



北京金融科技产业联盟  
BEIJING FINTECH INDUSTRY ALLIANCE

# 金融行业 GaussDB 运维白皮书

北京金融科技产业联盟

2025 年 9 月

## 版权声明

本报告版权属于北京金融科技产业联盟，并受法律保护。转载、编摘或利用其他方式使用本报告文字或观点的，应注明来源。违反上述声明者，将被追究相关法律责任。



## 编制委员会

### 主任：

聂丽琴 赵永田

### 编委会成员：

胡达川 张明永 窦德明

### 编写组成员：

张 浩 彭 烱 任 政 孙海霞 李 星 张 政  
韩培亮 葛伟豪 苏积辉 曾小龙 张旱文 李胜杰  
谢家凯 刘占明 赵 琼 饶昊泉 尹宏明 薛德龙  
王晓华 王 辉 陈 曦

### 编审：

李明艳 吴 辉 洪学敏 肖 永

### 参编单位：

北京金融科技产业联盟  
中国工商银行股份有限公司  
中国银行股份有限公司  
华为云计算技术有限公司  
交通银行股份有限公司  
中国邮政储蓄银行股份有限公司  
华夏银行股份有限公司

## 序言

随着信创进程的加速与金融行业数字化转型的深入，核心数据库作为承载金融业务数据、支撑关键业务运转的“心脏”，其安全、稳定、高效及自主可控的重要性日益凸显。华为云 GaussDB 作为国内领先的金融企业级数据库，历经多年发展与实践检验，已在众多金融机构实现了从边缘业务应用到核心系统承载的历史性跨越，成为支撑金融行业数字化转型与核心技术自主创新的坚实基础。

在此背景下，数据库的运维体系面临着从传统集中式向现代分布式架构演进的全新挑战。为应对海量数据、高并发交易、业务连续性、安全合规等复杂需求，构建一套标准化、专业化、智能化的运维管理体系迫在眉睫。为此，在金融科技产业联盟的牵头组织下，中国工商银行、中国邮政储蓄银行、华夏银行、交通银行、中国银行等多家国内领先的金融机构，与华为云计算技术有限公司通力合作，结合各机构在 GaussDB 大规模应用中的深厚积累与宝贵经验，共同编撰《金融行业 GaussDB 运维白皮书》。

本白皮书旨在系统性地总结 GaussDB 在金融核心业务场景下的运维理念、最佳实践与技术规范。它面向广大金融行业的数据库管理员、运维工程师、技术负责人、架构师以及致力于数据库运维体系规划与建设的决策者与管理者，期望为正在或计划使用 GaussDB 的机构提供从部署规划、日常监控、性能优化、高可用保障、容灾备份到安全合规等全生命周期的运维指导。

本白皮书的编写与发布具有三大重要意义：

一是凝聚行业共识，树立运维标杆：汇聚头部金融机构的集体智慧，形成行业认可的 GaussDB 运维标准框架，降低学习与试错成本，助力产业整体运维水平提升。

二是保障系统稳定，促进业务创新：通过规范、高效的运维实践，确保 GaussDB 数据库的稳定可靠运行，为金融业务的连续性和敏捷创新提供强大底层支撑。

三是推动自主可控，共建产业生态：夯实国产数据库的落地应用能力，加速金融核心系统的安全可控进程，共建健康、繁荣的数据库产业新生态。

我们衷心希望本白皮书能成为金融行业同仁在华为云 GaussDB 运维道路上的“灯塔”，为保障我国金融系统的安全稳定运行、推动金融科技高质量发展贡献一份力量。未来，我们也将持续迭代更新，与产业共同进步。

金融科技产业联盟

二〇二五年九月

## 目 录

一、 金融行业数据库运维挑战与GaussDB应用概况.....	7
(一) 金融行业业务特点与 IT 行业挑战.....	7
(二) 金融行业数据库运维发展机遇与挑战.....	10
(三) GaussDB 铸就金融数据库底座 .....	17
二、 GaussDB金融级运维体系总览 .....	19
(一) 运维目标与原则.....	19
(二) GaussDB 运维管理体系 .....	21
(三) GaussDB 运维关键技术 .....	29
三、 GaussDB典型运维场景与实践案例.....	41
(一) 高可用与容灾运维.....	41
(二) 运维效率提升.....	43
(三) 容量管理运维.....	48
(四) 巡检.....	51
(五) 备份恢复.....	54
四、 运维模式演进与技术发展展望 .....	57
(一) GaussDB 数据库技术智升金融行业运维展望 .	57
(二) 运维模式发展趋势.....	59
(三) 共同挑战与应对建议.....	62

## 一、金融行业数据库运维挑战与 GaussDB 应用概况

### （一）金融行业业务特点与 IT 行业挑战

在数字化浪潮的推动下，金融行业正经历着前所未有的变革。作为现代经济的核心，金融业务的特殊性决定了其信息技术系统必须满足极高的标准。尤其是核心业务系统，其稳定运行直接关系到国家经济安全和公众利益。因此，数据库作为支撑金融业务的关键基础设施，在可靠性、一致性、性能及合规性等方面面临着极致要求。本节将深入分析金融行业的业务特点，并探讨 IT 行业在金融业务快速发展背景下所面临的挑战。

#### 1. 金融行业的业务特点

金融行业的业务特点主要体现在以下几个方面：

##### （1）高可靠性与业务连续性要求

金融业务涉及资金流动和客户资产安全，任何系统故障都可能导致重大经济损失和声誉风险。因此，金融核心系统必须实现高可用性和业务连续性，通常要求系统全年无故障运行时间达到 99.999%（即每年停机时间不超过 5 分钟）。例如，银行的核心交易系统、支付清算系统等必须具备容灾备份和快速故障恢复能力，以应对硬件故障、网络中断甚至自然灾害等突发情况。

##### （2）强一致性与数据准确性

金融交易对数据一致性要求极高，任何数据不一致都可能引发账务错误甚至金融风险。ACID（原子性、一致性、隔离性、持久性）特性是数据库支持金融交易的基本要求。尤

其是在分布式架构下，如何保证跨节点事务的强一致性成为关键挑战。

### (3) 高性能与低延迟

金融业务场景中，高并发和低延迟是核心需求。例如，证券交易系统在高峰时段需处理每秒数万笔交易，且响应时间需控制在毫秒级别。支付系统在大型促销活动（如“双十一”）期间也面临瞬时流量激增的考验。数据库必须能够高效处理海量数据请求，同时保持稳定的性能表现。

### (4) 严格的安全与合规性

金融行业受到严格监管，数据安全和隐私保护是重中之重。国内外多项法规（如《网络安全法》、GDPR、PCI-DSS等）对数据的存储、传输和使用提出了明确要求。数据库需要提供完善的安全机制，包括数据加密、访问控制、审计日志等，以满足合规要求。

### (5) 业务复杂性与快速创新

金融业务类型多样，涵盖存贷、支付、理财、风控等多个领域，业务逻辑复杂。同时，随着金融科技的发展，新的业务模式（如移动支付、区块链、智能投顾等）不断涌现，要求数据库具备良好的扩展性和灵活性，以支持业务快速迭代和创新。

## 2. IT 行业在金融领域的挑战

在金融行业业务不断发展的背景下，IT 行业面临着多方面的挑战：

### (1) 传统架构的局限性

许多金融机构仍采用集中式数据库架构（如大型机+传统关系型数据库），虽然成熟稳定，但存在扩展性差、成本高昂、维护复杂等问题。随着数据量的爆炸式增长和业务模式的多样化，传统架构难以满足新时代的需求。

### （2） 异构数据库转型的技术复杂性

金融行业异构数据库迁移技术复杂度极高，需确保数据一致性、业务连续性及合规安全。涉及不同数据库 schema、类型映射与语法转换，需处理海量历史数据并最小化停机窗口。迁移过程中存在数据失真、应用兼容性及性能风险，需设计周密上线方案与应急逃生机制。

### （3） 合规与安全压力

金融行业的合规要求日益严格，IT 系统需不断适应新的监管政策。数据本地化、隐私保护、审计追踪等要求增加了系统设计和运维的复杂度。此外，网络安全威胁日益猖獗，如何防范数据泄露和恶意攻击成为持续挑战。

### （4） 运维保障与人才短缺

金融核心系统的运维保障要求极高，需具备完善的监控、预警和应急响应机制。然而，分布式数据库的运维难度较大，需要专业人才负责性能调优、故障诊断和容灾管理。目前，兼具金融业务知识和分布式技术能力的复合型人才严重短缺。

### （5） 成本与效率的平衡

金融机构在追求技术先进性的同时，需综合考虑成本效益。传统数据库 license 费用高昂，而分布式数据库的硬件成

本和运维投入也可能较大。如何在保障系统性能和安全的前提下实现降本增效，是 IT 管理的重要课题。

### (6) 生态兼容与迁移风险

数据库替换涉及业务迁移、数据同步、应用适配等一系列工作，风险较高。尤其是金融核心系统，迁移过程中必须保证业务连续性和数据一致性。此外，现有应用通常基于传统数据库开发，如何实现平滑迁移和生态兼容也是巨大挑战。

## (二) 金融行业数据库运维发展机遇与挑战

### 1. 数据库运维发展状况及趋势

数据库运维从早期的人工手动运维，逐步发展到如今融合自动化、智能化技术的运维模式，其发展历程与数据库技术的演进紧密相关。数据库运维技术支撑体系从单一工具向平台化、集成化发展，不断适应着日益增长的数据管理需求和复杂多变的应用环境。数据库运维模式的迭代如表 1 所示，每个阶段的核心特征与技术支撑存在显著差异：

表 1 数据库运维模式演进

运维模式	特点	核心技术支撑	痛点或优点
人工手动运维 (1980s-2000s)	依赖DBA经验、 手动操作	命令行工具 (SQL*Plus)、 Excel	痛点： 备份失败率高； 故障响应慢； 人力成本高
自动化运维 (2000s-2015s)	脚本/工具驱 动、任务自动化	Shell/Python 脚 本、Zabbix	痛点： 缺乏全局决策； 异常根因定位难； 策略调整滞后
智能化运维 (2015s-至今)	AI算法赋能、预 测式运维	AI 运维平台（如 Moogsoft）、机器学	优点： 性能提前预测；

		习框架	故障分钟级修复; 资源动态优化
--	--	-----	--------------------

在数据库发展的初期阶段，数据量相对较小，应用场景也较为简单，数据库运维主要依赖人工手动操作。数据库管理员（DBA）需要承担从数据库的安装、配置、数据备份与恢复，到性能监控与优化等一系列工作。在数据库安装过程中，DBA 需要根据硬件环境和应用需求，手动配置数据库的各项参数，如内存分配、存储路径等，任何一个参数设置不当都可能影响数据库的性能。在数据备份方面，DBA 需要定期手动执行备份操作，并将备份数据存储到安全的介质中，这种方式不仅耗费大量时间和精力，而且容易因人为疏忽导致备份失败或数据丢失。在性能监控方面，DBA 需要通过各种命令行工具，手动收集数据库的性能指标数据，如 CPU 使用率、内存占用、磁盘 I/O 等，并根据经验进行分析和优化，这对 DBA 的专业技能和经验要求极高。随着数据量的逐渐增加和应用系统的日益复杂，人工手动运维的局限性愈发明显，难以满足高效、稳定的数据管理需求。

为了提高运维效率，降低运维成本，自动化运维逐渐成为数据库运维的重要手段。自动化运维通过编写脚本、使用自动化工具等方式，实现了数据库运维任务的自动化执行。利用自动化脚本可以实现数据库的自动安装和配置，大大缩短了数据库部署的时间，并且能够确保配置的一致性和准确性。在数据备份与恢复方面，自动化工具可以按照预定的策略，定期自动执行备份操作，并在需要时快速恢复数据，有效提高了数据的安全性和可用性。自动化监控工具能够实时

采集数据库的性能指标数据，并通过设定阈值，在指标异常时自动发出警报，使 DBA 能够及时发现和处理问题，减少了因故障导致的业务中断时间。一些企业还建立了数据库自动化运维平台，将各种自动化工具和脚本进行整合，实现了对数据库全生命周期的自动化管理，从数据库的创建、部署、监控、维护到退役，都可以通过平台进行统一调度和管理，进一步提升了运维效率和管理水平。

随着人工智能技术的不断发展，智能化运维成为数据库运维的新趋势。智能化运维利用机器学习、深度学习、大数据分析等技术，对数据库的运行数据进行深度挖掘和分析，实现对数据库性能的智能预测、故障的自动诊断和根因分析，以及运维策略的自动优化。通过机器学习算法对历史性能数据进行学习，可以建立性能预测模型，提前预测数据库在未来一段时间内可能出现的性能瓶颈，为 DBA 提供预警信息，以便提前采取优化措施。在故障诊断方面，智能化运维系统可以实时收集数据库的各种运行数据，包括日志信息、性能指标、系统状态等，并利用深度学习算法对这些数据进行分析，快速准确地定位故障原因，自动生成解决方案，缩短了大量的故障分析处理时间。在运维策略优化方面，智能化运维系统可以根据数据库的实时运行状态和业务需求，自动调整数据库的参数配置、资源分配等运维策略，以实现数据库性能的最优。智能化运维的应用，使得数据库运维从传统的被动响应式运维向主动预防式运维转变，极大地提高了数据库运维的效率和质量，降低了运维成本和风险。

## 2. 金融行业运维模式转变

以某大型商业银行的核心交易数据库为例，不同运维模式下的关键指标差异如表 2 所示：

表 2 某银行 2018-2023 年运维年报

运维模式	故障平均修复时间 (MTTR)	数据备份失 败率	业务中断年 均次数	运维人力成本 (人 / 年)
人工手动运维	4.2 小时	8.7%	12 次	15 人 / 年
自动化运维	1.5 小时	2.3%	5 次	8 人 / 年
智能化运维	0.4 小时	0.5%	1 次	4 人 / 年

在金融行业，数据库运维发展状况的演变有着诸多典型案例。例如，某大型商业银行引入备份管理工具实现数据库每日定时备份，并通过自动化监控工具实时监测数据库性能指标。当 CPU 使用率超过 80% 时自动发出警报，运维人员能及时介入优化，这使得数据库故障预警及时性大幅提升，业务中断时间从过去平均每次数小时缩短至半小时以内，同时避免了客户投诉。

又如，一家金融公司，业务涵盖多种创新金融产品，数据处理复杂且对实时性要求极高。通过对数据库和系统资源进行预测，避免业务量剧增情况下资源问题拖垮性能。同时对贷款申请等复杂查询极易导致数据库查询响应变慢的场景，通过诊断优化给出索引推荐、执行计划或 CPU、内存资源配置的优化建议，保障业务高峰期系统的稳定运行，客户贷款申请审核的平均处理时间缩短了 30%，极大提升了客户体验和业务竞争力。

### 3. 金融行业数据库面临的运维挑战

金融行业作为国民经济的核心领域，其数据库承载着客户账户信息、交易流水、信贷数据等关键资产，且需支撑7×24小时不间断的业务服务（如实时转账、证券交易、保险理赔等），这使得数据库运维面临着远超其他行业的严苛要求。从实际业务场景来看，当前金融行业数据库运维的核心挑战集中在高可用保障、安全防护与运维效率等多重挑战，每一项挑战的突破与否，都直接关系到金融业务的连续性、客户权益的安全性及机构运营的稳定性。

#### （1）数据安全与合规性要求带来的运维压力

金融行业涉及大量敏感信息，如客户的个人身份信息、账户信息、交易记录等，数据安全至关重要。同时，金融行业受到严格的法规监管，如《通用数据保护条例》（GDPR）、《支付卡行业数据安全标准》（PCI DSS）等，合规性要求极高。在数据库运维过程中，需要采取一系列措施确保数据安全和满足合规性要求。例如，为了防止数据泄露，金融机构需要对数据库进行严格的访问控制，采用最小权限原则等措施。同时，需要对数据进行加密存储和传输，定期进行数据备份和恢复演练。在合规性方面，运维人员需要确保数据库的操作记录完整、可追溯，满足监管机构的审计要求。

#### （2）数据库高可用挑战：零中断目标下的多重压力

金融业务对“连续性”的要求近乎苛刻——银行转账系统中断1分钟可能导致数百万笔交易失败，证券交易系统故障10分钟可能引发市场波动，保险理赔系统宕机则会

直接影响客户体验与机构公信力。然而，金融数据库在实际运行中，需面对硬件故障、网络波动、数据量激增、业务峰值冲击等多重风险，高可用保障成为运维工作的“重中之重”，具体挑战体现在以下三方面：

a. 核心业务场景下的“零 RTO/RPO”实现难度

RTO（恢复时间目标）与 RPO（恢复点目标）是衡量数据库高可用的核心指标，金融行业对这两项指标的要求已逐步向“零”靠拢。但在实际运维中，实现这一目标面临诸多障碍：一方面，金融数据库多采用“主从架构”或“集群架构”，当主节点因硬件故障或软件 BUG 下线时，需依赖自动故障转移机制切换至备节点，但切换过程中可能出现“脑裂”、数据同步延迟等问题，导致 RTO 超出预期或出现部分数据丢失；另一方面，对于分布式数据库，其节点数量通常达数十甚至上百个，某一节点故障可能引发数据分片重新分配、一致性协议重试等连锁反应，若运维监控不及时或故障转移策略不合理，极易延长恢复时间。

b. 业务峰值与数据量激增的冲击：

金融业务具有显著的“周期性峰值”特征：每逢节假日、电商促销、证券交易日开盘 / 收盘时段，数据库需承受远超日常数倍的并发压力。若运维团队未提前做好资源扩容、性能优化，极易导致数据库出现 CPU 使用率飙升、内存溢出、磁盘 I/O 饱和，进而引发交易超时、查询卡顿甚至系统宕机。此外，金融数据量正以每年 30%-50% 的速度增长，传统的“垂直扩容”（如升级服务器配置）已无法满足存储需求，而

“水平扩容”（如增加节点、数据分片）则需兼顾数据一致性与业务兼容性，运维复杂度大幅提升。

c. 灾备体系建设的“成本与效率”平衡难题：

为应对地震、火灾导致的数据中心级故障等极端场景，金融机构需搭建跨地域灾备体系，但这一过程面临“成本高”与“效率低”的双重矛盾。一方面，灾备建设需投入大量硬件、运维人员配置等；另一方面，灾备演练的“真实性”与“业务影响”难以平衡：若演练场景过于简单则无法验证灾备体系的有效性；若采用“真实业务切换”则可能因演练过程中的异常影响正常业务。

(3) 数据库运维效率挑战：复杂度与人力的不匹配困境

随着金融数字化转型的深入，数据库架构从“集中式”向“分布式”演进，数据库类型从“单一关系型”向“多元混合”扩展，运维对象从“数十个节点”向“数百个甚至数千个节点”增长，而运维团队规模却难以同步扩张，“运维复杂度”与“人力效率”的矛盾日益突出，具体挑战体现在以下三方面：

a. 多元数据库环境的“碎片化”管理

为满足不同业务场景需求，金融机构往往搭建“多元数据库环境”：核心交易系统采用传统关系型数据库，移动支付等高并发业务采用非关系型数据库，行情分析、风控建模采用时序 / 列存数据库，客户关系管理采用图数据库。据调研，某大型银行已部署 12 种不同类型的数据库，节点总数超

500 个。这种“碎片化”环境给运维带来两大难题：一是“工具不兼容”——不同数据库的管理工具、监控指标、备份策略差异极大，运维人员需切换不同工具完成工作，效率低下；二是“知识门槛高”——运维人员需同时掌握关系型数据库的事务处理、非关系型数据库的分布式一致性、时序数据库的压缩算法等知识，而当前金融行业兼具“多数据库技能”的 DBA 仅占 20%，人员缺口显著。

### b. 运维流程的“自动化”覆盖不足

金融行业在数据库运维领域，仍有大量工作依赖人工操作：例如，数据库参数调整需手动修改配置文件，数据迁移手动执行脚本。人工操作不仅效率低，还易因“人为失误”引发故障。再如，运维人员在执行数据备份时，漏选关键表，导致恢复时数据缺失。此外，金融业务的“敏捷迭代”要求数据库运维也需具备“快速响应”能力，而人工操作难以满足这一需求，成为业务创新的“瓶颈”。

## （三） GaussDB 铸就金融数据库底座

GaussDB 扎根金融行业五载，历年推出的核心金融解决方案见图 1。



图 1 GaussDB 历年推出的特性示例

从应用规模来看，GaussDB 已实现从国有大行到区域银行的全层级覆盖，服务客户类型包括：国有大型银行、股份

制银行与城商行、证券保险、信贷等。以下做部分应用情况介绍：

### 1. 国有大型银行

工商银行：工商银行结合金融科技发展与业务创新需求，应用 GaussDB 不同解决方案落地全球信贷、企业手机银行、多渠道收费等 200+业务系统、上线 2500+节点。

中国邮政储蓄银行：邮储银行顺应数字化发展趋势，响应国家“十四五”政策要求，以新一代个人业务核心系统建设为契机，以金融科技赋能高质量发展，打造服务乡村振兴和新型城镇化的领先的数字生态银行。其核心系统的统一查询系统部署在全国最大规模的分布式集群，支撑 6.5 亿用户十年的历史交易明细查询和收支分析。

中国银行：中国银行是中国全球化和综合化程度最高的银行，中国银行应用 GaussDB 数据库目前已完成 100+业务系统 1300+节点投产上线，包括快捷支付、网银两个一类核心业务系统，支撑行内业务平稳快速完成 Oracle 数据库替换，满足全球范围内高可用、高性能和高安全的国产化数据服务。

交通银行：交通银行股份有限公司 GaussDB 已经完成金融市场交易管理 A 类业务系统上线。该系统包含前台交易、中台风控和后台清算，承载交通银行 X 万亿资产，年贡献利润占行内总利润的 1/3。该系统上线实现了应用、数据库和存储等核心部件自主可控。

### 2. 股份制银行与城商行

华夏银行：华夏银行作为全国性股份制银行，始终高度重视金融科技发展，开始重构底层核心系统，布局产业数字金融。依托 GaussDB 实现了业界领先的云计算 IT 架构，上线借记卡等十余个核心业务稳定运行，性能提升 1.5 倍。

### 3. 证券保险机构

国信证券是具有各类业务资格的“全牌照”券商，也是经中国证监会核准的首批证券承销与保荐机构。国信证券基于 GaussDB 打造新一代核心交易系统，实现极速订单委托时延，订单入库性能提升 43%，整体清算效率提升 4 倍，数据转换效率提升 6 倍。

## 二、GaussDB 金融级运维体系总览

### （一）运维目标与原则

金融行业的特殊性决定了其数据库运维工作不能仅限于保证数据库的“可用”，而必须围绕金融业务的本质需求，确立更高阶的目标，并遵循现代化的运维原则。

#### 1. 运维目标

GaussDB 金融级运维体系的构建旨在达成以下四大核心目标，全方位支撑金融业务的稳定运行：

##### （1）保障业务连续性

这是金融运维的生命线。目标是通过高可用、容灾、备份恢复等技术手段，确保数据库在任何情况下（包括硬件故障、网络中断、数据中心级灾难等）都能快速恢复，实现极低的恢复点目标（RPO）和极速的恢复时间目标（RTO），最大限度保障业务不中断、数据不丢失。

## （2） 保障性能稳定

核心交易系统对延迟和吞吐量有极致要求。运维体系需确保数据库性能能够满足业务高峰期的需求（如秒杀、年终结算、批量处理等），并且能够长时间（7x24 小时）以预期的性能水平运行，不会出现剧烈波动。这意味着需要对性能进行持续监控、容量规划和精细化的调优，避免性能波动影响用户体验和业务处理效率。

## （3） 保障安全合规

安全是金融业的基石。运维工作必须严格遵循国家及行业法规（如《网络安全法》、《数据安全法》、个人金融信息保护规定、银保监会相关指引等）。目标是通过全方位的安全管控，包括但不限于网络隔离、访问权限控制、数据加密（静态和动态）、安全审计、漏洞扫描与修复，构建纵深防御体系，确保数据的机密性、完整性和可用性，并满足合规审计要求。

## （4） 提升运维效率

在满足上述三项要求的前提下，通过工具和流程的优化，降低运维工作的复杂度和人力成本，提升变更、部署、扩缩容等日常操作的效率和准确性，实现“降本增效”，使运维团队能更专注于高价值的设计和 optimization 工作。

## 2. 运维原则

为实现上述目标，GaussDB 金融级运维体系的建设与执行应遵循以下四项核心原则：

### （1） 标准化

制定并严格执行统一的运维标准和管理规范，涵盖环境规划、安装部署、参数配置、监控指标、操作流程、文档模板等。标准化是自动化的基础，能有效减少人为错误，提升操作的一致性和可预测性。

### （2） 自动化

将重复性、规律性的运维工作（如巡检、备份、部署、扩缩容、故障切换等）尽可能自动化。通过自动化脚本或平台工具，减少人工干预，提升操作效率与准确性，并实现 7x24 小时的无人值守操作，是保障大规模数据库集群稳定运行的必由之路。

### （3） 可观测

运维不应仅限于监控简单的“是否存活”状态，而应具备深度可观测能力。这意味着需要从指标、日志和链路三个维度，全方位、细粒度地采集和分析数据库及其所在基础设施的运行状态，能够快速洞察系统内部隐含的异常、精准定位故障根因，并预测潜在风险。

### （4） 智能化

在可观测性产生海量数据的基础上，引入 AI(人工智能)能力，实现智能化的运维升级。具体包括：异常检测、智能预警、根因分析、自愈恢复、性能调优推荐、容量预测等。智能化是运维体系从“被动救火”走向“主动预防”和“自主优化”的高级阶段。

## （二） GaussDB 运维管理体系

随着金融行业数字化转型进入深水区，数据库作为承载核心业务与数据的基石，其稳定性、安全性与高性能至关重要。本章节旨在系统性地阐述基于 GaussDB 的金融级数据库运维管理体系。该体系深度融合了金融行业的严苛要求和 GaussDB 在金融领域多年的最佳实践，涵盖从上线到灾备、从安全到变更的全生命周期管理，为金融机构构建“稳、敏、智”的数据库运维能力提供全面框架，见图 2。



图 2 GaussDB 运维体系全景

## 1. 数据库全生命周期上线保障

为保障数据库迁移上线过程的平滑、安全与高效，我们总结并提出“金融行业十步迁移法”，形成标准化操作规程，如图 3 所示。



图 3 金融行业十步工艺迁移法

(1) 转型评估：全面调研现有业务架构、数据规模、SQL 兼容性及性能基线，评估迁移可行性与风险，制定总体策略。

(2) 方案设计：设计详尽的迁移技术方案，包括资源规划、网络架构、高可用方案、回退机制等。

(3) 对象迁移：使用专业迁移工具，自动化完成表、视图、索引、存储过程等数据库对象的转换与创建。

(4) 数据迁移：采用全量+增量同步技术，在保证数据一致性的前提下，实现海量数据的低延迟、无损迁移。

(5) SQL 审核：对应用 SQL 进行自动化审核与优化，识别不兼容语法、性能瓶颈及潜在风险，确保 SQL 在 GaussDB 上的最佳性能。

(6) 对比测试：在测试环境进行数据一致性校验，确保源端与目标端数据 100%一致。

(7) 功能测试：全面验证业务功能在 GaussDB 上的正确性，确保业务逻辑无误。

(8) 性能测试：通过压力测试、负载测试等手段，验证系统在高并发、大数据量下的性能表现，并针对性调优。

(9) 流量录放：录制生产环境真实流量，在测试环境进行回放，进一步验证系统稳定性和性能。

(10) 灰度上线：采用分批发布、蓝绿部署等策略，先切分少量非核心业务流量，验证无误后逐步完成全量上线，最大化降低风险。

## 2. 数据库容灾设计

基于金融行业多活/同城双活、两地三中心的主流容灾架构，设计 GaussDB 的容灾方案。利用其内置的逻辑复制、流复制等技术，实现跨 AZ、跨 Region 的数据同步。定期组织容灾演练，模拟真实故障场景，验证 RPO（恢复点目标）、RTO（恢复时间目标）指标，确保故障时能快速、准确切换。

## 3. 数据库备份设计

根据金融监管要求，制定多层次备份策略：

全量备份：定期执行（如每周），作为基线。

增量备份：高频执行（如每天），减少备份窗口。

日志备份：实时归档，支持时间点恢复（PITR）。

备份数据需异地、异介质存放，并定期进行恢复演练，验证备份数据的有效性和可恢复性。

## 4. 数据库安全管理

身份认证与权限管理：基于角色的最小权限分配，支持三权分立，避免权限过度集中。

数据加密：支持传输加密和存储加密（如透明数据加密 TDE），符合金融监管要求。

数据脱敏与审计：提供动态/静态数据脱敏功能，并记录所有敏感操作日志，实现全量 SQL 审计，满足合规审计与溯源需求。

安全加固：遵循 CIS 等安全基线，对操作系统和数据库进行安全配置加固。

## 5. 数据库版本管理

建立严格的版本管理制度，对 GaussDB 数据库内核、驱动、补丁等进行统一纳管。制定清晰的升级策略和流程，在充分测试的基础上，选择业务低峰期进行升级操作，并确保升级失败后可快速回退，保障业务连续性。

## 6. 数据库日常与深度巡检

日常巡检：通过自动化工具每日检查数据库核心健康度指标，如连接数、CPU/内存/磁盘使用率、慢 SQL、锁等待、复制状态等，并生成健康报告。

深度巡检：定期（如月度/季度）进行系统性深度检查，包括空间增长趋势分析、性能基线对比、参数配置优化、存储引擎碎片整理等，主动发现并消除潜在风险。

## 7. 数据库故障管理

建立标准化的故障应急响应流程，包括：故障发现与通告 -> 快速定位与诊断 -> 恢复与处置 -> 复盘与改进。利用 GaussDB 的智能 DBA 助手和运维平台，实现故障的快速发现与根因分析，并通过预案快速恢复业务。事后进行深度复盘，完善预案，形成知识库，避免同类故障再次发生。

## 8. 应急预案

在金融行业数据库运维实践中，生产问题的恢复效率往往至关重要。在金融行业，普遍要求达到 1-5-10 的快速响应能力，即 1 分钟发现，5 分钟定界，10 分钟恢复。因此对于故障场景的应急预案，显得极为关键。应急处理以快速恢复设备的正常运行与业务的正常提供为核心指导思想，其总体处理流程如图 4 所示。

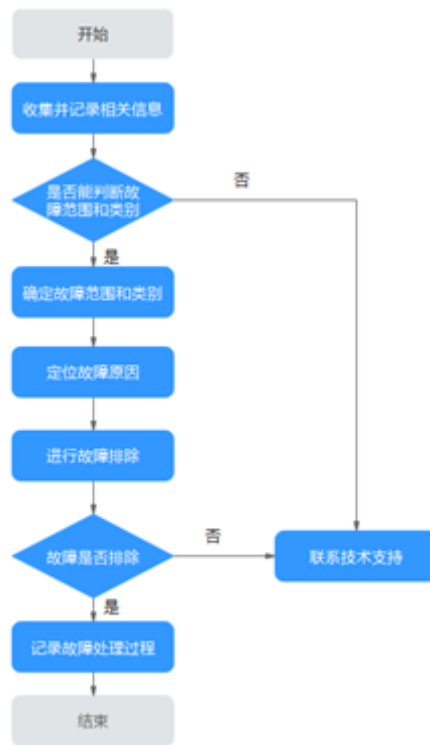


图 4 应急预案处理流程

### (1) 应急处理原则

由于重大故障很容易业务受损等严重后果，具有很大的危害性。为提高重大故障的处理效率、并尽最大的限度降低此类故障的损失，在维护本设备之前，应充分考虑并遵循以下应急处理的基本原则。

应急处理以快速恢复设备的正常运行与业务的提供为核心，因此，提高重大故障处理效率的关键是：客户应参考应急处理手册及时制定各种重大故障的处理预案，并定期组织相关管理人员与维护人员进行学习、演练。

以客户业务尽快恢复，对客户影响最低为原则。在此前提下，进行问题定位恢复和数据收集。

维护人员在上岗前必须接受必要的应急处理培训，学习判断重大故障的基本方法、掌握处理重大故障的基本技能。

## （2）业务快速恢复原则

业务恢复应综合考虑相应操作恢复业务成功的可能性和相应操作时间代价。参考的操作排序优先级如下：

优先：耗时比较短，成功可能性比较大的操作。

次之：耗时比较短，成功可能性比较小的操作。

最后：耗时比较长，成功可能性比较大的操作。

## （3）故障信息收集

故障信息是故障处理的重要依据，系统维护人员应尽可能多的收集故障信息。

## （4）故障判断

在获取故障信息后，需要对故障现象作出判断，确定故障的范围与种类。主要类型如下：

资源类故障：如 CPU 使用率 100%，线程池使用率 100%，动态内存使用率 100%，磁盘只读等。

网络故障：如关键服务之间网络无法互通、网络中断、网络质量差等。

底层故障：如物理机/虚拟机故障（内存、磁盘等）等。

业务类故障：如实例状态异常等。

业务缓慢：P80/P95 异常等。

## （5）故障定位

故障定位就是从众多可能原因中找出故障原因的过程。通过一定方法或手段分析、比较各种可能的故障成因，不断排除非可能因素，最终确定故障发生的具体原因。故障一般伴随服务日志、告警、错误提示和监控状态等信息，用户根

据关键服务日志信息、告警处理建议和应急预案，一般可以定位故障。

#### （6）故障排除

定位故障原因后，需要进行故障排除。故障排除是指采取适当的措施或步骤清除故障、恢复系统的过程。如检查进程状态、查看日志、修改相关配置文件，业务限流和重启服务等。根据不同的故障按照不同的操作规程操作，进行故障排除。例如资源类故障，在 SQL 异常或不优引发数据库资源瓶颈的异常场景，使用 GaussDBSQL 全链路分析、SQL 限流、SQL patch 调优进行快速的定位和恢复。

#### （7）确认故障是否被排除

故障排除之后要注意进行检测，以确保故障彻底被排除。通过查看日志、重复引发故障发生的动作和告警等方法确认系统已正常运行，并进行功能测试，以确保故障排除，业务正常。

#### （8）记录故障处理过程

故障排除后应回顾故障处理全过程，记录故障处理要点，给出针对此类故障的防范和改进措施，避免同类故障再次发生。

### 9. GaussDB 重大问题响应流程

华为公司深耕金融行业数十年，其服务团队深刻理解金融行业客户的业务诉求与监管要求。华为建立了全球化的专业服务支持体系，特别是在重大问题响应方面，其成熟的支撑体系确保了客户问题能得到最高效的解决。

7x24 小时全球技术支持中心（GTAC）：提供不间断的远程技术支持。问题分级与升级机制：根据问题影响范围紧急程度，自动触发升级流程，调动相应层级的技术专家资源。“一线授权”机制：对于紧急重大故障，一线支持工程师可快速申请所需资源（如研发专家），以最快速度恢复业务。根因分析与知识沉淀：问题解决后，团队会输出详细的根因分析报告，并将解决方案沉淀至知识库，用于优化产品和服务，形成闭环。

### （三） GaussDB 运维关键技术

#### 1. GaussDB 关键金融特性

##### （1） 高可用架构设计

不同的容灾诉求，催生多种拓扑组网，GaussDB 支持跨中心单集群、跨中心双集群、两地三中心、多地多中心等多种高可用架构，满足不同容灾等级需求，如图 5 所示。



图 5 GaussDB 高可用组网架构

普通业务：满足成本约束条件下，RPO 尽量小，RTO 尽量小，整系统可用度达到 99.95%，一般采用本地高可用或者异步复制容灾。

一般关键业务：RPO 尽量小，RTO 尽量小，整系统可用度达到 99.99%，一般采用同城双活。

核心关键业务：RPO=0，RTO 尽量小，整系统可用度达到 99.999%，一般采用两地三中心或者多地多中心。

## （2）应用无损透明

应用无损透明（ALT, Application Lossless and Transparent）提供了一种数据库状态变化的主动消息通知机制。JDBC 驱动也向应用程序提供了数据库实例状态变化的回调函数注册接口。应用程序可以针对某些数据库连接、向 JDBC 驱动注册状态变化的回调函数。当数据库实例状态发生变化时，JDBC 驱动会对注册的函数进行调用，通过注册回调函数可以很方便地在业务侧实现数据库状态变化的邮件通知、告警平台上报等运维管理操作。在发生故障切换（failover）时，能够向应用屏蔽可恢复的故障，通过驱动透明地自动切换到新主节点，恢复会话状态，并重新执行被中断的事务，使应用在事务处理上感知不到 I/O 异常和数据不一致，只是中间会有短暂的停顿，就像是执行了一条慢 SQL。此特性不仅降低了应用侧 DAO 层的开发复杂度，也降低了业务报错的概率，提升了运维效率。

## （3）备份与恢复

GaussDB 支持数据库实例的备份和恢复，以保证数据可靠性。当数据库或表被恶意或误删除，虽然 GaussDB 支持高可用，但备机数据库会被同步删除且无法还原。因此，数据被删除后必须依赖于实例的备份保障数据安全。GaussDB 支持实例级全量备份、表级全量备份、实例增量备份、表级增量备份，支持实例级、库表级、租户级恢复，并支持恢复到时间点。在误删实例、误删表、误删数据库或误操作表中数据的场景下，数据库运维工程师都能够很便捷地使用备份恢复功能恢复数据。提供灵活的备份方式切换功能，允许用户根据业务需求和实际情况，在快照备份与物理备份之间进行切换。数据量较大时，物理备份时间较长，可以切换为快照备份。

#### （4） In-place Update（原地更新）行存储引擎

相比于 Append Update（追加更新）行存储引擎，In-place Update（原地更新）行存储引擎可以提高数据页面内更新的 HOT UPDATE 的垃圾回收效率，有效降低多次更新元组后存储空间占用的问题，基于历史版本直接进行回收，实现了自治式的空间管理机制，减少了 I/O 时的性能抖动。同时实现了多个后台线程的并发访问，降低并发业务冲突竞争，从而提高性能。

#### （5） 在线 DDL

传统 DDL 操作若涉及重建数据，例如修改列类型、移动表空间、VACUUM FULL 等操作，会对操作的表加长时间的排他锁，阻塞对该表的并发读写操作，影响用户业务运行。

在线 DDL 只需要修改元数据，不涉及重建数据，操作时长较快，对业务影响较小，提高业务运行的连续性和数据库的可用性。

#### （6）容灾滚动升级

容灾滚动升级基于主实例先提交升级，备实例再提交升级的方式，可以优先使用到新版本特性，并且在特性不兼容或者故障场景下，可以快速切换旧版本的备实例恢复业务。

#### （7）按分片自动升降副本

金融等业务需要提供半数及以上节点数据故障业务快速恢复能力。当一个分片故障了半数及以上节点时，DN 执行写操作会超时，主要是同步备中有节点故障了，无法执行写操作。为了降低分片故障对业务的影响，需要对分片上可用节点进行降副本操作，当检测到故障恢复后，自动触发升副本操作。

#### （8）GaussRecoder 防勒索容灾不丢数

数据库主流的容灾方案是在异地建立容灾集群，异地容灾集群通过同步主集群的日志数据，确保主集群所在数据中心发生故障时能接管服务。但同步和异步复制方案难以同时实现数据库 RPO=0 和高性能。GaussRecoder 容灾方案能够兼顾 RPO=0 和高性能，将数据黑匣子和主集群部署在一起，主集群日志同步写入黑匣子后，业务才返回成功。灾难发生时，容灾集群可以通过远程连接黑匣子获取全量的日志数据，确保 RPO=0。由于主集群和黑匣子距离近，日志同步写入的时延低，不影响主集群的性能。同时，通过软硬结合实现了

三层防御：GaussDB 数据库和 GaussRecorder 双向认证实现防仿冒，利用鲲鹏 TEE 机密计算和 Dorado 机密存储实现了防攻击和防篡改。

#### （9）透明数据加密

透明数据加密（Transparent Data Encryption, TDE）提供表级数据加密存储功能。当用户使用本特性提供的语法创建加密表后，数据库向磁盘写入加密表数据前，会自动将其加密；同时，数据库从磁盘读取加密表数据后，会自动将其解密。可有效解决攻击者绕过数据库认证机制直接读取数据库文件引起静态数据泄露的问题。

#### （10）支持 WDR 诊断报告

WDR(Workload Diagnosis Report)基于两次不同时间点系统的性能快照数据，生成这两个时间点之间的性能表现报表，用于诊断数据库内核的性能故障。基于 SNAPSHOT 的性能基线，从多维度做性能分析，能帮助 DBA 掌握系统负载繁忙程度、各个组件的性能表现及性能瓶颈。SNAPSHOT 也是后续性能问题自诊断和自优化建议的重要数据来源。

## 2. 运维工具建设

### （1）迁移工具

GaussDB 提供了高效、可靠的数据库迁移服务，其中数据复制服务（DRS）和数据库与应用迁移服务（UGO）是两个核心工具。DRS 专注于数据的实时迁移、同步和校验，而 UGO 则专注于异构数据库的对象迁移和语法转换。两者结合可以为企业提供完整的数据库迁移解决方案，涵盖从评估、迁移到校验的全流程，支持多种数据库类型和迁移场景，大幅降低了迁移成本和风险，整体方案见图 6。

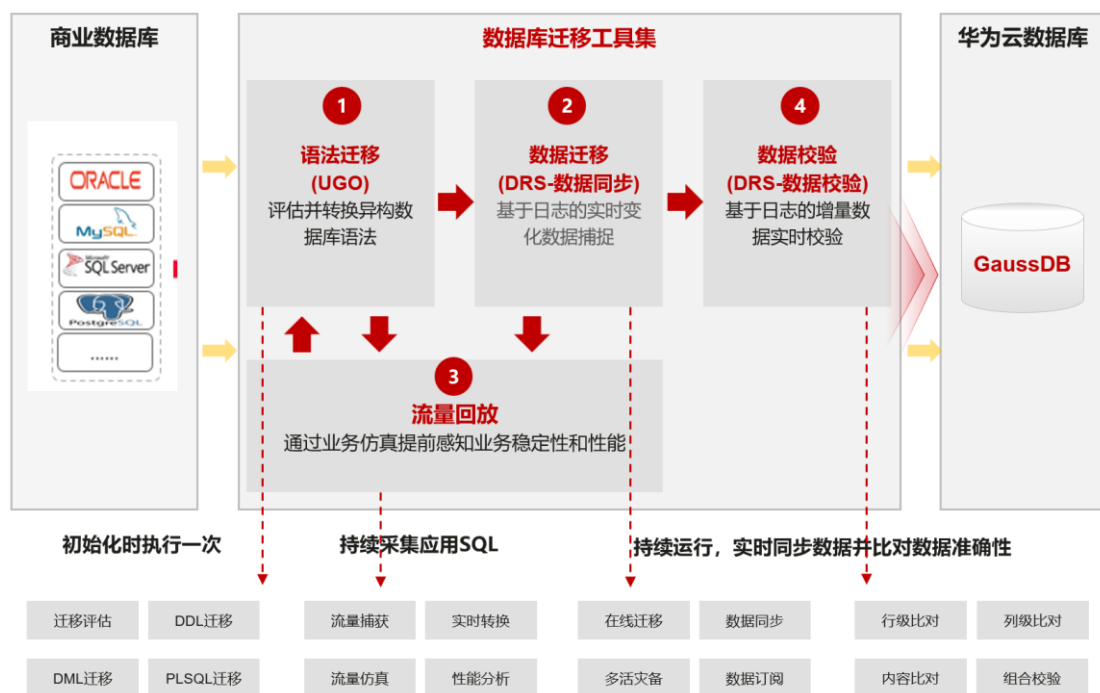


图 6 迁移工具

DRS 在数据迁移方面的支持将数据从多种主流数据源持续拷贝到 GaussDB，并保持数据一致性。迁移、同步过程采用全量+增量迁移模式，在全量数据迁移完成后，增量迁移任务会持续解析和回放源数据库的增量数据，直到追平当前数据。这种模式能够将业务中断时间最小化，实现零停机迁移。同时，DRS 对源数据库的资源消耗极低，进一步减少了对业务性能的影响。DRS 提供了多种数据校验方式，能够快

速、准确地识别数据差异，为用户提供可靠的数据一致性保障。DRS 还具备高性能和高可靠性的特点。通过并行技术加速数据迁移过程，支持迁移过程中源数据库的 DDL 操作，并提供操作审计、时延监控、在线对比等功能，使整个迁移流程 100%透明可控。此外，DRS 任务运行在独立的虚拟机上，租户之间数据隔离，确保了迁移任务的安全性和可靠性。

UGO 主要用于数据库结构迁移和 SQL 语法转换，能够自动化地完成数据库对象的评估、转换和迁移，大幅降低了异构数据库迁移的难度和成本。UGO 的迁移评估功能可以帮助用户在迁移前全面了解迁移工作的难度和投入，可以进行工作量评估和风险识别。UGO 的结构迁移功能负责将源数据库的对象定义（如表结构、索引、视图、存储过程等）转换为目标数据库支持的格式，支持自动化语法转换、对象校正和验证、持续集成支持。

DRS 和 UGO 在数据库迁移过程中各有侧重，两者可以协同工作，形成一个完整的迁移解决方案：UGO 首先对源数据库进行评估和结构迁移，将表结构、存储过程等对象转换为目标数据库支持的格式，在 UGO 完成结构迁移后，DRS 负责进行全量+增量的数据迁移，并在迁移过程中进行数据校验，确保数据一致性。

在实际迁移项目中，DRS 和 UGO 的协同工作流程如下：

- 迁移评估：使用 UGO 对源数据库进行评估，生成迁移工作量报告和风险分析报告。

- 结构迁移：使用 UGO 将源数据库的对象结构转换为目标数据库的语法，并在目标数据库中创建这些对象。

- 数据迁移：使用 DRS 进行全量+增量的数据迁移，将数据从源数据库实时同步到目标数据库。

- 数据校验：使用 DRS 的数据对比功能，校验源数据库和目标数据库的数据一致性。

- 业务切换：在数据一致的前提下，进行业务切换，并将流量导向目标数据库。

- 后期优化：使用 UGO 的 SQL 审核功能对目标数据库中的 SQL 语句进行优化，确保迁移后的性能。

另一方面，DRS 流量录制回放支持数据库迁移前的兼容性与性能验证、异构数据库对象迁移验证、数据库大版本升级验证、数据库重大问题复现与定位，是一个强大的数据库业务仿真验证工具。通过捕获真实的生产 SQL 流量，并在目标数据库上模拟执行，为用户在数据库迁移、升级和重大变更前提供了至关重要的测试验证环节，能有效暴露兼容性问题、性能隐患和功能差异。

通过 DRS 和 UGO 的协同工作，可以大大降低数据库迁移的难度和风险，实现高效、可靠的迁移上云。无论是同构迁移还是异构迁移，无论是数据迁移还是对象迁移，GaussDB 都提供了相应的工具和服务来满足用户的需求。

## (2) 自动化运维平台

GaussDB 提供图形化界面的统一管控运维平台，满足可观测、智能化的运维目标，如图 7 所示：



图 7 运维管控平台

核心功能包括：1) 实例生命周期运维：一键部署与扩缩容、参数配置、备份恢复等。2) 容灾管理：支持手动/自动的主备切换、容灾切换演练，及容灾监控。3) 日志安全：日志下载、日志分析、审计等。4) 智能运维：资源环境系统或者数据库系统的基础运维以及 AI4DB 智能运维能力。以及管理员拥有的平台运维相关能力。

智能运维从功能视角主要包括监控、告警、诊断优化以及会话管理，主要功能能力：

**监控大盘：**集成丰富的监控面板，实时展示集群健康状况、性能指标（TPS/QPS、延迟、连接数、资源利用率等），并支持灵活配置告警阈值和通知策略。

**智能巡检：**GaussDB 支持 SQL 类、功能类、指标类等近百项指标巡检项，同时支持自定义预警类巡检项，智能巡检和报告满足用户对历史的巡检，评估未来的风险。另一方面，预防性维护是保障业务连续性和数据安全的关键。趋势预测

功能通过对核心资源指标进行智能化增长趋势分析与预测，帮助运维团队提前识别潜在风险，实现从被动响应到主动预防的运维模式转变。趋势预测功能针对数据库大小、服务器内存使用率、磁盘占用率及动态内存使用情况等关键指标，基于历史数据构建预测模型，生成未来一段时间内的指标变化曲线。通过将 AI 预测技术与运维实践深度融合，为金融行业提供了一套符合监管要求的风险预判机制，建议用户结合自身业务特点，在容量规划、性能优化和应急演练等场景中深化应用。

**SQL 诊断功能：**对于安全、稳定和性能要求极高的金融行业核心系统，慢 SQL 诊断旨在自动分析 SQL 执行性能瓶颈的根本原因，并给出优化建议。慢 SQL 的一站式诊断调优功能，支持从内核视图和持久化数据中对过去时间内的慢 SQL 查询，获得慢 SQL 的执行时间、具体文本、执行计划等详细信息。同时可以根据唯一 SQL ID 下钻直接进行调优，该功能通过融合慢 SQL 关键指标，构建特征模型，推算慢 SQL 根因，并执行相应的优化，比如 SQL 限流、索引优化、SQL PATCH 等。帮助金融运维人员，一方面快速定位与优化，自动化的根因分析大大缩短传统 DBA 手动分析慢 SQL 的时间；另一方面，提供智能决策支持，提供数据驱动的优化建议，辅助运维人员制定更有效的数据库优化策略。

**索引推荐：**对金融行业而言，数据库查询性能直接影响交易效率与用户体验，索引是影响性能的关键因素之一。GaussDB 的索引推荐功能是一项基于 AI 的智能优化能力，

旨在自动分析 SQL 语句执行计划，识别性能瓶颈，并推荐最优索引策略以提升查询效率。该功能通过机器学习算法对数据库负载进行深度分析，综合考虑查询模式、数据分布和系统资源等因素，一方面，显著降低 SQL 执行时间，提升联机交易处理效率。另一方面，智能识别并建议清理冗余索引，减少无效存储空间。同时，改变依赖人工经验的传统索引优化模式，通过自动化分析降低对资深 DBA 的依赖，提升运维效率，尤其适用于业务规模大、SQL 语句复杂的场景。

### （3） AI 运维助手 GaussDB Doer

随着业务规模的扩大和数据量的激增，数据库运维的复杂性也随之增加。传统运维方法在确保数据高可用性、系统稳定性、性能优化及故障快速响应方面，面临着诸多挑战，往往需要投入大量人力和时间成本，且效果依赖于运维人员的个人经验和直觉。在此背景下，智能运维技术，特别是基于大模型的解决方案，为数据库运维带来了新的突破。

大语言模型（LLM, Large Language Model）是基于海量的通用语料进行数据训练而成的超深度学习神经网络，包含数十亿甚至更多参数，模型规模可以达到数百 GB 甚至更大，通过对大规模数据集进行训练，展现出强大的表达和学习能力。大语言模型在数据库运维领域的应用，具有对运维知识理解、分析流程编排以及工具调用的能力及优势。GaussDB 基于盘古大模型构建数据库从咨询、开发、优化、运维等阶段的各种 AI 应用，如图 8 所示。

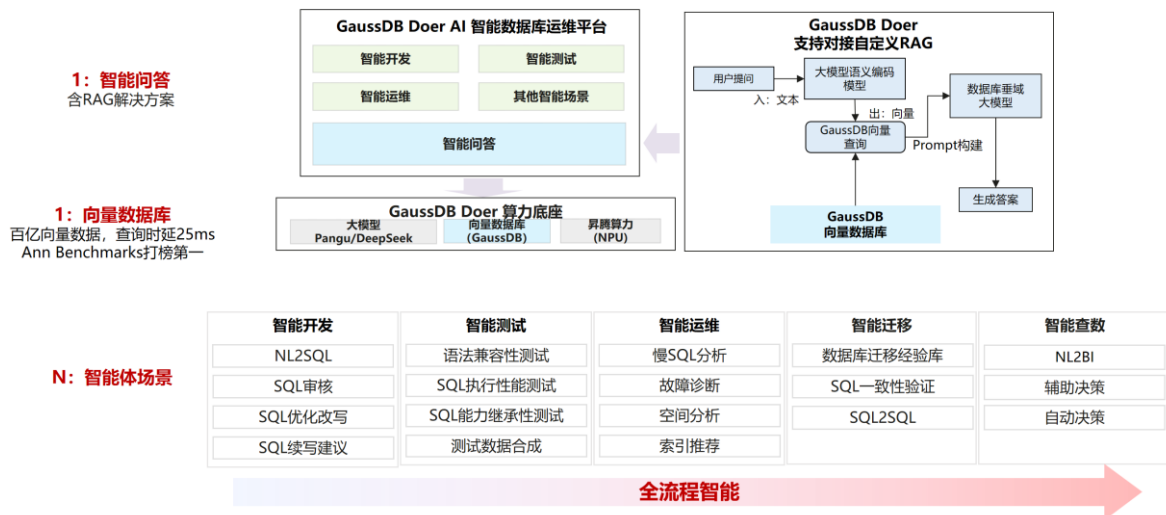


图 8 GaussDB Doer 智能助手

**智能问答：** GaussDB Doer 基于 RAG（Retrieval-Augmented Generation，检索增强生成）和 LLM 技术构建的智能知识问答能力，通过整合产品文档、社区资料、使用手册和运维手册等私域资料，运用自然语言处理和知识检索技术，GaussDB 智能问答助手能理解并解答用户关于 GaussDB 的各类问题，帮助用户快速找到解决方案，在安装配置、性能优化、操作使用以及故障排查方面，提供精确有效、交互式的回答。

**智能体场景：** GaussDB Doer 以 LLM 为决策大脑，按照案例知识召回、故障分析计划编排、故障分析执行、故障根因总结等步骤，结合运维工具集进行工作。一方面，通过 AI 算法构建了系列异常检测、慢 SQL 诊断、集群故障根因诊断等关键运维能力。通过构建运维工具 API 集合，在智能体分析故障流程中，实现不同故障运维场景下大模型能够调用合适的工具。另一方面，根据故障现象，通过相似度搜索，故障智能体在运维案例记忆中召回相关运维案例，组装成提示

词，输入到大模型进行分析流程编排。对于不在预设 TOP 场景的故障类型，则通过大模型会根据现有案例以及运维知识，推理生成规划，形成对未知故障的学习能力，同时调用 Function call 能力，从工具集当中选择各个步骤的所需的工具参数，以及对应的工具接口 API，生成的结果可以通过 restful/Dubbo 等接口协议进行执行，并且获取接口调用结果。最后，经过完整的计划编排以及分析执行后，这些步骤的结果将会汇总输入到 LLM 输出根因定位总结报告。在用户界面上，展示完整的分析流程以及对应的根因定位总结报告。

通过以上流程，可以减少运维人员分析故障的复杂度，节省故障定位时间，同时能够直观展示分析流程以及总结报告。

### 三、 GaussDB 典型运维场景与实践案例

#### （一） 高可用与容灾运维

在金融业务数字化转型深化的今天，数据库的持续稳定运行已成为维系客户信任与市场声誉的生命线。任何计划外的服务中断都可能引发链式反应，导致重大的财务损失与运营风险。因此，构建跨越单点、单机房、单地域的多层次高可用与容灾体系，已不再是技术选项，而是金融行业的刚性需求与合规基准。

#### 1. 案例：某行采用 GaussDB 流式容灾一拖二部署方案

##### （1） 背景与挑战

该行核心业务系统在国产数据库改造过程中，核心诉求是：1) 满足算力和容量诉求，算力规模达到 18.6 万 MIPS。

但原选型的数据库架构资源消耗严重：只支持物理机部署，相同算力下要求下消耗近 7000 台物理机。2) 业务不改造诉求。原选型数据库若用 NvmeSSD，单库不得超过 2T，超过 2T 需要分布式改造；且 Oracle 语法兼容差，迁移过程中应用改造工作量大。3) 满足金融级高可用诉求。极端环境下继续对外提供服务，RTO、RPO 要求极高

## (2) 解决方案

GaussDB 内核高度兼容 Oracle 语法，迁移过程配套 UGO 语法评估转换工具，一键式进行语法转换。

该银行核心业务系统根据容灾等级和数据量要求采用 GaussDB 集中式一主两备 BMS+Dorado 存算分离架构，单库支持 16T 以上数据量，只需普通的 SSD 性能可达万级 TPS，无需购买昂贵的 NvmeSSD，且可避免过早引入分布式改造工作量。以三集群方式部署在两个城市，如图 9 所示：

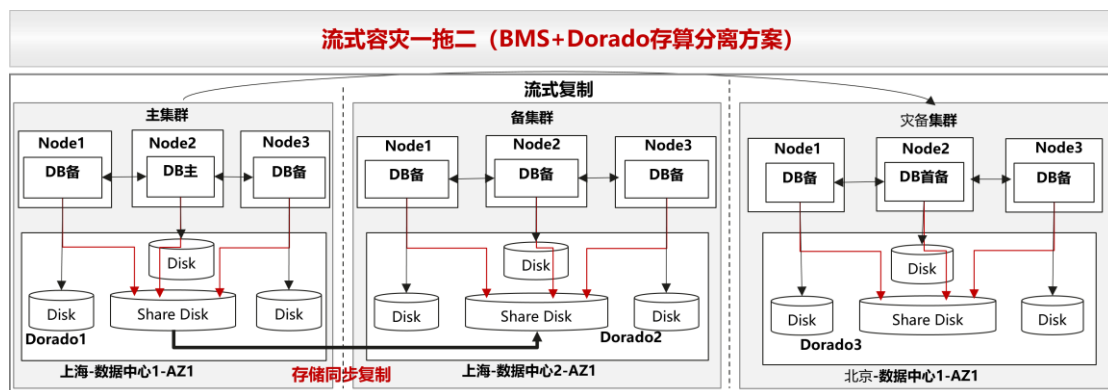


图 9 流式容灾组网

三个 GaussDB 集群搭建容灾关系：同城 Region 的两个 GaussDB 实例搭建 Dorado 容灾关系，异地跨云的一个 GaussDB 实例与同城的主实例搭建流式容灾。

同城双集群与异地容灾集群部署架构都采用了计算节点和存储节点分离的部署方式，计算节点使用 BMS，存储节点使用 Dorado。此架构下，容灾搭建时，同城主备集群之间使用全量 Build 同步数据。搭建完成后，同城主备集群通过 Dorado 存储增量同步数据，异地集群通过流式增量复制数据。DN 备节点故障，不中断业务的进行；DN 主节点故障，自动进行主备切换。当同城双中心出现自然灾害等原因发生故障时，异地灾备集群可以通过升主操作进行业务高可用保障。

### （3） 成果效益和创新

基于流式复制技术的多集群，提供两地三中心部署，同城双集群 RPO=0、RTO<=120s，同城双集群与异地容灾集群的 RPO<=10、RTO<=120s，在双集群的容灾能力的基础上进一步保障了主库业务的容灾能力，满足监管要求。主备集群间异步传输日志，不影响主集群业务提交。1 拖 2 其实是星型部署，可扩展多个备，即使出现单备机故障降级为双集群，也能保证集群级的高可用。

### （二） 运维效率提升

传统的数据库运维高度依赖资深 DBA 的经验，在面临突发故障、性能劣化等复杂场景时，往往响应滞后、定位耗时，甚至可能因人工误操作引发二次风险。

GaussDB 作为国产金融级分布式数据库，不仅提供了卓越的性能和高可用特性，更在其生态中引入了以 GaussDB Doer 为代表的 AI 智能运维助手，为上述痛点提供了全新的解决方案。本章节将结合两个在金融客户现场真实发生的联

创案例，深入探讨 AI 智能运维如何将运维效率从“人驱”转变为“智驱”，旨在为各位同行提供可参考、可落地的实践思路。

## 1. 案例：某行业务高峰期 CPU 瞬时飙升故障的自动定位与自愈

### (1) 背景与挑战：

某全国性商业银行的认证授权网关系统，在某个工作日上午的业务高峰期，业务访问量出现瞬时激增。数据库监控系统随即告警：数据库连接数飙升超过两倍，CPU 使用率瞬间打满至 100%。该情况导致部分业务请求响应缓慢，用户体验受到直接影响。传统运维模式下，DBA 团队需紧急介入，手动抓取现场信息（如连接来源、正在执行的 SQL、系统视图等），排查过程如同“大海捞针”，故障恢复时间（MTTR）存在巨大不确定性。

### (2) 传统处理方式与痛点：

**手动排查耗时：** DBA 需要分别检查活跃会话、慢查询日志、操作系统资源监控等，流程繁琐。

**信息关联性差：** 连接数暴增和 CPU 飙升可能是果，其背后的因可能是低效 SQL、锁等待、资源争用或外部工具异常，难以快速建立关联分析。

**现场难以保留：** 故障瞬间的状态可能稍纵即逝，若未能及时捕获有效信息，事后定位将极其困难。

### (3) GaussDB Doer 的智能运维实践：

在与该银行的联创项目中，我们利用 GaussDB Doer 对此次故障场景进行了仿真复现。Doer 并非单一工具，而是一个多智能体协同系统。在故障触发后：

**智能体协同诊断：** Doer 内部的日志分析智能体、指标分析智能体、SQL 分析智能体等被自动激活。

**根因快速定位：** 智能体们并行工作，在数分钟内完成了对海量监控指标和日志信息的关联分析，成功捕捉到在业务连接激增的同时，数据库内部发生了大量配置重载（reload）操作。正是这两个事件的并发，导致了 CPU 资源的彻底耗尽。

**报告输出与自愈：** Doer 自动生成了结构化的根因分析报告，清晰展示了从现象->分析->结论的完整证据链，与事后人工复盘结论完全一致。更进一步，基于此根因，可以预设自动化应对策略，未来若再次出现相同特征的并发事件，系统可自动执行限流或调度策略，实现故障自愈，避免对业务造成影响。

#### （4） 客户价值：

**效率提升：** 将可能需要数小时的人工排查过程缩短至分钟级。

**准确性保障：** 避免了人工遗漏和误判，结论基于全量数据证据。

**知识沉淀：** 分析报告可作为知识库案例，赋能整个运维团队。

## 2. 案例：某行主备切换与慢 SQL 性能劣化问题的自动化 workflow 诊断

### (1) 背景与挑战：

某商业银行在运维过程中面临两个典型难题：

**主备切换定位困难：** 数据库发生高可用切换后，其背后根因可能涉及网络抖动、硬件故障、操作系统异常或数据库本身问题。排查需要深入比对主备节点的日志、时间线和技术指标，过程极其复杂，难以通过固定脚本覆盖所有场景。

**慢 SQL 性能回溯难题：** 业务反馈慢 SQL 频发，但 DBA 难以回答“为什么这条 SQL 现在变慢了？”——是统计信息过时？执行计划变更？还是资源争用？缺乏自动化工具来对比 SQL 执行计划在不同时间点的细微差异。

### (2) 传统处理方式与痛点：

**高度依赖专家：** 两个问题均需专家级 DBA 投入大量时间深度挖掘。

**脚本开发滞后：** 为特定场景编写排查脚本开发周期长（1-2 个月），且通用性差，无法应对千变万化的实际故障。

**无法快速响应：** 业务迫切希望得到答案时，运维团队往往因流程长、手段有限而承受巨大压力。

### (3) GaussDB Doer 的智能运维实践：

GaussDB Doer 的创新性在于其提供了“自然语言编译 workflow”的泛化能力。

**自然语言驱动：** 面对主备切换难题，运维工程师无需编码，只需通过自然语言向 Doer 描述意图：“请分析最近一次

主备切换的根本原因，对比主备节点的日志和关键监控指标，输出一份分析报告。”

**workflow自动编译与执行：** Doer 内置的 GaussMaster 模块将此需求自动编译成一个可执行的工作流。该工作流会自动执行日志采集、时间轴对齐、错误模式识别、指标比对等一系列复杂操作，最终生成一份详尽的根因分析报告。

**敏捷开发新场景：** 对于慢 SQL 性能对比场景，同样通过自然语言描述：“对比某条关键 SQL 在故障时段和正常时段的执行计划差异，并检查相关表统计信息是否更新。” Doer 可快速编译并执行一个包含执行计划提取、差异比对、统计信息验证等节点的工作流。开发周期从传统的“月级”缩短至“天级”甚至“小时级”，实现了运维场景的快速覆盖和敏捷响应。

#### （4） 客户价值：

**降低技术门槛：** 普通运维人员也可处理复杂故障，释放专家资源。

**极致敏捷：** 新故障场景的应对策略实现“按需随建”，极大提升了运维团队的响应能力和自信心。

**标准化输出：** 所有诊断过程通过标准化工作流完成，输出格式统一，便于管理和知识传承。

金融运维的未来必将是智能化的。拥抱像 GaussDB Doer 这样的 AI 助手，不仅能提升效率和稳定性，更能让您的运维团队从繁重的重复劳动中解脱出来，聚焦于更具价值的架构优化和业务创新之中。

### （三）容量管理运维

#### 1. 案例：某行统一查询数据库扩容

##### （1）背景与挑战

某行经历三代核心系统，积累了大量的历史数据，目前已经上线新一代核心系统，历史数据呈现加速增长态势。在银行不断能力开放和数字化转型的大背景下，用户对查询历史明细时间跨度和服务体验要求也日益提高，统一查询系统配套新核心建设，实现历史数据查询功能，提升邮储银行个人客户查询服务体验。

该银行原本大规模使用开源 PG 数据库，但 PG 数据库单节点容量偏小，清理机制导致 IO 抖动，切换速度慢等一系列问题，传统 IT 架构已无法满足需求。该银行从 2019 年开始，从架构上进行创新性变革，通过先进的分布式技术驱动核心系统升级换代，同时沉淀关键技术构筑全行级领先信创服务能力。

##### （2）扩容实践

采用“新建集群，DRS 数据迁移”的方式，即新建扩容集群，接入历史明细和登记簿集群，其中单集群在区域 A 增加 64+1 台机器（DC1 32 台，DC2 32 台，仲裁节点 1 台），区域 B 增加 32 台机器，再基于 DRS 数据迁移服务将数据从现有集群同步至新扩容集群，同步完成后执行业务切换，完成集群扩容。

迁移时长评估：以历史明细集群为例，32 个分片节点同时并发进行数据迁移，每 DN 每小时迁移 150GB，4 小时迁

移 600G, 32DN\*600G=18T, 迁移实践窗内集群总使用容量达到约 280T, 裸数据约 154T, 整个系统预计 15 天完成扩容。

迁移方案：对原系统数据采用全量数据+增量数据迁移的策略。使用 DRS 数据复制服务。对于全量数据迁移，通过 select 的方式将源库的全量数据导出；调用 GaussDB 的 copy 接口批量写入目标库；可对单表根据主键进行进一步分片，每个分片作为独立的同步单元，并发执行同步效率 100MB/s。对于增量数据迁移，通过解析数据库日志，DRS 将抓取到的数据落盘存储，读取落盘数据，重构成对应的 sql 语句在目标库回放。

在数据迁移的过程中，主要遇到的问题有三个：

- 1) 迁移配置需进行优化。
- 2) 增量迁移表需要有唯一索引。
- 3) flink 实时业务报错且存在数据消费积压。

针对第一个问题：在任务启动前，采取对正向任务进行分片优化，检查运维面 ops 参数，配置任务规格参数等方式进行性能优化；在全量同步阶段，停止暂停、重启模块、修改限流条件等操作；在增量同步阶段，创建索引前暂停所有子任务的增量回放进程，任务进入增量同步阶段后，配置结束任务（结束任务前需要配置）需要的连接参数。

针对第二个问题：通过增加主键或唯一键的表；纯技术备份表进行 DROP 等方式进行解决。

针对第三个问题，通过 DN 主备切换清理 DN 上的异常线程进行解决。

### (3) 成果效益和创新：

统一查询是新一代核心系统，采用 GaussDB 分布式架构承载全行级交易明细、收支分析和登记簿业务，构筑了创新的全行级数据服务中心，为全行 6.5 亿用户、4 万网点、13 万智能机具提供实时十年期详单查询、登记簿及收支分析业务。

分布式架构解决传统数据库海量数据存储和处理的性能瓶颈，具备良好的可扩展性，提升新核心系统数据查询能力。首次投产上线节点 147 节点，扩容工程完成后部署节点超过 300 个，承载数据超过 500TB，足够支撑邮储全行未来三年内的业务增量，是目前中国金融行业最大的同城双活集群。

统一查询系统上线充分体现了 GaussDB 分布式架构的高并发、高性能、大容量承载能力，同时提供大数据量在线平滑迁移能力，相关架构设计及工程实践为金融行业信建设提供了重要参考案例。

主要的创新点有：

1) 数据库和业务均采用两地三中心同城 A-A 双活部署，提升了硬件资源使用率和业务并发支撑力；

2) 通过在内核实现分布式并行逻辑解码，采用 DRS 进行搬迁，DRS 支持全量和增量搬迁，整个迁移过程无需停业，DRS 自动追增量；

3) 迁移过程中出现任何问题，都可回退，原实例数据不受损，避免在原集群上扩容遇到问题影响业务，实现真正的平滑扩容；

4) DRS 支持流控，迁移时间段可控制，避免在业务高峰期迁移，严格控制数据搬迁速度，实现对业务无感迁移；

5) 实现了海量数据的一致性比对机制，提升比对效率，确保迁移完数据一致性。

6) 采用自适应 Upsert 机制处理全量数据与增量数据的冲突，使冲突解决性能提升 20+倍。

#### (四) 巡检

##### 1. 案例：某行深度使用巡检能力加强风险管理

###### (1) 背景与挑战

某行核心信贷系统及多个重要应用已全面上线 GaussDB。系统承载着日均数千万笔的交易量，涉及资金量巨大，对数据库的稳定性要求达到了最高级别。初期，运维团队面临巨大压力：人工巡检效率低、覆盖不全、无法深入分析，疲于应对各种“救火”式故障。

为此，该行采用 GaussDB 管控提供的一套“日常巡检+深度巡检”的双层防御工具，主动发现和解决潜在问题。

###### (2) 关键实践

日常巡检：日常巡检如同系统的定期全身体检，目标是快速发现并处置显性的、即时的风险。该行制定了标准的巡检计划和要求：在周一凌晨业务低峰期，开展每周一次的“全身体检”；通过 GaussDB 管控平台定时触发巡检任务，生成标准化的巡检报告。

通过巡检分析发现潜在问题的场景比如：1) 集群状态与关键进程检查：巡检工具首先自动核查集群状态是否为

“Normal”，并检测各 CN、DN 进程的启动时间。一次巡检中，系统发现某 DN 进程虽运行正常，但其连续运行时间远超其他节点，疑似存在过载后重启但未告警的历史情况，避免了“静默性故障”。团队据此深入排查，发现该节点所在物理机存在间歇性内存故障，及时进行了硬件更换，消除了一个重大隐患。

2) 关键性能指标洞察：巡检报告重点监控 80%/95% SQL 响应时间、在线会话率、线程池使用率及最长事务时间。某次巡检后，报告提示“95% SQL 响应时间”较上周同比上升了 35%。团队立即锁定该时间段内上线的一个新业务功能，其一条查询语句缺乏有效索引，导致全表扫描。在周末业务高峰来临前，快速完成了索引添加，避免了可能发生的系统性性能瓶颈，保障了客户体验。

3) 空间与日志管理：自动化检查数据库大小增长率、xlog 数量和未落盘脏页数量。系统曾预警某业务库 xlog 生成速率异常飙升，是平日的 2 倍。经分析，是某个批量代收付业务因上游系统重试机制缺陷，在短时间内发起了多轮重复操作。不仅提前预警了磁盘空间爆满的风险，更间接发现了业务逻辑的漏洞，促进了业务系统的优化，体现了运维对业务的反哺价值。

**深度巡检：**深度巡检是在日常体检基础上的专家会诊，侧重于性能趋势分析、瓶颈根因定位及潜在风险挖掘。每月例行一次或重大业务上线前、重大护网保障前的“专家会诊”。DBA 专家结合工具输出进行深度分析。

通过深度巡检回溯历史隐患的场景例如：1) 通过 WDR 性能报告深度分析：每月初，系统会自动生成上个月的 WDR

报告。该行 DBA 团队会重点分析报告中的 SQL 响应时间波动趋势、TOP 等待事件、Buffer 命中率和关键资源（CPU、内存、I/O）的使用趋势。在一次重大促销活动上线前的深度巡检中，WDR 报告预测当前负载下，业务峰值时线程池可能耗尽。团队提前与开发团队协同，优化了连接池配置，并进行了弹性扩容，确保了促销活动期间数据库的平稳运行。2）错误日志挖掘分析：工具会自动聚合扫描周期内的所有 ERROR 及以上级别的日志，按错误类型、发生节点、频率进行归类。某次月度深度巡检中，发现大量“死锁超时”错误，但因其未达到日常告警阈值而被忽略。通过对这些“轻微”错误的集中分析，定位到两个高频转账账户间存在非最优的业务处理顺序。协同开发修改代码后，基本消除了该类死锁，提升了系统吞吐量。3）系统资源健康度前瞻性判断：深度分析数据盘读写花费时间（ms）的时序变化趋势。通过数月的趋势对比，发现某台宿主机磁盘的读取延迟正在缓慢但持续地升高。判断出该磁盘寿命即将到期，在其未出现硬件故障告警前，就制定了计划内的数据迁移和磁盘更换窗口，实现了一次“无感”的硬件维护，避免了计划外的宕机风险。

### （3） 价值效益

该行的实践充分证明，一套嵌入到运维流程中的自动化的体系化巡检，能够带来显著价值：

从被动到主动：将运维模式从“救火”转变为“防火”，极大降低了突发故障率和应急处理成本。

从模糊到精确：基于数据的巡检报告，使决策更加科学，资源优化和容量规划更有依据。

从孤立到协同：巡检发现的问题促进了运维与开发、基础设施团队的深度协同，共同提升系统健壮性。

通过构建巡检体系，能为 GaussDB 稳定运行筑牢坚实基础，从容应对金融业务发展的各项挑战。

## （五）备份恢复

在金融行业，数据是毋庸置疑的核心资产。数据库的高可用架构（如主备切换）能够应对硬件或网络层面的故障，但无法防范诸如误删除数据、恶意篡改、应用逻辑缺陷等“软性”灾难。一旦发生此类问题，高可用架构会将这些错误操作同步到所有副本，导致数据错误“传染”全域。

因此，备份与恢复体系是保障金融数据安全的最后一道，也是最至关重要的一道防线。它不仅是技术实现，更是一种风险管理策略。本章将分享某大型商业银行（以下简称“某行”）在 GaussDB 生产环境中的备份恢复实践案例，为同业提供一份经过实战检验的参考范本。

### 1. 案例：某行新一代核心系统备份系统构建

某行新一代核心系统及其周边多个关键应用（如收支分析、明细查询、登记簿业务）均部署在 GaussDB 分布式集群上，总数据量超过数百 TB。在系统建设初期，团队曾过度依赖 GaussDB 自身的高可用特性，直至在一次内部演练中，模拟了“误删除重要业务表”的场景，才发现仅凭主备架构无法挽回数据损失，恢复过程充满挑战。

此次演练敲响了警钟，某行立即将备份恢复体系的建设提升到最高优先级，与 GaussDB 技术团队合作，制定并落地了一套符合金融监管要求、满足业务连续性目标的备份策略。

### （1） 核心挑战与策略选型

数据量巨大：单业务模块数据量即达 186TB，传统的每日全量备份在时间窗口和存储成本上均不可行。

恢复时间目标（RTO）严苛：核心业务中断不得超过小时级，要求备份系统必须具备快速恢复的能力。

恢复点目标（RPO）逼近零：必须尽可能减少数据丢失，理想状态下应能恢复到故障前任意时间点。

经过深入调研与 POC 测试，某行采纳了 GaussDB 推荐的“全量备份 + 增量备份”组合策略，并针对不同规模数据库进行了差异化配置：

对于大型数据库（>10TB，如核心业务），策略选择每周一次全量备份+每日一次增量备份。全量备份虽耗时较长（约 4 小时），但每周一次可将业务的影响控制在可接受范围内。每日增量备份仅备份 24 小时内的变化数据（约 300GB），耗时短（<30 分钟），存储效率高。此策略在保障数据安全的同时，完美平衡了性能、成本和操作复杂性。

对于中小型数据库（<10TB），策略选择每周一次全量备份+每 30 分钟一次增量备份。数据量较小，增量备份频率可以更高，RPO 可缩短至 30 分钟，更能满足对数据 freshness 要求极高的业务场景（如实时交易核对）。此策略参考了人行等机构的实践。

## （2） 真实场景下的实战价值：

某周四下午，某行开发人员在测试环境进行数据修复脚本验证时，误操作连接到了生产环境的数据库，执行了一条错误的 UPDATE 语句，影响了数十万条客户交易明细记录的状态。

### 紧急响应与恢复过程：

a. 立即止损：运维监控平台第一时间发出“大量数据更新”的告警，DBA 立即介入，终止了该会话，并锁定相关账号。

b. 评估损失：与业务部门共同确认，这是一次影响广泛的数据错误，必须进行恢复。

c. 恢复方案制定：由于错误操作发生在每日增量备份之后，团队决定使用周三的全量备份 + 周四的增量备份进行恢复，即可将数据库恢复到错误操作前的状态。

### d. 执行恢复：

快速拉取周三的全量备份数据。

应用周四生成的增量备份数据。

整个恢复过程耗时约 4.5 小时（与备份时间基本持平，符合预期）。

e. 业务验证：数据恢复完成后，业务与技术团队联合对数据完整性和准确性进行了验证，确认数据已完全恢复正常。

## （3） 价值体现

这次事件虽未造成实际业务损失，却是一次完美的“实战演练”。它充分验证了备份策略的有效性：

**RPO 达成：**数据成功恢复到故障前一刻，实现了 RPO≈0 的目标。

**RTO 可控：**数小时的恢复时间满足了业务部门的期望，将影响降到了最低。

**体系可靠性：**整个备份、存储、恢复流程完全自动化、工具化，避免了人工操作的不确定性，证明了整个备份体系的可靠性。

某行的实践深刻揭示了备份体系的核心价值：它不是成本中心，而是业务连续性的“保险单”。通过构建稳健、高效的备份恢复体系，方能真正实现“心中有数，遇事不慌”，为金融业务的永续运行奠定坚实的数据基石。

#### 四、 运维模式演进与技术发展展望

##### （一） GaussDB 数据库技术智升金融行业运维展望

金融行业数据库运维正经历从“人工密集型”向“智能自治型”的跃迁。华为云 GaussDB 作为金融企业级数据库的代表，通过透明多写、Serverless 模式创新、智能引擎升级、多模态融合、安全合规治理多条技术路径，重新定义金融数据库运维的边界，推动行业从“人工干预”向“自治进化”跨越。

##### 1. 透明多写，实现细粒度弹性、极致性能和高可靠，减少运维干预

GaussDB 通过对等读写、多租户管理重构金融数据库运维的弹性与韧性，以抵御传统数据库资源利用率低、扩容周期长，难以应对互联网型业务（如手机银行、数字支付）的潮汐式流量冲击。

(1) 对等读写：所有节点提供业务读写能力，高负载下性能无忧

3 节点硬件环境下，GaussDB 的性能达到 380 万 tpmC，这意味着差不多能支持 40 万个用户同时进行账户查询操作。GaussDB 高性能数据库稳定运行，降低因性能不足导致的系统崩溃或延迟的业务连续性问题，运维人员无需频繁处理因性能问题导致的故障。

(2) 多租户管理：隔离风险与资源复用

金融机构需同时支持核心交易、互联网业务、数据分析等多类负载，传统单库多租户模式易引发资源竞争与安全风险。GaussDB 数据库内核级多租户隔离，实现“一库多租、租户自治”。租户间资源硬隔离（CPU/内存/QPS 配额控制），避免关键业务被非核心负载抢占。租户级权限管理满足金融数据安全合规要求。

透明多写帮助金融机构实现资源的弹性供给、风险的隔离控制与成本的优化配置，推动运维模式从“静态管理”向“动态治理”转型。

2. 优化执行引擎：满足多模复杂业务，统一运维能力

在金融行业的 TP 业务中，常常伴随着分析型业务，比如金融的在线交易和实时数据分析(风控、即席分析和监管报送)诉求，而实时分析是数据密集型和计算密集型负载，在集中式下对 OLTP 业务造成很大冲击。GaussDB 一体化 HTAP 架构，客户端请求过来后，首先基于代价、轻量化 AI 和规则自动路由到 TP、AP 引擎，业务无需改造，TP 负载无阻塞；

突破列式内存引擎，通过设计缓存池实现实时行列转换技术，支撑 TP 负载性能低影响，TP 提交即刻对 AP 可见，保障分析结论时效性。无需运维人员干预，兼顾性能和实时性，一站式的同时解决用户的 TP 和 AP 业务诉求。未来分布式 HTAP 实现数据、计算密集型实时分析能力的水平扩展，基于分区实现高效剪枝减少无效计算，利用多节点算力消解算力瓶颈。同时基于主机资源管控与调度算法、备机读物理隔离、云原生的计算节点弹性扩展能力，实现混合负载资源 (CPU、内存、连接等) 全链路管控，确保资源利用率最优，运维无忧。

### 3. AI4DB 安全防护：从“被动防御”到“主动免疫”

金融数据安全关乎国计民生，传统安全策略（如防火墙、权限控制）难以应对新型攻击（如 SQL 注入、数据泄露）。GaussDB AI4DB 通过行为建模与异常检测，实现主动安全防护。通过构建用户/应用行为基线：学习合法用户的访问模式（如登录时间、查询频率、数据范围）。检测异常行为：识别异常操作（如非工作时间批量导出数据、越权访问敏感表），触发阻断并启动审计溯源，避免数据泄露风险。

## （二）运维模式发展趋势

在金融行业数据库运维的早期阶段，运维工作主要依赖于命令行界面（CLI）和脚本化操作，这一阶段被称为“黑屏化”运维。DBA 需要凭借深厚的技术积累和丰富的经验，通过手动输入命令或编写脚本完成日常的监控、备份、扩容、故障处理等操作。这种模式运维效率高度依赖个人能力，操

作风险较大，且缺乏标准化流程。尤其是在金融行业，数据库规模庞大、业务连续性要求极高，任何人为失误都可能引发严重的业务中断或数据风险。

为应对黑屏化运维的挑战，数据库运维逐渐进入“白屏化”阶段。这一阶段以图形化运维平台为标志，通过可视化的方式对数据库集群进行统一管理。运维人员可通过界面完成大部分日常操作，显著降低了操作门槛，提高了工作效率。在金融场景中，诸如 GaussDB 所提供的管理控制台（如 DAS、OPS 等）支持一键部署、监报告警、自动备份与恢复、性能分析等功能，极大提升了运维的规范性和可重复性。

更重要的是，白屏化平台推动了从“单点运维”到“关联运维”的演进。平台能够整合多维度数据（如性能指标、日志、拓扑关系等），通过关联分析快速定位问题根因。例如，某交易系统响应延迟上升，运维系统可自动关联分析数据库负载、网络流量以及业务请求分布，快速识别是否因某个 SQL 请求或资源竞争导致。这种模式不仅提升了故障处理的效率，也为系统性优化提供了数据支撑。

此外，白屏化平台逐步引入了一定的自动化能力，如基于规则的风险拦截、弹性扩缩容和智能索引建议等，进一步减少了人为干预。但此时的自动化仍较多依赖于预置规则与专家经验，尚未具备真正的认知与判断能力。

随着人工智能技术特别是大语言模型（LLM）的突破，数据库运维正迈向以“问答式交互”和“AI 辅助决策”为特征的下一代模式。在这一模式下，运维人员可通过自然语言提

出问题，如“为什么昨晚批量作业执行时间异常？”或“如何优化账户查询事务的响应速度？”，系统基于知识库与实时数据自动生成诊断报告或优化建议。

这种模式的核心依赖于融合多源信息（运维日志、性能数据、知识文档、历史故障库）构建的行业知识库，以及基于大模型的语义理解与推理能力。在金融行业这类强监管、高风险需求的场景中，知识库可纳入包括合规要求、业务SLA、架构最佳实践等内容，使运维决策不仅基于技术数据，也符合业务与风控目标。

同时，运维自动化将进一步向“智能化”发展。例如系统可基于历史数据与实时状态预测潜在故障，或自动制定并执行优化方案（如索引调整、资源调度等）。而运维人员角色也逐渐从重复性操作中解放，转向规则制定、策略优化和异常处理等更高价值的工作。

在这一过程中，金融客户可依托如 GaussDB 所提供的 AI 功能平台，结合自身业务数据与运维经验，构建企业内部的智能运维体系。但需注意的是，人工智能目前仍处于辅助阶段，关键决策还需与人的判断相结合，尤其在资金安全与合规方面需设置必要的人工审核机制。

从黑屏到白屏，再到未来的问答与智能化，数据库运维模式的演进体现了技术发展对金融行业运维效率、可靠性和成本的深刻影响。基于 GaussDB 的运维体系正在这一趋势中不断融合工具平台、知识库与 AI 能力，帮助金融用户构建面向未来的运维能力。建议各机构在参考本白皮书的基础上，

结合自身组织特点与业务场景，逐步推进运维模式的升级，打造高效、稳定且具备持续进化能力的数据库运维体系。

### （三）共同挑战与应对建议

#### 1. 数据库新技术应用带来的运维挑战与对策

数据库新技术，比如存算分离、AIGC、多元异构统一纳管（包括部署形态异构、业务形态异构）、安全内生化的技术的应用一方面带来了降本增效与产品升级，另一方面也带来“复杂性上升、合规收紧、安全边界外扩”的运维挑战。应对方向是统一纳管、标准化与自动化、智能可观测与安全闭环。

以金融行业视角和 GaussDB 数据库视角详细展开分析见表 3、表 4：

表 3 金融行业视角的运维挑战与对策

挑战	典型表现	对策
部署形态异构与运维断层	数百到数千套，集中式/分布式，国产/开源并存，日志、监控、链路追踪与运维割裂	统一纳管与DBPaaS：一套平台覆盖多租户、多租户；策略/告警/变更统一编排
业务形态异构与运维断层	事务型、分析型业务融合；基于数据构建智能应用的需求正在得到快速推广，图、文本、视频、音频等结构化、半结构化、非结构化数据融合	数据库选型时兼顾HTAP、向量等融合能力；数据库能支撑好AI创新性业务
信创迁移与适配风险高	性能验证、异构同步与上线风险高	标准化迁移流水线：压测、异构同步与校验、灰度与回滚、演练常态化
分布式架构复杂定位难	分片、副本、一致性与全局事务带来定位与恢复难	全局一致性保障：基于分布式软快照的数据管理，统一备份/恢复/回滚；透明多写三层池化技术，业务无需分

		布式改造
监报告警噪声大	指标碎片化、阈值难统一、误报多	智能可观测：关联分析、异常检测、合理调整告警阈值、告警降噪、根因定位与自愈
安全与合规愈加严格	权限越界、操作不可追溯、变更高风险	细粒度权限与审计：最小权限、操作留痕、高危拦截、合规基线与工单闭环 软硬协同安全防护：软件、硬件协同治理
DBA获取难与人力成本高	一专多能要求高、人工操作占比高	平台化 + 知识库：专家经验沉淀、模板化操作、开展技能培训与认证

表 4 GaussDB 视角的运维挑战与对策

挑战	典型表现	对策
工具对运维支撑的覆盖度不足	不同行业、企业需要定制告警项；只能使用限定的方式开展备份恢复	可编排的运维原子能力：提供灵活的运维选择，例如支持自定义告警项，支持指定节点或指定AZ备份。
优化器泛化能力不足，业务场景无法完全覆盖	不同行业、不同业务数据访问模型、数据存储模型差异较大，优化方法和经验无法直接复用	智能优化器：内置AI算法，提供更准确的行数估算、代价模型、路径搜索等能力。
运维的自动化与可观测性不足	大部分操作依赖人工判断，智能化运维效果难以衡量	可观测性与 AI Ops：提供异常检测、根因分析与自愈等能力，开放关键指标与日志，用户可以基于开放的数据构建自己内部的运维模型
数据向量化、多模化	一套数据库需要同时计算、存储不同类型的数据	全数据类型支持：文本、视频、图片、json等数据类型向量转化，支持不同模态混合查询。 内置大模型语义算子：使用SQL即可无缝接入RAG框架，助力企业快速构建AI应用

容灾组网场景多样性	满足金融监管不断提升的容灾要求,多地多中心、多层灾备等容灾组网形态,以及对接Dorado/物理机等不同设备的容灾方案不断涌现	支持多地多中心容灾形态,可灵活编排的容灾架构;支持容灾链路策略控制,灵活选择同步模式、同步时延及容灾链路的启停与限速
安全合规风险闭环	权限与审计覆盖不足、高危操作风险高	内置合规基线:最小权限、操作审计、高危拦截、策略模板与自动化巡检
经验可沉淀可复制	大型企业期望提供平台化+方法论的一体化运维体系	产品即服务:平台+工具+培训+SLA,联合治理与行业标准输出

## 2. 金融行业数据库实现自治运维的挑战与应对建议

### (1) 金融行业数据库运维的共同挑战

尽管各金融机构的业务规模、系统架构不尽相同,但在数据库运维层面,尤其是在追求极致稳定、高性能和高可用的核心生产环境中,面临着高度相似的挑战:

系统复杂性指数级增长,运维响应速度要求极高。分布式架构的普及、微服务化的改造,使得数据库实例数量激增,拓扑关系复杂;任何单点故障(如案例中的CPU飙升、主备切换)都可能被迅速放大成系统级故障,要求运维团队必须在分钟级甚至秒级内完成定位与响应,传统依赖人工排查的方式已难以为继。

故障根因定位深度依赖专家经验,知识难以沉淀和复制。如案例所示,复杂的故障往往是多个事件并发、连锁反应的结果,其排查需要DBA具备深厚的知识储备和丰富的实战经验。然而,资深专家资源有限,其经验难以标准化、工具化地复制给整个团队,导致运维能力瓶颈突出,且存在人员流动带来的知识流失风险。

运维自动化场景覆盖不足，“脚本海洋”难以维护。为提升效率，运维团队编写大量自动化脚本，但这些脚本往往针对特定场景、固化特定逻辑（如案例二中客户面临的困境），泛化能力差。当业务变化或出现新故障场景时，脚本需要大量修改甚至重写，开发周期长、维护成本高昂，形成了新的“自动化债务”。

保障业务连续性的压力巨大，试错成本极高。金融系统 7x24 小时服务的特性意味着任何运维操作都如履薄冰。尤其是在进行性能优化、故障处理时，缺乏有效的仿真预演和精准影响分析能力，使得运维人员不敢轻易尝试，往往选择保守但低效的方案，从而错失了通过优化提升系统稳定性和性能的机会。

## （2） 应对挑战的智能化路径与建议

面对上述挑战，拥抱智能化已成为提升运维效率、保障金融业务稳健运行的必然选择。基于 GaussDB 的智能运维实践，有以下几点建议：

**战略规划层面：**将智能运维纳入企业数字化智能化转型的核心组成部分。决策层应充分认识到，智能运维不再是锦上添花的工具，而是保障核心业务系统稳定、高效运行的“数字基础设施”。应制定长期的智能运维演进路线图，从资源投入和组织保障上给予充分支持。

**技术选型层面：**选择原生集成智能运维能力的数据库平台。选择一款将 AI 能力内生于数据库内核及其运维生态的平台至关重要。这确保了运维功能能直接利用数据库的内部

指标和日志，实现更深度的洞察和更精准的控制，避免后期集成各种独立工具带来的数据隔阂和运维复杂度。

实施路径层面：采用“由点及面、循序渐进”的落地策略。

初期：从效率提升最显著的场景切入。优先将 AI 助手应用于告警降噪与聚合、慢 SQL 根因自动分析、快速健康诊断等高频、高耗时的场景，快速展现价值，建立团队信心。

中期：构建场景化自动化 workflow。借鉴如 GaussDB Doer 的“自然语言编译 workflow”理念，鼓励运维团队将重复性的复杂排查动作（如案例中的 CPU 飙升分析、主备切换分析）沉淀为可复用的自动化 workflow 模板，逐步形成企业自身的智能运维知识库和“数字员工”。

长期：迈向预测性与自治性运维。在积累了足够的数据和信任后，逐步利用 AI 的预测能力，实现对容量风险、性能劣化趋势的提前预见，并最终在部分场景实现无需人工干预的自优化、自修复、自安全的自治运维目标。

组织与文化层面：培养“人机协同”的新型运维模式与文化。智能运维并非取代 DBA，而是将其从繁琐重复的劳动中解放出来。团队需要转型，学习如何向 AI“提问”、如何解读 AI 产出的报告、如何决策 AI 提供的建议方案。这意味着需要培养既懂数据库技术又具备数据思维又懂 AI 的新型运维人才，构建人机协同、相互赋能的新型运维体系。

展望未来，金融行业正在从数字化向智能化转型的关键时期，数据量、交易量、系统复杂度必将持续增长，数据价

值更加凸显，给数据库运维带来的挑战只会加剧，不会消失。唯有主动拥抱变化，利用如 GaussDB 这样融合 AI 能力的先进数据库技术，将运维工作从一门“艺术”转变为基于数据的“科学”，才能在未来激烈的竞争中，构筑起坚实、高效、可靠的数据基石。