

金篆信科 GoldenDB 数据库 V6 技术白皮书

2024-12-18 发布

2024-12-18 实施

金篆信科有限责任公司发布

法律声明

若接收金篆信科有限责任公司（以下称为“金篆信科”）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。

本文档版权所有金篆信科有限责任公司。保留任何未在本文中明示授予的权利。文档中涉及金篆信科的专有信息。未经金篆信科事先书面许可，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。

在未经金篆信科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金篆信科产品和技术的进步将不断更新，金篆信科不再通知此类信息的更新。

金篆信科有限责任公司

目录

1 文档概述	6
1.1 文档使用范围.....	6
1.2 术语.....	6
1.3 缩略语.....	7
2 产品介绍	8
2.1 产品定位.....	8
2.2 产品特点.....	9
3 体系结构	11
3.1 架构设计.....	11
3.1.1 逻辑架构.....	11
3.1.2 系统数据流向.....	13
3.1.3 分布式/集中式一体化.....	14
3.1.4 混合负载（HTAP）能力.....	15
3.2 部署模式.....	16
3.2.1 集中式部署.....	16
3.2.2 分布式部署.....	16
3.2.3 混合部署.....	17
3.2.4 物理机部署.....	18
3.2.5 容器化部署.....	18
3.3 可靠性方案设计.....	19
3.3.1 单管理节点.....	19
3.3.2 本地单中心高可用.....	19
3.3.3 本地同城双中心高可用.....	20
3.3.4 两地三中心高可用.....	21
3.3.5 两地四中心高可用.....	21
3.3.6 三地五中心高可用.....	22
4 功能介绍	23
4.1 分布式事务控制.....	23

4.1.1 隔离级别	24
4.1.2 分布式事务控制	24
4.2 基础功能	26
4.2.1 数据类型	26
4.2.2 函数功能	26
4.2.3 索引机制	27
4.2.4 存储过程	27
4.2.5 触发器功能	27
4.2.6 视图功能	28
4.2.7 分区功能	28
4.2.8 执行计划	28
4.2.9 字符集	28
4.2.10 国产化适配	28
4.3 高级功能	29
4.3.1 扩展性	29
4.3.2 数据重分布	29
4.3.3 读写分离	30
4.3.4 导入导出	32
4.3.5 Oracle 兼容性	33
4.3.6 压缩功能	34
4.3.7 日切卸数	34
4.4 产品性能	35
4.4.1 缓存管理能力	35
4.4.2 并发控制机制	35
4.4.3 数据复制优化技术	35
4.4.4 批量协议	36
4.4.5 并行执行	36
4.4.6 流式查询	36
4.4.7 复杂查询并发	36
4.4.8 大容量数据处理能力	36
4.4.9 分布式批处理	37
4.5 安全性	37

4.5.1 安全审计	38
4.5.2 密码安全	39
4.5.3 数据安全	39
4.5.4 运维安全	43
4.6 数据库管理功能	44
4.6.1 安装与升级	44
4.6.2 数据配置	45
4.6.3 运维	45
4.7 GoldenDB 工具集	46
4.7.1 CACTool 迁移评估工具	47
4.7.2 Sloth 数据迁移工具	48
4.7.3 Replay 回放工具	50
4.7.4 InsightTool 巡检诊断工具	51
5 应用场景	53
5.1 金融行业	53
5.2 电信行业	54
5.3 政务行业	55
5.4 交通行业	56
5.5 医疗行业	57

1 文档概述

1.1 文档使用范围

本文针对中兴通讯 GoldenDB 数据库的市场定位和特点，帮助用户、合作伙伴等从产品的体系架构、组件基本原理、产品功能和应用场景上了解本产品。

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 数据库建立初步的了解和认识。

1.2 术语

本章节对 GoldenDB 数据库 V6 技术白皮书中常用的术语、定义以及英文缩写进行汇总说明，便于用户阅读。

表 1-1 术语表

术语/定义	说明
分布式数据库	具备分布式事务处理能力、可平滑扩展、分布于计算机网络且逻辑上统一的数据库。
集中式数据库	一种经典、传统的数据库结构，由一台或少数几台机器联合管理数据，一般不对数据进行分片，没有分布式事务。
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。
计算节点	GoldenDB 数据库中负责计算的组件，对外提供数据库实例接入服务；负责 SQL 优化、SQL 路由、数据节点的负载均衡、分布式事务的调度等。
数据节点	GoldenDB 数据库的数据存储组件，实现对数据的存储。提供单分片、多分片两种存储形式。分片包含多副本，保证数据安全。
全局事务管理器	全局事务管理，用于协助计算节点进行分布式事务管理，主要包括生成、释放全局事务 ID (GTID)、维护活跃事务以及当前活跃 GTIDs 的快照。
租户实例	GoldenDB 数据库提供多租户数据库实例管理，在 GoldenDB 中，可以创建多个数据库实例，分配给不同的业务使用。租户实例也称作集群，实例由计算节点、数据节点、全局事务管理器构成，可以是单分片或多分片。租户之间资源、数据相互隔离。
分片	GoldenDB 将数据依据分发策略存储在不同数据节点上，数据节点简称为分片。
DBGroup/安全分片组	多个数据节点组成一个安全分片组，其中有一个主节点(Master)，一个或多个备节点(Slave)，主备之间进行数据复制，实现数据的多副本存储。

术语/定义	说明
原子性	事务中的所有操作要么全部执行成功，要么全部失败回滚，没有中间状态。
一致性	事务开始前和结束后，数据库的完整性约束没有被破坏。
隔离性	并发执行的事务之间是相互隔离的，一个事务的执行不能影响其他事务的执行。
持久性	一旦事务提交，其结果应该是永久性的，即使系统发生故障，事务的结果也不应丢失。
Insight	GoldenDB 数据库运维平台的名称，负责数据库实例的创建、运维操作，提供监控、告警等运维功能。
JDBC	Java 数据库连接,是 Java 语言中用来规范客户端程序如何来访问数据库的应用程序接口，提供了诸如查询和更新数据库中数据的方法。
ODBC	开放式数据库连接接口，开发人员能够使用相同的 API 来连接、查询和操作不同类型的数据库。
OCI	Oracle 调用层接口，Oracle 公司提供的由头文件和库函数等组成的一个访问 Oracle 数据库的应用程序编程接口。它允许开发人员在第三代编程语言中通过 SQL 操纵 Oracle 数据库。
SQL	结构化查询语言，一种数据库查询和程序设计语言，用于存取数据以及查询、更新和管理数据库。
DDL	数据库模式定义语言 DDL(Data Definition Language)，用于定义或改变数据库或表的结构等初始化工作。
DML	数据操纵语言 DML(Data Manipulation Language)，用于管理和检索数据库中的数据。
DQL	数据查询语言 DQL(Data Query Language)，用于查询数据库表中记录。
DCL	数据控制语言 DCL(Data Control Language)，用于定义数据库的访问权限和安全级别。
gSync	GoldenDB 使用的数据节点之间的同步复制技术。
CACTool	GoldenDB 迁移评估工具。

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
CN	Computer Node	计算节点
DN	Data Node	数据节点

缩略语	英文	中文
GTM	GlobalTransactionManager	全局事务管理节点
MDS	MetaDataServer	元数据管理服务
PM	ProxyManager	计算节点管理服务
CM	ClusterManager	集群管理服务
JDBC	Java Database Connectivity	Java 数据库连接
ODBC	Open Database Connectivity	开放数据库连接
OCI	Oracle Call Interface	Oracle 调用层接口
SQL	Structured Query Language	结构化查询语言
API	Application Programming Interface	应用程序接口
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： 原子性（Atomicity） 一致性（Consistency） 隔离性（Isolation） 持久性（Durability）
OLTP	On-Line Transaction Processing	联机事务处理
OLAP	On-Line Analytical Processing	联机分析处理
HTAP	Hybrid Transaction/Analytical Processing	混合事务分析处理
DDL	Data Definition Language	数据定义语言
DML	Data Manipulation Language	数据操纵语言
DQL	Data Query Language	数据查询语言
DCL	Data Control Language	数据控制语言

2 产品介绍

2.1 产品定位

GoldenDB 是安全、自主、可控的集中式/分布式数据库产品，产生背景：

- 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
- 主管部门从信息安全等角度对 IT 基础设施提出了安全可靠的要求
- 日趋严重的 IT 成本控制压力对 IT 行业提出更高要求

GoldenDB 是中兴通讯自主研发的集中式/分布式数据库产品，作为成熟稳定商用领先的金融级数据库产品，深耕银行(国有大行、股份制银行、城商行等)、证券、保险、电信运营商等多个关键领域。产品融入中兴通讯多年的数据库研发经验，为用户提供了强一致、高可用、高可靠、高安全、高性能、可扩展的一体化的数据库解决方案；满足 OLTP 和 OLAP 不同类应用要求，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 通过全套数据库解决方案，可对需要关系型数据库的应用系统提供统一的数据库服务，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。GoldenDB 数据库在功能、高可用、安全性、高可用提供的主要特性如下：

- 主要功能特性：

- 1) 强一致性分布式事务。

GoldenDB 数据库从数据库层面实现分布式事务的强一致性，对应用透明，无需应用改造，提升应用迁移效率。采用分布式事务管理器和已提交事务自动回滚相结合的技术，解决分布式事务的强一致性问题。使用全局事务 ID 解决分布式事务的并发控制，严格保证事务 ACID 特性，在实现强一致性的同时确保性能优异。

- 2) 灵活的数据切片技术。

GoldenDB 数据库分布式场景下支持哈希、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。

- 3) 可扩展性。

GoldenDB 软件架构分层设计，基于 Share-Nothing 架构，采用集群方式部署，支持集中式和分布式混合部署。其中分布式集群实现各组件的灵活扩展，从而提供高性能的数据库服务。同时结合数据动态重分布和读写分离等技术，实现性能的线性扩展。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。通过增加计算节点，可以提升系统的业务处理能力；通过增加数据节点和在线数据重分布功能，可以提升系统的存储能力。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。

- 高 SQL 兼容性。

GoldenDB 支持 SQL92、SQL99、SQL2003 标准语法，兼容 MySQL 语法，兼容常用 Oracle、DB2 语法。支持多级分片、存储过程、全局唯一索引、MVCC、闪回、XA 接口等高级特性。

- 数据备份恢复。

GoldenDB 提供在线热备功能，支持一键恢复到任意时刻、指定库表备份恢复，克隆在线数据恢复，并支持分布式场景下的全局一致数据恢复。

- 支持读写分离，提升读写效率。

提供灵活的读写分离功能，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。支持读主节点、读备节点、主备节点权重配置，在线权重调整，支持灵活的负载均衡模式，灵活提升系统读性能。

- 高可靠性

GoldenDB 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。支持多地多中心（AZ）组网，任何节点不存在单点故障，可以支持多种组网架构，创新研发了 gSync 数据复制技术，针对不同的业务场景灵活配置不同的策略来满足不同的可用性和可靠性要求，提高系统吞吐量的同时，实现同城 RPO 为 0。满足金融级数据高可靠、服务高可用要求。GoldenDB 支持机房级故障自动切换、支持异地带载演练和异地带压演练。

- 云化支持能力

GoldenDB 借助云技术来实现数据库即服务，具备集中管控、快速部署、高效运维，以及弹性扩缩容等优点。GoldenDB 可对接云化平台标准接口，提供网络和存储持久化方案，把云平台 and GoldenDB 高可靠机制相结合，实现云化场景下的数据库多副本，节点自愈、故障隔离等能力。

- 支持混合负载

GoldenDB 支持混合负载（HTAP），实现了分布式并行计算处理能力，具有更快速的关系型数据复杂计算处理速度。应用程序可使用 SQL 语句查询 PB 量级的数据，可在毫秒级或秒级返回 AP 类查询响应，用于数据分析、海量数据聚合和报表生成等联机分析处理的使用场景。GoldenDB 进一步扩展列存存储 VEngine 模式，适用于交互式 SQL 数据仓库应用场景，用于提升业务报表统计和多维分析的处理性能。

- 自主、安全、可控。

GoldenDB 数据库自主研发，核心代码 100%掌握，并在金融、电信等多个关键行业核心业务大规模商用。GoldenDB 数据库通过安全可靠测评、EAL4+、商用密码认证，可信数据库安全专项等多项安全和自主可控相关的专项测试。

- 高性能

北京国家金融科技认证中心、国家工业信息安全发展研究中心、建设银行、中信银行、银联数据、中国移动等众多权威机构和重要客户组织的测试中性能排名第一。

3 体系结构

3.1 架构设计

GoldenDB 是自主研发的数据库系统，整体由计算节点、数据节点、全局事务管理器、管理节点四种核心模块组成，整体架构采用高可靠性设计，无单点故障。计算节点为无状态多节点部署，数据集群内由多个安全分片组构成，每个安全分片组内数据节点主备多机部署，全局事务管理器主备多机部署。

3.1.1 逻辑架构

GoldenDB 数据库总体逻辑架构如图 3-1 所示

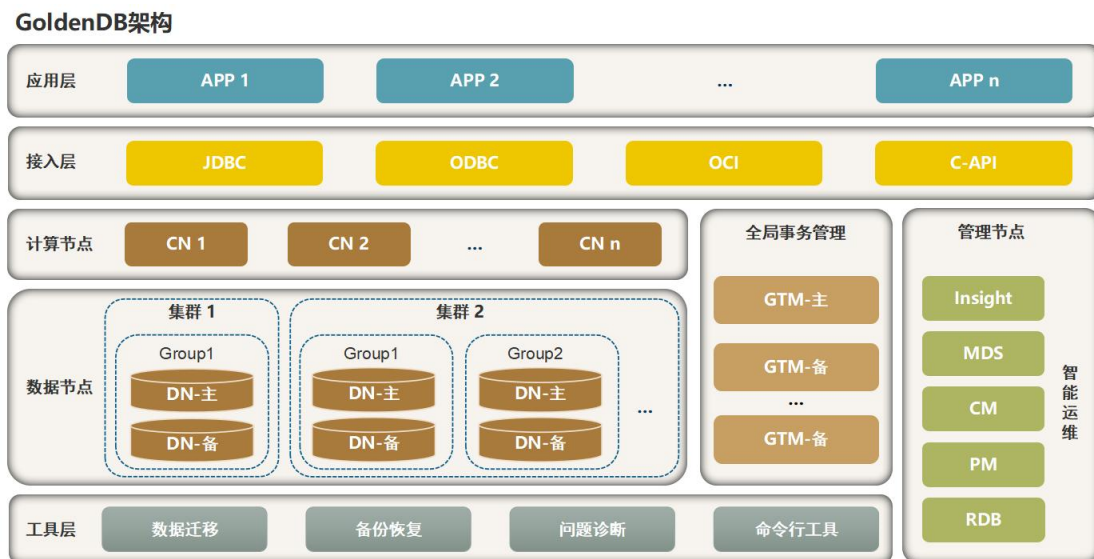


图 3-1 GoldenDB 逻辑架构图

GoldenDB 数据库由计算节点、数据节点、管理节点和全局事务管理节点四个部分组成。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。

计算节点

计算节点集群是分布式数据库的核心层，由无状态的计算节点组成。计算节点从应用层驱动或者管理节点接收 SQL 请求（结构化查询语言 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；在执行分布式查询计划时，通过持续地访问数据节点，完成 SQL 请求的最终计算。

数据节点

数据节点集群是数据的最终存储模块，由多个安全组组成，每个表中的数据按照特定策略进行横向分片后存放对应的安全组中。

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。

数据分片策略包括复制策略、哈希策略、范围策略、列表策略、多级分片。

- 哈希策略：适用于将数据均匀的分布到预先定义各安全组上，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。
- 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。
- 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。

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。管理节点具备如下功能：

- 可视化统一运维管理。GoldenDB 的管理平台（Insight）提供 GoldenDB 数据库产品的统一操作维护入口，用户可以在 Insight 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。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。元数据管理器还保存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。管理工作包括计算节点的创建、启用、禁用、删除和高可用调度。
- 数据节点管理，包括数据节点、安全组、数据节点集群的创建、变更、删除和数据节点异常、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 系统数据流向

GoldenDB 的计算节点和数据节点是多对多关系。一个计算节点可以接入多个数据节点子集群，提高计算节点利用率；一个数据节点子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服务端口来建立的。服务端口是 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子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子集群的访问接入。应用向绑定了该服务端口的计算节点的对应端口发送 SQL 请求，该 SQL 请求会被计算节点路由到该服务端口定义的数据节点子集群。

当不同应用接入相同数据子节点集群时，可以定义不同的服务端口，达到接入权限、连接资源的隔离。

GoldenDB 应用访问关系如下图所示。

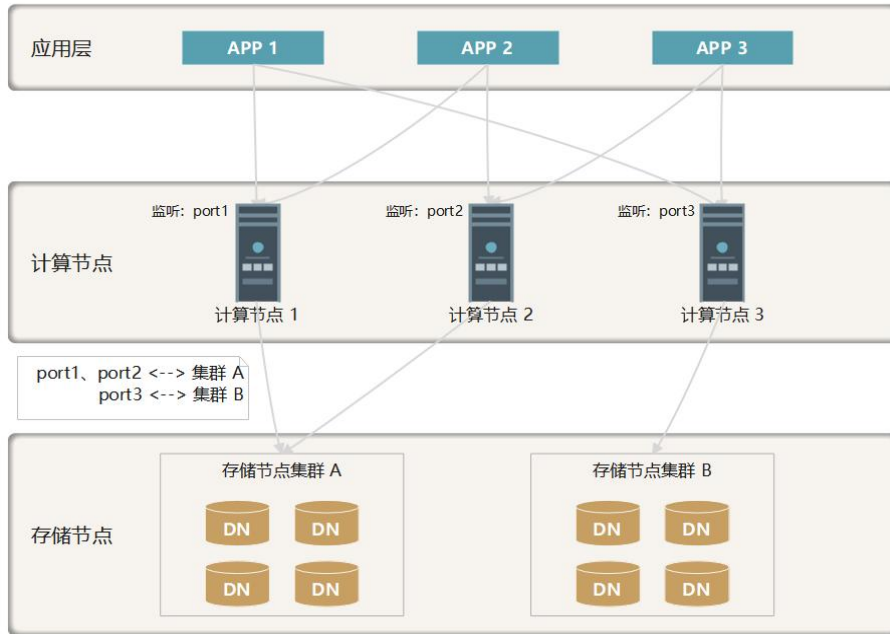


图 3-2 应用访问关系图

“图 3-2 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. 数据节点完成该请求的本地执行，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。

3.1.3 分布式/集中式一体化

GoldenDB 支持多种安装部署模式，适应不同的需求和场景。实现分布式

/集中式一体化运维管理。

多种部署模式：

- GoldenDB 提供了集中式部署模式。在集中式部署模式下的租户只包含单分片租户(CN/DN)且不带 GTM 节点。这种模式适用于数据规模比较小，不用按分片存储的业务。它只有普通事务，没有分布式事务。
- GoldenDB 支持分布式部署模式。分布式部署模式下的租户只包含多分片租户(CN/DN)且带 GTM 节点。这种模式适用于数据规模比较大，单个分片存储不下的业务。因为数据分布在多个分片，事务执行可能跨多个分片，即存在分布式事务，需要全局事务节点管理分布式事务。
- GoldenDB 还提供了集中式+分布式混合部署模式。在这种模式下，包含多种类型的租户，其中租户可分为单分片租户且不带 GTM 节点以及多分片租户且带 GTM 节点。这种模式适用于不同的数据规模业务要求，较为灵活。

关键部署特性：

- GoldenDB 支持同一集群的多租户特性，同一集群支持 MySQL 和 Oracle 两种兼容模式租户，兼容不同存量业务的快速对接和迁移。从而可以有效地支持多租户环境下的数据库服务。
- 在部署场景方面，GoldenDB 支持多种配置。可以选择单管理节点部署，适用于较小规模的应用；也可以选择本地单中心高可用部署，提高系统的可用性和容错性；还可以选择本地同城双中心高可用、两地三中心高可用、两地四中心高可用等复杂场景，满足多地域、多数据中心的需求。并且支持数据库租户的可用资源的灵活可配，包括 CPU、内存等单机资源和可用数据库节点数等。
- GoldenDB 可以部署在物理机，也可以进行容器化部署。容器化部署可以提供更高的灵活性、可移植性和资源利用率，使得 GoldenDB 能够更好地适应云计算环境和容器编排平台。

总之，GoldenDB 提供了多种安装部署模式、适应不同的场景和平台，为用户提供了可靠、高效的数据管理解决方案。

3.1.4 混合负载（HTAP）能力

GoldenDB 支持混合负载处理能力，同时满足交易类应用系统和分析类应用系统的数据库要求。其中 TP 部分由计算节点(CN)处理，并可增加 AP 组件(Coordinator 和 Worker)进行 SQL 统计查询，具有更快速的关系型数据复杂计

算处理速度，适用于在毫秒级、秒级或分钟级返回查询响应，用于比较复杂的数据分析、海量数据聚合和报表生成等联机分析处理的使用场景。

GoldenDB 可选装列存存储 VEngine 节点，适用于交互式 SQL 数据仓库应用场景，用于提升业务报表统计和多维分析的处理性能。通过 VEngine 节点，可创建行存模式表、列存模式表和行列混合表三种表类型。GoldenDB 对行列混合表同时在行存存储节点和列存存储节点创建表，并通过日志同步组件自动进行数据准实时同步。

应用的 SQL 请求统一通过 CN 接入，CN 根据 SQL 所访问表的存储类型进行自动路由：CN 可直接访问行式存储，或者通过 AP 的 Coordinator/Worker 调度，对行式表或列式表进行查询计算。

3.2 部署模式

3.2.1 集中式部署

集中式部署指的是单分片租户，且不带 GTM 节点。GoldenDB 的集中式部署模式适用于数据规模较小，不需要按分片存储的业务。在这种模式下，仅需支持单机事务即可。

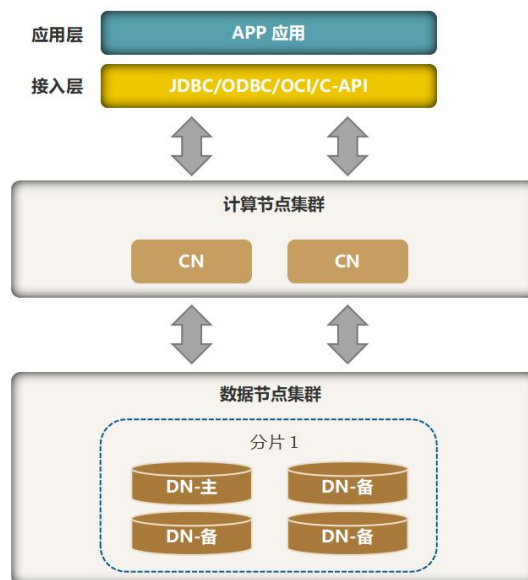


图 3-3 集中式部署

3.2.2 分布式部署

分布式部署指的是多分片租户，且带 GTM 节点。GoldenDB 的分布式部署

模式适用于数据规模较大,单个分片无法满足存储需求的业务。在这种模式下,数据分布在多个分片上,事务可能涉及多个分片,因此支持分布式事务,并通过全局事务节点管理分布式事务。

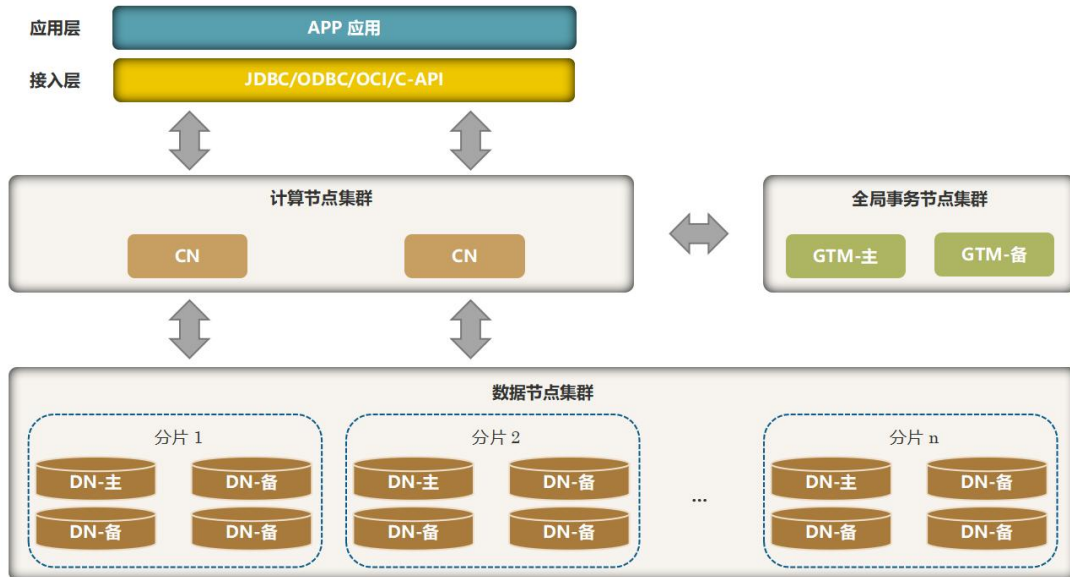


图 3-4 分布式部署

3.2.3 混合部署

混合部署指的是既有单分片租户,也有多分片租户。GoldenDB 的集中式+分布式混合部署模式提供了更大的灵活性,适用于不同规模和需求的业务。

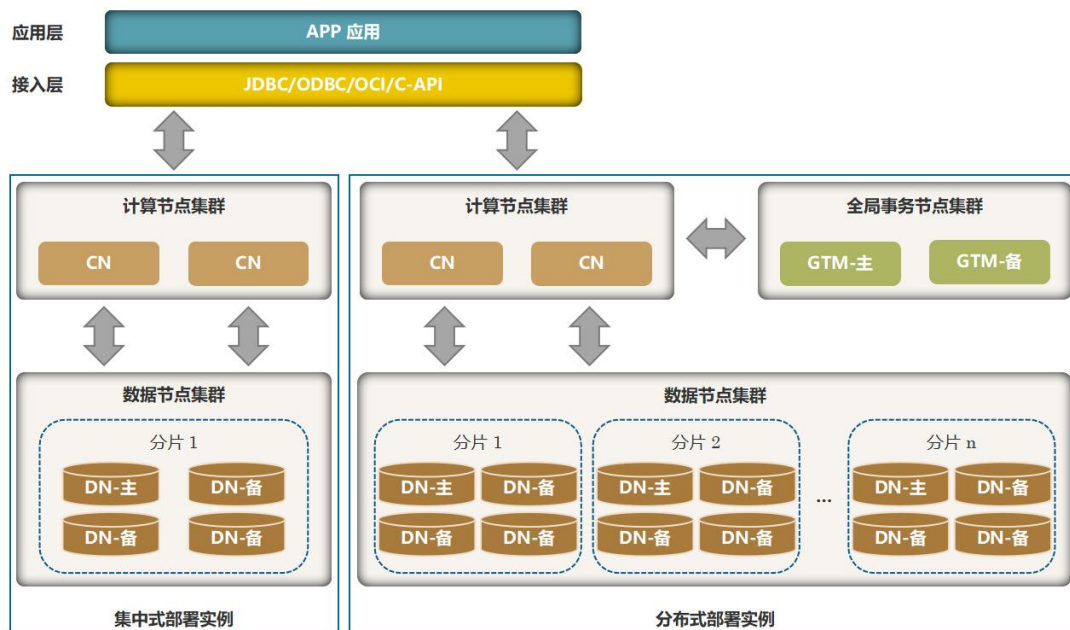


图 3-5 混合部署

3.2.4 物理机部署

GoldenDB 部署载体分为物理机部署和容器化部署。物理机部署是指 GoldenDB 管理节点 (CM/PM/MDS/InsightServer 等) 和非管理节点 (CN/DN/GTM/LDS 等) 都部署在物理机。GoldenDB 物理机部署的优势在于提供了灵活性、可靠性和性能优势。

- GoldenDB 物理机部署提供了较高的可靠性。由于物理机部署能够将数据库系统与硬件资源进行隔离，一旦出现故障或者需要进行维护，只需对相应的物理机进行操作，而不会影响其他部署在其他物理机上的数据库系统。这种隔离性能够提供更高的可靠性，确保数据库系统的稳定运行。
- GoldenDB 物理机部署还能够提供更好的性能。通过合理配置和管理物理机的硬件资源，可以充分利用处理器、内存和存储等资源，提高数据库系统的运行效率和性能。同时，物理机部署还能够根据实际需求进行横向和纵向扩展，以满足不同规模和负载的数据库管理需求。

综上所述，GoldenDB 物理机部署的优势在于可靠性和性能优势。通过合理配置和管理硬件资源，GoldenDB 能够提供稳定、高效和可靠的数据库管理解决方案，满足不同场景下的需求，并提供优质的数据库服务。

3.2.5 容器化部署

容器化是指 GoldenDB 管理节点 (CM/PM/MDS/InsightServer 等)、非管理节点 (CN/DN/GTM/LDS 等) 都部署在容器中。容器化部署带来了许多优势，使得数据库管理更加灵活、高效和可靠。

- 容器化部署提供了更高的灵活性。GoldenDB 容器可以在不同的环境中轻松部署，无论是在本地服务器、云平台还是私有数据中心。容器化部署使得 GoldenDB 可以与其他容器化应用程序或服务一起运行，并且可以根据需要进行快速扩展或缩减，以适应不同规模和需求的工作负载。
- 容器化部署提供了更高的可移植性。GoldenDB 容器可以在不同的操作系统和硬件平台上运行，而无需重新编写或调整代码。这意味着 GoldenDB 可以轻松地在不同的环境中迁移和部署，不受特定操作系统或硬件的限制，从而提供更大的灵活性和便捷性。
- 容器化部署还提供了更高的效率和资源利用率。GoldenDB 容器可以以轻量级的方式运行，减少了资源的消耗，同时也提供了更快的启动时间和更快的部署速度。容器化部署还可以通过自动化和集中化的管理，简化了 GoldenDB 的部署和维护过程，减少了人工操作的需求，提高了运维效率。

综上所述，容器化部署提供了更高的灵活性和可靠性，同时可以确保安全性，GoldenDB 容器可以通过隔离机制确保应用程序和数据的安全性，避免了因为一个容器的故障或安全问题而影响整个系统。此外，容器化部署还可以实现快速的回滚和恢复，保证数据库服务的高可用性和业务连续性。

3.3 可靠性方案设计

GoldenDB 支持单数据中心高可用、同城双中心、两地三中心、多地多中心等容灾架构模式。GoldenDB 基于多副本冗余技术、一致性复制技术和灵活高可用策略实现多种容灾方案，单台服务器故障自动切换到本数据中心其他机器的副本上；双中心以上的机房故障时，可快速切换到其他机房或者城市灾难中心，最大程度保证业务连续性。

3.3.1 单管理节点

单个 AZ 且管理节点数量 1。这种部署模式有着简单的组网结构，适合一些小规模或者测试环境的应用。优点：组网简单，只需要在单个可用区中部署一个管理节点，即可满足基本的数据库需求。缺点：由于只有一个管理节点，一旦该节点发生故障，服务将会中断，无法提供连续的服务。

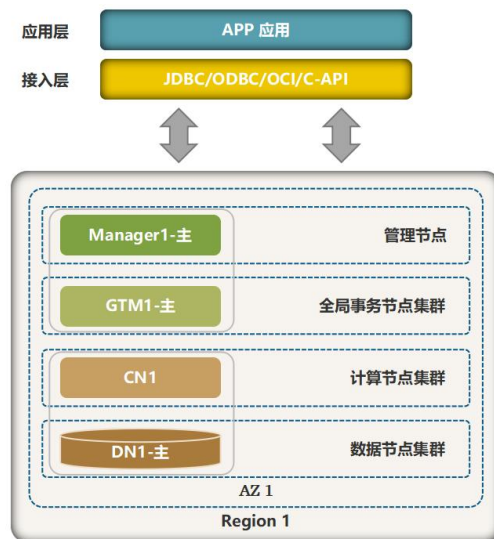


图 3-6 单管理节点组网图

3.3.2 本地单中心高可用

单个 AZ 且管理节点数量大于 1。这种部署模式可以提供更高的可用性，适用于对高可用性有一定要求的应用场景。优点：可以实现高可用性，通过在

同一个可用区内部署多个管理节点，当其中一个节点发生故障时，其他节点可以接管服务，保证数据库的连续性和可用性。这样可以减少单点故障带来的影响，提高应用的可靠性。缺点：由于所有的管理节点都部署在同一个可用区内，一旦整个可用区发生故障，所有的管理节点都会受到影响，导致服务无法提供。这样就会对应用的可用性产生风险，特别是在面临自然灾害或者硬件故障等情况下。

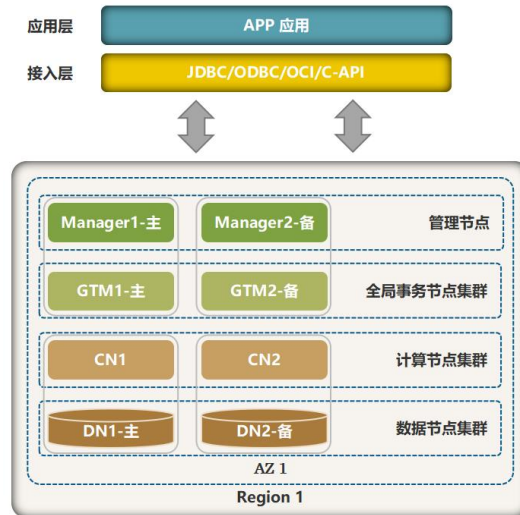


图 3-7 本地单中心高可用组网图

3.3.3 本地同城双中心高可用

双 AZ 且管理节点数量大于 1，分布在同一城市。这种部署模式可以提供更高的可用性，适用于对高可用性有较高要求的应用场景。优点：可以实现更高的可用性。通过将管理节点分布在不同的可用区中，当其中一个可用区发生故障时，其他可用区的节点可以接管服务，保证数据库的连续性和可用性。这样可以有效降低单点故障的风险，提高应用的可靠性。缺点：由于所有的管理节点都部署在同一城市内，一旦整个城市的可用区由于自然灾害或者硬件发生故障，所有的管理节点都会受到影响，导致服务无法提供。

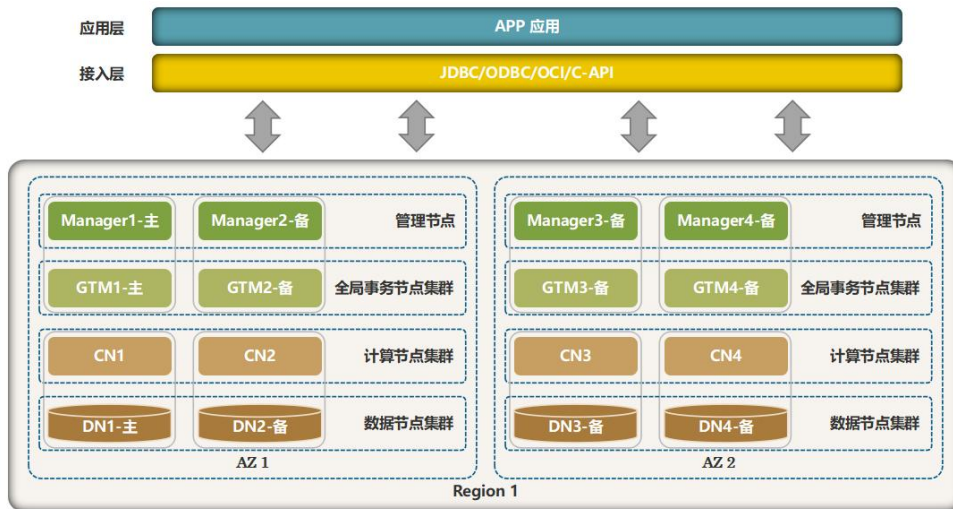


图 3-8 本地同城双中心高可用组网图

3.3.4 两地三中心高可用

3 个 AZ 且管理节点数量大于 1，分布在两个不同的城市。这种部署方式具有高可用性的优点，即使其中一个城市或可用区发生故障，系统仍然可以正常运行。然而，这种部署方式需要更多的资源，包括服务器和网络设备等，以支持多个节点的运行。因此，需要仔细考虑资源投入和成本效益之间的平衡。同时，还需要确保两个城市之间的网络连接稳定和可靠，以确保数据同步和节点之间的通信正常运行。

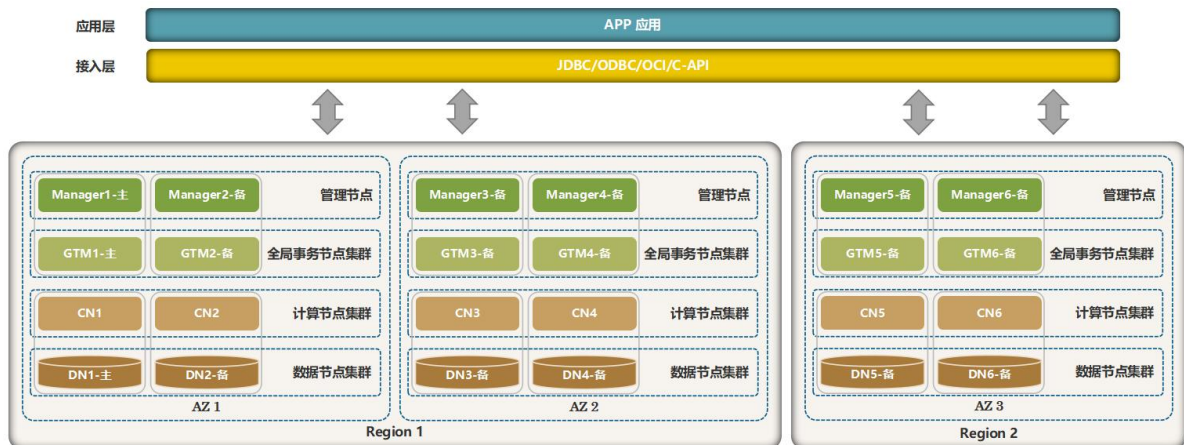


图 3-9 两地三中心高可用组网图

3.3.5 两地四中心高可用

4 个 AZ 且管理节点数量大于 1，分布在两个不同的城市。可以提供高可用性，即使一个城市或可用区发生故障，仍可以确保数据库的正常运行。然而，这种部署方式需要更多的资源，因为需要维护多个管理节点，并且需要在不同

的城市之间进行数据同步和通信。尽管有这些缺点，但高可用性的好处通常是值得的，特别是对于对数据库可用性要求较高的应用场景。



图 3-10 两地四中心高可用组网图

3.3.6 三地五中心高可用

这种部署意味着在三个不同的地理位置建立五个数据中心来确保数据库系统的高可用性。这种部署架构在一个地区或数据中心出现故障时，仍能保持数据库系统的正常运行，从而确保业务的连续性。每个数据中心都有多个冗余节点，包括管理节点、数据库节点以及存储节点。这种多重冗余性可以降低系统发生单点故障的风险。数据中心之间通过数据同步和复制机制实现数据的异地备份。这确保了即使某个地区或数据中心发生灾难，仍能够恢复数据并保持业务的连续性。尽管需要更多的资源来维护多个数据中心和节点，但通过合理的资源管理和负载均衡策略，可以最大限度地利用资源，并确保系统的性能和可用性。

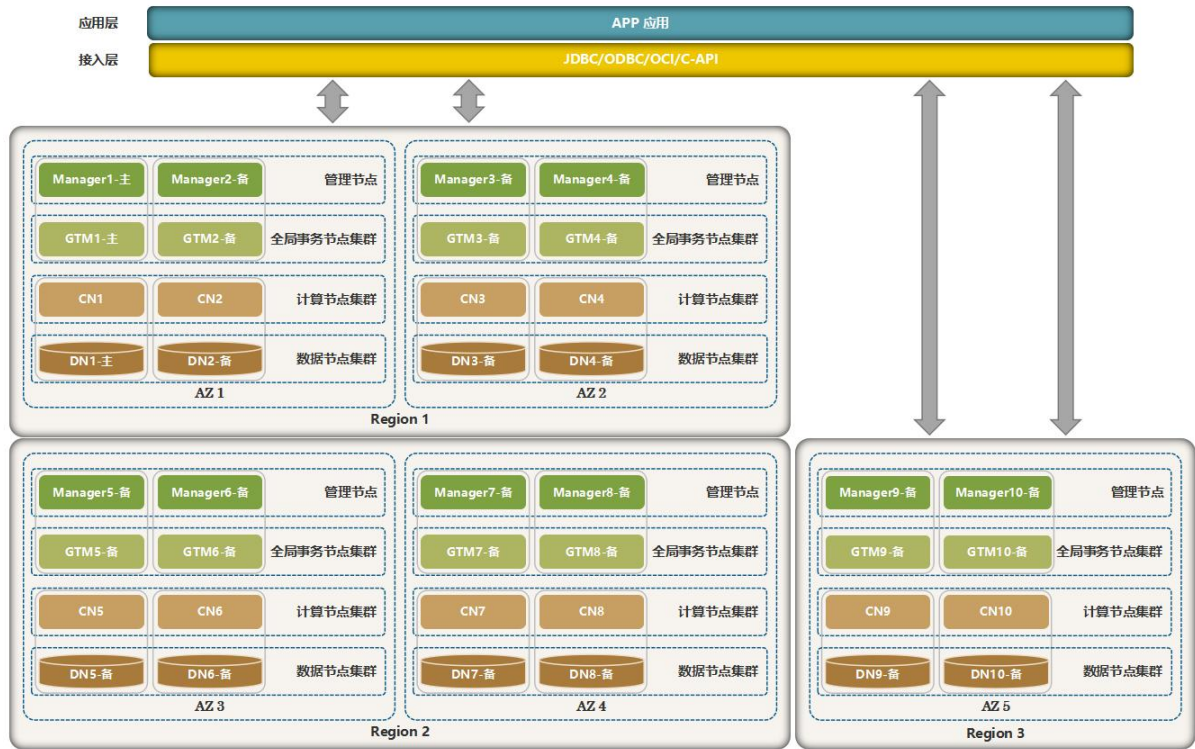


图 3-11 三地五中心高可用组网图

4 功能介绍

GoldenDB 是一款适用于 OLTP 场景，同时兼顾 OLAP 的关系型数据库。

GoldenDB 兼容 SQL92、SQL99、SQL2003 标准，同时兼容常用 Oracle、DB2、MySQL 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、DQL 数据查询语言、DCL 数据控制语言、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。

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，统一运维管理方便用户进行数据库集群的操作管理。

4.1 分布式事务控制

GoldenDB 作为事务型数据库，支持事务的 ACID 特性。

4.1.1 隔离级别

GoldenDB 支持读未提交（Read Uncommitted）、读已提交（Read Committed）、可重复读（Repeatable Read）、串行化（Serializable）的隔离级别，推荐使用 Read Committed 的隔离级别。

4.1.2 分布式事务控制

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 实现了强一致性解决方案，即当用户完成数据更新后，后续所有读操作都能且只能读到更新以后的值。

GoldenDB 分布式事务的核心思想是引入全局事务管理器，记录当前所有正在执行的全局事务标识及其状态，作为分布式事务的全局标识，通过全局事务机制实现全局事务的隔离性，并协助实现全局事务的原子性。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记录在全局事务管理器中。

- 全局事务管理器为每一个分布式写事务分配一个全局唯一的全局事务 ID，并根据事务的存活情况维护对应 ID 的生命周期。
- 全局事务 ID 加到每条数据记录中，更新数据时同时更新全局事务 ID，将其作为分布式环境下表的全局锁。

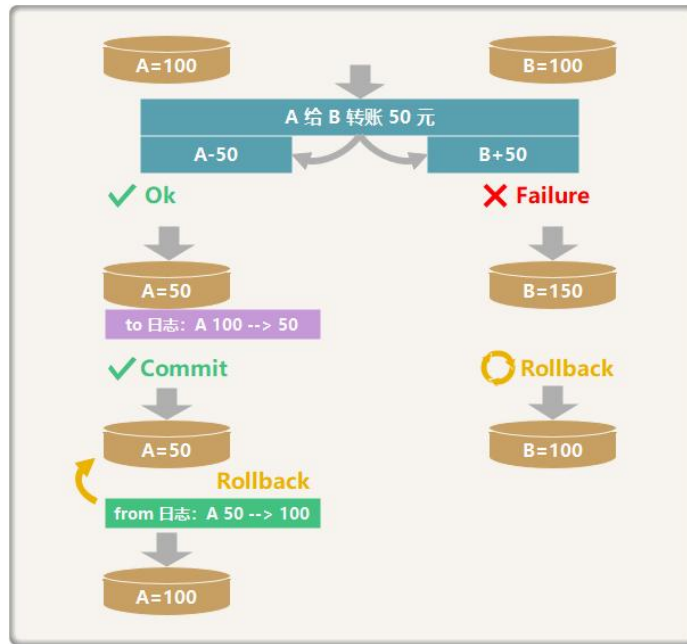


图 4-1 转账交易示意图

通过已提交事务回滚，解决分布式一致性问题。以转账交易为例，如图 4-1，假设 A、B 账户原来各自拥有 100 元，A 给 B 转账 50 元。A、B 属于不同节点，各自执行对应更新事务，A 执行成功，B 执行失败。B 失败后，进入单机事务异常回滚流程，A 成功后，进入已提交事务回滚流程，保证了分布式事务的一致性。

另外，为了提升效率，减少计算节点和全局事务管理器交互次数，将申请单个全局事务 ID 优化为批量申请全局事务 ID，释放单个全局事务 ID 优化为批量释放全局事务 ID。

为了避免全局事务管理器异常导致数据库不可用，全局事务管理器采用主备架构，主出现异常时，自动切换到备上继续提供服务。

为了降低全局事务管理器异常影响到数据库所有租户的风险，全局事务管理器扩展为全局事务管理器集群架构，多个全局事务管理器和多个租户绑定，某个全局事务管理器出现异常后发生切换，那么切换过程中影响的也仅仅是故障全局事务管理器绑定的租户，其它租户依旧可以正常运行。

4.2 基础功能

GoldenDB 支持通用数据库标准协议，支持 JDBC、ODBC、OCI、CAPI 等常见开发语言所使用的接口驱动，包括 C、C++、Java、.Net、Python 等。

GoldenDB 兼容 SQL92、SQL99、SQL2003 标准语法，兼容 MySQL 语法，兼容常用 Oracle、DB2 语法，支持视图、函数、索引、存储过程、触发器、分区等功能，帮助用户从 Oracle、DB2、MySQL 等数据库快速地迁移到 GoldenDB 数据库，节省应用改造的成本，提升迁移效率。

4.2.1 数据类型

GoldenDB 支持日常应用所能涉及的各种数据类型，可以满足企业数据管理的需要，部分数据类型如下：

- 数值类型：BOOL、BOOLEAN、BIT、TINYINT、SMALLINT、MEDIUMINT、INT、INTEGER、BIGINT、DECIMAL、DEC、NUMERIC、FLOAT、DOUBLE、DOUBLE PRECISION。
- 字符类型：CHAR、VARCHAR、BINARY、VARBINARY、TINYBLOB、TINYTEXT、BLOB、TEXT、MEDIUMTEXT、LONGBLOB、LONGTEXT、ENUM、SET。
- 日期类型：DATE、DATETIME、TIMESTAMP、TIME、YEAR。
- JSON 数据类型和空间数据类型。

4.2.2 函数功能

GoldenDB 支持的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、转换函数等，部分函数如下：

- 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、CHAR_LENGTH、COALESCE、CONCAT、CONVERT、IN、INSTR、LENGTH、LOWER、REPLACE、LTRIM、RTRIM、TRIM、UPPER、LEFT、RIGHT 等。
- 数学函数包括 ABS、CEIL/CEILING、FLOOR、ROUND、MOD、POW 等。

- 日期和时间函数包括 DATE、DATE_ADD、DATE_FORMAT、DATE_SUB、DATEDIFF、DAY、DAYNAME、DAYOFMONTH、DAYOFWEEK、DAYOFTYEAR、CURDATE、CURRENT_DATE、CURRENT_TIMESTAMP、CURTIME、LAST_DAY、MINUTE、MONTH、NOW、TIME、TIMEDIFF、TIMESTAMPADD、TIMESTAMPDIFF、TO_DAYS、TO_SECONDS、WEEK、YEAR 等。
- 转换函数包括 CAST、CONVERT 等。
- 聚合函数包括 AVG、COUNT、MIN、MAX、SUM 等。
- 位运算函数包括 &（按位与）、|（按位或）、！（按位非）、^（按位异或）、<<（按位左移）、>>（按位右移）等。

4.2.3 索引机制

GoldenDB 支持多种索引类型，包括唯一索引、复合索引等，通过使用索引可以快速定位数据，提高查询性能。GoldenDB 支持全局索引，在分布式数据库中，要在多个分片间实现索引功能，需要建立分布式全局索引。全局索引能够快速定位记录所在的分片信息，根据分片信息下发查询语句到正确的分片上，然后利用单数据节点索引进行数据快速查询。

4.2.4 存储过程

GoldenDB 支持存储过程。存储过程是在数据库中定义一些 SQL 语句的集合，存储过程既可以有返回值，也可以没有返回值，主要用于执行操作。GoldenDB 的存储过程是在整个集群上执行，存储过程定义和解析放在计算节点上，由计算节点统筹判断存储过程体中的语句是否跨数据分片，是否需要在计算节点处理数据，是否可以将语句下压到存储节点执行。

4.2.5 触发器功能

GoldenDB 支持触发器功能。

- 触发时间支持操作前触发（BEFORE）与操作后（AFTER）触发。
- 触发事件支持表的插入（INSERT）触发，表的更新（UPDATE）触发，表的删除（DELETE）触发等。

4.2.6 视图功能

GoldenDB 支持视图功能，视图保存复杂查询，简化用户操作，支持单表和多表视图。GoldenDB 支持创建视图、删除视图、修改视图等视图基本操作。

GoldenDB 支持物化视图功能，物化视图是数据库中的一种特殊视图。普通视图仅保存查询语句，物化视图将查询的结果存储在数据库中。物化视图可以显著提高复杂查询的性能，避免了每次查询时需要重新计算结果。

4.2.7 分区功能

GoldenDB 支持在分片的基础上再进行分区操作，支持创建分区、删除分区等操作。分区表按照哈希算法进行分区，将数据均匀地分布在多个分区中，从而实现并行查询，查询操作只需要处理特定的分区，从而减少扫描的数据量，提高查询效率。分区表也可以按照时间范围进行分区，每个分区包含特定时间段内的数据，这使得按时间范围查询变得更高效率。

4.2.8 执行计划

GoldenDB 支持查看数据查询、数据插入、数据更新和数据删除等操作的执行计划，并且可以根据不同的需要输出不同的内容，可以仅输出计算节点的执行计划，也可以仅输出数据节点的执行计划、还可以同时输出计算节点和数据节点的执行计划等。通过查看执行计划可以了解 SQL 的执行顺序、执行瓶颈和进行辅助优化。

4.2.9 字符集

GoldenDB 当前支持 latin1、gbk、utf8、utf8mb4 和 gb18030-2022 等字符集。

4.2.10 国产化适配

GoldenDB 具备全栈国产化适配能力，具体已完成适配包括但不限于：

- CPU：鲲鹏、飞腾、海光、兆芯、龙芯、申威、珠峰等。
- 操作系统：麒麟、统信、新支点、凝思、普华、深度、方德、龙蜥等。

- 工具：英方、爱数、迪思杰、新炬网络、新数科技、鼎甲科技等。
- 中间件：宝兰德、东方通、中创、普元、金蝶等。
- 业务：高阳、神州数码、长亮、宇信、赞同、同方软银等。

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：

- 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或增加安全组数目，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；
 - 增加分片。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，扩展组建新的子集群。可以在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。

GoldenDB 数据重分布根据不同的重分布场景会生成不同的重分布执行计划；比如在数据分布策略发生变化时，如范围表变为哈希表时，GoldenDB 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对比，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 Insight 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服务端口，设置不同的读写分离模式，包括以下几种：

- 本区域/同区域查询

读操作在本区域和同区域的主、备机间根据配置的权重做读负载均衡；

- 异区域查询

读操作仅在异区域备机间做读负载均衡；

- AZ 级查询

读操作在区域间根据配置的权重做读负载均衡；

- 最新副本

读操作在指定机房的最新副本间做读负载均衡。

除了上述服务端口级别的读写分离模式设置，GoldenDB 还支持：

- SQL 级别的读写分离

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 级别的优先级高于服务端口级别。

- 事务级读写分离

只读事务内的第一条读语句根据读写分离配置的规则选取备机，后续读语句绑定该备机执行；

- 会话级读写分离

当前会话设置为只读后，事务内只支持读语句，并根据事务级读写分离规则做读负载均衡。

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。

结合上述服务端口和 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，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，即仅将满足条件的记录导入至系统中。

- 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。

- 支持对原始文件记录末尾的空值进行处理。

4.3.5 Oracle 兼容性

GoldenDB 兼容常用 ORACLE 语法，大大降低业务从 Oracle 迁移至 GoldenDB 的工作量。

- 数据对象兼容
 - 支持序列 SEQUENCE 的创建、删除、获取；
 - 支持 DBLINK 的创建、删除、使用；
 - 支持 BEFORE、AFTER、INSTEAD-OF 等触发器的创建、删除、触发；
 - 支持包 Packages、自定义函数、存储过程的创建、删除、调用等；
 - 支持 REF Data Types、VARRAYS 等自定义类型的创建、删除、使用；
- SQL 语法兼容
 - 支持 ROWID、ROWNUM、DUAL 等；
 - 支持 FULL JOIN、(+) JOIN、MINUS、START WITH、MERGE INTO 等；
 - 支持 FLASHBACK BEFORE DROP、FLASHBACK AS OF TIMESTAMP 等；
- 内置函数兼容
 - 支持字符串函数，包括：LPAD、RPAD、SUBSTR、TO_CHAR、INSTR、TRIM、LTRIM/RTRIM、||拼接等；
 - 支持数值函数，包括：ROUND、TRUNC 等；
 - 支持日期和时间函数，包括：SYSDATE、TRUNC、MONTHS_BETWEEN、ADD_MONTHS 等；
 - 支持转换函数，包括：TO_CHAR、TO_NUMBER、TO_DATE、TO_TIMESTAMP、TO_CLOB、ROWTOHEX 等；
 - 其他函数，包括 NVL、NVL2、DECODE 等；
- PL/SQL 兼容
 - 支持 %TYPE、%ROWTYPE 等数据类型；

- 支持 LOOP、GOTO、FOR LOOP、CASE WHEN... END CASE 等流程控制语法；
- 支持 FORALL、INTO、:=变量赋值等特有语法；
- 系统视图兼容
 - 支持 V\$DATABASE、V\$INSTANCE、V\$SESSION、V\$MYSTAT 等常用系统视图；
- 高级包兼容
 - 支持 DBMS_OUTPUT、DBMS_SQL、DBMS_RANDOM、DBMS_UTILITY、DBMS_JOB、DBMS_CRYPT、DBMS_STATS、DBMS_LOB 等常用高级包；

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，通过对表或表分区进行数据压缩来减少磁盘存储空间；普通表支持透明页压缩，分区表支持行格式压缩；压缩算法支持 Zlib 和 LZ4，压缩率在 40%到 90%之间。

备份压缩，是指对备份出来的数据进行压缩来减少空间占用，GoldenDB 支持 Quicklz 和 LZ4 算法对备份文件进行压缩，默认使用 Quicklz 压缩算法。

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

有些核心业务场景，需要能准确地将会计日当天的业务数据卸载出来，不受会计日切换的影响。使用传统数据库存在卸数丢失跨天事务、分析报告不准确的问题，需要事后进行手工修复数据。GoldenDB 提供日切卸数功能，可以卸载准确的会计日当日业务数据，更快得到正确的分析报告，无需在事后手工修复数据。以卸载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数据补录或清洗处理，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）。

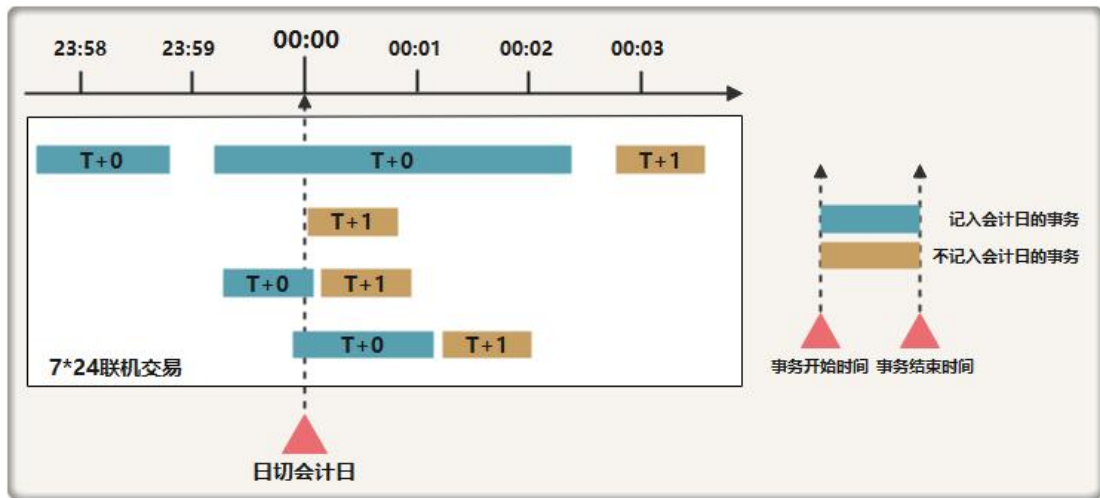


图 4-2 日切卸数示意图

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、批量协议、并行执行、流式查询等手段，对整体性能进行提升，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、多版本并行控制技术（multiversioning）、全局乐观锁+自动补偿机制。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系统可用性，为了数据的安全可靠，

数据必须在满足拥有一定数量的副本之后，才返回处理结果给客户端。GoldenDB 采用自研的 gSync 复制技术，在配置的副本同步策略满足后，主库返回操作结果给客户端；通过线程池、非阻塞式同步、并行复制等关键技术可实现在确保数据 RPO 为 0 的同时保证系统的吞吐量。

4.4.4 批量协议

针对 PREPARE 模式，GoldenDB 引入批量协议，支持在 PREPARE 时只需要携带一组 VALUE，而在 EXECUTE 时可以携带任意多组 VALUE。批量协议有效提高 PREPARE 模式使用灵活性和成功率，大大提升 PREPARE 模式下 GoldenDB 的处理性能。

4.4.5 并行执行

SQL 语句在 GoldenDB 计算节点到各个数据节点间、计算节点到单个数据节点内，都支持并行执行，且保证分布式事务一致性。这种并行执行模式，能有效提高计算节点执行性能。同时 GoldenDB 数据节点支持并行查询和并行 DDL，进一步提高对应场景下的处理性能。

4.4.6 流式查询

GoldenDB 支持流式查询，数据节点和计算节点之间、计算节点和客户端之间，都不需要等待结果集全部处理完成才回响应，而是按照设置分批回响应，且结果集不落盘。在大结果集查询场景下，全流程“execute-fetch”有效减少响应等待时间，整体性能大幅提升，并能降低系统 CPU、内存资源占用。

4.4.7 复杂查询并发

GoldenDB 自研 OLAP 引擎可以有效提升复杂规模查询计算效率，充分利用集群中多台服务器的 CPU 以及内存资源，对海量数据进行快速查询分析。计算节点统一接收业务请求，判断涉及 OLAP 类查询或者列存查询时，自动转发给 AP 节点，由其承担复杂查询和复杂分析的任务。自研 OLAP 引擎基于分布式执行架构，实现多机并行执行；使用更丰富的数据统计信息，应用维度更科学的代价模型，在处理复杂查询时生成更优的执行计划；采用列式存储，实现全面向量化，充分利用单机多核计算能力并发执行，全面提升查询速度。

4.4.8 大容量数据处理能力

GoldenDB 具备完备的海量数据管理功能，提供 PB 级的数据容量支持，

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。

4.4.9 分布式批处理

GoldenDB 提供分布式架构下批处理功能，满足大数据量批处理需求；支持流式处理和存储过程对数据进行批处理，减少客户端与数据库的访问次数，批量返回数据集并进行批量处理。

4.5 安全性

GoldenDB 作为一款安全、自主、可控的数据库产品，在安全性方面取得了显著成就，并得到了众多权威机构的广泛认证与认可。

全面的安全性保障

- **安全审计：**GoldenDB 支持界面化的安全审计功能，能够详尽记录并检测数据库管理操作及用户数据库对象操作，为数据库的安全监控提供了坚实保障。
- **密码安全：**GoldenDB 具备密钥管理与标准密码运算能力，支持符合国家商用密码标准的加密算法，确保密码的安全可靠。
- **数据安全：**GoldenDB 通过精细的授权管理员角色、数据脱敏查询、数据传输加密、数据加密存储、数据运行安全等功能，以及强大的备份和恢复机制，全面确保数据的完整性和可恢复性。
- **运维安全：**GoldenDB 通过适合分布式架构的一致性元数据管理能力、异常问题诊断能力、监控告警能力和数据一致性管理能力等，确保 GoldenDB 的运维安全，全面保障数据库的稳定可靠。

权威机构认证

- **安全可靠测评：**于 2024 年 9 月 30 日，首批通过由中国信息安全测评中心、国家保密科技测评中心、国家知识产权组织的分布式数据库的安全可靠测评，充分证明了 GoldenDB 在安全可靠方面的卓越表现，为用户提供了有力的安全保障。
- **信息技术产品安全分级评估 EAL4+ 认证：**由公安部第三研究所于 2023 年 7 月 9 日颁发，证明了 GoldenDB 在信息安全方面的卓越能力和高标准。
- **GM/T0028 商用密码产品认证证书：**由国家密码管理局商用密码检测中

心于 2024 年 1 月 12 日颁发，证明了 GoldenDB 在商用密码产品方面的合规性和安全性。

- **关系型数据库安全专项证书：**由中国信通院于 2024 年 7 月 16 日颁发，证明了 GoldenDB 在关系型数据库安全能力方面均符合标准要求。
- **TLCP 测试软件检测报告：**由鼎铉商用密码测评技术有限公司于 2023 年 8 月 10 日出具，进一步验证了 GoldenDB 密码技术的先进性和可靠性。
- **网络安全等级保护核心系统等级测评报告（4 级）：**GoldenDB 承接的项目由中国金融电子化公司测评中心于 2022 年 7 月 28 日出具，证明了 GoldenDB 核心系统在网络安全方面具有较高的防护能力。

通过一系列先进的安全技术和措施，如身份认证、权限管理、数据加密、审计日志、数据备份与恢复等，GoldenDB 全方位地保障了数据的安全性和完整性。同时，GoldenDB 还具备灵活扩展、高性能、高可用等特点，能够满足各种复杂业务场景的需求，为用户提供安全、可靠、高效的数据库服务。

4.5.1 安全审计

GoldenDB 支持界面化的安全审计功能，包括配置、查阅、备份、导出功能。该功能支持启停管理，并具备对数据库管理操作、用户数据库对象操作等的检测能力，同时能创建与用户关联的安全事件记录。还具备清理审计数据的能力，确保审计信息的时效性和系统性能，为数据库的安全监控提供全面保障。通过提供详细的审计日志和报告，帮助组织迅速响应风险事件，追溯问题根源，并优化数据访问控制策略。

- 数据库管理操作

GoldenDB 的设计支持完整的安全管理功能，包括用户权限控制，审计等。并且审计事件中包括主体身份，可以通过审计事件追溯到该事件对应的数据库用户。

- SQL 日志审计

审计是对数据库用户执行操作的记录和审核。GoldenDB 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。

- 审计数据可用性保证

GoldenDB 支持审计记录的查看、导出操作，但是必须具备对应权限的管理员或用户执行，无授权的用户无法查看或导出这些审计记录。审计数据具备加密存储的保护能力。

- 审计数据清理

GoldenDB 支持按策略周期清理最早的审计记录。

4.5.2 密码安全

GoldenDB 提供了密钥管理和密码运算功能的调用机制，以维护数据库管理系统中数据资产在存储和传输过程中的加密保护需求。GoldenDB 具备密钥管理与标准密码运算能力。

- 密钥生成

GoldenDB 的密钥由安全管理员通过管理页面进行初始化和更新，使用符合 GM/T0028 标准的密钥生成算法，生成 CN 和 DN 工作密钥。

- 密钥变更

安全管理员通过管理界面进行密钥的更新，备份与销毁。

- 密码运算

GoldenDB 根据国家商用密码产品二级以上的特定的密码算法和密钥长度来执行密码运算列表，支持 SM2/SM3/SM4 加密算法。

- 密码验证

GoldenDB 支持通过国密算法对用户名密码进行解密校验。数据库设置了密码校验规则，包括长度、密码不能与用户名相同、英文字母大小写、特殊字符及数字等组合限制，如果设置的密码不符合密码校验规则数据库会失败返回。

4.5.3 数据安全

GoldenDB 支持用户数据保护的能力，不仅支持精细的授权管理员角色，实现职责分离和角色约束，确保只有被授权的用户才能访问和操作数据库对象，而且支持数据库中的表列进行脱敏查询和数据传输加密功能，无论是与外部应用之间的交互还是数据库内部的数据传输，都能得到高级别的加密保护。此外，GoldenDB 提供事务处理和存储介质故障的数据恢复机制，以保障数据的完整性和可恢复性，而且支持数据副本间的准实时一致性，确保数据的实时同步，并在数据不一致时及时发出告警，为用户提供全面、可靠的数据保护服务。

4.5.3.1 访问控制

- 基于角色的访问控制

GoldenDB 中有针对不同权限级别的各类系统表。GoldenDB 对不同的角

色赋予或撤销不同的操作权限,有些角色赋予所有权限,有些角色只赋予读权限,有些角色只赋予写权限。角色赋予的权限会保存在对应的权限系统表中并持久化。创建用户后,将用户设置为一种角色,用户就拥有了该角色所对应的权限。

- 身份鉴别标识

用户登录数据库支持身份鉴别,身份鉴别标识包含用户名、密码等信息,在登录过程中需要对关键信息进行加密传输。

GoldenDB 通过'用户名'作为唯一标识。当用户登录后,根据唯一标识,从全部的权限系统表中加载该用户的全部权限。当执行某个 SQL 时,会根据执行的 SQL 语句类型、访问的库表等来汇总需要的全部权限,然后将该需要的权限集合与用户所拥有的权限做比较,如果当前用户所拥有的权限满足需要的权限,则允许 SQL 执行,否则终止 SQL 执行并返回报错。通过设置不同的权限实现了精细化的权限访问控制。

- 会话状态控制

GoldenDB 通过客户端与 CN 端口建立会话,GoldenDB 必须支持会话状态管理、SQL 语句监控和活跃事务数监测,以及最大会话数控制,确保数据库操作的合规性。同时,GoldenDB 必须支持多种会话终止方式、会话的锁定/解锁功能,并支持存储会话信息以追溯会话历史,提供全面的会话管理功能。这些措施共同保障 GoldenDB 会话管理的有效性,增强数据库访问的安全性和可管理性。

- 白名单客户端控制

GoldenDB 还支持白名单控制,通过 IP 白名单的方式过滤连接请求,只有在 IP 白名单中的客户端才可以访问,非白名单客户端连接报错。

4.5.3.2 数据查询脱敏

GoldenDB 支持对数据库中的表列进行脱敏查询,例如:模糊化信用卡号或身份证号中的某些数字。

4.5.3.3 数据传输保护

GoldenDB 内部传输支持 SSL 加密。SSL 是在传输通信协议(TCP/IP)上实现的一种安全协议,采用公开密钥技术,有效保护数据的传输安全。GoldenDB 内部传输使用 SSL 加密。原端将原始 SQL 语句进行 SSL 加密处理,然后将密文发送给目的端,目的端收到密文后进行解密,得到原始 SQL 语句进行正常解析和处理。此外 GoldenDB 也支持客户端使用 TLCP 加密的

链路与 CN 交互。

4.5.3.4 数据加密存储

GoldenDB 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。

- 列级加密

对数据库中的某一列数据进行加密，这种加密方式具有较高的灵活性和选择性，可以根据实际需求选择需要加密的列。列加密支持加密算法包括 AES256 和 SM4。

- 表级加密

对整个数据表进行加密，这种加密方式适用于对整个表的数据进行保护。表加密支持的加密算法包括 AES256 和 SM4。

- 整个数据库级加密

对整个数据库进行加密，包括数据库中的所有表、视图、索引等对象。这种加密方式提供了最高级别的安全性，但也可能对性能产生较大影响。库加密实现上依赖于表加密。

4.5.3.5 数据运行安全

GoldenDB 支持数据运行安全，致力于为用户提供全方位的安全保障。GoldenDB 具备拦截大事务、拦截大结果集查询和预防误操作等多种机制来预防潜在的危险操作，确保数据在系统运行过程中的稳定性和可靠性。

- 拦截大事务操作

GoldenDB 能够智能地拦截大事务操作。在数据库环境中，大事务往往会对系统性能产生严重影响，甚至可能导致系统崩溃。通过拦截大事务，GoldenDB 能够避免这类操作对系统造成的冲击，确保系统能够持续稳定运行。

- 拦截大结果集查询

GoldenDB 还能够拦截大结果集查询。在某些情况下，用户可能会执行返回大量数据的查询操作，这不仅会消耗大量系统资源，还可能导致系统响应时间变长。通过拦截大结果集查询，GoldenDB 能够提醒用户注意查询操作的合理性，并引导用户采取适当的措施来优化查询性能。

- 预防误操作

GoldenDB 还具备预防误操作导致系统异常的能力。GoldenDB 通过预操作发现并避免可能导致系统异常的误操作。这一特性不仅提高了系统的安全性，还为用户提供了更加可靠的数据使用环境。

4.5.3.6 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：可以在需要时立即发起备份任务

GoldenDB 备份的功能特点：

- 备份任务管理

GoldenDB 支持在 Insight 上进行备份任务的管理，包括备份任务策略的配置、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，如 IBM TSM、NBU、S3、COS 等，可以直接将备份结果存储到第三方存储介质。

- 备份文件存储策略

GoldenDB 支持实例级存储配置、分片级存储配置、AZ 级存储配置、节点级存储配置等备份文件的存储策略，可以灵活应对局点的不同业务需求和网络隔离场景进行选择 and 配置。

GoldenDB 数据恢复功能的特点：

- 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快于数据库日志的回放速度，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快照：

- 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及要恢复到的时间点，即可进行自动化的数据恢复。

4.5.4 运维安全

GoldenDB 在运维安全领域提供了全面、高效且智能的解决方案，旨在为用户打造稳定、可靠且高效的数据库服务体验。

4.5.4.1 元数据管理能力

GoldenDB 支持元数据统一管理，并具备满足一致性要求的分布式数据库元数据管理能力。它不仅全面涵盖了元数据的创建、修改、删除、查询、可视化以及持久化等核心管理环节，还提供了对结构信息和分片信息等复杂元数据的管理功能。此外，GoldenDB 还具备对组件元数据的统一管理能力，能够修复各组件元数据不一致问题，确保各组件元数据的完整性和一致性。这一特性使得用户能够更加方便地进行元数据管理和维护。

4.5.4.2 异常问题快速诊断

GoldenDB 支持异常问题快速诊断功能。除了常规的集群服务器、数据库实例、慢 SQL、锁等待、死锁等问题诊断分析外，GoldenDB 还增加了对分布式事务、跨节点数据一致性等复杂问题的诊断能力。通过异常诊断的能力，GoldenDB 能够快速定位问题根源，提供准确的解决方案建议，有效缩短故障恢复时间。

4.5.4.3 监控管理告警能力

GoldenDB 支持一键向导式部署，使得部署过程更加简便快捷。图形化管理平台提供了丰富的监控信息，包括系统概况、运行情况、关键指标、集群情

况、吞吐情况等，同时还支持自定义监控指标，满足用户个性化需求。当系统出现问题时，GoldenDB 能够实时发出告警，并支持快速跟踪和定位问题，确保系统稳定运行。

4.5.4.4 一键巡检及故障诊断

GoldenDB 支持一键巡检及故障诊断功能，能够按需收集日志及诊断信息，并进行智能分析和处理。支持日志打包及切片功能，有效的提升了问题诊断效率，同时 GoldenDB 还提供了丰富的诊断报告能力，包括按指定时间生成集群和数据库诊断报告等，为用户提供了全面的系统健康状况分析。

4.5.4.5 主备节点数据一致性检查

GoldenDB 支持主备节点数据一致性检查功能，能够可视化实现数据库存储节点副本间主备数据的一致性检查。通过智能高效的数据处理能力，GoldenDB 能够实时监测主备节点数据的一致性状态，并在发现不一致情况时及时发出告警，确保数据的完整性和可靠性。同时，GoldenDB 还提供了丰富的数据修复工具和方法，帮助用户快速解决数据一致性问题。

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，提供了全面的运维管理功能，以满足不同场景下的数据库管理需求。

GoldenDB 提供 Insight 统一运维管理平台，提供 WEB 界面形式的统一管理入口，支撑各种场景的运维需求，简化用户对集群的管理，满足不同的运维管理需求。Insight 同时提供丰富的 API 接口，供其他系统集成，满足客户构建自身统一运维平台的需求。

4.6.1 安装与升级

GoldenDB 支持如下安装与升级功能：

- Insight 提供资源管理功能，将物理机、虚拟机纳入资源池统一管理和监控；物理机安装模式，从物理机或虚拟机资源池分配资源进行安装。对于云化，由云化管理平台纳管服务器资源分配容器进行安装，安装后由 Insight 纳管容器租户。
- 提供多样化的安装形式，包括：界面引导式白屏安装、一键式黑屏安装、API 接口安装，具备自动推荐安装参数快速安装，也支持客户自定义安装

参数，支持多种场景的安装需求；支持从一个节点往多个节点部署的能力。

- 容灾安装模式，支持租户跨多地多可用区的灾备部署能力，如两地三中心、同城多中心。
- 支持 CPU 架构：X86 与 ARM。
- 支持多种主流操作系统，如 Redhat、麒麟。
- 支持命令行、界面化启停操作。
- 安装过程记录日志，可从界面查看安装日志。
- 支持一键式命令升级与界面引导式升级，支持任务编排，支持不停服务升级，保证升级期间的业务连续性。
- 提供扩缩容 API 接口，支持其他系统通过标准的 API 接口使用 Insight 的资源，完成扩缩容操作。

4.6.2 数据配置

GoldenDB 支持如下配置功能：

- 提供参数配置功能。支持查看、变更组件参数配置，具备对单个节点或同类组件多节点批量变更的能力；支持参数动态生效与静态生效。
- 提供参数比对功能。同类组件的多个节点间可以进行参数比对，发现差异。可以实时、定时进行比对，提供比对报告。
- 提供参数模板功能。参数模板即一组预先设定好的参数，在安装时进行应用，安装完成的参数就是预设值。
- 提供参数变更历史查询功能。
- 提供存储配置功能。可以配置数据存储目录、日志目录、备份目录。
- 提供资源规格配置功能。

4.6.3 运维

GoldenDB 支持如下运维功能：

- 支持统计监控功能。包括统计集群节点的 CPU、内存、磁盘、网络的使用情况。支持慢 SQL 统计，如语句类型、用户名、数据库名、次数、总耗时、平均耗时等。支持性能状态统计，如每秒事务数和查询数、SQL 平均响应时间，高频 SQL 等。Insight 平台内置丰富的监控报表，支持自

定义监控报表。提供 API 拉取接口，支持主动上报 Kafka、Prometheus。

- 支持告警功能。包括支持事件告警和阈值告警。数据库运行中遇到重要事件、异常事件会产生事件告警；对于监控指标，设置告警基线，超过基线产生阈值告警。支持告警信息的实时展示与历史告警查询。告警上送支持多种形式，包括：邮件、REST 接口、Kafka、ActiveMQ、Syslog、SNMP。
- 支持实时监测 SQL 执行过程中资源使用情况；提供查询计划的缓存管理功能；提供 SQL 改写的优化建议。支持在线会话的分析诊断功能。支持 SQL 限流功能。支持历史 DML 操作在线查询功能。支持超长 SQL 与超大事务的自动查杀功能。
- 支持对各类事件如配置修改、表结构变更等进行日志记录的功能，支持多种维度的日志，如错误日志、慢日志、general 日志、DDL 日志、审计日志等，可通过日志查看操作内容、执行过程和结果。Insight 运维平台记录完整的操作日志，可用于审计。
- 支持基于角色、组、用户的权限管理模型，支持权限隔离，保障资源的使用安全。
- 提供丰富的接口供局点系统调用，包括 REST 接口、SQL 接口、命令行接口，可以查询状态、统计等多种信息，可以实现管理操作，也可以支持问题分析诊断。
- 系统具备高可用能力，自动检测故障，故障主动自愈修复。
- 支持巡检诊断功能，提供诊断报告。

4.7 GoldenDB 工具集

GoldenDB 的工具集极为丰富，专为支持其高效运行与管理而设计，全面覆盖了数据库迁移、测试、同步、诊断及优化等多个关键领域。其中，CACTool（Capture Analyze Convert Tool，简称：CAC）迁移评估工具，为业务迁移提供关键决策、改造工作量预估以及整体迁移策略的参考；Replay 回放工具则能够解析并回放生产流量，助力企业在不中断业务的前提下进行性能测试与故障排查；Sloth 迁移同步工具则实现了 GoldenDB 与其他数据库之间的全量及增量数据迁移与同步，确保了数据的准确性与完整性；而 InsightTool 巡检诊断工具，则以其快速、准确的特点，为数据库系统的故障诊断提供了强有力的支持。这些工具共同构成了 GoldenDB 强大的工具集，为企业提供了全方位、一站式的数据库管理解决方案。

4.7.1 CACTool 迁移评估工具

CACTool 迁移评估工具是 GoldenDB 数据库系统中一款重要的工具，旨在帮助用户在数据库迁移前评估业务迁移的可行性，并为迁移计划、工作量评估等提供关键支持。以下是对 CAC 的详细介绍：

- **CAC 功能概述**

CAC 主要用于对业务迁移进行深入的可行性评估，具备信息采集、对象转换、分片键推荐、语法兼容性分析和生成迁移评估报告等功能。其目标是帮助用户全面了解从当前数据库（如 Oracle、DB2、MySQL）迁移到目标数据库 GoldenDB 的可行性和潜在的不兼容项。通过自动整改不兼容的语法，CAC 能够帮助定制迁移策略、简化迁移方案，提高迁移效率。

1. 语法兼容性评估：CAC 能够识别并报告出当前数据库中与目标数据库不兼容的语法元素。CAC 支持对表结构、序列、函数、存储过程、同义词以及 SQL 语句进行兼容性分析。通过对源库采集的 DDL、DML、DQL 等进行分析，CAC 工具可以自动进行 DDL 转换，并生成兼容性评估报告。

2. 数据量分析：CAC 能够统计和分析当前数据库系统的数据量，包括表和其他对象的数量等。

3. 迁移计划制定：基于评估结果，CAC 能够协助用户制定详细的迁移计划。

- **CAC 产品优势**

1. 提高迁移效率：通过自动整改不兼容的语法并生成评估结果协助用户制定详细的迁移计划，CAC 能够显著缩短迁移周期，提高迁移效率。

2. 降低迁移风险：通过全面的语法兼容性评估和数据量分析，CAC 能够帮助用户提前发现并解决潜在的问题，降低迁移过程中的风险和不确定性。

3. 优化资源分配：基于迁移计划，CAC 能够指导用户合理分配资源，确保迁移工作的顺利进行。

4. 易用性：提供图形化界面评估报告，降低理解难度并增强了用户体验。

- **CAC 使用场景**

在数据库迁移项目启动前，使用 CAC 分析源端数据库（如 Oracle、DB2、MySQL）的历史运行数据，并对目标数据库 GoldenDB 进行全面的评估，能够为迁移决策、改造工作量预估以及整体迁移策略的制定提供关键性的参考。

4.7.2 Sloth 数据迁移工具

数据库迁移是数据库管理中的一项重要任务，它涉及到将数据从一个数据库系统转移到另一个数据库系统的过程。这一过程可能由于多种原因发生，如技术升级、成本节约、性能提升或业务需求变化等。以下将详细介绍 GoldenDB 的数据库迁移工具 Sloth。

● Sloth 功能概述

Sloth 是一个集管控、运维、编排于一体的数据迁移工具，提供了多种数据处理能力，如存量数据的迁移能力、增量业务的实时同步能力、数据的快照比对能力、数据的实时同步比对能力、数据的修复能力等。用户可根据具体的诉求，按照操作指导，编排自己需要的任务，达成目的。还提供了任务监控、模块部署、表画像等能力。

1. 数据迁移：支持将源端数据库的存量业务数据和实时增量业务数据无缝迁移至目标端数据库。

2. 数据库同步：支持源端数据库指定位点之后产生的增量业务数据，实时同步到目标数据库。

3. 同步比对：支持在数据同步过程中对同步完成的数据进行源端和目标端数据库差异实时校验功能。

4. 快照比对：支持基于快照点的全量数据校验功能。

5. 数据修复：支持同步比对或者快照比对的差异数据进行数据修复，保证数据一致性。

6. 数据发布：支持增量数据按一定格式同步到第三方存储设备 Kafka。

7. 运维管理：Sloth 支持多种运维能力，包括但不限于权限管理、动态黑名单、全加增状态扭转、在线性能观测及调优等。

8. 监控功能：Sloth 支持多种监控能力，如采集、回放、比对组件状态监控等。基于状态机变更，Sloth 配套实现了系统级高可用、单模块高可用。

9. 部署功能：Sloth 支持图形化部署采集、回放、比对组件，可以在编排完任务时，即拿即用。

10. 表画像：Sloth 数据迁移及数据同步支持绘制表维度的数据画像，即表画像，在运维管理界面可以直观查看表维度的可视化描述。Sloth 表数据画像包含多个元素：表的基本信息、数据统计、采集、回放、比对组件的时间统计和时延统计、性能（流量、QPS）统计、类型/列/行过滤/HEX 转换等特殊配置、剔除某张表、关闭某张表的比对、一键关闭任务级三无表比对、暂停回放、暂停比对、绑定执行计划等运维操作。

11. 动态加减表：Sloth 支持增量同步数据过程中动态修改同步表的黑白名单。

12. 断点续传：Sloth 具备断点续传能力，数据迁移或数据同步过程中发生故障，导致数据传输中断，在故障恢复后，系统能够自动从故障点继续迁移或同步。

13. 对象迁移：Sloth 支持迁移源库中的数据对象定义(表、函数、存储过程、触发器、视图、同义词等)，Sloth 的对象迁移发生在业务数据迁移之前，通过对象迁移保证源端数据库和目标端数据库对象结构的一致(异构数据库字段类型等通过映射来保证一致)，确保业务数据能够从源端数据库中完整、准确地迁移到目标数据库中。

14. DDL 同步：Sloth 支持在增量业务数据同步过程中同步 DDL。

● Sloth 产品优势

Sloth 是专注于异构或同构数据库与 GoldenDB 进行数据交互的专业服务与解决方案，旨在助力客户实现同构或异构数据库向 GoldenDB 的实时数据迁移和数据同步。Sloth 迁移或同步解决方案，配置流程简单便捷、风险低、运维成本低、可靠性高、性能突出。

1. 可视化集中管理，操作便捷快速：Sloth 无需单独安装驱动程序或应用程序，无需对源数据库做大幅改动，只需在 Sloth 提供的可视化管理界面进行简单配置即可开启数据传输。Sloth 管理界面中，只需花费几分钟就能设置一个迁移任务。您可以在迁移任务中定义用来执行迁移的各项参数，其中包括设置源数据库和目标数据库的连接，以及迁移类型和对象。Sloth 支持上千个节点的集群规模，可同时运行并集中管理大量数据迁移同步任务。

2. 不停服同步，业务无感知：Sloth 的解决方案能够助力客户将数据库迁移至 GoldenDB，且几乎不需要停机拷贝。源数据库在迁移期间发生的所有数据更改都会复制到目标数据库，因此迁移期间不影响源数据库对外提供服务。在数据复制完成后，源数据库和目标数据库将保持继续同步，您可自由选择业务切换时间。

3. 高性能数据迁移：Sloth 使用高规格服务器来保证每条迁移同步链路都能拥有良好的传输性能。在数据迁移方面，数据传输服务底层采取了多种性能优化措施：相对于传统的数据迁移工具，极大提升了传输性能。

4. 支持多种数据源：Sloth 目前已支持 Oracle、DB2、SQLServer、MySQL、PostgreSQL、Kafka 等类型的数据源向 GoldenDB 数据库进行迁移或同步，但具体的功能因源数据库类型不同而有所区别。

5. 内置多种过滤器：Sloth 支持正则表达式过滤库表；支持配置库、表、

字段、SQL 表达式等多种方式在数据同步过程中对源数据事件进行过滤。

6. 故障自动恢复: Sloth 具备高可用功能, 支持迁移任务在不同节点自由调度, 少量工作节点宕机并不会影响进行中的任务。并且支持机房级故障自动修复的高可用功能。

7. 数据一致性校验: Sloth 支持多种数据一致性校验方式, 如快照比对、迁移/同步比对, 高效识别数据差异并保证数据迁移的质量。

8. 差异数据修复: Sloth 支持差异数据的修复功能, 保证数据迁移的一致性。

- **Sloth 使用场景**

1. 数据迁移和同步: 在数据库迁移项目中, 支持同构数据库(如 GoldenDB 到 GoldenDB, 包括不同的分片数量场景)、异构数据库(如 Oracle、MySQL 等)之间的全量数据、增量数据的迁移和同步。

2. 数据比对: 数据库基于快照点的全量校验和数据同步过程中的数据实时校验, 确保迁移数据的准确性和完整性。

3. 数据修复: 针对源端数据库和目标端数据库数据不一致的场景, 进行差异数据修复。

4. 数据发布: 针对特殊业务场景, 支持将数据同步发布给第三方 Kafka。

5. 对象迁移: 支持数据库对象迁移。

4.7.3 Replay 回放工具

Replay 回放工具是一款功能强大、易于使用的数据库测试工具, 能够帮助开发人员和测试人员高效地回放生产流量, 对 GoldenDB 进行全面的测试。通过使用该工具, 可以提高 GoldenDB 的稳定性和可靠性, 降低故障率和维护成本。以下是对 Replay 回放工具的详细介绍:

- **Replay 功能概述**

Replay 回放工具是 GoldenDB 工具集中的一部分, 主要用于将生产流量解析为可回放的格式, 并在测试环境中, 回放解析的生产流量, 以模拟真实用户行为, 从而对数据库系统进行全面的测试。

1. 流量录制: Replay 回放工具支持将生产流量(包括各种 SQL 语句、事务操作等)解析为可回放的格式。

2. 流量回放: 在测试环境中, 该工具可以回放之前录制的生产流量, 以模拟真实的用户行为, 对数据库系统进行压力测试、性能测试等。

3. 测试验证：通过回放流量，可以验证数据库系统在特定场景下的表现，包括响应时间、错误率等关键指标，从而评估系统的稳定性和可靠性。

- **Replay 产品优势**

1. **高效性：**Replay 回放工具能够高效地回放数据库流量，减少测试时间和成本。

2. **准确性：**由于回放的是真实的生产流量，因此测试结果更加准确可靠。

3. **易用性：**该工具提供丰富的集成选项，方便用户进行配置和使用。

4. **可扩展性：**Replay 回放工具支持 Oracle、GoldenDB 等数据库，可以根据实际源端数据库类型进行定制回放给目标端 GoldenDB。

- **Replay 使用场景**

1. **性能测试：**在版本更新或数据库迁移前，使用 Replay 回放工具对数据库系统进行性能测试，确保在同等数据量与并发压力下 GoldenDB 的稳定性和可靠性。

2. **故障排除：**当生产环境出现问题时，根据捕获的相关数据库流量并在隔离的环境中重放，以便开发人员安全地调试问题，定位并修复故障。

3. **回归测试：**在软件开发生命周期中，使用 Replay 回放工具进行回归测试，确保新版本的数据库系统不会引入旧版本中已经修复的错误。

4. **A/B 测试：**可以将流量发送到两个数据库系统（如新版本和旧版本、异构数据库迁移等），比较它们的表现，以便做出数据驱动的决策。

4.7.4 InsightTool 巡检诊断工具

InsightTool 巡检诊断工具是一款针对 GoldenDB 数据库系统设计的独立部署工具，它提供了全面的巡检、问题诊断以及体检报告式的诊断报告功能。以下是对 InsightTool 的详细介绍：

- **InsightTool 功能概述**

GoldenDB 提供独立部署的 InsightTool 巡检诊断工具，支持巡检、问题诊断，提供体检报告式的诊断报告。支持组件日志的一键采集。支持网络监控、磁盘性能下降监控。

1. **独立部署：**InsightTool 作为独立的软件工具，可以单独安装在服务器上。

2. **全面巡检：**自动对 GoldenDB 的各个组件进行巡检，包括管理节点、存储节点、计算节点等，确保 GoldenDB 的整体健康状态。

3. 精准诊断：内置丰富的诊断规则和算法，能够准确识别并定位 GoldenDB 中的潜在问题，如性能瓶颈、组件故障、可疑 SQL 等。

4. 体检报告式诊断报告：提供详细的体检报告式诊断报告，报告内容涵盖了巡检结果、问题诊断、建议措施等多个方面，帮助用户全面了解 GoldenDB 的健康状况。

5. 一键日志采集：支持对 GoldenDB 各个组件的日志进行一键采集，方便用户进行日志分析和故障排查。

6. 网络监控：实时监控 GoldenDB 的网络状态，包括网络延迟、带宽利用率等，确保系统的网络性能稳定可靠。

7. 磁盘性能监控：监控 GoldenDB 所在磁盘的性能指标，如 IOPS、吞吐量等，及时发现并处理磁盘性能下降的问题。

8. 一键卸载：支持脚本方式或者命令行方式对 InsightTool 进行卸载的能力。

● InsightTool 产品优势

1. 高效巡检诊断：InsightTool 能够自动化地完成巡检和诊断任务，大大提高了运维人员的工作效率。

2. 诊断报告精准：内置的诊断规则和算法确保了问题的精准识别与定位，减少了误报和漏报的情况。

3. 报告内容全面：提供了全面的巡检和诊断功能，涵盖了 GoldenDB 的各个组件和性能指标，确保系统的整体健康状态。

4. 支持定制与扩展：InsightTool 支持与其他系统或工具的集成安装，方便用户根据实际需求进行定制和扩展。

● InsightTool 使用场景

1. 日常巡检：运维人员可以使用 InsightTool 进行日常巡检，全面了解 GoldenDB 的健康状态，及时发现并处理潜在问题。

2. 故障排查：当 GoldenDB 出现故障时，运维人员可以使用 InsightTool 进行问题诊断，快速定位故障原因，并采取相应的解决措施。

3. 性能优化：开发人员和运维人员可以利用 InsightTool 提供的体检报告，分析数据库系统的性能瓶颈，进行相应的优化调整，提高系统的运行效率和响应速度。

5 应用场景

5.1 金融行业

金融尤其是银行核心 IT 应用，对性能、稳定性以及安全性有着极高的要求。随着移动互联网和电子商务的发展，金融行业的交易量激增，数据量呈现几何式增长。传统 IOE（IBM 小型机、Oracle 数据库、EMC 存储设备）集中式架构已无法满足持续增长的高并发和快速响应的需求。在这种背景下，GoldenDB 因其自主可控、高性能、高可靠等特点，在金融行业中的应用显得尤为重要。

1. 实现核心系统主机下移

优势：降低成本；灵活扩容；高可靠性

2. 构建新一代金融信息系统

优势：开放灵活；快速创新；平台化管理

应用场景：

- 新一代金融信息系统平台，支持金融科技与互联网金融业务的创新发展，推动业务数字化转型。

3. 建设数据库云服务平台

优势：云服务标准化；简化运维；资源优化

应用场景：

- 金融行业 IT 基础设施上云，采用全栈云服务方式，包括 IAAS、数据库、微服务框架、运维管理工具等，实现开发、部署、扩容、运维和升级的一系列优化。

4. 提升金融交易系统快速决策能力

优势：高效存储；混合处理；实时分析

应用场景：

- 实时风控系统：通过实时数据分析，提升风控能力，防范金融风险。
- 反欺诈系统：利用实时数据分析能力，快速识别欺诈行为。
- 实时经营分析：提供实时经营数据，支持快速决策。

应用案例：

工商银行对公核心系统原先运行在 IBM 大机上，使用 DB2 数据库承载业

务。随着金融核心系统业务量和数据量不断增长，原有的基于大机打造的集中式封闭架构已无法满足工商银行对该核心系统未来增长需求，工商银行核心业务系统亟需向分布式开放架构转型。工商银行于 2020 年开展对该核心系统下主机工作，进入分布式架构转型和国产化转型阶段，着眼于未来，借助分布式开放平台实现对该核心系统安全可控、高并发、高容量、易扩展等特性。

工商银行选择 GoldenDB 开展对该核心系统主机下移项目，目前已成功将该核心系统迁移至 GoldenDB，由 GoldenDB 独立支撑该核心业务运行，该系统已完全退出大机，GoldenDB 支撑工商银行上千万用户账户，日均处理 2 亿服务调用量，提供高效稳定的数据库服务，日均处理近 1.6 亿笔的交易量，承载交易金额日均超万亿。

GoldenDB 在金融行业尤其是银行核心 IT 应用中具有明显的优势，包括实现核心系统主机下移、构建新一代金融信息系统、建设数据库云服务平台以及提升金融交易系统快速决策能力。通过这些应用，GoldenDB 不仅帮助金融机构降低成本、提高效率，还能确保系统的高可用性和安全性，满足金融行业对高性能、高可靠性的严格要求。随着金融行业不断推进数字化转型，GoldenDB 将继续发挥其在数据库领域的领先地位，助力金融机构实现业务创新和发展。

5.2 电信行业

电信行业是数据库产品主要市场之一，也是对数据库产品有很高要求的行业之一，尤其是在容量、性能和 Oracle 兼容性等方面有极致的要求。

电信业务运营支撑系统主要包括 BMO（BSS\MSS\OSS），其中计费/CRM/ERP 等数据库关系着运营商消费者的核心利益，是运营商的最重要数据，对性能、稳定性以及安全性都有着极高的要求，传统集中式架构一方面无法满足大容量、高并发、高可扩展、高可用等关键需求，另一方面，为降低业务改造成本，对国产数据库有很高的语法兼容和工具集要求，具体需求如下：：

- 1) 高度兼容 Oracle 需求。CRM&BOSS 不仅使用大量的复杂 SQL，还使用大量的 Oracle 高级对象特性，如存储过程、包、自定义函数、触发器、匿名块、系统视图、Sequence、自治事务、JOB、XA、物化视图等；
- 2) 因无法进行全覆盖测试，必须通过 SQL 录制及回放工具进行覆盖性验证测试；
- 3) CRM&BOSS 不仅数据量大，每到月初月末，存在大量出账跑批类业务；
- 4) 完整的数据库工具集，包括：迁移工具（全量+增量）、回放工具、数据同步工具、数据库开发工具、巡检功能等等。
- 5) 本地同城高可用，RPO=0，RTO<30S；
- 6) 反向同步，支持故障回切。RPO=0；

目前，GoldenDB 已在运营商大面积商用，且在很多关键系统割接上线，如浙江移动账务库、山东移动营业及账务库、河北移动账务库/计费库/订单库、CMIOT 计费结算中心、中移动全网客户中心、安徽移动产商品中心等落地案例。各系统割接上线后，系统运行稳定。在整个国产化替换过程中，除了数据库驱动、导入导出工具等需要替换成 GoldenDB 的驱动或者导入导出工具外，业务核心代码几乎无改动。

5.3 政务行业

各级政企用户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；长期以来政府机关企事业单位使用的多是国外数据库 Oracle，对 Oracle 的语法、运维手段和性能都非常熟悉。政企客户对于数据库的应用场景和要求如下：

1.强大的 Oracle 兼容性

GoldenDB 支持 Oracle 的大部分语法和功能，包括存储过程、包/包体、自定义函数、自定义数据类型以及游标等。这种兼容性使得现有基于 Oracle 的应用可以直接迁移到 GoldenDB 上，减少了重新开发的成本和复杂性。

优势：无缝迁移；业务无感

2.强劲的性能表现

GoldenDB 在同等资源配置下，能够提供超越 Oracle 的性能表现，尤其是在高并发场景下。

优势：高并发处理；快速查询

3.配套数据同步工具

GoldenDB 提供了配套的数据同步工具，支持核心系统与下游数据分析 OLAP 平台之间的数据同步，能够实现海量数据准实时同步。

优势：数据一致性；实时性

4.合理规划资源投入

GoldenDB 具备多租户能力和动态缩扩容能力，可以根据不同租户的需求合理分配资源，并在紧急情况下进行性能扩容。

优势：资源优化；灵活性。

5.高可用能力

GoldenDB 具备健壮的高可用能力，支持多地多中心的灵活故障切换，确保核心系统的高可用性。

优势：故障切换；数据安全。

6.完善的服务支撑体系

GoldenDB 提供多级保障体系，帮助政企用户建立健全数据库运维管理能力。

优势：技术支持；运维管理

场景案例

海关主要负责监管经各口岸进出境的运输工具、货物、行李物品、邮递物品和其他物品，征收关税和其他税费，查缉走私并编制海关统计和办理其他海关业务。IT 系统从使用场景上分为“对内”和“对外”。

海关业务系统数据体量没有金融行业庞大，但是业务场景更加复杂。其核心类系统有：通关、报关、物流、检验检疫、智能海关等。现阶段核心系统为第 2 代核心系统，主要运行在 Oracle 数据库上，小部分运行在 SQLServer 数据库上。

从 2023 年 2 月份开始，GoldenDB 参与海关新核心数据库选型 POC 测试，截止目前已经完成两个核心业务系统的兼容性对接和性能测试，其中，业务兼容性对接结果整体兼容率 95%以上。在充分完成业务兼容性对接测试之后，启动业务性能测试，目前无论是业务量还是耗时都完全超越了原有的 Oracle 数据库。

GoldenDB 作为一款数据库产品，在政务云环境中具有显著的优势，特别是在需要高并发、大容量数据处理和高性能的场景下。其强大的 Oracle 兼容性、强劲的性能表现、配套的数据同步工具、合理的资源规划、高可用能力和完善的服务支撑体系，使其成为政企客户理想的数据库解决方案。通过 GoldenDB 的应用，各级政府企业可以实现信息共享和业务协同的智能化政务服务，提高政务系统的效率和稳定性。

5.4 交通行业

交通行业是一个高度依赖数据管理和实时处理的领域，涵盖公共交通、道路运输、航空、铁路、港口等多个子领域。交通行业数据库的应用场景非常广泛，从高速过路费到旅客票务服务，从道路监控到港口装卸等，都需要高效、

可靠的数据库支持。以下是交通行业中几个典型的数据库应用场景：

1、交通收费管理系统

交通收费管理系统用于管理和处理各种交通收费，如高速公路收费、停车场收费等。

2.交通规划与设计系统

交通规划与设计系统用于支持交通基础设施的规划和设计，需要处理大量的地理信息数据。

3.交通数据分析系统

交通数据分析系统用于对交通数据进行深入分析，支持决策制定和优化。

4.旅客信息服务系统

旅客信息服务系统为乘客提供实时的交通信息，如公交、地铁、火车、飞机的时刻表、延误信息等。

5、港口管理系统

自动化集装箱码头操作系统（以下简称“TOS”）作为自动化集装箱码头的“大脑”，是集装箱码头实现自动化运行的关键所在。

场景案例

各省交通集团主要从事交通基础设施的投资建设运营、装备制造、城市综合开发等，为客户提供投融资、咨询规划、设计建造、管理运营一揽子解决方案和综合一体化服务。信息化管理，主要包括数字施工、自动化设备管理以及数字化管理系统等，以及对客业务，包括计费功能，面对来往车辆提供快速、准确的服务。

在数据规模方面，核心收费业务相关的数据量大概在 2TB 左右，单表数据量大概在 1000 万行，整体规模不大。现网业务系统（如综合运营管理系统）主要是基于 Oracle。

GoldenDB 的业务兼容性测试和性能测试满足交通集团国产数据库选型要求，后续将针对部署方案和高可用方案进一步深化讨论，进行投产落地。

5.5 医疗行业

医疗行业的核心系统包括医院信息系统(HIS)、电子病历系统(EMR)、实验室检验管理系统(LIS)、医学影像传输系统(PACS)、放射学信息系统(RIS)、临床信息系统(CIS)、合体检信息系统(PEIS)、医院资源管理系统（HRP）、

医院客户关系管理系统（HCRM）等

医疗行业中使用的数据库种类较多，大多是随系统产品采购引入，具体使用哪种类型的数据库，还需要根据具体系统的需求来决定。

场景案例

HIS 系统是医院管理的核心系统，它集成了医院各个科室的医疗信息，包括病人信息、医嘱信息、药品信息、检验信息等，为医护人员提供全面的医疗信息支持。

HIS 系统应用数据库使用特点如下：

- 1) 用 Oracle 和 SQLServer 数据库的 HIS 占了绝大多数；
- 2) 数据类型基本覆盖关系数据库常见的各种类型，以及 clob, blob 等超大字段；
- 3) 严重依赖普通视图、嵌套视图、物化视图、函数、存储过程、索引、触发器、序列等数据库功能，大量业务逻辑在数据库中实现
- 4) OLTP 与 OLAP 共存；而且数据量大。

目前 GoldenDB 已与一些行业内的医疗系统 ISV 进行了对接，与部分业务系统进行了认证与适配。