

阿里云 表格存储Table Store

产品简介

文档版本：20191024

法律声明

阿里云提醒您在使用或阅读本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云文档中所有内容，包括但不限于图片、架构设计、页面布局、文字描述，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 禁止： 重置操作将丢失用户配置数据。
	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告： 重启操作将导致业务中断，恢复业务时间约十分钟。
	用于警示信息、补充说明等，是用户必须了解的内容。	 注意： 权重设置为0，该服务器不会再接受新请求。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明： 您也可以通过按Ctrl + A选中全部文件。
>	多级菜单递进。	单击设置 > 网络 > 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
##	表示参数、变量。	<code>bae log list --instanceid Instance_ID</code>
[]或者[a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ }或者{a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

法律声明.....	I
通用约定.....	I
1 什么是表格存储.....	1
2 适用场景.....	4
3 常见问题.....	9
4 专家服务.....	12

1 什么是表格存储

表格存储（Table Store）是阿里云自研的NoSQL多模型数据库，提供海量结构化数据存储以及快速的查询和分析服务。表格存储的分布式存储和强大的索引引擎能够支持PB级存储、千万TPS以及毫秒级延迟的服务能力。

为什么选择表格存储

- 全托管

表格存储是一种全托管的数据库。使用表格存储您只需专注于业务研发，无需担心软硬件预置、配置、故障、集群扩展、安全等问题，在保证高服务可用性的同时，极大地减少了管理及运维成本。

- 模型丰富

表格存储支持多种数据库模型，包括Wide column、Timeline、Timestream、Grid。其中，Wide column模型是一款经典模型，目前绝大部分半结构化、结构化数据都存储在Wide column模型系统中；Timeline模型是表格存储自研的模型，主要用于消息数据，适用于IM、Feed和物联网设备消息下推等消息系统中消息的存储和同步，目前已被广泛使用；Timestream模型可应用于时序数据、时空数据等核心数据场景。

- 无缝扩展

表格存储通过数据分片和负载均衡技术，实现了存储无缝扩展。随着表数据量的不断增大，表格存储会进行数据分区的调整从而为该表配置更多的存储。

- 查询能力强

除了支持主键查询，表格存储还支持多元索引、全局二级索引。全局二级索引相当于给主表提供了另外一种排序方式，即对查询条件预先设计了一种数据分布，可加快数据查询的效率。多元索引基于倒排索引和列式存储，支持多字段自由组合查询、模糊查询、地理位置查询、全文检索等，可解决大数据的复杂查询难题。

- 高可靠

表格存储将数据的多个备份存储在不同机架的不同机器上，并会在备份失效时进行快速恢复，提供99.9999999%（10个9）的可靠性。

- 数据强一致

表格存储保证数据写入强一致，并保证数据3副本均写入磁盘，且所有数据保持一致。写操作一旦返回成功，应用程序就能立即读到最新的数据。

表格存储使用场景

表格存储适用于元数据、大数据、消息数据、时空数据、时序数据等场景下的系统搭建，具体方案参见[适用场景](#)，代码实现参见[表格存储Github仓库](#)。

表格存储入门

您可以在控制台快速体验表格存储，更多参见[快速入门手册](#)。

相关概念

- **数据生命周期**

数据生命周期（Time To Live，简称TTL）是数据表的一个属性，即数据的存活时间，单位为秒。表格存储会在后台对超过存活时间的数据进行清理，以减少您的数据存储空间，降低存储成本。

- **地域**

地域（Region）是指物理的数据中心，表格存储服务会部署在多个阿里云地域中，您可以根据自身的业务需求选择不同地域的表格存储服务。详情参见[表格存储已经开通的Region](#)。

- **读/写吞吐量**

读写吞吐量的单位为读服务能力单元和写服务能力单元，简称CU（Capacity Unit），是数据读写操作的最小计费单位。

使用表格存储

- **表格存储控制台**

阿里云提供了Web服务页面方便您管理表格存储。您可以登录[表格存储控制台](#)操作表格存储实例。

- **表格存储SDK**

阿里云提供了SDK接口方便您灵活使用表格存储。详情参见[表格存储 SDK 参考](#)。

- **表格存储命令行工具**

阿里云提供了命令行工具Tablestore CLI方便您灵活地管理表格存储。详情参见[#unique_10](#)。

计算分析

您可以对表格存储中的数据进行计算分析：

- 如何在同一个云账号下实现表格存储和MaxCompute之间的无缝连接，参见[使用MaxCompute访问表格存储](#)。
- 如何可视化表格存储的数据，参见[#unique_12](#)。

迁移同步

您可以将异构数据平滑迁移同步到表格存储，还可以将数据同步到OSS等服务中：

- 如何将MySQL中的数据无缝迁移到表格存储，参见[同步MySQL数据](#)。
- 如何将MaxCompute中的全量数据同步到表格存储，参见[#unique_14](#)。
- 如何将HBase中的全量数据同步到表格存储，参见[#unique_15](#)。
- 如何将表格存储中的数据同步到OSS，参见[将表格存储数据同步迁移到OSS](#)。

表格存储定价

表格存储计量项包括数据存储量、预留读/写吞吐量、按量读/写吞吐量和外网下行流量。更多信息，参见[计量项和计费说明](#)。

表格存储及相关资源的价格信息，参见[云产品定价页](#)。

表格存储视频介绍

2 适用场景

本文为您介绍表格存储的适用场景以及可借鉴的开发指导、经典案例。通过本文提供的场景和应用案例，您可以更好地了解表格存储使用方式、设计思路、架构实现等，为您的方案选型提供参考依据。

本文中的经典场景都提供相应的实现代码：[表格存储Github仓库](#)

热门案例

- [海量结构化数据存储技术揭秘：Tablestore存储和索引引擎详解](#)
- [现代IM系统中的消息系统架构 - 架构篇、模型篇、实现篇](#)
- [结构化大数据分析平台设计](#)
- [Feed流系统设计-总纲](#)
- [数据同步-从MySQL到Tablestore](#)
- [数据中台之结构化大数据存储设计](#)
- [亿级消息系统的核心存储：Tablestore发布Timeline 2.0模型](#)
- [基于Tablestore管理海量快递轨迹数据架构实现](#)
- [Tablestore+Delta Lake \(快速开始\)](#)
- [基于大数据的舆情分析系统架构 - 架构篇](#)
- [Lambda plus：云上大数据解决方案](#)

控制台可视化样例

- [电商订单](#)
- [共享汽车](#)
- [快递轨迹追踪](#)
- [Wifi设备管理](#)
- [店铺搜索](#)
- [地理围栏](#)
- [智能元数据](#)

场景案例

- [元数据](#)

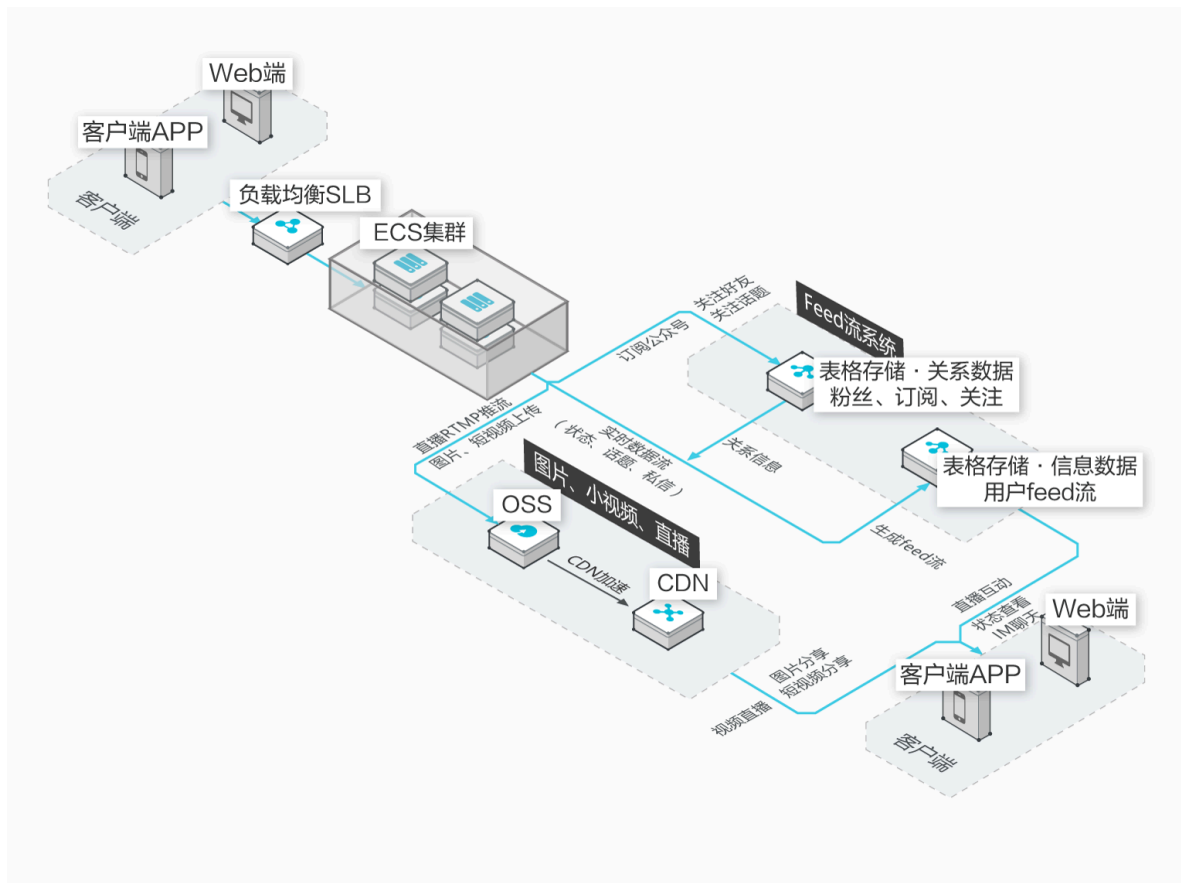
用户存储海量的文档、媒体文件等数据的同时，对文件元数据的存储和分析不可或缺。此外，电商的订单、银行流水、运营商话费账单也需要存储及分析大量的元数据。表格存储单表支持

PB级存储、千万QPS，以及多种索引方式（全局二级索引、全文索引、倒排索引以及时空索引），满足不同场景在线的不同的查询需求，可以帮助您轻松实现高效的元数据管理。

- 基于Tablestore打造亿量级订单管理解决方案
- 基于Tablestore多元索引打造亿量级店铺搜索系统
- 海量智能元数据管理系统实现解析
- 基于Tablestore的海量保险单查询平台

· 消息数据

表格存储自研的Timeline模型主要用于消息数据，能够抽象出支撑海量Topic的轻量级消息队列，可以存储大量社交信息，包括IM聊天，以及评论、跟帖和点赞等Feed流信息，接口简单易用。目前表格存储Timeline模型已被应用在众多IM系统中，例如支撑钉钉海量消息同步等。此外，表格存储采用按量付费，能够以较低的成本满足访问波动明显、高并发、低延时的需求。



- TableStore Timeline：轻松构建千万级IM和Feed流系统
- 如何打造千万级Feed流系统
- 高并发IM系统架构优化实践

- 轨迹溯源

表格存储提供了面向轨迹类场景的Timestream模型，可提供PB级存储、千万TPS以及毫秒级延迟的服务能力、以及多种索引方式（全局二级索引、全文索引、倒排索引以及时空索引）。使用表格存储Timestream模型，您可以轻松管理、分析跑步、骑行、健走、外卖等轨迹数据。

- [基于Tablestore实现海量运动轨迹数据存储](#)
- [药品监管系统架构揭秘：海量溯源数据存储与查询](#)
- [基于Tablestore实现海量摩托车的轨迹管理](#)
- [超级快递——如何用系统来保证快递准时送达](#)

- 科学大数据

多维网格数据是一种科学大数据，在地球科学领域（气象、海洋、地质、地形等）应用非常广泛，且数据规模也越来越大。相关的科学工作者有快速浏览数据的需求以及在线查询的需求，查询种类丰富、延迟要求高。表格存储是一款阿里云自研的分布式NoSQL服务，可以提供超大规模的存储容量，支撑超大规模的并发访问和低延迟的性能，可以轻松解决科学大数据的海量存储规模和查询性能问题。

[基于Tablestore的海量气象格点数据解决方案实战](#)

- 互联网大数据

热点新闻及娱乐八卦可以在短短数分钟内，有数万计转发，数百万的阅读，如何能够实时的把握民情并作出对应的处理对很多企业来说都是至关重要的。此外，商品在各类电商平台的订单量，用户的购买评论也都对后续的消费者产生很大的影响。商家的产品设计者需要汇总统计和分析各类平台的数据做为依据，决定后续的产品发展，公司的公关和市场部门也需要根据舆情作出相应的及时处理。表格存储单表提供PB级存储、千万QPS，以及多种索引方式，可以帮助您轻松实现百亿级互联网舆情存储及分析。

- [百亿级全网舆情分析系统存储设计](#)
- [海量监控日志基于EMR Spark Streaming SQL进行实时聚合](#)

- 物联网

表格存储单表提供PB级数据存储规模，无需分库分表，同时支持千万QPS，可以轻松满足IoT设备、监控系统等时序数据的存储需求，大数据分析SQL直读以及高效的增量流式读接口让数据轻松完成离线分析与实时流计算。

- [基于Tablestore的共享汽车管理平台](#)
- [基于Tablestore的Wifi设备监管系统架构实现](#)
- [基于TableStore的物联网元数据管理](#)
- [如何高效存储海量GPS数据](#)

开发指南

产品对比

- [深入对比 HBase 与阿里云的表格存储服务](#)
- [从SQL到NoSQL—如何使用表格存储](#)

数据模型

- [表格存储数据模型和查询操作](#)
- [TableStore数据模型 - WideColumn和Timeline](#)
- [Tablestore Timestream：为海量时序数据存储设计的全新数据模型](#)

使用指导

- [海量结构化数据存储技术揭秘：Tablestore表设计最佳实践](#)
- [详解TableStore模糊查询——以订单场景为例](#)

数据迁移

- [通过Datahub将本地的CSV文件导入Tablestore](#)
- [基于Tablestore Tunnel的数据复制实战](#)

专家服务

表格存储提供专业的免费的技术咨询服务，欢迎加入我们的钉钉讨论群。

表格存储技术交流群

2665人



3 常见问题

本文列举了阿里云表格存储的常见问题，帮助您快速了解表格存储。

一般性常见问题

· 什么是表格存储？

表格存储是阿里云自研的NoSQL多模型数据库，提供海量结构化数据存储以及快速的查询和分析服务。表格存储的分布式存储和强大的索引引擎能够支持PB级存储、千万TPS以及毫秒级延迟的服务能力。具有全托管、多模型、无限存储等特点，具体参见[什么是表格存储](#)。

· 为什么选择表格存储？

表格存储具有以下优势：

- 全托管

表格存储是一种全托管的数据库。使用表格存储您只需专注于业务研发，无需担心软硬件预置、配置、故障、集群扩展、安全等问题，在保证高服务可用性的同时，极大地减少了管理及运维成本。

- 模型丰富

表格存储支持多种数据库模型，包括Wide column、Timeline、Timestream、Grid。其中，Wide column模型是一款经典模型，目前绝大部分半结构化、结构化数据都存储在Wide column模型系统中；Timeline模型是表格存储自研的模型，主要用于消息数据，适用于IM、Feed和物联网设备消息下推等消息系统中消息的存储和同步，目前已被广泛使用；Timestream模型可应用于时序数据、时空数据等核心数据场景。

- 无缝扩展

表格存储通过数据分片和负载均衡技术，实现了存储无缝扩展。随着表数据量的不断增大，表格存储会进行数据分区的调整从而为该表配置更多的存储。

- 查询能力强

除了支持主键查询，表格存储还支持多元索引、全局二级索引。全局二级索引相当于给主表提供了另外一种排序方式，即对查询条件预先设计了一种数据分布，可加快数据查询的效

率。多元索引基于倒排索引和列式存储，支持多字段自由组合查询、模糊查询、地理位置查询、全文检索等，可解决大数据的复杂查询难题。

- 高可靠

表格存储将数据的多个备份存储在不同机架的不同机器上，并会在备份失效时进行快速恢复，提供99.99999999%（10个9）的可靠性。

- 数据强一致

表格存储保证数据写入强一致，并保证数据3副本均写入磁盘，且所有数据保持一致。写操作一旦返回成功，应用程序就能立即读到最新的数据。

- 表格存储是否有存储和扩展限制？

表格存储没有存储和扩展限制。表格存储通过数据分片和负载均衡技术，实现了存储无缝扩展。随着表数据量的不断增大，表格存储会自动进行数据分区的调整从而为该表配置更多的存储。

- 如何设计表格存储的表？

合理的表设计能够发挥出数据库的最优性能，关于如何设计表格存储的表，参见[海量结构化数据存储技术揭秘：Tablestore表设计最佳实践](#)。

运维管理

表格存储如何运维，运维成本高么？

表格存储是一种全托管的数据库。表格存储可以自动根据您的业务量的增长进行扩容，还可自动进行故障检测和迁移，对您的应用程序屏蔽机器和网络的硬件故障，为您完成一切运维任务。使用表格存储您只需专注于业务研发，无需担心软硬件预置、配置、故障、集群扩展、安全等问题，在保证高服务可用性的同时，为您省去运维成本。

使用场景

表格存储有哪些主要使用场景？

表格存储适用于以下场景：

- 元数据

用户存储海量的文档、媒体文件等数据的同时，对文件元数据的存储和分析不可或缺。此外，电商的订单、银行流水、运营商话费账单也需要存储及分析大量的元数据。表格存储单表支持PB级存储、千万QPS，以及多种索引方式（全局二级索引、全文索引、倒排索引以及时空索引），满足不同场景在线的不同的查询需求，可以帮助您轻松实现高效的元数据管理。

- 消息数据

表格存储自研的Timeline模型主要用于消息数据，能够抽象出支撑海量Topic的轻量级消息队列，可以存储大量社交信息，包括IM聊天，以及评论、跟帖和点赞等Feed流信息，接口简单易

用。目前表格存储Timeline模型已被应用在众多IM系统中，例如支撑钉钉海量消息同步等。此外，表格存储采用按量付费，能够以较低的成本满足访问波动明显、高并发、低延时的需求。

- 轨迹溯源

表格存储提供了面向轨迹类场景的Timestream模型，可提供PB级存储、千万TPS以及毫秒级延迟的服务能力、以及多种索引方式（全局二级索引、全文索引、倒排索引以及时空索引）。使用表格存储Timestream模型，您可以轻松管理、分析跑步、骑行、健走、外卖等轨迹数据。

- 科学大数据

多维网格数据是一种科学大数据，在地球科学领域（气象、海洋、地质、地形等）应用非常广泛，且数据规模也越来越大。相关的科学工作者有快速浏览数据的需求以及在线查询的需求，查询种类丰富、延迟要求高。表格存储是一款阿里云自研的分布式NoSQL服务，可以提供超大规模的存储容量，支撑超大规模的并发访问和低延迟的性能，可以轻松解决科学大数据的海量存储规模和查询性能问题。

- 互联网大数据

热点新闻及娱乐八卦可以在短短数分钟内，有数万计转发，数百万的阅读，如何能够实时的把握民情并作出对应的处理对很多企业来说都是至关重要的。此外，商品在各类电商平台的订单量，用户的购买评论也都对后续的消费者产生很大的影响。商家的产品设计者需要汇总统计和分析各类平台的数据做为依据，决定后续的产品发展，公司的公关和市场部门也需要根据舆情作出相应的及时处理。表格存储单表提供PB级存储、千万QPS，以及多种索引方式，可以帮助您轻松实现百亿级互联网舆情存储及分析。

- 物联网

表格存储单表提供PB级数据存储规模，无需分库分表，同时支持千万QPS，可以轻松满足IoT设备、监控系统等时序数据的存储需求，大数据分析SQL直读以及高效的增量流式读接口让数据轻松完成离线分析与实时流计算。

表格存储为您提供了各个场景下的经典使用案例供您参考，具体参见[适用场景](#)。

计费

表格存储的费用是多少？

使用表格存储，您可以按实际使用量付费，先使用，后付费，以较低的成本满足访问波动明显大并发低延时的需要。您也可以预先购买资源包，之后使用资源时，扣除相应的额度。您也可以预先购买资源包，之后使用资源时，扣除相应的额度。一般情况下，资源包更加优惠。详细的价格信息请参见[阿里云产品定价](#)。

4 专家服务

如果想对表格存储有进一步的了解，除了官网文档外，您还有更多途经来对表格存储有一个更深入的了解。

- 阅读[适用场景](#)。
- 关注我们的[开发者社区](#)。
- 加入我们的钉钉公开交流群（搜索钉钉群号『11789671』或群名称『表格存储公开交流群』）。

