

# Rakuten 乐天积分系统从 Cassandra 到 TiDB 的选型与实战

原创

李震

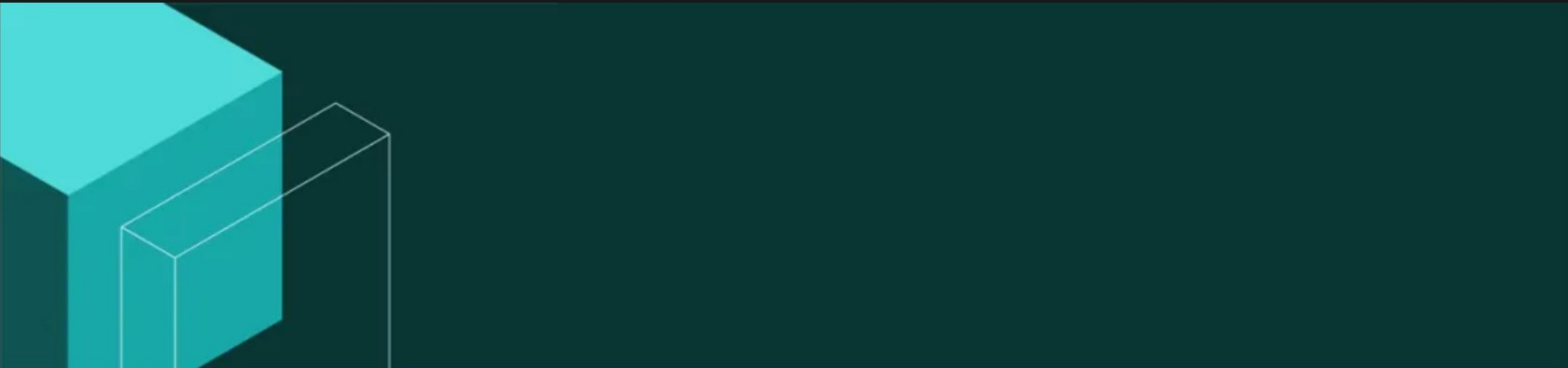
PingCAP

2025年01月02日 19:31

上海



微信扫一扫  
关注该公众号



## ■ 导读

自 2002 年推出以来，乐天积分系统的流量稳步增长，每天发放的积分从数百万到数千万不等。近年来，不仅仅是展示积分，提升用户体验的重要性也日益增加。

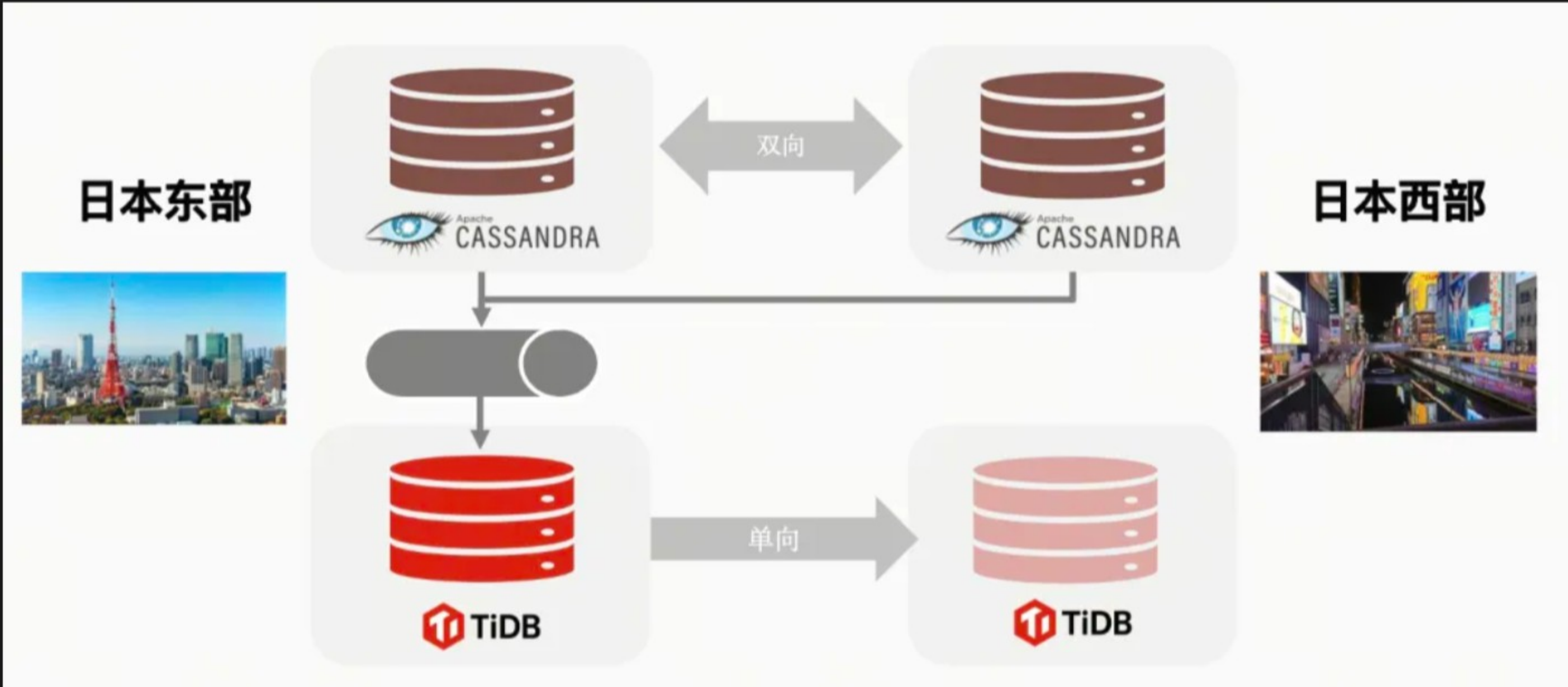
随着数据量的不断增加，挑战在于提供灵活的功能的同时保持成本低廉。本文根据 Rakuten 乐天 INPD 高级软件开发工程师李震在 TiDB 社区活动大连站的分享整理，介绍了乐天选择 TiDB 的思考和实际的应用体验。

# 1 Rakuten 乐天及其积分系统介绍

Rakuten 乐天是一家成立于 1997 年的日本公司，总部位于东京，员工总数超过 30,000 人，业务遍及 100 多个国家和地区。除了电商业务外，乐天还涉及电信、金融等多个行业。

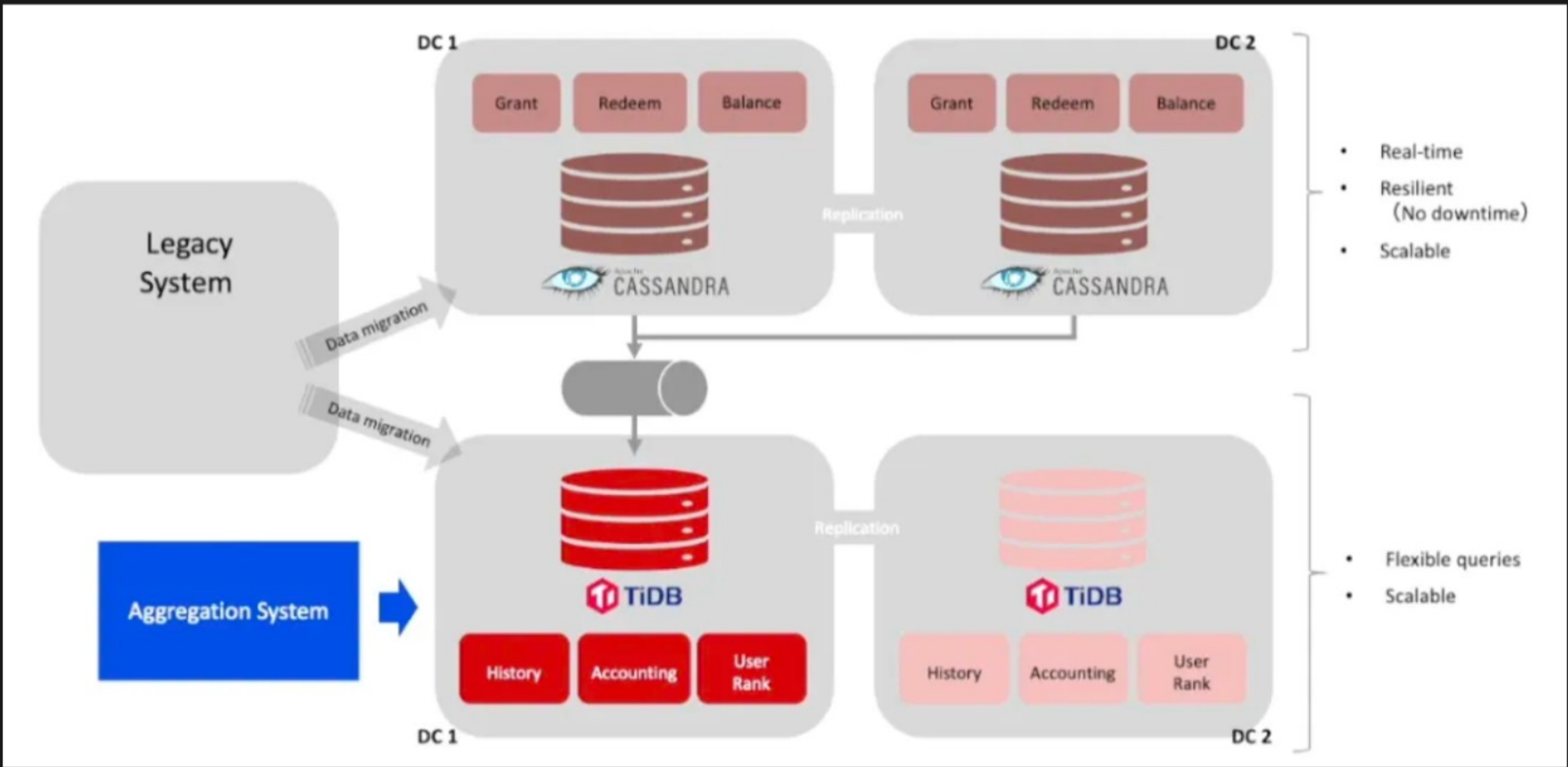
乐天的积分系统在日本非常普及，用户通过使用乐天的服务，如银行卡、手机卡等，可以获得积分，并在乐天商城中使用这些积分进行购物或抵扣，我们每个季度还会举办促销活动、会有大量用户高频访问积分系统，因此，乐天积分服务平台的性能、延迟和可用性要求极高。

乐天的数据中心主要分布在日本东部和日本西部，东部数据中心大致位于东京地区，而西部数据中心位于大阪地区。目前，我们的主要数据库使用的是 Cassandra，通过批量处理周期性地 将主数据库中的数据同步到 TiDB 数据库中。TiDB 目前采用的是主从方案，主数据库位于日本东部，而数据同步则在日本西部进行。



下图展示了乐天正在开发的下一代积分平台的架构图，该平台旨在提供更高性能和低延迟的服务，以满足用户对高可用性的需求。





## 2 Cassandra 的限制

过去，积分系统虽然主数据库使用的是 Cassandra，但作为一款 NoSQL 数据库，它存在一些限制。例如，它对事务的支持较弱，也没有传统关系数据库的关系模型。

### 业务层面 Cassandra 的限制

首先，就开发而言，Cassandra 存在一致性缺失的问题。尽管 Cassandra 宣称其一致性可调，但事实上无法完全确保数据库的一致性。对于积分项目而言，倘若数据一致性无法得到保障，可能会对业务造成影响。此外，Cassandra 对事务的支持力度较弱，致使我们的应用程序为了在 Cassandra 基础上实现分布式事务，需投入诸多额外的精力。与其他 NoSQL 数据库相同，Cassandra 不支持 JOIN 操作，这对于需要灵活查询的业务来说，构成了一种限制。

### 运维层面 Cassandra 的限制

从运维的角度来看，Cassandra 的一致性问题是主要的考量点。这要求我们定期进行数据修复操作，即 data repair，这是一个非常耗时且资源密集的过程。我们需要每天对我们的周数据进行修复。此外，Cassandra 社区生态发展也存在诸多限制。由于我们有时需要对数据库中的数据进行分析以指导决策，Cassandra 在数据集成方面的支持不足。

## 3 探索 NewSQL 解决方案

总结来说，我们对积分系统的数据库要求如下：

- 数据存储**：数据库需要能够存储大量的交易数据和衍生数据。
- 查询能力**：数据库应支持跨多个列的条件搜索，并且查询速度要足够快，以支持在线计算。
- 可扩展性**：随着数据量的增加，数据库应能够扩展而不会导致停机。
- 易用性**：数据库的操作应该简单易学，以便新成员能够快速上手。
- 技术支持**：数据库软件需要持续更新，并且能够提供技术支持。

当时，我们对比了多种基于 MySQL 的分布式解决方案，但这些方案总让人感觉不太舒适，可能是因为 MySQL 最初设计时并非作为分布式数据库而设计，导致后续的分布式解决方案不尽人意。除了 MySQL 的解决方案和 Cassandra 之外，我们也在寻找其他能满足业务需求的解决方案。

后来，注意到了 NewSQL 这一概念。NewSQL 的主要优势在于：

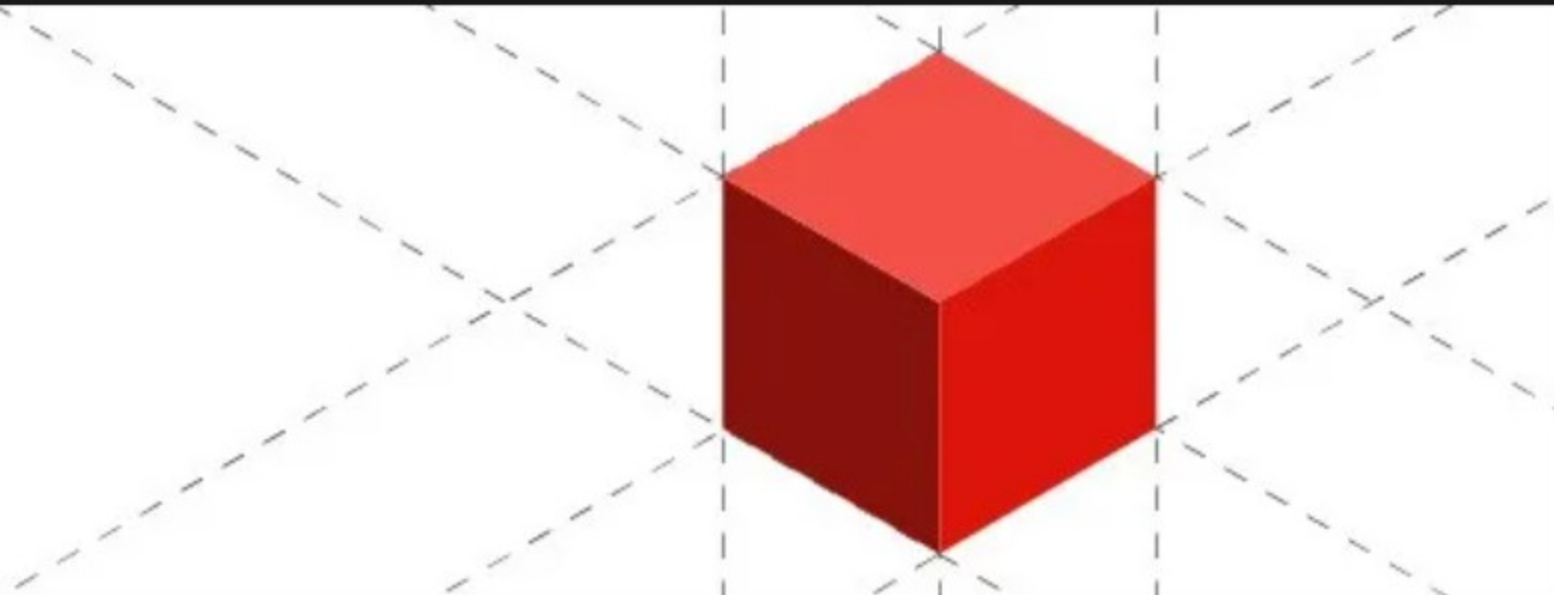


1. **关系模型**：NewSQL 支持关系模型，这与乐天对数据库能够存储大量交易数据和衍生数据的要求相符合。
2. **ACID 特性**：NewSQL 提供 ACID（原子性、一致性、隔离性、持久性）保证，这对于确保数据的准确性和可靠性至关重要。
3. **SQL 支持**：NewSQL 完全支持 SQL，这意味着它可以利用 SQL 的强大查询能力，满足乐天对数据库进行条件搜索和在线计算的需求。
4. **查询灵活性和索引使用**：NewSQL 支持复杂的查询，这与乐天对数据库查询能力的要求相匹配。
5. **可扩展性**：NewSQL 在添加集群时扩展容易，这满足了乐天对数据库随着数据量增加能够扩展而不会导致停机的要求。
6. **分布式数据库**：NewSQL 是分布式数据库，这意味着它可以在多个节点上存储数据，提高了数据的可用性和容错性。

Feature	RDBMS(SQL)	NoSQL	NewSQL
Relational	Yes Relational model	No Key-value storage	Yes Relational model
ACID	Yes	No Applied at application-level	Yes
SQL	Yes	No ⌘Limited support	Yes
Query flexibility Index use	Yes ⌘Supports complex queries	No ⌘Cannot support complex queries	Yes ⌘Supports complex queries
Scaling by adding clusters	Difficult	Easy	Easy
Distributed DB	No	Yes	Yes

我们考察了市场上的一些主要 NewSQL 候选者，包括 CockroachDB、TiDB 等，并进行了横向对比和测试，在众多数据库解决方案中，我们最终选择了 TiDB。

## 4 选型 TiDB 的前期准备



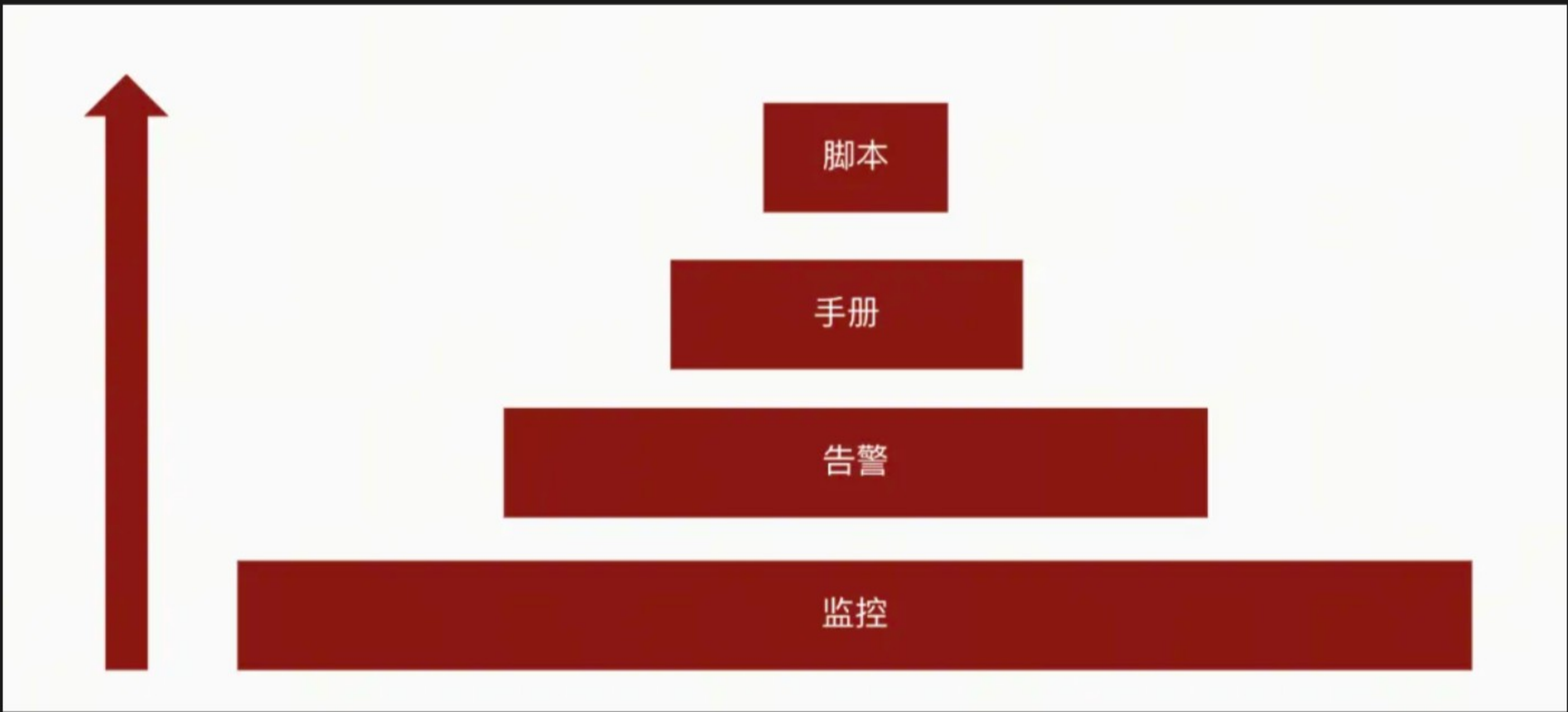
数据库也被称为“真理之源”，如果数据库出了问题，那我们的业务肯定也就不在了。因此，我们选择 TiDB 是基于对数据库的高性能和稳定性的要求。

在引入新技术时，尤其是数据库，必须遵循一个原则：在没有充分理解和测试之前，绝不应将其应用于生产环境。在前期我做了充分的准备：

- **深入学习 TiDB 的官方文档**：作为一款开源数据库，TiDB 的文档不仅支持中文，还提供了优质的英文文档以及日文文档，这对我们团队来说尤为重要。
- **理论与论文调研**：一个优秀的产品不仅需要良好的工程实现，还需要坚实的理论支撑，因此我们也深入研究了 TiDB 的一些论文，还进行了源码层面的学习；
- **测试验证**：我们从理论、性能和运维三个层面对采用 TiDB 的可行性进行了充分的测试验证。首先进行 POC（概念验证），以确保 TiDB 能够满足我们的业务需求；再使用压力测试工具对 TiDB 进行压力测试；最后进行运维测试，类似于混沌工程：QA 团队会随机关闭某些 TiDB 节点，观察系统的表现，并评估如何从故障中恢复。

此外，我们还将 TiDB 集成到现有的监控系统中，设置了指标监控和日志监控。在监控的基础上，我们还建立了告警规则，以便在系统出现问题时，及时将告警信息发送到值班人员的手机上。每个告警信息都附带了描述信息，指明了发现的问题，并提供了一个链接指向我们的操作手册。这样，无论新员工还是老员工，无论是否接触过 TiDB，都可以根据手册一步一步操作来恢复系统。手册中包含了脚本链接或 Jira 链接，通过点击链接、填写相应参数并点击按钮，即可恢复系统。





## 5 自主研发的 TiDB Master-Slave 切换脚本 - TiMS

这里简单介绍一下我们自主研发的一个小工具，名为“TiMS”，与北美咖啡品牌“Tim Hortons”重名。这是一个 TiDB Master-Slave 的切换工具。在阅读了 TiDB 和 TiCDC 的源码之后，我实现了这个工具，用于管理和切换 TiDB 的主从集群。

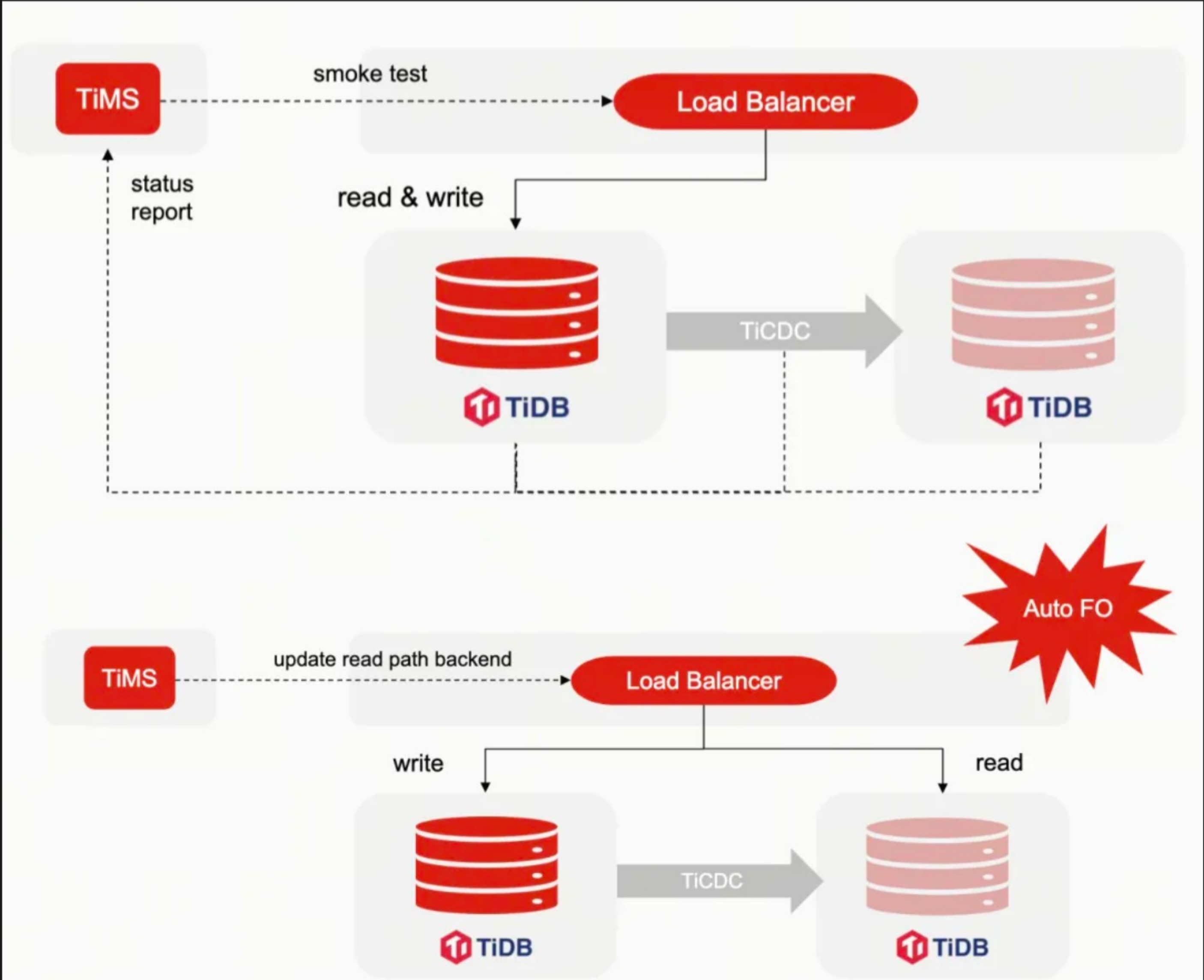
### TiMS 工具的功能

TiMS 首先会通过 Load Balancer 运行一个小测试来确认 TiDB 的主从集群是否正常。因为 Load Balancer 只能探测连接，无法确认数据库是否运行正常。通过 TiDB 的接口，TiMS 可以获取 TiDB 集群的状态信息，检查整个主从集群的状态。

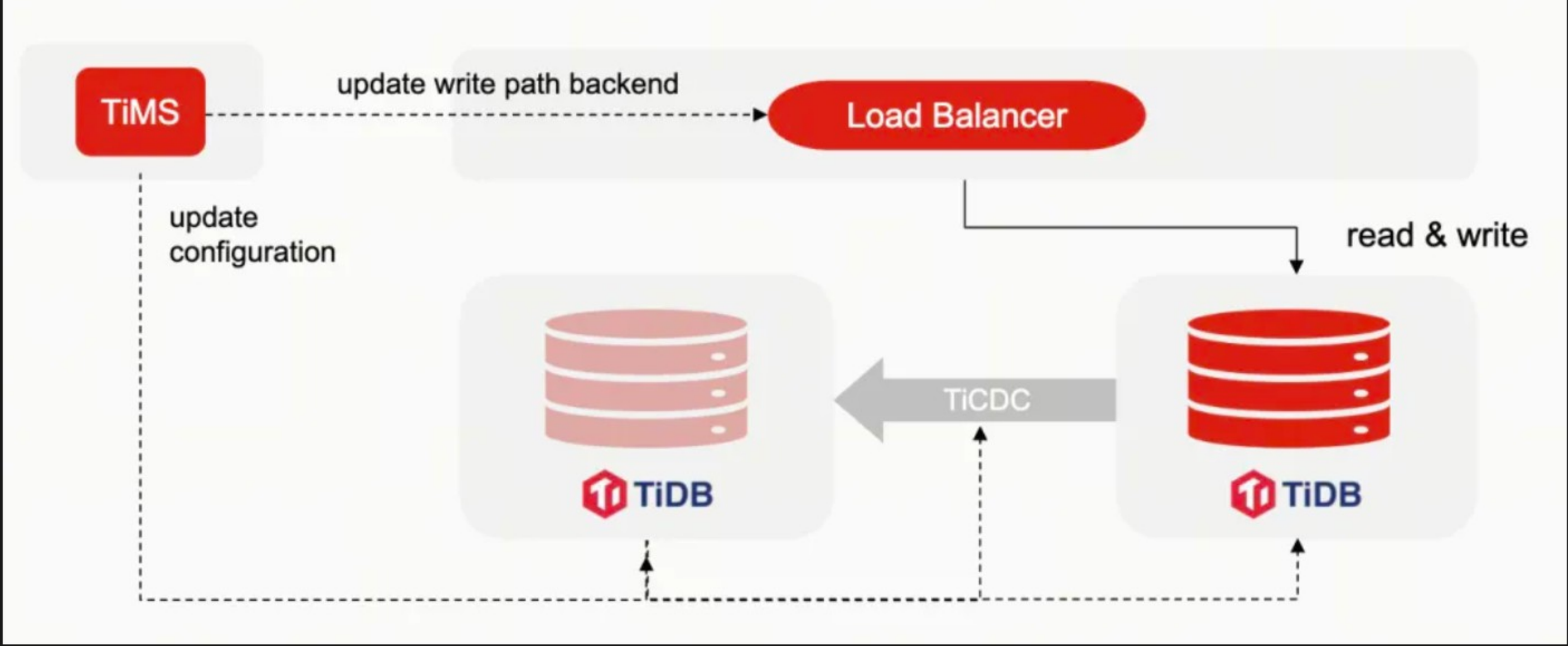
TiMS 还实现了自动故障转移（Auto Failover），当检测到数据库有问题时，它会修改 Load Balancer 的配置，将读流量切换到 Slave 集群上。这个过程是自动的，适用于对延迟不敏感的操作。对于写操作，由于业务要求不能自动执行故障转移，但 TiMS 可以简化故障转移或切换的过程。

### TiDB 主从集群切换流程

首先，我们将 TiDB 的主从集群进行修改，一边设置为可读写，另一边设置为只读。然后停止 TiCDC 的数据同步方向，应用 redo log 将数据传输路径从原来的 slave 转移到 master。接着修改负载均衡器的配置，将读写流量切换到 standby 集群。

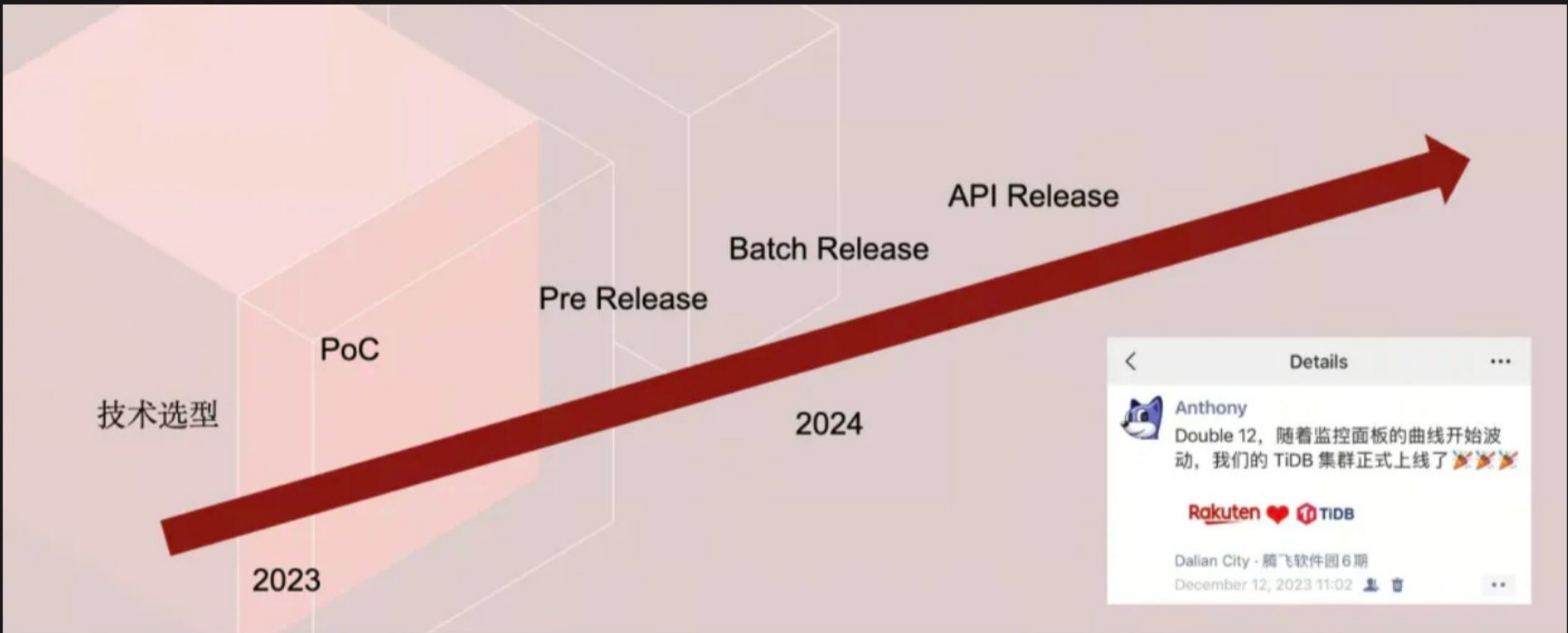







## 6 TiDB 的应用效果

从 2023 年开始，我们进行了技术选型，并经过近一年的 POC（概念验证），才正式将 TiDB 上线。右下角是我朋友圈的截图，显示去年 12 月 12 日，我们的 TiDB 正式 release，至今不到一年。




我们首先上线的是批处理功能，然后逐步上线实时流量。TiDB 在积分应用场景中，主要用于积分的历史统计分析和用户级别的积分记账，这些业务几乎都需要非常灵活的搜索才能实现。




### Point history

Shows when point activity happened and what point activity is upcoming




### Statistics

Point acquisition graph, lifetime points, and per-service point information



### User rank

Calculated based on points earned and other factors



### Accounting

Totals of grant and redeem transactions per service  
Used to generate invoices

以下是一些截图，虽然内容是日文，但可以看到包括统计分析和历史数据在内的信息，这些都是从我们的 TiDB 中查询出来的，并最终展示给终端客户。





# 7 TiDB：无负担的运维体验

最后，我想分享一下我个人对 TiDB 的印象。

我认为 TiDB 是一个没有什么存在感的数据库，但这并不是一件坏事。它的语句与 MySQL 兼容，使用的工具如连接池、Hikari 等也是 MySQL 兼容的。对于运维来说，TiDB 的批量处理吞吐量高，实时流量延迟低，即使部分节点到期也能保持高可用性，为用户提供服务。我可以像使用 MySQL 一样使用 TiDB，同时无需考虑扩缩容等需求，没有任何负担。


/ 相关推荐 /



### B 站 x TiDB：助力海内外业务增长，核心场景百套 TiDB 应用实践

行业实践

B 站数据库负责人赵月顺：助力海内外业务增长，百套 TiDB 的选型与运维实战



### 知乎 PB 级别 TiDB 数据库集群管控实践

行业实践

知乎 PB 级别 TiDB 数据库集群管控实践

💡 点击文末 **阅读原文**，立即下载试用 TiDB ！





订阅号  
@PingCAP  
点击关注



服务号  
@TiDB Club  
点击关注



视频号  
@PingCAP  
点击关注



PingCAP 官网  
pingcap.com  
点击访问



Bilibili  
@TiDB\_Robot  
点击关注



知乎  
@PingCAP  
点击关注



微信搜一搜



TiDB

[阅读原文](#)