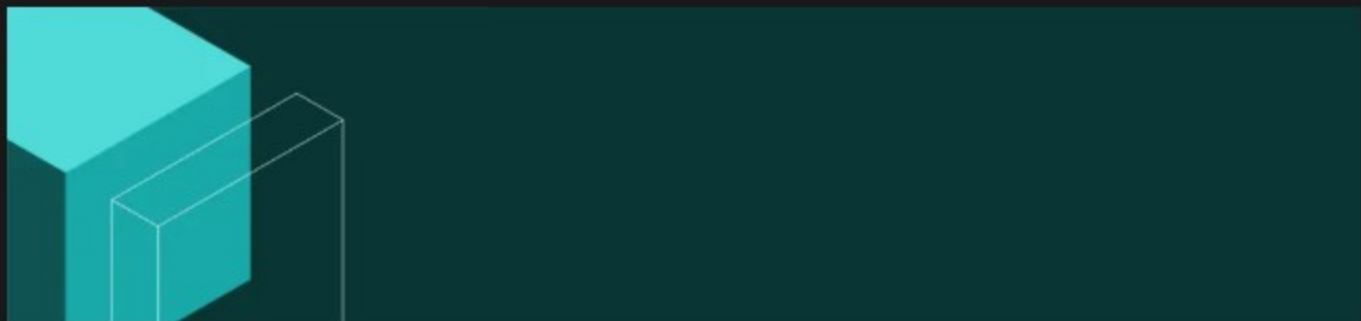


微众银行：大规模 TiDB 运维体系建设 & 金融级稳定性保障漫谈

原创 黄蔚 PingCAP 2025年04月29日 19:30 北京



■ 导读

微众银行自 2019 年起引入 TiDB，逐步构建起以分布式架构为核心的金融级数据库平台，支撑了贷款、理财、风控等多条关键业务线。在业务快速发展的同时，如何保障系统的高可用性与运维效率，成为数据库团队的核心课题。通过构建自研管理平台、容量中心、自动巡检与智能诊断体系，微众银行实现了大规模 TiDB 集群的稳定运行与持续演进。

本文介绍了微众银行在应用 TiDB 开展运维体系建设及稳定性保障的经验。

以下是来自微众银行数据库团队的黄蔚老师在 TiDB 社区深圳站交流活动中的现场分享实录。

1 TiDB 在微众银行的场景演进

微众银行自 2019 年开始引入 TiDB，以解决传统单机架构在全局管理单元中所暴露的性能瓶颈问题。随着业务快速增长，系统在容量、并发、灵活性等方面面临挑战，经充分评估多种分布式数据库解决方案后，微众银行最终选定 TiDB，并逐步扩大其应用范围。



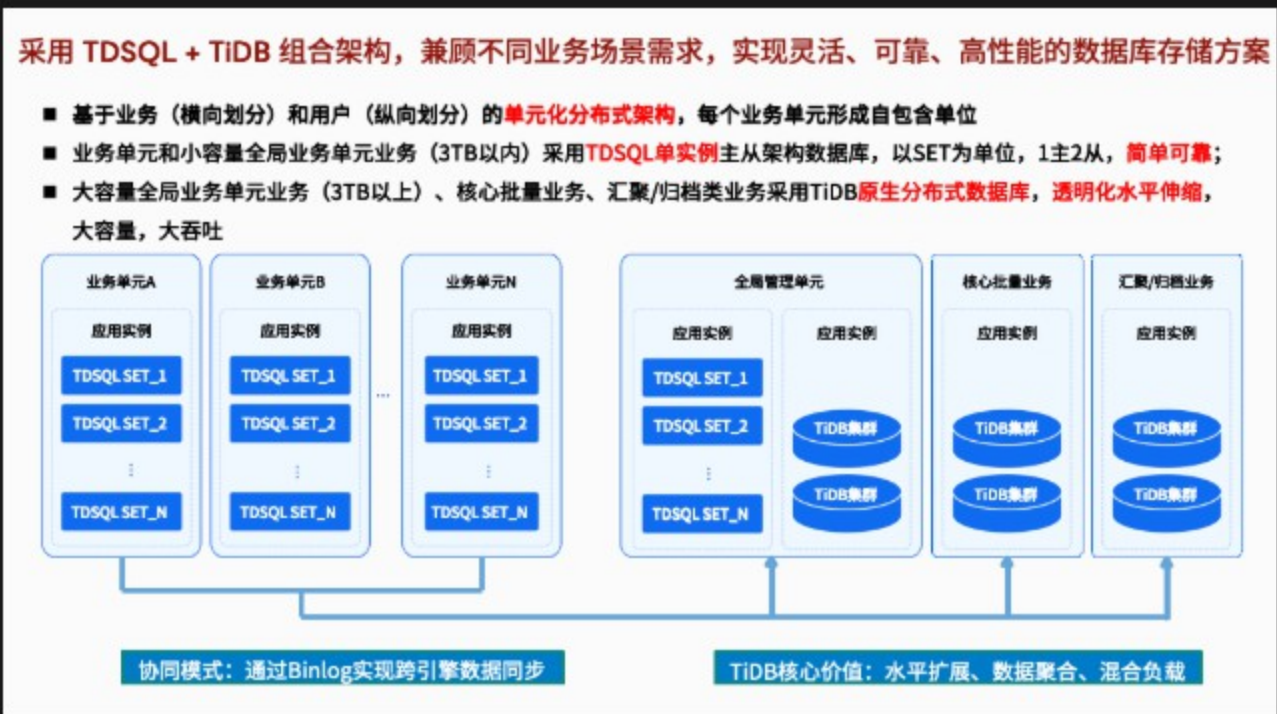
TiDB 从 2.0 到 8.0 的快速演进，为微众银行在多个关键业务场景中提供了强有力的技术支持：

- **TiDB 2.0** 提供良好的水平扩展能力，解决了早期业务扩容需求；
- **TiDB 4.0** 引入乐观 DDL 和 DM 2.0，支持多源数据的汇聚与归档，显著提升批量处理效率；
- **TiDB 5.4** 实现单集群 200TB 以上容量的支持，满足大规模归档与分析型业务场景；
- **TiDB 6.5** 增强金融级事务处理能力，通过悲观锁和 RC（Read Committed）隔离级别保障账务系统的一致性要求；
- **TiDB 7.0/8.0** 强化向量化执行引擎与多租户支持，更好服务于敏态业务与资源复用需求。

得益于 TiDB 的持续演进，微众银行在数据库架构层面获得了显著的灵活性与扩展性支持，支撑了贷款、存款、理财、财富管理等多个业务的数字化发展。TiDB 被广泛应用于联机交易、批量处理、实时归档、风控、反洗钱、客服支撑等多元业务场景，已成为微众银行一体化数据架构的重要组成部分。

在实践中，微众银行也与 TiDB 的核心开发团队 PingCAP 建立了紧密的技术合作关系。一方面，业务需求推动了新功能的落地与稳定性提升；另一方面，TiDB 的架构优化和新特性，如多租户支持、向量化处理能力等，反哺了业务的敏捷性和创新能力，实现了数据库与业务发展的良性共振。

统一架构设计：TDSQL + TiDB 双引擎组合



为了应对多元化业务对数据存储和处理能力的不同要求，微众银行在数据库架构上采用了 **TDSQL 与 TiDB 结合的双引擎方案**。这一架构设计在兼顾稳定性与扩展性的同时，实现了对各类业务场景的灵活支持。

系统整体遵循“单元化分布式架构”设计理念，基于业务类型和数据规模进行横纵向划分：

- **小容量业务单元（< 3TB）**以及小规模全局管理单元，采用 TDSQL 单实例主从架构部署。此类架构简单、可靠，适用于轻量级联机交易等场景；
- **大容量业务单元（> 3TB）**，如核心账务批量处理、归档、数据汇聚类场景，则使用原生分布式的 TiDB 集群。TiDB 的水平扩展能力使其更适合支撑高并发、大吞吐、海量数据的处理需求。

这一组合架构实现了架构设计上的“因需而定”，同时在业务发展过程中展现出显著价值：

- **业务开发维度**：TiDB 支持透明的水平扩展，开发人员无需进行分库分表改造，显著降低系统演进成本；
- **运维管理维度**：在需要扩容或缩容时，TiDB 能实现对业务透明、无侵入的在线变更，避免了传统单体数据库中常见的拆分与迁移操作；
- **架构演进维度**：业务增长阶段可直接扩展现有 TiDB 集群，无需重新进行单元化拆分或重构，降低了架构演进复杂度与风险。

截至目前，微众银行内部已部署多个大型 TiDB 集群。其中，**单集群最大吞吐量达到 237K QPS，数据容量突破 200TB**，充分验证了 TiDB 在处理核心级别负载方面的可行性与稳定性。

2 大规模运维挑战和稳定性保障

随着 TiDB 在微众银行内部应用规模的持续扩大，数据库运维体系也面临着新的挑战。特别是在金融行业这一对稳定性要求极高的场景中，构建一套符合“金融级”标准的数据库架构，成为保障系统可靠性的核心课题。

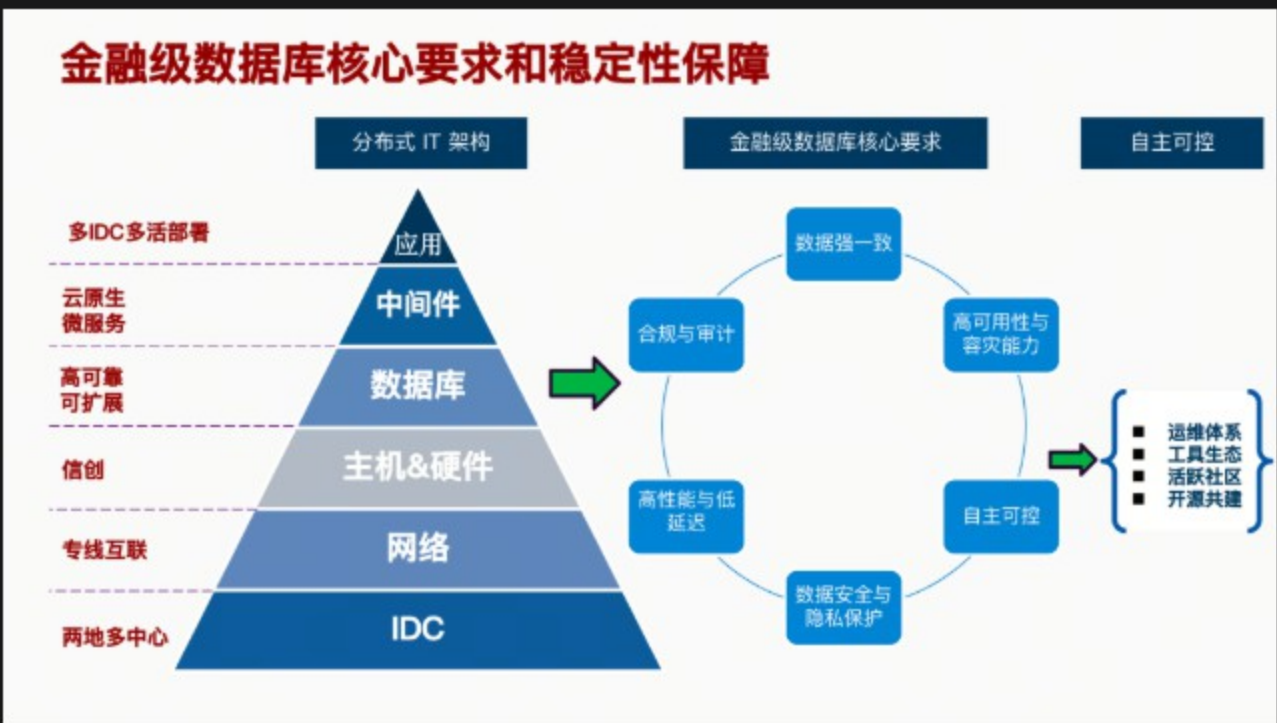
什么是“金融级”数据库？

“金融级”并非仅仅是一个营销术语，其背后代表着一整套可量化、可落地的高可用性标准与系统性能能力。例如：

高可用性目标如 99.999%（五个九）；数据恢复点目标（RPO=0）需满足 等于 0 故障恢复时间目标（RTO）需要尽可能缩短至秒级。

要达成这些目标，必须依赖于覆盖数据中心、网络、存储&计算、数据库、中间件与应用架构各层的协同设计。

金融级数据库核心要求和稳定性保障



以微众银行为例，其金融级数据库系统的设计在各层面体现如下要求：

- **网络层**：为保障数据强一致性，所有 IDC（数据中心）之间通过 专线互联，确保网络时延毫秒级，且稳定可靠，以满足高频交易场景下的数据传输要求；
- **数据库层**：要求具备**强一致性、弹性扩展、高可用**等特性，满足核心账务系统对交易精确性和可靠性的严苛要求；
- **中间件层**：基于单元化架构的服务间交互需依赖于高可靠的消息总线，确保交易流程跨系统联动时的稳定性；
- **应用层**：部署需支持**多 IDC 多活架构**，在城市级或节点级故障发生时，能够实现快速切换和服务连续。

此外，数据合规、安全性、隐私保护，以及**低延迟、高性能**的访问能力，都是现代金融系统所必备的能力基础。

要构建真正意义上的“金融级”分布式数据库体系，仅依赖功能特性远远不够，更需从以下四个维度实现可持续发展：

1. **运维体系完善**：在小规模阶段可依靠人工运维，但在系统复杂度提升后，必须构建完整的自动化运维体系，包括部署、监控、扩缩容、诊断、升级等全生命周期能力。
2. **生态系统成熟**：新引入的数据库系统必须兼容原有生态，如 MySQL 协议支持、备份恢复工具、告警系统等，确保平滑迁移与一致性运维流程。
3. **社区活跃度**：活跃的技术社区为使用方提供经验借鉴与问题交流平台。例如，TiDB 的 TUG 社区提供了丰富的实战案例与技术交流，有助于加快问题定位与解决。
4. **开源共建机制**：开源是实现技术透明与深度可控的重要基础。开源数据库不仅支持源码级定制与调优，还能通过与厂商合作共建工具链，提升平台能力，降低依赖风险。

通过在上述四个方面的持续投入，微众银行实现了以 TiDB 为核心的分布式数据库平台在金融场景下的高可用性与稳定性保障，确保了复杂业务系统的长期可持续发展。

TiDB 在微众银行的使用规模迅速增长

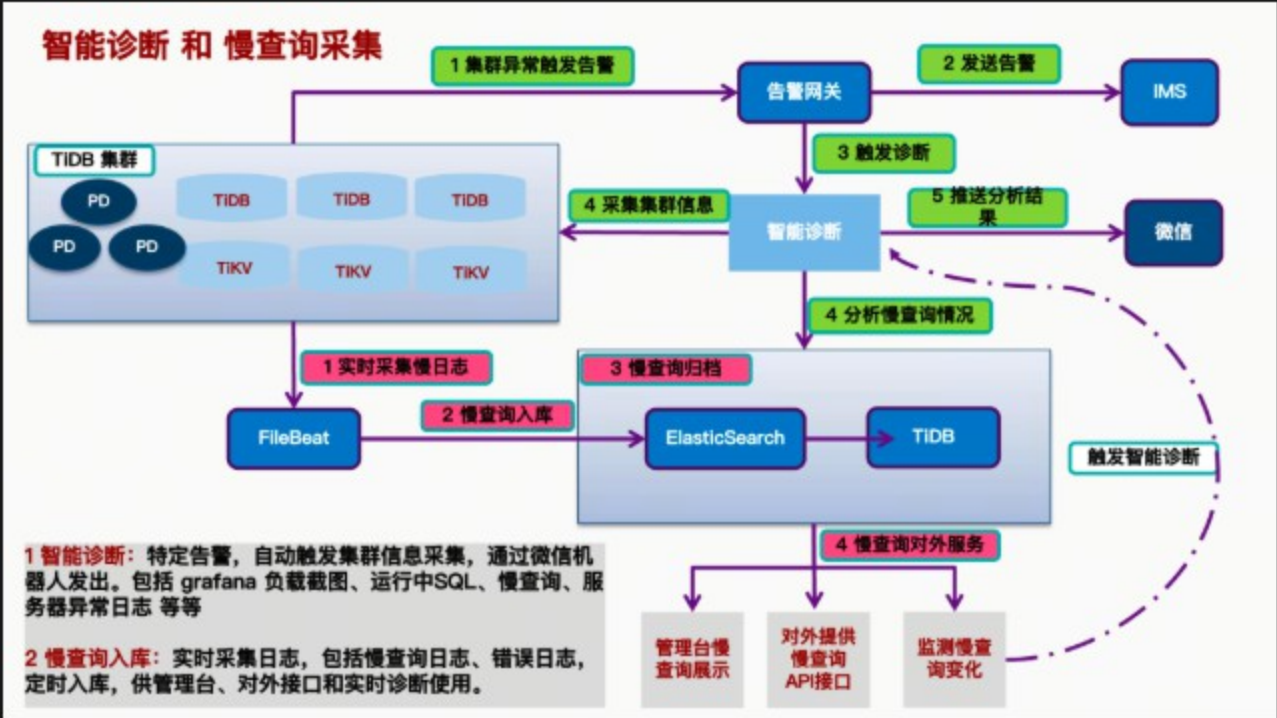




智能诊断和慢查询采集

在实际运维过程中，智能诊断与慢查询采集功能显著提升了微众银行数据库运维的效率。以早期版本为例，当 TiKV 的内存使用率达到 80% 时，系统可能会触发告警。过去在处理此类问题时，需要先连接 VPN，再登录跳板机，通过相应工具查看 TiKV 的日志，操作流程较为繁琐。随后，还需分析是否由于某条 SQL 查询导致内存占用过高，例如是否存在未命中索引等问题，并据此进行进一步处理。整个过程步骤较多，排查路径较长，效率相对较低。

通过引入智能诊断与慢查询采集机制，相关信息能够集中展示和自动采集，运维人员可快速定位问题原因，从而缩短故障处理时间，提升响应效率。



为应对分布式数据库中告警响应和问题排查效率不高的问题，微众银行自研了智能诊断和慢查询采集功能，显著提升了运维处理的自动化和精度。

智能诊断

针对常见的资源告警场景，系统上线了自动化的智能诊断机制。当告警触发时，系统会自动采集相关机器的信息，包括 Prometheus 中的性能视图、服务器上的错误日志，以及 TiKV 节点的关键运行状态。

系统会自动识别日志中的异常关键字，并对 TiKV 组件进行初步分析。采集完成后，所有性能视图和基础信息会被统一整合，并通过微信机器人推送给相关运维人员。这一机制实现了告警与初步诊断结果的同步送达，帮助运维人员快速定位问题所在。

慢查询采集

慢查询采集功能的开发，源于早期版本在慢查询日志处理上的性能问题。当时通过 Dashboard 查看慢查询，但由于 Dashboard 与 PD 模块耦合较深，且系统是实时从本地日志中加载慢查询信息，在日志量较大时，可能会引发 OOM 或网络异常问题。

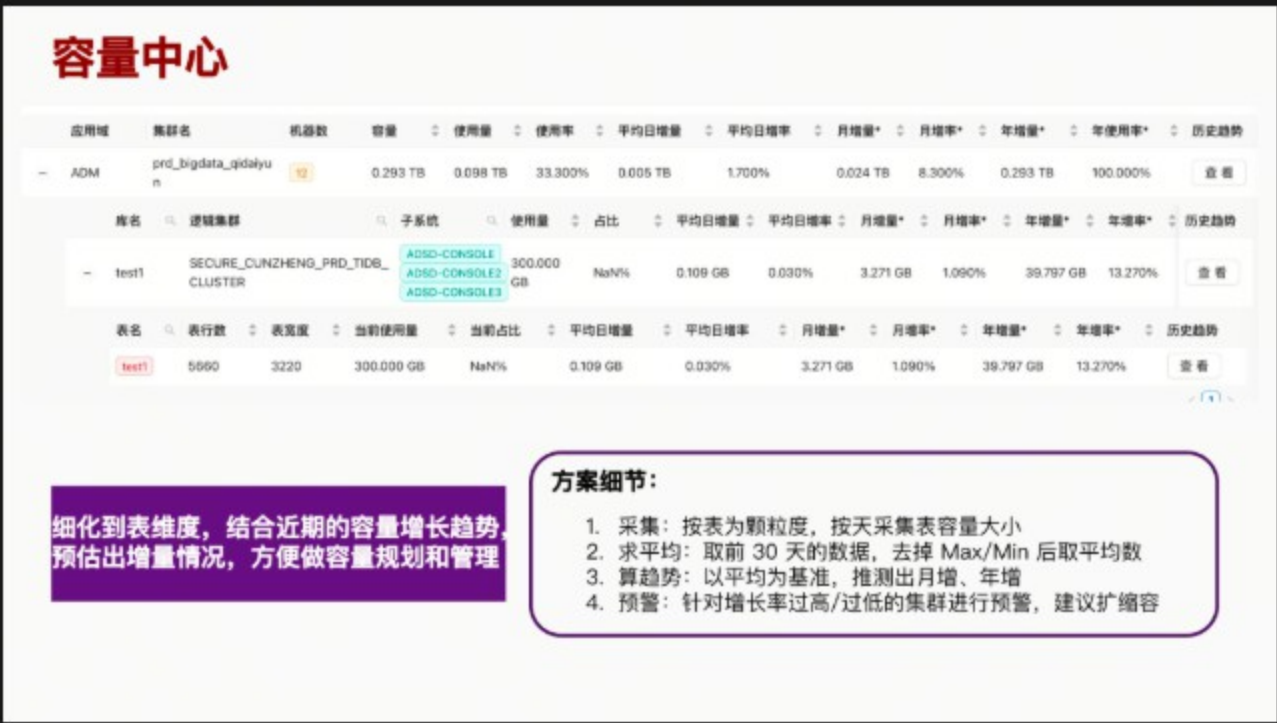
为此，微众银行将慢查询采集功能进行了独立建设。系统通过 ELK（ElasticSearch + Kibana）架构，实时采集 TiDB 节点上的慢查询日志，完成预处理后写入 ElasticSearch，再进行聚合分析，最终将结果归档到 TiDB 中。

这些数据最终可通过多个渠道提供服务：

- 管理平台支持按条件查询慢查询日志；
- 对外提供接口，供运维团队或行内智能化运营系统调用，用于系统运行状况分析与故障根因定位；
- 同时具备告警联动能力：当系统检测到慢查询数量在一定时间段内超过阈值时，会触发告警并启动诊断流程，辅助判断是否由负载过高或服务器异常引起。

通过这两个工具，微众银行在慢查询定位与资源异常处理方面实现了流程自动化和响应加速，进一步提升了大规模数据库环境下的运维效率。

容量中心



容量中心功能主要用于支持 TiDB 集群的容量管理。在此功能上线之前，扩容往往缺乏充分的数据支撑，业务团队通常只是基于当前压力提出扩容需求，缺乏对长期趋势的判断。

引入容量中心后，系统可以通过算法预测未来 30 天或一年内的容量增长率。如果检测到增长率异常偏高，可能说明业务在数据生命周期管理方面存在问题，例如未定期清理历史数据。此时，运维团队会建议业务方进行分区表操作，如添加或删除分区，以控制数据增长。

若增长确实源于业务发展带来的数据增加，则会根据预测结果提前规划扩容需求。通过这一机制，容量中心不仅提升了容量管理的主动性，也为后续的资源优化和降本工作提供了数据支持。

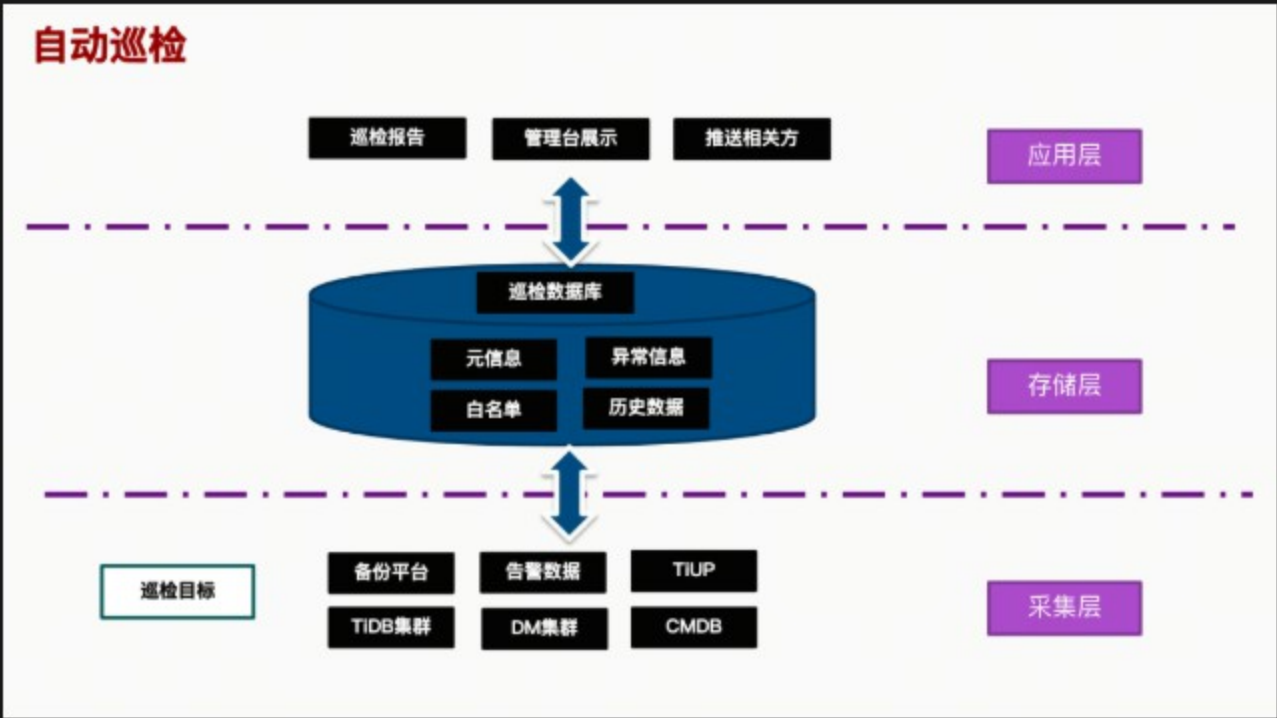
自动巡检

自动巡检的核心目标是**提前发现系统潜在问题**，从而减少运行过程中的突发告警。在早期未建立巡检机制时，系统经常在运行期间出现问题，例如磁盘空间不足、内存使用率过高、慢查询频发等，导致运维团队每天需要处理大量告警事件。

为解决这一问题，微众银行逐步建立起完整的自动巡检机制。目前已整合了**50 余项巡检项目**（早期为 40 项），覆盖多个关键维度，包括：

- TiDB 集群自身运行状态；
- DM（数据迁移）任务健康状况；
- 备份与高可用性配置；
- 安全性与合规性相关检查等。

通过系统化的巡检流程，运维人员能够在问题影响业务前提前识别风险，从而提高整体系统的稳定性和运维效率。



在实际应用中，自动巡检涵盖了多个维度，包括 TiDB 集群自身、DM 数据迁移组件、备份高可用配置以及安全性等方面的检查。

以 TiDB 容量监控为例，系统会通过 Prometheus 获取相关容量指标数据，并对近 24 小时的容量趋势进行分析。如果系统设置的告警阈值为 80%，而当前容量已达到 78%，系统将预判该集群在未来一天内可能会突破阈值，从而提前发出预警。

通过这一机制，巡检系统不仅能发现当前问题，还能实现趋势预测，帮助运维人员及时采取措施，避免潜在风险进一步扩大。

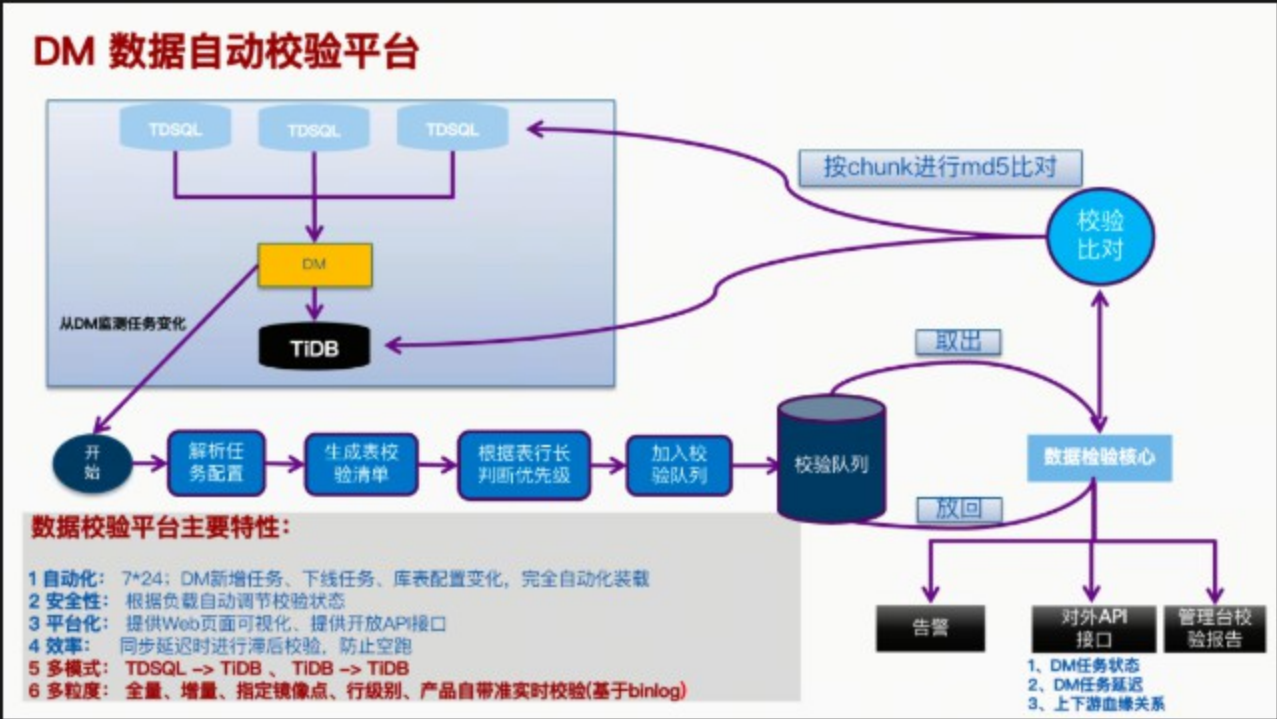
自动巡检-巡检项			
共计 40+ 巡检细项			
巡检维度包括:			
• TiDB集群			
• DM集群			
• 容量			
• 告警			
• 备份			
• 安全			
• 高可用			
• 批量耗时			
持续迭代			
分类	子分类	项名	项描述
集群自身	可用性	集群自身可用性	检查集群自身可用性
	可用性	集群自身可用性	检查集群自身可用性
	可用性	集群自身可用性	检查集群自身可用性
	可用性	集群自身可用性	检查集群自身可用性
	可用性	集群自身可用性	检查集群自身可用性
	可用性	集群自身可用性	检查集群自身可用性
	可用性	集群自身可用性	检查集群自身可用性
	可用性	集群自身可用性	检查集群自身可用性
	可用性	集群自身可用性	检查集群自身可用性
	可用性	集群自身可用性	检查集群自身可用性
DM	DM 集群	DM 集群	检查 DM 集群
	DM 集群	DM 集群	检查 DM 集群
	DM 集群	DM 集群	检查 DM 集群
	DM 集群	DM 集群	检查 DM 集群
	DM 集群	DM 集群	检查 DM 集群
	DM 集群	DM 集群	检查 DM 集群
	DM 集群	DM 集群	检查 DM 集群
	DM 集群	DM 集群	检查 DM 集群
	DM 集群	DM 集群	检查 DM 集群
	DM 集群	DM 集群	检查 DM 集群
备份	备份平台	备份平台	检查备份平台
	备份平台	备份平台	检查备份平台
	备份平台	备份平台	检查备份平台
	备份平台	备份平台	检查备份平台
	备份平台	备份平台	检查备份平台
	备份平台	备份平台	检查备份平台
	备份平台	备份平台	检查备份平台
	备份平台	备份平台	检查备份平台
	备份平台	备份平台	检查备份平台
	备份平台	备份平台	检查备份平台
告警	告警数据	告警数据	检查告警数据
	告警数据	告警数据	检查告警数据
	告警数据	告警数据	检查告警数据
	告警数据	告警数据	检查告警数据
	告警数据	告警数据	检查告警数据
	告警数据	告警数据	检查告警数据
	告警数据	告警数据	检查告警数据
	告警数据	告警数据	检查告警数据
	告警数据	告警数据	检查告警数据
	告警数据	告警数据	检查告警数据
安全	安全配置	安全配置	检查安全配置
	安全配置	安全配置	检查安全配置
	安全配置	安全配置	检查安全配置
	安全配置	安全配置	检查安全配置
	安全配置	安全配置	检查安全配置
	安全配置	安全配置	检查安全配置
	安全配置	安全配置	检查安全配置
	安全配置	安全配置	检查安全配置
	安全配置	安全配置	检查安全配置
	安全配置	安全配置	检查安全配置

在日常工作中，运维人员可以通过查看巡检报告，提前发现存在风险的实例。例如，当发现某个实例存在触发告警的可能时，可以主动采取措施，如清理日志或进行扩容，从而**避免因问题发展而引发被动告警**，有效减轻运维压力。

在银行场景中，无论是 TDC（交易数据中心）还是 TDP（交易数据平台），通常都要求**跨 IDC 的高可用部署**。如果系统发现某些节点未按规则部署，例如三个副本节点只分布在两个物理机房中，巡检系统也会将此类问题识别并触发预警，提示相关团队及时整改。

截至目前，微众银行已建立了 **50 余项巡检项目**。这些内容在实践中不断完善，并结合与同业的交流和对比，已能有效发现并预警多类潜在系统隐患，成为保障系统稳定运行的重要手段之一。

DM 数据自动校验平台



在数据汇聚场景中，数据一致性至关重要。微众银行使用 DM 工具将多个上游 TDSQL 实例的数据同步至 TiDB，但由于 DM 本身作为第三方工具，如何确保同步过程中的数据一致性一直是关键挑战。

对此，微众银行自研了一套数据校验平台。该平台的核心功能是**自动化处理大规模数据同步任务的校验流程**，避免人工操作的不现实性与不确定性。

目前系统已覆盖 **100 多条同步链路**。一旦新任务被创建，平台会自动加载配置，解析任务信息，生成对应的校验任务，且 **7*24 小时持续运行**。任务完成后，平台会自动生成校验报告，并在发现异常时触发预警，供相关人员及时处理。

该平台支持多种数据源与目标的组合校验方式，包括：

- **TDSQL 到 TiDB 的数据同步校验**
- **TiDB 到 TiDB 的跨集群数据校验**

同时，支持多种校验模式与力度：

- **全量校验**
- **增量校验**
- **指定时间范围与组件维度的校验**

此外，DM 在 6.5 版本中也已原生支持基于 binlog 的准实时校验功能，为数据一致性保障提供了更多支持手段。

通过这套校验平台，微众银行能够在多维度、复杂链路场景下有效确保数据同步过程中的一致性与完整性，为各类数据汇聚业务提供了可靠支撑。

大规模金融级 TiDB 运维实践总结

大规模金融级 TiDB 运维实践总结

TiDB 在分布式数据库具备 5 大能力

- 金融级特性：强一致性 / 高可用 / 安全合规 / 自主可控
- 弹性扩展：透明水平扩容支撑 200TB 级规模业务，免分库分表
- 高并发处理：十万级 QPS 业务无侵入，不需单元化改造
- 云原生架构：持续演进适配业务快速迭代需求
- 运维体系：本地化 TIUP 工具 + 通用监控，高效构建运维体系

TiDB 管理平台能力建设总结

- 全生命周期管理（部署 / 升级 / 扩缩容 / 监控 / 退役），建立规范模板与风险管控
- 表级资源可观测性，容量预测 / 实时画像，智能资源可观测性粒度细化
- 智能巡检体系，健康 / 安全 / 备份 / 高可用 / 跑批多维度监测，自动化报告预警
- 智能诊断优化，慢查询分析（执行计划 / 索引 / 锁竞争），自动推送根因分析

从产品能力来看，TiDB 在微众银行的大规模应用中展现出以下五项核心优势：

1. **金融级特性**：包括在 6.5 版本中引入的悲观锁和 RC（Read Committed）隔离级别，持续优化的事务处理时延，满足金融系统对一致性与稳定性的严苛要求；
2. **弹性扩展能力**：支持水平扩展，无需分库分表，即可应对业务数据量与访问压力的快速增长；
3. **高并发处理能力**：能够支撑十万级 QPS 的业务负载，满足核心系统的性能要求；
4. **持续的版本演进**：通过功能迭代不断提升产品稳定性与适用性；
5. **高效运维体系支撑**：依托 Prometheus 等工具构建可观测性基础，提升整体可运维性。

在平台能力建设方面，微众银行也构建了围绕 容量预测、自动巡检、智能诊断 等关键能力的体系化运维平台。通过产品特性与平台建设的双重支撑，TiDB 能够有效支撑银行业务规模的快速发展，保障系统运行的稳定与高效。

3 未来探索展望

降本增效探索

近年来，降本增效已成为各行各业共同关注点，微众银行也在 TiDB 使用过程中积累了多项降本实践经验。针对 TiDB 新版本（如 7.0 和 8.0），**资源管控**与**副本放置策略**是其中两项重要的新特性，为资源优化和成本控制提供了新的技术手段。

降本增效探索：Resource Control（资源控制）和 Placement Rules（放置规则）

特性	Resource Control	Placement Rules
核心目标	逻辑层面的 资源分配公平性 （CPU、内存、IO 等）	物理层面的 数据副本位置控制 （Region/节点）
作用对象	用户会话或资源组（Workload Group）	表、分区或数据库（数据副本）
管理维度	资源配额、优先级、限流	数据存储位置、副本数、副本分布策略
典型场景	避免高负载查询耗尽资源，保障核心业务稳定性	冷热数据分离、跨机房容灾、合规性存储

典型协同场景：

场景 1：混合负载资源隔离
收益：OLTP 和 OLAP 混合部署时，避免资源争抢

场景 2：跨地域多活部署
收益：数据本地化存储，降低跨地域访问延迟。

场景 3：冷热数据分层存储
收益：将冷数据存储到低成本节点，释放高性能节点资源

➡

协同示例：

假设一个金融系统需要同时处理实时交易（OLTP）和历史数据分析（OLAP）：

1.Resource Control:

- 创建两个资源组 oltp_group 和 olap_group，分别分配 80% 和 20% 的 Request Unit (RU)
- 确保 OLTP 交易即使在高负载下也能优先获得资源

2.Placement Rules:

- 将 OLTP 业务表的数据副本固定在本地机房的 TiKV 节点（低延迟访问）
- 将 OLAP 分析表的数据副本放置到远端机房的 TiFlash 节点（高吞吐分析）

3.联合效果

- OLTP 业务既享有高优先级资源配额，又通过就近存储降低延迟
- OLAP 分析任务受资源配额限制，且数据物理隔离避免干扰在线业务

其中，**资源管控**功能支持在同一集群内通过资源配额限制的方式，避免多个租户之间发生资源争抢问题。通过对 CPU、内存等资源进行限制与隔离，系统可以实现更加精细化的资源调度，在保障关键业务性能的同时，提高资源利用率。

副本放置策略作为 TiDB 的新特性之一，提供了对数据副本位置的灵活控制能力，可用于实现更精细的资源调度与成本优化。通过该策略，系统可以根据业务需求将数据副本**有选择性地放置在指定存储介质或节点上**。例如，在冷热数据分离场景中，可将热数据放置在性能更高的 SSD 硬盘上，将冷数据迁移至成本更低的 HDD 介质，以降低整体存储成本。此外，还可以将特定数据类型或 leader 副本优先放置在**靠近应用的计算节点**，减少网络延迟，提升访问效率。

这一机制具备良好的灵活性和扩展性，为未来在资源分级调度、跨机房部署、业务分层管理等方面的深入探索提供了基础。

更多探索与展望

重点规划

降本

冷热分离
Resource Control
Placement Rules

探索

TiDB 向量化能力
全链路故障分析与定位

国产化

ARM
操作系统

运维提效

版本升级
自动化运维

通过社区共建与技术迭代，持续推动分布式数据库在金融核心场景的深度应用

除了降本方向，微众银行也在积极探索**向量化处理**、**AI 技术**等前沿领域应用。例如，基于自研 AI 工程化平台“模型可插拔”等特性，微众银行成为业内率先成功部署 DeepSeek-R1 满参数模型的机构之一，验证了 AI 工程化平台开放式架构的灵活、敏捷性优势。

在**国产化方面**，微众银行在 ARM 架构和国产操作系统适配方面也有丰富实践经验。未来，我们期望和国产 TiDB 数据库继续深化安全可控技术。

在**效率提升方面**，版本升级与自动化运维是两项关键工作：

- 持续版本迭代不仅带来新特性，还大幅提升系统可靠性与稳定性；
- 自动化运维体系建设也在不断推进，以进一步降低人工干预成本，提高整体运维效率和一致性。

最后，面向未来发展，我们希望通过社区共建与技术迭代，持续推动国产分布式数据库在金融核心场景中的深度应用，不断提升系统的安全可控能力与业务支撑能力。

/ 相关推荐 /




4.98 亿月活背后的国产数据库： 咪咕视讯携手 TiDB 攻克内容分发 核心系统挑战

行业实践




4.98 亿月活背后的国产数据库：咪咕视讯携手 TiDB 攻克内容分发核心系统挑战



攻克多版本运维难题：爱奇艺百套 TiDB 集群升级至 v7.1.5 实战宝典来袭！

行业实践



攻克多版本运维难题：爱奇艺百套 TiDB 集群升级至 v7.1.5 实战宝典来袭！

💡 点击文末【[阅读原文](#)】，立即下载试用 TiDB ！

 <p>订阅号 @PingCAP 点击关注</p>	 <p>服务号 @TiDB Club 点击关注</p>	 <p>视频号 @PingCAP 点击关注</p>
 <p>PingCAP 官网 pingcap.com 点击访问</p>	 <p>Bilibili @TiDB_Robot 点击关注</p>	 <p>知乎 @PingCAP 点击关注</p>



微信搜一搜

 TiDB

[阅读原文](#)