腾讯云数据仓库 TCHouse-C

产品简介

腾讯云

版权所有：腾讯云计算（北京）有限责任公司

【版权声明】

©2013-2026 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【商标声明】

腾讯云

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。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【服务声明】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。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。

【联系我们】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或95716。

文档目录

产品简介
- 产品概述
- 基本概念
- 部署形态与版本对比
- 集群架构
- 产品优势
- 应用场景

产品简介
产品概述

最近更新时间：2026-09-08 17:37:00

腾讯云数据仓库 TCHouse-C 是基于开源 ClickHouse 构建的全托管实时分析型数据仓库，面向 PB 级海量数据场景，提供亚秒级多维分析查询能力。依托 MPP（大规模并行处理）架构，单查询峰值吞吐可达每秒数 TB，查询性能最高可达传统数据仓库的 10 倍以上。
如果需要一个开箱即用、弹性扩展、成本可控的 ClickHouse 云端集群，TCHouse-C 是腾讯云为此量身打造的托管方案。

核心能力

分钟级交付，零运维负担

从下单到集群就绪，最快 5 分钟完成 TB 至 PB 级数据仓库的部署。控制台提供一站式集群管理——节点管理、规格变配、参数调优、版本升级均可在线完成，省去自建集群的部署调试和日常运维开销。

卓越查询性能

基于 ClickHouse 列式存储引擎和向量化执行框架，结合分布式 MPP 并行计算架构，充分利用多核 CPU 和高速 I/O 能力：

- **查询速度**：同等数据规模下，查询性能最高可达传统行式数据仓库的 10 倍以上。
- **峰值吞吐**：单查询峰值扫描速率可达每秒数 TB。
- **压缩效率**：数据压缩比可达 10:1，大幅降低存储占用。

存储计算分离架构

支持存储计算一体与存储计算分离两种架构，存储计算分离架构下支持对计算节点和存储节点进行独立配置与弹性伸缩：

- **计算层**：可根据查询负载按需扩缩计算节点，无需搬迁数据。
- **存储层**：独立数据存储，完全按实际数据量付费。
- **成本优势**：无需多副本机制即可保证数据高可用，利用对象存储进一步降低存储成本。

可视化运维监控

提供完善的集群监控面板，覆盖 CPU、内存、磁盘 I/O 等近百项核心指标，支持自定义告警规则，帮助团队实时掌握集群健康状态，快速定位性能瓶颈。

企业级安全保障

- **网络隔离**：每个集群独立部署于用户 VPC 内，通过私有网络隔离保障数据访问安全。

版权所有：腾讯云计算（北京）有限责任公司

- **数据可靠**：支持多副本冗余存储，异常情况下自动容灾切换，用户无感知完成故障恢复。
- **合规加密**：支持云硬盘加密（部分地域），保障静态数据安全。

部署形态

云数据仓库 TCHouse-C 提供标准版与弹性版两种部署形态，您可在创建集群时按业务需要选择：

- **标准版（存算一体）**：计算与存储一体部署，架构简洁、查询性能稳定，适合数据规模与负载相对稳定的分析场景。
- **弹性版（存算分离）**：计算节点无状态、数据持久化到共享托管存储，支持以计算资源组为维度独立弹性伸缩。

相比标准版，弹性版在以下方面更具优势：

- 计算弹性伸缩：按计算资源组维度快速扩缩计算节点，从容应对业务高峰与低谷；
- 存储按量、成本更优：海量数据落共享托管存储，按存储资源用量计费，可用存储容量包抵扣；
- 高可用、免数据迁移：计算节点无状态，故障时快速替换、无需搬迁数据。

两种形态的详细差异与选型建议，请参见 部署形态与版本对比 。

适用对象

角色	典型诉求
数据工程师	需要高吞吐的实时数仓，替代离线 T+1 报表方案
数据分析师	需要亚秒级 Ad-Hoc 查询，支持多维灵活分析
运维 / SRE	需要 PB 级日志的实时检索和监控告警
技术管理者	需要降低自建 ClickHouse 的运维成本和稳定性风险

相关产品

如果您在评估不同的数据分析方案，以下腾讯云产品也值得了解：

- TCHouse-D：基于 Apache Doris 内核构建的企业级分析型数据仓库，适合实时多表关联和复杂 SQL 场景。
- 数据湖计算 DLC：Serverless 数据湖分析与计算服务，适合多源异构数据联合查询。
- 流计算 Oceanus：实时流处理服务，可作为 TCHouse-C 的上游实时数据管道。

版权所有：腾讯云计算（北京）有限责任公司

基本概念

最近更新时间：2026-09-17 17:59:30

本文介绍使用 TCHouse-C 前需要了解的核心概念。分为**基础设施层**和**数据引擎层**两个部分，帮助您快速建立对产品的整体认知。

基础设施概念

地域（Region）

指 TCHouse-C 集群所在的物理地理区域（如北京、上海、广州、新加坡等）。不同地域之间网络完全隔离，跨地域数据访问需通过公网或专线。集群创建后**不可更换地域**，请根据业务用户分布就近选择。

可用区（Availability Zone）

同一地域内电力和网络相互独立的物理数据中心，在同一地域内选择不同可用区可提升容灾能力。

集群（Cluster）

TCHouse-C 的基本服务单元。一个集群由以下组件构成：

组件	说明
ClickHouse 计算节点	负责数据存储和查询计算，可包含 1 个或多个分片
ZooKeeper/ClickHouse Keeper 管理节点	负责分布式协调，管理副本同步、DDL 操作等元数据
CLB 负载均衡	提供统一的集群访问入口，自动分发查询请求

分片（Shard）

TCHouse-C 将海量数据水平切分后分散到不同节点，每个节点承载数据的一部分。分片是数据水平扩展的基本单位。

模式	分片与节点的关系
单副本（非高可用）	1 个分片 = 1 个节点
双副本（高可用）	1 个分片 = 2 个节点（互为副本）

副本（Replica）

为保障数据安全和服务高可用，TCHouse-C 支持将同一分片的数据冗余存储在两个节点上。两个节点互为副本，数据实时同步。

高可用（High Availability）

高可用模式下，每个分片配置两个副本。当某个节点发生故障时，查询自动路由到副本节点，实现**用户无感知的故障切换**，保障服务连续性。

> **说明：**
> 生产环境请务必选择高可用模式。非高可用模式（单副本）适用于开发测试或对可用性要求不高的场景。

数据引擎概念

以下概念源自 ClickHouse 数据引擎，是理解 TCHouse-C 数据模型的关键。

表引擎（Table Engine）

表引擎决定了数据如何存储、索引、是否支持副本等核心行为。TCHouse-C 支持 ClickHouse 的完整引擎家族，最常用的包括：

引擎	说明	适用场景
`MergeTree`	基础列式存储引擎，支持主键排序、分区、数据生命周期管理	单副本场景下的通用分析表
`ReplicatedMergeTree`	MergeTree 的副本版本，通过 ZooKeeper 实现跨节点数据同步	高可用场景下的标准选择
`Distributed`	分布式路由表，将查询分发到所有分片并汇总结果	跨分片的全局查询入口
`MaterializedView`	物化视图，数据写入时自动触发聚合计算并存储结果	预聚合加速、实时指标计算

本地表 vs 分布式表

类型	说明
本地表（Local Table）	实际存储数据的表，每个分片上各有一份，使用 `MergeTree` 或 `ReplicatedMergeTree` 引擎
分布式表（Distributed Table）	逻辑路由表，不存储数据，将查询自动分发到各分片的本地表并合并结果

> **说明：**
> 数据写入推荐直接写入本地表（或通过分布式表路由写入），查询通过分布式表进行，以获得全局视图。

版权所有：腾讯云计算（北京）有限责任公司

分区（Partition）

在表引擎层面，数据可按某个字段（通常是日期）进行分区。分区的好处：

- 查询时自动裁剪无关分区，减少扫描数据量；
- 支持按分区快速删除历史数据（`ALTER TABLE ... DROP PARTITION`）；
- 分区粒度建议按月或按天，避免分区数过多导致元数据膨胀。

数据压缩

ClickHouse 默认对列式数据进行高效压缩，TCHouse-C 支持 LZ4（默认，速度优先）和 ZSTD（压缩比优先）两种算法。实测典型业务数据压缩比可达 **10:1**，大幅降低存储成本。

概念关系总览

```
--基础设施层（集群 = 三层架构）--
集群 (Cluster)
├── 接入层：CLB 负载均衡 + TCP/MySQL/HTTP/HTTPS 端口
├── 计算层：
│   ├── 分片 1 (Shard 1) - 副本 A、B（计算节点， 本地表 + 数据）
│   ├── 分片 2 (Shard 2) - 副本 A、B（计算节点， 本地表 + 数据）
│   └── .......
└── 协调层：ZooKeeper / ClickHouseKeeper x 3

--数据引擎层（逻辑对象、运行于集群之上）--
分布式表（Distributed Table）
├── 引擎：Distributed
├── 不存储数据
├── 写入：按 sharding_key 路由到各分片的本地表
└── 查询：分发到所有分片并执行，汇总结果返回
```

术语字典

术语	说明
标准版	存算一体部署形态，计算与存储一体部署。
弹性版	存算分离部署形态，计算节点无状态、数据持久化到共享托管存储，支持按计算资源组弹性伸缩。
计算资源组	弹性版中计算能力的组织与弹性伸缩单元；用户可新建自定义资源组以隔离不同业务负载。命名规范：6-24 个字符，仅支持小写字母/数字/下划线，以字母或数字开头和结尾，创建后不可修改。

版权所有：腾讯云计算（北京）有限责任公司

共享托管存储	弹性版持久化数据的存储层，多个计算节点共享访问，按存储资源用量计费。
存储资源用量	弹性版托管存储中实际占用的数据量，作为存储计费依据；因统计机制存在约 2~3 小时的更新延迟。
存储容量包	用于抵扣存储资源用量的预付费资源包，购买后按容量抵扣托管存储用量。
Shard（分片）	数据的逻辑分片。弹性版采用单 Shard 结构。
Replica（副本）	计算副本，承载查询与写入。弹性版副本由系统统一管理。
缓存空间	弹性版计算节点本地盘/云硬盘，用于缓存热数据加速查询，不持久化数据本体。
Common 节点	承载协调/公共服务的节点，弹性版仅支持 ClickHouse Keeper。
ClickHouseKeeper / ZooKeeper	分布式协调服务，管理副本元数据与一致性。弹性版强制高可用部署。
冷热分层	依数据访问热度自动分层存储的能力，弹性版默认开启。

版权所有：腾讯云计算（北京）有限责任公司

部署形态与版本对比

最近更新时间：2026-10-12 10:49:00

云数据仓库 TCHouse-C 提供标准版与弹性版两种部署形态，分别面向不同的业务规模与弹性诉求。本文帮助您快速了解两种形态的核心差异，选择合适的部署形态。

形态总览

部署形态	一句话定位
标准版	存算一体架构，计算与存储部署在同一节点，性能稳定、部署简单，适合数据规模与负载相对稳定的分析场景。
弹性版	存算分离架构，计算节点无状态、数据持久化到共享托管存储，支持以计算资源组为维度独立弹性伸缩，适合负载波动大、数据量大、对弹性和成本敏感的场景。

核心对比

对比维度	标准版	弹性版
架构特征	存算一体（Shared-Nothing）	存算分离（Shared-Storage）
数据存储	数据存储在节点本地云盘	数据持久化到共享托管存储，节点本地盘仅作缓存加速
分片与副本	多 Shard、每 Shard 可多副本，按分片打散数据	单 Shard 多 Replica，副本由系统统一管理
计算弹性	以节点/副本为单位整体扩缩，扩容涉及数据重分布	以计算资源组为维度独立弹性伸缩，计算节点无状态，扩缩快、无需数据搬迁
计算资源组	不涉及	支持多个计算资源组，读写负载可按组隔离
冷热分层	可选配置	默认开启（数据天然分层至托管存储）
高可用	依副本配置	计算节点无状态，故障可快速替换；协调服务（Keeper）强制高可用
存储计费	随节点云盘计费	按存储资源用量计费，可购买存储容量包抵扣
计算计费	节点包年包月/按量	按计算资源组包年包月/按量计费
典型场景	稳定负载的实时分析、BI 报表	潮汐/大促弹性负载、海量数据低成本分析、高可用在线分析

版权所有：腾讯云计算（北京）有限责任公司

如何选择

- **选择标准版：**业务负载与数据量相对稳定，追求高性能查询性能与简单运维，无需频繁弹性伸缩。
- **选择弹性版：**负载存在明显波峰波谷、需要按需弹性扩缩计算资源；数据量大、希望存储按量付费降低成本；对节点故障快速恢复、免数据迁移有较高要求。

> ⓘ **说明：**
> 两种形态在 购买页 “**产品版本**”处切换选择，创建后不支持形态互转。

相关文档

- 集群架构
- 计费概述
- 快速入门

集群架构

最近更新时间：2026-09-08 17:37:02

TCHouse-C 基于开源 ClickHouse 构建，在保持与社区版架构一致的前提下，针对**云端场景**在易用性、安全性和稳定性方面进行了深度增强。

标准版架构

集群架构拓扑图展示

TCHouse-C 标准版集群由三个核心层组成：

接入层

组件	职责
负载均衡（CLB）	提供集群统一访问地址，自动将连接分发到可用的计算节点
TCP 端口（9000）	ClickHouse 原生协议，适用于 `clickhouse-client` 及各语言 SDK
MySQL 端口	使用标准的 MySQL 命令行工具或编程语言的 MySQL 驱动来查询

（9004）	ClickHouse
HTTP 端口（8123）	HTTP 端口，适用于 REST 调用、JDBC/ODBC 连接

计算存储层

计算层由多个**分片（Shard）**组成，每个分片包含 1~2 个**副本（Replica）**节点：

- **数据存储**：每个副本节点上运行 ClickHouse Server 进程，管理本地表数据的写入、压缩、索引和查询。
- **并行计算**：查询通过分布式表（Distributed Table）自动分发到所有分片并行执行，最终汇总结果返回。
- **弹性伸缩**：支持在线增加分片数量以扩展计算和存储容量。

计算节点规格选型参考：

节点类型	节点存储	适用场景
标准型（**推荐**）	支持挂载增强型 SSD 云硬盘和高性能云硬盘，支持独立扩容、最大单节点支持 320TB，支持透明数据加密（部分地域），写入云硬盘的数据自动加密。	大多数分析场景
高性能型	搭载 NVMe SSD 本地盘，提供更高 I/O 性能，节点磁盘空间固定、不支持独立扩容磁盘。	对 I/O 延迟极敏感的高频查询
大存储型	搭载大容量 HDD 本地盘，节点磁盘空间固定、不支持独立扩容磁盘。	海量冷数据存储、日志归档

协调层（ZooKeeper 或 ClickHouse Keeper）

TCHouse-C 使用 3 节点 ZooKeeper 或 ClickHouse Keeper 集群负责分布式协调：

- **副本同步**：协调 ReplicatedMergeTree 引擎的跨副本数据同步。
- **DDL 管理**：确保分布式 DDL 操作（如建表、改表）在所有节点上一致执行。
- **Leader 选举**：管理合并（Merge）任务的 Leader 选举。

> ⓘ **说明：**
> ZooKeeper 或 ClickHouse Keeper 节点由 TCHouse-C 全托管，用户无需关心其部署和运维。

高可用 vs 非高可用架构对比

维度	高可用模式	非高可用模式
副本数	每分片 2 副本	每分片 1 副本
表引擎	`ReplicatedMergeTree`	`MergeTree`

故障切换	自动：节点故障时查询自动路由到副本	不支持：节点故障则分片不可用
数据安全	双副本冗余，单节点损坏不丢数据	依赖云硬盘三副本保障（非本地盘）
成本	约为非高可用模式的 2 倍	更低
推荐场景	**生产环境**	开发测试、非关键业务

弹性版架构

弹性版采用存算分离架构，整体由计算层、存储层、协调层三部分组成，计算与存储实现解耦：

- 计算节点无状态、可弹性伸缩，数据统一持久化到共享托管存储；
- 查询与写入请求进入计算副本后，经 ReplicatedMergeTree 引擎写入并通过协调服务在副本间同步元数据，数据持久化到共享托管存储；
- 查询时优先命中本地缓存空间，未命中则从托管存储读取并回填缓存。

计算层：单 Shard 多 Replica

- 弹性版采用单 Shard 多 Replica 结构：一个逻辑分片下部署多个计算副本（Replica），副本由系统统一管理，无需用户手动配置。
- 计算节点无状态：节点本地云盘仅作为缓存空间（File Cache）加速查询，数据本体不落本地盘，因此节点可被快速替换、快速扩缩。
- 计算能力以计算资源组为单位组织：集群创建时默认生成 `default_cluster`，用户可按需新建多个计算资源组，将不同业务的读写负载隔离到不同资源组。

存储层：共享托管存储

- 集群数据统一持久化到共享托管存储，多个计算节点/资源组共享访问同一份数据；
- 存储容量按存储资源用量弹性计费，无需为存储预留固定节点云盘，海量数据成本更优；
- 冷热分层默认开启：数据依访问热度自动分层，热数据经本地缓存加速、冷数据沉淀至托管存储。

协调层：强制高可用

- 弹性版依赖分布式协调服务（默认支持 ClickHouse Keeper）管理副本元数据与一致性，采用高可用部署（≥3 节点），保障元数据服务不中断。
- 计算副本基于 ReplicatedMergeTree 引擎 + 协调服务实现数据一致性与自动同步。

标准版 vs 弹性版架构对比

架构维度	标准版	弹性版
架构模式	存算一体	存算分离

版权所有：腾讯云计算（北京）有限责任公司

分片结构	多 Shard	单 Shard 多 Replica
数据存储位置	节点本地磁盘（本地盘/云硬盘）	共享托管存储（本地磁盘作缓存）
计算弹性单元	节点 / 副本	计算资源组
扩缩容	• 新增/减少 Shard 分片 • 扩容涉及数据重分布，需手动 rebalance 历史数据	• 增减计算资源组 • 计算节点无状态，扩缩无需搬数据
协调服务	双副本强制配置	必须高可用（≥3）
冷热分层	可选	默认开启

云端增强能力（相比自建 ClickHouse）

TCHouse-C 并非简单的 ClickHouse “搬上云”，而是在以下方面做了深度增强：

能力维度	TCHouse-C	自建 ClickHouse
部署交付	控制台一键创建，分钟级就绪	需手动部署配置，通常耗时数小时
集群管理	控制台可视化管理，内置监控告警	需自行编写运维脚本和监控
版本升级	支持滚动升级，降低业务影响	需停机手动升级，风险较高
弹性扩容	控制台一键操作，支持**存算一体**和**存算分离**两种架构	扩容需手动添加节点、迁移数据
高可用	内置 Keeper 全托管，高可用一键开启	需自行搭建 ZooKeeper、配置副本
安全合规	VPC 隔离 + 云硬盘加密，开箱即用	需自行配置网络隔离和加密

版权所有：腾讯云计算（北京）有限责任公司

产品优势

最近更新时间：2026-09-08 17:37:02

选择 TCHouse-C 而非自建 ClickHouse 或其他方案的核心理由。

卓越查询性能

核心指标：查询性能最高可达传统行式数据仓库的 **10 倍以上**，单查询峰值扫描可达 **每秒数 TB**。
技术根基：

- 列式存储 + 向量化执行引擎，充分利用现代 CPU 的 SIMD 指令集。
- MPP 分布式架构，查询自动分发到所有分片并行执行。
- 智能索引（主键排序、跳数索引），精准定位目标数据块，避免全表扫描。

分钟级交付，零运维负担

核心指标：从下单到集群可用，最快 **5 分钟**。
对比自建：

操作	TCHouse-C	自建 ClickHouse
初始部署	分钟级，控制台一键完成	数小时~数天（含 ZK、网络、监控）
日常运维	全托管，内置健康检查和自动告警	需专人维护，处理 OOM、合并阻塞、磁盘告警
版本升级	支持滚动升级，自动兼容性检查	停机手动升级，回退复杂
参数调优	控制台在线修改，一键应用到集群	修改配置文件后逐节点重启

灵活弹性伸缩

支持在线扩缩容，业务不停服：

- **水平扩展：**增加分片数量，线性提升存储容量和并行计算能力。
- **垂直升配：**单节点规格升级（CPU / 内存），提升单查询性能。
- **存储独立扩容：**云硬盘容量可独立调整，不需要变更计算规格。
- **缩容支持：**业务低谷期可缩减节点降低成本。

数据压缩，成本更低

核心指标：数据压缩比可达 **10:1**，综合使用成本可比传统方案降低 **50% 以上**。
成本优势构成：

成本项	如何省

存储成本	ClickHouse 列式压缩算法（LZ4/ZSTD），10TB 原始数据仅占约 1TB 磁盘
计算成本	按需选择节点规格，支持存储计算分离，避免资源捆绑浪费
运维成本	全托管架构，无需配备专职 DBA / 运维人员
弹性成本	支持按量计费 + 弹性伸缩，低峰期自动省钱

企业级安全与高可用

核心能力：生产环境开箱即用的安全和容灾体系。

安全保障

- **网络隔离**：集群部署在用户 VPC 内，通过安全组精细控制访问权限。
- **数据加密**：支持云硬盘透明加密（AES-256），保障静态数据安全。
- **访问控制**：支持 ClickHouse 原生的用户 / 角色权限体系。

高可用保障

- **双副本冗余**：高可用模式下每分片两副本，副本数据实时同步。
- **自动故障切换**：节点故障时查询自动路由到副本，服务不中断。
- **跨可用区部署**：副本可分布在不同可用区，应对机房级故障。

弹性版专属优势

计算弹性伸缩

以**计算资源组**为维度快速扩缩计算节点：业务高峰时增加节点分摊读写压力，低谷时缩容节省成本。计算节点无状态，扩缩过程无需数据搬迁，弹性响应快。

低成本海量存储

数据持久化到**共享托管存储**，按**存储资源用量计费**，无需为存储预留固定节点云盘；配合冷热分层，冷数据自动沉淀，海量数据分析成本更优。

高可用免迁移

计算节点无状态，节点故障时可快速替换、无需数据迁移；协调服务强制高可用部署，保障元数据服务持续可用。

负载隔离

可创建多个计算资源组，将不同业务（如实时写入与离线查询）隔离到不同资源组，互不争抢资源。

版权所有：腾讯云计算（北京）有限责任公司

应用场景

最近更新时间：2026-04-21 18:26:52

TCHouse-C 适用于需要对海量数据进行实时多维分析的各类场景。以下是六个典型应用方向。

用户行为分析

业务痛点

网站、App、游戏等产品积累了海量用户行为数据（点击、浏览、停留时长等），传统 T+1 离线分析无法满足实时决策需求，且面对数百维度的灵活查询时响应缓慢。

TCHouse-C 方案

- 将用户行为数据实时写入 TCHouse-C，构建包含数百个属性列的用户特征大宽表。
- 借助列式存储和向量化引擎，任意维度组合的 Ad-hoc 查询均可亚秒级响应。
- 支持漏斗分析、留存分析、路径分析等典型分析模式。

效果

- **查询延迟**：10 亿行大宽表多维聚合查询 < 1 秒。
- **业务价值**：精准营销决策从 T+1 提速到实时，显著提升用户转化率。

企业经营分析与实时报表

业务痛点

企业需要对经营数据（销售额、订单量、库存、财务指标等）进行多维度实时分析，传统数据仓库查询慢，无法支持管理层随时随地的个性化分析需求。

TCHouse-C 方案

- 业务数据库通过 CDC 实时同步至 TCHouse-C。
- 构建多层数据模型（ODS → DWD → DWS → ADS），利用物化视图自动完成实时预聚合。
- 对接 Grafana、Superset 等 BI 工具，构建实时经营看板。

效果

- **分析体验**：数亿行数据、数百维度自由组合，查询响应亚秒级。
- **业务价值**：经营数据「随问随答」，辅助管理层快速决策。

实时日志分析

业务痛点

运维团队需要对应用日志、系统日志、安全审计日志进行实时检索和分析。日志量每天数十 GB 到数 TB，传统方案（如 ELK）在超大规模下存储成本高、查询性能下降明显。

TCHouse-C 方案

- 日志数据按日期分区存储，自动淘汰过期分区，存储成本可控。
- ClickHouse 的字符串搜索、正则匹配、全文检索性能远超通用搜索引擎。
- 压缩比高达 10:1，大幅降低日志存储成本。

效果

- **查询速度**：TB 级日志的关键词搜索秒级返回。
- **成本节约**：相比其他开源方案，存储成本可降低 60%+。

实时数仓与 ETL 管道

业务痛点

传统离线数仓（Hive / Spark）只能提供 T+1 数据，无法满足业务对实时数据的需求。而流处理框架（Flink / Spark Streaming）擅长实时计算但不擅长交互式查询。

TCHouse-C 方案

- 作为实时数仓的存储和查询引擎，承接 Flink / Spark 的实时计算结果。
- 利用物化视图实现轻量级实时 ETL（如实时去重、聚合、宽表关联）。
- 同时支持实时写入和高并发分析查询，兼顾数据新鲜度和查询性能。

效果

- **数据延迟**：从 T+1 缩短到秒级 ~ 分钟级。
- **架构简化**：减少 Lambda 架构中的批处理链路，降低系统复杂度。

广告投放与归因分析

业务痛点

广告投放业务需要对海量曝光、点击、转化数据进行实时归因和效果分析。数据量大（日均数十亿条）、维度多（渠道、素材、地域、设备等），且需要实时调整投放策略。

TCHouse-C 方案

- 实时接入广告曝光和转化数据，构建归因分析宽表。
- 利用 ClickHouse 的 `AggregatingMergeTree` 引擎实现实时 UV / PV 去重计算。
- 支持多种归因模型的灵活查询（首次点击、末次点击、线性归因等）。

版权所有：腾讯云计算（北京）有限责任公司

效果

- **分析实时性**：投放效果秒级可查，快速优化 ROI。
- **处理能力**：日均数十亿行数据的实时写入和查询无压力。

IoT 与时序数据分析

业务痛点

物联网场景下，传感器、设备产生持续高频的时序数据，需要实时监控设备状态、检测异常，同时支持历史趋势分析。

TCHouse-C 方案

- 时序数据按时间分区写入 TCHouse-C，利用列式压缩大幅降低存储开销。
- 利用 ClickHouse 丰富的时间函数和窗口聚合能力，实现实时趋势分析和异常检测。
- 配合 TTL（数据生命周期管理）自动清理过期数据，存储成本可控。

效果

- **写入吞吐**：单集群支持百万级 TPS 高频写入。
- **查询响应**：亿级时间点的趋势聚合查询秒级返回。

场景选型速查

场景	数据规模	实时性要求	推荐集群规格
用户行为分析	10 亿+ 行	秒级	标准型，≥4 分片
企业经营分析	1 亿~10 亿行	秒级 ~ 分钟级	标准型，2~4 分片
实时日志分析	日增 TB 级	秒级	标准型 / 高性能型，≥4 分片
实时数仓	10 亿+ 行	秒级	标准型 / 高性能型，≥4 分片
广告归因分析	日增数十亿行	秒级	高性能型，≥8 分片
IoT 时序分析	百亿+ 数据点	秒级	标准型，≥4 分片

版权所有：腾讯云计算（北京）有限责任公司