

云数据库 MariaDB

产品简介

产品文档



腾讯云

【版权声明】

©2013-2018 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模式的承诺或保证。

文档目录

产品简介

产品概述

系统架构

产品优势

标准与认证

独享集群数据库解决方案

典型应用场景

服务等级协议

分布式版本

产品简介

产品概述

最近更新时间：2018-10-15 11:45:56

1.简介

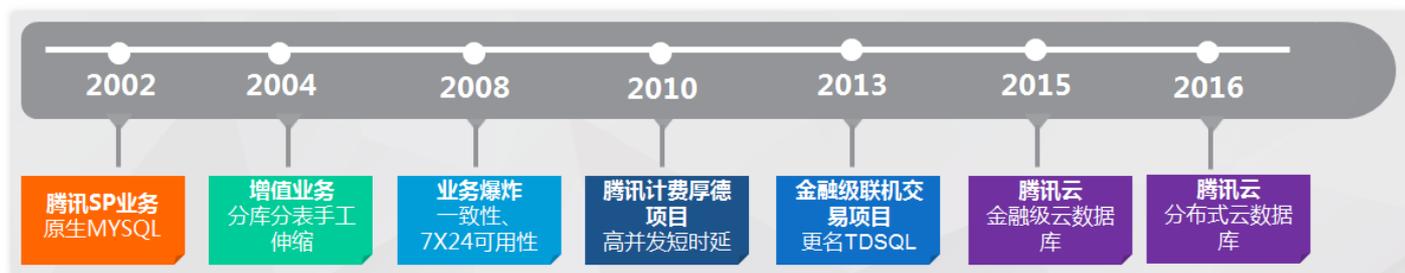
云数据库TencentDB for MariaDB (TDSQL) 定位于OLTP场景下高安全性的企业级云数据库，十余年来一直应用于腾讯计费业务，[MariaDB\(TDSQL\)兼容MySQL语法](#)，拥有诸如线程池、审计、异地容灾等高级功能，同时具有云数据库的易扩展性、简单性和性价比。

2.发展历史

TDSQL从立项到上线腾讯云，经历了十余年：

- **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MariaDB (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，并选择MariaDB作为内核。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.内核说明

MariaDB (TDSQL) 内核基于**MariaDB 10.X**，具体版本参见实例具体说明。

4.分布式架构

请参见分布式云数据库DCDB详细说明。

系统架构

最近更新时间：2018-10-15 11:47:32

1. 系统架构前言

1.1 高可用架构

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。

由于数据库中记录了数据，想要在三台数据库中切换，数据必须是同步的，所以**数据同步技术是数据库高可用方案的基础**；当前，数据复制方式有以下三种方式：

- **异步复制**：应用发起更新（含增加、删除、修改操作）请求，Master 完成相应操作后立即响应应用，Master 向 Slave 异步复制数据。因此异步复制方式下，Slave 不可用不影响主库上的操作，而 Master 不可用有概率会引起数据不一致。
- **强同步复制**：应用发起更新请求，Master 完成操作后向 Slave 复制数据，Slave 接收到数据后向 Master 返回成功信息，Master 接到 Slave 的反馈后再应答给应用。Master 向 Slave 复制数据是同步进行的，因此 Slave 不可用会影响 Master 上的操作，而 Master 不可用不会引起数据不一致。

****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。

- **半同步复制**：半同步复制是google提出的一种同步方案，他的原理是正常情况下数据复制方式采用强同步复制方式，当 Master 向 Slave 复制数据出现异常的时候（Slave 不可用或者双节点间的网络异常）退化成异步复制。当异常恢复后，异步复制会恢复成强同步复制。半同步复制意味着 Master 不可用有概率会较小概率引起数据不一致。

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.MariaDB(TDSQL)架构简介

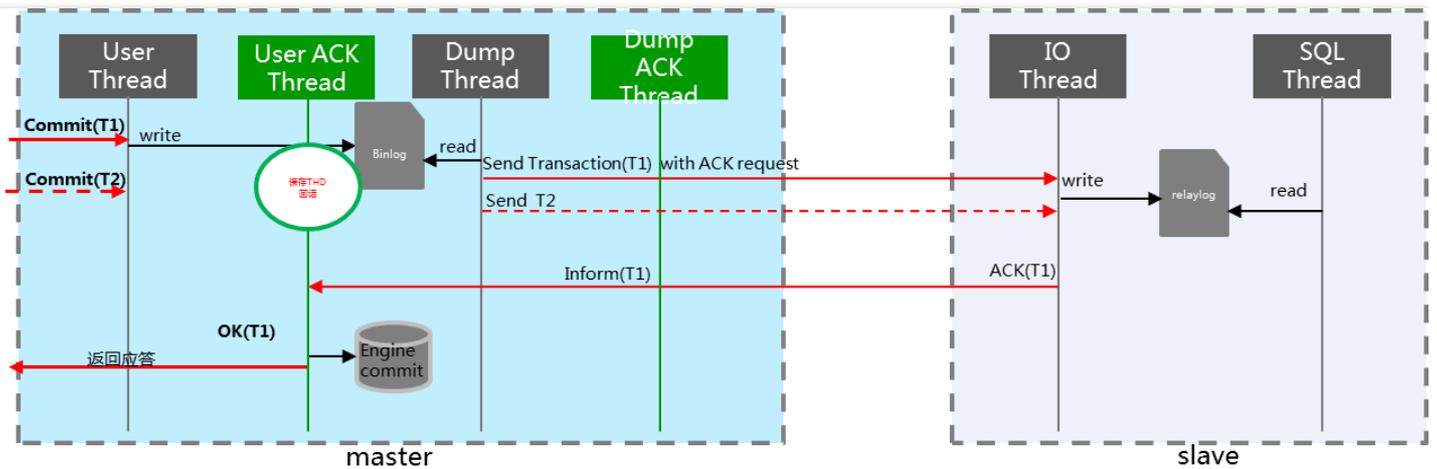
2.1 MariaDB(TDSQL)异步多线程强同步复制技术（强同步复制）

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。

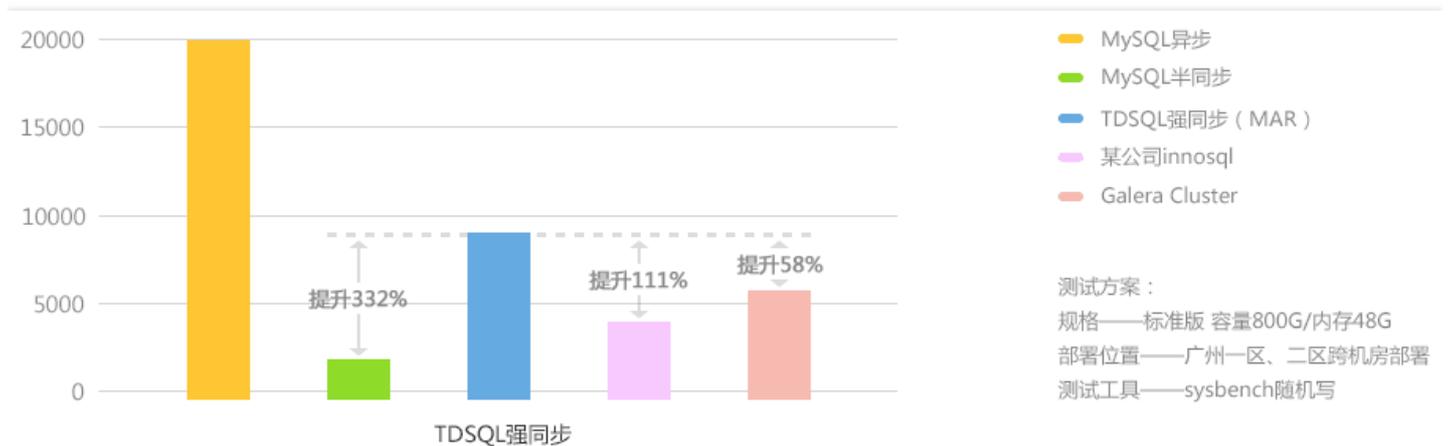
在腾讯内部业务中的多年积累，自主研发出数据库异步多线程强同步复制方案（Multi-thread Asynchronous Replication MAR），相比于Oracle的NDB引擎，Percona XtraDB Cluster和MariaDB Galera Cluster，其性能、效率和适用性更据优势。简单来说，MAR强同步方案强同步技术具有以下特点

- 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：

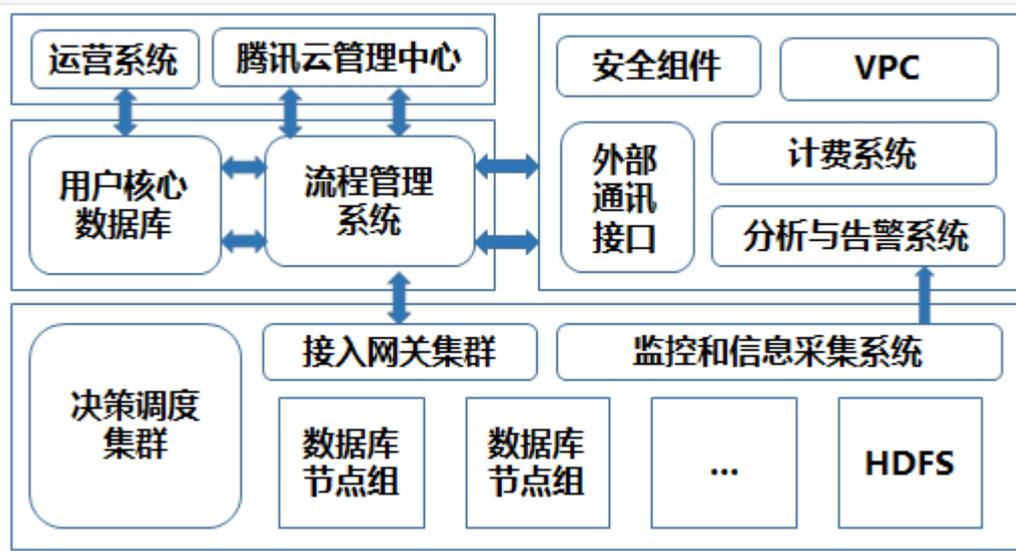


从性能上优于其他主流同步方案，通过对比，在跨可用区(IDC机房)同样的测试方案下，我们发现其MAR技术性能优于MySQL半同步约5倍，优于MariaDB Galera Cluster性能1.5倍（此处测试使用sysbench标准用例测试）。

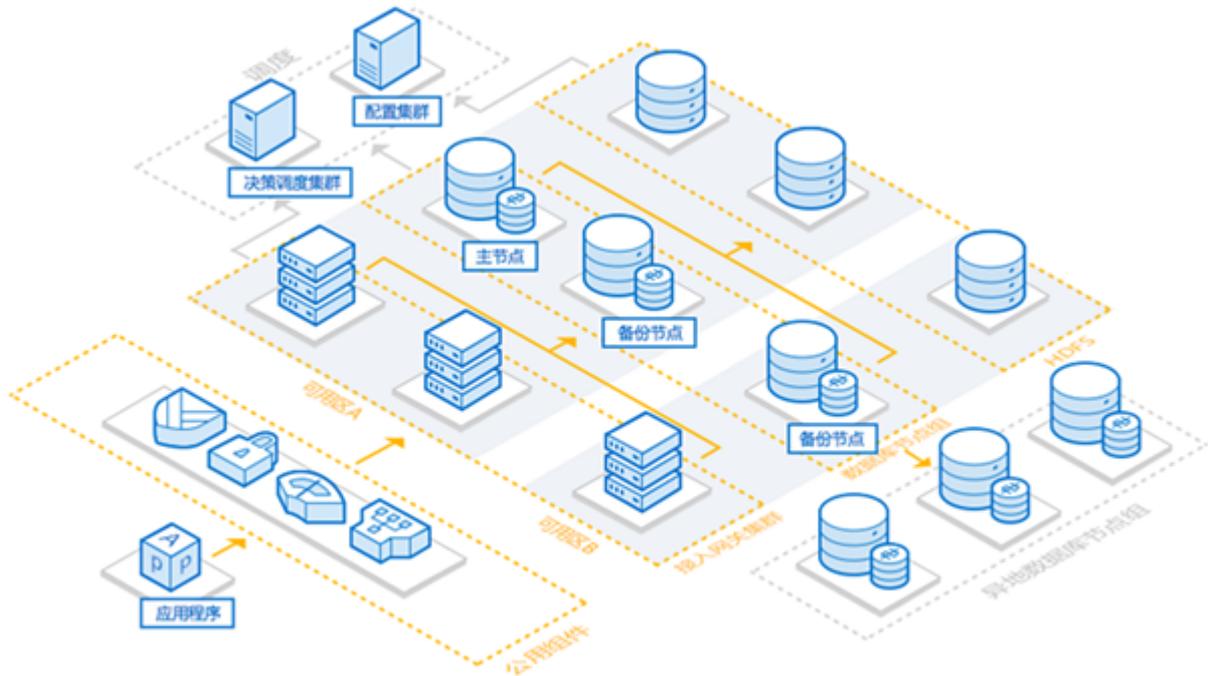


2.2 MariaDB(TDSQL)集群架构

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



其中，MariaDB(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 (SET) :** 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 (Tschedule) :** 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，其包括
 - 调度作业集群 (MariaDB(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，MariaDB(TDSQL)把Schedule、zookeeper、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 (ZooKeeper) : 它是MariaDB(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, ZooKeeper部署数量需大于等于3台并。
 - 运维支撑系统 (OSS) : 基于MariaDB(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(TProxy非腾讯云网关TGW)。
 - 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 (ZooKeeper) 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, 对用户仅提供唯一——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。

2.2 MariaDB(TDSQL)的高可用特性

2.2.1 物理高可用

根据您的购买实例配置不同, MariaDB(TDSQL)通常采用一主一备双节点、或一主两备三节点架构。每个节点安装在独立跨机架部署的物理机上, 确保确保不会因为单一设备、机架网络故障或断电影响数据库服务。

2.2.3 网络高可用

MariaDB(TDSQL)的每个节点物理机均采用双网卡双连交换机配置, 物理网络安全可靠。而在实际使用时, TProxy前对接腾讯云网关TGW; SET节点出现故障后, TProxy最快200ms内切换DB路由; 若TProxy故障, TGW在1s内负载到其他存活的TProxy; 切换不改变访问VIP (虚拟IP), 以便屏蔽物理服务器变化带来的影响。

2.2.4 备份恢复服务

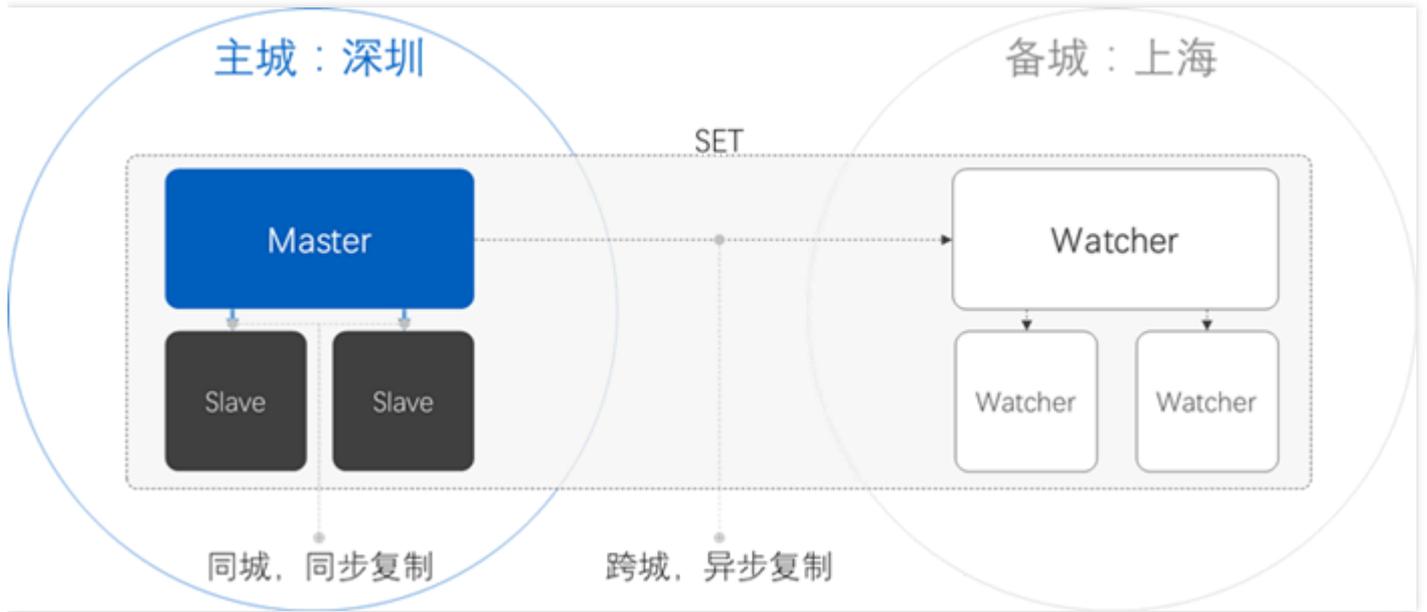
备份服务: 备份模块负责定时对MariaDB(TDSQL)进行 (物理) 备份和二进制文件 (Binlog), 备份文件将上传到安全等级更高的分布式文件集群中 (HDFS) 中, 通常情况下, 备份总是在备节点上面发起, 以避免对主节点提供的服务带来冲击。

恢复服务: 恢复服务又名回档恢复, 有恢复模块负责将HDFS上面的备份文件恢复临时实例上, 以使用户检查或调整而不影响主实例的运行。

备份下载: 您可以将备份文件转储和下载到指定位置, 如价格更低的腾讯云COS。

2.2.4 两地三中心

MariaDB(TDSQL)两地三中心部署架构——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:



产品优势

最近更新时间：2018-10-15 11:48:30

1.数据强一致性

支持配置强同步复制，不同于MySQL的异步复制，在主备架构下，强同步确保主备数据强一致，避免您的数据库在主备切换时丢失数据。当然，您也可以通过修改配置，关闭强同步功能以提高性能。

2.更高安全性

MariaDB (TDSQL)可以为您的数据库提供全面的安全防护：

- 防DDoS攻击：当用户使用外网连接和访问 MariaDB (TDSQL) 实例时，可能会遭受 DDoS 攻击。当 MariaDB (TDSQL) 安全体系认为用户实例正在遭受 DDoS 攻击时，会自动启动流量清洗的功能。
- 系统安全：即使在内网，MariaDB (TDSQL) 也处于多层防火墙的保护之下，可以有力地抗击各种恶意攻击，保证数据的安全。另外，物理服务器不允许直接登录，只开放特定的数据库服务所需要的端口，有效隔离具有风险的操作。
- VPC网络隔离：可支持VPC网络，以安全隔离内网其他设备的访问。
- 内网风控：腾讯云数据库团队无法直接访问到MariaDB (TDSQL) 物理机或数据库实例，必须通过腾讯云运维管理平台访问，即使是排查问题也必须在安全设备上，且通过内部风控系统的严格管理。
- 对象粒度的权限管控：用户可定义到表级的权限，并允许配置访问 MariaDB (TDSQL) 的 IP 地址，指定之外的 IP 地址将被拒绝访问。
- 数据库审计：支持配置数据库审计，记录管理员或用户的操作历史，用于出现风险后的管控。
- 数据库防火墙（即将支持）：支持自定义恶意行为，并对恶意行为、会话予以阻断。
- 操作日志：系统记录用户访问腾讯云WEB管理中心操作MariaDB (TDSQL) 的全部记录，常用于事后追溯。

3.更强大的功能

- 可支持多源复制 (Multi-source Replication) ：为复杂的企业级业务（例如保险的前台>中台>后台>数据仓库）提供有力支持。
- 可支持XtraDB、TokuDB等更高级的存储引擎，引入group commit for the binary log等技术，有效提高业务性能并减少存储量。
- 支持线程池 (Thread pool) 等MySQL、审计日志等企业版拥有的能力。
- 时钟精确到微秒级别，可用于对时间要求精确度较高的金融交易类业务。
- 提供虚拟列 (函数索引) ，可有效提供数据库分析统计运算性能。

4.更高可用性

MariaDB (TDSQL) 的设计旨在提供高于99.99%的可用性，提供双机热备、或一主两备2个备机用于透明的故障转移，还提供故障节点自动修复、自动备份、回档等功能，帮助业务更稳定、安全地运行。

5.更高性能

基于PCI-E SSD、强大IO性能保障数据库的访问能力，存储固件采用NVMe协议，专门针对PCI-E SSD设计，更能发挥出性能优势，高IO型单实例最大可支持6TB容量、480GB内存和22万以上QPS(每秒访问次数)，性能优势让您以较少的数据库实例支撑更高的业务并发。

所有 MariaDB (TDSQL) 实例内核都非原版MariaDB内核，而是经过腾讯顶尖数据库研发基于实际需求修改。而且默认参数都是经过多年的生产实践优化而得，并由专业DBA 持续对其进行优化，确保 MariaDB (TDSQL) 一直基于最佳实践在运行。

6.与 MySQL 兼容

MariaDB (TDSQL) 使用InnoDB存储引擎并与 MySQL5.5、5.6兼容。这意味着，已用于 MySQL 数据库的代码、应用程序、驱动程序和工具，您只需对其进行少量更改甚至无需更改，即可与MariaDB (TDSQL) 配合使用。

7.便宜易用

支持即开即用：您可以通过腾讯云官网MariaDB (TDSQL) 规格定制，下发订单后自动生成MariaDB (TDSQL) 实例。将MariaDB (TDSQL) 配合CVM一起使用，在降低应用响应时间的同时还可以节省公网流量费用。

按需升级：在业务初期，您可以购买小规格的MariaDB (TDSQL) 实例来应对业务压力。随着数据库压力和数据存储量的变化，您可以灵活调整实例规格。

管理便捷：腾讯云负责 MariaDB (TDSQL) 的日常维护和管理，包括但不限于软硬件故障处理、数据库补丁更新等工作，保障 MariaDB (TDSQL) 运转正常。您也可自行通过腾讯云控制台完成数据库的增加、删除、重启、备份、恢复等管理操作。

标准与认证

最近更新时间：2018-10-15 11:48:49

MariaDB(TDSQL)现已代表腾讯云云数据库通过多项国家或国际认证，包括但不限于：

- 软件著作权
- ISO22301认证
- ISO27001认证
- ISO20000认证
- ISO9001认证
- 可信云服务认证
- 信息安全等级保护
- STAR认证

MariaDB(TDSQL)部分功能设计标准参考：

- GBT 20273-2006 信息安全技术 数据库管理系统安全技术要求
- JRT 0072-2012 金融行业信息系统信息安全等级保护测评指南

独享集群数据库解决方案

最近更新时间：2018-05-25 16:40:03

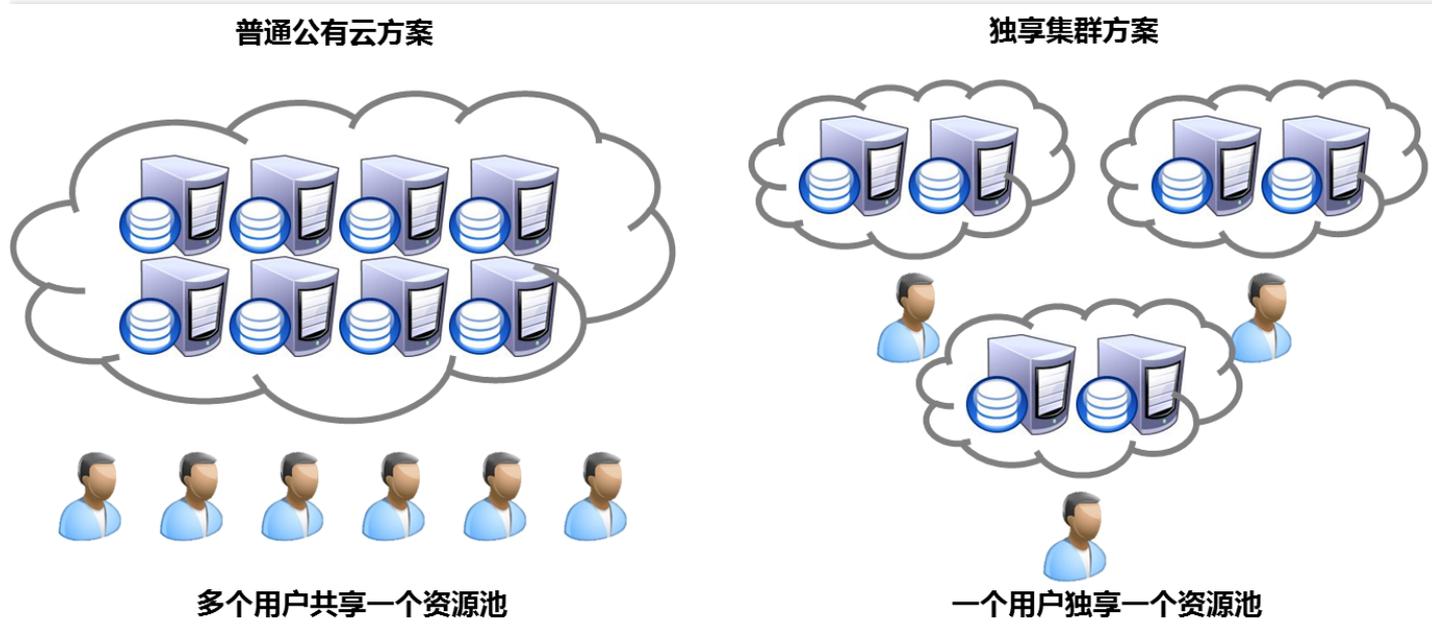
1. 独享集群数据库简介

独享集群数据库（Database Dedicated Cluster），简称独享集群，可以让您以**独享物理集群（完整设备）**资源方式购买、创建数据库，以满足您对资源独享、物理安全、行业监管等需求；购买独享集群后，您可以在其上灵活创建多种自定义规格的云数据库。

- 现已支持：MySQL、TDSQL;
- 现已支持：北京、上海、深圳、广州的公有云、金融云、金融云围拢中部署独享集群数据库；
- 现已支持：HIOY5机型（CPU48核，内存512GB，SSD磁盘7200GB，万兆双网卡）。**请注意：数据库实例实际可用规格仅机型实际配置的70~90%。**

2. 独享集群数据库的架构

独享集群于公有云多租户的方案可以用下图表示：



独享集群享用云数据库的全部能力，但在部署架构上有如下区别：

- 数据库服务器：承载数据库的物理主机，是您使用数据库的主要部分，您将完全独占。

- 腾讯云综合运营管理系统：部署在腾讯云更高等级的管理网络中，是您提供灵活的运维能力的基础，您将与其他用户共享。
- 备份和日志存储：基于HDFS技术架构的备份中心，您可以申请（限金融客户）独占。
- 腾讯云网关/安全网关：部署在数据库前端的提供智能负载均衡和安全管理虚拟网络设备，为您的数据库提供VIP（唯一虚拟IP）、VPC、主备切换、安全防护等能力，您将与其他用户共享。

备注，独享集群（金融围拢方案）与公有云独享集群方案在隔离性上有一定不同，详情请咨询对应架构师。

3.独享集群数据库的使用步骤

1.（以集群设备维度）[购买独享集群](#)，购买时请注意

- 选择与您服务器相同的地域和可用区，以为您提供最佳体验
- 选择数据库内核，详情请参见[TDSQL与MySQL兼容性说明](#)。
- 选择机型与主从架构。目前支持一主多从架构，不同从机数量代表可用性的不同，一主一从理论可用性为99.95%，一主二从理论可用性为99.99%以上。
- 选择购买组数，多组设备可以组成一个集群，进而为集群提供更多冗余资源池已确保可用性。

1.在[集群列表页](#)查看、管理独享集群详情；单击“操作-分配实例”，即可在独享集群上分配数据库实例；



1. 进入MySQL/TDSQL实例列表页，查看刚分配的数据库实例，将[实例初始化](#)后即可正常使用。

如下图所示，标志为 **主实例（独享）** 均为独享集群分配的实例：

关系型数据库 < MySQL-实例列表 全部项目 华南地区 (广州)

云数据库MySQL支持按量计费阶梯价，用得越久越便宜。内存和硬盘可按需搭配。了解详情

+ 新建 对比监控 续费 更多操作

请输入IP(独行分租)或实例名

ID/实例名	监控	状态	实例类型	所属项目	所属地域	配置	数据库版本	操作
cdb-...		运行中	主实例	默认项目	华南地区 (广州)	高IO版 1000MB/25GB/1000QPS	MySQL5.6	登录 管理 升级
cdb- pma_...		运行中	主实例 (独享)	默认项目	华南地区 (广州)	高IO版 2000MB/50GB/2400QPS	MySQL5.6	登录 管理 更多
cdb- cdb1...		运行中	主实例 (独享)	默认项目	华南地区 (广州)	高IO版 4000MB/100GB/4400QPS	MySQL5.6	登录 管理 更多

4. 常见问题

独享集群意味着您独享一个独立的物理资源池，那么您需要知道我们对独享集群的一些建议：

1、为什么需要为每个实例和独享集群本身都预留一定冗余

数据库的可用性将直接影响整个业务系统可用性，我们建议您为每一个实例都预留一定冗余，同时也为独享集群预留一些没有被分配的剩余资源，甚至剩余设备，以备不时之需。

2、为什么大多数分配实例时支持指定设备，部分情况下又无法指定设备？

腾讯云为您提供将实例分配到某组指定设备的能力，而如果您能明确预判每个实例的负载，请将实例负载均匀的分配到不同的设备组中（系统有限制的除外）。然而分布式数据库等应用场景下，节点负载人工可能无法预判，所以

3、冷备数据和日志我应该怎么申请独享

公有云暂不支持冷备备份和日志存储独享。金融围拢方案下，请先联系您架构师。

4、为什么设备配置≠实例最大配置，但我无法使用到100%的资源

操作系统、磁盘raid、系统必要日志空间等都会占用一定设备资源，腾讯云根据多年积累，设定了一个单机最大实例规格值（该值可能调整）

5、独享集群是否支持VPC网络

独享集群不支持，但独享集群分配的实例支持分配到不同的VPC网络。

典型应用场景

最近更新时间：2018-10-15 11:49:12

数据云灾备（异地灾备）

数据是企业运营的重要组成，信息化带来便利的同时，电子数据、存储信息极易毁损、灭失的特点也暴露了出来。而在自然灾害、系统故障、员工误操作和病毒感染面前，任意一次事故就有可能导致企业的运营完全中断，甚至灾难性损失。因此，确保数据的安全完整性，特别是核心数据库的安全和完整性是每个企业必须考虑的。

而企业自建异地数据灾备中心通常耗资巨大，通常包括一笔开支巨大的机房硬件，软件成本，还包括每年运营费用的持续维护投入。而为了小概率事件买单，往往与企业的财务需求不符。

因此，利用云数据库和云接入产品，可以直接建立云上的一套数据灾备中心，将主数据中心的数据，通过安全专网，实时同步到云上的异地备份中心。不仅解决海量数据的运营管理的问题，而且性价比高。

业务系统上云

如果您的业务系统还未上云，那么您可能会遇到这些情况，例如：

- 业务发展非常快，每年若按峰值准备服务器，其增长规模都是非常大的一笔开销。
- 而新业务部门为了，经常需要快速上线新业务，如果每次都要做准备和采购势必影响效率。
- 几乎每个业务系统都遭遇过访问量激增，再创历史新高，后端资源无法支撑的情况。
- 而不少企业领导认为IT部门就是成本中心，每天关心的核心问题不是推进业务，而是一直在解决问题，要不是系统不稳定，要不性能不足。

面对业务系统上云等一系列挑战，腾讯云云数据库通过多年的积累，能够为用户提供：

- 安全、开放的数据库解决方案；
- 提供高可用的方案，采用强同步复制技术和高可用（HA）架构实现高容灾；
- 支持弹性伸缩；

混合云

云数据库TDSQL支持专有云部署方案，可以部署在用户自建机房。业务系统和数据通过专线（或VPN）进行安全同步，构建易扩展的混合云架构。

读写分离

云数据库所有实例的备机均默认支持读写分离策略——即支持备机开放只读；

- 支持通过SQL语法或只读帐号实际只读
- 如果您选择的配置有多个备机，将自动负载的只读策略
- 可通过升级配置，增加更多备机

开发测试

您可能需要维护多个软件版本环境进行测试，甚至需要大量资源进行压力测试。

传统方案是自建服务器和数据库来支撑该需求，然而这样就会浪费大量的硬件资源，因为开发人员不会时时刻刻的使用测试资源，因此测试资源往往是闲置的，而利用云服务器、云数据库的弹性伸缩，可以有效的解决基于测试资源不足或测试资源浪费的问题。

服务等级协议

最近更新时间：2017-06-05 15:18:16

服务等级协议详见《政策与规范》-[云数据库服务等级协议](#)。

金融云用户请联系接口腾讯云商务或架构师同事了解金融定制版（TDSQL）服务等级协议。

分布式版本

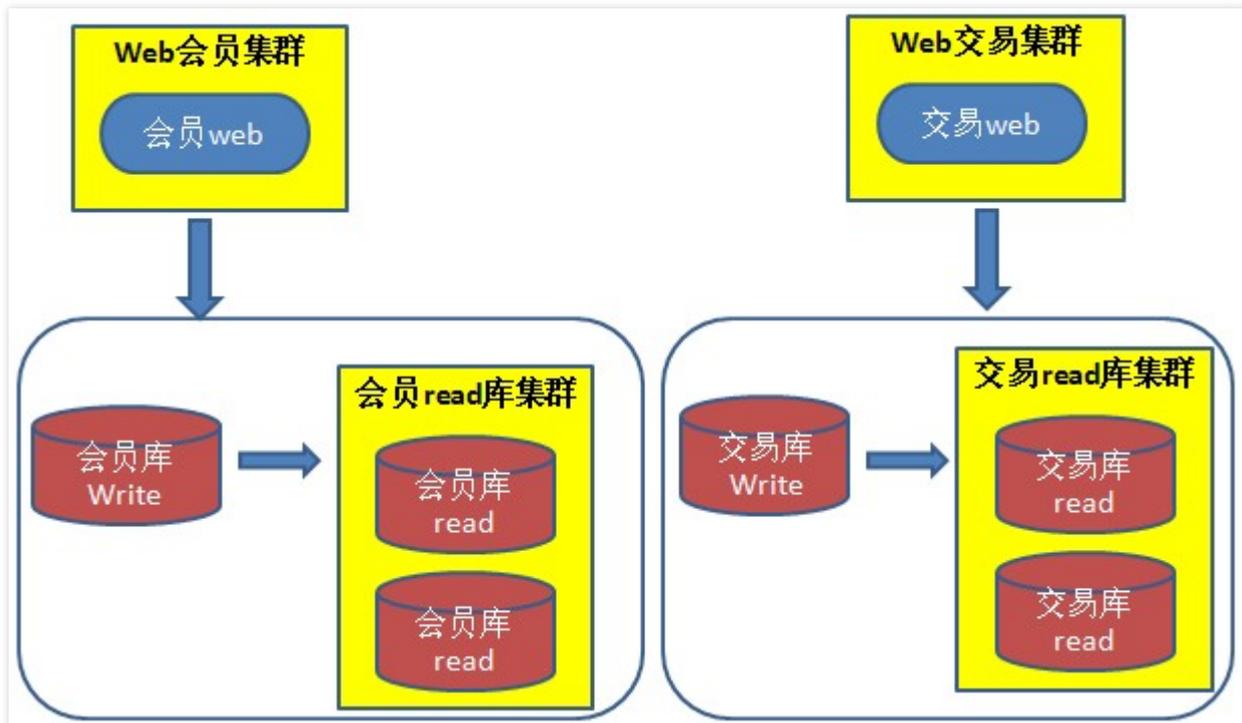
最近更新时间：2018-10-15 11:49:39

1 概述

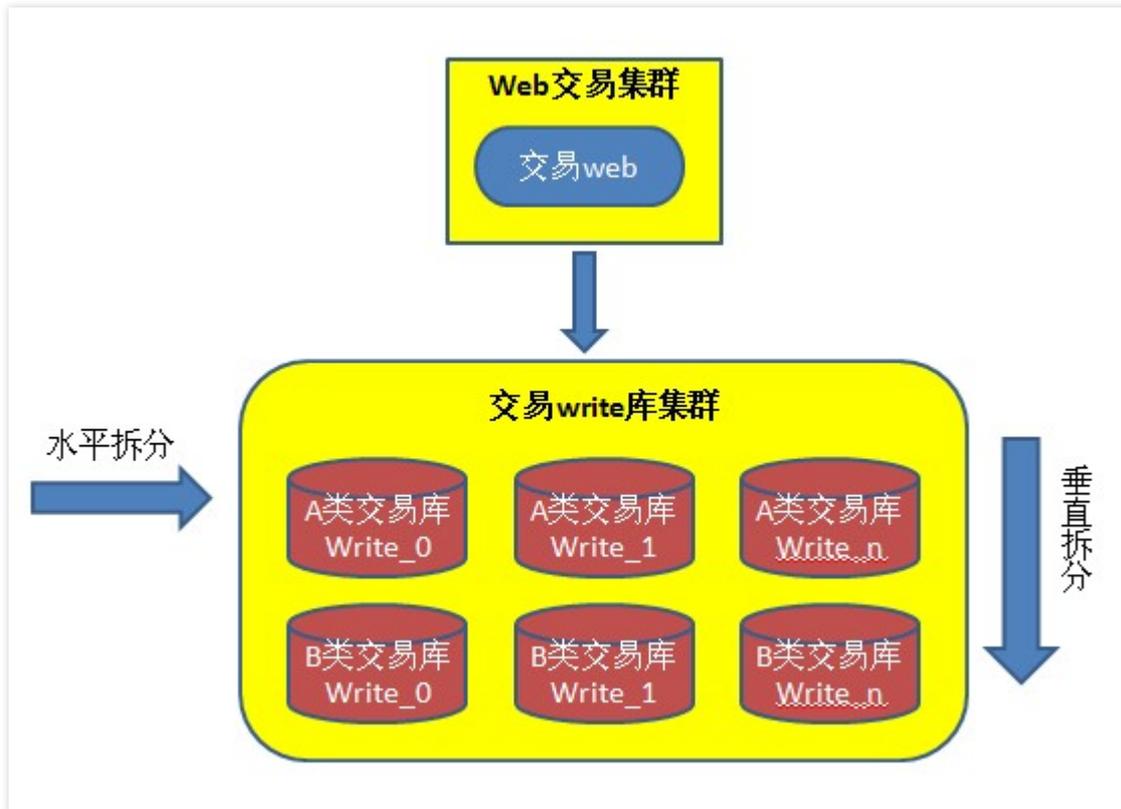
1.1 数据切分/分片(sharding)

在高性能并发互联网架构中，性能瓶颈往往出现在数据库服务器，特别是当业务(用户)达到百万级用户规模以后，通过在数据层进行合理的数据切分（sharding），可以有效解决数据库性能、可伸缩等问题。数据库切分同样是从两个维度考虑：垂直切分(按功能切分)和水平切分。

- **垂直切分**也就是按功能切分，这种切分方法跟业务紧密相关，实施思路也比较直接，比如“京东JD”等电商平台，将数据按功能切分为会员数据库、商品数据库、交易数据库、物流数据库等。而垂直拆分并不能彻底解决压力问题，因为单台数据库服务器的负载和容量也是有限的，随着业务发展势必也会成为瓶颈，解决这些问题的常见方案就是水平切分了。



- **水平切分**是按照某种规则，将一个表的数据分散到多个物理独立的数据库服务器中，这些“独立”的数据库“分片”；多个分片组成一个逻辑完整的数据库实例。



1.2 分片规则

关系型数据库是一个二维模型，数据的切分通常就需要找到一个分片字段（shardkey）以确定拆分维度，再通过定义规则来实现数据库的拆分。如何找到合适的分片规则需要综合考量业务，这里介绍几种常见的分片规则：

- 基于日期顺序。如按年拆分，2015年一个分片，16年一个分片。
 - 优势：简单明了；易于查找
 - 劣势：当期（16年）的热数据的服务器性能可能不足，而存储冷数据性能却闲置。
- 基于用户ID求模，将求模后字段的特定范围分散到不同库中。
 - 优势：性能相对均衡；相同用户数据在一个库中。
 - 劣势：可能导致数据倾斜（如设计的是商户系统，京东一个商户数据能比几千个小商户的数据还多得多）
- 将主键(primary key)求模，将求模后字段的特定范围分散到不同库中。
 - 优势：性能相对均衡；不容易出现数据倾斜的问题；相同主键的数据在一个库中；
 - 劣势：数据随机分散，某些业务逻辑可能需要跨分片join却不能直接支持。

另外，在分片的数据源管理方面，目前也有两种思路：

1. 客户端模式：由业务程序模块中的配置来管理多个分片的数据源，分片的读写与数据整合在业务程序内进行。
2. 中间件代理模式：在分片数据库前端搭建一个中间件代理，后端多个分片数据库对前端应用程序透明。

2 分布式云数据库DCDB for TDSQL

2.1 自动水平拆分（分库分表）

DCDB for TDSQL是部署在腾讯云公有云上的一种兼容MySQL协议和语法，支持自动水平拆分的分布式数据库。分布式数据库即业务获取是完整的逻辑库表，后端却将库表均匀的拆分到多个物理分片节点。目前，DCDB for TDSQL默认部署主备架构且提供了容灾、备份、恢复、监控、迁移等方面的全套解决方案，适用于TB或PB级的海量数据库场景。

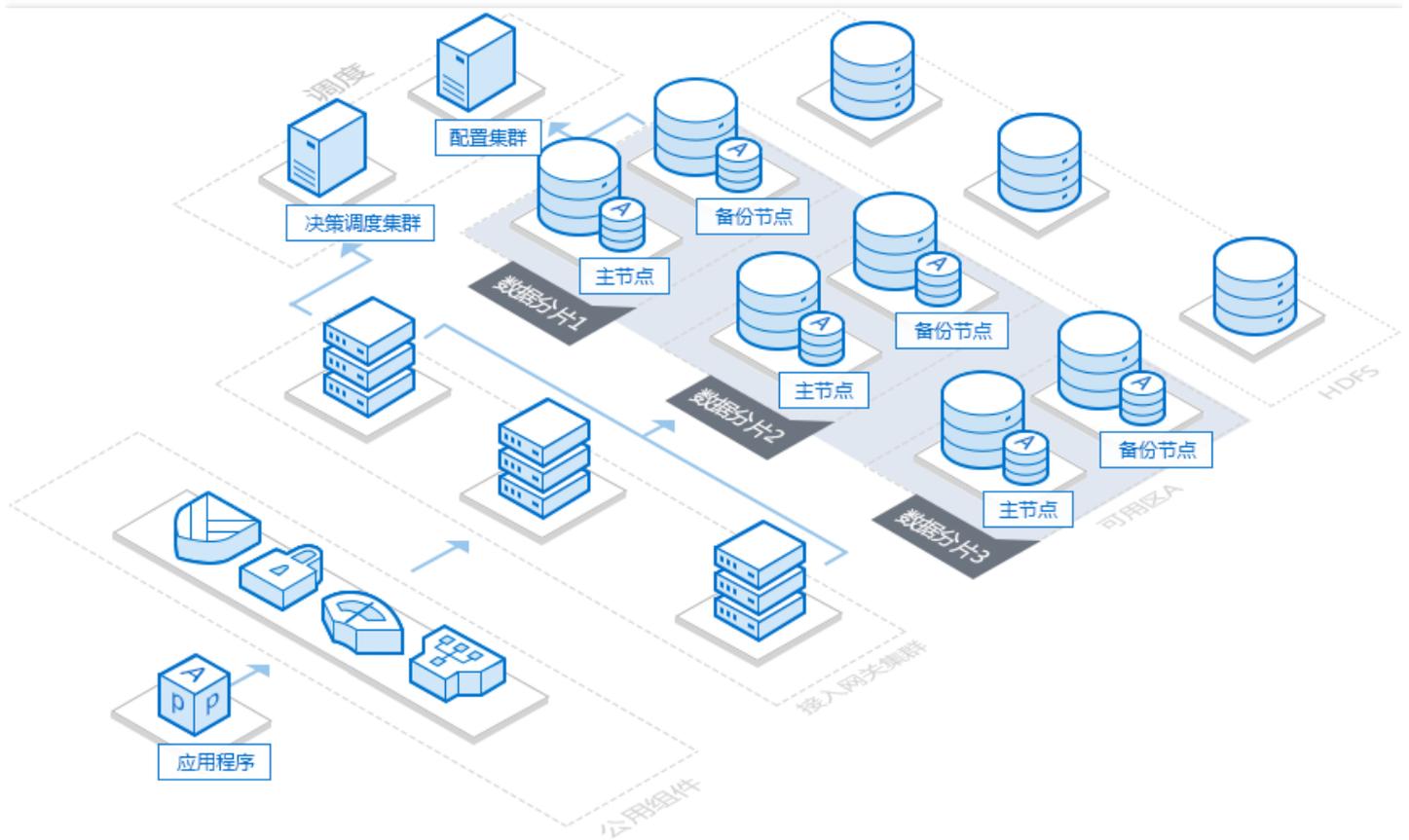
DCDB for TDSQL的历史可以追溯到2004年，腾讯互联网增值业务开始爆发，业务量的爆炸给MySQL数据库带来了巨大的扩容压力，当时即开始引入分库分表机制来解决难题——针对大的表，按照ShardKey预先拆成多个子表，分布在不同的物理机器节点上。而时至今日，TDSQL后端存储数据量已经是庞然大物，以“米大师”为例，分布式数据库TDSQL承载米大师各个渠道的100亿的账户，用户数量接近9亿，日交易金额超过10亿人民币。

为什么分布式数据库TDSQL能够轻松支撑起海量业务，笔者总结其有如下优势：

- **自动拆分表**：TDSQL支持自动的分库分表，结合数据的统一调度机制，实现按需的容量伸缩。因为TDSQL通过网关屏蔽了内部的分库分表细节，对于开发人员来说，不再需要关心如何切分数据、如何路由请求等待，只需初始化分片字段（shardkey），直接面向逻辑库表进行编程、专注业务逻辑的实现即可，大大降低了程序的复杂度。
- **自动容灾切换**：无论是物联网，还是大数据，或是支付服务，任何存储了海量数据的业务都对后台存储数据库的可用性要求非常高。而通常解决方案是，容灾切换需要业务检测和配合，与业务程序深度耦合，而且切换过程复杂，甚至需要人工介入。业务恢复后还需要对切换过程中可能出现的错误数据进行手工修复，运维起来非常痛苦。TDSQL数据节点和网关实现多点容灾，自动检测实例的运行状态。当发现主节点不可用时，会自动触发主备容灾切换流程，保障在主机故障、网络故障、IDC故障等灾难情况下数据库的高可用性。这个容灾切换过程对业务完全透明，且无需人工干预，在确保用户体验的同时，也极大的简化了运维工作。
- **数据高一致性保障**：如果您对数据的丢失或错乱零容忍。TDSQL在MySQL原有异步复制和半同步复制的基础上创新的使用了多线程强同步复制机制，确保每一笔交易在返回用户应答前，在集群中至少有两个拷贝，然后再通过一系列切换机制保障在节点故障发生切换时，数据不会丢失或错乱。
- **集群化管理，自动扩容**：业务的请求量会因上线新功能或者做营销活动等原因，峰值请求量激增至平时的几倍甚至几十倍。之前，需要提前知晓业务测的动向，由DBA提前手工扩容，通常，大多数分布式数据库扩容过程繁杂，且中间过程手工操作多，极容易出错。而TDSQL在集群层面实现了自动部署、自动容量伸缩、自动备份恢复、数据定点回档以及多维监控等功能。当扩容需求时，DBA只需要在前台单击按钮发起扩容流程，即可自动完成扩容工作。TDSQL的集群化运营体系，极大提升了DBA的工作效率，同时大大降低了手工操作可能带来的失误。

2.2 产品架构

DCDB for TDSQL一个实例的基本架构如下：



数据分片：兼容MySQL的数据库引擎、监控和信息采集（Tagent）组成。

请注意，分布式数据库默认为每个分片配置两个节点，即“一主一备”，每个分布式数据库需至少两个分片。

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。

接入网关集群（TProxy）：在网络层连接管理SQL解析、分配路由，即可以理解为开源分布式数据库的中间件。

请注意，为不让proxy成为性能瓶颈，分布式数据库proxy数量通常等于分片数量。

备份集群：腾讯云云数据库数据备份集群。

请注意，分布式数据库备份默认保存5天，两份。