

表格存储服务

产品介绍

文档版本

10

发布日期

2020-06-23



华为技术有限公司



版权所有 © 华为技术有限公司 2020。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

目录

1 什么是 CloudTable.....	1
2 应用场景.....	3
3 集群模式功能简介.....	9
3.1 集群管理.....	9
3.2 HBase.....	10
3.3 OpenTSDB.....	10
3.4 GeoMesa.....	11
4 与其他云服务的关系.....	12
5 CloudTable 权限管理.....	14
6 计费说明.....	16
7 使用限制.....	18
8 配额说明.....	19

1 什么是 CloudTable

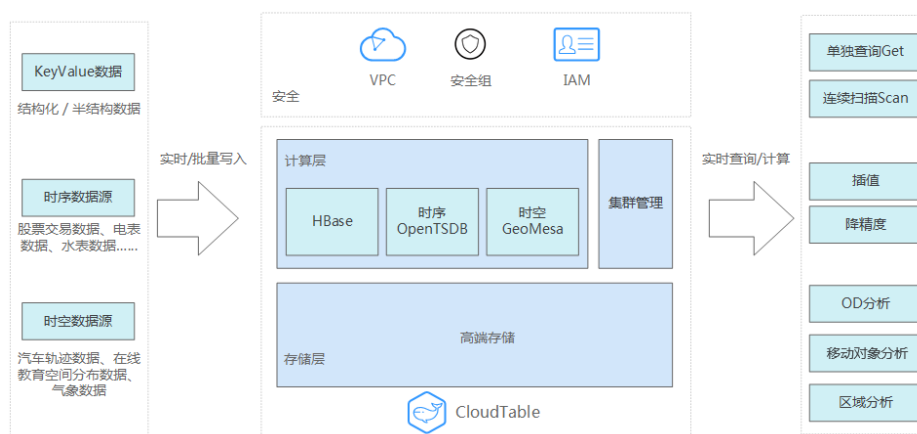
表格存储服务（CloudTable Service，简称CloudTable）是基于Apache HBase提供的全托管NoSQL服务，提供毫秒级随机读写能力，适用于海量(半)结构化，时空，时序数据存储，可被广泛应用于物联网、车联网、金融、智慧城市、气象等行业。

CloudTable集群模式为用户提供专属集群，即开即用，适合业务吞吐量大，时延要求低的用户。

产品架构

CloudTable服务的产品架构如下图所示：

图 1-1 产品架构



- **生态开放**：CloudTable兼容HBase/OpenTSDB/GeoMesa原生接口。
- **HBase**：支持Key-Value数据模型。架构高可用，HMaster为两个节点，主备模式，HA实时检测。计算单元故障，region可以秒级转移，保证业务的高可用。存储和计算分离保证数据的高可靠，存储采用多备份机制，存储可靠性不低于99.999999%。
- **时序数据库**：集成OpenTSDB，提供时序数据的高效读、写、查询和计算能力，读写性能提升30%-60%。支持插值、降精度、聚合强大分析能力，10:1高压缩比，成本更低。

- **时空大数据**：集成地理大数据处理套件GeoMesa，帮助物联网存储和分析海量时空(spatio-temporal)数据，提供区域分布统计、区域查询、密度分析、聚合、OD分析等功能，超越传统空间数据库功能。
- **存储层**：高可用，坏盘不影响数据读写，不影响可靠性。计算与存储分离，存储按使用量计费，成本更低。
- **集群管理**：提供可视化管理平台，支持集群管理、修改配置参数、按CU进行在线扩容等功能。支持可视化监控运维，提供常用的监控信息（如CPU使用率、IOPS、流入流出的流量等）和自定义告警规则，简化系统运维。
- **安全**：CloudTable使用IAM服务实现身份认证功能。CloudTable的底层计算单元运行在一个VPC网络中，VPC即虚拟私有云，是通过逻辑方式进行网络隔离，提供安全、隔离的网络环境。安全组用来实现安全组内和安全组间的访问控制，用户可以自定义安全组规则来实现对CloudTable的访问控制。

2 应用场景

物联网 IoT 设备监控

应用场景：

梯联网、燃气、水务、电力、化工、互联网等IoT设备通过物联网套件服务接入上云，设备数据和分析结果实时高效写入到CloudTable的时序数据库OpenTSDB中，通过OpenTSDB接口将时序结果输出到用户的前端监控系统进行展现，实现物联网设备实时监控分析系统。

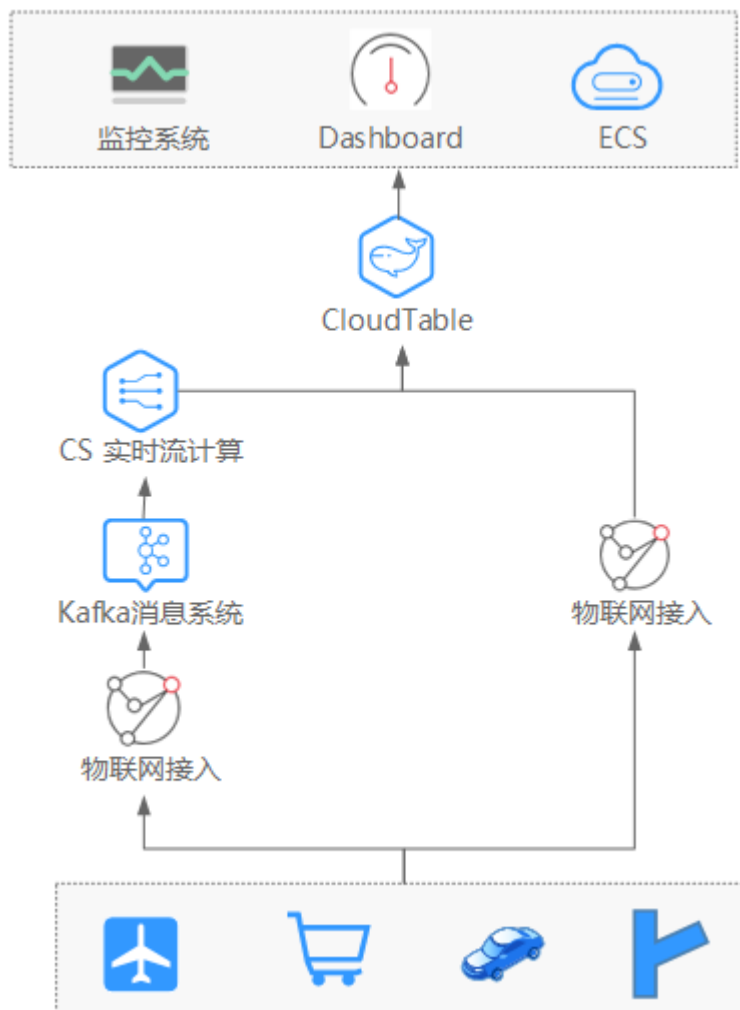
优势：

- 易接入
CloudTable(OpenTSDB)开放协议，轻松对接消息系统、实时流计算系统，降低开发难度。
- 高性能读写
千万级的时序数据写入吞吐量，百万数据点3秒查询时延。相比开源OpenTSDB，针对读性能提高30%~60%，写并发提升60%。
- 支持聚合能力
提供插值、降精度、丰富的聚合函数能力。
- 低成本
存储按使用量计费，弹性扩容，从容应对业务的不确定性。10:1的高压缩比，成本更低。

建议搭配使用：

实时流计算服务 CS + 数据接入服务 DIS + 数据湖探索 DLI + IoT平台 + 对象存储服务 OBS

图 2-1 IoT 设备监控



消息日志类数据存储和查询

应用场景：

消息数据、报表数据、推荐类数据、风控类数据、日志数据、订单数据等结构化、半结构化的Key-Value数据均可以存储和查询。

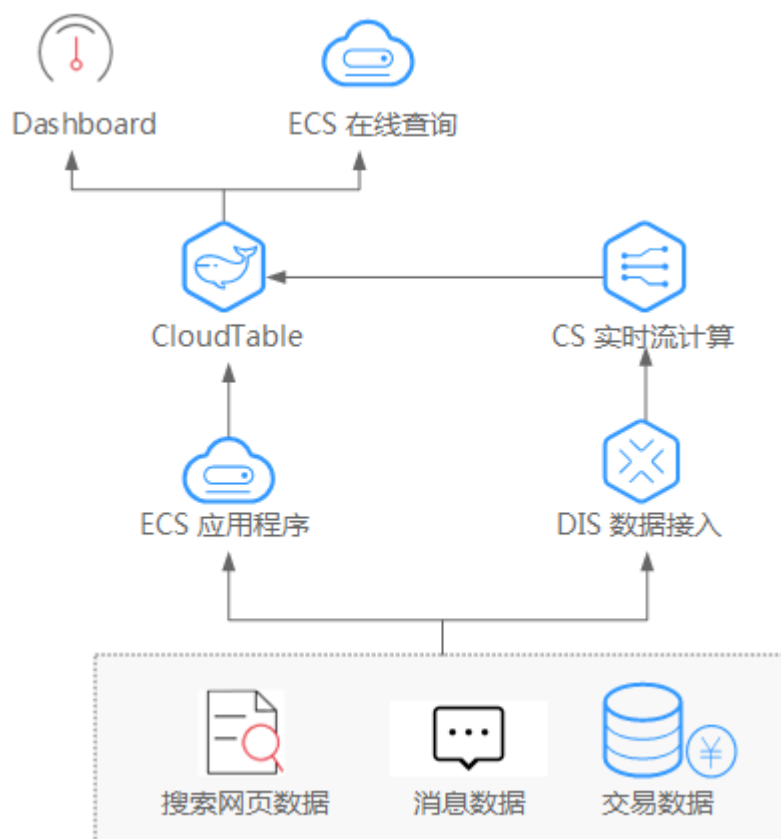
优势：

- 海量存储
支持离线、在线海量Key-Value数据存储，存储容量可扩展。
- 高性能读写
亿级写入吞吐量、ms级查询，用于在线应用和报表展现。
- 生态丰富
基于Hadoop生态组件丰富，与华为云产品有高度的整合能力。

建议搭配使用：

数据接入服务 DIS + 实时流计算服务 CS

图 2-2 消息日志类数据存储和查询



车联网：位置大数据应用

应用场景：

在车联网中，通常会包含几类数据：车辆、驾驶员等基本信息，车况、电池、电机等监控数据，车辆路线。CloudTable引入了地理大数据处理套件GeoMesa，可以帮助物联网存储和分析海量时空(spatio-temporal)数据，提供区域分布统计、区域查询、密度分析、聚合、OD分析等功能。

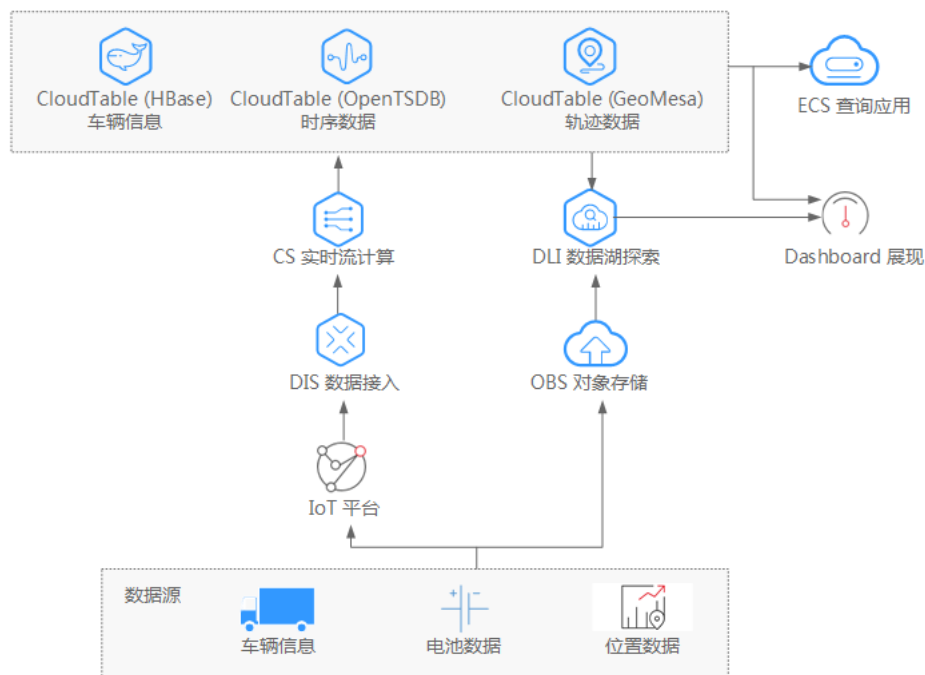
优势：

- 多模数据库能力
针对不同的数据提供不同的索引能力，为其提供最优的性能和查询分析能力。
- 丰富的查询分析函数
GeoMesa提供区域分布统计、区域查询、密度分析、聚合、OD分析等功能。
- 无缝对接分析能力
时空数据库可以和DLI服务无缝对接，提供对于时间+空间数据的良好分析能力，例如热力图等。

建议搭配使用：

实时流计算服务 CS + 数据接入服务 DIS + 数据湖探索 DLI + IoT平台 + 对象存储服务 OBS

图 2-3 位置大数据



画像数据存储和查询

应用场景：

画像通常用一些标签来刻画自然人/物的特征，而每一个自然人/物所拥有的标签集合是不确定的，数据更新非常频繁，这类数据被广泛应用于市场决策、推荐以及广告系统中。

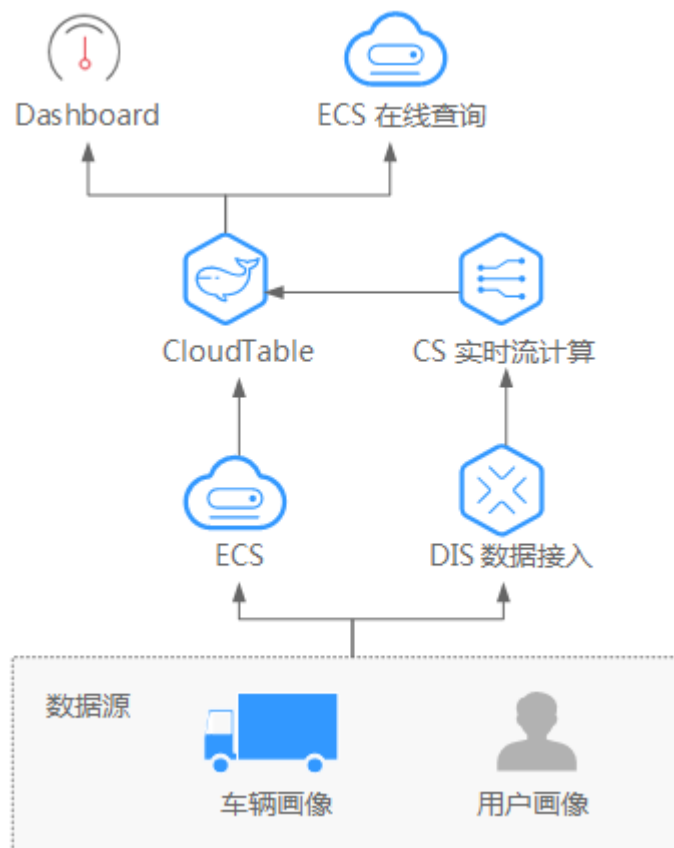
优势

- 稀疏矩阵
HBase的稀疏矩阵模型，天然适合非结构化数据的存储，数据表无需预先定义 schema，行与行之间不需要严格的列定义。
- 支持任意更新
支持行的任意更新，无性能损耗。而且利用HBase自身的多版本机制，支持保存数据的多个历史版本。

建议搭配使用：

数据接入服务 DIS + 实时流计算服务 CS

图 2-4 画像数据存储和查询



Serverless Web/移动应用后端

应用场景：

使用CloudTable和函数服务FunctionGraph结合，用户可以快速构建高可用、自动伸缩的Web/移动应用后端。

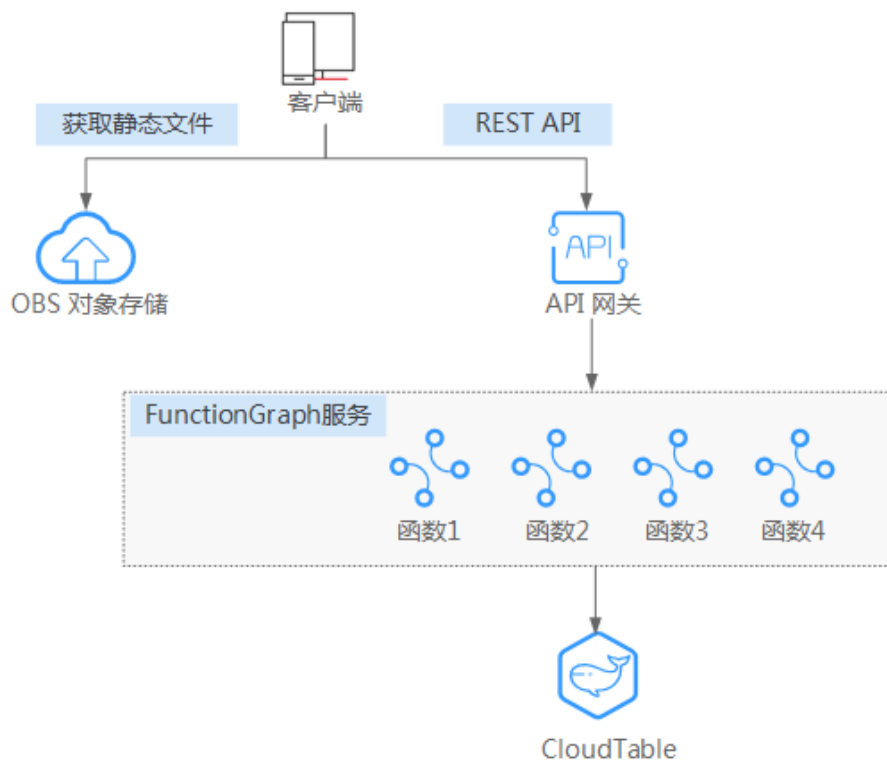
优势：

- 高可用
利用CloudTable，OBS的高可用性实现网站数据的高可靠性，利用API Gateway和FunctionGraph的高可用性实现网站逻辑的高可用。
- 超高性能
单盘最高20000 IOPS，350MB/s吞吐量。
- 灵活扩展
业务爆发时可以自动调度资源运行更多函数实例以满足处理需求。
- 低成本
只针对函数处理文件数据的时间进行计费，存储按使用量计费，弹性扩容，对于非峰值处理，无需购买冗余的资源。

建议搭配使用：

对象存储服务 OBS + 函数服务 FunctionGraph

图 2-5 Web/移动应用后端



3 集群模式功能简介

3.1 集群管理

CloudTable服务是华为云提供的一项分布式、可扩展的Key/Value数据存储服务。CloudTable提供Web官网服务界面，CloudTable集群模式的具体功能如下：

- 创建集群：在CloudTable服务界面完成集群的创建。支持按照用户创建集群时选择的计算单元个数、实际使用存储容量计费。用户可以自主地选择服务提供的高级特性，独立安装、独立收费。扣费时余额不足先提醒用户续费，在保留期冻结集群资源，续费后再解冻。通过计算存储分离架构和计算资源动态调整，尽可能的为用户降低成本。
- 扩容集群：支持集群扩容计算单元。
 - 扩容计算单元：用户根据实际需要或业务情况，动态的增加计算单元的个数，保证读写性能。集群自适应的实现负载均衡，保证业务不中断，平滑扩容。扩容计算单元将会产生额外的费用。
- 高级特性：当前CloudTable集群模式支持OpenTSDB和GeoMesa高级特性。用户可选择性启用高级特性。启用OpenTSDB高级特性将会产生相应的费用。
- 管理集群：对创建的集群进行管理。
 - 指标监控：集群运行中，收集各项监控数据，上报云监控（Cloud Eye），向用户以图形化的方式呈现集群运行状况。当指标出现异常时，通过消息通知用户和管理员，及时人工介入。
 - 删除集群：当用户不再需要集群时，可选择删除集群。此操作为高危操作，删除集群可能导致数据丢失，因此在执行删除操作之前，请确认不存在正在运行的业务，所有数据都已经保存。
 - 重启集群：当修改完集群HBase参数后，或者因长时间不重启导致系统运行缓慢等情况下，用户需要重启集群。重启操作将有可能导致正在运行中的业务数据丢失，如果需要执行重启操作，请确定不存在正在运行的业务，所有数据都已经保存。
 - 查询告警：集群运行异常或系统故障时，CloudTable服务会收集故障信息并上报网管系统，维护人员可根据用户提供的告警信息定位问题原因。
 - 日志查询：记录用户对集群操作信息，便于集群运行异常时定位分析问题原因。

3.2 HBase

HBase是一个稳定可靠，性能卓越、可伸缩、面向列的分布式云存储系统，适用于海量数据存储以及分布式计算的场景，用户可以利用HBase搭建起TB至PB级数据规模的存储系统，对数据轻松进行过滤分析，毫秒级得到响应，快速发现数据价值。

HBase适用场景有：

- 海量数据存储
适用于TB ~ PB级以上的数据存储，提供动态伸缩能力，方便用户在性能或容量需要改变时，改变集群资源，轻松构建企业海量数据存储系统。
- 实时查询
HBase的列式KeyValue存储机制，适用于企业用户明细数据即时查询，基于主键的低时延点查，响应时延一般为秒级或毫秒级，方便用户对数据的实时分析。

HBase的架构和详细原理介绍，请参见：<https://hbase.apache.org/book.html>

3.3 OpenTSDB

CloudTable集群模式还提供了基于OpenTSDB时序数据库能力。

OpenTSDB是基于HBase的分布式的，可伸缩的时间序列数据库。它存储的是时间序列数据，时间序列数据是指在不同时间点上收集到的数据，这类数据反映了一个对象随时间的变化状态或程度。

产品优势

- 低成本
时间戳采用delta编码进行压缩，数据值采用XOR进行压缩。
存储与计算解耦，为IoT场景海量数据、动态热点的数据特征量身打造，方便按照并发度和存储量按需独立扩容。
- 企业级
分布式架构，横向水平扩展。
高压缩率算法，节约成本的同时，提升查询速度。
- 兼容性
兼容OpenTSDB社区2.3.0版本。
兼容OpenTSDB原生接口，业务迁移应用“0”改动。
- 时序数据计算
插值，缺失的数据点，支持线性插值数据补全。
降精度，支持预降精度和实时降精度计算，满足高效查询需求。
空间聚合，支持按照不同的Tag进行空间聚合和分组计算。
丰富的聚合函数，提供AVG、SUM、MAX、MIN等聚合函数。

OpenTSDB 适用场景

OpenTSDB支持海量时间序列数据的存储、索引和查询，可用于系统监控和测量、物联网数据、金融数据和科学实验结果数据的监控和分析。

- 对于运维工程师而言，OpenTSDB可以获取基础设施和服务的实时状态信息，展示集群的各种软硬件错误，性能变化以及性能瓶颈。
- 对于管理者而言，OpenTSDB可以存储和分析系统监控数据，衡量系统的SLA，理解复杂系统间的相互作用，展示资源消耗情况。集群的整体作业情况，可以用以辅助预算和集群资源协调。
- 对于开发者而言，OpenTSDB可以通过存储和分析系统监控数据，发现集群的主要性能瓶颈。

何时开启 OpenTSDB

当需要存储和查询时间序列数据时，可开启OpenTSDB。

3.4 GeoMesa

CloudTable集群模式还提供了基于GeoMesa的时空大数据查询、分析能力。

GeoMesa是一个基于HBase的分布式的，可伸缩的开源时空数据库。GeoMesa利用高度并行化的索引策略，为用户提供基于HBase的空间与时间数据查询与处理能力。

GeoMesa 适用场景

GeoMesa支持毫秒级时空查询，可用于存储，索引和查询各类用户GIS数据。

GeoMesa适用场景有：

- 海量地理位置信息数据存储
基于HBase的存储能力，GeoMesa适用于存储TB~PB级地理位置信息数据。
- 实时查询
基于特殊设计的索引结构，GeoMesa提供了快速查询时空数据的能力。

何时开启 GeoMesa

当需要存储和查询地理空间与时间数据时，可开启GeoMesa。

4 与其他云服务的关系

统一身份认证 IAM

CloudTable使用统一身份认证服务（Identity and Access Management，简称IAM）实现认证功能。

了解更多IAM的相关信息，请参见[统一身份认证服务](#)。

弹性云服务器 ECS

CloudTable使用弹性云服务器（Elastic Cloud Server，简称ECS）作为服务实例的节点，每个弹性云服务器是服务实例中的一个节点。

了解更多ECS的相关信息，请参见[弹性云服务器](#)。

虚拟私有云 VPC

CloudTable使用虚拟私有云（Virtual Private Cloud，简称VPC）为集群提供网络拓扑，实现多个不同集群互相隔离并控制访问。

了解更多VPC的相关信息，请参见[虚拟私有云](#)。

对象存储服务 OBS

CloudTable使用对象存储服务（Object Storage Service，简称OBS）存储集群数据的备份和快照，实现安全、高可靠和低成本存储需求。

了解更多OBS的相关信息，请参见[对象存储服务](#)。

云数据迁移 CDM

CloudTable使用云数据迁移（Cloud Data Migration，简称CDM），可以将云上云下或第三方云上的多种数据源的数据迁移到CloudTable集群的HBase表中。

了解更多CDM的相关信息，请参见[云数据迁移](#)。

云搜索服务 CSS

CloudTable使用云搜索服务（Cloud Search Service，简称CSS）中的Elasticsearch搜索引擎来补充全文检索能力。用户在HBase的建表语句中，通过METADATA来定义索

引字段的Elasticsearch schema并指定云搜索服务集群的连接地址，在建表时会自动连接云搜索服务集群并在Elasticsearch中创建索引。用户的源数据存储在CloudTable的HBase中，索引数据存放在云搜索服务的Elasticsearch中。

了解更多CSS的相关信息，请参见[云搜索服务](#)。

云审计服务 CTS

CloudTable使用云审计服务（Cloud Trace Service，简称CTS）为用户提供CloudTable资源操作请求及请求结果的操作记录，供用户查询、审计和回溯使用。

了解更多CTS的相关信息，请参见[云审计服务](#)。

云监控服务 CES

CloudTable使用云监控服务CES监控集群中的多项性能指标，从而集中高效地呈现状态信息。云监控支持发送自定义告警，用户可第一时间获取异常通知。

了解更多CES的相关信息，请参见[云监控服务](#)。

5 CloudTable 权限管理

如果您需要对华为云上购买的CloudTable资源，为企业中的员工设置不同的访问权限，以达到不同员工之间的权限隔离，您可以使用统一身份认证服务（Identity and Access Management，简称IAM）进行精细的权限管理。该服务提供用户身份认证、权限分配、访问控制等功能，可以帮助您安全的控制华为云资源的访问。

通过IAM，您可以在华为云账号中给员工创建IAM用户，并授权控制他们对华为云资源的访问范围。例如您的员工中有负责软件开发的人员，您希望他们拥有CloudTable的使用权限，但是不希望他们拥有删除集群等高危操作的权限，那么您可以使用IAM为开发人员创建用户，通过授予仅能使用CloudTable，但是不允许删除集群的权限，控制他们对CloudTable资源的使用范围。

如果华为云账号已经能满足您的要求，不需要创建独立的IAM用户进行权限管理，您可以跳过本章节，不影响您使用CloudTable服务的其它功能。

IAM是华为云提供权限管理的基础服务，无需付费即可使用，您只需要为您账号中的资源进行付费。关于IAM的详细介绍，请参见[IAM产品介绍](#)。

CloudTable 权限

默认情况下，管理员创建的IAM用户没有任何权限，需要将其加入用户组，并给用户组授予策略或角色，才能使得用户组中的用户获得对应的权限，这一过程称为授权。授权后，用户就可以基于被授予的权限对云服务进行操作。

CloudTable部署时通过物理区域划分，为项目级服务。授权时，“作用范围”需要选择“区域级项目”，然后在指定区域（如华北-北京1）对应的项目（cn-north-1）中设置相关权限，并且该权限仅对此项目生效；如果在“所有项目”中设置权限，则该权限在所有区域项目中都生效。访问CloudTable时，需要先切换至授权区域。

如表5-1所示，包括了CloudTable的所有系统角色。由于华为云各服务之间存在业务交互关系，表格存储服务的角色依赖其他服务的角色实现功能。因此给用户授予表格存储服务的角色时，需要同时授予依赖的角色，表格存储服务的权限才能生效。

表 5-1 CloudTable 系统角色

系统角色	描述	类别	依赖关系
cloudtable Administrator	表格存储服务的管理员权限。	系统角色	依赖Tenant Guest和Server Administrator角色，在同项目中勾选依赖的角色。

表5-2列出了CloudTable常用操作与系统权限的授权关系，您可以参照该表选择合适的系统权限。

表 5-2 CloudTable 操作与系统策略关系

模式	操作	cloudtable Administrator
集群模式	创建集群	✓
	重启集群	✓
	扩容集群	✓
	删除集群	✓
	参数配置	✓
	开启OpenTSDB	✓
	查看CloudTable服务集群列表、 集群详细信息	✓
	查看监控信息	✓
	查看审计日志	✓

相关链接

- [IAM产品介绍](#)
- [创建用户并授权使用CloudTable](#)

6 计费说明

表格存储服务（CloudTable）支持按需计费，同时您也可以选择更经济的包年包月折扣套餐计费模式。为了便于您便捷的下单购买，我们在控制台购买界面中已经为您计算好了整个CloudTable集群的价格，您可一键完成整个配置在购买。

计费项

CloudTable服务根据您选择的计算单元规格进行计费。CloudTable服务的价格详情，请参见[产品价格详情](#)。您可以通过CloudTable服务提供的价格计算器，选择您需要的计算单元规格，来快速计算出购买CloudTable集群的参考价格。

计费模式

CloudTable服务提供两种计费模式供您选择：按需计费和包年包月折扣套餐计费。

- **包年包月折扣套餐**：根据计算单元的购买时长，一次性支付费用。最短时长为1个月，最长时长为1年。
- **按需付费**：计算单元按实际使用时长计费，计费周期为一小时。

变更配置

在购买CloudTable集群时，您可以根据业务需要选择合适的集群规格或计算单元个数。当集群创建成功后，CloudTable提供如下几种变更配置的方式。

- **扩容**
当集群创建成功后，随着业务量的增长，您可以对集群进行扩容，从而增加计算单元的数量，详情请参见[扩容集群](#)。
- **开启OpenTSDB**
如果您在创建CloudTable集群时没有勾选OpenTSDB，当需要使用OpenTSDB时，可以在集群创建完成后开启OpenTSDB，详情请参见[开启OpenTSDB](#)。开启OpenTSDB至少需要增加2个计算单元，您可以根据实际需求增加OpenTSDB所需的计算单元个数。OpenTSDB开启后，暂不支持关闭。

若CloudTable服务提供的变更配置方式无法满足您的需求，您也可以重新创建CloudTable集群。

续费

目前CloudTable提供按需和包年包月购买方式，按需是每小时扣费，如果余额不足导致欠费，而包年包月是集群时长用完前需续费。如果您未能续费，华为云不会立即停止您的业务，订单转入保留期，此时集群将终止服务，数据仍然保留。

保留期的时长由客户等级而定，具体请参见[保留期](#)。

如需续费，请进入[续费管理](#)页面进行续费操作。

欠费

- 包年包月集群，没有欠费的概念。
- 按需购买的集群是按每小时扣费，当余额不足，无法对上一个小时的费用进行扣费，就会导致集群欠费，集群欠费后有保留期。您续费后解冻集群，可继续正常使用，请注意在保留期进行的续费，是以原到期时间作为生效时间，您应当支付从进入保留期开始到续费时的服务费用。

服务到期

- 按需购买的集群，没有到期。
- 包年包月集群到期后进入保留期，此时无法在CloudTable管理控制台进行该集群的操作，相关接口也无法调用，自动化监控或告警等运维也会停止。如果在保留期结束时您没有续费，集群将终止服务，系统中的数据也将被永久删除。

7 使用限制

使用CloudTable前，您需要认真阅读并了解以下使用限制。

- CloudTable集群必须创建在VPC子网内。
- 建议使用支持的浏览器登录表格存储服务。
 - Google Chrome: 36.0及更高版本
 - Mozilla FireFox: 35.0及更高版本
 - Internet Explorer: 9.0及更高版本

当使用Internet Explorer 9.0时可能无法登录CloudTable管理控制台，原因是某些Windows系统例如Win7旗舰版，默认禁止Administrator用户，Internet Explorer在安装时自动选择其他用户如System用户安装，从而导致Internet Explorer无法打开登录页面。请使用管理员身份重新安装Internet Explorer 9.0或更高版本（建议），或尝试使用管理员身份运行Internet Explorer 9.0。

 - Microsoft edge: 13.0及更高版本
- 创建CloudTable集群时，会为您创建一个默认安全组，请勿随意删除或更改默认安全组。删除或更改默认安全组可能导致集群异常，影响CloudTable集群的使用。
- CloudTable集群使用的安全组请勿随意放开权限，避免被恶意访问。
- CloudTable不会保存您设置的登录节点的初始密码，请您设置并保管好密码。为避免被恶意攻击，建议设置复杂度高的密码。
- 集群节点仅用于运行CloudTable集群，其他客户端应用程序、用户业务程序建议申请独立弹性云服务器部署。

8 配额说明

表格存储服务用户默认的基础资源如下：

- 实例数
- CPU数量
- 内存容量(GB)
- 磁盘数
- 磁盘容量(GB)

上述资源的配额查看及修改，请参见[关于配额](#)。