

数据库及应用系统迁移 指南

(2020 年)

中国信息通信研究院云计算与大数据研究所
2020 年 7 月

版权声明

本白皮书版权属于中国信息通信研究院云计算与大数据研究所，并受法律保护。转载、摘编或利用其它方式使用本研究报告文字或者观点的，应注明“来源：中国信息通信研究院云计算与大数据研究所”。违反上述声明者，本院将追究其相关法律责任。

牵头编写单位及人员：

中国信息通信研究院云计算与大数据研究所：刘思源、魏凯、姜春宇、马鹏玮、李俊逸、田稼丰

参与编写单位：

阿里云计算有限公司：樊文凯、杨霖、王佳毅

云和恩墨（北京）信息技术有限公司：盖国强、李轶楠、杨廷琨、黄宸宁

中兴通讯股份有限公司：朱业、张玲东、李磊、李强周、刘勇

天津南大通用数据技术股份有限公司：李辉、高燕崧、李世辉、刘辉

迪思杰（北京）数据管理技术有限公司：方军、刘磊

北京海量数据技术股份有限公司：黄晓涛、张超、舒正英、林慧莹

上海新炬网络信息技术股份有限公司：程永新、梁铭图、杨光

北京金山云网络技术有限公司：赵景波、蒋琪、王嘉楠

北京华胜信泰数据技术有限公司：刘淼、李正效、杨光

上海爱可生信息技术股份有限公司：陈书俊、黄炎、路新英、徐爽

PostgreSQL 中文社区：胡辉

前 言

随着数据库技术与高速网络通信、云计算、人工智能等信息技术的不断深化融合，数据库和应用系统的存在形式愈发丰富，数据库之间的相互迁移已成为常态化的趋势。从同构到异构数据库之间的迁移，从数据到数据及应用系统的迁移，从本地间到上云迁移，迁移的表现方式不一而足。

不论是对数据库的架构改造，还是数据库上云搬迁，亦或是数据库产品变更，都涉及数据库及应用系统的迁移。目前，针对此领域的指导性工程方法、权威交付流程和专业评价体系的发展仍在起步阶段。

本指南首先介绍了数据库及应用系统迁移过程中的痛点和风险，然后梳理了迁移交付过程中常见的流程指南，并对企业迁移能力的评价进行多维度建议。最后，总结了迁移的原则、策略和典型问题，并分享行业经典案例。

本指南的编写得到了行业内数据库厂商，数据库专业服务公司以及数据库开源社区的大力支持，在此表示由衷感谢！由于时间仓促，水平所限，错误和不足之处在所难免，欢迎各位读者批评指正，意见建议请发送至 liusiyuan@caict.ac.cn。

目 录

一、数据库及应用系统迁移概述.....	1
（一）本指南内容概述.....	1
（二）本指南的价值.....	1
（三）本指南的适用范围.....	1
（四）本指南的指导对象.....	2
二、数据库及应用系统迁移痛点.....	2
（一）源数据库与应用调研的痛点.....	2
（二）兼容性评估和风险评估的痛点.....	3
（三）可行性验证的痛点.....	3
（四）全面业务改造的痛点.....	3
（五）迁移执行的痛点.....	4
（六）业务验证的痛点.....	4
（七）上线割接的痛点.....	4
（八）护航保障的痛点.....	4
（九）项目总结的痛点.....	5
三、数据库及应用系统迁移流程指南.....	5
（一）源数据库及应用系统调研.....	5
（二）兼容性和风险评估.....	7
（三）可行性验证.....	11
（四）全面业务改造.....	12
（五）全面业务测试.....	16
（六）割接演练.....	18
（七）迁移执行.....	19
（八）业务验证.....	23
（九）正式割接.....	24
（十）护航保障.....	24
（十一）项目总结.....	24

四、迁移能力建议.....	25
（一）迁移服务能力建议.....	25
（二）迁移工具能力建议.....	27
五、总结.....	31
（一）迁移原则.....	31
（二）迁移策略.....	32
（三）典型问题.....	34
六、附录：代表性案例.....	35
（一）阿里云在国内某大型传统零售 ERP 软件案例.....	35
（二）云和恩墨在某世界 500 强企业迁移案例.....	40
（三）中兴通讯在某股份制银行核心业务迁移案例.....	42
（四）南大通用在人社部电子认证系统迁移案例.....	45
（五）海量数据在某交通科研单位证书认证系统案例.....	47
（六）迪思杰在某四大银行客户信息整合系统迁移案例.....	49
（七）新炬网络在某运营商重要系统迁移案例.....	52

图 目 录

图 1	某 ERP 业务系统架构图.....	37
图 2	某 ERP 软件迁移流程图.....	39
图 3	某系统架构图.....	41
图 4	核心系统数据架构图.....	44
图 5	电子认证系统总体架构图.....	46
图 6	证书认证系统平台总体架构图.....	48
图 7	某四大行 ECIF 迁移架构图.....	50
图 8	ECIF 迁移工作流程图.....	51
图 9	数据库双平面运行架构图.....	51

表 目 录

表 1	某两种数据库日期时间类函数映射关系示例.....	14
表 2	某数据库数据存储结构及范围示例.....	14
表 3	某两种数据库隔离级别示例.....	14

一、数据库及应用系统迁移概述

（一）本指南内容概述

本指南首先从数据库及应用系统迁移各时期所面临的痛点出发，针对不同时期的痛点，总结迁移流程相关的方法论，并提出迁移服务能力与工具能力的建议，最后总结出迁移过程中的典型问题和误区，希望读者在日后更加专业、顺利、完整地实施迁移工程。

（二）本指南的价值

本指南的主要价值是：

- 通过方法论为用户平滑、无损地将源数据库迁移至目标环境提供指南；
- 提供迁移服务能力与迁移相关工具能力的建议；
- 迁移过程中典型问题分析与常见误区的解析；
- 为读者提供一些各个行业成功的迁移案例。

（三）本指南的适用范围

本指南主要针对异构数据库之间的数据库及应用系统迁移工作，从技术难度上讲，异构数据库迁移远远大于同构数据库迁移，所以，同构数据库的迁移亦可作为部分参考。本指南的适用范围是：

- 不同数据库、平台之间的异构迁移。
- 不同应用架构的数据库之间的整体迁移。整体迁移是指，针对应用系统及周边依赖系统和数据库产品整体架构的迁移。

（四）本指南的指导对象

本指南的指导对象大致为：

- 有数据库及应用系统迁移意向、迁移需求的用户；
- 数据库及应用系统厂商；
- 高等院校师生及从事或有意愿从事数据库行业的相关人员。

二、数据库及应用系统迁移痛点

数据库及应用系统迁移需要完备的规划方案和谨慎的实施操作，稍有不慎，就会导致迁移失败，甚至丢失宝贵数据，针对整个迁移工作中不同阶段的实际问题，整理出如下痛点：

（一）源数据库与应用调研的痛点

1. 兼容性痛点

- 源数据库使用的特定的开发接口在目标数据库是否兼容；
- 应用绑定了某些特性（例如IP地址、操作系统包、目录结构等）；
- 源数据库特有的特性会增加迁移的难度。

2. 应用改造痛点

- 旧应用的改造会大幅提升迁移难度；
- 缺乏应用开发商的相关支持；
- 应用系统架构和数据库系统之间的关系梳理困难。

3. 资源规划痛点

- 目标数据库成本、资源核算，如何明确对应的资源；

（二）兼容性评估和风险评估的痛点

1. 兼容性评估痛点

- 源数据库到目标数据库迁移，有多少不兼容的语法、对象等，是否有方案；
- 数据库对象兼容性问题的确认，存在不同数据库之间库、表改造难度的判定。

2. 风险评估痛点

- 评估是否有可行的回退方案；
- 目标数据库的可用性、性能是否能够满足原应用需求；
- 评估迁移的成本及工作量；

（三）可行性验证的痛点

- 目标数据库架构选型是否能够满足新应用的需求；
- 数据库、应用改造工作量巨大，如何可以通过工具或产品实现半自动化或自动化改造；
- 应用的逻辑是否依赖数据库的特有功能，该功能是否能在目标数据库兼容，不兼容是否能够通过应用改写等方式满足需求。

（四）全面业务改造的痛点

- 除了可行性验证，其他相关业务可能会带来新问题，影响整体的时间进度和项目计划；
- 应用改造和目标数据库改造的分工可能不明确。（示例：原应

用架构的功能在目标数据库不支持时，需要界定改造工作是由应用方或数据库方完成）；

- 由于考虑不周，全面业务改造存在遗漏，只完成部分业务改造。

（五）迁移执行的痛点

- 迁移的时间窗口能否满足；
- 丰富的数据库对象，如何实现快速迁移；
- 数据一致性检查是否准确快速；
- 需要迁移的系统多，迁移工具能否自动化实现。

（六）业务验证的痛点

- 业务验证是否能够覆盖全面功能；
- 是否存在可量化的指标评估新系统具备上线的条件；
- 业务连续性是否能够满足需求。

（七）上线割接的痛点

- 是否能够在时间窗口内顺利完成迁移割接；
- 是否能够在时间窗口内（甚至不停机的情况下）完成平滑地迁移和数据一致性效验；
- 能否保证数据回流一段时间，以防止割接后应用异常需要回退；
- 上线割接完成后，如何快速性能调优。

（八）护航保障的痛点

- 应急响应是否可行；
- 数据库性能监控及保障，如何快速响应。

（九）项目总结的痛点

- 项目总结是否及时、必要；
- 项目总结后的方法论是否有价值。

三、数据库及应用系统迁移流程指南

本章通过梳理数据库及应用系统迁移流程，为读者介绍迁移相关方法论。需要说明的是，各阶段的顺序不是僵硬不变的，有时需要在不同阶段之间向前和向后移动。这取决于每个阶段的结果和下一个阶段的具体任务。接下来，本指南将详细介绍每个阶段的内容。

（一）源数据库及应用系统调研

对源数据库及应用系统的调研，有助于后续深入评估改造点和工作量，有利于定位系统迁移过程中的难点和风险。这里将源数据库和应用系统调研分为源数据库、应用系统、数据库和应用系统三个方面。方便读者在迁移时选择合适的目标数据库和应用系统架构模型，以及整套迁移解决方案。

1. 源数据库调研

数据库迁移需要考虑数据结构、数据类型、数据库性能、数据使用场景四部分基础数据。数据库结构和数据类型是静态数据，可以通过语法、语义比对完成调研；数据库性能和数据使用场景是动态数据，

与其对应的业务属性、数据库硬件资源、数据库自有能力、数据属性和应用系统属性直接相关。

2. 应用系统调研

应用的调研主要是发现应用和数据库之间的调用关系和调用方式，理清应用各个模块与数据库调用 SQL 的兼容特点，明确应用在各个模块转换的改造点。

应用调整的内容取决于源、目标数据库架构切换的状态以及源、目标数据库的兼容度。通过调用 SQL 详情实现改造点定位，即交互 SQL 点定位，如果应用系统采用成熟框架，可以在框架中直接完成修改；如果在代码中直接嵌入 SQL，若有成熟工具快速完成改造点定位和半自动化实现改造样本，将会极大方便改造工作。

3. 数据库和应用系统调研

复杂的业务生态包含多个数据库、应用的同时运行，在计划数据库架构转型时需要综合考虑数据库与应用系统的关系，由易到难选择影响较小的数据库和业务系统。任何企业信息系统架构的转型绝不是一蹴而就的，需按照科学的方法论全局思考，有计划地完成数据库和应用系统迁移。

数据库和应用之间的关联关系包含但不限定以下种类：

- 数据复制链路
- DBlink 链路
- API 结构调用

- Proxy

- Http

等

综上，数据库、应用系统、数据库和应用的架构关系梳理是迁移流程的重中之重，需要投入大量人力物力。目前，现有专业公司推出智能工具可以协助完成数据库、数据库和应用架构关系梳理，快速缩短时间，掌握现有数据库和应用的结构、架构、性能、关系拓扑，有助于后续决策。

（二）兼容性和风险评估

对源数据库及应用系统的调研完成后，接下来应该对数据库和应用系统迁移工作进行兼容性和风险评估。不同数据库之间，虽然都遵循一定的标准，但实现方法千差万别，有的差别较大，这些差别包括但不限于：

- 架构层面，底层存储、内存处理架构、整体流程架构都有不同。
- SQL 语法方面，创建各种对象的语法、异常处理方式、特别是不同系统存储过程、函数、触发器等差别都较大。

这些差别都需要进行可行性评估，需要评估的内容包括：

- DDL 语法

包括数据库、表空间、表、视图、序列、同义词、触发器、链路等对象的创建语法对比，最好用表格方式列出差异

- DML 语法

包括增、删、查、改各种语法比较，尤其是查询语法的多表联合

查询、层次查询、开窗函数等，特别注意对空值的处理细节。

- 自有非标准对象

如自增列、伪列、伪表、物化视图等等，源库支持，目标库不支持的，有无替代方法，如何替代。

1. 数据库结构语法兼容

数据库结构语法是否兼容，直接关系数据对象迁移的结构，常规分析需要将每一个对象进行语法拆解分析，针对特殊对象需要进行逐一测试，要求分析研发人员要非常了解原、目标数据库的语法结构。

2. 数据库结构语义分析

数据库结构语法正确，但不一定语义正确，语义的分析需要结合数据和应用的语义。语义分析相对语法分析更复杂，针对每一个表的特殊字段类型要依据表的数据和业务需求进行转换。

3. 上下文环境兼容

数据对象之间有依赖关系，其中一个对象不兼容可能会影响其它数据库对象的创建。在分析兼容性时需要对整库作全景画像，通过全景画像确定上下文关系和数据对象之间的依赖关系，决策数据对象兼容结果。

综合兼容性分析，是对所有用户自建的对象在目标数据库的梳理。也建议读者在目标数据库选择时进行不同数据库的兼容性分析，依据最终不同数据库的兼容结果选择合适的目标数据库，有计划地将数据和应用系统的改动量降低到工作预期。同时，目标数据库选定后，

需要依据源数据库性能和业务访问数据库的特性以性能置换原则估算出目标数据库的规格，有助于了解本次切换需要的成本，将诸多迁移中不确定因素变为确定因素。

4. 风险评估

迁移的常见风险评估工作主要包括几个方面：

（1）目标库性能风险

不同数据库的并发、锁、隔离级别、事务控制方法不同，同时一些复杂的 SQL 语法也存在差异，业务迁移后，需要针对这些差异，适当调整数据库配置和应用代码，否则有可能存在性能不够的风险。

（2）数据一致性风险

海量数据迁移，很难精确核对数据一致性和完整性，同时 24 小时不间断服务的业务系统还需要处理增量数据，也可能引出数据不一致的问题。该问题可能导致系统中断，或引起业务偏差。

（3）应用改造风险

一般来说不同数据库的存储过程、函数存在较大差别，业务迁移到目标库，会对存储过程和函数作较大改造，如从支持游标回滚的数据库，迁移到不支持游标回滚的数据库，就必须用列表等其他方法来实现相同的功能，如果这些改造未经充分测试，可能会引起业务偏差。

（4）时间窗口风险

数据库迁移项目，在准备工作阶段，必须先计算迁移的数据量，

选择迁移工具，然后经过几轮测试，计算出数据迁移的可能耗时。制定迁移计划时，为可能存在的无法预测的问题，预留足够时间，否则可能导致在窗口期内无法完成迁移任务。

（5）上线误操作风险

数据迁移执行阶段，宜集中业务经验和数据库技术经验丰富的人员一起工作，宜让业务经验丰富且工作严谨的人员直接操作，以避免误操作。

5. 性能评估

结合源数据库的 CPU、内存、IO 等综合性能和单粒度事件性能为基线，按照目标数据库单粒度性能基线评估出目标数据库所需要的性能资源。

6. 制定回退方案

数据库及应用系统在迁移过程中或迁移完成后，可能由于特定原因，必须回退到原系统。为确保顺利回退，需要制定详细的回退方案，同时需要充分考虑回退处理的时间，将回退处理的时间纳入计划，并留有余量。制定回退方案一般遵循最小改动原则，可将回退方案依据数据库变化而制定。

数据库及应用系统迁移的回退方案，主要包括以下两种情形：

- 目标数据库中未产生新的业务数据

系统迁移过程中，或者迁移完成后尚未正式对外提供服务，如果需要回退，可以直接停止新系统，使用原系统即可。

- 目标数据库中已经产生了新的业务数据

如果迁移完成后，已经正式对外提供服务，在新系统中产生了业务数据的情形下回退，需要提前规划好新系统中产生的业务数据如何处理，是通过技术手段还是通过业务处理的方式来进行回退。

如果数据库源、目标的结构变化不大，仅仅是表结构语法、语义变动，可以通过数据库同步工具进行数据库回流，在数据层实现回退方案；如果数据库源、目标的结构变化大，例如集中式数据库切换为分布式数据库，数据库层有分表分库和字段调整，最优的方式是实现应用层双写，在紧急情况可以将数据切换至源数据库。

回退方案需确保数据的最终一致性，保证用户系统的稳定运行。

（三）可行性验证

在对源数据库和应用系统做充分调研，对目标数据库的兼容性做充分兼容性、风险评估后，需要选取待迁移业务中有代表性的交易模块，对目标数据库使用 POC 的方式进行验证，证明迁移的可行性。

1. 典型交易模块的选取

选取待迁移业务中有代表性的交易模块，通过 POC 的方式在目标数据库验证迁移可行性。

典型交易模块的选取除了需要有涵盖性外，还需遵循如下原则：

- 是重要的业务功能；
- 模型的实现较为复杂；
- 是频繁发生的交易；

- 涉及对源数据库与目标数据库差异敏感的交易；
- 不具备上述特征，但是经专家评议需要纳入验证的交易。

2. 纯 SQL 验证

在典型交易模块与数据库对接验证前，先从选取的典型交易模块中提取对应的SQL语句，通过压测工具直接对数据做压测，验证目标数据库的处理能力是否满足要求。

3. 典型交易模块验证

对选取的典型交易模块涉及的应用程序、库表结构、库表数据根据目标数据库的特性进行分析、设计和改造，验证其迁移至目标数据库后的功能和非功能是否满足系统要求。相对纯SQL验证阶段，此阶段通过高度仿真业务实现，验证应用迁移至目标数据库涉及的技术点。

（四）全面业务改造

经过前面调研评估和验证后，迁移工作进入全面业务改造阶段，由于数据库产品除了共同遵守通用规范和标准，还都有各自的一些语法和对象，细节处理各不相同，因此数据库产品无法完全互相兼容，同一数据库产品的不同版本也无法全面向下兼容。因此业务改造可以依次从如下方面进行：

1. 数据类型

不同数据库产品有各自的数据类型定义，一般分成数字类型、字

符类型、日期时间类型等。这里以某国外数据库和某国内数据库的相互迁移为例进行介绍：

- 数字类型

NUMERIC(p,s)：完全映射至 **NUMBER(p,s)**。如果 **p** 未指定，则默认为 **38**。

- 字符类型

CHAR, **VARCHAR** 和 **LVARCHAR** 是常用的三个字符类型，其中 **CHAR** 和 **LVARCAHR** 的长度都是 1-32767，**VARCHAR** 的长度是 1-255。

- 时间类型

常用时间类型有 **DATE** 和 **DATETIME**，这是两个精度不同的时间类型。**DATE** 只包含日期数据。**DATETIME** 除了日期之外，还可以通过不同的定义来定义更精确的时间。

2. 函数

不同数据库产品有各自的内置函数，异构数据库改造应对源数据库用到的每个函数，在目标数据库中映射到对应函数，并调整参数个数、参数顺序、参数类型等，常见改造场景如下：

- 函数同名，但是在源数据库和目标数据库中功能不同或不完全相同；
- 函数同名，但参数隐式转换规则不同；
- 函数同名，但参数个数或者参数类型不同；
- 函数不同名，但功能相同；
- 无对应函数，需通过其他方式实现。

这里以某两种不同数据库互相迁移为例，示范对比日期时间类函数的映射关系：

表 1 某两种数据库日期时间类函数映射关系示例

数据库 A	数据库 B
CURRENT_DATE	TODAY
CURRENT_TIME STAMP	CURRENT

3. 语法规则

数据库产品语法规则差异较大，虽然应用系统在建设期初会尽量使用遵循 ANSI 或者 ISO 的 SQL 标准语法，但是数据库方言的使用依然无法避免。因此改造过程中，需要将源数据库自身支持的语法规则调整为目标数据库的语法规则。

4. 数据存储

数据库改造和迁移前，应确认源数据库产品和迁移目标数据库产品的数据存储结构及范围，这里以某数据库为例进行介绍：

表 2 某数据库数据存储结构及范围示例

存储相关限制	最大值
一个表或者分片最大的行数	4,277,659,295
一个表或者分片最大的数据页数	16,775,134

5. 隔离级别

表 3 某两种数据库隔离级别示例

数据库 A	数据库 B	ANSI
DIRTY READ	READ ONLY	READ UNCOMMITTED
COMMITTED READ 默认	READ COMMITTED 默认	READ COMMITTED

REPEATABLE READ	SERIALIZABLE	REPEATABLE READ
REPEATABLE READ	SERIALIZABLE	SERIALIZABLE

6. 标识符

数据库对象的名字由标识符来表示，如表名、字段名、索引名、视图名、存储过程名称等等。

7. 关键字

数据库产品的关键字（保留字）有着大量的不同，具体请参考数据库产品的关键字（保留字）列表，并进行改造。

8. 大小写

数据库产品对大小写敏感不尽相同，在改造过程中需要针对具体数据库产品适当处理。

9. 表定义

数据库产品的建表语句通常相似，一些细节需要修改，例如在建表语句中指定 **DEFAULT** 值时，某些数据库的字符型字段可以不加引号，但某些数据库必须用引号标注。

10. 序列

不同数据库产品序列功能基本相同，一般支持 **DML** 和 **DDL** 语句操纵序列，多用户可以并发访问同一个序列，**GRANT** 和 **REVOKE** 语句可以用来更改序列的访问权限。

11. 用户、角色等

用户、角色等一般沿用源数据库业务规则，通过自动结合人工方式在目标数据库中创建并继续使用。

12. 其他数据对象

上文提到的数据类型、函数、语法规则等，单独或者组合存在于下列数据对象中，因此在改造过程中请遵守共同的原则，细节不再赘述。

- 索引
- 约束
- 视图
- 触发器

等

数据库结构订正完成后，用户数据库专家宜完成数据库对象结构订正确认，确保所有订正后的对象语法逻辑正确。业务系统测试人员全面对业务系统功能、性能进行测试，确保业务代码的逻辑正确性。

（五）全面业务测试

为降低数据库及应用迁移的风险，需要在迁移进行全面测试。如果条件允许，测试环境应尽量接近生产环境，确保测试结果能够接近生产环境的实际运行效果。

1. 功能测试

进行应用系统功能的全面测试，在测试过程中如果发现问题，应及时通知相关人员解决，并做回归测试，最终通过功能测试。

2. 性能测试

选取系统在实际运行中，对软、硬资源利用率较高、交易频繁的业务，对已对接目标数据库的应用系统做性能验证。

3. 稳定性测试

根据应用系统的实际需求，在满足性能指标的前提下，对新系统持续运行一段时间，定时监控各项资源使用指标是否有明显增长，系统是否可以保持高效、稳定地运行。

4. 可靠性测试

根据原系统在运行过程各类故障经验，模拟在极端场景下新系统是否可以应对各类突发问题。

5. 扩展能力测试

如果目标数据库为分布式数据库，则必须提供在线横向扩展能力。需要测试是否可以在应用系统及用户无明显感知的基础上，实现扩容动作，解决现有系统在性能、存储容量等方面逐渐出现的瓶颈问题。

6. 安全能力测试

目标数据库完整的安全能力可以有效防止应用系统中敏感数据泄露，防范信息安全风险的发生。建议测试项如下：

7. 数据库及应用系统迁移测试

通过多次数据库及应用系统迁移测试，发现并解决迁移过程中发

现的问题，对迁移工具、迁移程序脚本、迁移过程等逐步优化完善，避免在数据正式割接过程中出现问题。

8. 回退方案验证

对数据库和应用系统迁移的回退方案进行验证，确保在数据库迁移失败的情况下，能够正常回退，避免对业务造成影响。

（六）割接演练

经过全面业务测试阶段后，迁移工程进入割接演练的环节。割接演练是针对正式迁移前，模拟真实上线环境下，对系统进行压力测试和破坏性测试。

1. 割接方案制定

准备制定数据库及应用系统割接方案，方案中应包含数据库及应用系统备份方案、割接方案、应急预案、回退方案，割接方案中，需要明确割接的操作步骤、操作时间和操作人员。

2. 压力测试和破坏性测试

在模拟新系统上线的真实环境测试中，对新系统实施压力测试和破坏性测试，模拟在最极端环境下，新系统功能的完整性、稳定性和高可靠性。

3. 测试总结

上线前，根据之前测试结果，出具下列测试报告：基础环境测试报告、系统源代码迁移报告、系统功能测试报告、系统性能测试报告、

系统稳定性测试报告、系统压力测试报告、系统扩展性测试报告、系统安全性测试报告和系统上线评估报告。

4. 新旧系统同步互备

制定备份、恢复及容灾措施，预防不可控因素带来的损失。正式割接前的备份工作必不可少，在新环境上线前务必做好旧程序包的保留和数据同步，以便在紧急情况可以快速回退。

5. 切换演练

协调实施容灾系统切换演练工作，制定切换演练总体方案，制定数据库切换检查清单，切换演练期间严密监控容灾数据库的系统负载、异常等待事件，并处理期间可能出现的问题。

（七）迁移执行

迁移执行过程会和业务改造交替反复迭代，对工具使用缺省参数进行迁移时候的一些报错可以首先人工进行参数调整，对于一些不能直接迁移的特性，需要再次进行业务改造，然后再进行迁移，直至所有迁移完毕。

迁移执行通常涉及如下环节：

1. 迁移环境检测

检测主机环境常见条目：

- 硬件配置检查
- 登录账号检查

- 系统空间检查
- 网络接口配置检查

检测网络环境常见条目：

- 路由检查
- 带宽检查
- 网络安全配置检查

检测数据库环境常见条目：

- 数据库环境变量
- 数据库日志模式
- 数据库字符集

2. 结构迁移

结构迁移是指将源数据库的建表语句，迁移到目标端不同数据库中，迁移完成保证源、目标数据库中的建表语句功能、性能等价使用。

结构迁移通常有三种方式：

- 利用开源工具完成结构迁移；
- 利用专业迁移工具强大默认支持的源库到目标库内置结构迁移转换引擎，根据迁移工具迁移流程进行结构迁移；
- 通过和最终应用开发厂商进行结构迁移语句的具体定制，把特殊结构迁移转换方式写入迁移转换引擎，根据迁移工具自身使用流程进行结构迁移。

3. 数据迁移

针对可停机、可停业务的迁移场景，且迁移时间窗口满足的情况下，可以利用数据库自带或第三方工具进行静态迁移；在不可停机、不可停业务，迁移时间窗口不满足的情况下，可以利用专业迁移工具实现存量数据在线迁移。

（1）全量数据迁移

将源数据库的存量数据，迁移到目标端数据库中，迁移完成保证源端、目标端数据库中的数据准确、可用、安全。全量数据迁移根据实际业务场景大致分为两种情况：

- 可停机停业务、迁移时间窗口满足

利用数据库自带工具进行静态数据迁移或利用第三方工具进行静态数据迁移。

- 不可停机停业务、迁移时间窗口不满足

利用专业迁移工具实现数据在线迁移。

（2）增量数据迁移

在全量数据迁移完成后，应用业务无法即时切换的情况下，产生的增量数据也需要实时迁移到目标数据库。

4. 对象迁移

在完成结构迁移、全量数据迁移、增量数据迁移后还需要把源数据库的其他对象迁移到目标数据库。

迁移方法分为三种：

- 利用开源工具完成对象迁移

- 利用专业迁移工具完成对象迁移
- 利用专业迁移工具和应用开发商互相配合、特殊定制完成对象迁移

5. 数据同步

在不可停机、停业务，迁移时间窗口较短的情况下，可以利用专业迁移工具实现增量数据同步。

6. 数据比对

在可停机停业务，比对时间窗口满足的情况下，可以利用第三方比对工具或者编写数据库的比对脚本来实现比对；在不可停机、停业务，比对时间窗口不满足的情况下，利用专业的比对工具完成数据的比对和修复，并且比对结果以报表的方式进行查询及展现，完成整体数据比对工作。

7. 迁移结束环境确认

在完成结构迁移、全量数据迁移、增量数据迁移、迁移数据比对后，已经具备了正式业务迁移的条件，业务迁移完成前还需要在目标端数据库完成：

- 序列的重建
- 触发器的启用
- 执行计划的收集

等

业务迁移完成后还需要在目标端数据库完成：

- SQL 的优化
- 索引的优化
- 整体数据库监控

等

8. 构建数据回流

为保证业务迁移后目标数据库切换成真正的生产库后，出现故障无法持续提供对外业务时，保证目标端已经变化的数据能迁移回流到原来的生产库，业务具备再次切换到原来的生产库正式提供业务的能力。

利用专业迁移回流工具完成：

- 目标端到生产端的根据具体时间点进行数据回流
- 目标端到生产端通过分析日志技术数据实时回流

（八）业务验证

数据库及应用系统的很多业务逻辑需要业务场景充分触发并验证，尤其是在大数据量和高并发等极端情况下，更是对数据库功能、性能和可靠性的检验。

1. 迁移数据验证

从源数据库导入到目标数据库中的历史数据文件可以按主键顺序进行组织，可以以文本文件的方式卸载迁移数据，并确保导出数据能够按照主键有序输出。对导入文件和导出文件分别进行比较操作，通过比对结果是否一致，完成迁移数据的一致性验证。

2. 业务功能验证

完整全面的业务验证是割接稳定成功的基础，需要业务方制定全面的功能验证检查列表，逐一通关完成，并依据制定的回退方案监控源、目标数据库数据一致性。测试是否全面覆盖，对于系统上线的稳定运行意义重大。

根据系统中不同业务交易的类型属性，将比对工作分为运行过程比对、运行结果比对、与静态数据比对三个部分。

（九）正式割接

通常割接前需要至少三次割接演练，确保割接过程中各个环节没有疏漏，并根据不同业务系统情况制定割接流程，分配每个流程责任人，通关制完成各个环节，最终实现正式割接顺利完成。

（十）护航保障

数据库及应用系统迁移完成后，最危险的环节是切换后生产环境的第一个业务高峰，需要配置专业的数据库专家，快速响应应用和数据库出现的突发问题。之后需要定期跟踪一定时间，保障业务系统的稳定运行。最终的护航时间，需要根据实际情况确定。如果遇到突发情况，在无法处理的情况下，依据回退方案和演练细则有步骤的完成方案的回退。

（十一）项目总结

项目成功后，宜整理交付文档和总结经验。项目总结阶段宜从如下方面梳理：

1. 交付文档

整个迁移工程各环节均应有文档记录，便于跟踪、分析问题；定期提供服务报告并及时更新所有相关文档，包括实施的服务、服务建议和存在的问题等。

2. 经验总结

迁移完成后，以项目里程碑为基准，与用户一起对系统迁移运行情况，服务情况进行总结，并提出总结与提升方案。总结内容包括但不限于：服务执行总结，系统运行情况及趋势总结，后续运维重点工作讨论，运维技术与方法交流等。

四、迁移能力建议

（一）迁移服务能力建议

建议迁移实施人员以团队参与方式，团队人员包括迁移涉及数据库的专业技术人员、基础硬件系统技术人员及应用开发、运维人员；团队须具有对替代数据库选型的能力，可提供合理的数据库选型建议方案，针对服务提供商的能力评估。结合迁移数据库的类型，考量服务提供商的能力。例如：服务时间、擅长优势、类似案例等；

1. 迁移场景

针对迁移场景，检查是否具有：

- 离线迁移
- 在线迁移

- 同构迁移
- 异构迁移
- 数据迁移
- 应用迁移

2. 迁移流程

针对迁移流程，检查是否具有：

- 成熟迁移体系
- 成功迁移流程
- 优化迁移步骤
- 丰富迁移管理

3. 迁移工具

针对迁移工具，检查是否具有：

- 自主迁移工具
- 精通其它迁移工具
- 集成不同迁移工具
- 统一迁移管理平台

4. 迁移人员

针对迁移人员，检查是否具有：

- 精通迁移流程
- 精通迁移工具
- 迁移数据库认证

- 丰富行业经验
- 大量迁移案例

5. 迁移案例

针对迁移案例，检查是否具有：

- 不同场景迁移案例
- 重大示范迁移案例
- 大量迁移案例累积
- 成功迁移案例比例

（二）迁移工具能力建议

1. 在线迁移

针对在线迁移，检查是否具有：

- 不停机
- 不停业务
- 不修改数据库参数
- 不占用系统大量资源

2. 旁路迁移

针对旁路迁移，检查是否具有：

- 非调用数据库接口
- 非生产主机安装

3. 增量迁移

针对增量迁移，检查是否具有：

- 实时 DML 增量迁移
- 实时 DDL 增量迁移
- 实时大交易增量迁移
- 实时频繁 DDL 增量迁移

4. 转换迁移

针对转换迁移，检查是否具有：

- 表名转换
- 字段名转换
- 增、减字段转换
- 复合计算转换
- 条件数据过滤
- SQL 语法定义及语义兼容性转换
- 数据精度适配，不同数据库之间字段精度适配及隐式转换

5. 迁移比对

针对迁移比对，检查是否具有：

- 全量比对
- 增量比对
- 全量+增量比对
- 定时比对
- 比对修复

- 异构比对

6. 迁移回流

针对迁移回流，检查是否具有：

- 数据回流
- 回流比对
- 回流切换
- 实时双活

7. 对象迁移

针对对象迁移，检查是否具有：

- 同构对象迁移
- 异构对象迁移
- 对象改造
- 对象比对
- 对象修复
- 迁移日志报告

8. 异构迁移

针对异构迁移，检查是否具有如下能力：

- 异构系统迁移；
- 异构数据库版本迁移；
- 异构数据库迁移；
- 异构数据迁移；

- 异构对象迁移。

9. 一体化迁移

是否具有如下能力：

- 对象数据迁移一体化；
- 全量增量迁移一体化；
- 全量增量迁移比对一体化。

10. 高性能迁移

是否具有如下能力：

- 全量数据迁移高性能；
- 增量数据迁移高性能；
- 迁移比对高性能。

11. 特殊场景迁移

是否具有如下能力：

- 特殊表（如大量 LOB、XML 等数据）迁移；
- 特殊系统（如 ERP、EBS、SAP 等）迁移；
- 特殊数据库（如 Cache 等）迁移；
- 特殊云（如上云、下云、云间等）迁移。

12. 系统画像能力

- 数据库和应用系统自动化分析能力，确定数据对象、类型、语法、语义。

五、总结

（一）迁移原则

为保证迁移前后数据完整一致，迁移项目成本可控，迁移过程数据零丢失，迁移后的系统稳定运行，本指南总结如下四项原则。

1. 数据完整一致原则

数据库和应用系统迁移需要遵循数据完整一致性原则。数据完整一致原则并不是要求源数据库和目标数据库的表结构、表数据一模一样，而是要求目标数据库与源数据库所支撑的业务系统结果一致。不能因为数据库的切换导致数据丢失、混乱、不一致，给用户带来不可估量的损失。

2. 业务稳定运行原则

业务稳定运行原则要求迁移后的目标数据库和应用系统要满足且不低于迁移前原系统的稳定运行要求。迁移不应该仅仅是一次架构平台的转变，更应该要求转变后的系统比原有系统更优，尤其是对用户正常商业行为运行的支撑。

3. 成本可控原则

数据库和应用系统迁移同时也要考虑原系统改造到目标系统的硬成本和软成本，遵循成本可控原则。如果整个改造的成本超过用户预期太多，那么改造的必要性就值得怀疑。在预估成本过程中，建议增加不可控因素，例如由业务变更带来的数据模型的不确定性，构建

成本控制模型，实现整个项目周期的成本可控。

4. 回退可信原则

回退可信原则即确保原系统和目标系统之间要有数据通路或业务通路，可以构建源、目标数据库的数据回流链路，或构建业务系统双写机制，在规定时间内确保源、目标数据库和应用系统架构可快速切换，不影响业务正常运行。

（二）迁移策略

通过对迁移流程的关键步骤分析，本指南总结出六条迁移策略。

1. 充分调研评估

数据库和应用系统迁移与否，是通过对源数据库和应用系统进行充分调研，并对目标数据库和应用的适配工作详细了解后所决定的。迁移方案的编订需要对数据库和应用通盘评估，对源数据库的数据结构、数据语法、上下文环境、性能和风险完成综合评估，对数据库之间、应用系统与数据库、应用系统之间关联关系评估，力图将不可控因素降至最低，便于选择合适的迁移方法，为制定解决方案提供有力依据。

2. 真实可行验证

对应用和目标数据库的评估结果进行可行性验证，验证手段包括但不限于 POC 测试、样例应用迁移等。验证完成后，需要对可行性验证报告进行多维度专家评审，评审通过，证明迁移工作具备可行性。

3. 系统整体改造

迁移中的改造环节需要相关研发人员参与，整个改造过程是一个不断试错的过程，需要对目标数据库和应用适配过程和原业务系统逻辑充分掌握。改造内容包括 SQL 语法、数据库架构、应用代码改造等。改造过程可通过相关工具辅助完成，实现自动化推进。

4. 稳步实施测试

应用系统和目标数据库适配完成后，需要对业务系统进行功能、性能、安全性等全方位测试。实施测试过程建议重点关注如下环节：首先，应用代码适配目标数据库需要改造模块进行代码级单元测试，确保代码可以正常运行；其次，所有业务系统改造完成后，要根据原业务系统的功能测试流程，由专业测试人员进行全流程测试，确保测试覆盖面要全；最后，依据原业务系统生产运行的峰值，设定业务系统性能基线和业务运行模型，对目标数据库和适配后的业务系统进行性能测试，测试结果不低于源数据库和业务系统的性能波峰。

5. 有序割接护航

业务系统的割接护航需要多部门协同，割接需要有护航保障，护航需要设定护航窗口，在时间窗口内监控业务系统的相关运行情况，对出现的问题进行重点优化，保证业务系统在割接后正常运行。如割接失败，依据回退可信原则稳步回退。

6. 全面迁移总结

在完成整体迁移工作后，需要对迁移各阶段的各项具体工作内

容、工作经验进行总结，优化迁移流程、迁移方法，分享迁移经验，不断完善整体迁移解决方案。

（三）典型问题

迁移过程中涉及的问题多种多样，本指南列举五个常见的典型问题，希望对读者有所启发和帮助。

1. 数据库资源消耗差异

源数据库和目标数据库执行相同业务所需资源可能存在差异，不同数据库支持相同业务的资源成本需尽可能精准评估，否则目标数据库可能无法承载源数据库的负载业务。

2. 数据库数值显隐式转换

数据库中存在较多显、隐式转换，不同数据库可能存在不同的隐式类型转换规则，如果源数据库与目标数据库隐式类型转换规则不同，可能造成应用不可用。

3. 时间日期类型转换

不同数据库时间日期类型的格式存在差异，对于时间格式要求比较敏感的业务系统，需要对时间格式进行充分测试。

4. 数据库字段长度差异

源数据库和目标数据库数据类型默认长度可能存在差异，会导致迁移前后系统展示结果不一致，需要对数据库字段长度进行充分测试。

5. 应用框架方言包适配

应用开发框架使用的数据库方言包可能存在差异。迁移过程中需要使用与目标数据库适配的方言包，使用不适配的方言包可能导致应用无法运行。此时可通过业务模块代码嵌入，绕过方言包。

六、附录：代表性案例

本指南梳理了部分国内数据库厂商和数据库专业服务商主导或参与的典型迁移案例。案例介绍框架分为迁移前基本情况和数据情况、主要面临的痛点和难点、平台总体架构、数据库及应用系统迁移实施步骤以及迁移后应用效果五部分。后续中国信通院还将继续征集更多的典型案例补充到未来的版本中，促进数据库迁移工作经验交流与信息共享。

（一）阿里云在国内某大型传统零售 ERP 软件案例

1. 迁移前基本情况和数据情况

该公司为国内外 3000 多家企业、近 30 万家门店提供数字化管理 ERP 系统定制化服务。其 ERP 系统使用 Oracle 数据库提供数据支撑。随着互联网经济的发展，企业业务的快速增长及业务种类高速迭代，原来 ERP 系统架构面对不断快速增长的数据，纵向扩展数据库已经不能满足客户业务需求，同时高配的数据库硬件和软件许可费用给客户带来较高支出，ERP 系统数据库及应用转型成为必然选择。项目于 2019 年 6 月立项，该公司内部研发工程师进行目标数据库选型调研，发现人工选型要调研的内容庞杂，至少需要 3 个月完成目标数据库选

型论证。综合多方因素，最终选择阿里云数据库和应用迁移服务。

2. 主要面临的痛点和难点

（1）目标数据库如何选择

开源数据库及其它商用和云数据库百花齐放，各有优劣，如何选择一款可以快速完成新 ERP 系统数据库及应用迁移改造的适配数据库比较困难。

（2）数据库和应用改造无从下手

业务系统数据库异构切换，数据库对象改造点、新 ERP 系统和目标数据库的适配改造点未知；在适配过程中存在的业务系统和数据库的风险未知。

（3）数据迁移及应用割接存在风险

数据迁移及应用割接在客户环境中进行，涉及到本地 IDC 链路和云环境链路。更重要的是异构数据库的数据迁移如何确保数据完整一致，业务稳定运行并能实现降本增效存在较大挑战。

3. 平台总体架构

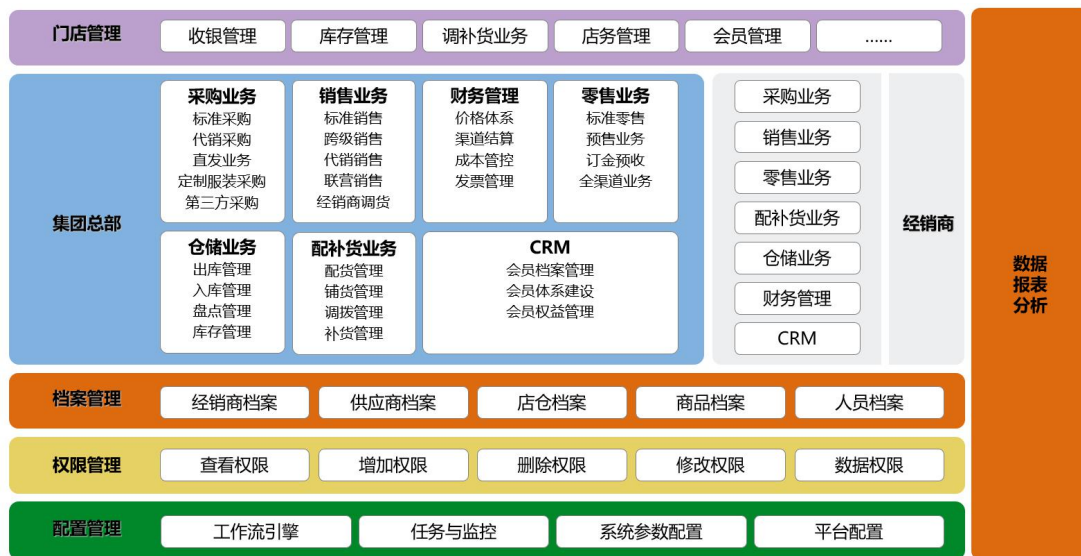


图 1 某 ERP 业务系统架构图

4. 数据库及应用系统迁移实施步骤

2019 年 6 月 25 日：阿里云数据库和应用评估产品进场，通过自动化智能采集该公司的数据库和应用系统结构、运行、计算资源性能信息，持续采集 7 天；

2019 年 7 月 2 日：该公司上传其数据库和应用的采集信息，经过数据库和应用评估（Advanced Database & Application Migration，以下简称 ADAM）产品智能分析，推荐选择 Oracle 数据库平滑迁移至 PolarDB-O 数据库可以确保迁移后应用和数据库的改造工作量最少，ADAM 产品数据库改造系统根据目标数据库的语法、运行规则给出了改造规则和样例，公司研发工程师根据改造规则进行修订即可；ADAM 产品数据库应用改造系统定位出需要配合改造应用代码的位置和需要修改的代码样例，公司研发工程师通过改造样例结合该 ERP 系统的上下文关系进行代码改造，完成与目标数据库的适配；

2019 年 7 月 8 日：选择该 ERP 系统的会员管理模块进行改造 POC

适配，结合 ADAM 产品的改造能力 3 天内完成了该模块在目标数据库的适配，确定方案可行，正式立项；

2019 年 7 月 15 日 - 18 日：公司研发工程师使用 ADAM 产品数据库结构迁移工具完成 Oracle 数据库的数据对象迁移订正工作。对于无法判定的修订对象，通过 ADAM 产品 SQL 在线翻译工具和阿里云数据库专家的线上支持，协助数据库工程师完成目标数据库对象迁移的订正工作；

2019 年 7 月 22 日：使用阿里云 DTS 数据迁移工具完成测试数据从 Oracle 数据库迁移至目标数据库测试环境，并进行数据增量迁移链路测试；

2019 年 7 月 25 日：ERP 系统研发工程师投入 3 人，持续 32 人/天完成所有 ERP 系统功能模块的改造适配工作；

2019 年 8 月 27 日：测试工程师正式进场，对所有业务模块进行功能性测试和性能测试，并反馈研发工程师进行相应代码的优化调整；

2019 年 9 月 10 日：在阿里云专家服务的支持下进行数据库及应用迁移割接演练，前后完成三次演练，制定生产系统割接流程表；

2019 年 9 月 22 日：在阿里云公共云配置目标数据库 PolarDB-O 数据库环境；开启 DTS 数据全量迁移链路，在不影响客户原业务系统正常使用的前提下，完成数据全量迁移；

2019 年 9 月 23 日凌晨：开启 DTS 数据增量服务，进行增量数据追平；

2019 年 9 月 24 日 - 25 日：获取客户生产系统割接授权，正式完成该公司其中一个客户的生产环境割接，持续时间不超过 1 小时；

2019 年 9 月 25 日 - 10 月 2 日：进行一周新系统运维护航，并整理相关数据库及应用迁移材料，进行项目总结；

2019 年 10 月 8 日：向上层汇报，业务系统改造完成；

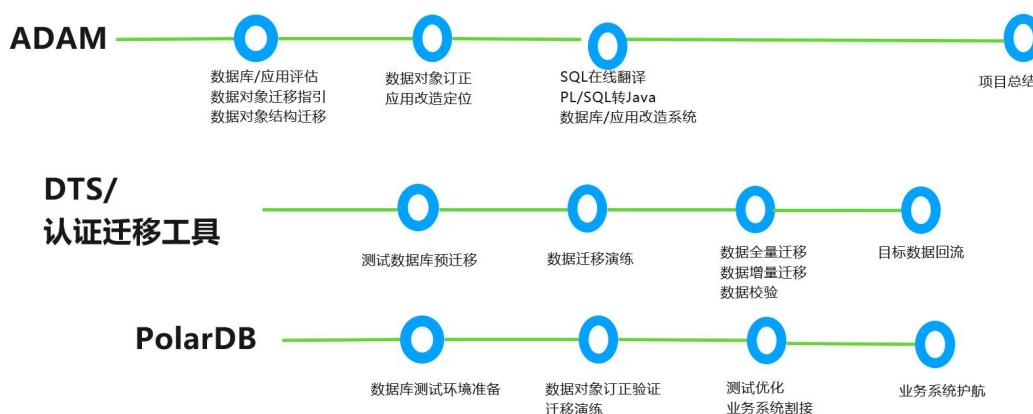


图 2 某 ERP 软件迁移流程图

5. 迁移后应用效果

迁移后的 ERP 系统利用 PolarDB-O 数据库计算存储分离的特性，在云的加持下，业务高峰期可以动态升级计算节点配置，最高支持 16 节点，在业务低峰期，保持 2 节点正常提供服务，同时单节点的纵向增配可在分钟级别实现。

新系统实现了门店管理、经销商、档案管理、权限管理、配置管理模块化，各个模块之间通过 API 方式进行调用，配合阿里云敏捷开发架构，根据客户需求可以快速迭代产品功能。帮助最终客户降低 40%+ 的系统成本，提高了 ERP 系统各模块间的使用效率，使门店间的业务往来更流畅。

（二）云和恩墨在某世界 500 强企业迁移案例

1. 迁移前基本情况和数据情况

某世界 500 强企业 Oracle 数据库数据量达到几十 T，改造前系统性能情况如下：TPS 峰值为 5000，QPS 峰值为 9000，数据库服务器 CPU 占用率为 60%。

因数据量增长带来的压力及高并发的业务需求，且数据库已经超出关系型数据库服务提供的能力，未来数据量会持续增长。因此，必须设计一套可持续扩展的分库分表方案，此方案需要包括对现有数据的拆分方案，以及在数据量增加到一定程度后，平稳地增加新的子库的方案，最终计划将数据由现在的 Oracle 数据库改造迁移至 MySQL 数据库。

2. 主要面临的痛点和难点

由于某业务表非常大，按照传统分库的方式进行拆分，对于 MySQL 而言，其大小仍然偏大。因此，可以在每个库中进一步拆分成多个子表，其对应的需求、用例、任务等表也根据同样规则拆成多个表。

此外，由于分库键值对应的数据量相差可能很大，如果采用通用的哈希值方式拆分，会导致拆分不均。因此，选择以自定义的形式进行拆分。新建一张分库配置表，存放每个分库键值对应的子库名字及子表后缀信息。应用访问时，通过此表获取特定的分库键值对应的子库及子表后缀，从而访问指定子库的指定表。

3. 平台总体架构

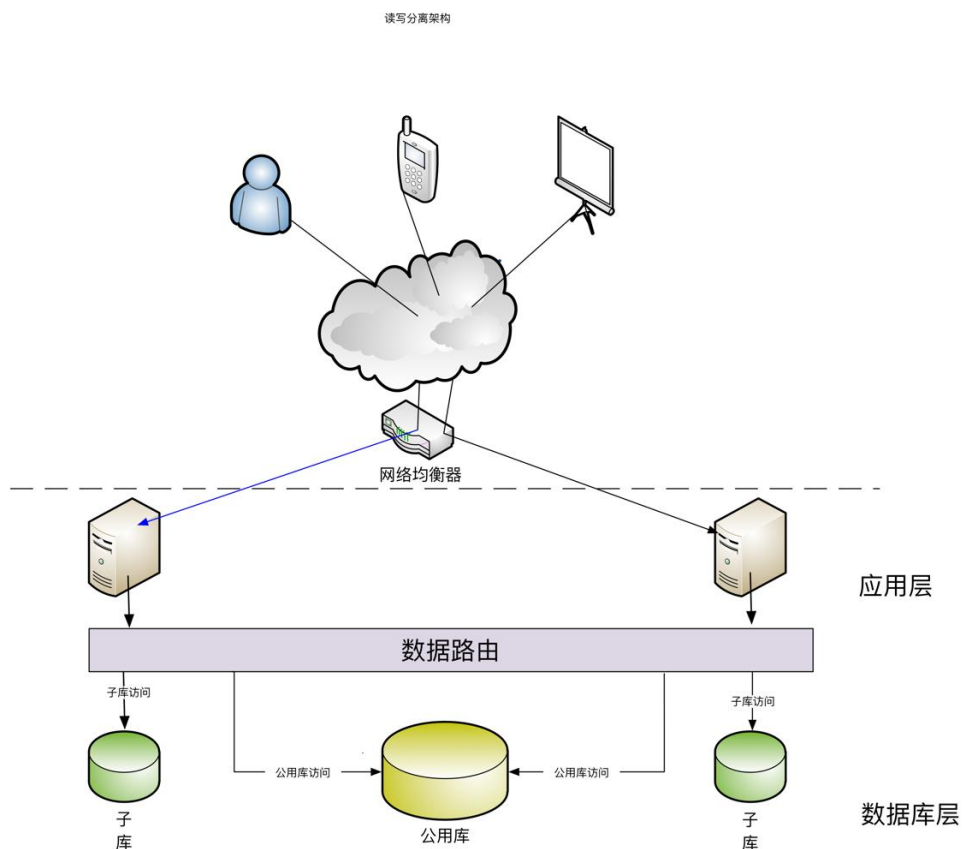


图 3 某系统架构图

4. 数据库及应用系统迁移实施步骤

2019年4月：制定系统分表分库架构，以及确定分库分表基本原则、确定具体需要分库的表；

2019年4月：创建空的分库，并建立通用表结构，程序按照新的架构创建新的项目分支并进行代码调整 and 开发；

2019年7月：设计迁移工具具体功能点；

2019年8月：编写迁移工具，迁移工具需要满足如下特色功能：

- 设计字段类型映射并自动创建表结构
- 并行进行数据的迁移

- 自动创建索引及约束
- 自动表结构及表数据比对
- 支持断点续传

2019年9月：调试、测试迁移工具；

2019年10月 - 12月：正式迁移，由于原架构涉及多个部署点，因此迁移分多次进行，迁移后的系统在新架构下运行，未迁移系统仍并行运行于旧系统。

2019年10月 - 2020年3月：对新架构下的MySQL进行持续调优。

5. 迁移后应用效果

迁移至新环境后，基于MySQL分库分表的新架构在业务承载能力上得到很大提升，应用程序在10万+级并发下，主要接口请求时间由平均大于500ms降低至50ms内，主要页面响应时间从平均大于10s降低至2s内。

（三）中兴通讯在某股份制银行核心业务迁移案例

1. 迁移前基本情况和数据情况

某股份制银行原信用卡核心系统运行在IBM AS400+DB2平台上，由于近年来数据飞速增长，客户量已接近亿级，数据量超过10T，源数据库平台能力已经无法应对未来的发展需要，尤其是在扩展性上无法满足业务需求。

为此该银行开发了新一代信用卡核心系统，数据库采用中兴通讯的GoldenDB分布式数据库。系统设计要求支持PB级数据量，百亿级别

的数据库表，联机交易时延稳定在40ms以内，日终批处理在1.5小时内完成，支持系统在线扩容，同时整体架构基于廉价服务器组成的开放平台。

2. 主要面临的痛点和难点

（1）业务系统

- 核心系统联机业务对时延特别敏感，以快捷支付业务为例，单笔业务有 30 多条 SQL 语句，时延必须小于 40ms。而旧系统中，一旦业务压力增大，业务时延随即增大，甚至出现拒绝交易的情况。目标数据库要求能应对数十万并发请求，并能在长期高并发、高负载的状态下减少业务锁冲突，保证系统稳定性。

（2）原系统达到瓶颈

- 旧核心系统批处理性能已达到瓶颈，批处理时间较长，需要利用分布式架构特性提升批处理业务的性能，同时尽可能减少业务代码入侵、减少迁移成本、提升系统效率。

（3）时间窗口小

- 联机业务需要能快速迁移到新系统中，特别是在追增阶段，需要完成下载、转码、迁移、验证等多项工作，时间窗口极其紧张，需要有完善、可行的系统迁移方案与实践经验。
- 亿级客户量的核心系统必须满足金融行业信息系统 6 级灾备等级要求，同城异地机房可进行无缝切换。

3. 平台总体架构

该行新核心数据库采用两地三中心的部署架构，结合GoldenDB特有的快同步技术与灵活的分组管理技术保证核心业务连续性，实现同城RPO为0，异地机房RPO秒级，在出现城市级别灾难时异地机房可快速接管联机业务，RTO小于5分钟。

新核心存储引擎配置的X86服务器性能较高，综合成本和可用性考虑，每个服务器中部署了2个存储引擎，不同服务器之间做交叉数据副本主备，保证单机异常不会影响到系统可用性。单机中两个存储引擎的文件系统相互独立，保证IO隔离。

在设计业务连接实例时，将不同业务连接的计算引擎做物理隔离，杜绝因个别业务故障引起的故障传染；由于不同业务对于数据库的隔离级别和运行模式的要求各不相同，通过统一配置管理平台，灵活配置不同的隔离级别与运行模式。

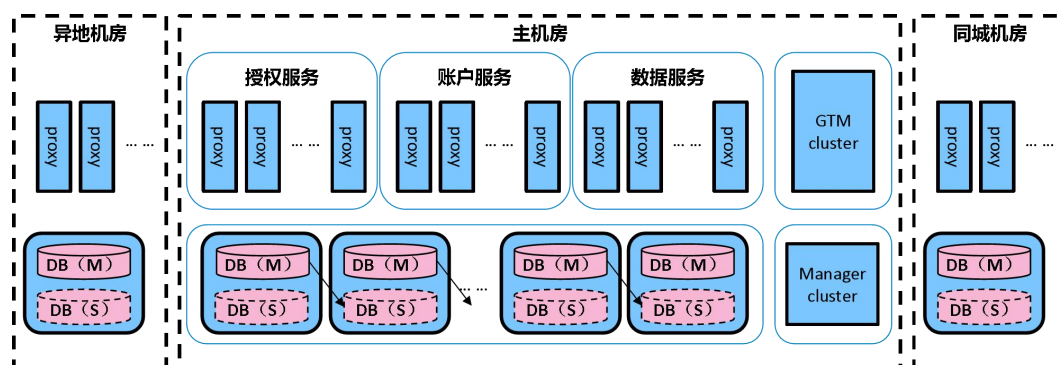


图4 核心系统数据架构图

4. 数据库及应用系统迁移实施步骤

2017年10月 - 2018年5月：产品选型测试；

2018年6月 - 2018年12月：产品功能对接测试；

2019年1月 - 2019年4月：业务功能、性能对接测试验证；

2019年3月 - 2019年6月：数据迁移性能指标达成；
2019年5月 - 2019年9月：核心系统跟账演练，正确率达标；
2019年8月 - 2019年10月：分别完成4次迁移演练与4次切换演练；
2019年8月 - 2019年10月：投产上线应急预案交付与实践；
2019年10月26号：信用卡新核心系统正式上线。

5. 迁移后应用效果

自2014年以来，中兴通讯与该银行共同研发分布式数据库GoldenDB，稳中求进，不断深入。在该银行冠字号系统、门户网站、金融同业平台、统一零售积分系统、电商管家、开放银行、用户权益系统等系统陆续成功投产。2019年10月26日该银行基于GoldenDB的新一代信用卡核心正式上线运行，日均交易量9000万笔，顺利通过双十一、双十二的业务峰值考验，数据库运行平稳。

经过5年不断打磨，GoldenDB分布式数据库经历了多个严苛的商用案例考验，已经具备在金融交易业务中大规模商用的能力。

（四）南大通用在人社部电子认证系统迁移案例

1. 迁移前基本情况和数据情况

为深化社会保障卡（以下简称“社保卡”）在人力资源和社会保障领域应用，为公众提供更为便捷安全的用卡方式，需利用人力资源和社会保障电子认证系统，以第三代社保卡为载体，加载 SM2 算法数字证书，为第三代社保卡在人社业务系统中的应用安全提供有力保障。本案例建设任务有两个：第一，以人力资源社会保障电子认证根

系统为行业信任源点，建设省级电子认证系统，实现与部电子认证系统的统一规范建设。第二，对所有使用数字证书的业务系统，完成基于 SM2 算法数字证书集成应用改造，为业务系统提供身份认证、签名验签和数据加解密等应用安全支撑。该系统记录全国人社社保人员及企业相关信息，总条目约为1亿8000万。

2. 主要面临的痛点和难点

证书认证发证速度缓慢，验证测试为2万条目/时，无法达到用户要求6万条目/时。

3. 平台总体架构

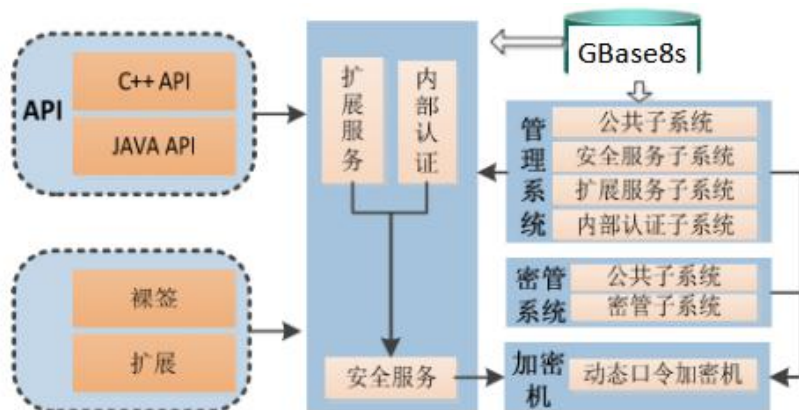


图 5 电子认证系统总体架构图

4. 数据库及应用系统迁移实施步骤

2017年12月- 2018年3月：进行普华应用软件适配开发工作，完成全部应用存储过程、函数、SQL语法迁移工作；

2018年4月- 5月：完成业务测试，包括功能、性能、稳定性、安全能力；

2018年6月：完成人社部数据库软件实施部署工作；

2018年7月- 8月：完成数据迁移、应用割接工作；

2018年9月- 10月：完成试运行业务验证工作；

2018年11月- 12月：完成重庆、成都数据库环境安装部署实施工作；

2019年1月- 2月：完成辽宁、山东数据库环境安装部署实施工作；

2019年3月- 4月：完成西藏、山西、湖北数据库环境安装部署实施工作；

2019年5月- 6月：完成山东、西藏、辽宁数据库软件升级工作；

2019年7月- 8月：完成新疆、海南、河南、山西（三期）数据库环境安装部署实施工作；

2019年9月：完成福建、广东、黑龙江数据库环境安装部署实施工作；

2019年10月：完成项目总结，交付文档。

5. 迁移后应用效果

大大提升发证速度，测试结果远超出用户要求，达到14万条目/时。

（五）海量数据在某交通科研单位证书认证系统案例

1. 迁移前基本情况和数据情况

时至今日，我国ETC系统已经逐步发展成为一个庞大的信息化生产系统，同时作为关系国计民生的重要生产系统和信息系统，保障网

络信息安全，确保系统安全运行尤为重要。因此要求为ETC业务提供数字证书的证书认证系统，需采用高可用及容灾的部署模式。由于响应国家在IT信息化建设的意见与要求，需要目标数据库并实现架构要求及数据的完整、可靠迁移。

2. 主要面临的痛点和难点

- 如何简单、便捷地完成数据库数据的迁移，达到业务使用要求；
- 目标数据库产品可以满足数据高可用及支持异构数据库间的数据灾备。

3. 平台总体架构

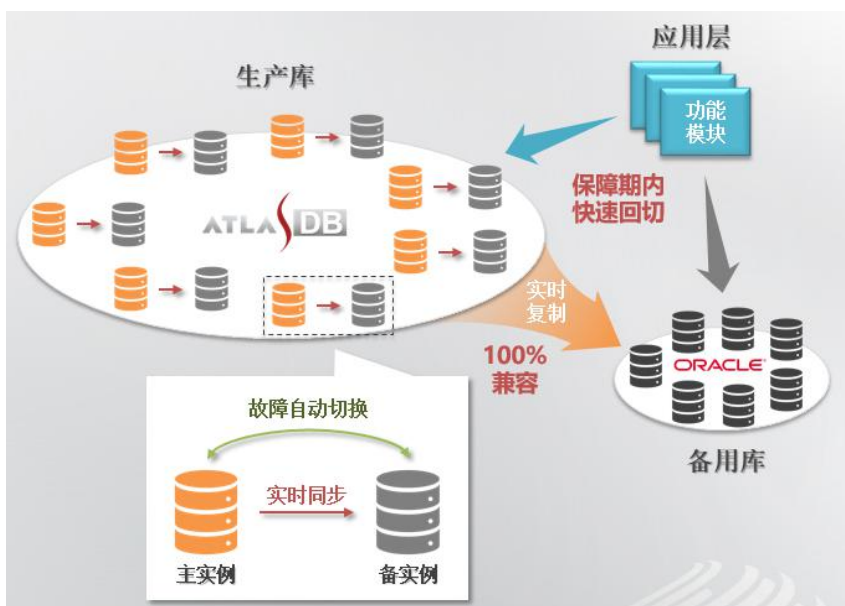


图6 证书认证系统平台总体架构图

4. 数据库及应用系统迁移实施步骤

2019年5月中旬：完成产品选型及产品验证工作，明确产品选型；

2019年5月下旬：完成迁移测试，基本确定迁移方法及风险测估；

2019年6月初：对整体迁移步骤的验证、整理及迁移方案讨论会，确定了正式的迁移方；

2019年6月中旬：完成正式数据迁移并进入保证阶段。

5. 迁移后应用效果

本项目建设实现了数据迁移及交付使用要求，同时也达到了本地数据高可用及数据异构容灾的要求。

（六）迪思杰在某四大银行客户信息整合系统迁移案例

1. 迁移前基本情况和数据情况

某四大银行的ECIF是企业级客户信息整合系统，主要目的是整合银行各个系统的客户信息。解决银行客户信息重复采集、共享性差，利用率低，操作风险高等问题，提供集中、统一、共享和权威的单一客户视图，为企业后台服务系统提供实时、完整、一致的客户信息。

随着各个系统数据量及业务需求的增加，急需建立新一代ECIF系统，完成将相应原有系统的数据，在较短时间内迁移到新系统上，并运行新的应用业务。

2. 主要面临的痛点和难点

（1）迁移数据量大、时间窗口少

迁移数据量为15T+，在不停机、不停业务的情况下，生产数据迁移至中间计算库要在1天内完成。

（2）不同数据库之间的数据迁移

各业务系统的数据库主要为Oracle数据库，新系统的核心数据库为DB2。异构数据库之间字段类型区别、数据校验比对方法存在痛点和难点。

（3）整体迁移分不同阶段，非一次性完成

整体迁移分多阶段，非一次性、一晚上迁移完成，持续近一个月。

3. 平台总体架构

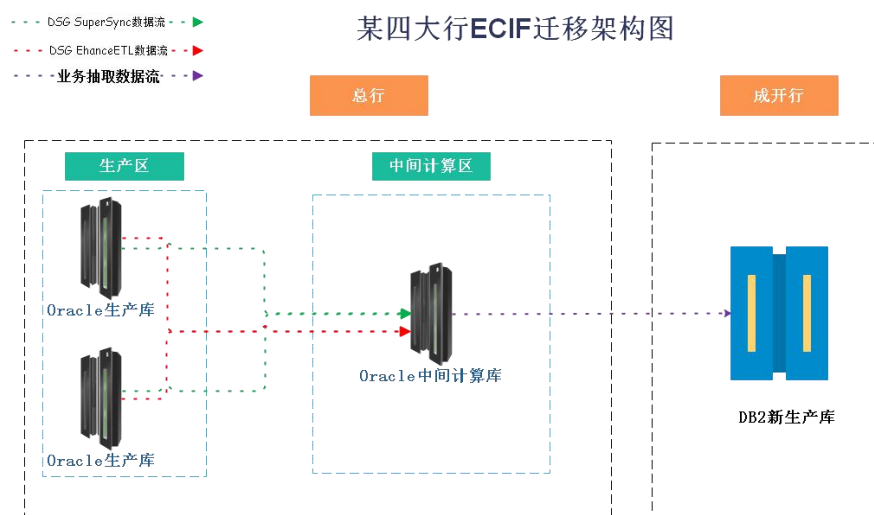


图7 某四大行 ECIF 迁移架构图

数据迁移策略如下：

- 生产数据导入到中间计算库；
- 中间计算库完成全量数据核准，并自动停止增量装载；
- 全量数据核准完成，打开增量装载开关，中间计算库追加部分实时增量数据；
- 关闭增量装载开关，中间库再次停止增量，进行部分数据核准；

- 利用独特生成的各表的轨迹变化表，再次进行数据校验、数据修复等工作。

4. 数据库及应用系统迁移实施步骤

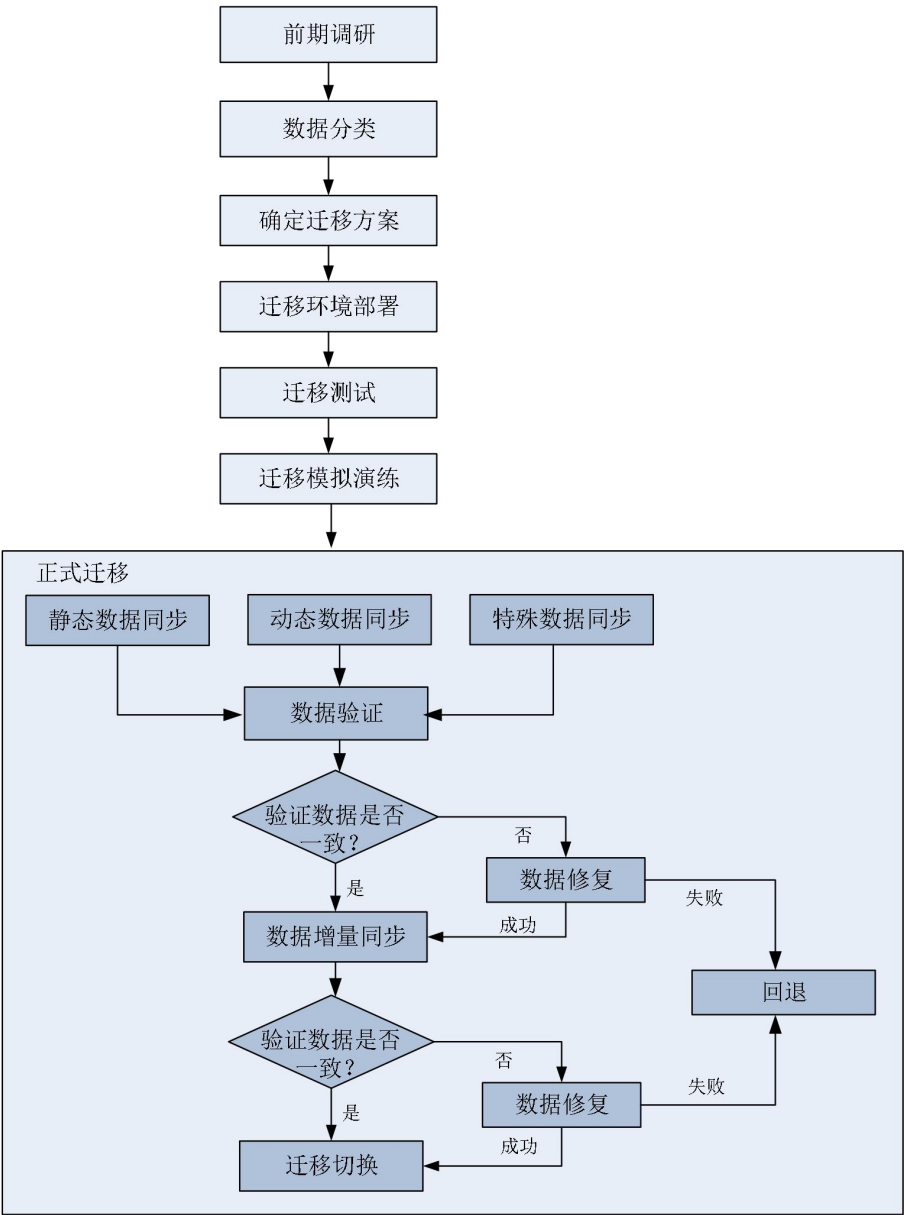


图 8 ECIF 迁移工作流程图

2015年1月1日 - 2015年1月2日：利用DSG全量同步工具进行数据初始化同步；

2015年1月2日 - 2015年1月3日：数据映射比对，映射备份，设置停

止装载SCN点或取全量数据装载点；

2015年1月3日 - 2015年1月6日：检核数据库全量数据装载；

2015年1月6日 - 2015年1月12日：成都分行开始工作从ECIF检核数据库中抽取全量数据到新一代生产主机；

2015年1月12日 - 2015年1月22日：追加增量数据。（该步骤为成开方从ECIF检核数据库中抽取增量数据到新一代生产主机。期间需要DSG将同步程序启停数次，并设置检查点。便于成都方面进行基于时间点的增量数据抽取）；

2015年1月22日 - 2015年1月24日：数据修复和比对；

2015年1月25日 - 2015年1月26日：新系统应用验证；

2015年1月27日 - 2015年1月29日：新业务应用压测；

2015年1月30日：新一代ECIF业务上线。

5. 迁移后应用效果

利用DSG自主迁移工具达到600G/小时的迁移速度，且利用DSG自主日志分析技术对每张表的每条记录变化生成额外变化轨迹表并增加特殊识别字段，为后续数据一致性比对、业务连续性校验提供技术实现基础。在近一个月的迁移过程后，新ECIF系统顺利上线并正常运行。

（七）新炬网络在某运营商重要系统迁移案例

1. 迁移前基本情况和数据情况

某运营商 O 域系统原本采用 Oracle+SUN 小机 RAC 两节点高可用

架构的组合搭建后台数据库，数据体量约 1.5TB，并搭建了一比一的 ADG 容灾环境。业务系统承载了大量数据采集和报表任务，每天数据库的归档日志量达 500G 以上，整体负载较高。

2. 主要面临的痛点和难点

原 Oracle 服务器服役年限较长，在业务高峰期间服务器的各项资源使用率较高，影响业务的高效运行。另外，在业务和数据库运行过程中，还遇到以下问题：

（1）原系统面临过保风险

SUN 小机服务器面临过保风险，硬件故障和硬件性能问题获得原厂工程师直接支持有难度，业务稳定性和健壮性无法得到保障。

（2）原系统资源未来难以满足扩容需求

同类型的服务器已不再采购，现有存储资源和服务器资源无法满足未来业务扩容的需求。

（3）原系统性能瓶颈突显

业务峰值期间的性能瓶颈明显，通过程序逻辑和数据库层优化效果有限，无法满足业务持续拓展的需求。

3. 平台总体架构

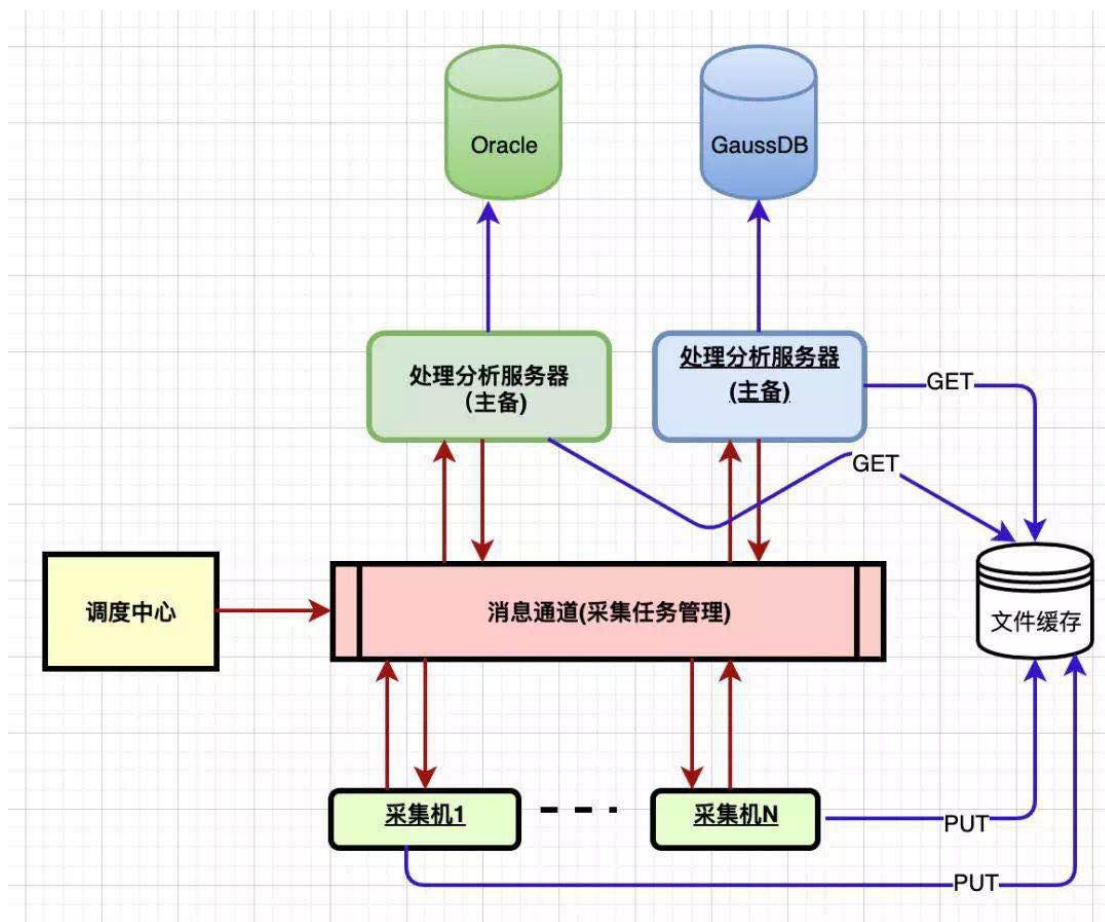


图9 数据库双平面运行架构图

4. 数据库及应用系统迁移实施步骤

2019 年 9 月，确认针对上述业务系统进行数据库架构转型，并确认时间表；

2019 年 10 月，服务器通电及配置完成，GaussDB 及组件安装配置完成；

2019 年 11 月，完成全量数据导入到 GaussDB 数据；

2019 年 12 月，完成业务程序改造；

2020 年 1 月，完成所有业务功能性测试；

2020 年 2 月至今，进行数据库双平台运行测试及性能测试。

5. 迁移后应用效果

在当前的数据库双平面运行期间，迁移后数据库录入相同数据的资源消耗率大大降低，高峰期间的 CPU 使用率从原来的 60% 下降到 10% 以内，且业务程序各项报表的生成时间，数据采集队列录入效率及数据查询效率等方面都有显著的性能提升，为后续业务的持续拓展提供了有力支持。

中国信息通信研究院 云计算与大数据研究所

地址：北京市海淀区花园北路 52 号

邮政编码：100191

联系电话：13691032906

传真：010-88665185

网址：www.caict.ac.cn

