基于 CPU 环境的多类型数据库同步方法、装置及设备

张 浩[1] 朱志强[1] 陈志强[1]
ZHANG Hao ZHU Zhiqiang CHEN Zhiqiang

摘 要 目前基于国产 CPU 和操作系统的全国产环境下应用生态日益完善，在国产环境下 WEB 应用系统需要支持神通、达梦、金仓、翰高等多种国产数据库，以满足不同客户的需求。日常 WEB 应用系统迭代开发过程中，存在不同数据库间存在关键字不统一、支持的数据类型不一致、SQL 语句语法格式不相同、自带的函数存在差异、支持的数据库操作不完全相同等问题，开发维护及问题定位困难。通过应用多类型数据库适配与同步方法，开发人员只需要输入基准数据库的连接信息，导出 XML 数据库结构描述模板，导出的 XML 文件包含数据库类型、数据表结构等信息，项目实施人员只需要输入数据库的基本信息后，将基准数据库生成的 XML 文件导入，点击生成 SQL。就可以生成与基准数据库有差异性的 SQL 语句，或者直接点击同步，也可以将目标数据库的表结构与基准数据库的表结构进行同步，简化了项目组人员的升级过程。

关键词 国产 CPU；数据库适配；微服务；跨平台

doi：10.3969/j.issn.1672-9528.2024.02.019

0 引言

近年来，国家大力扶持具有自主知识产权的全国产软硬件的发展，涌现了以国产操作系统和 CPU 为代表的众多具有自主知识产权的基础软硬件产品。中标麒麟系统、深度操作系统等国产操作系统生态环境日趋完善，龙芯、飞腾等具有自主知识产权的高端通用芯片蓬勃发展，技术水平达到或接近同类产品的世界先进水平。

随着国产基础软硬件的蓬勃发展，国产基础软硬件的推广和使用带来了前所未有的机遇。数据库作为生态环境中重要的一环，基于国产操作系统和 CPU 的环境下涌现出了多种国产数据库 [1]。不同数据库间存在关键字不统一、支持的数据类型也不完全一致、SQL 语句语法格式不一致、自带的函数不相同等问题。开发人员需要学习并掌握各种国产数据库的 SQL 语法，针对不同数据库上编写对应的 SQL 语句并进行测试，并向版本控制系统提交各种数据库对应的 SQL 文件，供开发人员间共享文件及方便后续发包人员获取升级文件 [2]。项目组开发成员的编程风格差异和数据库操作语句 SQL 的灵活性，导致 SQL 写法各异、SQL 规范性较差，后续开发维护及问题定位困难。发包人员需要从版本控制系统检出不同数据库对应的 SQL 文件，根据要发送的需求编号整理 SQL 文件并打包发给现场升级人员，需要 SQL 执行正确的顺序。现场升级人员需要提前备份每个待升级数据库，根据数据库类型，选择要执行的 SQL 并关注每条 SQL 执行结果 [3]。

另外，将 SQL 汇总到文件中目录复杂，SQL 文件不方便浏览和信息查找，增加了后续维护及问题定位的困难。一次性更新大量需求时，数据库改动较大，需要维护大量的 SQL 文件，由于 SQL 数量多、SQL 关联性强，SQL 文件整理极其复杂。由于无法通过 SQL 文件方便地查看某个数据库表的历史需求修改，当更新历史需求集合时，易出现漏更或者旧数据库表结构、表数据等错误覆盖。

1 技术方案

1.1 多类型数据库适配方法概述

多类型数据库适配方法是指将不同类型的数据库系统进行互操作，使它们能够在同一环境中协调工作。这种适配方法可以用于将关系型数据库与非关系型数据库、分布式数据库与传统数据库等进行集成，以满足不同应用场景的需要。

基于国产 CPU 环境的多类型数据库适配方法主要的目标是实现国产环境下不同数据库结构的兼容适配，支持数据结构的导出与同步。目前已适配神通、金仓、达梦、翰高等国产数据库 [4]。

该适配方法具有支持微服务环境下多数据库（模式）的结构、支持多种数据库模式同步适配、支持默认模式设置、提供模式检测等功能。在微服务环境下，服务与数据库模式之间的多为一对一的关系，不同的项目服务名相同，但所对

1. 浪潮软件集团有限公司 山东济南 250000

应的数据库模式不尽相同[5]。为了方便管理，适配器支持服务与模式的对应关系用字典项维护，以该对应关系作为模式导入导出的默认选项，提高用户体验。

该适配方法采用 Java 语言实现，支持跨平台运行，满足多平台下数据库结构同步的需求。

1.2 适配器技术实现

适配器技术实现一般涉及适配器模式（Adapter Pattern），是一种结构型设计模式，它允许不兼容的接口之间进行交互。适配器像是两个不同系统或类之间的桥梁，让原本由于接口不兼容而不能一起工作的类可以协同工作。多类型数据库的适配器技术是通过编写特定的适配器程序来实现数据库之间的互操作。适配器充当一个中间层，将不同类型的数据库系统的不同语法、协议和数据模型转换为统一的接口，使得应用程序可以无缝地与这些数据库进行交互。

（1）API 适配器：通过为不同类型的数据库编写适配器 API，将其封装成统一的数据库访问接口。这种适配器技术通常需要实现不同数据库系统的 API 调用和命令的转换，以实现对各种数据库的访问。

（2）SQL 转换器：SQL 转换器是一种特殊的适配器，它能够将一个类型的数据库系统的 SQL 语句转换为另一个类型数据库系统所需的 SQL 语句。这种转换包括语法转换、函数转换、数据类型转换等，以实现跨不同数据库的查询和操作。

（3）数据格式转换器：数据格式转换器用于将一个数据库系统中的数据格式转换为另一个数据库系统中所需的数据格式。这种转换包括对数据类型、数据编码、数据存储结构等的转换，以确保数据在不同数据库系统之间的兼容性。

（4）数据同步器：数据同步器是一种特殊的适配器，它能够实现不同数据库系统之间的数据同步和复制。通过捕获源数据库系统的变更日志或操作记录，并在目标数据库系统上执行对应的操作，实现数据的同步和一致性。

需要根据具体的多类型数据库环境和需求选择合适的适配器技术实现。重要的是确保适配器能够处理不同数据库系统的语法差异、数据格式差异和性能差异，保证数据的准确性和一致性。此外，进行充分的测试和性能优化也是确保适配器技术实现成功的重要步骤。

本文采用 JDBC 实现数据库结构适配与同步，虽然国产数据库大部分都有 JDBC 规范接口的实现，但是各自的实现所支持的语法格式却不尽相同，比如不同数据库间存在关键字不统一、支持的数据类型不一致、SQL 语句语法格式不相同、自带的函数存在差异、支持的数据库操作不完全相同等。

为了实现数据库结构的同步适配，所描述方法依赖倒置的原则，采用工厂方法模式设计出了数据库适配模块，DBADAPTER 是默认的实现类，KingbaseAdapter、PostgreSQLAdapter、DamengAdapter、ShentongAdapter 等为不同国产数据库本身具体的实现类，通过重写 DBADAPTER 类中的模板方法来定义自身具体语法。通过 Java 的多态原理来实现数据库的兼容适配，屏蔽不同数据库语法的差异性，运行期间，根据应用的数据库连接环境，工厂类 DbAdapterFactory 动态生成具体数据库适配实体类，来实现不同数据库语法结构的适配兼容，如图 1 所示。

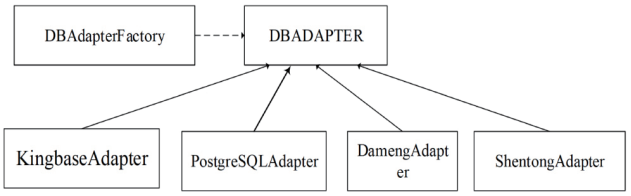

图 1 实体类关系

为实现适配器的高扩展性，采用开闭原则设计，针对新类型的数据库，只需要继承默认数据库适配器默认实现类模板，重写相关语法模板方法即可，便可方便地扩展更多的国产数据库

两个数据库之间的数据结构适配，采用符合标准规范的 XML 作为中间产物，从而屏蔽不同步数据之间的语法结构差异，将多个数据库的语法适配关系由多对多的关系变为一对一的关系，具体数据库只需要关注与 XML 文件的交互即可，如图 2 所示。

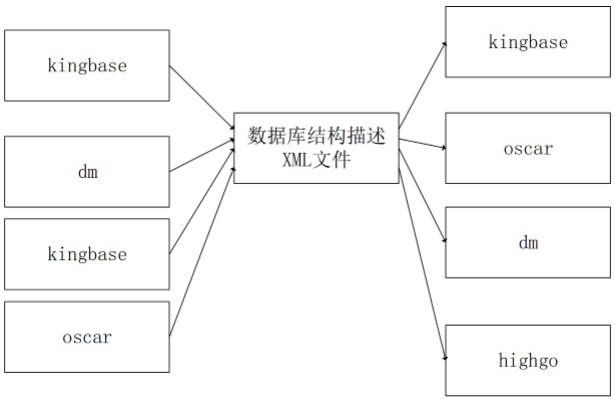

图 2 实体类关系

该 XML 包括数据库常用的表结构（表结构信息、主键信息、索引、约束、视图等信息）来完整的描述数据库结构信息，如图 3 所示。

该 XML 相关子结构信息如下。

模式集合：数据库类型名称。

模式：模式下的表结构信息，视图信息。

表结构信息：表名、主键信息、列集合信息、索引集合信息、约束集合信息。

主键信息：主键名、主键相关列及其索引顺序。

列信息：包括列名、类型 (JDBC 标准规范下 SQL 字段标准类型，未知类型为 -999)、类型名、字段长度、精准度、非空描述、默认值信息。

索引信息：索引名、索引类型（包括 B-TREE 索引，CLUSTERED 聚簇索引、HASHED 哈希索引、OTHER 其他类型的索引）、唯一索引标志、索引列集合。

索引列信息：索引设计列及其索引顺序。

外键约束：约束名、外键列、外键名、主键表、主键列、更新 / 删除规则。

图 3 XML 文件

数据库结构信息的导入导出都以 XML 文件为模板，将不同数据库语法差异的关系由多对多变为一对一，通过符合规范的 XML 数据库结构描述模板为中转站，以此屏蔽不同数据库之间的语法结构差异，达到跨平台的效果，不同数据库之间结构适配流程，如图 4 所示。

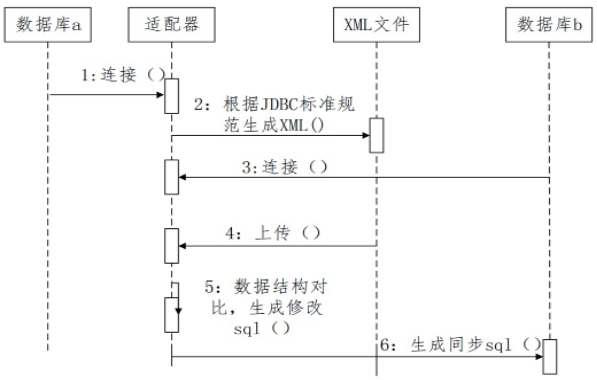

图 4 结构适配流程

以数据库 a（基准数据库）向数据库 b（目标数据库）同步数据库结构为例。

（1）选择数据库 a 作为基准数据库，输入驱动类名称、URL 名称、用户名、密码等信息后，使用适配器连接数据库 a。

（2）选择需要同步的相关模式，导出标准的 XML 数据库结构描述模板。

（3）选择数据 b 作为目标数据库，输入驱动类名称、URL 名称、用户名、密码等信息后，使用适配器连接数据库 b。

（4）向适配器上传步骤 2 导出的 XML 数据库结构描述文件。

（5）选择需要同步的模式，生成相关差异性的 SQL 语句。

（6）在数据库客户端执行 SQL，或者直接同步，至此，数据库 b 的数据库结构已与数据库 a 同步。

1.3 多类型数据库同步方法具体实施流程

1.3.1 数据库初始描述文件导出

数据库初始描述文件导出的描述是指从数据库中提取出的用于描述数据库结构和内容的文件。这个文件可以包含数据库表、列、索引、约束等的定义信息，以及初始数据的导出。通常，数据库初始描述文件的格式可以是 SQL 脚本、XML 文件、JSON 文件等，具体格式取决于数据库管理系统的支持和要求。

开发人员运行数据库适配器工具 JAR 包，首先，输入应用界面网址 http://IP: 端口号 /dbAdapterView/out，进入数据库结构导出操作界面，选择一个数据库作为基准数据库，输入数据库连接信息，如驱动类名称、URL 名称、用户名、密码等信息，点击测试连接，提示已连接，说明数据库连接成功，反之连接失败。然后，选择导出的服务名（或手动增加数据库服务模式名）对应的模式名，勾选后点击导出文件，即可导出选中的模式对应的数据库描述 XML 文件，该文件就是当前连接数据库对应模式的数据结构 XML 描述文件，该 XML 文件包含数据库的表结构、数据、视图等信息。数据表结构导出 XML 界面，如图 5 所示。

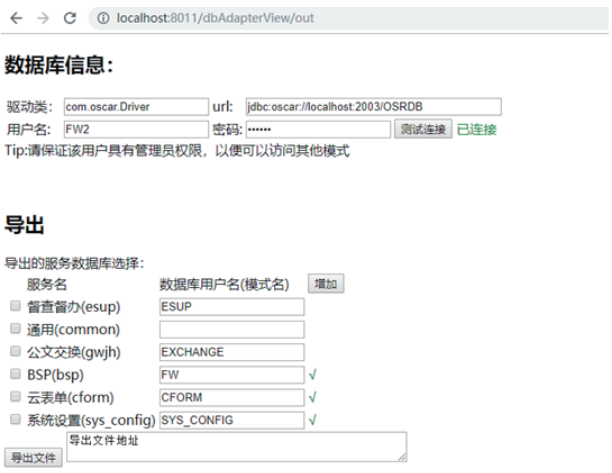

图 5 XML 导出界面

1.3.2 数据库结构同步升级

项目实施人员运行数据库适配器工具 JAR 包，输入应用界面网址 http:// http://IP: 端口号 /dbAdapterView/syn，输入需要同步的数据库连接信息，如驱动类名称、URL 名称、用户名、密码等信息，点击测试连接，提示已连接说明数据库连接成功，反之连接失败。连接成功后，点击选择文件，将最新的 XML 数据库结构文件导入，页面将会显示该 XML 文件中包含的所有模式，选中需要更新的模式，点击生成 SQL 文件按钮，即可生成将数据库同步到基准数据库所需要执行的 SQL 语句，将该 SQL 语句在数据库管理工具中执行即可；或点击直接同步按钮，即可直接以 XML 文件结构为标准，对目标数据库进行表结构的同步。数据库表结构导入界面，如图 6 所示。

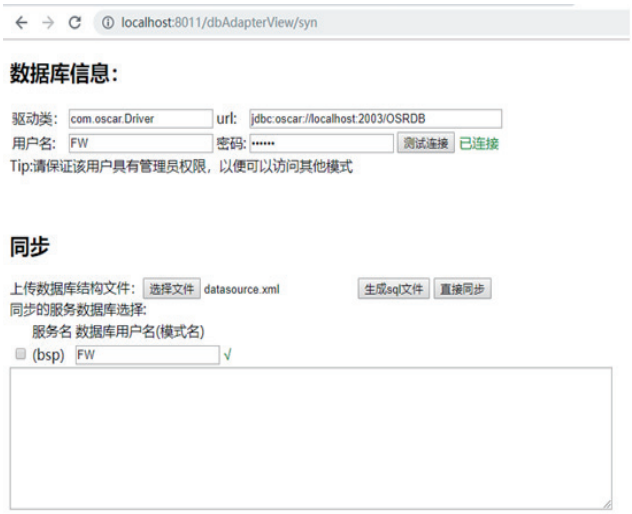

图 6 数据库表结构导入

2 结语

在多类型数据库环境中，实现数据的同步和适配是非常重要的任务。基于 CPU 环境的多类型数据库同步方法、装置及设备可以为数据库系统之间提供高效、可靠的数据传输和转换策略，在保障数据一致性和安全性的同时，提升了数据库系统整体性能和可操作性。

通过选择适当的数据库同步方法、装置及设备，可以帮助用户有效地实现数据库之间的数据同步和适配。在相应的过程中，需要充分考虑不同数据库类型的差异和兼容性问题，并结合实际需求制定合适的同步策略。

本文在基于 CPU 环境的多类型数据库同步方法、装置及设备的研究中，以更高效、更可靠的数据同步技术为目标，取得了一系列显著的成果。通过深入分析和探索，成功开发出了一种创新的数据库同步方案，为用户提供了更全面、更灵活的数据同步解决方案。

本文的研究不仅聚焦于单一类型数据库之间的同步，还致力于实现多类型数据库之间的无缝同步，包括关系型数据库、非关系型数据库、图数据库等。本文提出了一种高效的数据转换和映射机制，使得不同类型数据库之间的数据能够得到准确地转移和同步，满足了用户在不同业务场景下的灵活需求。

在实验和测试中，本文验证了该方法和装置的良好性能，与传统的数据库同步方法相比，在数据同步效率、可靠性和灵活性方面都得到了显著提升。本文的研究对于促进数据管理和应用的发展，提高数据库运维效率，具有重要的实际意义和应用前景。

综上所述，本文的研究不仅深入探索了基于 CPU 环境的多类型数据库同步方法，还设计和实现了相应的装置和设备。这些成果将为数据库同步技术的进一步发展提供有力支撑，同时也为促进数据科学与人工智能的融合发展做出了积极贡献。期待后续研究能够在更广泛的领域为用户提供更优质的数据管理和应用解决方案。

参考文献：

[1] KASAHARA H, NARITA S. Practical multiprocessor scheduling algorithms for efficient parallel processing[J].IEEE transactions on computers, 1987,33(11):1023-1029.

[2] 寇媛媛，王晓明 . 数据库同步技术的研究与应用 [J]. 传感技术学报 , 2016, 29(6):927-933.

[3] LIU Y, ZHANG X, LI H, et.al. Allocating tasks in mult-i core pro-cessor based parallel systems[C]//Proc of IFIP International Confer - ence. Piscataway:IEEE, 2007:748-753.

[4] 许华虎，季永华 . 基于 XML 的分布式异构数据库数据同步系统的研究 [J]. 计算机工程与应用 , 2005, 41(5): 184-186.

[5] 张月琴，袁新坤 . 一种异构数据库同步技术的研究与实现 [J]. 微计算机信息 , 2008, 24(33):182-184.

【作者简介】

张浩（1989—），男，硕士，工程师，研究方向：人工智能、数据库、大数据。

（收稿日期：2023-11-28）