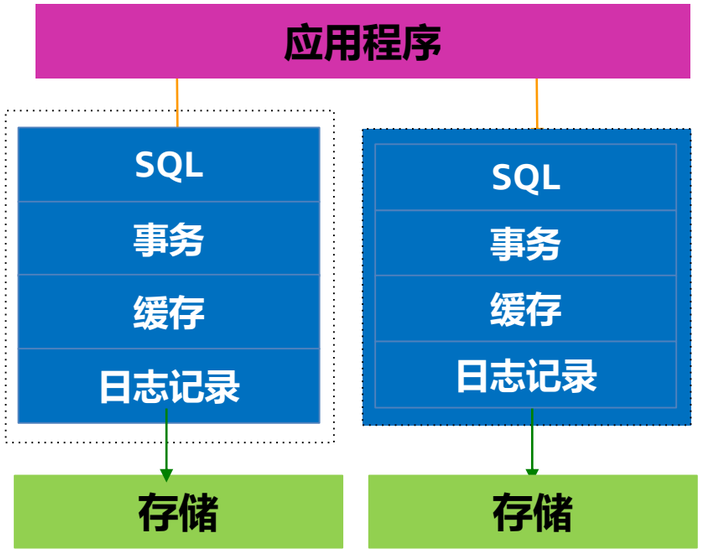
将存储与计算解耦开来



传统分布式数据库堆栈

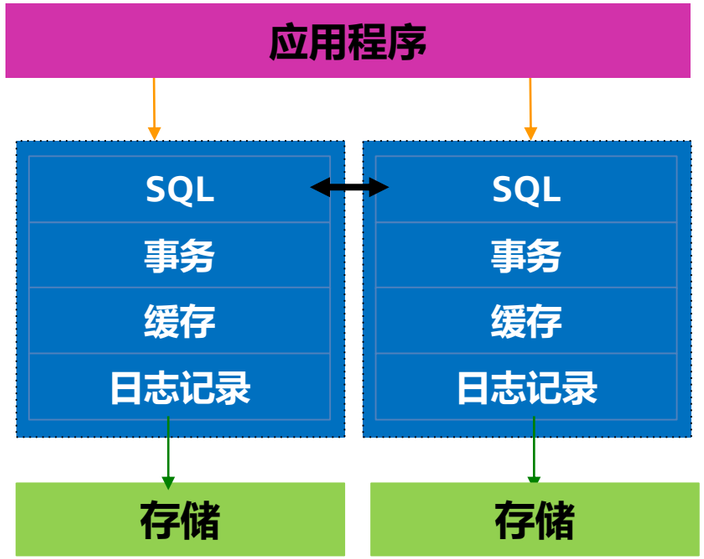
分片

应用层耦合



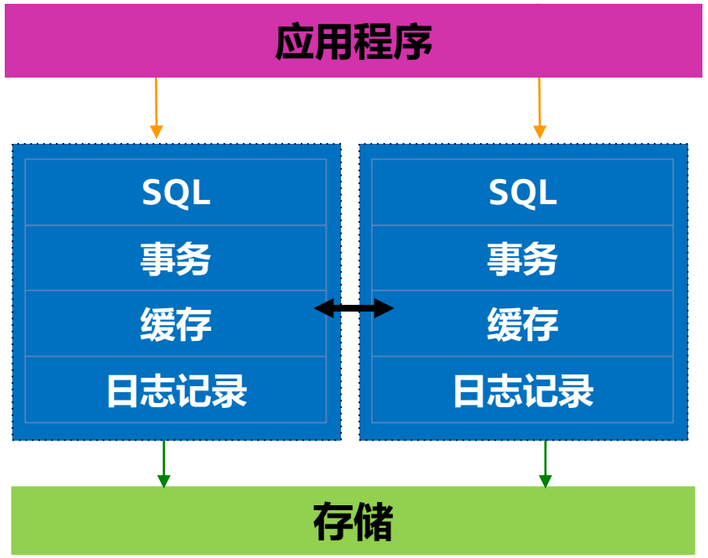
无共享

SQL 层耦合



共享磁盘

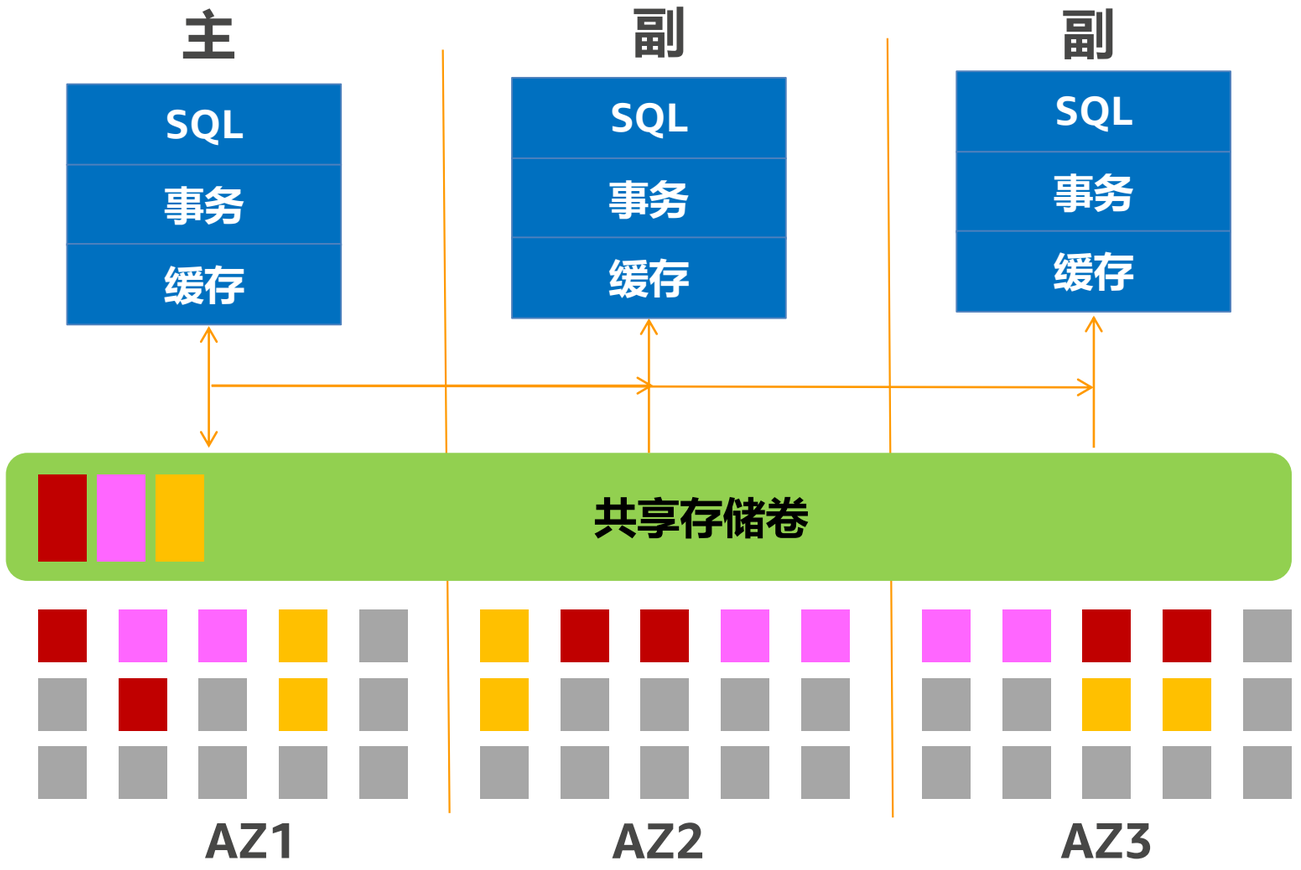
缓存与存储层耦合



- **相同的整体式堆栈**
- **分布式共识算法表现不佳**

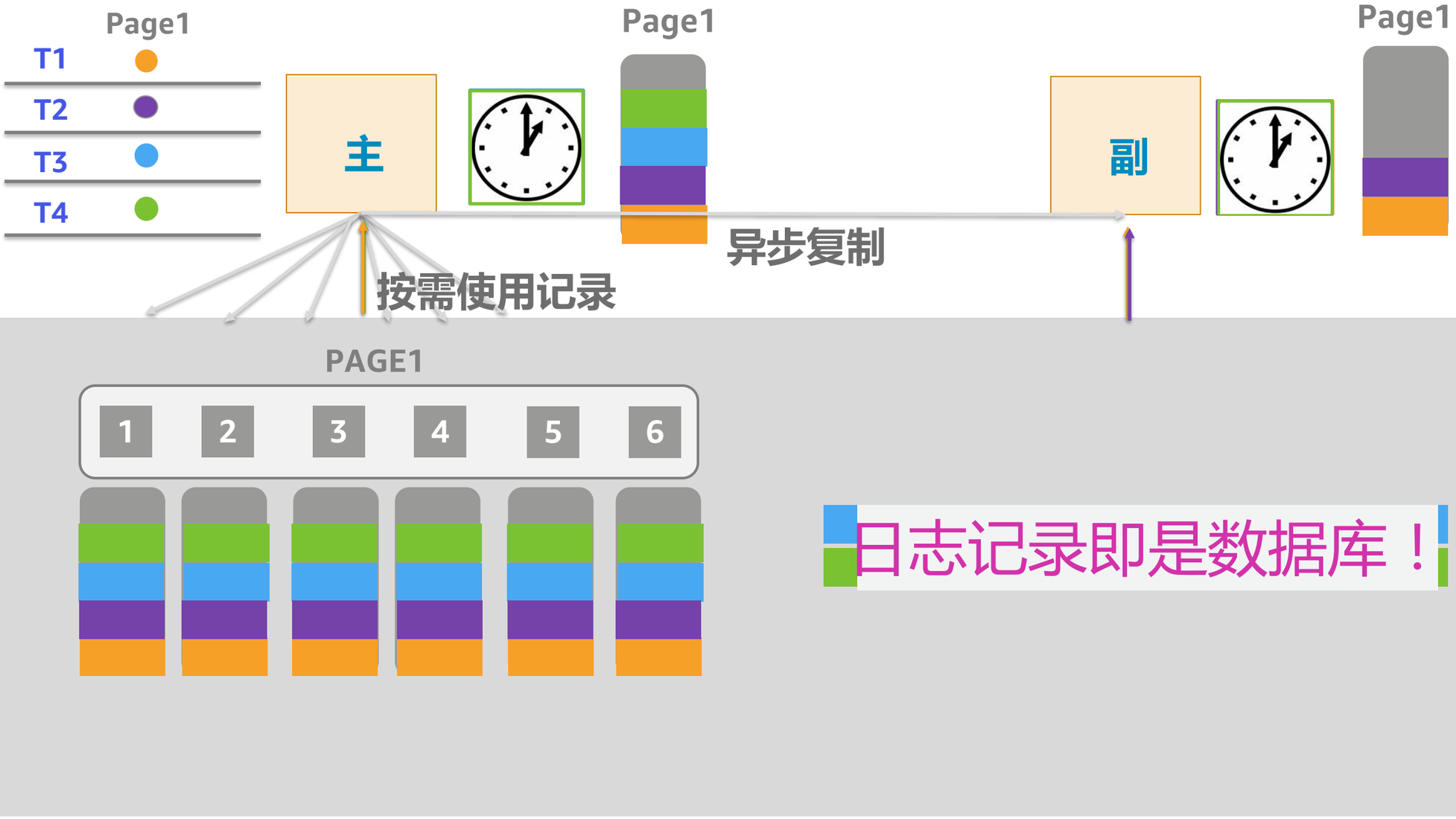
Aurora: 横向扩展、分布式架构

- 将 Log 机制推送至存储层
- 4/6 写入仲裁与本地跟踪
  - ✓ 写入性能
  - ✓ 读取横向扩展
  - ✓ 可用区+1容错机制
  - ✓ 即时数据库重做恢复

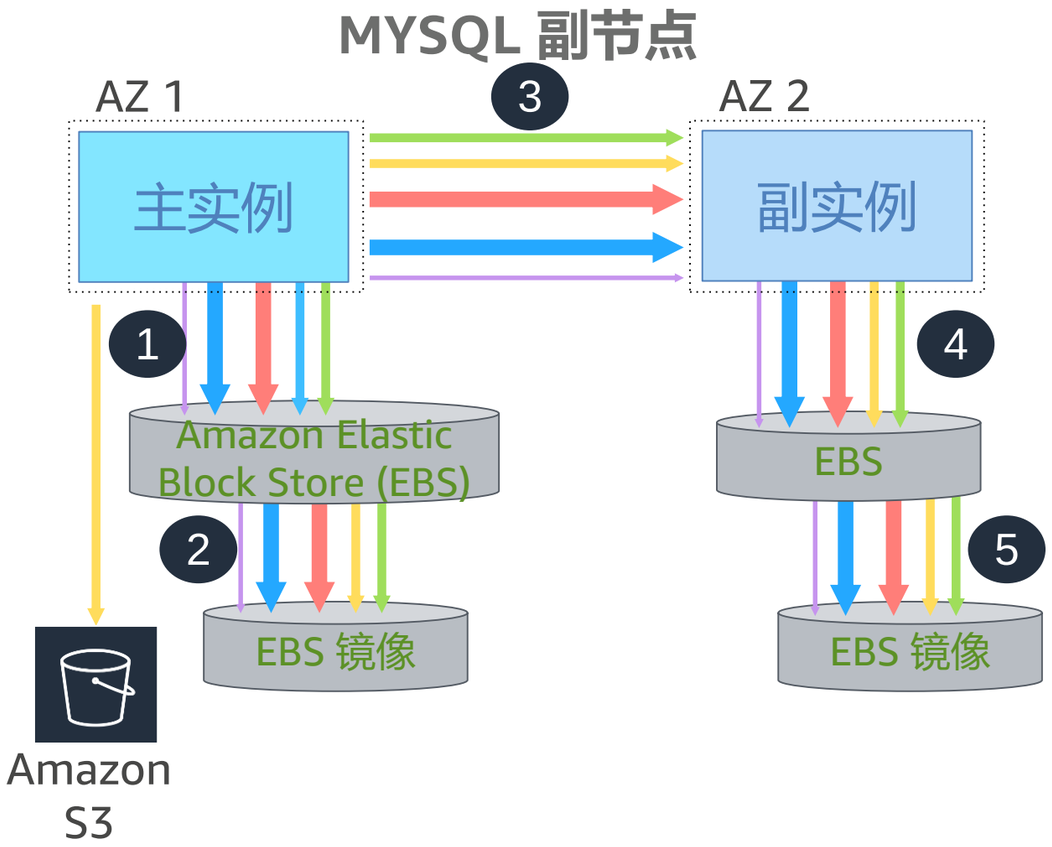


不再需要妥协！

实际运行中的日志应用情况



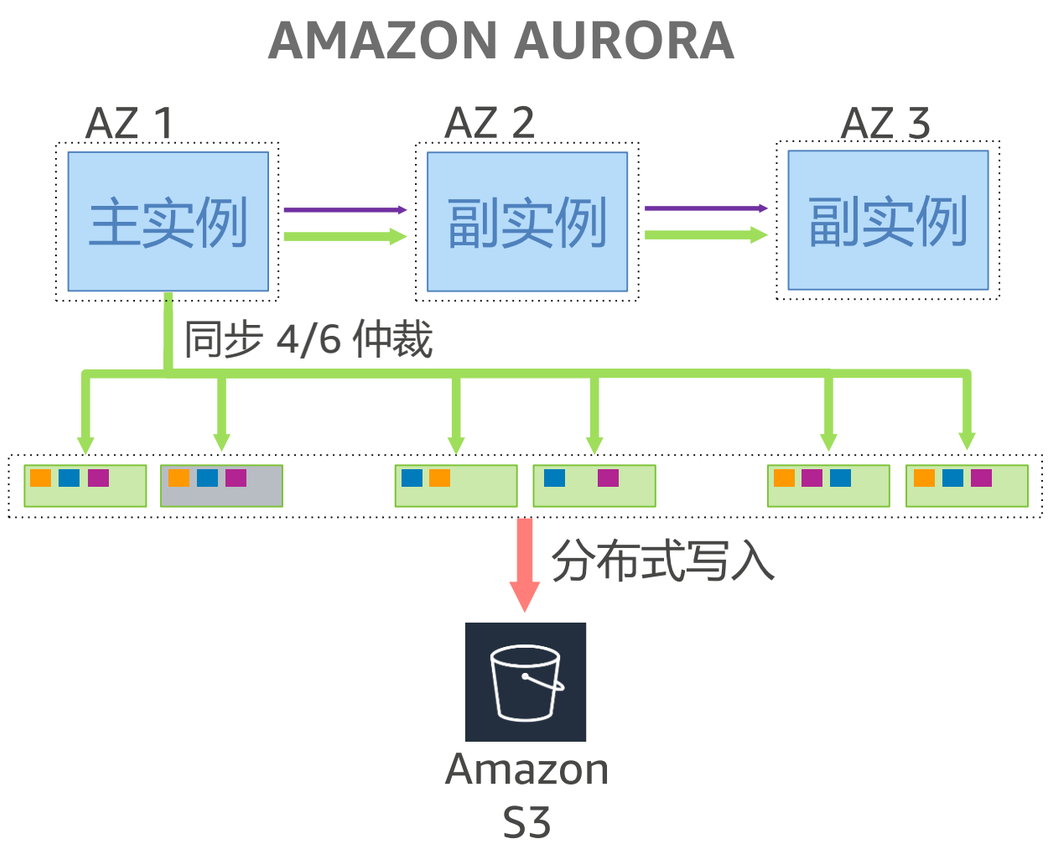
MySQL 与 Aurora 的 I/O 特征差异



MySQL I/O配置运行Sysbench 30分钟

0.78MM 事务

每事务7.4次I/O



Aurora IO 配置运行 Sysbench 30分钟

27MM事务

每事务0.95次I/O

达到前者的35倍

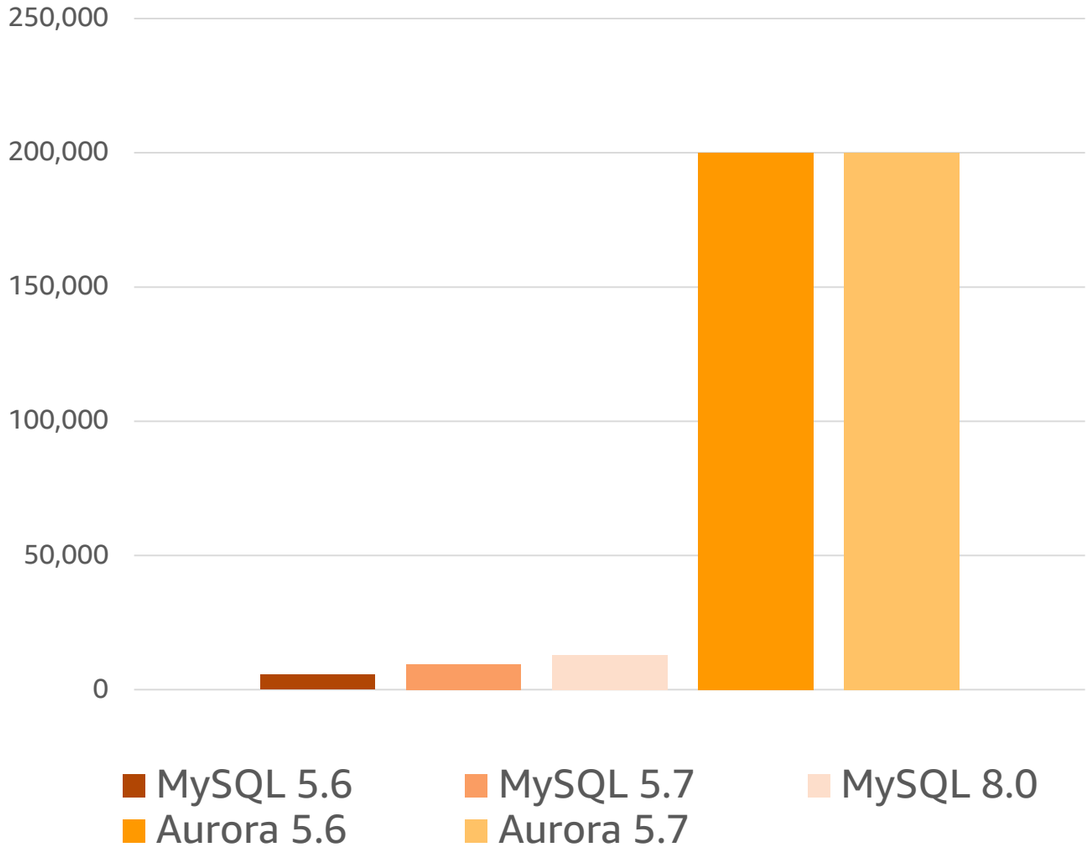
仅为前者的1/7.7

写入类型

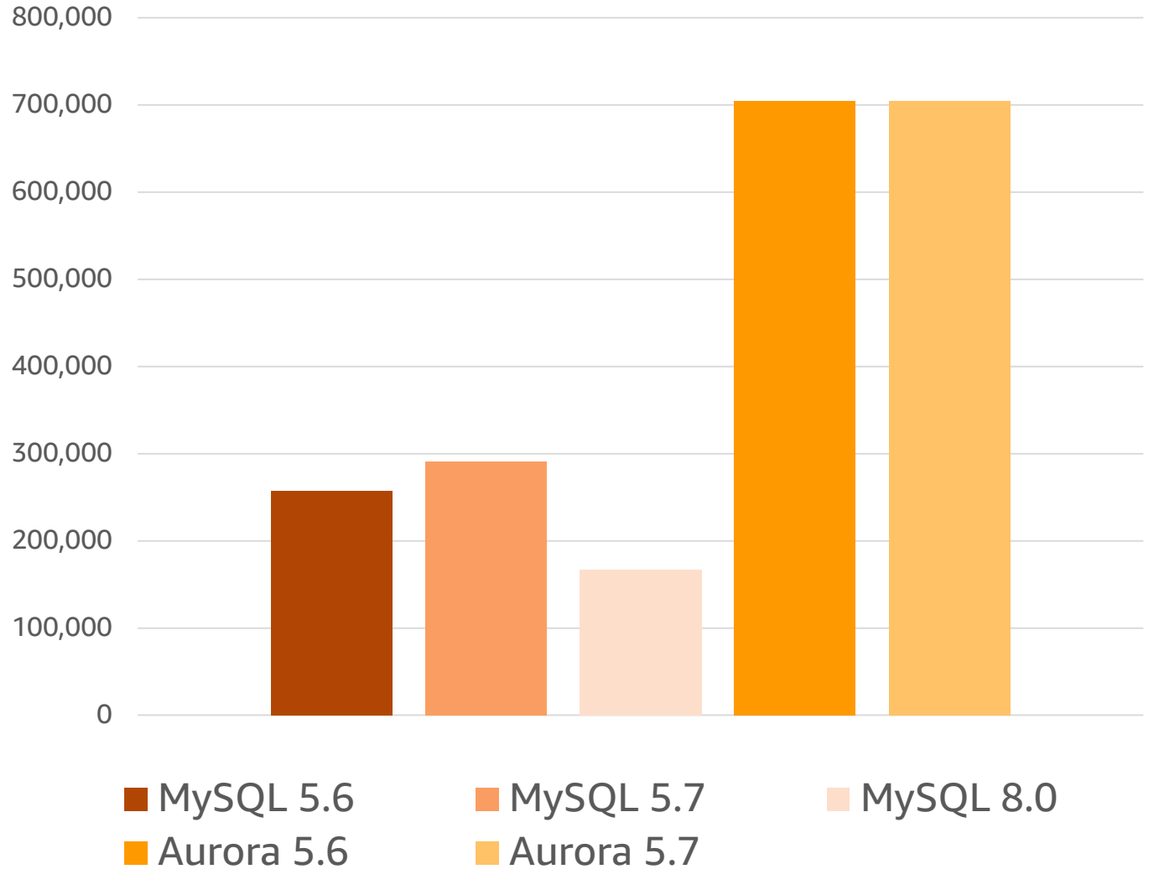
日志　二进制日志　数据　双写入　FRM文件

写入与读取吞吐量

Aurora MySQL 的速度可达 MySQL 的5倍



写入吞吐量

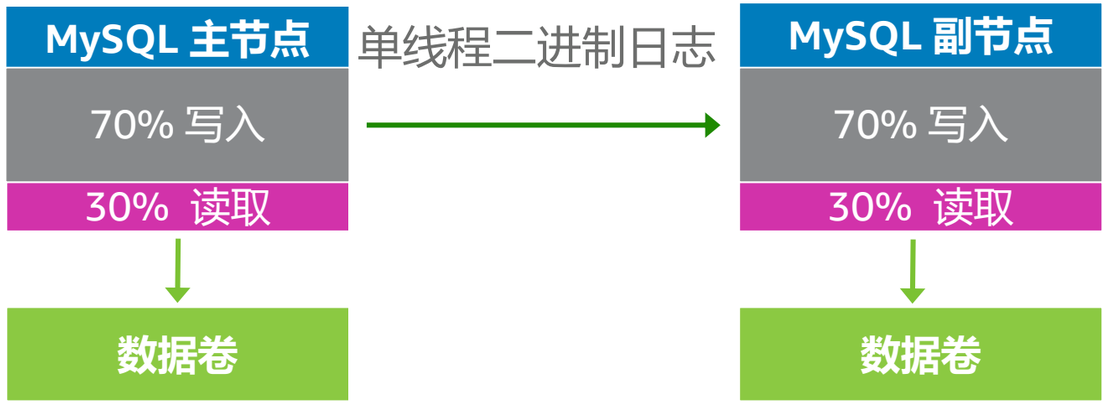


读取吞吐量

在 R4.16XL 实例上运行 Sysbench 处理250份表，每份表包含20万行

读取横向扩展

MYSQL 读取扩展



使用**完整的**变更逻辑

相同的写入工作负载

独立存储

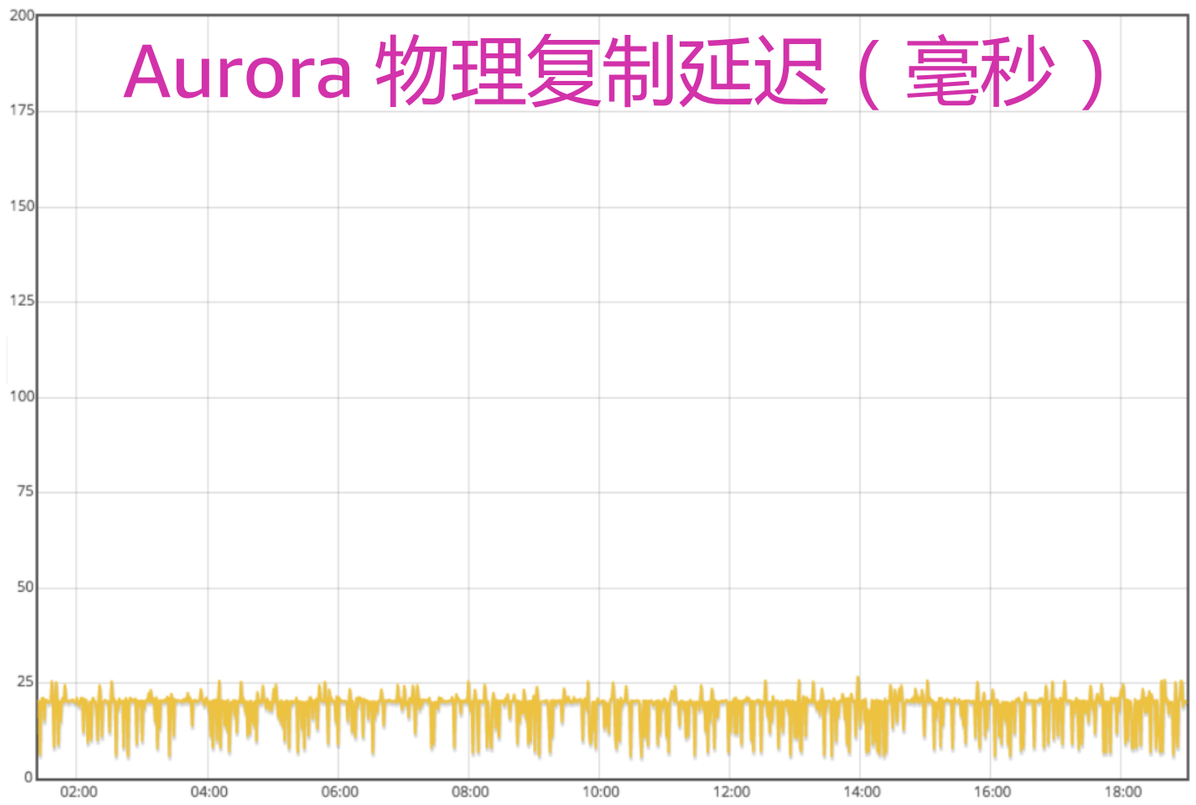
AMAZON AURORA 读取扩展

以物理方式使用 **delta** 变更

不向 replica 写入

共享存储

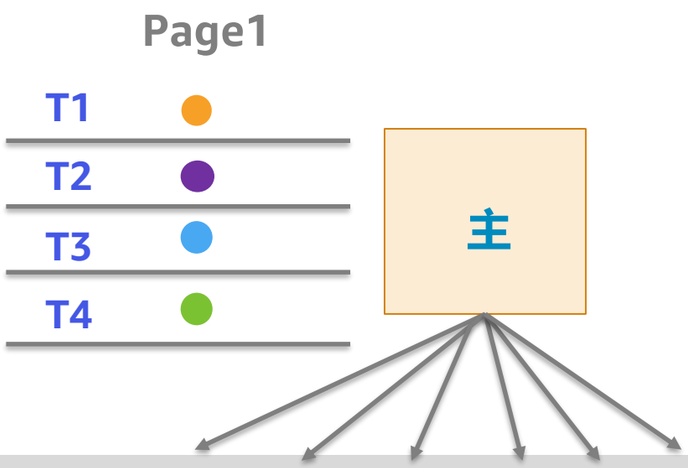
Aurora MySQL 逻辑与物理复制延迟



GoGuardian

“在 MySQL 当中，我们看到复制延迟达到近12分钟。从实际应用的角度来看，这显然是种近乎荒谬的情况。利用 Aurora，4个副本的最大读取延迟从未超过20毫秒。”

运行中的 Quorum 与本地跟踪

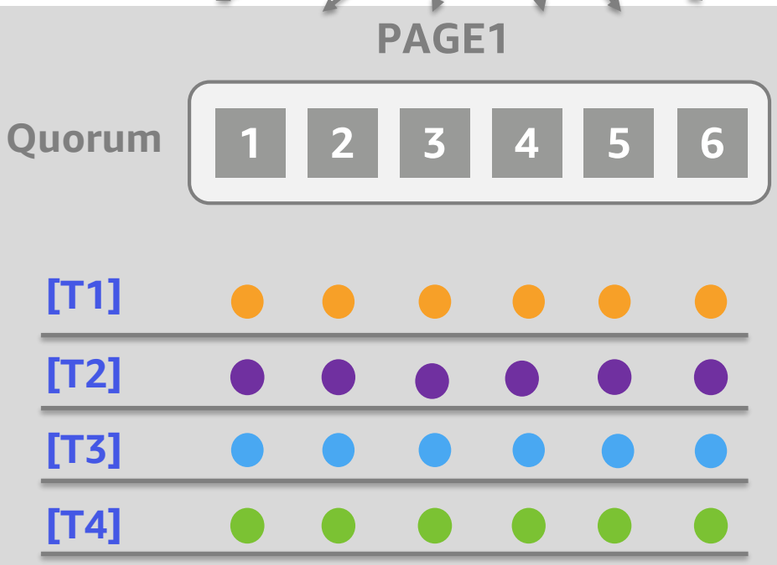


持久性: 1 2 3 4 5 6

跟踪:

等待Tx: 无 T3 T4 T4

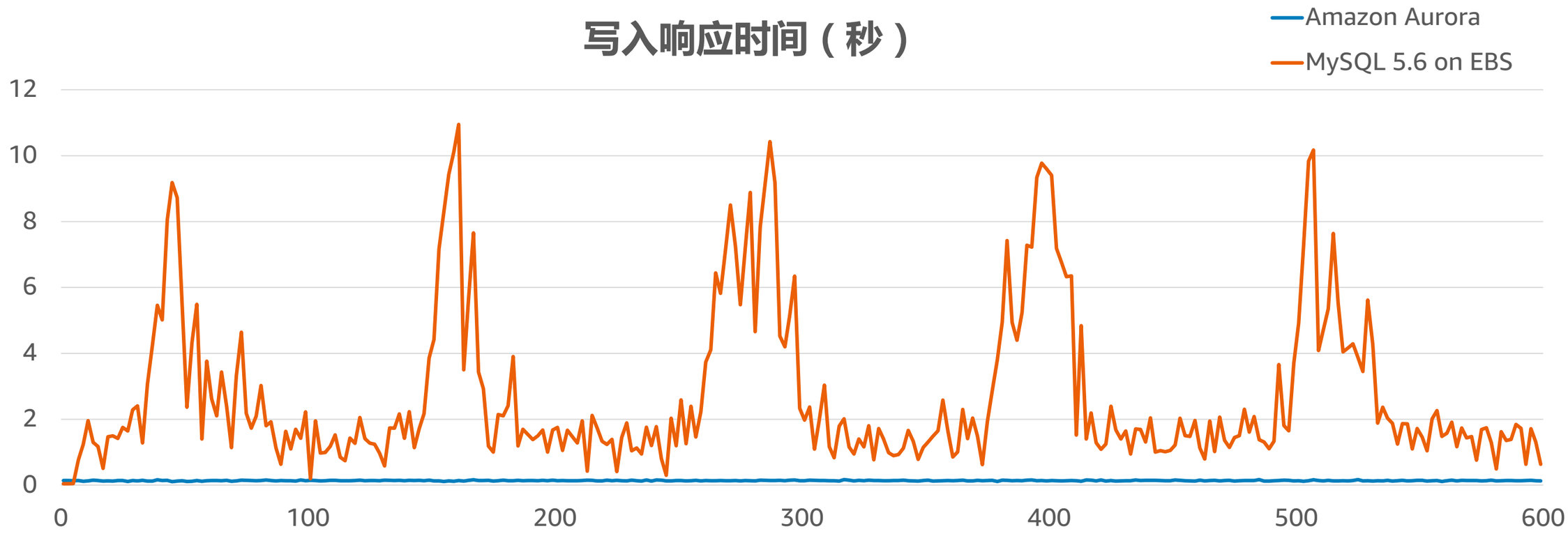
已提交Tx: 无 T1 T2 T3 T4



无需量级
分布式提交

负载条件下的性能变化

Amazon Aurora >一致性提升200倍

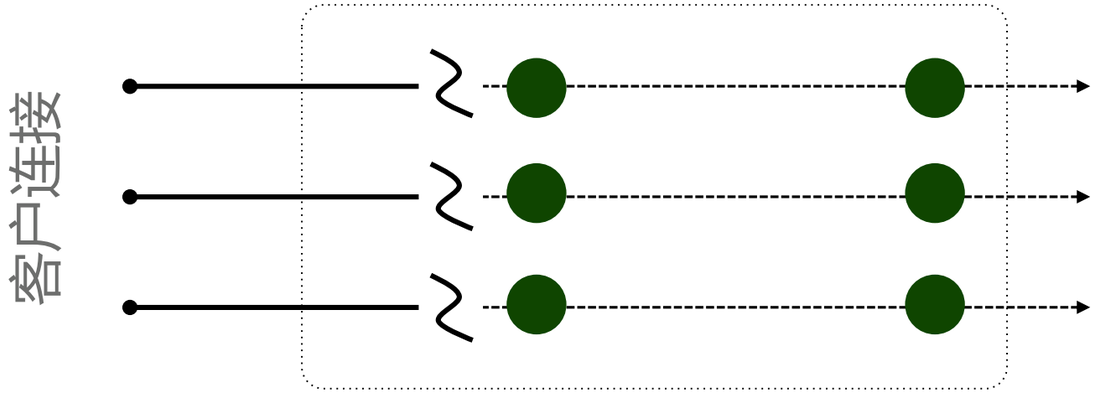


时间以秒为单位

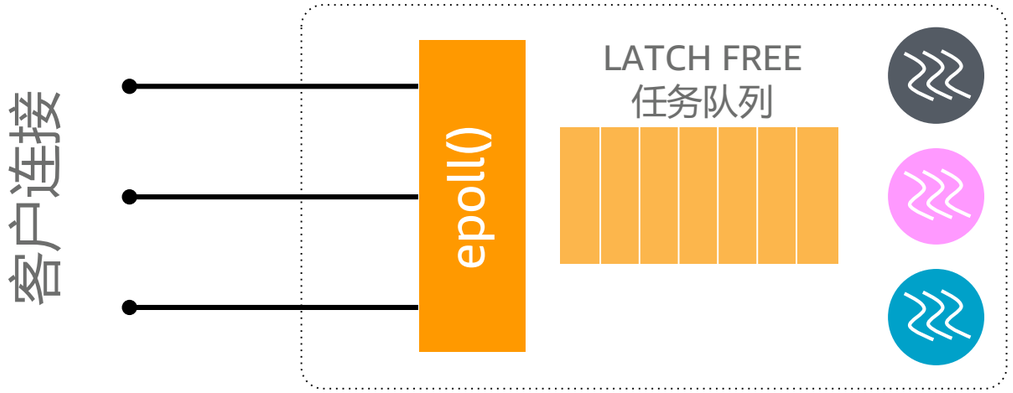
SysBench OLTP（只写）工作负载与250份表，每表20万行，采用 R4.16XL 实例

我们还做出哪些其它努力以提升吞吐量？

MYSQL 线程模型

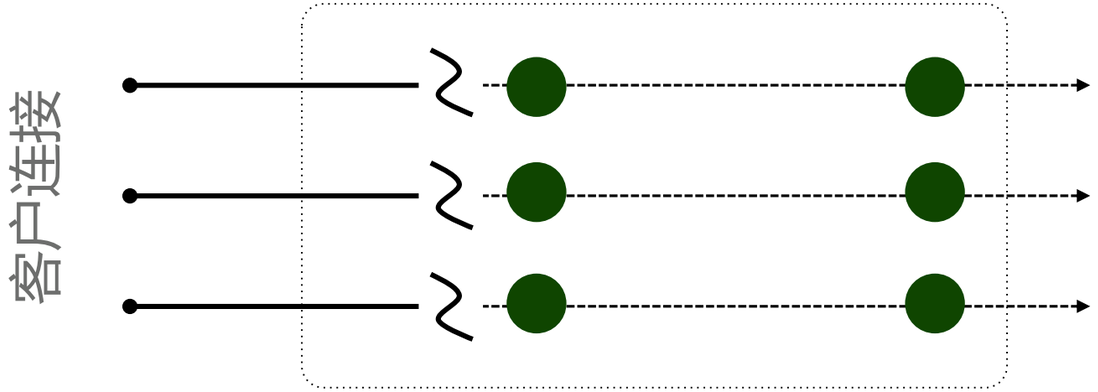


AURORA 线程模型

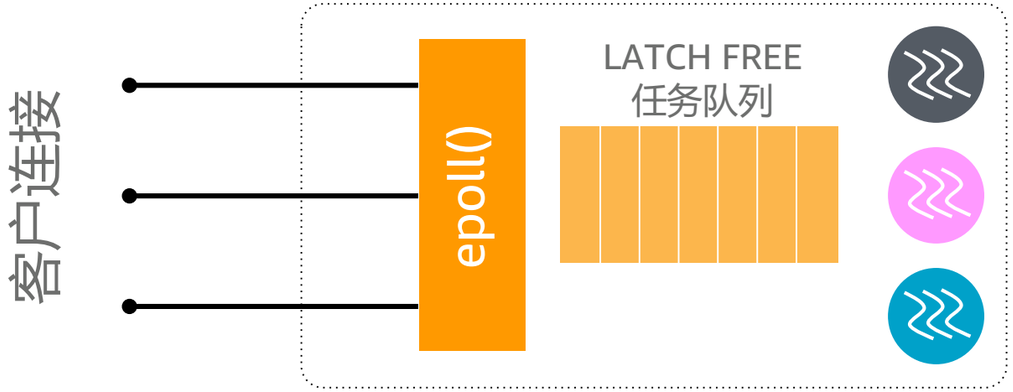


我们还做出哪些其它努力以提升吞吐量？

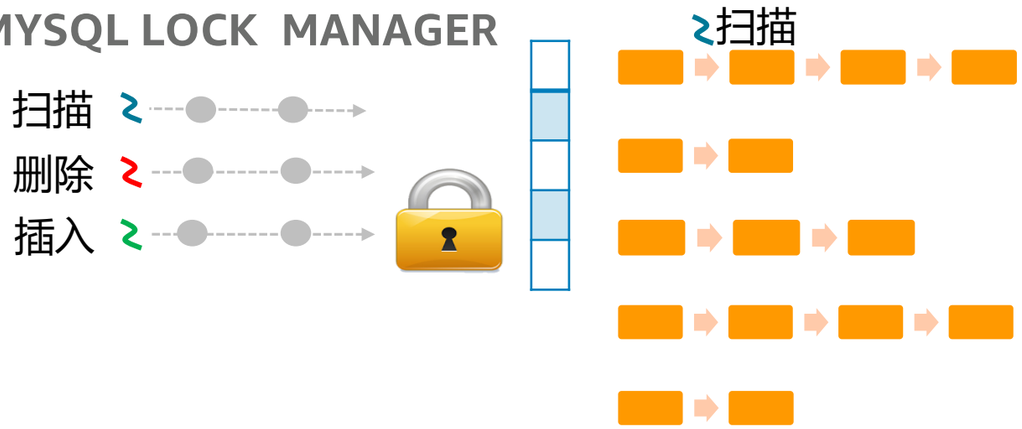
MYSQL 线程模型



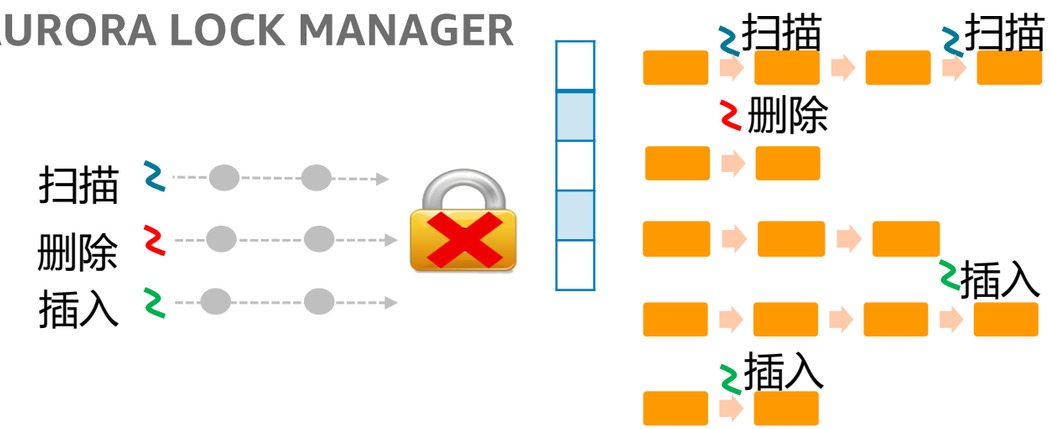
AURORA 线程模型



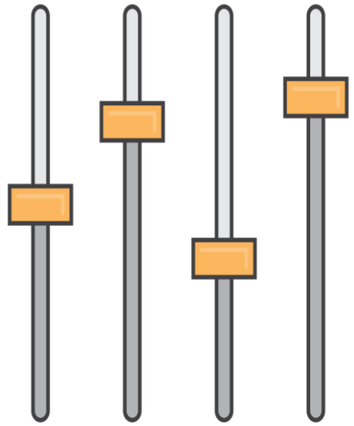
MYSQL LOCK MANAGER



AURORA LOCK MANAGER

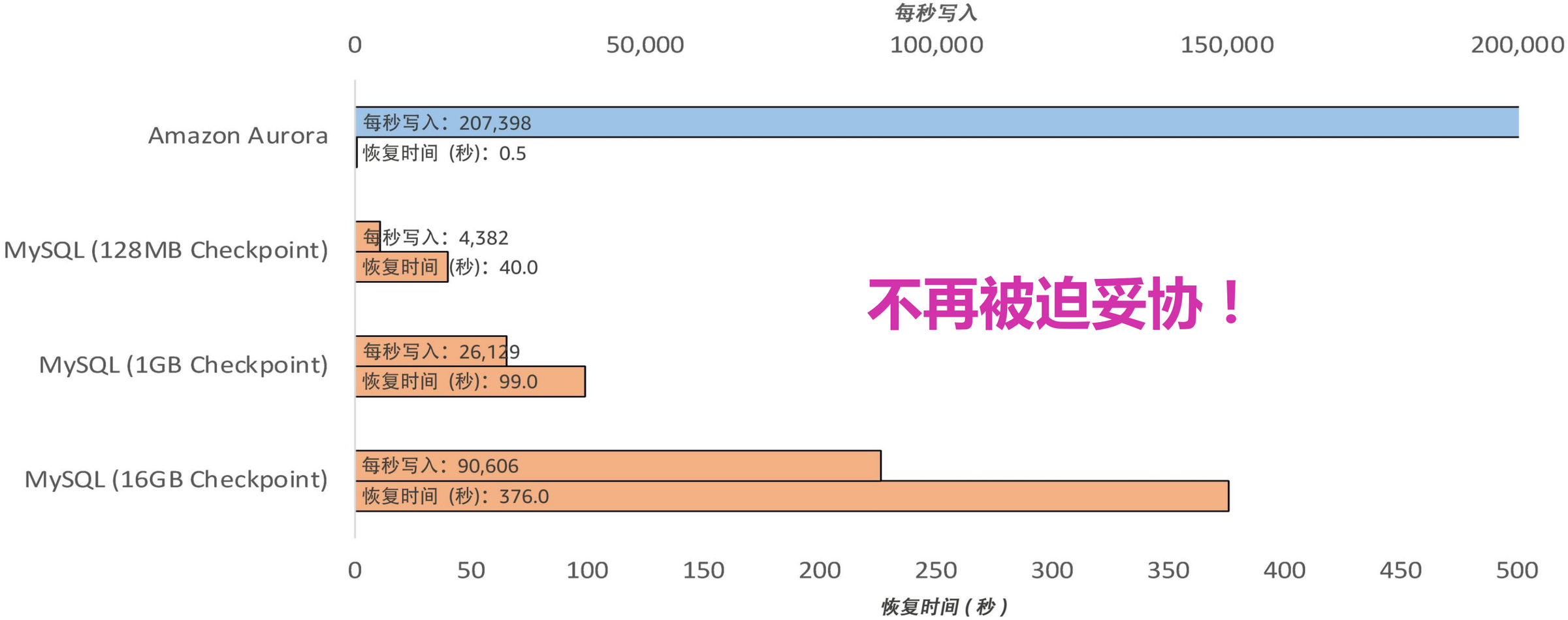


具体性能参数？



针对不同硬件配置进行**预调优**或**自动调优**

恢复时间与写入性能

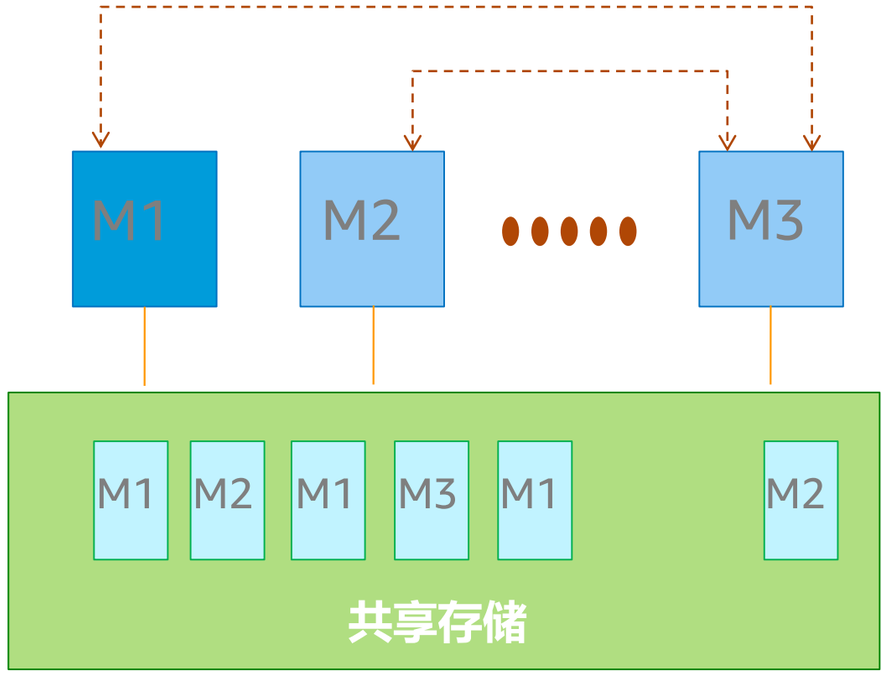


不再被迫妥协！

SysBench OLTP 只写）10 GiB工作负载，250 份表，每表20万行

现有多主解决方案

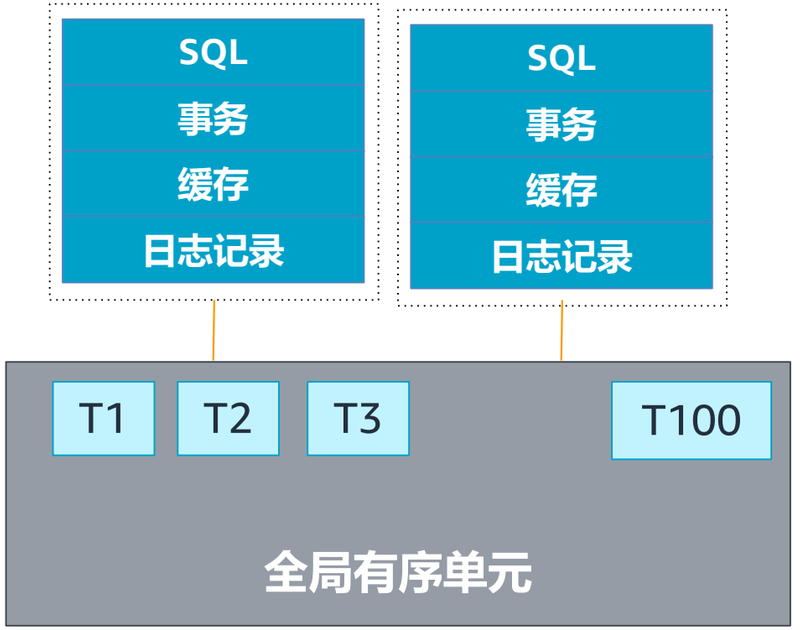
分布式 lock manager



重量级同步：悲观锁与强制锁的扩展表现

例如 – Oracle RAC, DB2 Purescale, Sybase

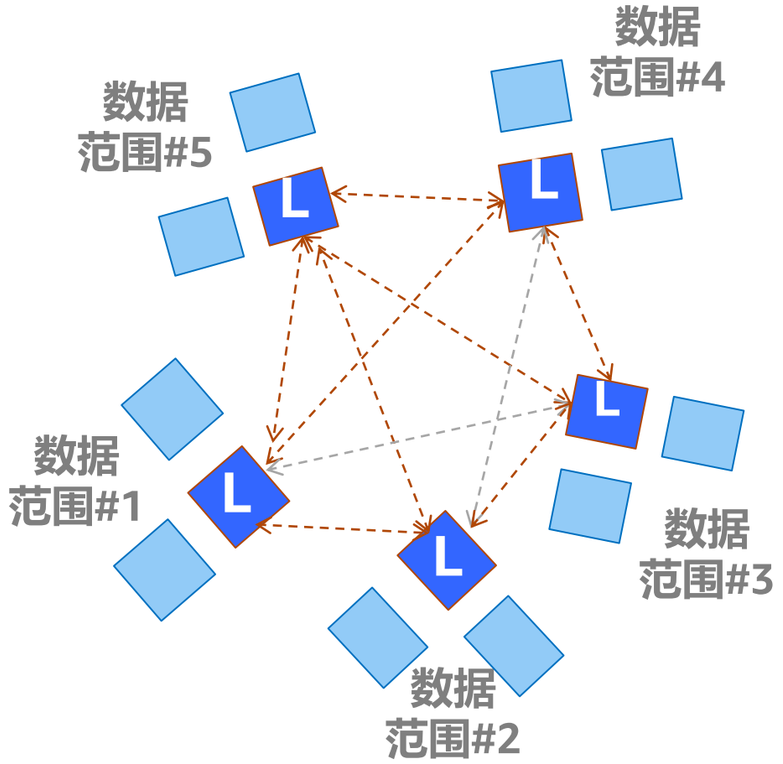
读取-写入集的全局排序



全局实体：规模伸缩瓶颈

例如 – Galera, TangoDB, FaunaDB

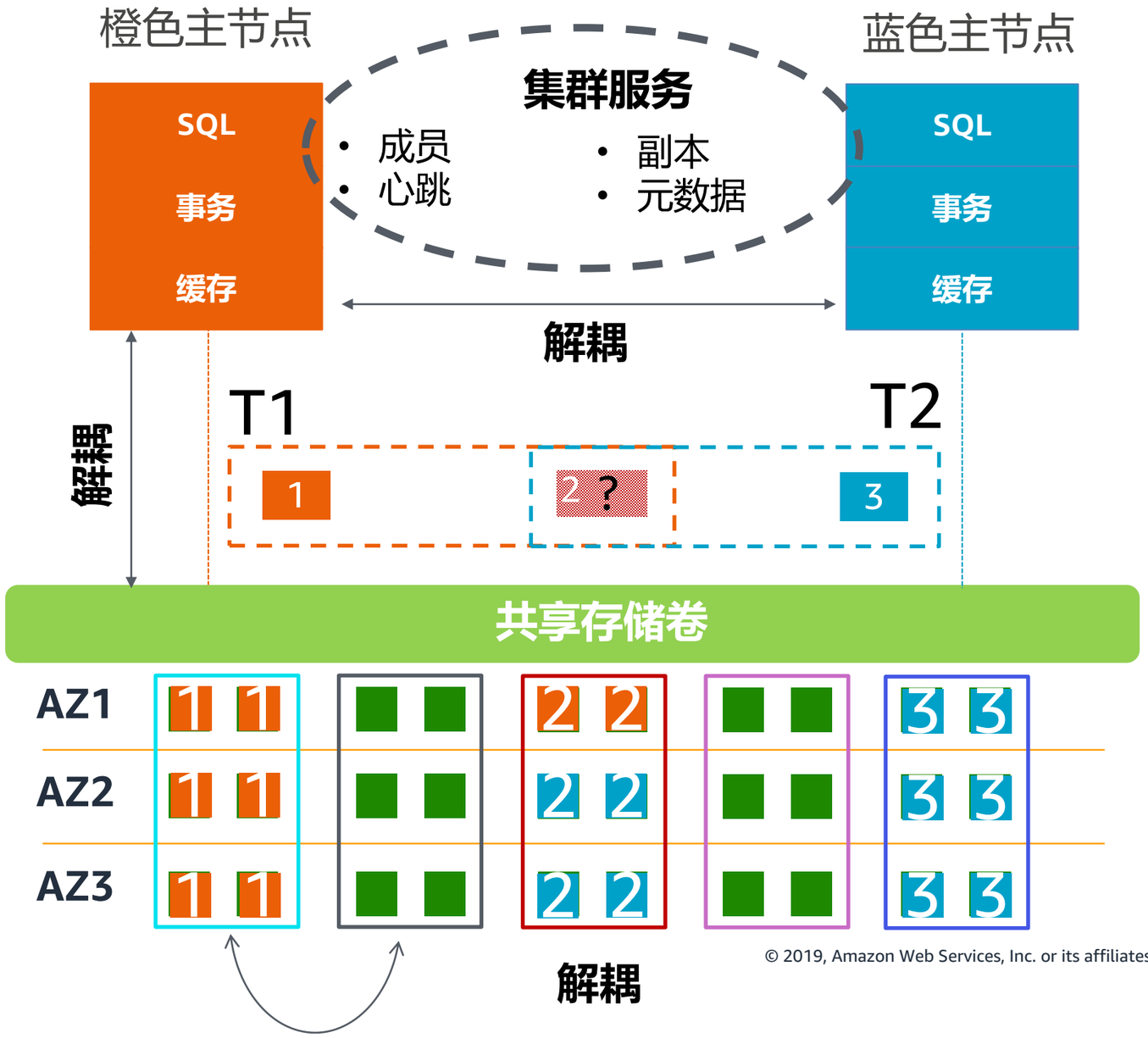
Paxos leader 与 2PC



重量级共识协议：热分区与跨分区查询难题

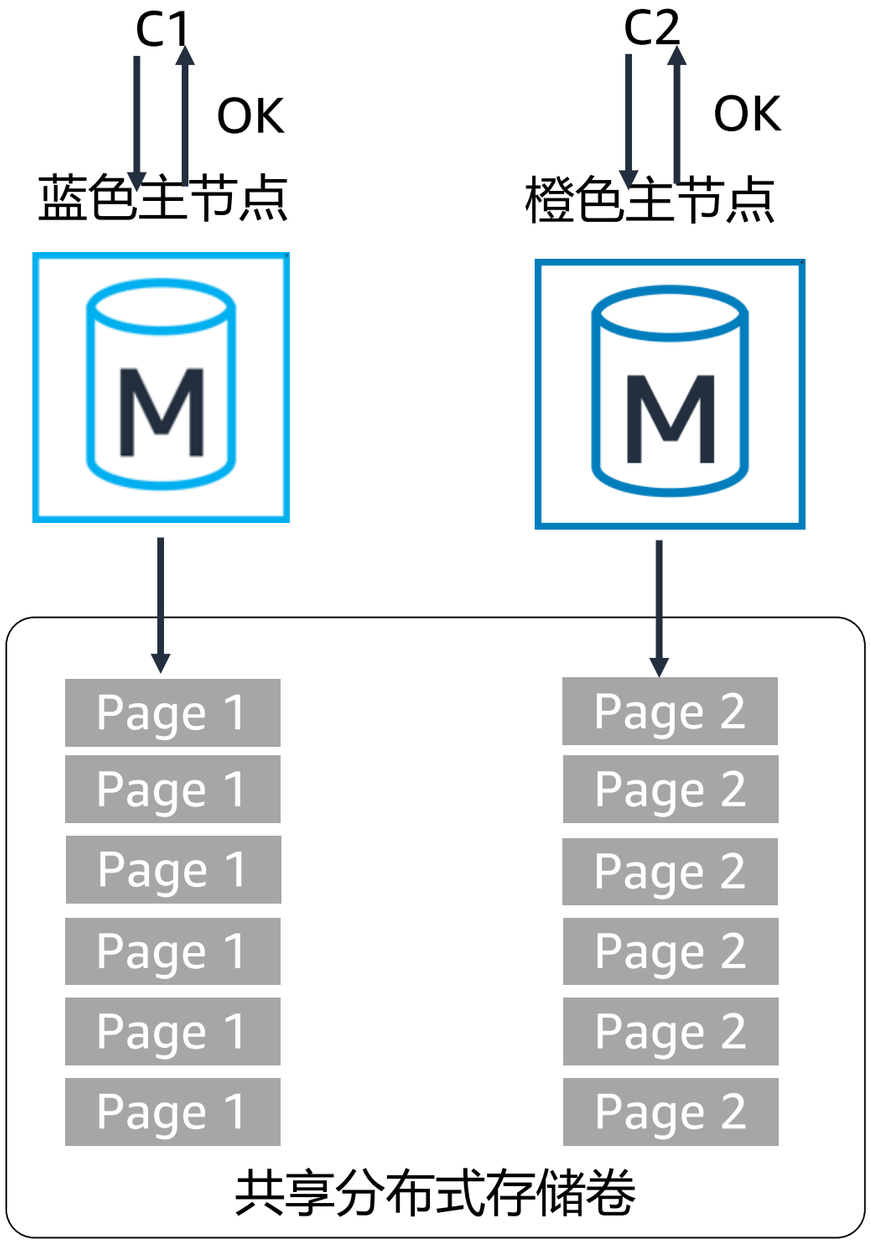
例如 – Spanner, CockroachDB, Ignite

Aurora 多主架构



- 无悲观锁
- 无全局排序
- 无全局提交协调
- 乐观冲突解决方案
- 解耦系统
- 微服务架构

多主架构——如何实现（happy path）

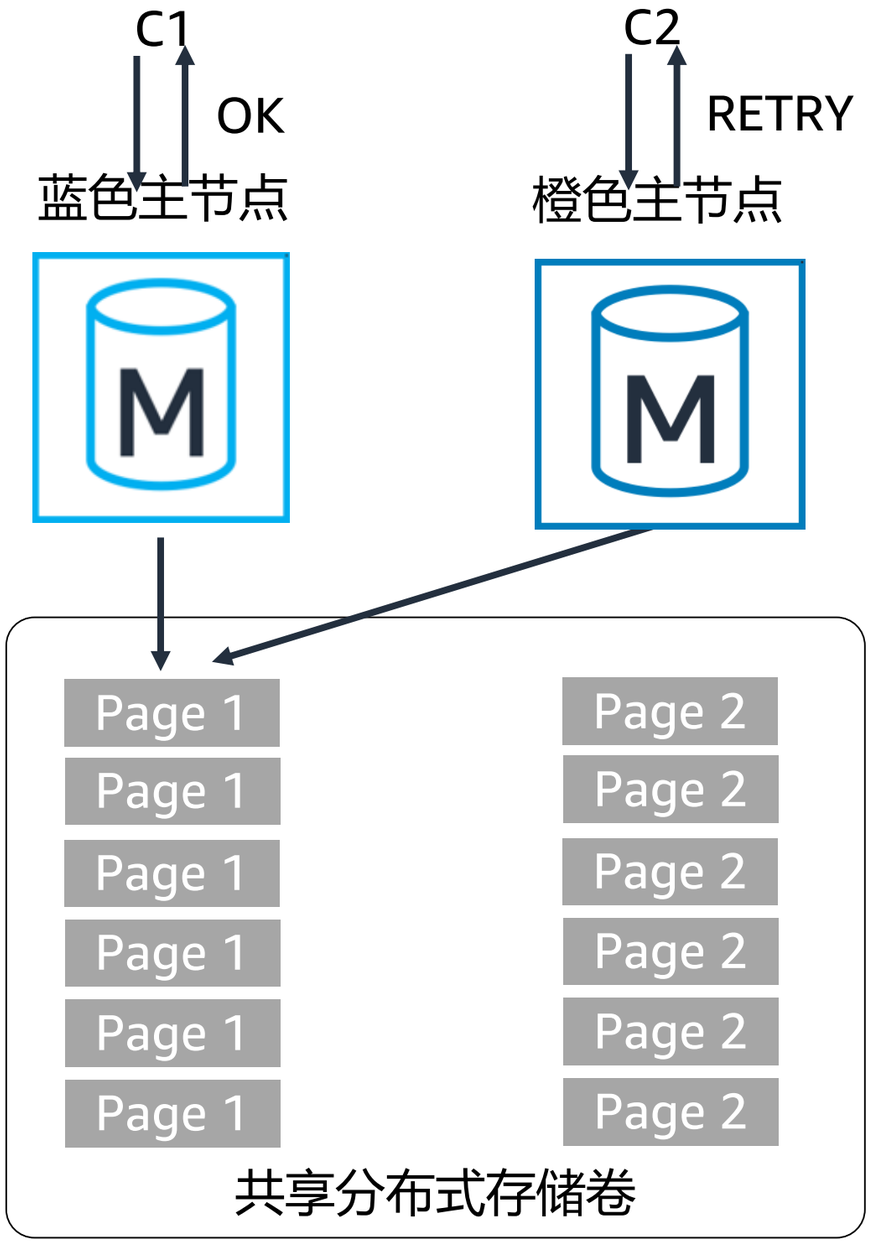


源自不同表上不同主节点的非冲突性写入

时间	蓝色主节点	橙色主节点
1	Begin Trx (BT1)	Begin Trx (OT1)
2	Update (table1)	Update (table2)
3	Commit (BT1)	Commit (OT1)

零同步

Aurora 多主架构——如何实现（物理冲突）

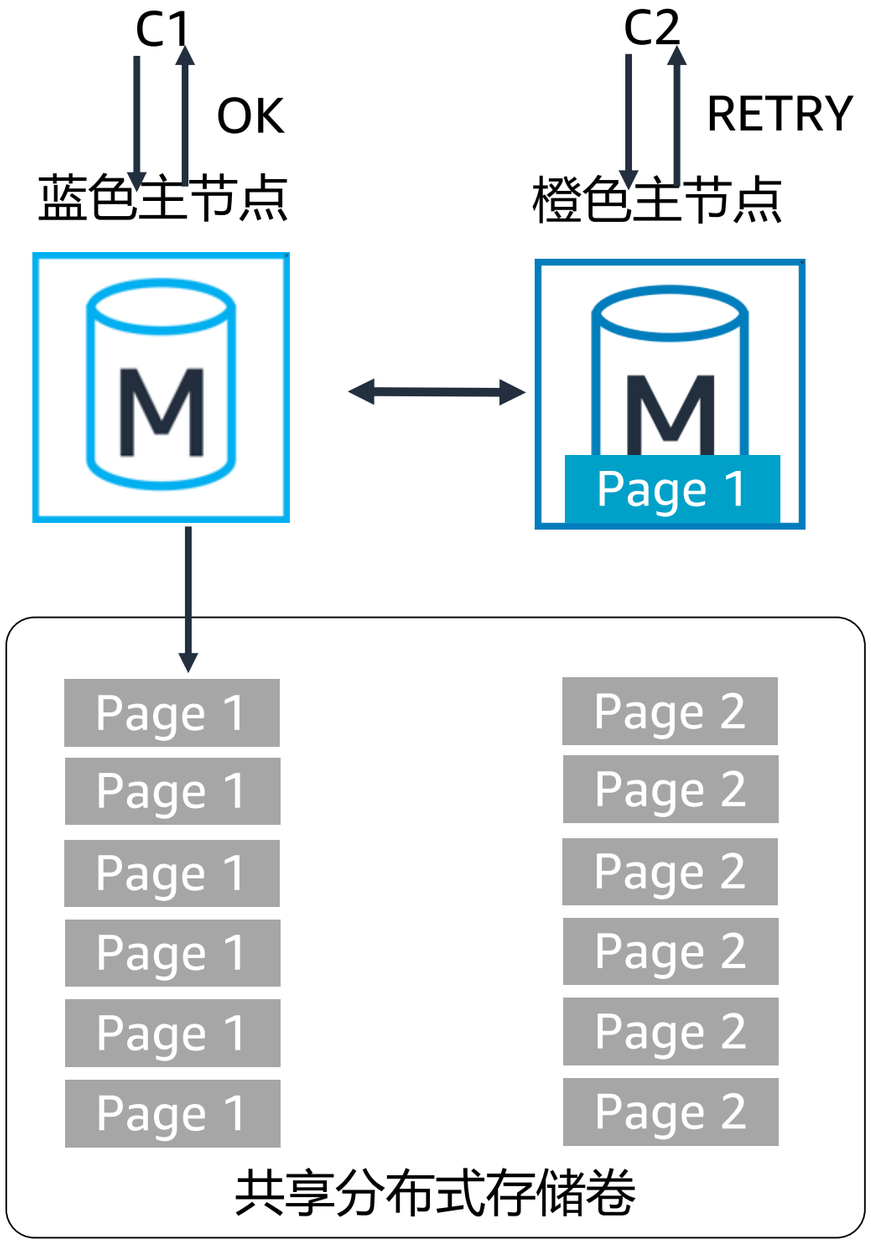


源自同一表上不同主节点的冲突性写入

时间	蓝色主节点	橙色主节点
1	Begin Trx (BT1)	Begin Trx (OT1)
2	Update (row1, table1)	Update (row1, table1)
3	Commit (BT1)	Rollback (OT1)

乐观冲突解决方案

Aurora 多主架构——如何实现（逻辑冲突）

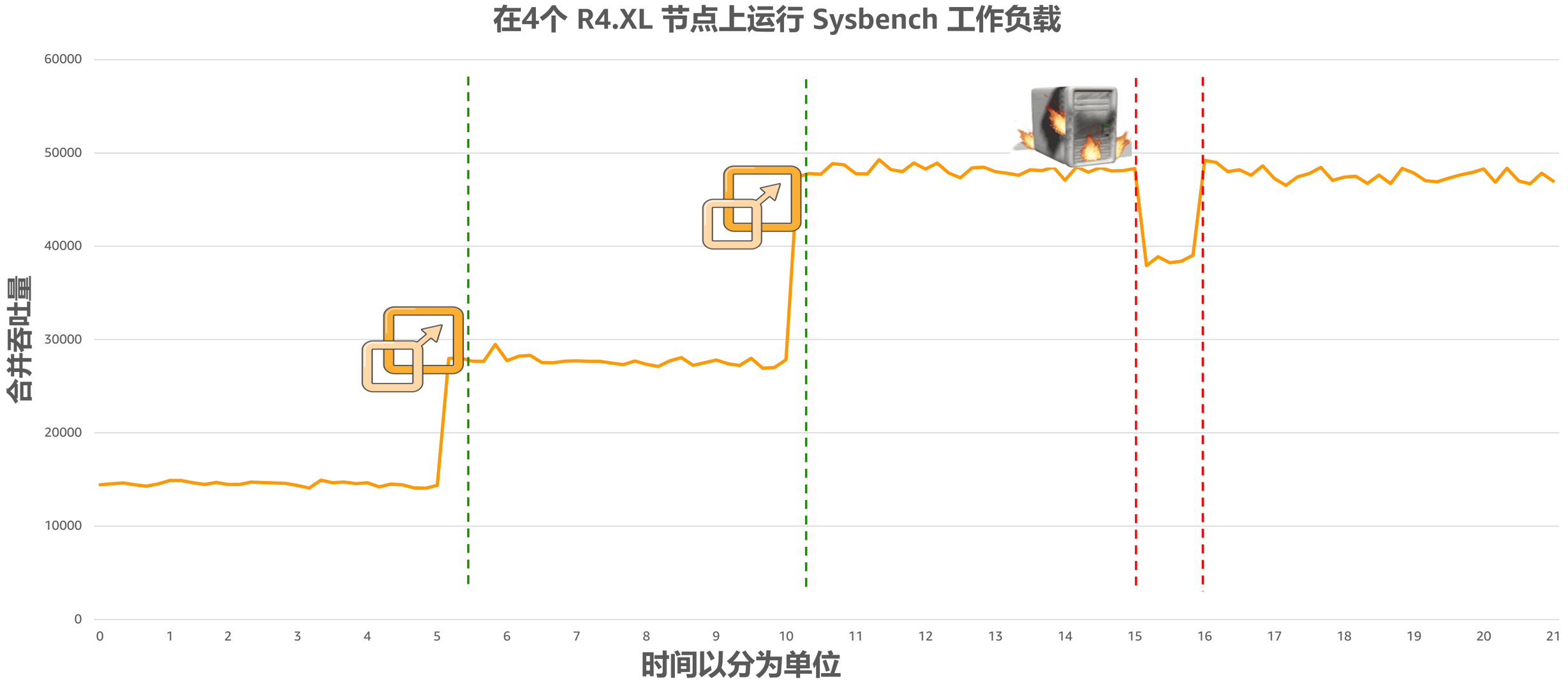


源自同一表上不同主节点的冲突性写入

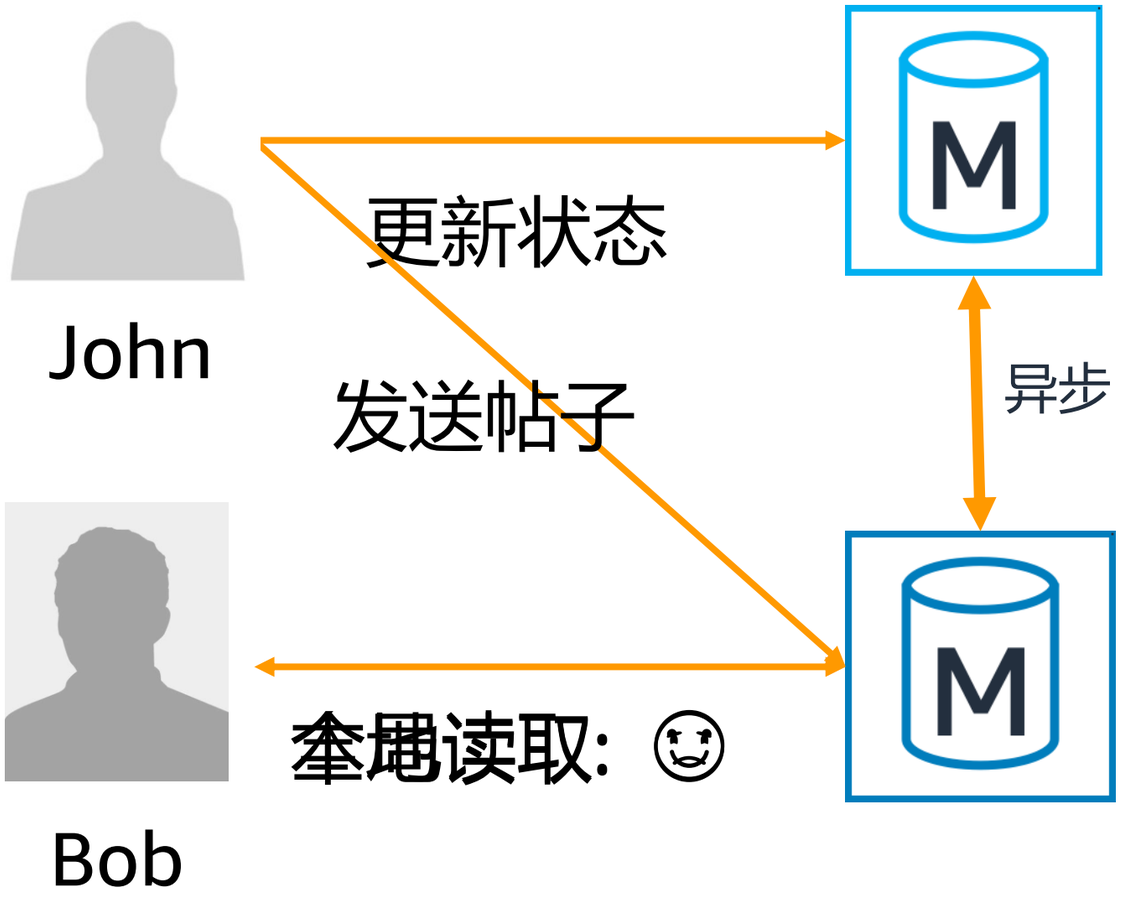
时间	蓝色主节点	橙色主节点
1	Begin Trx (BT1)	Begin Trx (OT1)
2	Update (row1, table1)	
3		Update (row1, table1) and Rollback (OT1)
4		

不需要分布式锁定！

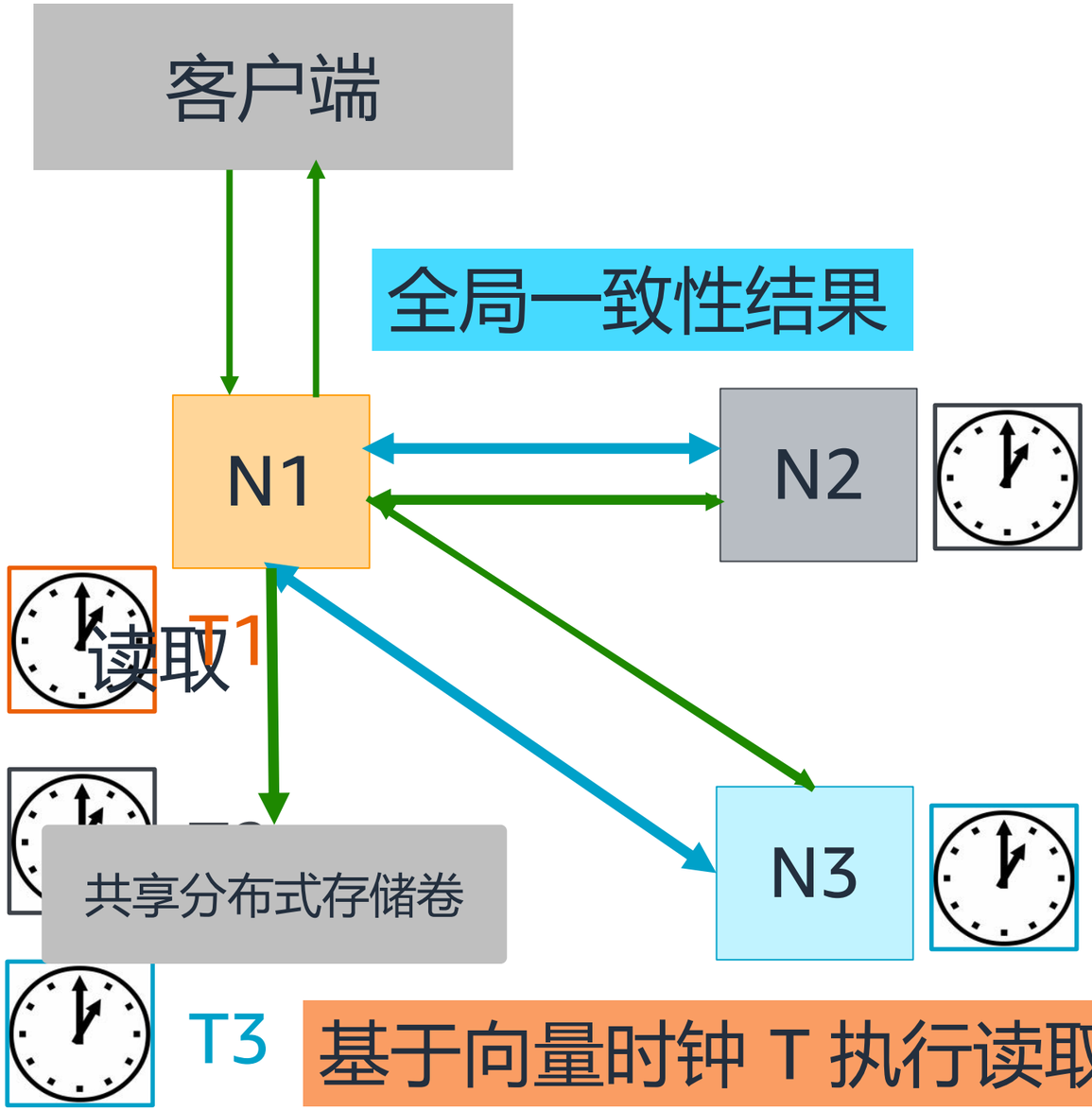
Aurora 多主架构——规模伸缩与可用性



Aurora 多主架构全局读取——是什么？



Aurora 多主架构全局读取——如何实现？



- 写入路径上不存在等待
- 仅给全局一致读取增加延迟
- 可基于会话配置

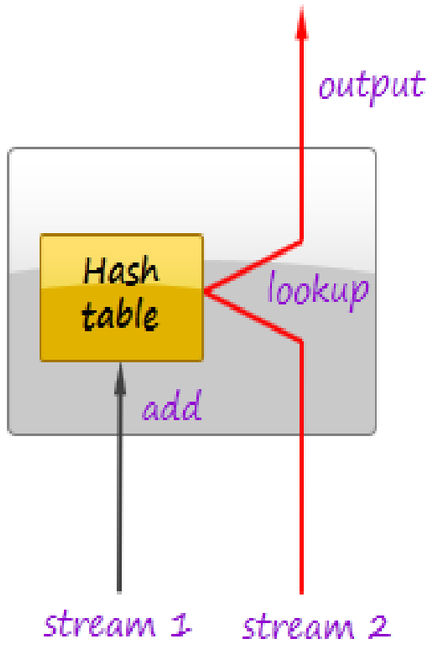
基于向量时钟 T 执行读取，T = (T1, T2, T3)

Aurora 多主架构——总结

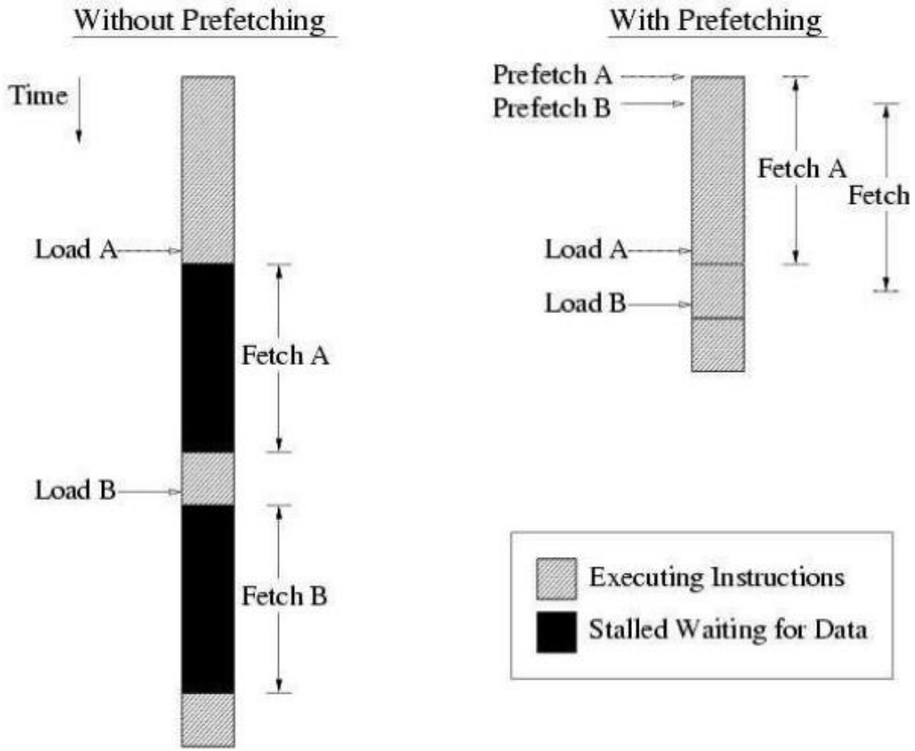
线性伸缩	→	乐观冲突解决方案
持续可用	→	微服务架构
企业级持久性	→	6份副本，每可用区2份
SQL 兼容性	→	支持索引、约束、触发器、过程、函数等

降低查询延迟

批量扫描



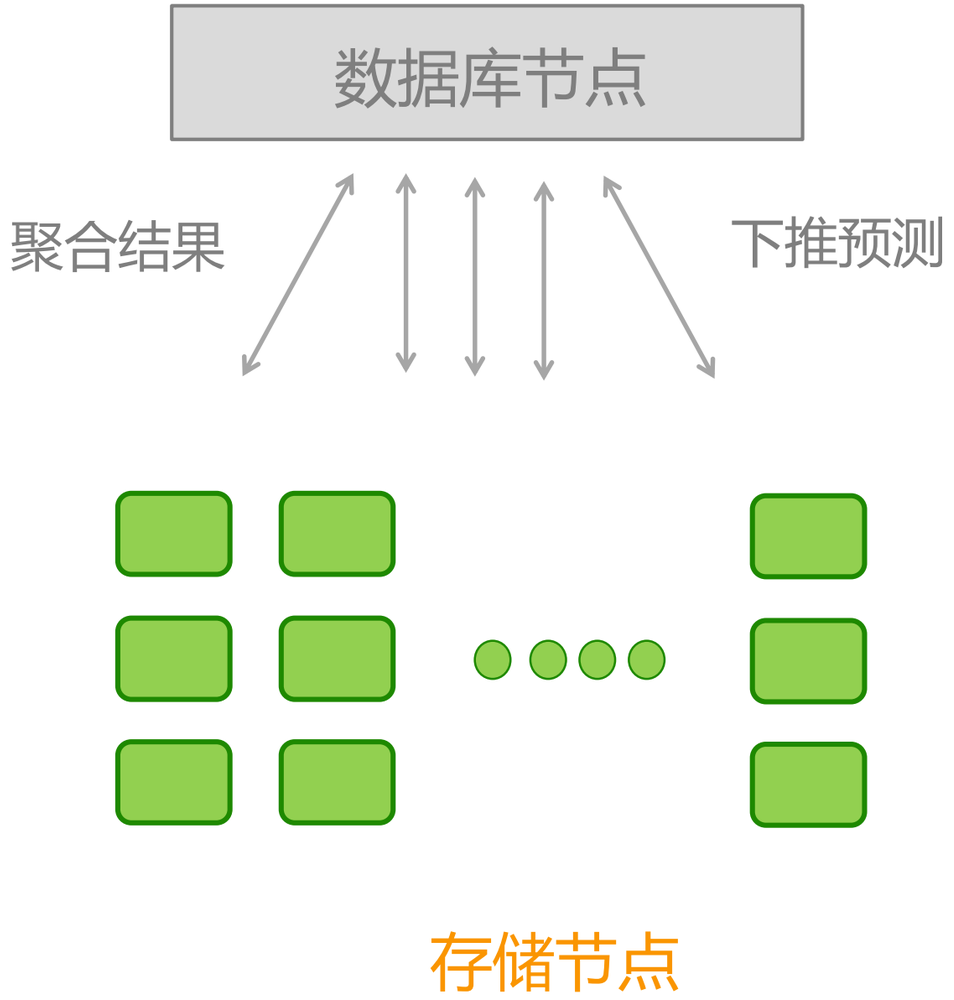
哈希连接



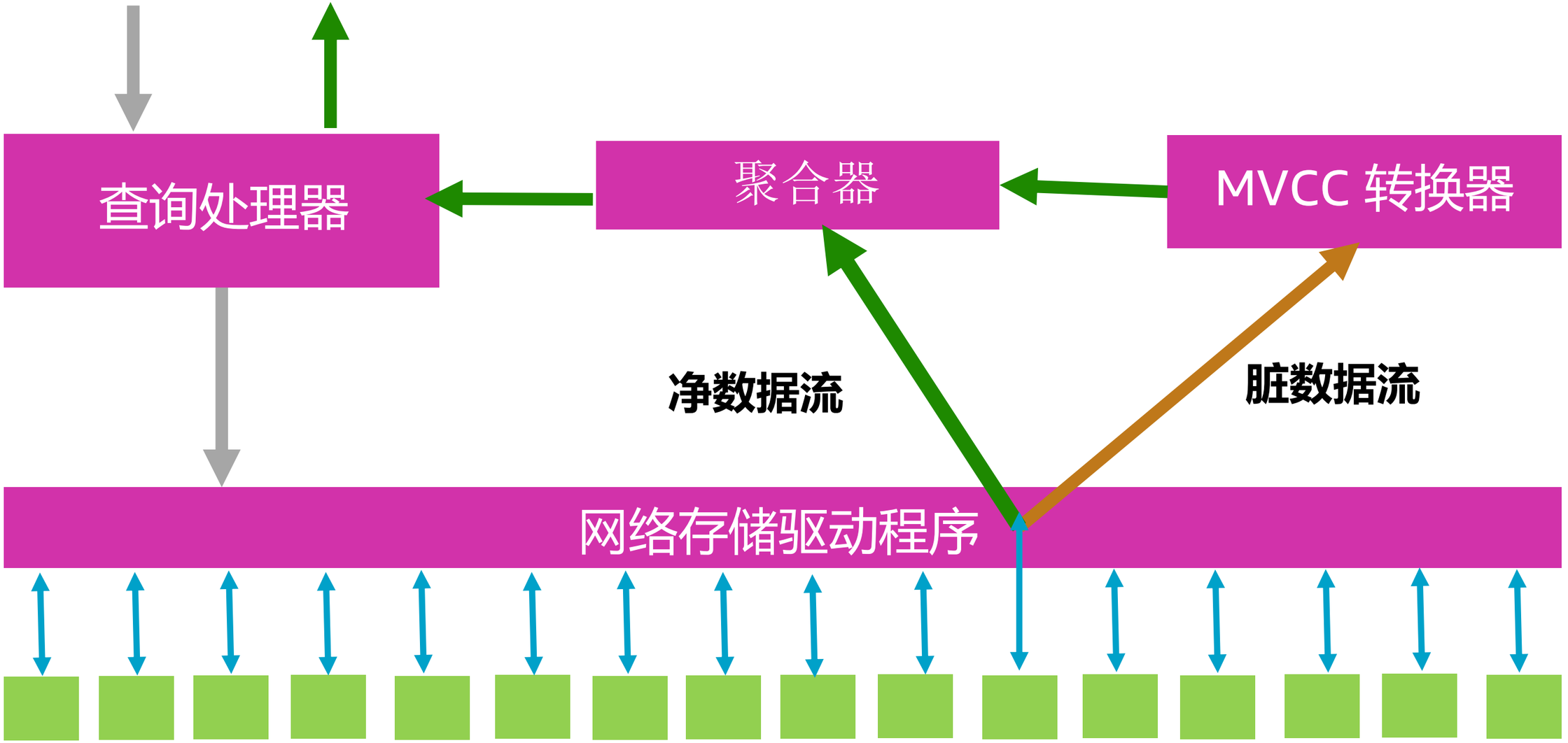
异步键预取

降低查询延迟——并行查询

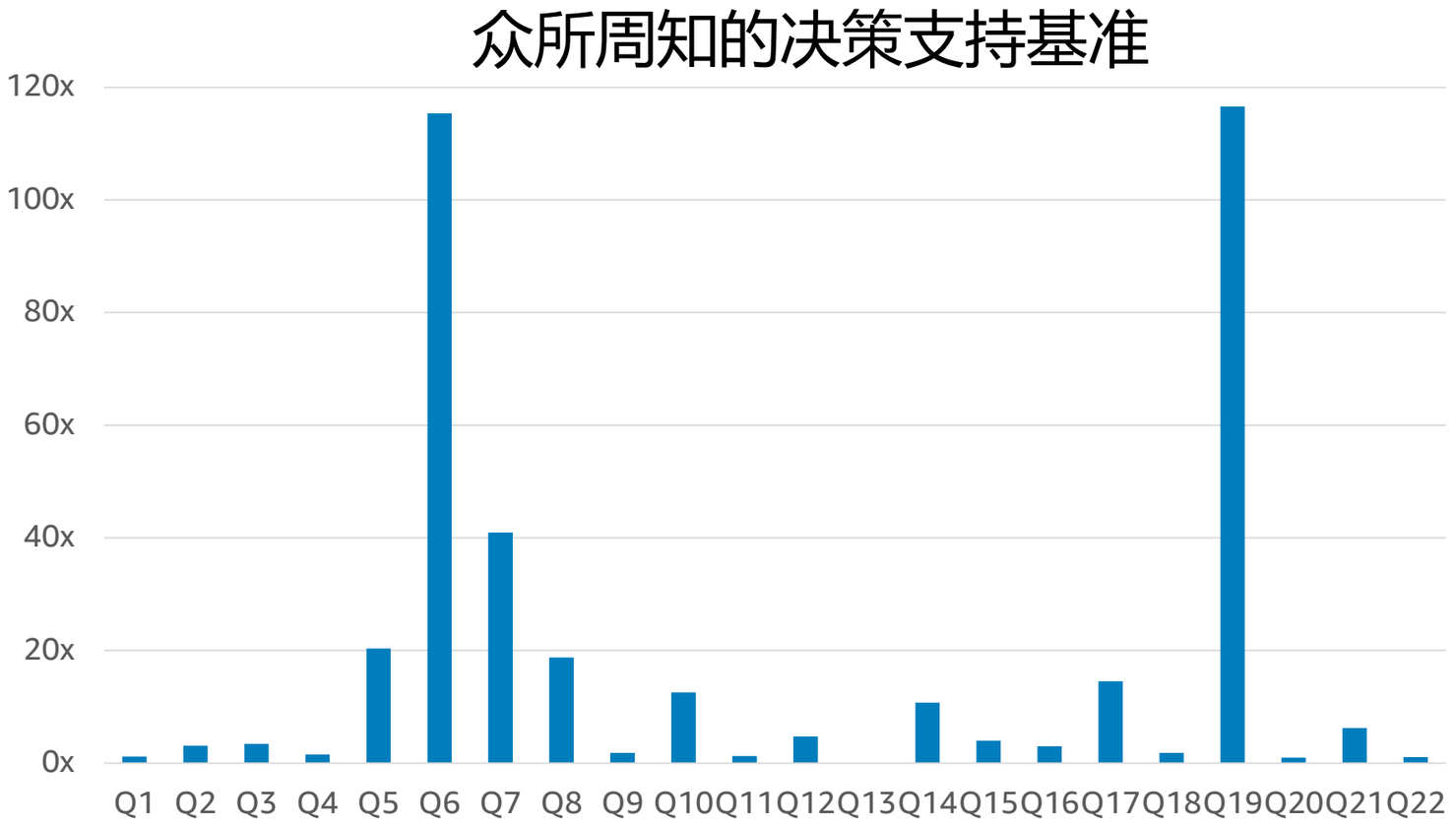
- 并行、分布式处理
- 让处理更接近数据
- 减少缓冲池污染



并行查询架构



并行查询——性能结果



减少查询响应时间

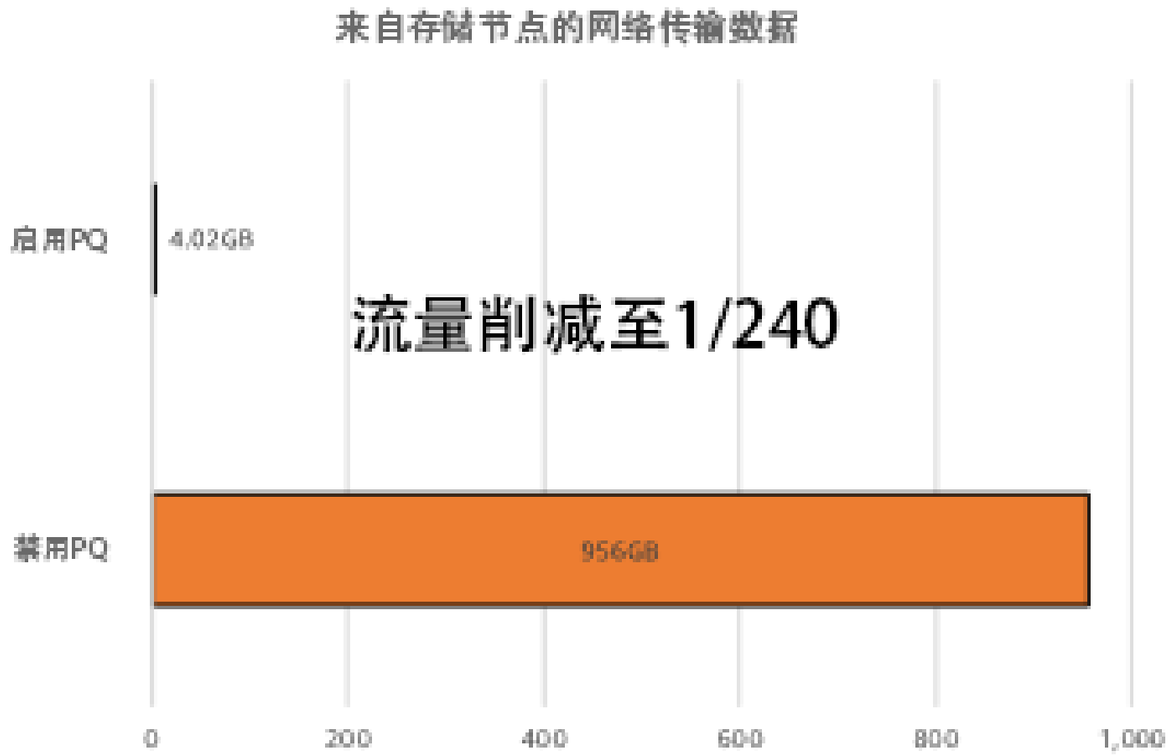
- 峰值加速约120倍
- 超过10倍加速：22项查询中的8项

NETFLIX

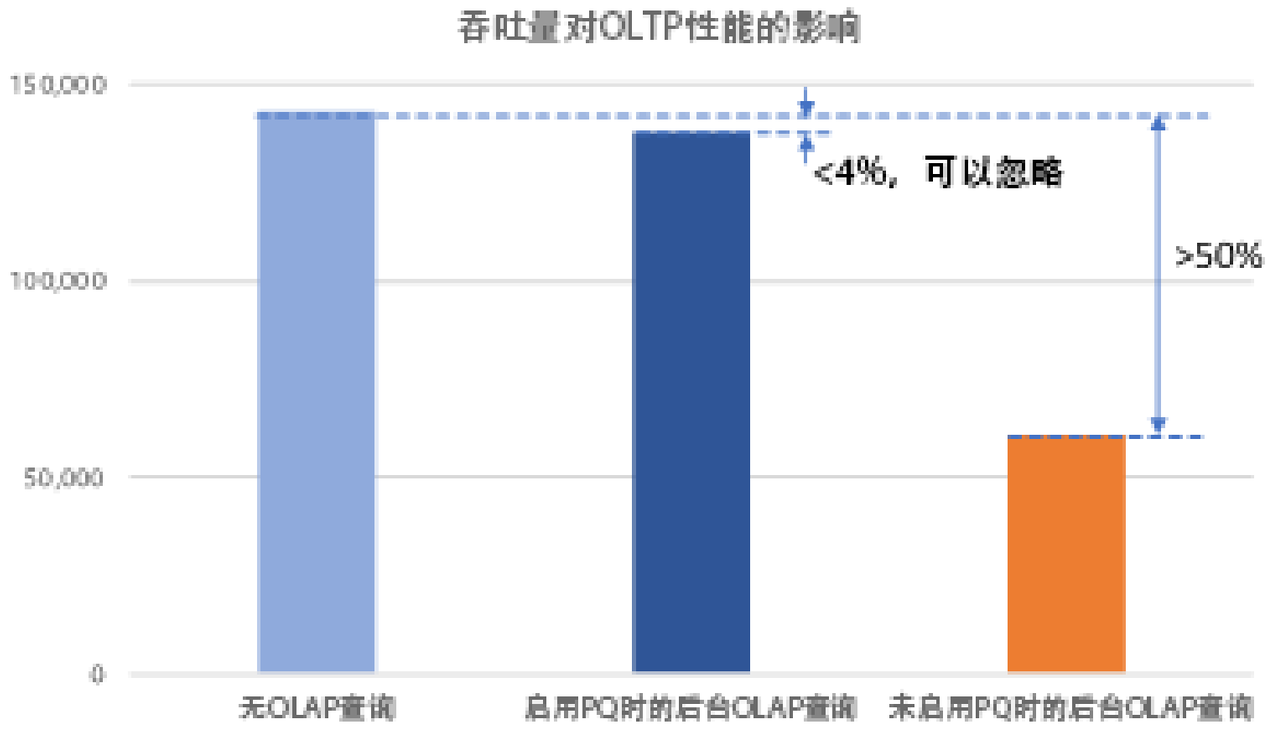
我们对 Aurora 的并行查询功能进行了测试，并发现性能提升效果非常显著。具体来说，我们得以将实例类型从 r3.8xlarge 降低至 r3.2xlarge。对于这一用例，并行查询代表着巨大的性能进步。

云数据架构师 Jyoti Shandil

并行查询——性能结果



已知决策支持基准
数据规模：1 TB，查询#6

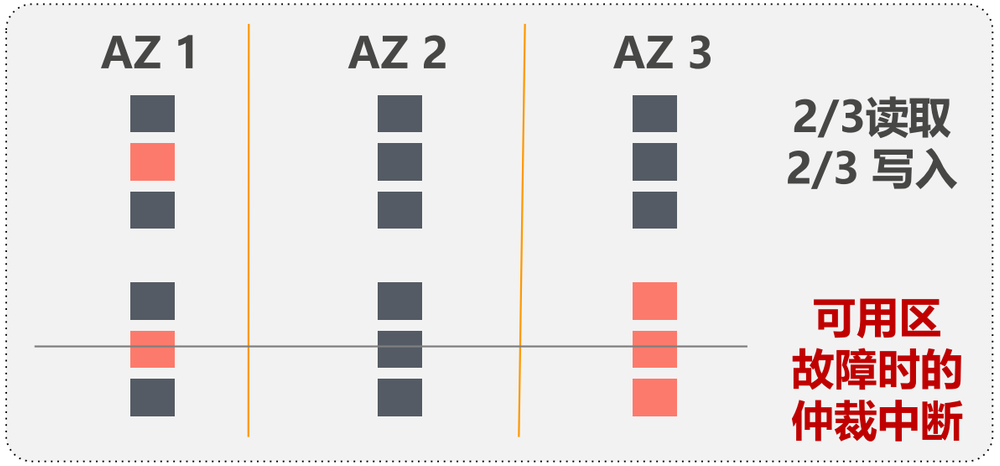


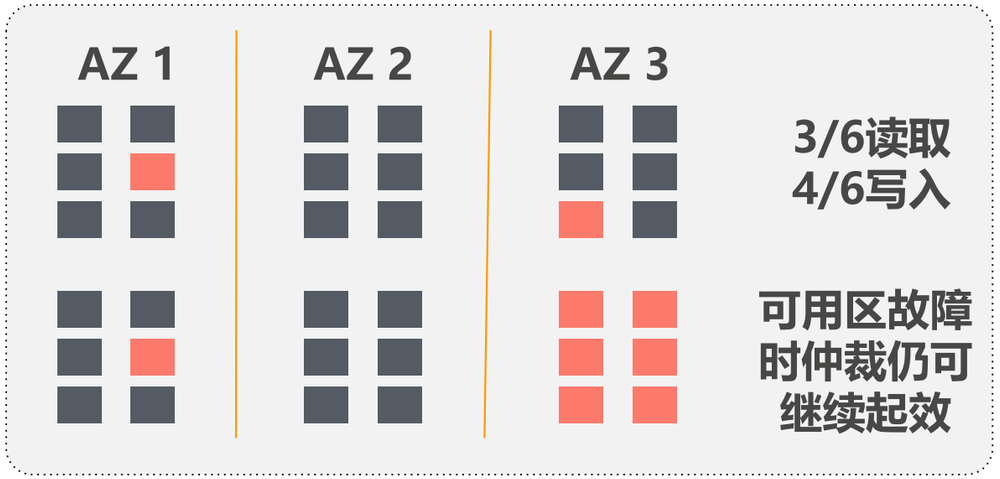
运行在r4.8xl上的Sysbench OLTP工作负载
数据规模：150 GB

可用性与持久性

aws INNOVATE
ONLINE CONFERENCE

“可用区+1”容错机制





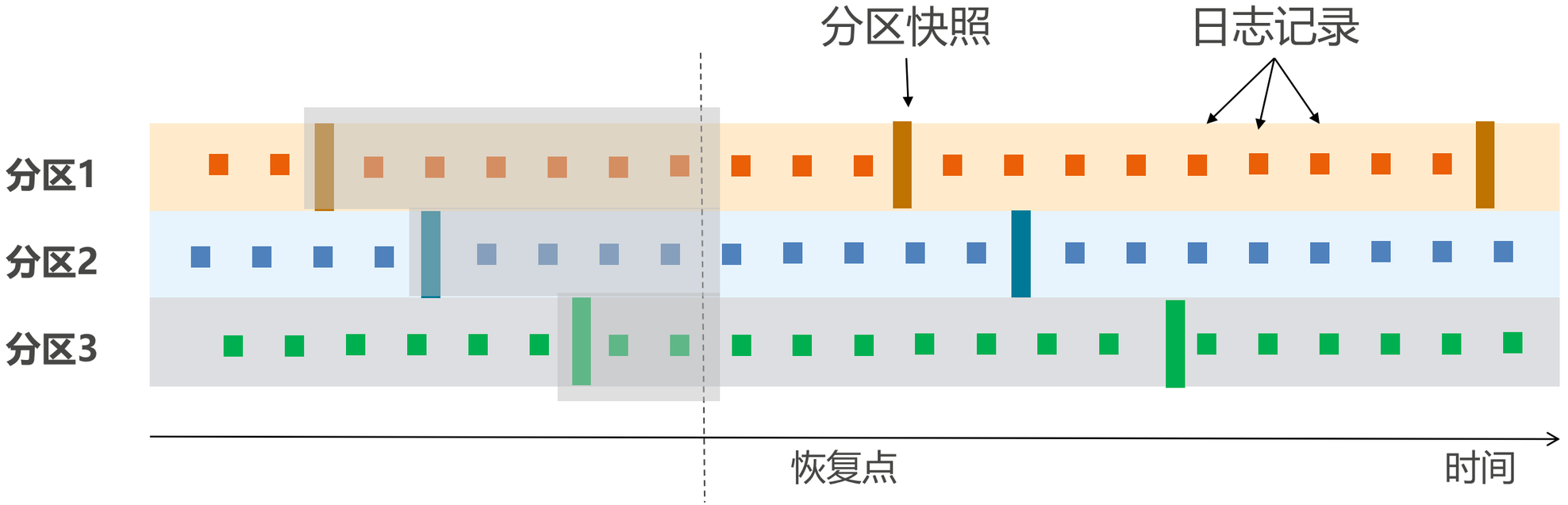
为什么？

- 在大规模集群体系中，故障总会出现
- 可用区故障是一种“命中注定”

如何实现？

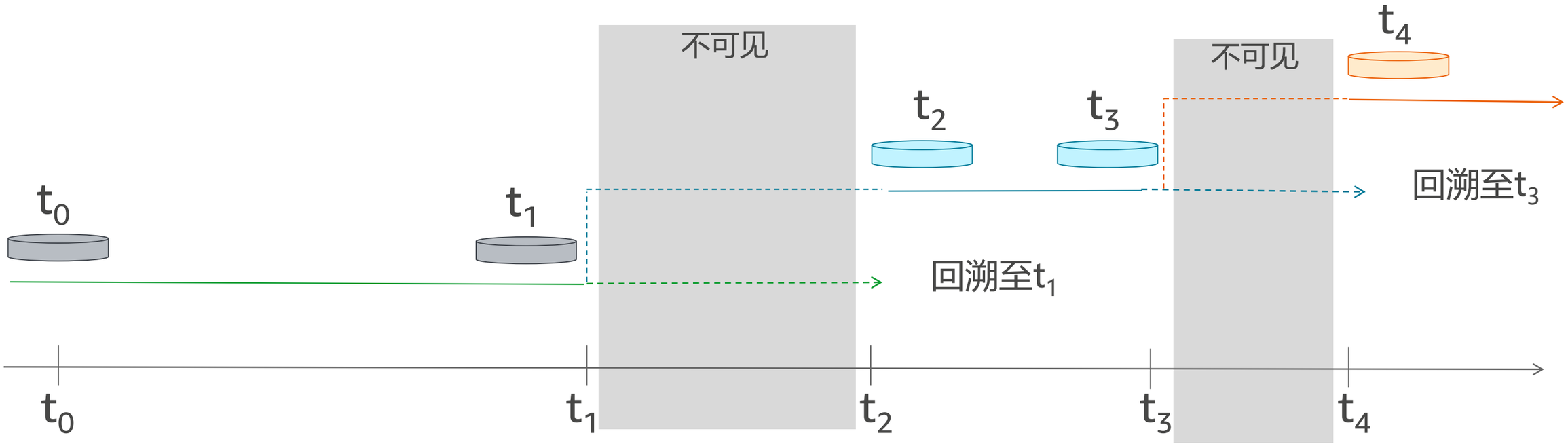
- 6份副本，每可用区2份
- 2/3仲裁无法满足需求

持续备份



- 并行保存每个分区的定期快照，将重做日志引流至 Amazon S3
- 持续执行备份，且不影响性能或可用性水平
- 在还原时，将对应的分区快照与日志流检索至存储节点
- 将日志流应用于并行与异步分区快照

数据库回溯



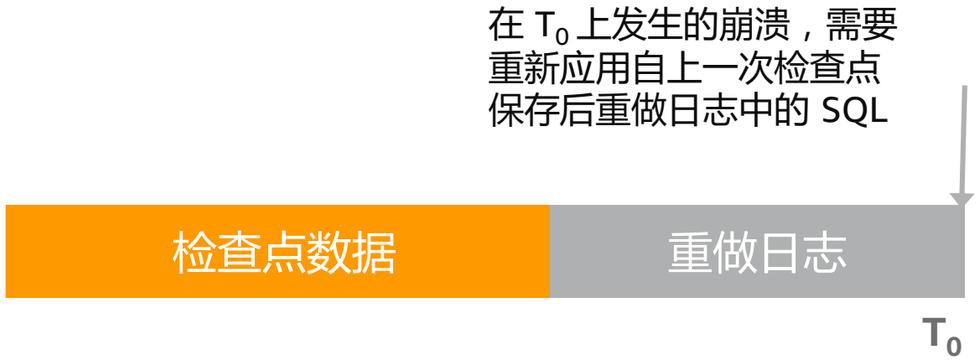
回溯机制代表将数据库恢复至某一时间点，而无需使用备份资源

- 从意外 DML 或 DDL 操作中回溯
- 回溯不具有破坏性，您可以多次回溯以找寻正确的时间点

即时崩溃重做恢复

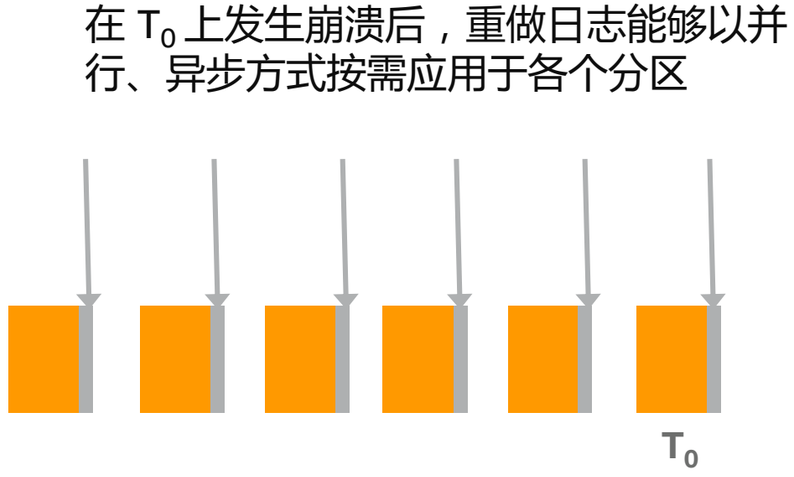
传统数据库

- 重放自上次检查点之后的日志
- 在单线程中慢速重放

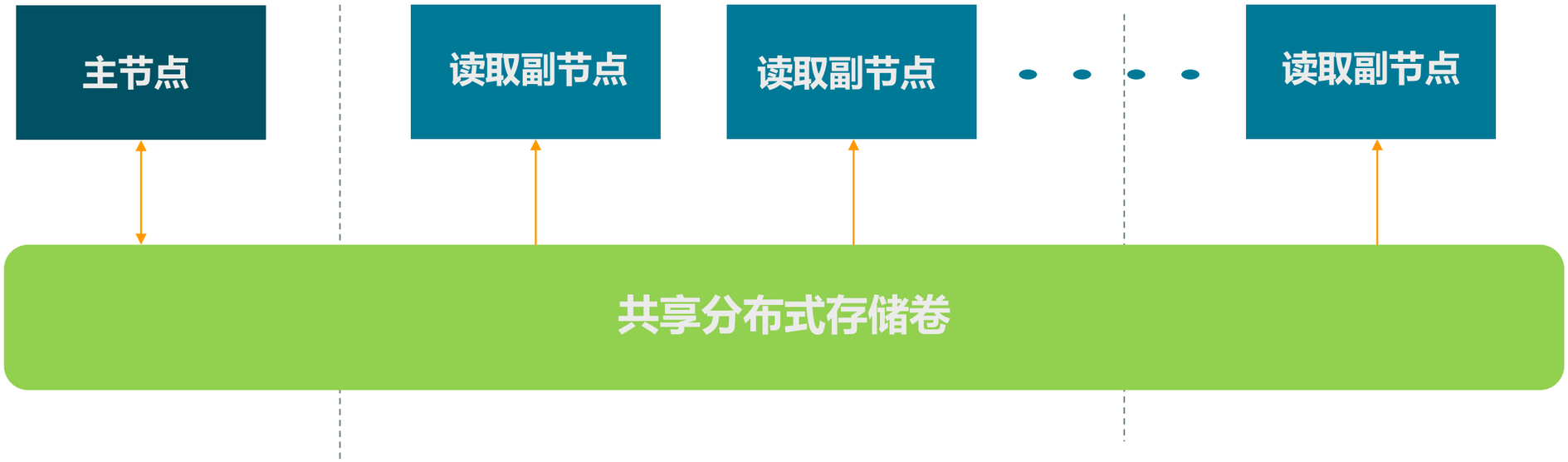


Amazon Aurora

- 无需检查点
- 无需重放启动准备



读取副节点与快速故障转移

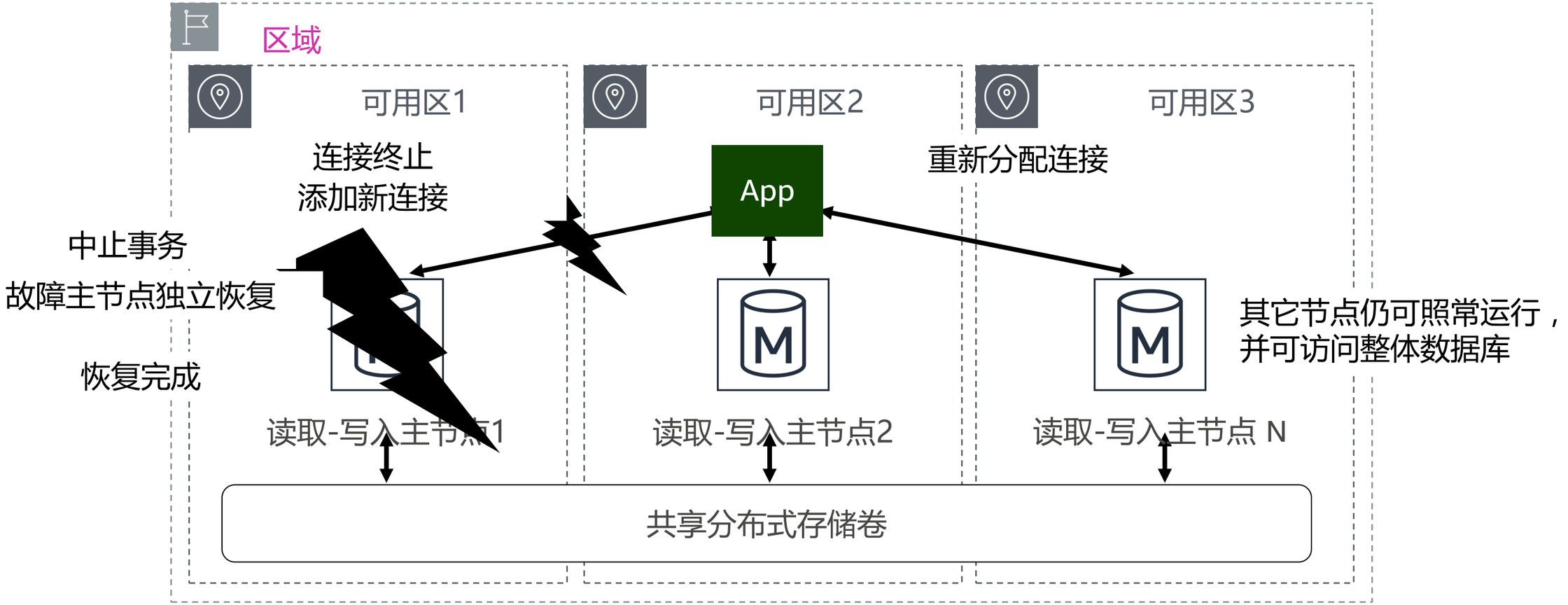


在多可用区内最多可提供15份可升级的只读副本

副节点与主节点共享存储——不丢失数据

可配置故障转移顺序

多主架构持续可用性



- 通过故障与计划内维护实现持续可用性
- 持续监控并自动恢复发生故障的主节点

全局复制

更快的灾难恢复速度与数据位置强化

可管理性

aws INNOVATE
ONLINE CONFERENCE

性能洞察能力 (Performance Insights)

仪表板显示数据库负载

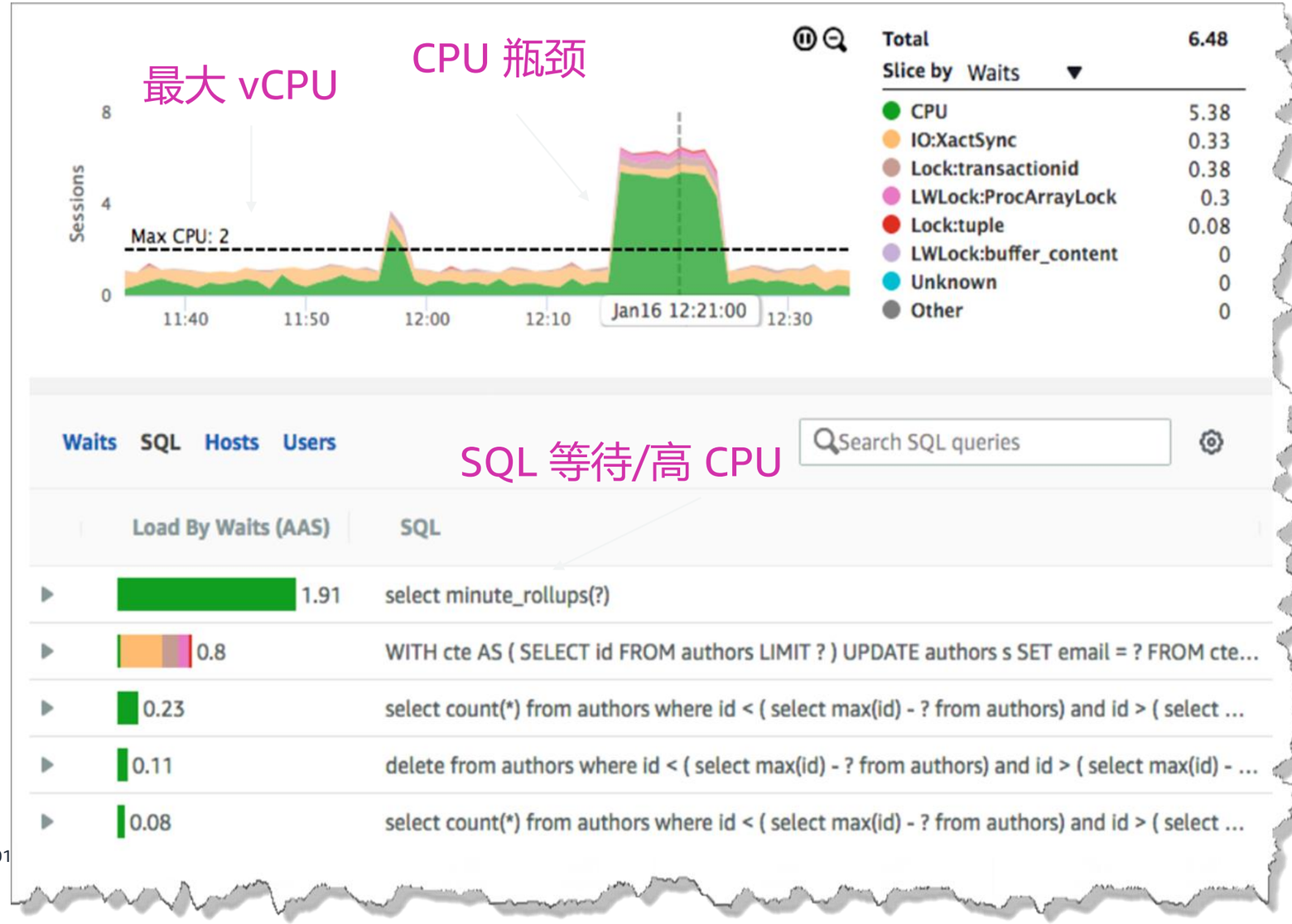
- 操作简单——支持拖放等直观操作
- 强大——可通过放大进行深入观察

确定性能瓶颈来源

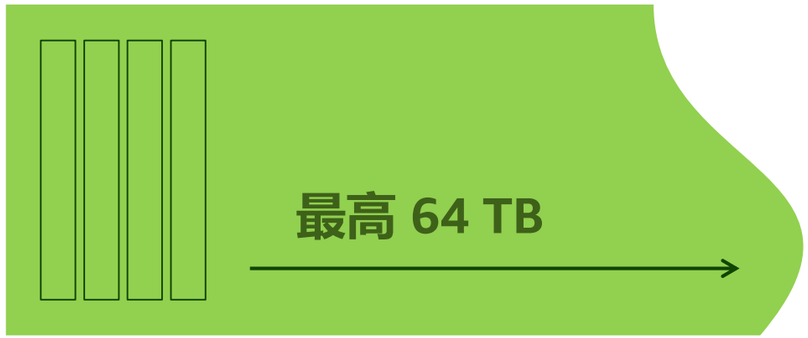
- 按首要 SQL 分类
- 按主机、用户、等待事件等分类

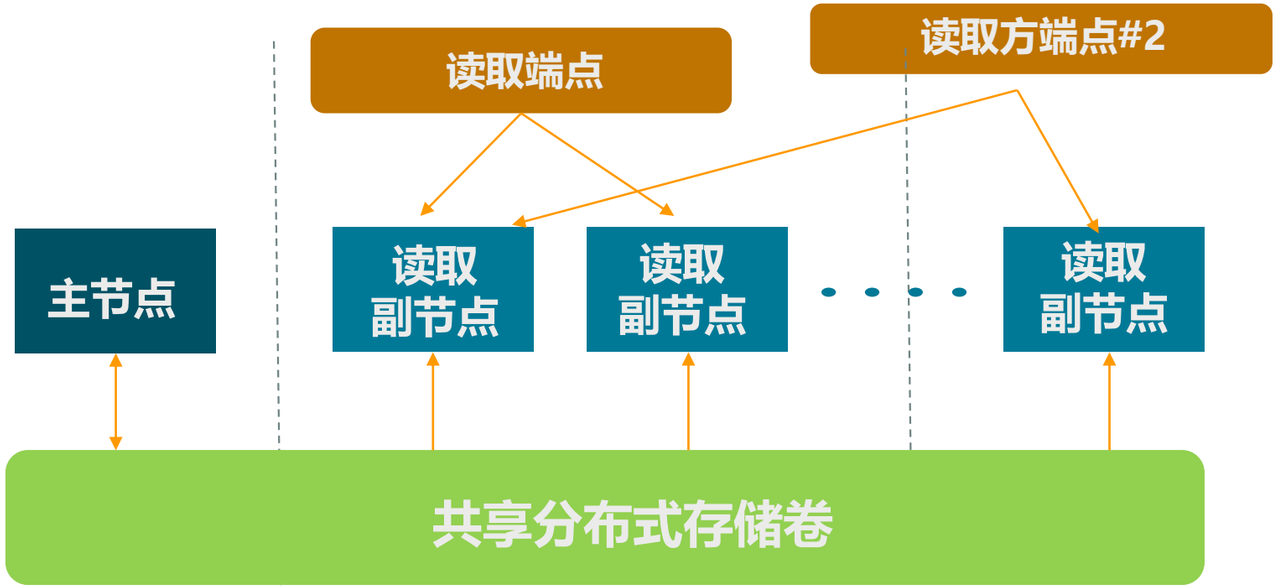
可调节时间范围

- 小时、天、周、月
- 最高保留近2年数据；免费保留近7天数据



简化管理





- 自动存储扩展，最高 64 TB
- 自动实现重新条带化、镜像修复、热点管理以及加密

- 读取端点具有负载均衡功能
- 读取端点自动伸缩 *新功能*
- 定制化读取端点

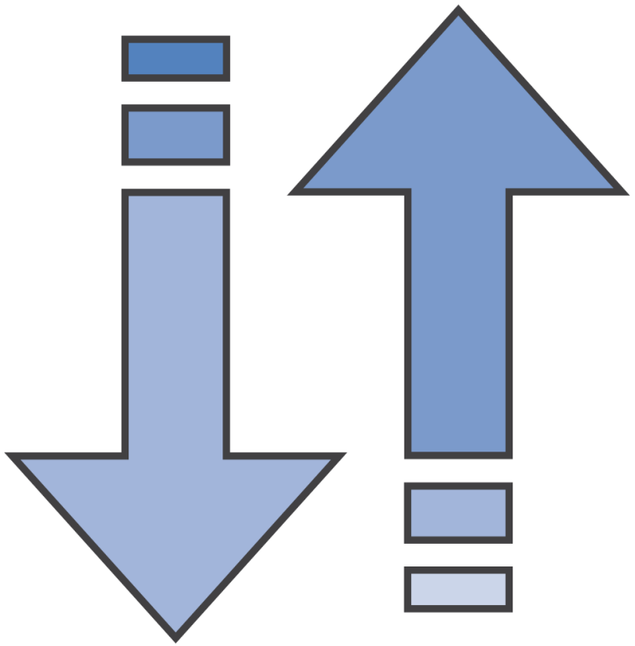
Aurora Serverless . . .

自动响应您的应用程序负载

在10秒以内实现容量规模伸缩

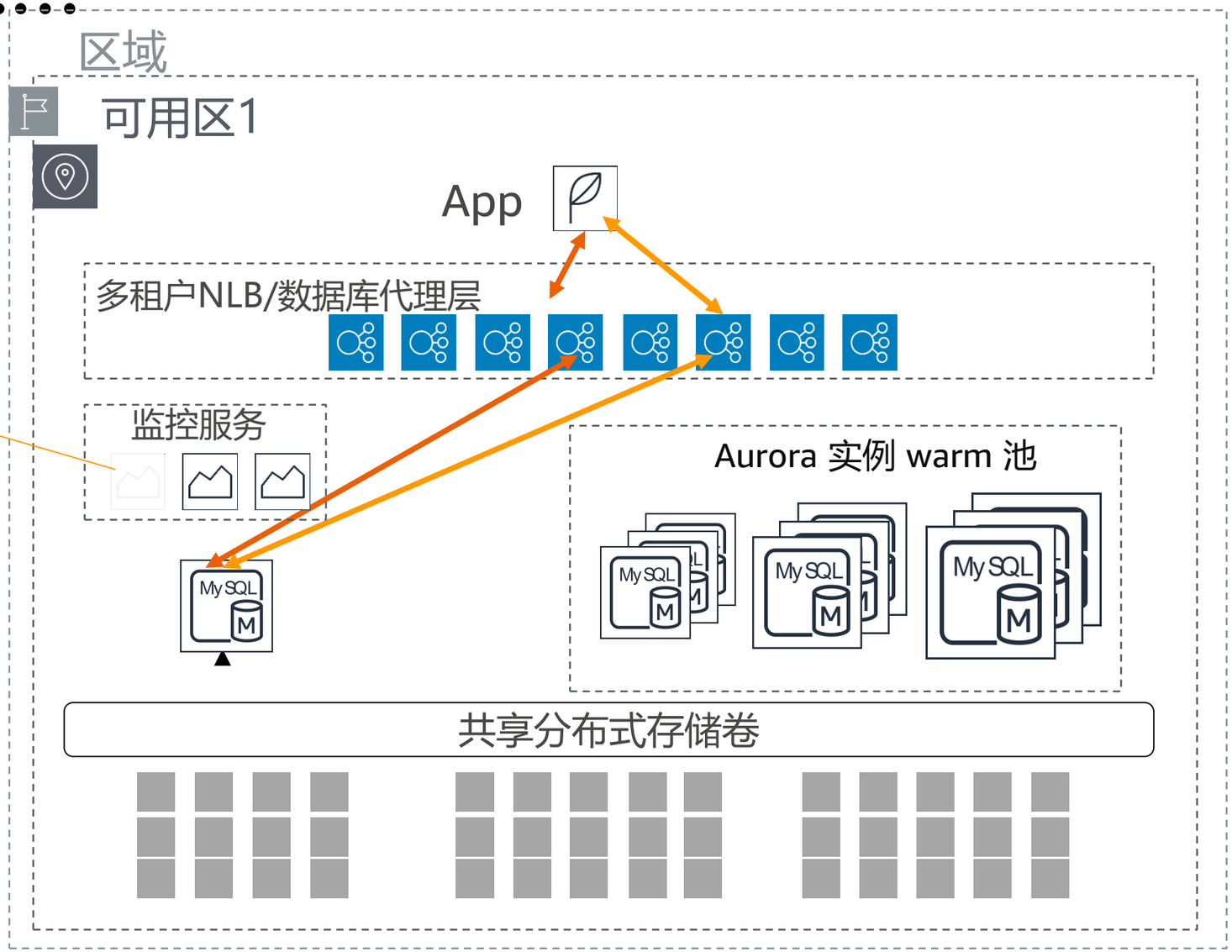
新实例默认配备 warm 缓冲池

多租户代理具有高可用性

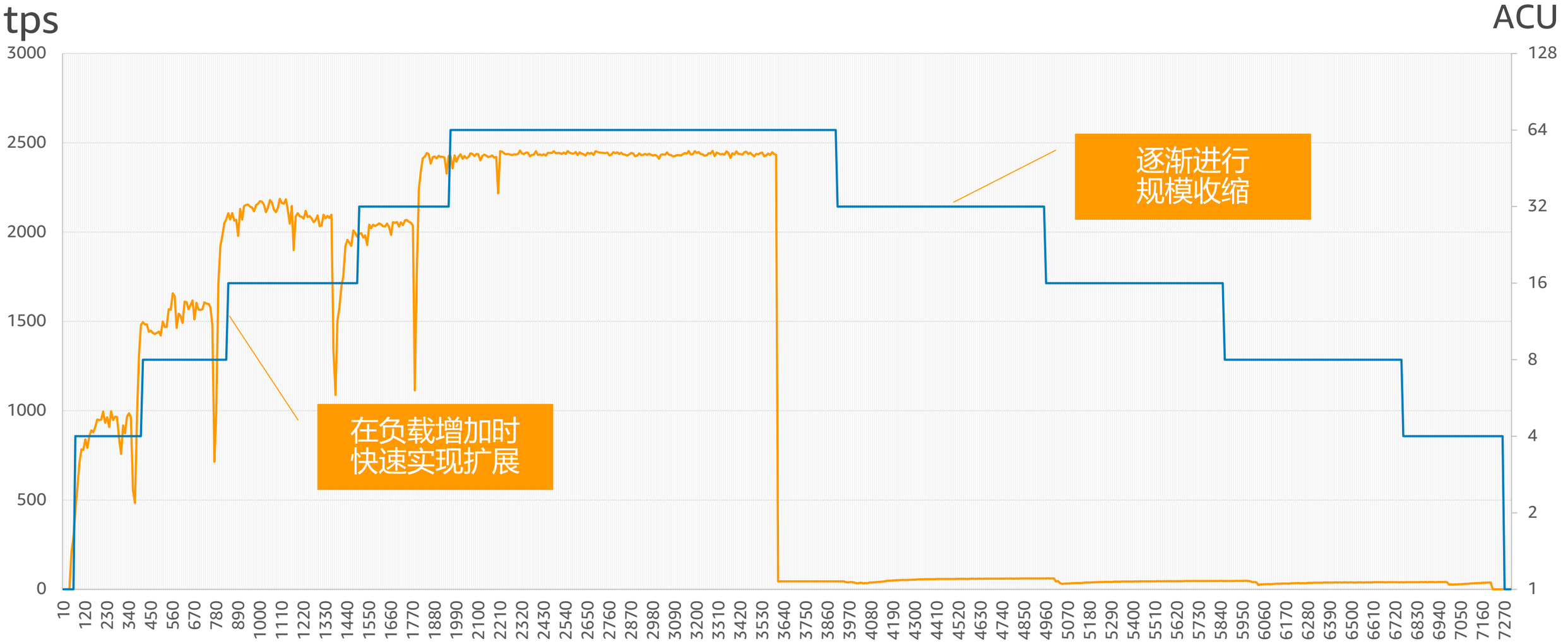


工作原理……

>70% CPU 利用率或者
>90% 最大连接



其在实践当中如何起效？



感谢参加 AWS INNOVATE 2019 在线技术大会

我们希望您在这里找到感兴趣的内容！

也请帮助我们完成**投票打分**和**反馈问卷**。

欲获取关于 AWS 的更多信息和技术内容，可以通过以下方式找到我们：


微信公众号：AWSChina


新浪微博：https://www.weibo.com/amazonaws/


领英：https://www.linkedin.com/company/aws-china/


知乎：https://www.zhihu.com/org/aws-54/activities/


视频中心：http://aws.amazon.bokecc.com/


更多线上活动：https://aws.amazon.com/cn/about-aws/events/webinar/

aws INNOVATE
ONLINE CONFERENCE