

云数据库TDSQL

产品简介

产品文档



腾讯云

【版权声明】

©2015-2016 腾讯云版权所有

本文档著作权归腾讯云单独所有，未经腾讯云事先书面许可，任何主体不得以任何形式复制、修改、抄袭、传播全部或部分本文档内容。

【商标声明】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。

【服务声明】

本文档意在向客户介绍腾讯云全部或部分产品、服务的当时的整体概况，部分产品、服务的内容可能有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

文档目录

文档声明.....	2
产品概述	4
使用场景	6
高可用架构设计.....	9
服务等级协议.....	15
分布式版本	16

产品概述

1. 简介

腾讯云云数据库TDSQL定位于OLTP场景下高安全性的企业级云数据库，十余年来一直应用于腾讯计费业务，TDSQL兼容MySQL语法，拥有诸如线程池、审计、异地容灾等高级功能，同时具有云数据库的易扩展性、简单性和性价比。

为了能支持金融级的业务需要，CDB for TDSQL的诞生经历了十余年。

为满足不同方向的需求，分布式云数据库DCDB for TDSQL和（金融级）云数据库CDB for TDSQL项目启动，通过腾讯云平台，将TDSQL划分了两个发展方向：

（金融级）云数据库TDSQL

：基于混合云（支持公有云和自建云），以金融级安全和数据一致性为前提，发展高安全、高容灾能力和高并发性能的综合解决方案。

分布式云数据库DCDB for

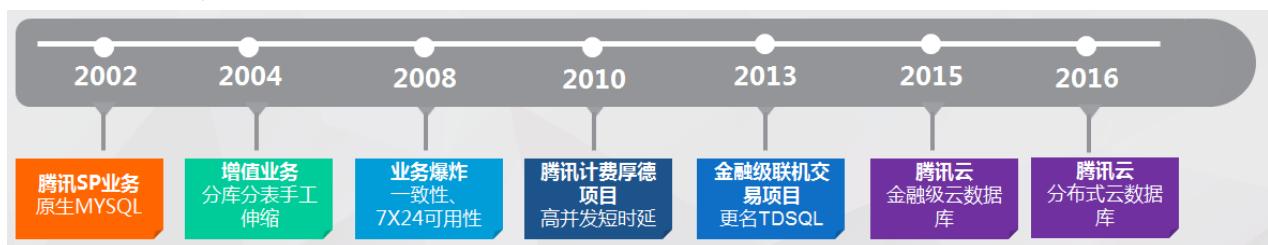
TDSQL

：基于公有云，兼容MySQL协议和语法，具有自动分库分表能力的分布式数据存储解决方案，存储TB或PB级的海量数据库场景。

2. 发展历史

- 2002年，基于运营商SP业务，腾讯数据库团队开始对MySQL进行改造；
- 2004年，腾讯互联网增值业务开始爆发，腾讯互联网增值业务开始爆发，业务量的爆炸给MySQL数据库带来了巨大的扩容压力，当时即开始引入分库分表机制来解决难题——针对大的表，按照ShardKey预先拆成多个子表，分布在不同的物理机器节点上。；
- 2008年，腾讯游戏、QQ空间、财付通等各类业务再次爆发，故障率变高（如硬件故障、操作系统故障、应用程序或服务故障），而传统的异步复制，主备之间部分数据并不一致，主备切换后不一致的数据就会出现问题，带来用户投诉和公司经济损失。腾讯数据库团队在优先保障数据一致性的理念下，通过同步线程异步化，孤岛检测、黑名单、多级切换等方案，最终诞生了强同步(MAR)技术。同时，通过多年发展，也做到自动扩容、故障检测、切换、恢复过程均由系统自动完成。

- 2010年，基于正在火热的互联网支付业务超高可用性、超高并发和极短响应的需求，腾讯厚德项目（HOLD项目）启动，目标是实现自主研发的，超高并发、超短时延的高一致分布式数据存储层。最终，厚德项目非常成功，后面会讲到的TDSQL的主要技术，包括跨机房强同步、异地容灾、切换一致性保障、数据自动的分片（sharding）、自动伸缩、集群自动管理等，都来自于该项目。
- 2012年，腾讯内部正式给这款百炼千锤的产品命名为TDSQL，考虑到后续云的大规模使用，这次项目直接以兼容MySQL引擎作为目标。在后续两年时间，陆续支撑米大师（Midas）、微众银行（WeBank）等多个兄弟业务的上线，并针对银行场景的数据关系模型设计了关系紧密的数据聚合，同时将跨节点的分布式架构转换扩展到单机架构，有效的覆盖了大中小多层次的用户。
- 2015年，TDSQL正式进驻腾讯云。



3.服务定位

3.1. (金融级) 云数据库CDB for TDSQL的定位

面向涉及交易、订单、资金流转、信贷、ERP等要求数据强一致，更高的安全等级的数据库服务，满足金融、电商平台、政府、集团企业等核心系统的数据库服务。

3.1. 分布式云数据库DCDB for TDSQL的定位

面向百万用户以上的大型应用，如电商、金融、O2O、社交、物联网等产生海量的数据，普通MySQL架构无法支撑业务增长，而必须采用水平拆分（即分库分表）的架构。

使用场景

下列场景适用于（金融级）云数据库CDB for TDSQL

1. 金融行业核心应用

无论是银行、保险、证券、基金，还是新兴金融企业如小额贷款、互联网理财等，其核心应用系统如资金交易、流转、账务等系统都需要一种高可用的数据库进行支撑，而监管要求也对数据库本身的安全性、标准化都提出了更高的要求。

面对这一系列挑战，腾讯云金融级云数据库TDSQL够为金融行业用户提供：

- 提供安全、开放的系统平台；
- 提供高可用的方案，如系统采用双主双活、多机热备份模式或集群服务等模式；
- 优先保障主备数据库强一致性；
- 提供多种容灾方案，数据可靠性至少99.99999%；
- 支持弹性伸缩；
- 支持物理专享，单一用户内部系统之间采用多租户模式，用户与用户之间物理隔离；
- 需满足一行三会等机构的监管要求。

2. 交易或订单系统

随着电子商务与在线交易的兴起，越来越多的企业支持或计划支持电子商务或在线交易能力，这就给企业的交易或订单系统带来了更多问题亟待解决。以知名B2C电商平台为例，从2012年开始，诸如618、双11等活动导致平台订单量则成倍增长，伴随着订单量的增加，系统面临的挑战与日俱增。

腾讯云来为腾讯及其合作伙伴提供基于交易或订单系统的金融级云数据库，并已经为交易或订单场景解决了以下问题：

- 解决复杂订单业务和处理流程下，数据不一致的问题；
- 解决海量交易情况下，数据库的性能瓶颈；

- 解决海量数据的运营管理，及时发现和解决数据库问题。

3. 公众服务系统

公众服务系统是政府、企业为公众服务的线上门户，诸如食品溯源、专利查询、O2O服务、病历查询、购票、办事预约等都是公众查询类系统的一种。该类系统的特点场景特点是用户基数大，信息查询分散，随着时间推移数据急剧膨胀。该类系统的运营特点就是确保数据库的安全完整性，在而自然灾害、系统故障、员工误操作和病毒感染面前，企业的正常运营都有可能变得脆弱及其脆弱，因此这类企业就需要一种数据永远不丢，稳定可靠的数据库管理系统，而这是腾讯云金融级云数据库正好擅长的：

- 跨机房（跨可用区）容灾，至少6个数据库副本，数据可靠性达到99.99999%
- 提供高可用的方案，随时随地为用户提供服务。
- 支持读写分离，大幅提高性能，支持某些系统读大于写等业务特点。

下列场景适用于分布式云数据库DCDB for TDSQL

4. 大型应用

近些年来，百万用户以上的大型应用，如电商、金融、O2O、社交应用，将持续产生海量的数据，数据规模为TB甚至PB级。面对海量信息内容，企业需在选

择之初，就考虑水平拆分（即分库分表），因为“水平拆分”能力将提供

- 无限大的存储空间（或PB级）
- 单表几亿~千亿行数据
- 更高效的读写性能

腾讯云分布式云数据库的自动分库分表（group-sharding）方案，利用普通硬件而非共享存储等方式，为用户搭建一个无限存储容量，超高读写QPS的数据库管理系统：

- 支持自动分库分表，PB级数据存储；
- 支持proxy代理网关，物理分库分表，逻辑整合，对DBA完全透明；
- 超高性能，异步多线程强同步复制方案(MAR)，物理读写分离，性能是同类方案的30%以上。

5. 物联网数据

在工业监控和远程控制、智慧城市的应用、智能家居、车联网等物联网场景下。及其庞大的传感器数量和密集的数据采样频率将会形成超大数据规模存储。

6. 平台文件索引

平台的图片、小文件、视频的数据极大，文件索引为亿级，该类数据通常只有新增、修改、读取、删除操作，分布式数据库可以有效提高文件索引的修改效率和稳定性。

高可用架构设计

1. 高可用架构

1.1 高可用架构基础

在生产系统中，通常都需要用高可用方案来保证系统不间断运行；数据库作为系统数据存储和服务的核心能力，其可用要求高于计算服务资源。目前，数据库的高可用方案通常是让多个数据库服务协同工作，当一台数据库故障，余下的立即顶替上去工作，这样就可以做到不中断服务或只中断很短时间；或者是让多台数据库同时提供服务，用户可以访问任意一台数据库，当其中一台数据库故障，立即更换访问另外数据库即可。

当然，由于数据库中记

录了数据，想要在多台数据库中切换，需要进行数据同步，所以数据同步技术是数据库高可用方案的基础。

1.2 常见高可用架构介绍

- 共享存储方案

：使用共享存储，如SAN存储。SAN的原理是多台数据库服务器共享同一个存储区域，这样多台数据库都可以“读写”同一份数据。当主库发生故障时，第三方高可用软件把文件系统在备库上挂起，然后在备库上启动数据库即完成切换。

- 日志同步或流复制同步：数据库最常见复制模式，如MySQL数据库。每当写入数据，MySQL Master Server将自己的Binary Log通过复制线程传输给Slave，Slave接收到Binary Log以后，依照Binary Log内容，写入相同数据到文件系统。目前MySQL已经提供：

- 异步复制：异步复制可以确保得到快速的响应结构，但是不能确保二进制日志确实到达了slave上，即无法保障数据一致性；
- 半同步复制：（由google提供的同步插件）半同步复制对于客户的请求响应稍微慢点，在超时等情况下，会退化为异步，即基本保障数据一致性，但无法保证数据完全一致性。

- 基于触发器的同步：试用触发器记录数据变化，然后同步到另一台数据库上。

- 基于中间件的同步

：系统不直接连接到底层数据库，而是连接到一个中间件，中间件把数据库变更发送到底层多台数据库上，从而完成数据同步。早几年，由于业务需求，数据库性能、同步机制等问题，某些软件开发商通常采用类似架构。

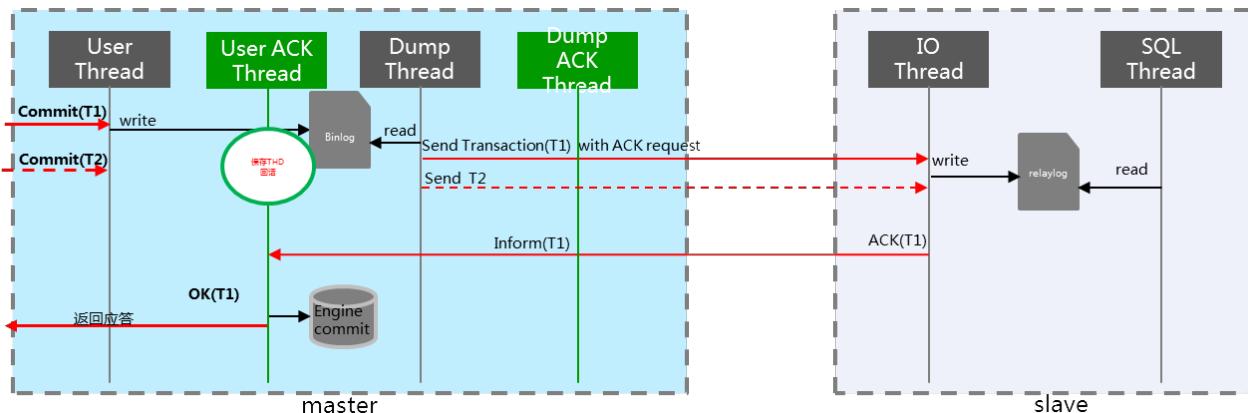
2.TDSQL高可用架构简介

2.1 异步多线程强同步复制方案（强同步方案）

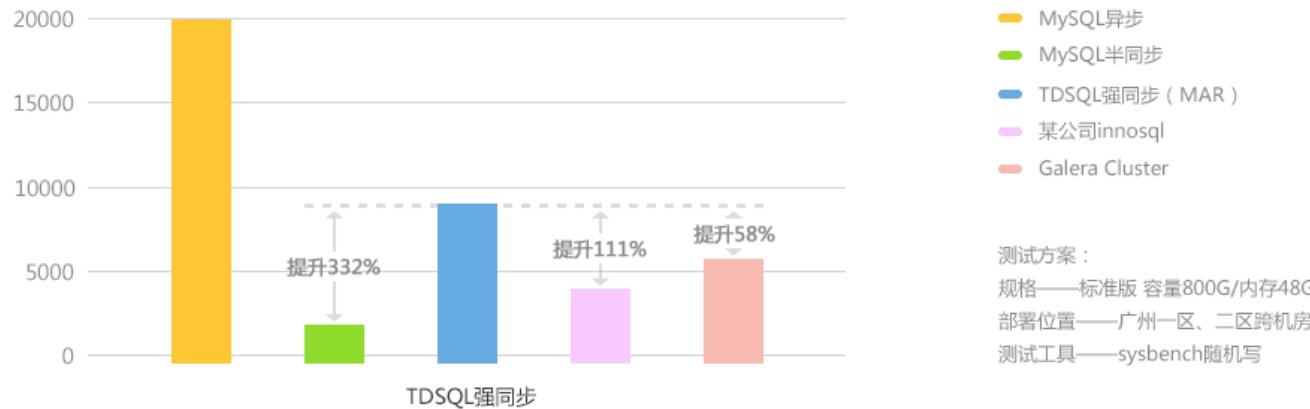
同步技术发展过程中，提供了异步复制、半同步等同步技术，这两种技术面向普通用户群体，在用户要求不高、网络条件较好、性能压力不大的情况下，能够基本保障数据同步；但通常情况下，采用异步复制、半同步机制容易经常出现数据不一致问题，直接影响系统可靠性，甚至出现丢失交易数据，带来直接或间接经济损失。在腾讯内部业务中的多年积累，自主研发出数据库异步多线程强同步复制方案（Multi-thread Asynchronous Replication MAR），相比于Oracle的NDB引擎，Percona XtraDB Cluster和MariaDB Galera Cluster，其性能、效率和适用性更据优势。简单来说，MAR强同步方案强同步技术具有以下特点

- 一致性的同步复制，保证节点间数据强一致性；
- 对业务层面完全透明，业务层面无需做读写分离或同步强化工作；
- 将串行同步线程异步化，引入线程池能力，大幅度提高性能
- 支持集群架构；
- 支持自动成员控制，故障节点自动从集群中移除；
- 支持自动节点加入，无需人工干预；
- 每个节点都包含完整的数据副本，可以随时切换；
- 无需共享存储设备

腾讯MAR方案强同步技术，只有当备机数据同步后，才由主机向应用返回事务应答，示意图如下：

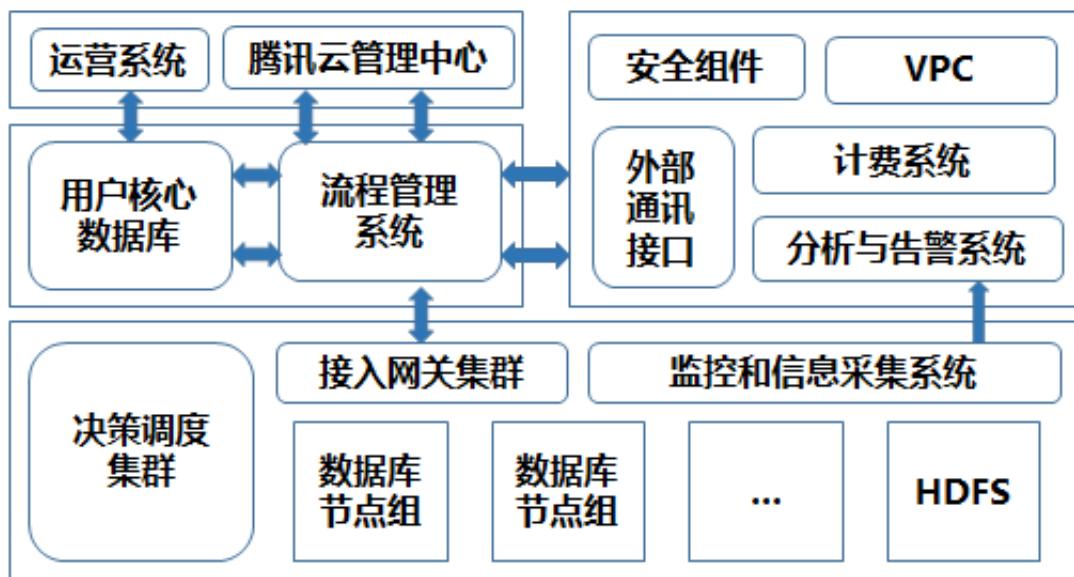


从性能上优于其他主流同步方案，通过对比，在跨可用区(IDC机房)同样的测试方案下，我们发现其MAR技术性能优于MySQL半同步约5倍，优于MariaDB Galera Cluster性能1.5倍（此处使用sysbench标准用例测试）。

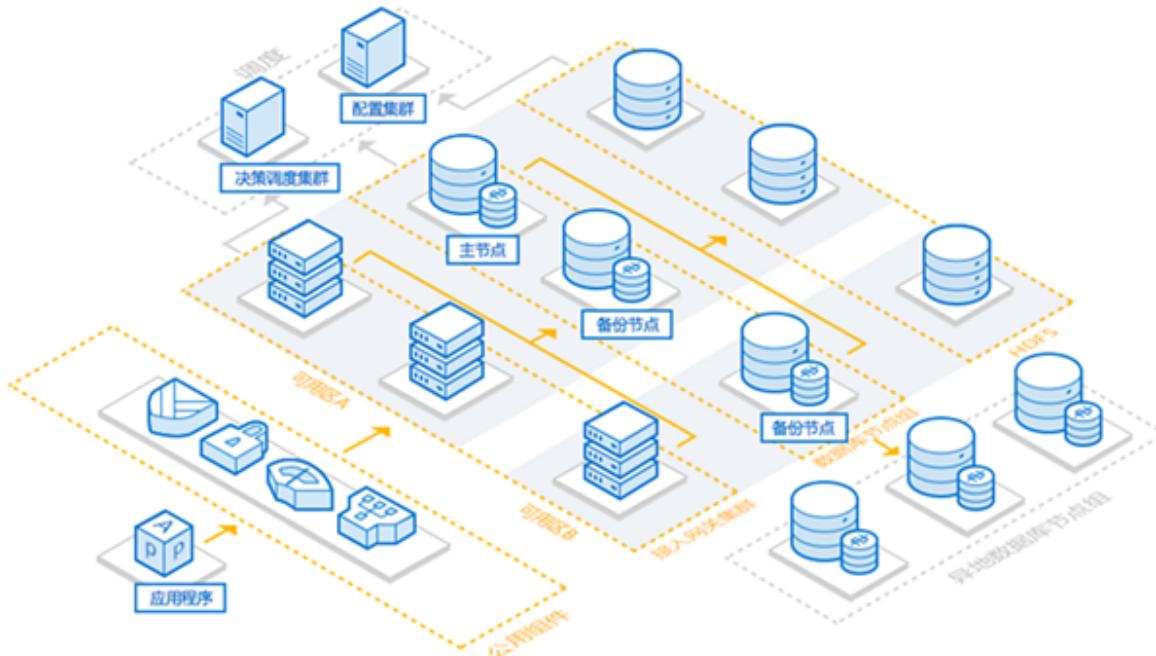


2.2 TDSQL集群架构

TDSQL采用集群架构，一套独立TDSQL系统至少需要十余系统或组件组成，架构简图如下



其中，TDSQL最核心的三个主要模块是：决策调度集群（Tschedule）、数据库节点组（SET）和接入网关集群（TProxy），三个模块的交互都是通过配置集群（TzooKeeper）完成。



- **数据库节点组（又称“分片”）**：由兼容MySQL数据库引擎、监控和信息采集（Tagent）组成，其架构由“一个主节点（Master）、若干备节点（Slave_n）、若干异地备份节点（Watcher_m）”，通常情况下：
 - 部署在跨机架、跨机房的服务器中；
 - 通过心跳监控和信息采集模块（Tagent）监控，确保集群的健壮性；
 - 分布式架构下，基于水平拆分，若干个分片（数据库节点组）提供一个“逻辑统一，物理分散”分布式的数据库实例。
- **决策调度集群**：作为集群的管理调度中心，主要管理SET的正常运行，记录并分发数据库全局配置，其中包括
 - 调度作业集群（TDSQL Scheduler）帮助DBA或者数据库用户自动调度和运行各种类型的作业，比如数据库备份、收集监控、生成各种报表或者执行业务流程等等，TDSQL把Schedule、z

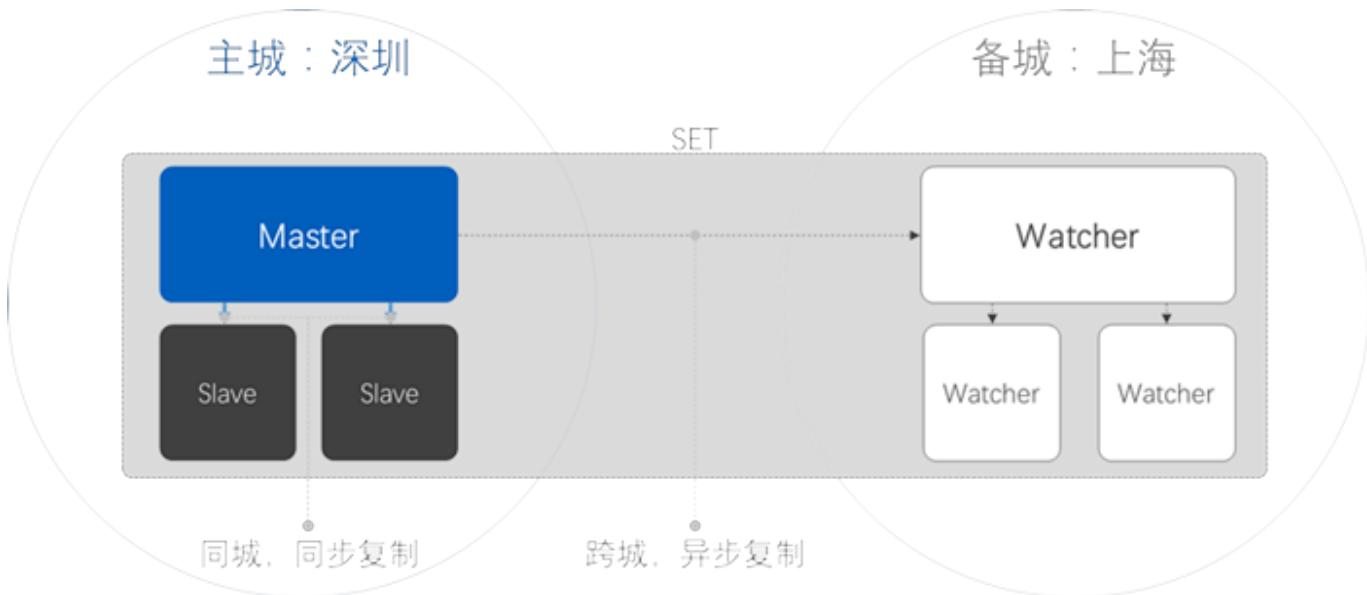
ookeeper、OSS（运营支撑系统）结合起来通过时间窗口激活指定的资源计划，完成数据库在资源管理和作业调度上的各种复杂需求，Oracle也用DBMS_SCHEDULER支持类似的能力。

- 程序协调与配置集群（TzooKeeper）：它是TDSQL提供配置维护、选举决策、路由同步等，并能支撑数据库节点组（分片）的创建、删除、替换等工作，并统一下发和调度所有DDL（数据库模式定义语言）操作。
 - 运维支撑系统（OSS）：基于TDSQL定制开发的一套综合的业务运营和管理平台，同时也是真正融合了数据库管理特点，将网络管理、系统管理、监控服务有机整合在一起。
 - 决策调度集群独立部署在腾讯云全国三大机房中（跨机房部署，异地容灾）。
-
- 接入网关集群（TProxy）：在网络层连接管理SQL解析、分配路由。
 - 与数据库引擎部署数量相同，分担负载并实现容灾；
 - 从配置集群（TzooKeeper）拉取数据库节点（分片）状态，提供分片路由，实现透明读写；
 - 记录并监控SQL执行信息，分析SQL执行效率，记录并监控用户接入信息，进行安全性鉴权，阻断风险操作；
 - TProxy前端部署为腾讯网关系统TGW，对用户提供唯一一个虚拟IP服务。

这种集群架构极大简化了各个节点之间的通信机制，也简化了对于硬件的需求，这就意味着即使是简单的x86服务器，也可以搭建出类似于小型机、共享存储等一样稳定可靠的数据库。而且，TDSQL保持了MySQL原有协议，保证基于MySQL开发的系统、应用都不需要修改。

2.2（金融级）高可用架构

TDSQL提供金融级的容灾架构，可以支持到两地三中心部署结构——同城节点直线距离大于10KM，异地节点直线距离大于100km，使用腾讯自主研发的高可用调度方案（High Availability HA）实现。示意图如下：



请注意:使用“强同步”复制时，如果主库与备库自建网络中断或备库出现问题，主库也会被锁住 (hang)，而此时如果只有一个主库或一个备库，那么是无法做高可用方案的。——

因为单一服务器服务，如果股指则直接导致部分数据完全丢失，不符合金融级数据安全要求。

因此，(金融级云)数据库CDB for

TDSQL的方案是使用“两个”备库，只要有一个备库是正常的，主库就不会被hang住。

除此之外，未避免人为原因造成数据误删除，CDB for TDSQL通过数据备份服务(HDFS)提供可配天数的，至少3份完全备份和增量备份，进一步保障数据安全，做到不丢，不断，不错。

服务等级协议

云数据库(TDSQL)服务等级协议详见《政策与规范》[云数据库服务等级协议](#)

分布式版本

腾讯云云数据库（TDSQL）即将支持分库分表能力，敬请期待！