

C³I 异构型数据库通信平台实现技术

张水平 万应辉 卢潇 陈靖 邱晓奕 (空军电讯工程学院 4系计算机教研室 710077)

摘要:本文主要论述了异构型数据库通信平台课题的设计思想与实现技术,该课题主要是针对异构型数据库集成与互操作问题而研制的。最终在异构网络环境下为用户提供了一个界面友好、操作方便灵活,集数据转储、数据操作为一体的集成化平台。

关键词:异构型数据库 异构型网络 数据转储 命令翻译

一、引言

我们研制的 C³I 异构型数据库通信平台目的是为解决网络环境下异构型数据库通信问题。即在异构计算机网络环境中,用户可以对多个异构数据库进行完全透明的访问,并且支持在各种不同数据库系统之间进行数据转储和数据交换。异构型数据库通信平台屏蔽了各种数据库在物理上和逻辑上的差异,具有用户界面友好,操作方便灵活,是一个对异构型数据库有较强操作能力的集成化平台。

二、系统构成

1. 硬件平台

由于被集成到 C³I 系统中的现存的各类数据库是在不同机型、不同操作系统下开发的,因此,C³I 系统的生存环境必须是能够连接不同机型的异构网络。目前,国内外市场成熟并流行的网络类型有 Decnet 网、Novell 网和目前逐步占领市场的 WindowsNT 及 Windows95,因此我们研制的平台也必须兼有上述各类网络操作系统的异构网。其硬件环境结构由图 1 所示。

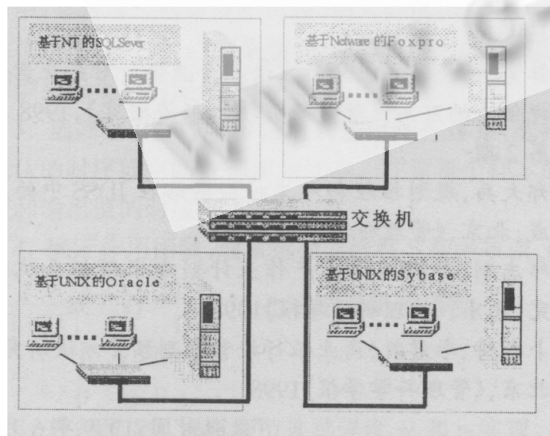


图 1 硬件平台

2. 方案介绍

异构型数据库通信平台要求提供对任意数据库资源的完全透明的访问,强调通用性和开放性,既要保证全局数据共享,又要保证各成员数据库的自治。在体系结构上可以有三种实现方案:

(1)对各种异构型数据库都建立用户交互接口,不进行任何模式的集成。这种方案虽然简单易行,但用户无法透明的访问数据,并且当增加一种新的异构数据库时,必须增加用户接口,非常繁琐,显然不可取。

(2)采用分布式数据库系统结构,将在物理上分布而逻辑上相关的异构型数据库通过分布式数据库系统来进行管理。分布式数据库系统具有独立性和分布透明性的特点,用户对任何数据库的操作都如同在本地进行,不必关心其数据模型、物理位置等细节。但是分布式数据库系统要求在各成员数据库之上建立一个全局模式,从而对整个系统实施统一的控制。该全局模式由所有成员数据库模式集成。若数据库是异构的,建立这个全局模式将相当困难。

(3)采用联邦数据库系统结构。这种方案是 Heimbigner 和 Mcleod 在 1985 年提出的。联邦数据库系统 FDBMS(Federated Database System),不采用全局模式,在维持局部成员数据库自治的前提下,对异构的成员数据库进行部分的集成,提供数据的共享和透明访问。它代表没有集成和完全集成的折衷方案,具有分布性、自治性和异构型的特征。

根据联邦数据库系统的特征,可以看出这种体系结构能较好地满足异构型数据库通信的要求,本平台的研制就借鉴了这种体系结构。

3. 功能构成

“异构型数据库通信平台”是一个利用面向对象技术开发的系统。作为一种新的发展方向,面向对象技术不

只是一种新的程序设计技术,它还是一种全新的设计和构造软件的思维方法。利用面向对象技术可以生产出模块化的、可维护性好的、可重复使用的应用程序,设计出通用性好的应用系统。

由于 ODBC 为访问数据库提供一个统一的标准。在此基础上,一个支持 ODBC 标准的应用开发工具可以相互存取任一支持 ODBC 标准的数据源(即可操作多种格式的数据)。所以在本平台的开发中,我们以 ODBC 作为异构型数据库访问的中间件技术,实现异构型数据库的访问。

在本平台的开发中,首要问题是实现转换处理器。转换处理器包括模式转换处理器、数据转换处理器和命令转换处理器。它们分别对成员数据库中的不同模式、数据和命令进行转换。

研制的软件平台结构由图 2 所示。

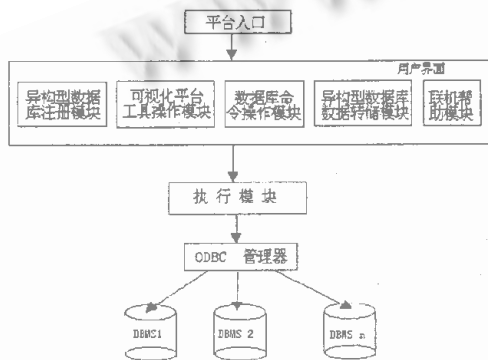


图 2 软件平台结构

4. 各功能模块介绍

(1)注册模块的功能。我们要求平台支持多种数据库管理系统,这就是平台的跨平台特性。当一种新类型的数据库要加入到平台下时,首先要使用平台所提供的数据库注册工具对这种新类型的数据库进行注册。即在客户端通过 ODBC 驱动器管理器对新类型的数据库进行静态配置。当一个数据库在网络的任何一个客户端注册后,就能够被网络上的所有用户所共享和访问。

(2)可视化数据操作模块的功能。可视化数据操作模块是平台提供给用户的一个可视化的数据操作工具,该模块具有友好的用户界面,用户可利用这些工具进行数据的追加、删除、修改的查询,所有的界面均采用 Windows 的标准风格,具有菜单操作和工具条操作两种方式,结果显示清楚直观,适合用户的浏览和查看。

数据操作模块由:数据源连接,数据删除,数据修改,数据追加及通用数据库查询五部分组成。其中连接数据库模块是其他模块的基础。而通用数据库查询模块可从大量的信息中按用户的要求提取所需的数据,该查询模块的特点如下:

①用户能够自由地选择要查询的内容。数据库中所有的信息都在用户的选择范围之内。信息能以比较自然的方式表示,用户不需知道数据库中表的结构,表之间的关系以及数据项的定义。所有数据库的实现细节对用户完全透明。

②用户能够方便、自由地定义查询条件。对日期型数据可按时间范围查询,对数值型数据可按照具体的值或范围查询,对字符型数据可以按照模糊条件查询。多个查询条件之间可以按照“并且”或“或者”的关系连接。对于指定的数据项能列出库中所有的值供用户浏览、选择。

③对于每一次特定的查询,用户可以在原来的基础上反复的定义、修改查询条件和查询内容。用户可以在上一次的查询结果之上继续增加查询条件,缩小查询范围,也可放宽查询条件,扩大查询范围。

④用户可对查询结果进行统计,如查询到的记录的个数,数据项的最大值、最小值、平均值。

⑤用户可对查询结果的输出格式进行定义,可进行直接显示,或将数据保存到数据文件之中,也可以报表形式直接进行打印,用户也可自行定义打印的格式。

(3)命令操作模块的功能。该平台的命令操作模块,提供给用户一个可编辑的命令窗口,用户选择命令操作语言类型后,可以在命令窗口中进行命令的输入,由平台对用户命令进行语法检查并进行翻译和执行,用户可以立即得到命令执行后的结果。

(4)数据转储模块的功能。异构型数据库数据转储是将一个数据库中的一个表或若干表的记录集传送并转储到另一种类型数据库的一个表中,在数据转储过程中要保持数据的完整性和正确性。我们所研制开发的异构型数据库转储模块,采用一种新的思路和方法,力求通用性,可以用于任意两种数据库系统中,解决了以前数据转储方法中所存在的问题。

(5)联机帮助模块的功能及实现。“异构型数据库通信平台”的联机帮助模块的功能是随时为用户提供详细的联机帮助信息,这些帮助信息包括平台各个功能模块的功能及使用方法介绍,平台可视化工具的使用说明及平台的相关技术支持等。通过使用联机帮助,用户可以

更安全、高效、方便地在本平台上进行异构型数据库的访问。

三、实现技术分析

1. 数据源的连接

在进行数据库操作之前必须首先进行数据源的连接,数据源的连接意味着向系统提交数据源的数据库管理系统类型等信息。只有获得了该信息,才能对数据源进行各种操作。因此连接数据源是进行数据操作的关键,也是整个平台其他功能模块的基础,其他模块运行环境及入口参数将在此模块正确执行后获得。

2. 库结构信息的获取

传统的 ODBC 编程过程比较复杂,各种参数不易理解,且直接获取返回的数据较困难。VC++ 5.0 的 MFC 类库对 ODBC 的 API 进行封装,部分简化了 ODBC 编程(尤其是对数据库记录集的操作),但单纯利用 MFC 类获取异构型数据库的结构信息仍然比较困难,因此需要将 MFC 和传统 ODBC API 编程结合起来。我们利用 ODBC 接口函数重载了 MFC 中 CRecordset 类的部分成员函数,创建 CTable 和 CColumns 类。利用这两个新创建的类,可以很方便的获取异构型数据库结构信息(库的表集合信息及表中的字段集合信息)。

3. 数据库的互操作

各种类型的数据库系统在彼此独立的情况下,只能使用本系统的命令来访问本数据库系统中的数据,如用 Foxbase 命令只能访问 Foxbase 的数据库。当各种不同的数据库系统集成在同一个网络环境中时,这种单一模式的命令访问方式将给用户使用带来许多不便之处。所以,实现操作语言的透明性是实现异构型数据库互访问的重要方面,即允许用户使用一种公共的语言就能够访问网络上的各种类型的数据库中的数据。要实现异构数据库系统的互操作,必须解决以下问题:

(1)命令集拆分:将用户输入命令窗口内的命令集拆分为若干个单条命令;

(2)命令翻译:将一种特殊类型的数据库命令翻译为可实现本命令功能的标准的 SQL 命令集或一个与该命令功能等价的程序段;为了完成命令的翻译,我们采用了命令解释器和命令词典等技术;

(3)DB 接口:执行翻译的 SQL 语句,通过 ODBC 接口从指定数据库中提取数据,并将命令操作的结果返回。

各种数据库命令的解释器都是不同的,它们主要进行各种类型数据库命令的翻译,将一种专用的数据库命

令翻译为各种数据库都能识别并执行的命令。根据 ODBC 理论,我们选择 SQL 作为公共的命令语言。

4. 异构型数据库数据转储的实现

异构型数据库相互通信问题中的另一个重要方面是异构型数据库的数据转储。异构型数据库数据转储也是实现数据库信息共享,充分利用已有资源的重要措施。在以前的文献中,在解决异构型数据库数据转储问题时有非常大的局限性,例如转储是单向的,即只能从 Foxpro 库向 Oracle 库进行转储,而不能进行双向的转储。出现这种问题的原因是转储算法本身就是一种针对某两种数据库系统而设计,并不考虑通用性所造成的。一般这种转储算法以分析被转储数据库的数据存储文件的格式为基础。这就造成只有数据文件格式公开的数据库系统才能作为源数据库,当数据库系统的数据存储格式是封闭的,则这种数据库系统就不能作为源数据库。这就是造成以前数据只能单向转储的根本原因。

我们所研制开发的异构型数据库转储模块,采用一种新的思路和方法,力求通用性,可以使用于任意两种数据库系统中,从而解决了以前数据转储方法中所存在的问题。为解决该问题,我们采用了如下方法:

(1)通过动态光标及动态 SQL 嵌入。应用程序可以根据需要动态构造 SQL 语句,建立相应的记录集,利用动态光标获取记录集中的数据信息,以便进行数据转储。

(2)通过动态数据窗口。由于异构型数据库数据转储中源数据库和目的数据库是根据用户需要来确定的,所以必须以动态数据窗口的方式显示和操纵数据库。通过动态数据窗口来确定动态的记录集。

(3)建立数据类型映射系统库。在异构型数据库数据转储中源数据库和目的数据库的字段类型和字段名称都可能存在着差别,这样在进行数据转储时还必须考虑到数据类型转换问题。为此,我们建立一个数据类型映射系统库,利用该库可以得知任意数据库间类型的转换关系。

四、本方案特点

本方案具有如下特点:

1. 灵活的扩充性:本方案通过命令转换字典管理系统管理员提供了扩充的工具软件。

2. 功能强大的互操作性:该平台真正作到了互操作,用户在本地结点上用自己熟悉的语言与网上任一异构库进行通信,而不必熟悉操作对象库的一套语言。

3. 各数据库成员的自治性:原有的应用程序极其数

数据库无须修改即可在本地运行,因而使其前序工作有了继承性。

4. 系统平台的高可靠性:由于转换目标程序借助了目标数据库管理系统的功能,因而其可靠性得到保障。

五、推广应用价值

本软件平台可以提供给用户多种数据库的访问方式,底层物理数据库之间的类型差异、语意差异等完全对用户透明。本软件平台具有友好的用户界面,用户使用方便、灵活。另外,本软件可用作客户机前端软件,可以作为一种底层数据库访问工具,即使不懂数据库的用户也可以使用本软件平台进行数据库的操作。因此,本软

件平台可以应用于 C³I 系统、各类办公自动化系统和校园网的建设中,具有很好的推广应用价值。

参考文献

- [1] 唐雪飞,熊平,刘锦得“异构数据库互操作性的概念及其面临的问题”《计算机科学》1994.NO 6
- [2] 孙志军,周明生“联邦式异构分布数据库的性能分析”1996 国际数据库会议纪要
- [3] 齐进,周伯鑫,徐南荣“实现异构型数据库集成的一种方法—原数据法”《计算机工程与应用》1998.NO 4

(来稿时间:1999 年 6 月)