

面向数据库的中间件

——数字图书馆关键技术

付鸿鹄

(中国科学院文献情报中心 北京 100080)

【摘要】 结合数字图书馆建设中数据库访问的问题,介绍了面向数据库的中间件这一关键技术,并对有关产品进行了分析比较。

【关键词】 数字图书馆 中间 数据库

Database—Oriented Middleware

—The Key Technology of Digital Library

Fu Honghu

(The Documentation and Information Center of CAS , Beijing)

【Abstract】 This article introduces one of the key—technologies of digital library: Database—oriented middleware and several products of it.

【Keywords】 Digital library Middleware Database

1 前言

自90年代初,数字图书馆被提出以来,关于数字图书馆的理论和实践研究已经有了极大的进展。数字图书馆将成为“未来的决策网络和应急知识网络”。随着Internet的发展壮大,数字图书馆的建设实施也在紧锣密鼓地进行中。网上信息源的分散和系统结构的差异给数字图书馆的建设带来很大的困难。因此,对各种不同结构的数据库实现透明访问是数字图书馆建设中的一个极为重要的问题。而面向数据库的中间件正是解决这一问题的必由之路。

2 面向数据库的中间件

2.1 中间件

中间件指的是一些系统软件,它们能使最终用户和开发人员觉察不到应用程序所使用的各种服务和资源上的差异。如果一个计算环境由多个开发商提供的产品组成,那么这些差异可能是开发商产品之间的差异或应用程序需求之间的差异造成的。中

间件的目的是通过为异质计算环境中的服务和资源提供统一、一致的观察结果,简化用户界面。中间件在为同一平台或不同平台上使用不同开发商产品的最终用户或者开发人员创造了浑然一体的连通性。因此,确切的说,它是透明的,开发人员和用户看不到它。中间件可以分为以下几类:

拷贝管理:数据提取、数据转换、复制、传播

网关:数据库和独立网关

程序到程序中间件:远程过程调用(RPC)、事务监控程序(TP 监控程序)、对象请求代理人

面向消息的中间件:提供可靠的、同步的、松散连接的交流服务,使得应用程序通过消息、队列、许可等交换信息。

2.2 面向数据库的中间件

面向数据库的中间件,简而言之,就是指一切连接应用程序和数据库的软件。与一般的中间件一样,面向数据库的中间件允许开发人员通过单一的、定义良好的API访问另一台计算机上的资源,如数据库服务器。

例如,如果一个关系数据库中的数据要被作为一个对象来访问,面向数据库的中间件可以将存储在关系数据库中的信息映射成源应用程序或目标应用程序可以访问的对象。对于其它类型的数据库

(如:层次型数据库、多维数据库等)也是如此。

面向数据库的中间件还可以提供对任意数量数据库的访问,而不需考虑数据库的模型和运行平台,这样无论是哪一种数据库,SQL Server、DB2、Oracle 还是 Sybase,都可以同时通过同一界面进行访问。通过这种机制,就可以把不同类型的源数据库和目标数据库映射成相同的模型,使他们易于集成。

面向数据库的中间件提供很多重要的功能,包括:

应用程序接口

将应用程序语言转化为可被目标数据库理解的语言,如:SQL

通过网络传递数据库查询请求

在目标数据库中执行查询

将响应集(查询结果集)通过网络返回到请求应用程序

将响应集转化为请求应用程序可以理解的格式

2.3 面向数据的中间件的类型

目前面向数据库的中间件有好几种类型,但基本上都属于本地中间件、呼叫层接口(Call Layer Interface, CLI)和数据库网关。

本地中间件是为特定数据库设计的,如 Sybase 设计的从 C++ 访问 Sybase 数据库的中间件就是一个面向数据库的本地中间件。由于是为特定的数据库设计的,面向数据库的本地中间件能提供最佳的访问性能。但也因此一旦用本地中间件建立了数据库连接,当要改变数据库时,需要对应用程序进行很大的修改。

呼叫层接口(CLI)如 ODBC 和 JDBC 提供多个数据库的统一界面。它可以把一般通用的接口呼叫转换成任意的数据库本地语言,也可以把响应集再转换成一致的表现形式以便向数据库发出请求的应用程序可以理解。

数据库网关能提供对大型系统内部数据的访问。它们可以从统一的应用程序接口集成多个数据库以便访问、重映射旧的数据库模型(如简单文件、ISAM、VSAM 等)并且对出入数据库网关的查询和信息进行转换。

3 几种常见的面向数据库的中间件

3.1 ODBC

ODBC 实际上并不是一个产品,而是微软在几

年前创建的一个标准。ODBC 是一个 CLI(Call Layer Interface),通过允许开发者制作一个在大多数关系数据库中可运行的简单的 API 调用,简化从 Windows(以及其它一些操作系统)到数据库的访问。

象所有的中间件一样,ODBC 提供一个定义良好的、不依赖于数据库的 API。使用 API 时,ODBC 通过一个驱动管理器来判定应用程序要连接的数据库的类型,并载入(或卸载)适当的 ODBC 驱动,这样,就实现使用 ODBC 的应用程序和数据库之间的相互独立。

ODBC 目前提供 32 位版本。大多数的数据库都有 ODBC 驱动。ODBC 是免费的,而其驱动不是。这些 ODBC 驱动可以从数据库供应商或第三方提供商购买。流行的应用程序开发工具大多通过 ODBC 提供对数据库的访问。实际上,ODBC 是微软 Visual Basic 以及其它开发工具与数据库连接的唯一途径。

3.2 JDBC

JavaSoft 的 JDBC 是第一个支持 Java 语言的数据库应用程序接口(API),功能上与 ODBC 相仿,提供 Java 开发人员一个从支持 Java 开发或支持 Java 应用程序运行的环境访问各种数据库的统一的接口。

JDBC 主要由两层组成:JDBC 应用程序接口(JDBC API)和 JDBC 驱动应用程序接口。JDBC API 提供从应用程序到 JDBC 管理器的通讯。开发者通过 API 使用标准 Java 机制访问数据库。数据库供应商提供 JDBC 驱动接口,也可以通过 JDBC-ODBC 连接桥使用传统的 ODBC 连接。

JDBC API 定义了一个 Java 类集,允许 Applet、Servlet、Java Bean 和 Java 应用程序连接数据库。通常,由这样一个小应用通过网络连接远程关系数据库服务器,如 Syabas、Oracle 或 Informix。这些由数据库供应商提供的原始的 Java JDBC 类与用户自定义的应用程序类共存,提供一种“纯 Java”的、可移植的数据库访问。这样就允许从任意支持 Java 的平台到任意数据库的连接。

JDBC 管理器和 ODBC 管理器一样,按 Java 小应用或应用程序的需要载入或卸载数据库驱动器。JDBC 支持单个或多个数据库服务器的连接。就是说,一个小应用可以同时连接本地的所有数据库和 Internet 上的公用数据库。

3.3 OLE DB

OLE DB 常被戏称为 ODBC 的“大哥”，它定义了一个数据访问服务器的集合，通过这些服务器可以很容易地连接到任意数目的数据源。这样，开发者就可以把多种不同的数据源作为单一的虚拟数据库来管理。OLE DB 允许使用标准的 COM 接口访问数据。

OLE DB 提供开发者访问关系数据库、文件、扩展表、电子邮件等数据的方法。通过 COM 接口，开发者可以使用 OLE DB 方便地集成面向对象的数据库和多维数据库。使用 OLE DB 时，数据库被简单的看作一个数据提供者组件。任何通过 OLE DB 使用本地数据格式和表现方法的组件都被视为数据提供者，包括关系数据库（使用 ODBC）、ISAM 文件、文本文件、E-mail、微软 Word 文件和数据流文件。这里的核心思想是生成独立的 OLE DB 组件对象用以调度数据提供者的各种附加特性和功能。这些独立的 OLE DB 组件叫做“服务提供者”。服务提供者类似于查询处理器允许应用程序使用它们实现不同数据联合（同族的或异族的）之间的互联。数据以单一的视图的形式存在，而不管它们的数据模型是关系型的、面向对象的还是多维的。

与数据提供者相对的是 OLE DB 数据消费者、单个数据提供者的应用程序或作用于任意多数据提供者的类属消费者。如微软的 Excel、Word、Project 都可以成为数据消费者。

3.4 数据库网关

数据库网关（也叫 SQL 网关）是一种应用程序接口（API），通过使用同一接口提供对运行在多种平台上的不同数据库的访问。它们类似于实际的数据库中间件产品，提供给开发者到任意数目数据库的访问接口，包括一些运行在典型的不易访问的环境下的数据库。如，通过一个 ODBC 接口和一个数据库网关，可以同时访问存储在大型主机环境下的 DB2 数据库、小型机上的 Oracle 数据库和 Unix 服务器上的 Sybase 数据库。开发者只需使用一个 API 调用，数据库网关就可以完成所有的其它工作。数据库网关把 SQL 调用解释成为标准 FAP（Format and Protocol）格式。FAP 格式通用的客户机和服务器连接，也是异质数据库和运行平台的通用联结。网关可以把 API 调用直接翻译成 FAP，把请求传递到

目标数据库并翻译以便目标数据库和平台做出响应。

目前市场上有很多数据库网关产品，如 EDA/SQL、RDA、DRDA 等。

RDA 并不是一种产品，它是一个开发者访问数据的标准。RDA 使用 OSI 并且支持动态 SQL 语句，允许客户端同时连接一个以上的数据库。但它不支持典型的事务相关服务。

DRDA 是 IBM 的一个数据库连接标准，它有许多数据库的支持如 Sybase、Oracle、IBI 和 Informix。与其它数据库网关一样，DRDA 提供便利的运行在多平台环境下的任意数目数据库的连接。DRDA 把数据库任务定义为远程请求、远程工作单元、分布式工作单元和分布式请求。DRDA 是一个定义良好的标准，它要求数据库符合标准的 SQL 语法，以便能够充分发挥 DRDA 在不同的系统、不同的情况下运行不同的数据库的功能。

4 结束语

随着计算机、通讯、多媒体、高密度存储等信息技术的发展和运用，尤其是网络的迅速普及和扩大，数字图书馆的建设已经成为各国信息水平的标志。而在数字图书馆的建设中，中间件的使用则是解决异质平台、异质环境、异质数据库的统一访问、统一存取的最佳方案，在数字图书馆的研究和建设已经有了应用，并起到了良好的效果。随着应用需求的发展，中间件的研究和功能也越来越被重视，应用也越来越多。面向数据库的中间件是数字图书馆的一项关键技术，必将在数字图书馆的建设中发挥更大的作用。

参考文献

- 1 刘晓白 孙卫 镇锡惠. 中国数字图书馆系统工程的宏观分析. 数字图书馆白皮书. 计算机世界报
- 2 David Linthicum. Database-Oriented Middleware. DM Review, 1999, (11)
- 3 Jean Bacon Richard Hayton Ken Moody. Middleware for Digital Libraries. D-Lib Magazine, 1998, (10)
- 4 Ivan Casanova. Message-Oriented Middleware: The Foundation for Business. DM Review, 1999, (10)
- 5 Joyce Bishchoff Ted Alexander 著. 成栋 魏立源译. 数据仓库技术. 电子工业出版社, 1998, 6