

ZTE中兴

# GoldenDB 分布式数据库 技术白皮书



## 法律声明

若接收中兴通讯股份有限公司（以下称为“中兴通讯”）的此份文档，即表示您已同意以下条款。若不同意以下条款，请停止使用本文档。

本文档版权所有中兴通讯股份有限公司。保留任何未在本文档中明示授予的权利。文档中涉及中兴通讯的专有信息。未经中兴通讯事先书面许可，任何单位和个人不得复制、传递、分发、使用和泄漏该文档以及该文档包含的任何图片、表格、数据及其他信息。

**ZTE**和**ZTE中兴**是中兴通讯的注册商标。中兴通讯产品的名称和标志是中兴通讯的商标或注册商标。在本文档中提及的其他产品或公司名称可能是其各自所有者的商标或注册商标。在未经中兴通讯或第三方权利人事先书面同意的情况下，阅读本文档并不表示以默示、不可反言或其他方式授予阅读者任何使用本文档中出现的任何标记的权利。

本产品符合有关环境保护和人身安全方面的设计要求，产品的存放、使用和弃置应遵照产品手册、相关合同或相关国法律、法规的要求进行。

本文档按“现状”和“仅此状态”提供。本文档中的信息随着中兴通讯产品和技术的进步将不断更新，中兴通讯不再通知此类信息的更新。

## 中兴通讯股份有限公司

**地址：** 中国深圳市科技南路 55 号

**邮编：** 518057

**网站：** <http://support.zte.com.cn>

**邮箱：** 800@zte.com.cn

## 目录

<b>1 文档概述</b> .....	<b>1</b>
1.1 文档使用范围.....	1
1.2 术语.....	1
1.3 缩略语.....	2
<b>2 产品介绍</b> .....	<b>3</b>
2.1 产品定位.....	3
2.2 产品特点.....	3
<b>3 体系结构</b> .....	<b>6</b>
3.1 架构设计.....	6
3.1.1 逻辑架构.....	6
3.1.2 物理架构.....	9
3.1.3 系统数据流向.....	11
<b>4 功能介绍</b> .....	<b>13</b>
4.1 分布式事务控制.....	13
4.1.1 隔离级别.....	13
4.1.2 分布式事务控制.....	13
4.2 基础功能.....	14
4.2.1 数据类型.....	15
4.2.2 函数功能.....	15
4.2.3 索引机制.....	16
4.2.4 存储过程.....	16
4.2.5 触发器功能.....	16
4.2.6 视图功能.....	16
4.2.7 执行计划.....	16
4.3 高级功能.....	17
4.3.1 扩展性.....	17
4.3.2 数据重分布.....	17
4.3.3 读写分离.....	18
4.3.4 导入导出.....	19
4.3.5 Oracle 兼容性.....	20

---

4.3.6 压缩功能.....	21
4.3.7 日切卸数.....	21
4.4 产品性能.....	22
4.4.1 缓存管理能力.....	22
4.4.2 并发控制机制.....	22
4.4.3 数据复制优化技术.....	22
4.4.4 大容量数据处理能力.....	23
4.5 安全性.....	23
4.5.1 数据安全.....	23
4.5.2 服务高可用.....	25
4.6 数据库管理功能.....	26
4.6.1 统一运维管理.....	26
4.6.2 运维数据分析挖掘.....	26
4.6.3 运维管理工具集.....	26
4.6.4 日志管理.....	27
<b>5 应用场景.....</b>	<b>28</b>
5.1 银行核心下移类应用.....	28
5.2 互联网金融类应用.....	29
5.3 政务云应用.....	30

# 1 文档概述

## 1.1 文档使用范围

本文针对中兴通讯 GoldenDB 分布式数据库的市场定位和特点，帮助用户从产品的体系架构、组件基本原理、产品功能和应用场景上全面了解本产品。

本文档适合初次接触本产品的用户，用于指导用户从宏观上对 GoldenDB 分布式数据库建立初步的了解和认识。

## 1.2 术语

本文使用的专用术语、定义，包括通用词语在本文档中的专用解释。术语具体说明参见下表。

表 1-1 术语表

术语	说明
分布式数据库	分布式数据库系统通常使用较小的计算机系统，每台计算机可单独放在一个地方，每台计算机中都可能存有 DBMS 的一份完整拷贝副本，或者部分拷贝副本，并具有自己局部的数据库。 位于不同地点的许多计算机通过网络互相连接，共同组成一个完整的、全局的逻辑上集中、物理上分布的大型数据库。
分布式事务	分布式事务，指事务的参与者、支持事务的服务器、资源服务器以及事务管理器分别位于不同的分布式系统的不同节点之上。
原子性	整个事务中的所有操作，要么全部完成，要么全部不完成，不可能停滞在中间某个环节。 事务在执行过程中发生错误，会被回滚（Rollback）到事务开始前的状态，就像这个事务从来没有执行过一样。
一致性	在事务开始之前和事务结束以后，数据库的完整性约束没有被破坏。
隔离性	两个事务的执行是互不干扰的，一个事务不可能看到其他事务运行中间某一时刻的数据
持久性	在事务完成以后，该事务对数据库所作的更改便持久的保存在数据库之中，不会被回滚。

### 1.3 缩略语

本文使用的专用缩略语以及说明参见下表。

表 1-2 缩略语表

缩略语	英文	中文
ACID	Atomicity、Consistency、Isolation、Durability	ACID，指数据库事务正确执行的四个基本要素的缩写。包含： <ul style="list-style-type: none"><li>➤ 原子性（Atomicity）</li><li>➤ 一致性（Consistency）</li><li>➤ 隔离性（Isolation）</li><li>➤ 持久性（Durability）</li></ul>
OLTP	On-Line Transaction Processing	联机事务处理
OLAP	On-Line Analytical Processing	联机分析处理

## 2 产品介绍

### 2.1 产品定位

GoldenDB 是安全、自主、可控的分布式数据库产品，产生背景：

- 云计算及互联网技术的兴起和发展，给各行各业带来大数据量的冲击，传统数据库难以应对应用层的高并发数据访问。
- 监管机构从国家信息安全的角度对 IT 基础设施提出了开源化、国产化、自主掌控的要求。
- 日趋严重的 IT 成本控制压力对 IT 行业提出更高要求。

GoldenDB 是中兴通讯的关系型分布式数据库产品，产品深耕金融、政企等行业，采用无共享架构，融入中兴通讯多年的数据库研发经验，为用户提供了高可用、高可靠、可扩展的“大数据+分布式数据库”解决方案；满足 OLTP 类应用，同时兼顾 OLAP 数据处理要求，提供统一的基础数据服务平台，有利于提升业务创新能力和用户体验。

### 2.2 产品特点

GoldenDB 分布式数据库产品主要特点如下：

- 提供数据库统一基础服务，实现各系统的扁平化。

通过分布式数据库解决方案，可对所有需要关系型数据库的应用系统提供统一的数据库服务，基于“大数据+分布式数据库”，实现数据的集中存放、统一管理和数据能力开放平台解决方案，为客户提供扁平化的基础数据处理平台。

- 自主、安全、可控。

完全自主研发，源代码全掌握，安全可控，已在金融行业成功商用，运行稳定、安全、可靠。

- 低成本。

支持 X86/ARM/OpenPower 等各类服务器的去中心化架构集群架构，成本低。

- 可扩展性。

GoldenDB 软件架构分层设计，计算节点、数据节点均可横向线性扩展，满足性能及容量的无限扩展需求。

- 高可靠性。

整个集群无单点故障，数据多副本，具备完善的数据备份恢复机制，支持双活数据中心，支持异地灾备。

- 强一致性分布式事务。

GoldenDB 具备完善的分布式事务处理机制，可保证读写及数据恢复的强一致。

- SQL 兼容性。

兼容 SQL 92、SQL99、SQL2003 标准，部分兼容 Oracle 语法。

- 分布式 SQL 优化。

内置大量的优化规则，对上百个场景进行优化，良好的复杂 SQL 语句兼容性和处理性能；支持 Prepare 预编译、执行计划缓存、数据集透传等功能，在保证数据一致性条件下实现高性能 SQL 处理。

- 分布式批处理。

提供分布式架构下批处理功能，满足金融、政企、运营商等行业日终大数据批处理需求，通过分布式 FetchSize 和存储过程功能对数据进行批处理，减少客户端与 DB 的访问次数，批量返回数据集并进行批量处理。

- 支持读写分离，提升读写效率。

系统根据负载情况及操作类型，把写操作发送到主库、读操作负载均衡到从库，提高从库的利用率。

- 灵活的数据切片技术。

支持哈希、范围、列表、复制、多级分片等多种数据分片规则，可以根据业务数据特征，选择最适合的分片技术把数据分别存储在多个数据安全组中；通过合理的数据分片规则，发挥分布式数据库的最佳性能。

- 智能高效的运维。

分布式数据库提供统一的管理门户和高效的运维工具，实现高效、智能、可视化运维管理。

- 分布式数据备份恢复。

支持在线热备，支持一键恢复到任意时刻，并支持全局一致的数据恢复。

- 在线数据重分布。  
支持数据节点的扩容、缩容，能高效地将数据均匀分布到数据库集群上；同时保证对在线业务影响小，且可操作性强。
- 分布式数据导入导出。  
数据导入导出与交易流分离，降低对联机交易影响。

## 3 体系结构

### 3.1 架构设计

分布式数据库架构采用了无共享的设计，从而在理论上获得了水平扩展不受限制的能力。

#### 3.1.1 逻辑架构

GoldenDB 分布式数据库总体逻辑架构如下图所示。

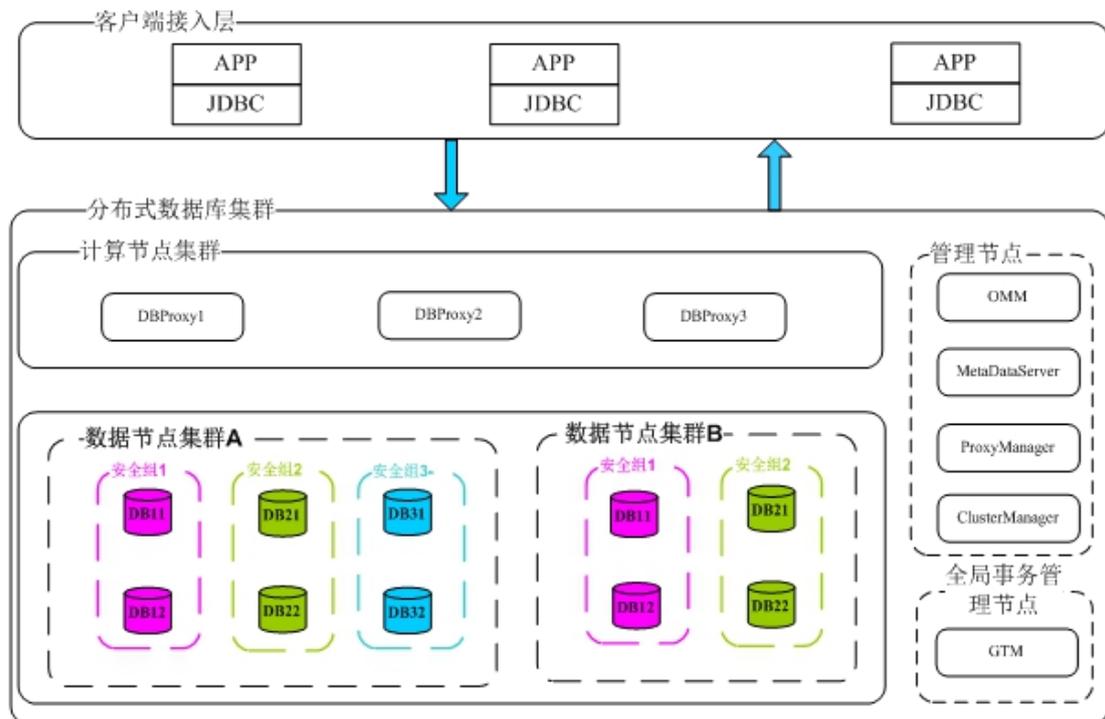


图 3.1-1 逻辑架构图

GoldenDB 分布式数据库由数据库驱动、计算节点集群、数据节点集群、管理节点和全局事务管理节点五个部分组成。

各个节点无需共享任何资源，都是独立自主的通用计算机节点，之间通过高速互联的网络通讯，从而完成对应用数据请求的快速处理和响应。

#### 数据库驱动

数据库驱动以集成的方式嵌入到具体应用中，和应用部署在一起。

GoldenDB 支持通用数据库标准协议，并能与 JDBC、ODBC 等常见开发语言所使用的接口驱动对接，包括但不限于 C、C++、Java、.Net、Python。GoldenDB

内部的拓扑结构、数据分布情况对应用透明。

从功能的角度来看，数据库驱动实现了计算节点的透明接入、负载均衡和故障透明转移。分布式数据库的计算节点负责处理应用的 SQL 请求，计算节点集群由多个计算节点组成，用户可以规划计算节点和应用之间的对应关系。

数据库驱动根据配置的规则将应用请求发送给合适的计算节点，并确保负载均衡地分配到这些计算节点上。当某些计算节点发生故障后，驱动层能够实施透明的故障转移，将应用的新请求发送给正常的计算节点，并在故障节点恢复后，能够将应用的请求重新路由到该节点上。

### 计算节点集群

计算节点集群层是分布式数据库的核心层，由无状态的计算节点 (DBProxy) 组成。

计算节点从驱动层或者管理节点接收用户的操作 (一般以结构化查询语言进行描述, Structured Query Language), 进行逻辑优化和物理优化, 生成满足分布式事务一致性的分布式查询计划。计算机节点在执行分布式查询计划时, 通过持续地访问数据节点, 从而完成用户的最终操作请求。

用户可以根据应用对可靠性、可用性、性能等因素的不同要求, 对计算节点进行合理的规划和划分。

### 数据节点集群

数据节点集群是应用数据的最终存储组件。所有的数据节点组成一个或多个数据库集群, 用户操作的事务不可以跨越多个数据库集群, 只能在一个数据库集群内进行。

“图 3.1-1 逻辑架构图”中数据库集群由一个或多个安全组 (DBGroup) 组成, 集群中每个表中的数据按照某种策略进行横向分片后存放到对应的安全组中, 分片策略包括复制策略、哈希策略、范围策略、列表策略。

- 复制策略: 适用于不常修改, 且频繁出现在关联或子查询中的小表。

复制策略下, 复制表的数据保存在每一个节点, 因此所有节点都需要这个表中数据的情况下, 可减少节点间网络数据的传输, 提高这种查询的性能。

- 哈希策略: 适用于将数据均匀的分布到预先定义的安全组上, 保证各安全组的数据量大致一致。

一般用于不需要关心分发字段的取值范围和具体含义, 且对该表的 SQL

操作基本都是等值操作的场景。

- 范围策略：适用于指定一个给定的列值或列值集合应该保存在哪个安全组上，常用于时间、日期、数值等类型的字段上，如数据按照自然月或自然天分布存储。
- 列表策略：适用于含有一系列限定性或枚举性的字段上，如数据按照机构代码、国家代码、地区代码分布存储。
- 多级分片：支持最多 5 层的分片规则设置，适用于需要精细化控制数据在集群中的分布形态的场景，如多法人场景，将不同法人的数据划分到不同数据安全组。

数据按照上述某种策略（复制策略除外）分片后，每个安全组上的实际数据在理论上只有总数据量的  $1/N$ （数据分布的均匀程度依赖于切分策略和真实数据分布的匹配程度）。

随着安全组数量的增加，每个安全组承载的数据量和读写负载会相应地减少，从而在数据节点集群内部具备了读能力和写能力的水平扩展。

安全组是由一个或多个数据节点构成的数据库节点组，组内的数据库节点拥有相同的数据。当安全组中存在多个数据节点时，其中一个数据节点为主用节点，其他数据节点都为备用节点，数据在主备节点之间实时复制。主用节点具备读写能力，备用节点可以提供读能力。安全组内的数据节点数量越多，可靠性就越高，读能力也得以进一步扩展。

## 管理节点

管理节点在分布式数据库中负责集群管理流程，不涉及业务的访问流程，无负载压力，一般采用两节点主备方式部署。管理节点按照功能分工，可分为如下几个子模块：

- 统一运维管理 OMM。

OMM 是 GoldenDB 分布式数据库产品的统一操作维护入口，用户可以在 OMM 上进行用户和权限管理、元数据管理、计算节点管理、数据节点管理、DDL 执行、节点扩容、备份恢复、系统安装、统计及告警管理等。

- 元数据管理器 MetaDataServer。

元数据指数据的元信息，如库、表、视图、触发器、存储过程、函数等数据模型的定义。

元数据管理器存放系统的全量元数据，是整个分布式数据库集群的元数据中心。

为了提高启动和运行效率，除了元数据管理器存有元数据定义外，计算节点和数据节点也会存放元数据定义，但计算节点和数据节点只存放本节点所涉应用的元数据定义，即当计算节点中的元数据和管理节点的元数据不一致时，会同步管理节点的元数据到本地。

此外，GoldenDB 元数据管理器还保存了整个集群的拓扑信息，因此是更广义的元数据管理。

- 计算节点管理 ProxyManager

负责管理计算节点集群。管理工作一般分为两类：

- 一类为集群的组建管理，包括计算节点的创建、启用、禁用和删除。
- 另一类为集群的应用管理，包括定义计算节点和应用的对应关系、计算节点异常后的数据恢复调度。

- 数据节点管理 ClusterManager

数据节点集群管理分为两类。

- 一类为集群的组建管理，包括数据节点、安全组、数据节点集群的创建、变更和删除。
- 另一类为集群的任务管理，包括数据节点异常、恢复后的调度管理、数据节点备份恢复的调度、数据重分布等功能的任务调度管理。

### 全局事务管理节点

全局事务管理器在分布式数据库中维护全局事务的全生命周期，提供申请、释放、查询全局事务的能力，采用集群方式部署。

### 3.1.2 物理架构

GoldenDB 分布式数据库总体物理架构如下图所示。

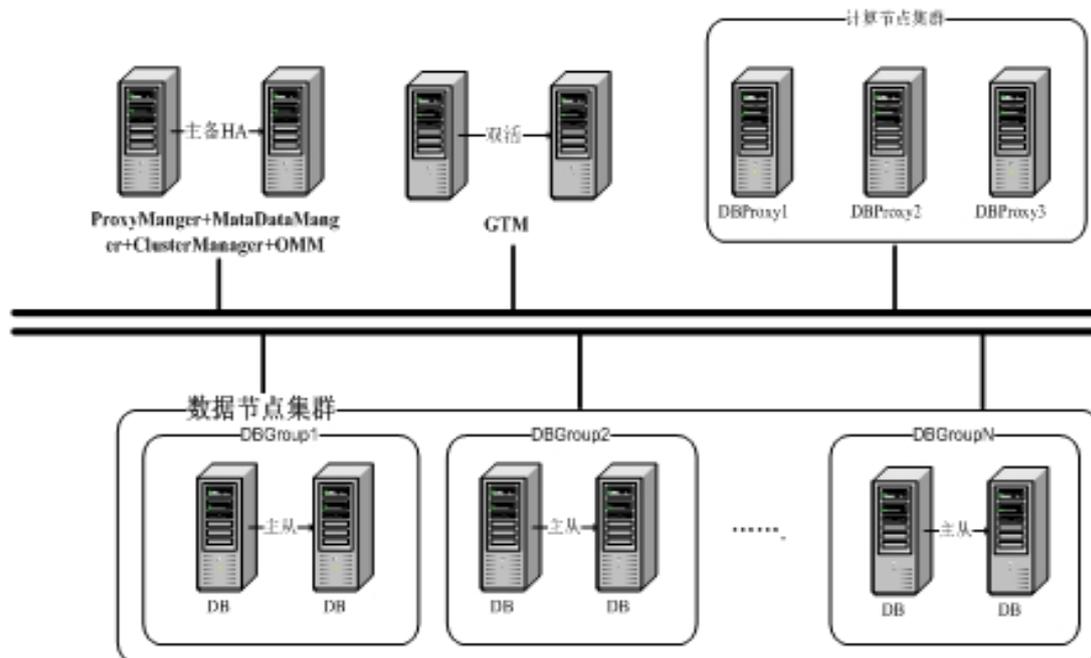


图 3.1-2 物理架构图

GoldenDB 分布式数据库推荐物理部署如下：

- 数据库驱动：可以直接使用标准数据库驱动接入 GoldenDB 集群，如 JDBC，可以支持按照最快响应时间和均匀随机方式将 SQL 请求下发至多个计算节点；当然也可以使用常用的负载均衡设备 LVS、F5 等接入 GoldenDB 集群。
- 计算节点 DBProxy：计算节点理论上可以单机部署，但通常要求部署 2 个以上，以便当某些计算节点发生故障后，驱动层能够实施透明的故障转移，将应用的新请求发送给正常的计算节点。
- 数据库集群：数据节点集群可以部署一个或多个，当业务之间的数据要求物理隔离时，可以将不同的业务数据存放在不同的数据节点集群中；当业务对隔离级别不高时，也可以多个业务的数据共享一个数据节点集群。

每个集群内部的安全组数目可以根据业务数据量、访问性能要求、硬件条件确定；理论上每个安全组内部的副本数目越多，可靠性就越高，但成本也越高。

- 管理节点：管理节点各模块既可以单机部署也可以使用 HA 软件进行双机冷备部署。
- 全局事务控制节点：分布式事务涉及节点间数据提交的一致性，该数据

的一致性主要依赖于全局事务的状态，所以全局事务节点必须是高可靠的，全局事务控制节点为集群模式，在发生异常时可以进行服务快速接管。

计算节点集群具备横向扩展能力，数据节点集群具备横向和纵向扩展能力。在高性能要求下，分布式数据库集群会涉及大量的计算节点和数据节点。

物理部署计算节点和数据节点时可以非常灵活，既可以部署在物理机器上，也可以部署在虚拟化平台上；同时每个节点既可以部署在一个设备上，也可以将多种类型的节点部署在同一设备上。不管采用哪种方式，一般都需要遵守如下原则：

- 计算节点需要分配更高的计算资源，数据节点需要分配更好的内存和存储资源。
- 归属于同一个应用的多个计算节点避免部署在同一台物理机器上，更进一步避免部署在同一机框或者数据中心。
- 归属于同一个安全组的多个数据节点避免部署在同一台物理机器上，更进一步避免部署在同一机框或者数据中心中。

### 3.1.3 系统数据流向

GoldenDB 的数据节点和计算节点是多对多关系。一个计算节点可以接入多个数据节点集群，提高计算节点利用率；一个数据节点集群也可以通过多个计算节点接入，多个计算节点间可以实现高可用，当某个计算节点异常后，不影响应用访问。

上述多对多关系是通过连接实例来建立的。连接实例是 GoldenDB 的端口访问控制机制，定义了计算节点端口信息至后端数据节点集群的一组对应关系，绑定了该连接实例的计算节点即提供对应数据节点集群的访问接入。应用向绑定了该连接实例的计算节点的对应端口发送 SQL 请求，该 SQL 请求实际会被计算节点路由到该连接实例定义的数据节点集群。

当不同应用接入相同数据节点集群时，可以定义不同的连接实例，达到接入权限、连接资源的隔离；当然，GoldenDB 也支持一个应用访问多个数据节点集群，但一般不推荐，用户可以自己选择。

GoldenDB 应用访问关系如下图所示。

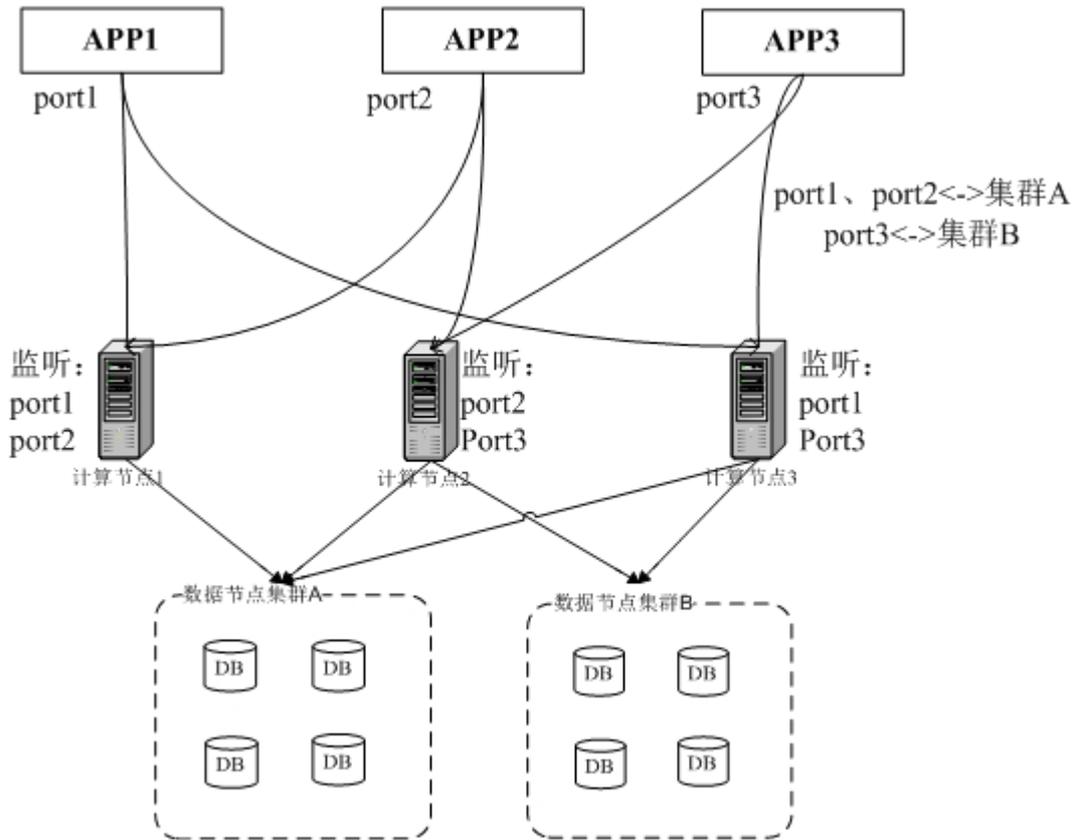


图 3.1-3 应用访问关系图

“图 3.1-3 应用访问关系图”说明：

- App1 和 App2 共享数据节点集群 A，App3 要求数据物理隔离，独享数据节点集群 B。
- 其中 APP1 使用 Port1 通过计算节点 1 和计算节点 3 接入集群 A，APP2 使用 Port2 通过计算节点 1 和计算节点 2 接入集群 A，APP3 使用 Port3 通过计算节点 2 和计算节点 3 接入集群 B。

结合上述对应关系，以 APP1 为例，一次 SQL 请求的数据流向说明如下：

1. 应用 APP1 使用 Port1 向计算节点 1 发起建链请求。
2. 建链成功后，APP1 在该链路上下发 SQL 请求。
3. 计算节点收到 SQL 请求后，进行语法解析、SQL 优化等，生成该 SQL 的分布式执行计划树；同时，通过查找 Port1 的连接实例定义信息，将子 SQL 请求下发至数据节点集群 A 的相关数据安全组。
4. 数据节点完成该子 SQL 的本地执行，并将结果集返回至计算节点 1。
5. 计算节点进行结果集汇总、计算，并将结果集返回至 APP1，过程中可能会在计算节点和数据节点之间产生多次的数据交互。

## 4 功能介绍

GoldenDB 分布式数据库是一款适用于 OLTP 场景，同时兼顾 OLAP 的分布式关系型数据库。

GoldenDB 分布式数据库完全兼容 MySQL 语法，兼容 SQL92、99、2003SQL 标准，同时兼容部分 Oracle 语法，其支持的语法的类别全面覆盖数据库语法类别，包括 DDL 数据定义语言、DML 数据操纵语言、数据管理语言、TCL 事务控制语言、常用函数、程序控制语言、批处理语法、序列等。同时，GoldenDB 对跨节点的复杂 SQL 操作支持全面，使得业务人员的开发工作量大幅降低，无需考虑大量的 SQL 改造。

在分布式数据库集群中节点数目和种类很多，为了方便管理，GoldenDB 各项产品化功能完善，统一运维界面 OMM 方便用户进行数据库集群的操作管理。

良好的 SQL 兼容性和可维护性，用户可以像使用单机数据库一样使用分布式数据库。

### 4.1 分布式事务控制

GoldenDB 做为事务型数据库，支持事务的 ACID 特性。

#### 4.1.1 隔离级别

GoldenDB 推荐使用 Read committed 的隔离级别。

#### 4.1.2 分布式事务控制

目前业界分布式数据库的数据一致性解决方案可分为强一致性、弱一致性和最终一致性。GoldenDB 实现了强一致性的解决方案，即当用户完成数据更新之后，后续所有读操作都能且只能读到更新以后的值，这是事务一致性的最高级别。

在 GoldenDB 的分布式事务解决方案中，数据节点作为分布式事务参与者要保证对自身数据操作的本地事务满足 ACID 属性，计算节点作为分布式事务协调者协调多个本地事务完成整个分布式事务控制，同时将全局事务状态实时纪录在全局事务管理器中。

基于 GTID 的分布式事务管理方案架构图如下所示。

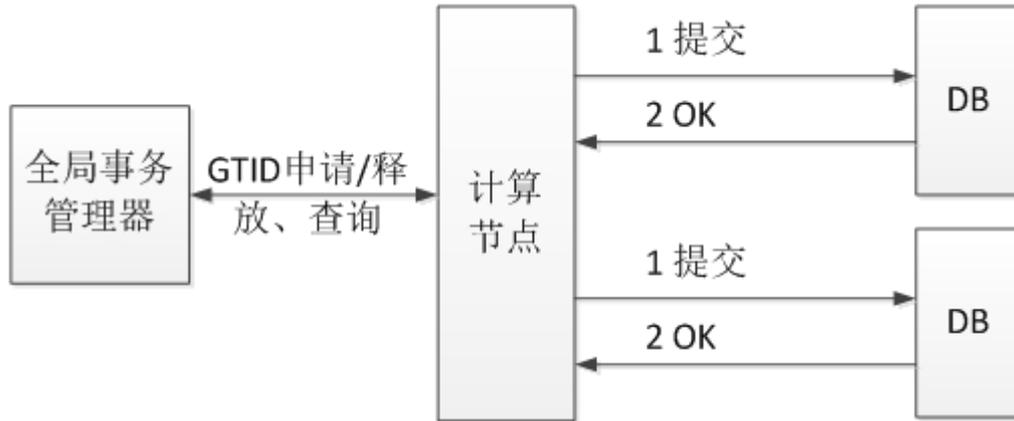


图 4-1 分布式事务管理方案

采用全局事务 ID 对分布式事务进行控制，其核心思想是全局事务控制和标签数据。

- 前者为每一个分布式写事务分配一个全局唯一的有序事务 ID，并根据事务的存活情况维护对应 ID 的生命周期。
- 后者在用户表中增加对应用透明的 GTID 列，并在该列中维护操作本行数据的最近一次分布式事务对应的全局事务 ID。

总体在流程交互上为“一阶段提交+补偿事务”的方式，如果事务在提交阶段有部分节点提交失败，本方案将采取回滚已成功提交的事务。

## 4.2 基础功能

GoldenDB 分布式数据库的 SQL 兼容 SQL92、SQL99、2003SQL 标准。GoldenDB 对 MySQL 语法兼容度高，同时也兼容 MySQL 工具生态圈，如 benchMarkSQL、Sysbench、NaviCat、DBVisualizer 等。

除了通用语法类型外，GoldenDB 还有些扩展语法，主要包括三类：

- 第一类是建表语法的扩展，如表分片信息的定义；
- 第二类是为了计算节点生成更好的执行计划优化而增加了一些 hint；
- 第三类是为了在原生语法上做了加固，更方便业务使用，比如我们在 Oracle 的 Sequence 序列语法中增加了批量获取多个 Sequence 值的接口。

GoldenDB 优秀的语法兼容性一方面使得业务开发者的语法使用习惯得以保留，另一方面也使得基于 Oracle 数据库开发的业务迁移至 GoldenDB 变得轻松，可以极大的降低开发工作量。

## 4.2.1 数据类型

GoldenDB 拥有完善的数据类型管理体系，且内置了日常应用所能涉及的所有数据类型，完全可以完全满足企业数据管理的需求。

内置的数据类型包括：

- 数字类型：INT、INTEGER、SMALLINT、TINYINT、MEDIUMINT、BIGINT、DECIMAL、NUMERIC、FLOAT、DOUBLE。
- 字符类型：CHAR、VARCHAR。
- 日期类型：DATE、TIMESTAMP、DATETIME、TIME、YEAR。
- 大对象类型：TEXT、BLOB。

## 4.2.2 函数功能

GoldenDB 函数支持的标量函数包括控制流函数、字符串函数、数学函数、日期和时间函数、转换函数。另外还提供了一些专用于 GoldenDB 的非 SQL 标准函数。

- 控制流函数

包括：CASE WHEN 控制、IF (expr1,expr2,expr3)、IFNULL(expr1,expr2) 等。

- 字符串函数

包括 CHAR\_LENGTH、COALESCE、CONCAT、CONVERT、IN、INSTR、LENGTH、LOWER、REPLACE、RTRIM 等字符串处理和比较函数。

- 数学函数

支持包括 ABS、AVG、CEIL/CEILING、CHAR、FLOOR、IN 等数学函数。

- 日期和时间函数

支持包括 DATE、DATE\_ADD、DATE\_FORMAT、DATE\_SUB、DATEDIFF、DAY、DAYNAME、DAYOFMONTH、DAYOFWEEK、DAYOFYEAR 等日期时间函数。

- 转换函数

支持 CAST、CONVERT 等数据转换函数。

### 4.2.3 索引机制

GoldenDB 支持 B 型树索引；拥有多种索引类型包括：唯一索引、复合索引、簇集索引。可以通过使用索引快速定位数据，提高查询性能。其中，针对唯一索引 GoldenDB 支持全局的唯一性约束。

### 4.2.4 存储过程

GoldenDB 支持存储过程功能，严格的说包含过程（PROCEDURE）和函数（FUNCTION），分别使用 CREATE PROCEDURE 和 CREATE FUNCTION 语句来创建。

### 4.2.5 触发器功能

GoldenDB 提供完备的触发器功能，支持插入触发器、更新触发器、删除触发器等触发器类型。

GoldenDB 触发器支持 GoldenDB 存储过程所支持的一切特性；

触发时间支持操作前触发（BEFORE）与操作后(AFTER)触发；

触发事件支持插入触发(INSERT)，更新触发(UPDATE)与删除(DELETE)触发。

### 4.2.6 视图功能

GoldenDB 支持视图功能。保存复杂查询，简化用户操作；提供单表和多表视图的支持。

GoldenDB 支持创建视图、删除视图、修改视图等视图基本操作。

### 4.2.7 执行计划

GoldenDB 支持 Explain 查看执行计划。计算节点支持查看分布式查询计划，并且可以多种格式输出，如仅打印计算节点的执行计划、仅打印 SQL 在数据节点的执行计划、同时打印计算节点和数据节点的执行计划等。

## 4.3 高级功能

### 4.3.1 扩展性

GoldenDB 的架构设计使得其具备高良好的扩展性，可以从如下几个角度来说明：

- DBProxy 计算节点集群是无状态的，可以动态任意扩展。
- 数据库集群可以根据需要进行如下几个维度的扩展：

- 在安全组内部添加数据节点增加数据副本数。

可以获得更高的可靠性，同时也可以通过读写分离功能达到读能力的线性扩展；

- 增加安全组数目。

获得数据节点存储和计算能力的线性扩展，在增加数据的分片时需要考虑如何将数据重新迁移到新分片上，即表数据的重分布，见下一节描述；

- 增加数据节点集群数目，扩展组建新的集群。

该方法可以在得到处理能力扩展的同时做到不同业务间数据的隔离性，用户可以根据要求选择。

GoldenDB 计算节点和数据分片的数目配比推荐为 1:1，性能可以随计算节点和分片数目的增加线性扩展，损耗控制在 5% 内。

若业务为 CPU 密集型，则计算节点数目要适当增加；若业务为 IO 密集型，计算节点数目可以适当减少。

### 4.3.2 数据重分布

业务的增长不可避免的需要对资源进行扩容，由于使用了分片技术，数据被切分成细小的分片分布在数据节点集群中。集群扩容后，原有的数据分片就面临着被打散重新分配的过程，这个过程就是数据重分布（ReSharding）。

GoldenDB 数据重分布功能采用后台数据迁移和日志回放相结合的方法，通过智能过滤，减少数据迁移量，并通过多表并发提升迁移效率，通过极限逼近和循环回放技术，减少新旧数据切换时间，实现不停机的在线数据重分布，达到数

数据库扩容、缩容对业务无感知。

GoldenDB 的数据重分布对在线业务影响小、且可操作性强，具备如下特点：

- 执行时间可控制，对业务影响秒级。

从前面的原理介绍中可以看出，重分布过程中数据的迁出迁入是个循序渐进的过程，仅在原分片和目标分片数据存在少量差距时才会禁写，因此对在线业务影响秒级，且不影响读服务；

- 支持多表并发重分布。

在分布式数据库的数据模型中，会将有关联关系的表采用相同的分片策略。在数据重分布过程中，这些相关表可以选择同步进行处理，以避免在数据重分布过程中由于相关表的分片策略临时不同而导致该 SQL 语句无法直接下压，影响 SQL 执行性能；

- 重分布数据智能识别。

数据重分布过程中会涉及到数据从源分片至目的分片的移动，开销很可观，GoldenDB 支持对表的源分片信息和目标分片信息进行详尽的分析比对，最大程度的减少数据在分片间的移动，提升重分布效率；

- 可操作性好。

GoldenDB 在统一运维界面 OMM 上提供重分布操作界面，全程可视化控制，包括执行、暂停、继续、取消、异常情况下的重试等；

- 产品通用性好。

对比需要预先做好分片规划的预 Sharding 方案，GoldenDB 的重分布方案对设计人员的要求相对较低，GoldenDB 支持任意分片策略间的重分布，彻底与业务模型解耦，无需在系统设计初期就要精确规划好未来的数据分布情况。

### 4.3.3 读写分离

读写分离是指利用数据节点集群安全组多副本，将部分读请求发往备用节点，提升系统的读能力。

在启动读写分离时，GoldenDB 的计算节点在收到应用 SQL 请求时，根据当前的语句类型和负载策略选择 SQL 下发的数据节点，将写操作发往主节点，将读操作发往备用数据节点。

在多个应用接入一个数据节点集群时，为了满足不同应用的需求，GoldenDB

支持对同一集群不同的连接实例，设置不同的读写分离模式，包括以下三种：

- 读主节点。

读操作默认向主发，当应用强制指定将读请求发往备机时，读请求即在备机间做负载均衡；

- 读备节点。

读操作仅在备机间根据配置的权重做读负载均衡；

- 主备节点。

读操作在所有副本间根据配置的权重做读负载均衡。

除了上述连接实例级别的读写分离模式设置，GoldenDB 还支持 SQL 级别的读写分离模式指定，应用可以在 SQL 语句后面添加 hint 信息强制发往主或备用节点，SQL 级别的优先级高于连接实例级别。

常见的使用场景如下：

- 由于数据在安全组内部的节点间同步存在时延，因此对实时性要求较高的 SQL 请求，应用希望将其发往主节点；
- 对于一些 SQL 如分析聚合类 SQL，应用希望将其发往备节点，减少对主节点的影响。

结合上述连接实例和 SQL 级别的读写分离模式设置，应用可以根据自己的希望设计出合理的读写分离策略。

#### 4.3.4 导入导出

数据导入导出一般用于系统间的数据迁移，小到导入导出一部分数据，大到数据割接、数据库迁移式升级、数据的分库、海量数据的迁移等，使用场景十分丰富，是数据库常用功能。

GoldenDB 支持将集群中符合查询条件的记录导出到指定的数据文件中，也支持将外部数据文件导入至集群合适的的数据分片中。

导入导出功能介绍如下：

- 外部接口。

GoldenDB 导入导出的外部接口为文件，待导入文件的格式要符合要求，一般为字段间用分隔符隔离的文本文件。导出结果也同样为文本文件。

- 工具插件。

导入导出可以认为是系统的一个功能插件，部署灵活，与在线交易流分离，可以降低对联机交易影响。

- 原始文件自动导入数据库集群。

GoldenDB 的导入导出工具，能自动根据导入表的分片规则，自动对原始数据文件记录分拣，将数据下发至对应数据节点，最终完成导入。

其中的各环节均支持并发处理，提升整体导入效率。工具也支持带条件导入，即仅将满足条件的纪录导入至系统中。

- 多分片数据统一出口导出。

GoldenDB 的导入导出工具，能自动控制各个分片并发进行数据导出，并可选择将导出文件上传至同一服务器；进一步，也可以将多个导出文件汇总成单个大文件，用户可以灵活选择控制导出过程。

- 容错处理。

导入操作耗时长，对系统资源占用多，且通常是将外部系统的数据文件导入到分布式系统中，而外部的文件往往会有异常数据。因此导入操作经常会遇到错误，GoldenDB 充分考虑导入异常的错误处理。

举例如下：

- 导入过程中遇到异常数据时，能将异常数据分拣出来，待后续处理。
- 流程处理的各阶段均有重试机制，最大程度保证导入成功。
- 系统异常时，能从异常时刻位置继续执行导入流程。
- 支持对原始文件纪录末尾的空值进行处理。

### 4.3.5 Oracle 兼容性

Oracle 有些非标语法的确会方便业务开发，同时目前大量现存业务是基于 Oracle 开发的，如果将这部分业务迁移至 GoldenDB，业务或多或少会涉及到改造工作，GoldenDB 将常用的 Oracle 语法支持起来可以降低业务迁移的工作量。部分 Oracle 语法列举如下：

- 控制流函数

包括：Decode 等。

- 字符串函数

包括 LPAD、RPAD、SUBSTR、TO\_CHAR、INSTR、TRIM、LTRIM/RTRIM、

||、 NVL 等字符串处理和比较函数。

- 数学函数

支持包括 TO\_NUMBER、ROUND 等数学函数。

- 日期和时间函数

支持包括 TO\_DATE、SYSDATE、TO\_TIMESTAMP 等日期时间函数。

- 系统关键字

支持 KEY、ROWNUM、PLUSJOIN 等系统关键字。

- 唯一序列

支持 SEQUENCE 相关的创建、获取等语法。

- 存储过程

支持 Oracle 存储过程特有语法。

### 4.3.6 压缩功能

GoldenDB 支持表数据压缩和备份压缩。表数据压缩时通过对表及分区表进行数据压缩来减少磁盘存储空间，在进行压缩操作前，可以对压缩率进行提前估算，从而做到有针对性的有效压缩，压缩算法采用 zlib 算法，压缩比在 2~11 之间。备份压缩，是指备份出来的数据会进行压缩来减少空间占用，GoldenDB 采用 quicklz 算法对备份文件进行压缩，压缩比在 2~10 之间。

GoldenDB 同时支持对表的数据空间进行归并整理，可以实现对通过数据压缩和归并整理优化出来的空间进行释放到表空间，从而实现最大限度的存储优化能力。

### 4.3.7 日切卸数

在金融领域，针对应用使用传统数据库存在批前卸数丢失跨天事务，分析报告不准确的问题，需要事后进行手工修复数据。GoldenDB 提供日切、卸数插件，可以获得准确的当日业务数据，更快得到正确的分析报告，无需在事后手工修复数据。以导出日切时刻的数据快照为例，使用 GoldenDB 日切卸数功能导出数据无需经过 ODS 处理，既能满足事务全局一致性，又能满足具有业务含义的数据一致性（在日切前发生的事务操作都保留在快照数据中，日切后发生的事务操作都不体现在快照数据上）。

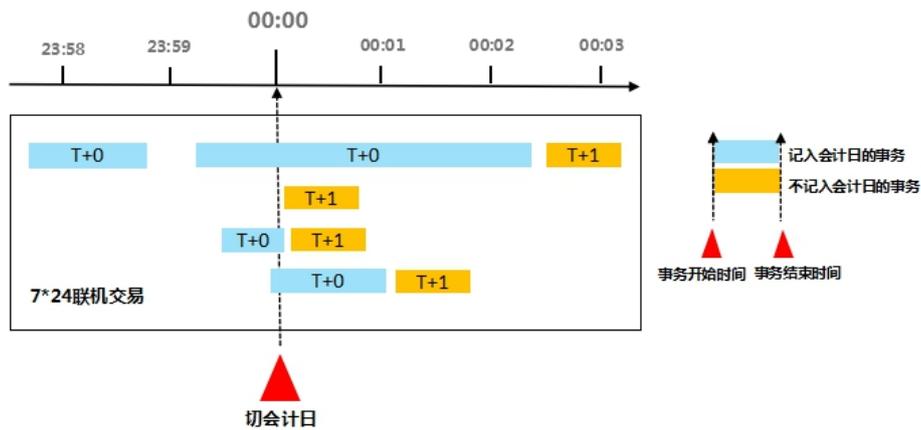


图 4-2 日切卸数示意图

## 4.4 产品性能

GoldenDB 通过执行计划缓存、并发事务控制、数据复制优化等手段，对整体性能进行提升，目前在银行核心业务系统中，TPS 能够达到 4 万，满足银行业务高并发低时延的特性。

### 4.4.1 缓存管理能力

GoldenDB 通过构建执行计划缓存、SQL 缓存，提高 SQL 解析效率、提高数据读取效率，在高频度读写系统中可减少磁盘 I/O 负担，提升整体系统效率。

### 4.4.2 并发控制机制

GoldenDB 在事务处理上通过采用成熟的主流技术来实现高效的事务处理，这些技术主要包括：以行级为主表级为辅的锁技术、单机多版本并行控制技术（multiversioning）、全局乐观锁+补偿事务。这些技术在保证事务 ACID 特征的前提下大大提高了事务的并发处理能力。

### 4.4.3 数据复制优化技术

分布式数据库通过增加副本数来提高高可用的系数，为了数据的安全可靠，数据必须在满足拥有了一定数量的副本之后，才能返回给客户端。GoldenDB 通过优化数据内核，将数据复制异步化，减少线程阻塞，提升利用率；同时通过分组管理、高低水位等技术，阻止了半同步复制向异步复制的切换，避免了数据丢失的可能；同时通过将同步副本上的回放线程的读写锁进行拆分，提升回放效率，

减少副本和主机的时延。

#### 4.4.4 大容量数据处理能力

GoldenDB 具备完备的数据库海量数据管理功能，提供 PB 级的数据容量支持，支持海量用户的大规模并发访问、具有高效索引和查询优化技术，具备海量数据处理能力并支持大数量用户的并发访问。

数据文件个数无限制，单表最大容量 40 亿，单表支持 1000 列。

### 4.5 安全性

#### 4.5.1 数据安全

##### 4.5.1.1 备份恢复

GoldenDB 拥有强大的备份和恢复功能，其备份类型分为两种：

- 全量备份：全量物理备份。
- 增量备份：最近一次全量备份后的增量部分备份（包含物理备份和逻辑备份）。

GoldenDB 的备份策略分为两种：

- 定时备份：可以指定每周或每天指定时刻进行备份
- 实时备份：可以在备份发起时进行备份

GoldenDB 备份的功能特点：

- 备份任务管理：GoldenDB 支持在 OMM 上进行备份任务的管理，包括备份任务的发起、备份历史操作的查阅等。备份任务的发起可以有二种模式。一种为实时备份，即用户配置好后立刻会发起一次备份任务；另一种为定时备份，用户可以设置定时备份策略，典型的策略为每周日备份全量数据，其他时间分别在周日全量备份的基础上做一次增量备份，如此，可通过全量数据和其他任一增量备份数据，快速恢复出想要的那天的数据。
- 备份文件存储：GoldenDB 的备份文件可在分片节点本地挂载 NFS 共享目录，将备份结果文件统一存放；同时 GoldenDB 也支持与外部备份系统对接，可以直接将备份结果存储到第三方存储介质。

GoldenDB 的数据恢复功能特点：

- 可恢复到任意时刻。由于 GoldenDB 有数据节点的全量及增量备份文件，同时有其运行过程中的 binlog 日志，借助这些数据可以将系统恢复至任意需要的时刻，但需注意全量及增量备份文件恢复的速度要大大快于 binlog 日志的回放速度，因此如无特殊要求，建议选择恢复到某次备份的结束时刻，以便更快的完成数据恢复；
- 一致性的数据恢复。由于 GoldenDB 备份了运行过程中的 binlog 日志及每个时刻的全局事务列表快照，因此可以根据这些信息恢复出一个全局一致性的数据副本；
- 恢复任务管理。GoldenDB 支持在运维平台上进行节点的一键恢复操作，用户直接在运维平台上选定要恢复的节点即要恢复到的时间点，即可进行自动化的数据恢复。

#### 4.5.1.2 加密与审计

为了确保数据的安全性，数据库需要对数据的全生命周期进行数据安全保障。分布式数据库系统支持的数据安全功能，内容如下：

- 访问鉴权

分布式数据库需要对来访用户进行鉴权，通过 IP 白名单的方式过滤连接请求，并通过国密算法对用户名密码进行解密校验。

- 通道加密

分布式数据库对客户端和数据库的连接进行 SSL 通道加密，防止通过抓包等手段查看业务数据。

- 数据加密

分布式数据库支持表级别和列级别的加密。通过列级别的加密，表中仅存储加密后的字符串，在不知道加密串的情况下，即使查询出来也会是乱码；表级别的加密可以加密整张表，在没有密钥文件的情况下，即使拷贝走表数据文件，也无法破解表内数据。

- SQL 日志审计

分布式数据库支持将所有的业务 SQL 写入审计日志文件，以便对 SQL 语句进行统计和审计，找出潜在的威胁。

## 4.5.2 服务高可用

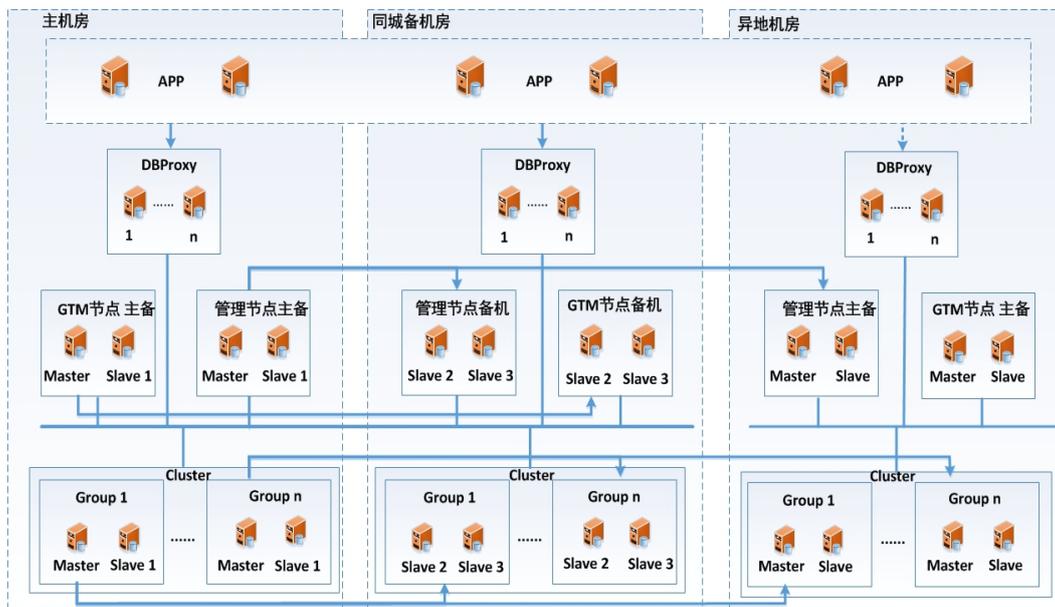
### 4.5.2.1 机房内高可用

在单机房内部，GoldenDB 具有模块级的高可用机制：

- 计算节点（DBProxy）无状态，通过集群部署，可随意伸缩；
- 全局事务管理节点集群部署，一主多备之间通过消息实时同步；
- 管理节点使用 HA 主备部署，通过数据库复制功能进行数据同步；
- 数据节点集群由多个分片（Group）组成，每个分片由多个数据节点组成，一主多备。

### 4.5.2.2 两地三中心

当进行两地三中心部署时，GoldenDB 拥有同城双活、异地灾备的高可用能力。



- 同城双活：依靠 GoldenDB 自身的强大集群功能，可实现准实时数据同步，保证同城主备机房数据一致性。同城双活采用全集群模式部署，同城灾备与主机房属于同一套集群，各管理节点主备机房分开部署。数据库集群采用一主多备模式，主机房中包括主机和备机，同城灾备机房部署备机，主机和备机之间采用数据库快同步复制方式，保障实时一致性。
- 异地灾备：本地、异地两个数据中心同时部署分布式数据库集群，根据网络性能可选快同步或异步方式进行数据同步，可用于距离远、网络条件不太好

的场景，能够支持孤岛演练、故障切换、异地回切等整体集成方案，并同时能保证数据的最终一致性。

## 4.6 数据库管理功能

GoldenDB 具备开放便捷的运维支撑体系，设计以智能化、最简化为目标。

### 4.6.1 统一运维管理

具备 GoldenDB OMM 统一运维管理平台。提供 web 界面形式的统一管理入口，支撑集群所有节点的一体化操作维护，简化用户对集群的管理。同时具备 API 接口供其他系统集成。主要功能如下：

- 一键安装，不停服务升级，保证升级期间的业务连续性，支持升级任务编排；
- 系统自动巡检，自动检测故障，系统主动自愈修复；
- 提供集群网络拓扑、用户管理、元数据管理、配置管理下发和配置比对；
- 数据重分布操作、备份恢复、主备一致性检查等可视化运维操作；
- 告警统一监控、管理；
- 数据库系统与用户权限控制。

### 4.6.2 运维数据分析挖掘

提供 GoldenDB InSight 运维数据分析挖掘组件。该组件是可插拔式的，从 GoldenDB 系统中采集数据，进行数据分析。对日志、空间、SQL、事务、配置等提供展示、分析、预测和辅助诊断功能，针对系统问题和瓶颈进行自动分析和预测，助力运维自动化和智能化。

### 4.6.3 运维管理工具集

提供 GoldenDB Tool Suite 运维管理工具集。供 DBA 定位问题时使用，包括操作管理命令行工具 dbtool、问题诊断工具、数据同步和导入导出工具等。其中：

- 命令行工具 dbtool

支持远程调用，工具可以输出统计、状态等多种信息，也封装了大量运维管

理操作接口，在系统维护、业务调试、问题诊断等多种场合均可以使用。大部分的 dbtool 功能同时支持 SQL 形式访问。

- 问题诊断工具 GoldenDB PD

可以提供从应用经计算节点到数据节点的自顶而下或自底而上的全链路交易跟踪及问题诊断能力。

- 双向数据同步组件 GoldenDB Sloth

提供 GoldenDB 与其他数据系统的双向实时同步功能。可用于平滑升级、数据实时分析等场景。该组件拓扑灵活，可扩展性好，同步过程中支持数据过滤和转换，并可保证同步数据一致性。

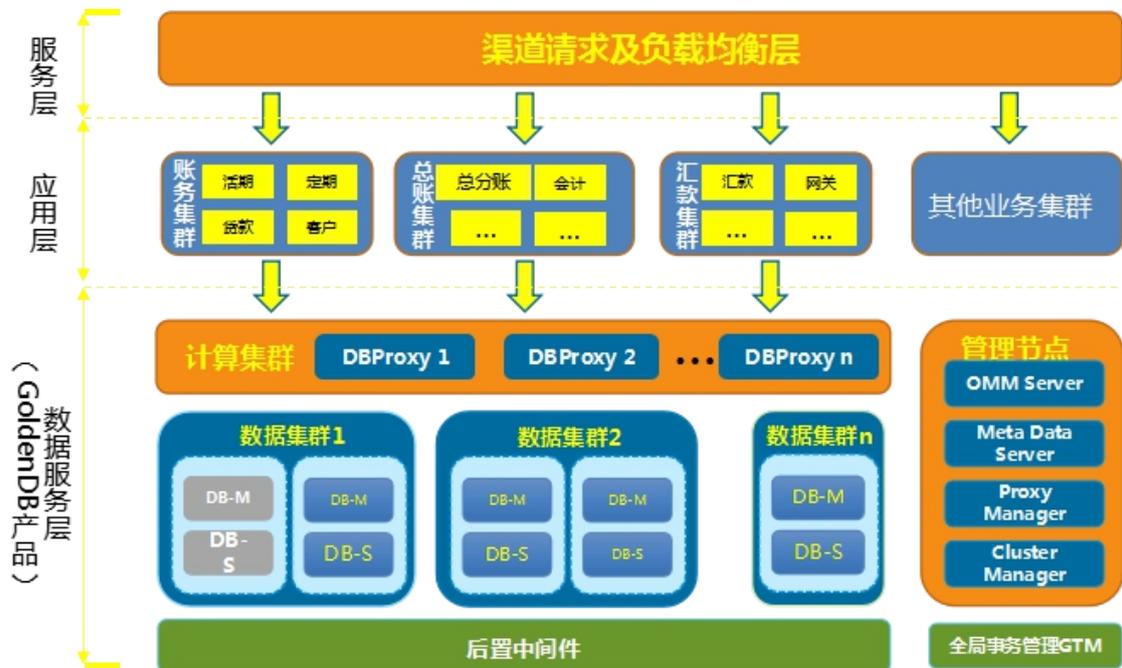
#### 4.6.4 日志管理

具备完善的运维日志集。包括运维日志、审计日志、一般查询日志、慢查询日志等，支持日志归集及全流程日志诊断分析。

GoldenDB 各个组件均有运维日志，并可以根据使用场景不同设置不同的日志输出级别，为降低分布式系统中的问题排查难度，GoldenDB 内部各组件支持全流程的日志输出，也提供接口和应用结合实现系统间的日志追踪。

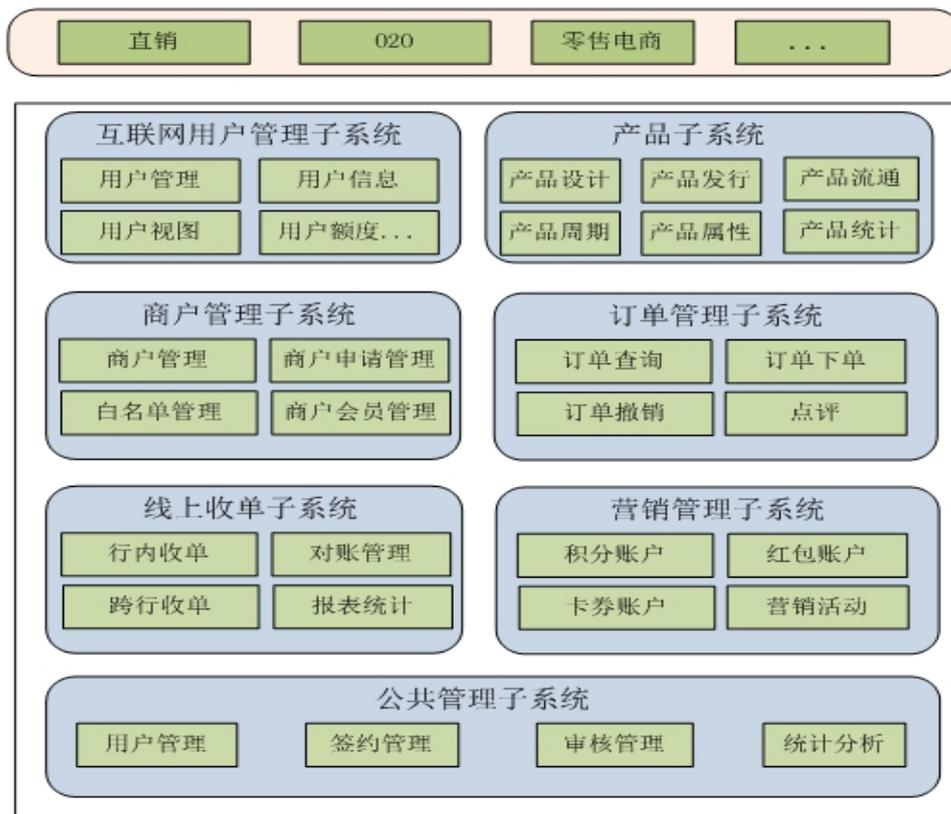
## 5 应用场景

### 5.1 银行核心下移类应用



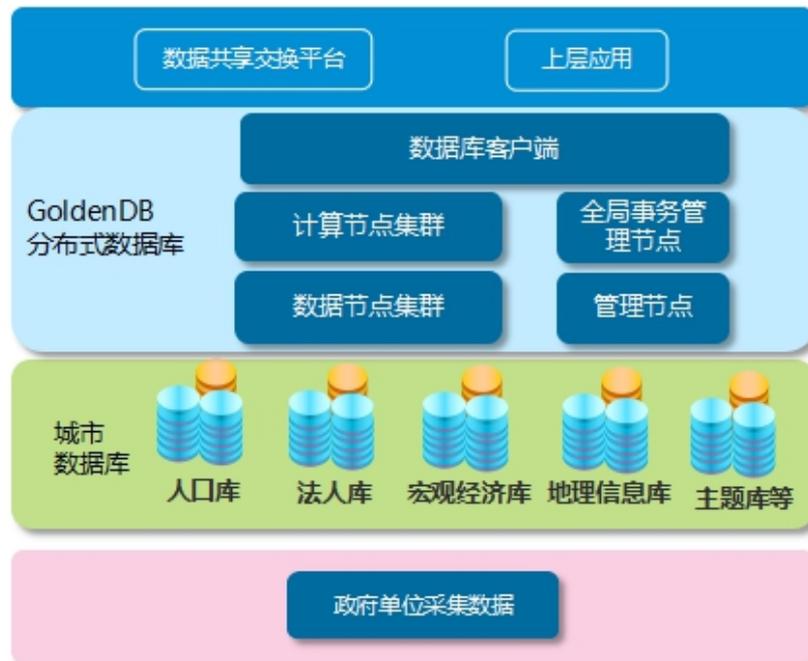
金融 IT 尤其是银行核心 IT 应用，对性能、稳定性以及安全性有着极高的要求，传统集中式架构无法满足持续增长的高并发和快速响应的需求。GoldenDB 分布式数据库是拥有自主知识产权的数据库产品，具有分布式事务一致性，容量、性能线性扩展、数据安全可靠以及易运维、易使用的特点。GoldenDB 支持中信银行进行核心业务下移改造，并已通过上线技术验证，达到了核心交易 4 万 TPS，同时 RPO，RTO 指标超出银监会对银行的监管要求。

## 5.2 互联网金融类应用



传统金融 IT 架构成本居高不下、交易处理能力存在瓶颈，结合互联网业务的特点及发展趋势，越来越多的互联网金融平台采用分布式架构进行系统建设工作，通过传统核心业务和互联网核心业务并行的方式，满足高并发的互联网业务发展。GoldenDB 在江苏省农村信用社联合社互联网金融平台的建设中，作为平台的基础数据库，为 2000 万用户提供数据服务，支撑了其 O2O、直销银行、收单等一系列的互联网金融应用场景。

### 5.3 政务云应用



在电子政务市场，各级政府为了打造信息共享和业务协同的智能化政务服务，需要建立统一的基础数据标准，并形成相关的基础信息库。在人口大省和大市，容量和性能给传统数据库带来压力。而分布式数据库的大容量、高并发特性，成为应对解决各级政务数据瓶颈的利器。使用分布式数据库海量数据的存储和快速查询的解决方案能够给为业务带来单机数据库所无法提供的数据可扩展性及性能可扩展性。