

文章编号: 1005-8451 (2008) 04-0044-04

主数据管理技术在高速铁路信息化上的应用

于卓桐, 刘峰

(北京交通大学 计算机与信息技术学院, 北京 100044)

摘要: 为解决跨应用程序系统中数据一致性的问题, 对 MDM 主数据管理技术进行研究, 介绍技术相关内容, 包括其概念、功能架构、核心思想和应用流程等。深入分析 MDM 思想的精髓, 结合目前高速铁路信息化的行业需求, 提出一种 MDM 的技术应用方案, 该方案拟在构建基于 MDM 的动车信息共享平台, 以实现不同机构对数据访问的需求, 保证铁路行业内各业务应用系统的数据一致性。通过深入调研, 设计出平台的物理结构, 并提出 MDM 数据管理模型以及相应的解决方案。

关键词: 主数据管理; 数据一致性; 动车共享信息平台; 数据管理模型

中图分类号: TP393.07 文献标识码: B

Application of master data management technology to informatization of high speed railway

YU Zhuo-tong, LIU Feng

(Computer and Information Technology Department, Beijing Jiaotong University Beijing 100044, China)

Abstract: In order to solve the old problem of data consistency across Application System, it was researched the technology of Master Data Management and introduced the relevant content including the concept, function framework, core principle and application flow. Analysed the kernel of the MDM method and the requirement of the high speed railway information age, it was outlined a MDM solution which purpose on constructing the motor car sharing platform for EMU in order to make sure the data consistency in the Different Railway Industry Application System. It also gave the physical construction, MDM data management model and solution.

Key words: master data management(MDM); data consistency; moto car sharing platform; data management model

目前, 越来越多的企业和行业面临跨业务应用系统的数据共享问题, 为保持行业范围内各个不同的业务系统更大的数据一致性, MDM (主数据管理) 技术应运而生。很多企业被这个 MDM 的想法所吸引, 并将其以中间件的形式运用到系统的基础框架中, 通过对主数据的统一管理, 维护整个企业内部跨地区不同业务应用系统的数据一致性, 这是至关重要的, 它对于企业的财务运营, 销售市场, 采购和供应链管理以及信息技术方面都有积极的影响。

收稿日期: 2007-08-02

作者简介: 于卓桐, 在读硕士研究生; 刘峰, 教授。

处理。本文所阐述的研究结果已成功应用于沈阳地铁 AFC 系统中。

参考文献:

[1] 上海市工程建设规范—城市轨道交通自动售检票系统通用技术规范 (DGJ08-1101-2007)[S].

1 MDM 技术

1.1 概念

MDM (Master Data Management) 是一系列的规则、方法、应用和技术, 通常应用于企业管理应用系统, 目的是为了协调和管理与企业的核心业务实体相关的系统记录和系统使用过程中的数据和元数据。主要针对跨遍布企业中各个角落的应用程序, 创建和维持准确、统一的核心商业实体视图的商业流程、应用程序和技术集成体系架构。MDM 可以通过生成或管理一个中心级数据库, 或通过建立一个存有所有需要数据属性的元数据层, 以此作为

- [2] DB 31/2004 城市轨道交通自动售检票系统通信数据接口规范第一部分: 车站终端设备与车站计算机/车站计算机与线路中央计算机/线路中央计算机与清分系统[S].
- [3] Michael Owens. The Definitive Guide to SQLite[M]. Apress, 2006.
- [4] 于明. 城市轨道交通自动售检票系统检票机应用软件架构设计与分析[J]. 城市轨道交通研究, 2007 (4): 37.

一个主记录系统使用，从而有效地保证各级数据库中不同数据的一致性。

1.2 MDM 功能架构

如图1，通过采用主数据管理技术，对分布式异构的各种不同的物理数据库中的各类信息进行合理有效地整合，并同时提供一个内容存储库，其中包含所有数据的元数据以及用于维护这些数据的统一策略、流程和工作流。此存储库默认为不受其他应用系统管理的主要信息，可部署为负责同步跨功能、跨地域应用程序的信息记录的虚拟系统或物理系统。



图1 MDM功能架构

MDM的管理范围包含了全面的结构化和非结构化主要信息（关于产品、客户和供应商等）、数据仓库信息和内容存储库。与其说MDM是软件，不如说它是业务战略，它跨越了软件、业务流程、信息工作、行业规则以及业务转换。

1.3 实现功能

(1) 实施一个可以从企业中任何位置访问与核心商业实体有关的信息的数据整合平台，使得分公司遍布世界各地的企业可以在任意终端的应用系统中实时访问到相应权限的总部核心数据，实现了整个企业一致的数据共享平台。

(2) 根据行业规则自动创建一个正确的逻辑视图，它正确反映了真实世界的需求，对于核心机构的

决策支持提供了唯一真实、准确、全面地数据来源。

(3) 实时为行业应用程序套件提供高质量的主数据，这些主数据记录着系统所有应用数据的属性，保证了主数据的一致性，也就有效保证了所有系统中应用数据的一致性。从而有效避免各个不同的业务系统产生信息孤岛，维持数据的高度准确性，实时性和完整性，并同时保证用户可以得到唯一真实的数据来源。

1.4 核心思想

实际上，主数据管理是一个联合的商业应用程序，同时又是一种方法学的体现。一方面，MDM已经包括在传统和打包的企业应用程序中。另一方面，它属于一个数据整合问题，要求企业业务数据能够跨应用程序实时可用，从而能够满足核心商业实体的需求。无论是降低成本、满足标准的合法需求、与合作伙伴交换“标准”数据、整合应用程序、实施新的应用程序还是从企业应用程序中创建更多的价值，都需要以统一的方式来处理主数据。为了管理好这部分主数据以及和其关联的所有其他业务数据，成功的关键就是数据整合开发和相关的商业流程联合的方法学。由方法学构成的整个解决方案需要利用现有的传统应用程序和企业应用程序，同时集成在整个企业中实时获得、转换和提供主数据需要的商业应用程序和集成业务。

这种引入了方法学式的数据整合思想是MDM成功的关键，除此之外又组织和捆绑了流程、技术、资源和交付件，这些过程协调一致组成了主数据整合业务的整个流程。其中，流程是指构建和实施数据整合解决方案需要的步骤以及步骤之间的关系；技术是指在每个流程步骤应用于处理数据的功能组件；资源是指应用技术和执行流程需要的技术人员，包括明确定义的角色和职责；交付件是指各个流程步骤和各个阶段生成的产品，实现数据的获得、转换和供应的自动化。

(1) 通过数据分类和分析，初步提取获得一整套主数据；

(2) 从技术和行业规则的角度来逐个源的了解主数据的当前状态，完成数据评估和验证；

(3) 数据联合定义基于特定的源来获得、标准化、验证和映射主数据需要的处理，以实现初步装载、批量更改、事务处理和实时更新；

(4) 定义跨多个源的主数据，依次完成复制、基于参考源的进一步增强、记录联合、整合、留存、转

换、聚合和映射规范, 以实现初步装载、批量更改、事务处理和实时更新, 以此作为数据一致性的保证;

(5) 解决方案实施构建、测试和实施主数据整合解决方案;

(6) 长期运行在运行设置中规定了主数据的处理、测量、维护机制和报告。

2 在高速铁路维修信息化中的应用

2.1 行业背景分析

随着高速铁路信息化建设与运营的逐步深入, 动车组维修、运用、管理等信息资源将会不断累积和膨胀, 对动车组信息资源进行整合和综合利用, 已经成为进行规范化、深化信息系统建设的关键和难点。传统的数据管理方法是通过各部门自己信息系统收集资产数据, 这样容易产生“信息孤岛”, 经常出现信息的重迭现象, 而这种情况也会随着业务部门和新的业务的增加而逐渐加剧。通过采用 MDM 技术对各段所得应用系统中数据进行集中管理, 铁道部就可以对动车组等资产信息获得一个“唯一的真实数据来源”, 所以整合以后的数据底层有统一的数据编码和数据模型。铁道部高速铁路动车组的管理部门就可以通过访问部中心的核心应用程序而得到动车组的运用、维修、履历、财务成本等资产信息。

2.2 数据分析

高速铁路系统的组织结构分为铁道部、二级管理机构和动车段(所)三个级别, 这三级组织成树形结构分布, 其中根节点为铁道部, 第 2 级为铁路局和客运公司, 最底层为 4 个动车段和 20 个运用所。在这个结构框架中, 每一层产生的数据在数量、粒度及面向应用的具体程度上各不相同, 且不同级别之间的数据具有明显的导入导出关系, 上级数据由下级导出。通常情况下, 上级定期或不定期的下发规范准则和决策类的信息, 而下级也不断地向上级提交实时业务反馈信息。因此, 维护不同源的数据的一致性和实时性成为信息平台的难点所在, 结合上文所述, 采用 MDM 主数据管理的方法, 来保证不同机构应用系统中访问数据的一致性, 以及铁道部中心获得唯一真实的数据来源。

2.3 MDM 应用模型逻辑结构

基于 MDM 的动车信息共享平台的模型架构包

含 3 个主要层次如图 2。

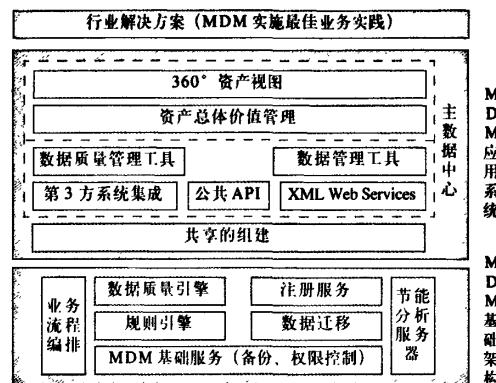


图 2 动车信息共享平台模型的层次

2.3.1 MDM 基础架构层

MDM 基础架构层为 MDM 应用程序提供基础的技术和服务, 主要包括数据质量引擎, 注册服务, 规则引擎, 数据迁移服务和其他一些重要的 MDM 基础服务, 如容灾、备份和权限控制等服务。同时也需要完成整个业务流程的编排和初步的智能分析。各组件详细介绍如下:

数据质量引擎主要服务于 MDM 应用程序层的针对行业业务实体的数据质量管理, 在此它只提供预备服务, 包括匹配、合并、清洗、配置和标准化等。其作用在于引导经过初步处理的数据进入上一层次的实体数据质管理阶段。

规则引擎主要负责决策以下情况, 集合同源信息, 处理源系统冲突以及根据数据质量规则调整数据的可见度使其应用于 MDM 的应用系统中。

数据迁移, 因为主数据是支持互操作性的环境的, 所以数据迁移为主数据管理提供了一个 ETL、SOA 和事件驱动过程的融合。它主要用于支持变化数据的捕获, 确定事件驱动的体系结构和提供 Web 服务。

注册服务用来维护元数据并同相关信息追溯到跨企业的数据源。它可以快速实时建立主数据的统一视图, 并且通过实时匹配和查询工具溯源得到真正业务数据, 并将其根据权限提交上层。

2.3.2 MDM 应用系统层

MDM 应用层主要包括针对动车组等资产建立的主数据中心和一些为本层和上层服务的共享组件。另外本层建立了 MDM 智能分析机制和预制的“端到端”的流程集成机制。本层主要负责数据质量

管理服务,整合服务,制定规则服务。在MDM基础机构层提供的服务的基础上,结合高速铁路行业规则,和动车组具体的核心数据模型,为主数据管理提供更进一步的服务。

本层中针对动车组建立的主数据中心是整个MDM模型的核心,从逻辑上大致分为3层,最低层为基础的核心数据模型,在本系统中体现为针对动车组资产特点建立的动车组资产数据模型;第2层是整个资产数据中心管理的核心层,主要负责对数据集成和对数据质量进行管理,具体还分为对外接口层和公共服务层。其中对外接口部分包括基于XML Web服务的,公共API接口的以及第3方系统集成的各类不同的接口方式,公共服务层主要是指基于MDM机制的数据质量管理工具和数据管理工具。第3层是客户应用界面层,主要为特定客户提供360°的资产视图以及对资产总体价值的管理功能,此部分功能可以集成在本系统中也可以独立开发一个单独的客户应用界面。

2.2.3 解决方案

由于各动车段所、铁路局和铁道部等不同机构,地理位置不同,业务系统不同,存储数据的数据库结构不同。因此,为实时维护数据的一致性,统一处理改变数据,经由MDM输入服务器,数据管理器,同步服务器后,将处理后的主数据提交MDM存储器。首先,某个业务应用系统触发了对某动车组资产主数据的改变,MDM机制将开始集中对主数据进行查重、合并、清洗、丰富和交叉关联等操作,最终从多个业务系统中整合最核心的、需要共享的数据(主数据),以服务的方式把统一、完整、准确的主数据分发给企业范围内需要使用这些数据的业务系统、业务流程和决策支持系统。方案流程如图3。

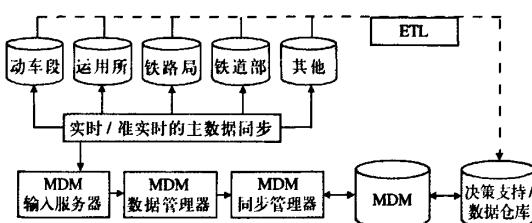


图3 解决方案流程

首先提取某个业务系统触发的改变的主数据,然后发送至预制的集成流程,根据流程过程,系统将统一格式的信息送至输入端口;MDM输入服务

器接受提交数据,根据数据来源的角色决定数据处理,将数据转化、映射、匹配后提交至MDM数据管理器,在此主要完成数据查重和合并,修改或丰富来源数据,而后交由MDM同步服务器处理;同步服务器完成转化和映射数据后将数据送至输出端口,其中转化主要是指将现有主数据转化为即将存入MDM存储库的主数据,并完成映射工作。处理后的数据重新回到流程中,流程触发客户端系统处理输入的数据,将整合之后完整、准确的主数据分发给所有相关的业务应用系统,同时为数据仓库提供了实时可靠的数据源。

经过以上载入主数据、整合主数据和分发主数据的处理过程,最终有效实现了数据集成,在分散的系统间保证主数据的一致性,使得铁道部和其他业务系统能够获得唯一真实的数据来源,并同时保证了各动车段所信息系统信息共享的实时性与一致性。另外MDM机制运用灵活,能够适应业务需求和IT架构的变化,有效避免了多种不同软件系统造成系统复杂性,同时更节约了高昂的维护成本。

3 结束语

MDM不是一种新兴的技术,而是一个解决老数据管理问题的新方法。传统的信息架构注重公司基础架构和系统的设计,并且一般都包含数据同步等功能。与传统方案相比,MDM包含数据管理技术以及MDM基础架构工具,为解决问题提供了一个更为灵活的方案。

掌握和运用MDM技术的关键就是针对应用行业业务特征建立相应的数据管理模型,本文分析了高速铁路维修行业的发展现状,应对其对数据资源共享的需求建立了动车信息共享平台,并建立了相应的数据管理模型,来维修不同地点的业务应用系统所访问数据的一致性,同时保证了铁道部能够得到唯一真实的数据来源,避免了传统的信息孤岛现象,实现了数据的一致性、实时性和准确定要求。

参考文献:

- [1] 李学伟, 汪晓霞. 中国铁路信息资源理论基础[M]. 北京: 清华大学出版社; 北京交通大学出版社, 2004.
- [2] Robert Rich. 使用IBM WebSphere的数据整合平台来管理主数据[Z]. 中国: IBM 白皮书.