

# 云数据仓库 PostgreSQL

## 产品简介

## 产品文档



腾讯云

**【 版权声明 】**

©2013–2022 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 】**

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。

## 文档目录

### 产品简介

产品概述

应用场景

节点规格

集群架构

# 产品简介

## 产品概述

最近更新时间：2022-03-15 14:58:16

云数据仓库 PostgreSQL ( Cloud Data Warehouse PostgreSQL ) ( 原 Snova 数据仓库 ) 为您提供简单、快速、经济高效的 PB 级云端数据仓库解决方案。云数据仓库兼容 Greenplum 开源数据仓库，是一种基于 MPP ( 大规模并行处理 ) 架构的数仓服务。借助于该产品，您可以使用丰富的 PostgreSQL 开源生态工具，实现对云数据仓库中海量数据的即席查询分析、ETL 处理及可视化探索；还可以借助云端数据无缝集成特性，轻松分析位于 COS、TencentDB、Elasticsearch Service 等数据引擎上的 PB 级数据。

## 产品优势

### 弹性伸缩

提供便利的弹性扩容能力，通过云控制台或云 API 简单操作便可以实现数百节点的伸缩或变配。可根据业务需求，选择计算单元、CPU、内存、存储空间的等比扩展，提高性能以适配业务的发展。

### 简单易用

通过控制台操作，即可实现集群管理、监控维护等工作，无需关注底层基础设施的繁重运维工作。完全支持 ANSI SQL 2008 标准，使用标准 SQL 即可构建企业级数据仓库。支持直接查询 COS 数据，无需提前数据预加载。

### 无缝集成

支持 COS 云存储扩展，实现存储空间的无限扩展。搭配多种工具及方案支持多源数据（例如传统关系型数据库、Ckafka、流计算等）高速导入，实现对云端多源数据的汇聚分析。

### 性能卓越

基于分布式大规模并行处理 MPP 框架，可线性扩展存储及计算能力。支持行列混合存储，可按业务需求选择最佳存储方案。查询引擎深度优化，查询效率数倍于传统数据仓库。

### 安全可靠

双节点同步冗余，实现用户无感的故障转移和容灾备份。分布式部署，计算单元、服务器、机柜三重防护，提高重要数据基础设施保障。用户集群独立部署，支持 VPC 隔离，数据访问安全多重保障。

# 应用场景

最近更新时间：2021-11-12 15:12:27

## 经营分析决策

在金融、零售等领域，需要对销售、资产、供应链等业务数据进行汇总分析，以便通过数据掌握公司经营情况，提高决策精准度及效率。

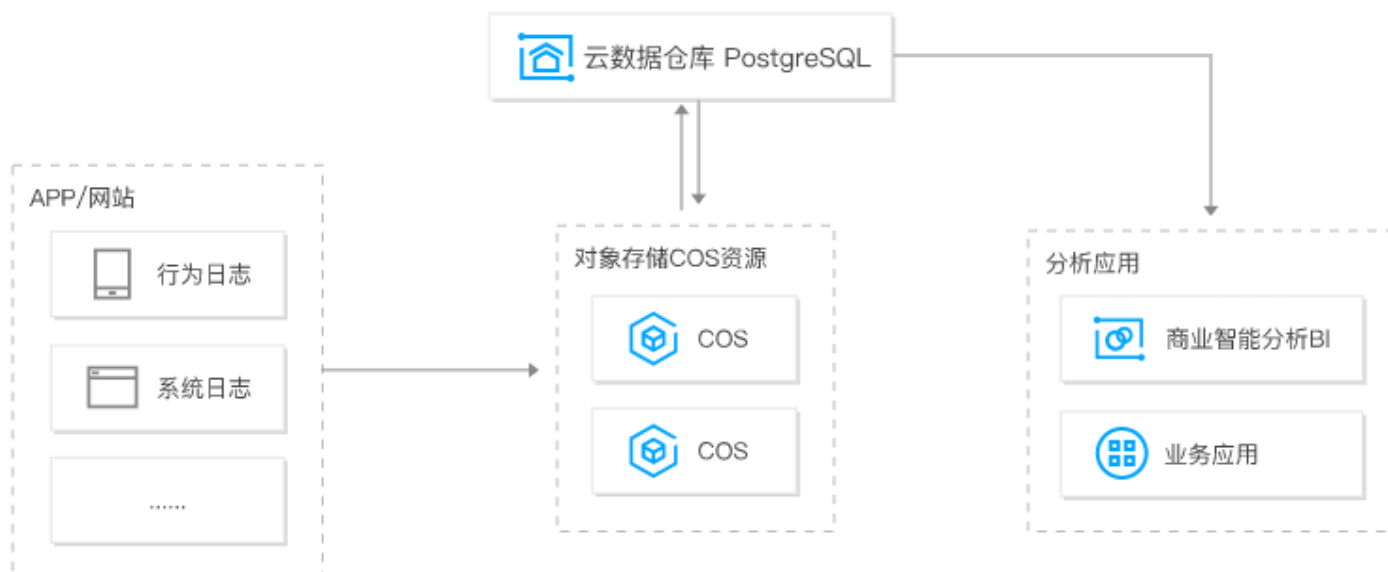
通过同步或 ETL 工具将分散在 TencentDB、Oracle、PostgreSQL 中的数据导入到云数据库 PostgreSQL 中，利用其对多源异构数据的分析能力，辅助业务决策。



## 海量日志分析

在互联网金融、游戏、O2O 等领域，经济高效地分析 PB 级用户行为、系统日志、订单等结构化或半结构化数据。

可先将海量数据直接存放于对象存储 COS 中，无需提前导入到云数据库 PostgreSQL，只需编写 SQL 便能对 COS 中数据进行快速分析及应用。



## 用户行为实时洞察

在互联网、游戏领域，对用户的行为进行实时分析、优化运营策略、提升资源运营效率。

可结合流计算服务，将用户实时操作数据进行过滤排序、汇聚分析，然后将结果实时同步到云数据仓库

PostgreSQL，借助于云数仓强大的 SQL 分析能力和查询效率，快速感知玩家行为趋势并及时调整运营策略。



# 节点规格

最近更新时间：2022-05-23 16:12:56

云数据仓库 PostgreSQL 支持如下两种节点规格：

- 计算密集型：可以提供更好的 I/O 能力，带来更高的性能。
- 弹性存储型：底层存储为增强型云盘，提供弹性伸缩扩展能力

在选择节点规格时，建议您根据需要的存储空间和计算能力进行选择。同时，云数据仓库 PostgreSQL 支持基于 COS 的外部表扩展，并可通过 gzip 实现外部存储上的数据压缩，不需要参与实时计算的数据可存放到外部存储中，进一步降低存储成本。

## 计算密集型节点规格列表

节点规格	CPU(核)	内存 (GB)	可用存储量 (TB)	双副本总存储量 (TB)	存储类型
nc2.large	16	64	1.8	3.6	1x本地 NVMe SSD

## 弹性存储型列表

### 节点规格

规格类型	CPU(核)	内存 (GB)
4C16G	4	16
8C32G	8	32
16C64G	16	64

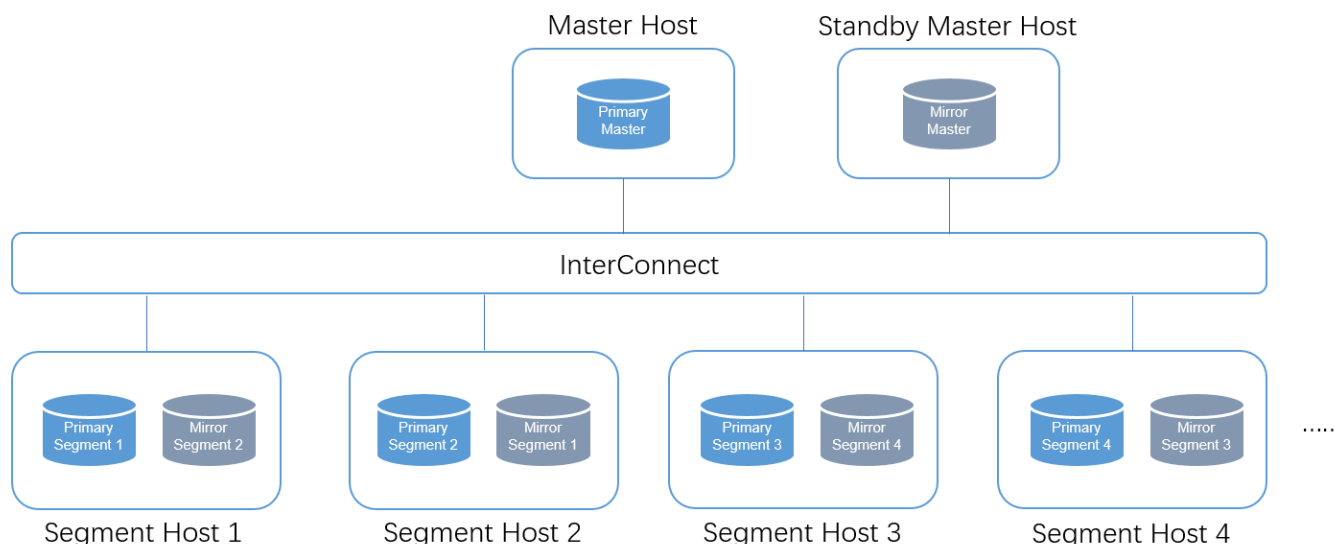
### 节点存储

磁盘类型	磁盘容量限制
增强型 SSD	200GB-6000GB
高性能云硬盘	200GB-6000GB

# 集群架构

最近更新时间：2021-11-12 15:12:44

云数据仓库 PostgreSQL 集群架构图如下图所示，采用 Shared-Nothing 的大规模并行处理架构，其中 Master 一主一备，分布在两台机器上。



- **Master 节点：**Master 节点不存储业务数据，只存储数据字典，负责生成、分发 SQL 执行计划到每个 Segment 节点，同时负责与客户端的交互及权限认证。
- **Segment 节点：**Segment 节点负责存储业务数据并执行由 Master 节点分发的 SQL 语句，同时为了保证每一个 Segment 服务在同一个性能水平上，每一个 Segment 节点机器有相同的资源配置，扩容时不作机型的改变。

为了保证集群高可用，每一个 Segment 节点上分别放置一个当前节点的 Primary Segment 和一个其他节点的 Mirror Segment 作为对应的备节点。当 Segment 节点不可用时，将采用镜像节点代替。当 Master 节点不可用时会自动切换至备 Master 节点（Standby Master），以保持集群可用状态，主节点恢复启动后会自动同步变更。

集群所有的业务数据存放于所有 Segment 节点的数据库上，每一张数据表都会被切片存放于各个 Segment 节点中，当进行分析时，所有 Segment 节点同时工作计算自己部分的数据，从而大大提升计算效率。

云数据仓库 PostgreSQL 支持数据备份及恢复功能，在开启数据备份后，每个 Segment 节点会创建包含重建数据命令的转储文件，在 Master 节点会创建几个包含系统目标、元数据文件、DDL 语句等信息的转储文件。在数据恢复时，所有的 Segment 都同时从本地备份文件恢复数据。