

整合异构地震数据库服务的新方法*

刘 坚¹, 李盛乐¹, 戴 苗²

(1. 中国地震局地震研究所(地震大地测量重点实验室), 湖北 武汉 430071; 2. 湖北省地震局, 湖北 武汉 430071)

摘要: 整合异构地震数据库服务的新方法通过数据层采用透明网关技术实现异构数据库互连, 应用层采取跨平台、跨语言的 ICE 中间件技术, 实现并提供各种与业务相关的二次开发接口, 用户通过该接口可以方便快捷操作业务数据; 实验结果表明, 该方法不仅屏蔽了异构数据库的服务差异, 同时也大大提高了整合数据服务的效率。

关键词: 异构数据库; 地震数据整合; 透明网关; ICE 中间件

中图分类号: P315-391

文献标识码: A

文章编号: 1000-0666(2013)02-00253-05

0 引言

目前, 多数单位都同时存在多个数据库平台, 而每个数据库平台都运行着相关的一套或多套应用, 如何在不影响现有应用运行的前提下, 快速有效地整合不同数据库平台上的数据, 实现不同数据库之间的数据共享, 是很多单位或企业面临的问题。

地震行业系统中, “十五”前兆数据管理系统使用的是 Oracle 数据库、“九五”前兆使用的是 MS-SQL Server 数据库、测震使用的是 MySQL 数据库。如何在这样的异构环境下迅速、有效、经济的通过单一的应用程序来部署可能存在于许多不同系统中的数据, 从而提供一个全面的数据视图, 而无须考虑何种数据库或者操作系统能达到各种数据库之间协同工作(苏淑文, 翁敬农, 2004), 是行业或单位进一步发展所需解决的问题。针对上述问题, 要实现网络环境下的信息共享, 就必须集成多个数据库系统, 实现不同数据库之间数据的共享, 而异构数据库服务整合是数据库集成中的难点。

1 整合方法与相关技术

在整合技术与方法上有比较多的选择, 任连仲(2007)提出基于过程存储的接口方法, 基于自置控件的接口方法以及直接编程的数据交换的方法等; 近年又推出基于 SOA 描述的多种可

用于集成的软件和方法, 为减少开发成本和实现“零维护”成本, 又有专家提出基于 WEB 2.0 技术进行系统前台端整合。经研究发现, 这些方法或方案皆需要二次编程, 或者有较高的开发成本, 或者有较高的技术门槛, 或者需要较长时间周期, 对一些轻量级的系统数据服务整合不易操作和开展。然而具体到 Oracle 数据库上, Oracle 10g 因提供了多种信息集成解决方案(Oracle 公司网站, 2008), 可满足各种各样的数据服务整合需求。

1.1 整合方法模式

(1) 集中式整合, 是将其它所有数据“剪切”或复制到一个数据库, 集中到一个中心数据库进行管理。虽然整合采用最简单形式进行信息集成, 但其实最难以实现。因为它要求把其它系统的所有数据都移动到一个数据库中, 再进行集中管理。所以旧系统程序可能需要重新写代码, 或因数据平台的改变, 要完全放弃旧系统而开发新的应用系统(蓝永健, 2008)。

(2) 联合式整合, 该模式将数据保留在原地进行更新与维护。表面上, 集成数据在一个虚拟的数据库中, 但实际数据仍然保存在原来的位置。Oracle 数据库联合解决方案允许访问 Oracle 和非 Oracle 数据, 同时隐蔽了应用程序数据的实际物理位置。此方式无须移动任何数据, 因此联合是信息集成的最快方式。然而, 由于是异地保存数据, 访问异构数据库必然带来时间上的消耗, 所以它难满足热点数据在访问性能与速度上的要求。

* 收稿日期: 2012-09-02.

基金项目: 中国地震局科研业务专项(1819)和中国地震局地震研究所所长基金(IS201146098)联合资助.

(3) 共享式整合, 该模式是指在多个应用程序数据存储器和数据库中同时维护多个信息副本。Oracle 提供了许多信息共享功能, 使用这些功能可搜索到经常访问的数据, 从而减少响应时间, 消除对远程数据存储的依赖性。

1.2 普通连接与透明网关技术

上述 3 种整合方法都涉及到异构数据库之间的连接, 目前连接方式有普通连接和基于 Oracle 的透明网关, 两种方式实际是异构服务中代理的两种类型, 其大部分强大处理能力皆已融合在本身数据库中, 并提供了能够完全利用 Oracle 数据库功能的信息集成的有效解决方案。

普通连接是通用的解决方案, 透明网关是一种专为非 Oracle 数据库系统进行编码、量身定制的解决方案。它提供一种比普通连接拥有更好性能与更多功能的优化解决方案。普通连接依赖于业界标准, 而透明网关使用它的本地接口访问非 Oracle 系统。透明网关亦经过端到端的认证。Oracle 拥有的透明网关可以连接许多非 Oracle 数据源, 如 Informix, Sybase, My sql, Microsoft SQL server 等。

2 将 ICE 中间件引入的新方法

2.1 异构数据库互连

共享式整合思想核心是在不影响现运行应用系统和数据的情况下, 将现有不同数据库数据进行整合, 达到异构数据库间的互连与互通, 访问多个不同类型数据库像是访问同一种类型数据库, 大大降低多个不同数据库数据整合难度, Oracle 透明网关技术已经提供了成熟的解决方案。如图 1 所示, 即通过 Oracle 访问、操作远程非 Oracle 数据库, 该技术的实现包括三部分: 第一部分, 异构服务模块是 Oracle 数据库的内核部分, 负责大部分异构连接的处理; 第二部分称为“代理”, 由代理通用代码和驱动组成, 其中代理通用代码是对所有基于异构服务产品的通用代码。而驱动是指与非 Oracle 系统直接交互的模块, 实现从异构服务 API 到特定非 Oracle 系统 API 的映射。代理的位置可以与非 Oracle 系统在同一台机器上, 或与 Oracle Server 在同一台机器上, 或者单独在一台机器上; 第三部分是指非 Oracle 系统的远程数据库。这样就实现了通过透明网关操作 Oracle 数据库, 操作其他远程非 Oracle 数据库数据。

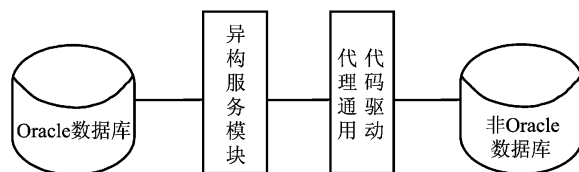


图 1 Oracle 数据库异构服务模块

Fig. 1 Oracle database heterogeneous services module

2.2 引入 ICE 中间件的作用

透明网关解决了异构数据库的互连互通问题, 本文讨论的是异构数据库数据整合, 且前提不能影响现有不同数据库平台上的应用系统, 这些应用系统可能是 B/S、C/S 结构, 运行的操作系统可能是 Linux、Windows 平台, 开发语言可能是 C++、C#、Java 等, 因此, 我们需要一种跨平台、跨语言且与具体实现无关的中间件, 通过它实现数据与业务之间的松耦合, 数据与业务分离, 互不影响; 只改变中间层的实现即可自适应数据与业务的变化。ICE 中间件可以轻易做到这些, 这也是将其引入用来解决数据异构服务整合目的之所在。

2.3 中间件 ICE 简介及其工作原理

ICE (Internet Communications Engine) 简写是一种现代的面向对象中间件, 可用于替代像 CORBA 或 COM/DCOM/COM + 这样的中间件 (Henning, Spruiell, 2009a, b)。它为各种有着苛刻的技术要求的应用提供了强大的网络基础设施; ICE 3.4 已实现对 C++, Java, Python, PHP, C#及 VB 的支持, ICE 定义了自己的规范语言 Slice, 它用于使对象接口与其实现相分离的基础性抽象机制; 客户和服务端可以用不同的编程语言编写, 可以运行在不同的操作系统和机器架构上, 并且可以使用多种网络协议进行高效通信。

相比其它的中间件技术, ICE 在架构上有如下的一些优势 (刘裕, 吴坚, 2004):

(1) 面向对象的语义。ICE 是完全建立在面向对象模型基础之上。

(2) 支持多个接口。对象可以提供多个不相关的接口, 同时又跨越这些接口、保持单一的对象标识。

(3) 支持同步和异步的消息传递。ICE 提供了同步和异步的操作调用和分派, 并且通过 ICE Storm 提供了发布—订阅消息传递机制。这样, 可以根据应用的需要来选择通信模型。

(4) 语言无关性。客户和服务端可以分别部署, 所用语言也可以不同。

(5) 无关性。客户及服务器与底层的机器架构屏蔽开来。

(6) 操作系统无关性。ICE API 完全是可移植的, 同样的源码能够在 Windows 和 UNIX 等操作系统上编译和运行等。

(7) 异步方法分派 (AMD) 是 AMI 的服务器端等价物, AMD 操作通常会把请求数据 (也就是回调对象和操作参数) 放入队列, 供应用的某个线程 (或线程池) 随后处理用。这样, 服务器就使分派线程的使用率降到了最低限度, 能够高效地支持大量并发客户。

总之, ICE 中间件使开发人员可以专注于提高应用程序的质量及开发效率, 可完成系统基础设施的部分任务 (汤晓燕, 徐竞, 2009), 其通信示意如图 2 所示。

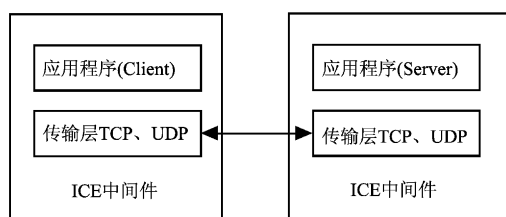


图 2 ICE 中间件通信示意图

Fig. 2 Schematic diagram of the ICE middleware communication

3 系统模型与系统实现

3.1 系统模型

本系统模型主要给出了在 Oracle 环境下如何

无缝连接 SQL Server 的实现, 当考虑 DB2、Informix、Mysql 和 Sybase 等其它异构数据库系统时, 因系统模型的实现原理也是相同的, 这里就不一一详述, 整个系统的架构如图 3 所示, 从图可以看出, 以 Oracle DB 为中心, 其上建立有连接 Ms-SqlServer 透明网关, 左边的现有的应用系统和新的应用皆通过 ICE 中间件访问各种新老业务接口, 这些接口通过 Oracle DB 透明网关操作 oracle 便可以方便快捷地访问远程的 Sql Server DB, 以此实现异构数据服务的跨平台、跨语言整合。

3.2 服务整合

所有的服务皆定制在服务器端, 通过如图 3 中间件接口向各种客户端应用 (服务客户 C/S 模式、基于浏览器的 B/S 模式) 提供新旧业务服务 (前兆分析预报数据服务、测震地震目录服务、共享信息网站服务等), 旧业务按旧业务定义服务接口, 新业务按新要求定义提供新的服务接口, 为此, 新旧服务皆可通过该接口访问, 从而统一提供服务, 这样就可有效将各种分散的应用服务整合至一起。

3.3 系统实现

从图 3 可以看出, 欲在整合的数据拓展业务应用, 两个关键点至关重要, 分别是透明网关的配置和 ICE 业务接口实现, 下面分别作具体实现介绍。

(1) 透明网关的配置流程

①在 Oracle Database 中附加安装 Oracle Transparent Gateway for Sql Server 服务, 且配置数据库别名;

②安装 Ms-SQL Server 客户端, 并添加别名;

③安装 Oracle Transparent Gateway for Ms-Sql Server 后, 可以在 \$ Oracle_ HOME 下看到 ora4msql

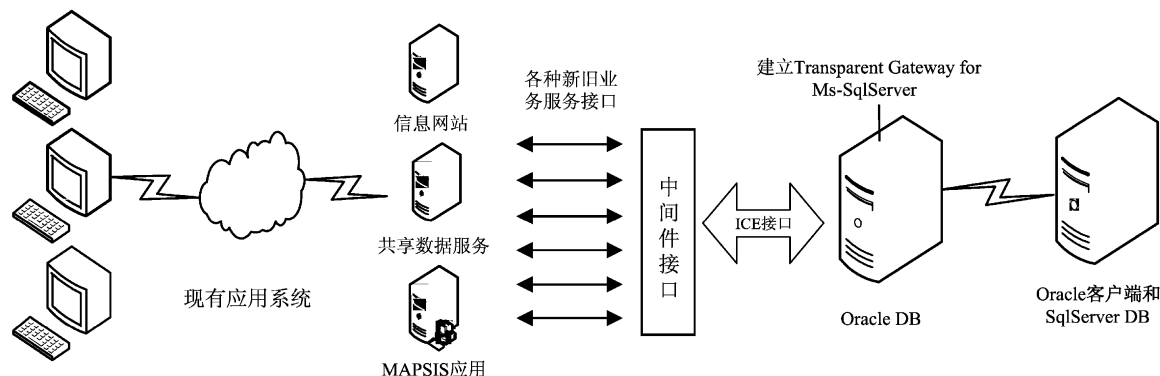


图 3 系统架构图

Fig. 3 System architecture diagram

目录, 编辑 \$ Oracle_ HOME/ora4msql/admin/initora4msql.sql 文件确认以下代码: EQ_ CONNECT_ INFO = 别名. 数据库名。如 “EQ_ CONNECT_ INFO = SqlServer_ Listener. DB_ EQ”, 其中 SqlServer_ Listener 是②中配置的别名信息, DB_ EQ 是①配置中的数据库名;

④编辑 Oracle 数据库相关文件参数。注意配置两个文件; \$ Oracle_ HOME/network/admin/listener.ora, 该文件是 Oracle 监听器进程的配置文件, \$ Oracle_ HOME/network/admin/tnsnames.ora, 该文件的配置主要提供的 tnsname 到主机名或者 ip 的对应。

⑤在 Oracle 上创建 Db Link。

(2) ICE 定义并实现接口流程

实现流程: ①编写 ICE 接口文件; ②选择开发语言, 生成相关类和接口; ③开发 Server 端; ④开发 Client 端; 最后部署 Server 端、运行 Client 端。几个关键部分实现介绍如下:

①Slice 定义

第一步是编写一个 Slice 定义, 其中含有应用所用的各个接口, 具体可参阅 ICE 官方文档说明, 定义的文件命名为 EQDBGateway. ice。

② Slice 定义编译映射至目标语言

Slice 定义编译可以映射至多种语言, 本文以映射到 C++ 为例, 编译上节的 Slice 定义, 生成 C++ 代理和骨架。在 UNIX 上, 编译定义为: \$ slice2cpp EQDBGateway. ice。slice2cpp 编译器根据这个定义生成两个 C++ 源文件: OperatorDB.h 和 OperatorDB.cpp。

OperatorDB.h 头文件含有与 OperatorDB 接口的 Slice 定义相对应的 C++ 类型定义。在客户和服务端源码中必须包括这个头文件; OperatorDB.cpp 文件含有 OperatorDB 接口的源码。所生成的源码同时为客户和服务端提供针对特定类型的运行时支持。例如, 它包含了在客户端整编参数数据的代码, 以及在服务端解编数据的代码。必须编译 OperatorDB.cpp 文件, 并把它链接进客户端和服务端。

③用目标语言实现服务器接口

因这部分接口的实现开发语言是可选的, 实现的方式也不尽相同, 每种语言实现方式的已经比较成熟, 相关参考资料也比较详尽, 具体实现代码在此省略。

4 采用 ICE 后的新旧应用对比

采用 ICE 后, 旧应用系统可以通过两种方式与原来数据库连通, 第一种应用紧耦合方式, 直接访问原来的数据库, 因此该种方式无需更改应用的任何代码; 第二种方式就是通过 ICE 访问 Oracle, 以 Oracle 透明网关去访问旧系统的非 Oracle 数据库, 从上述的原理已经得知, 将原系统实现功能代码原封不动拷贝至 ICE 服务端接口实现部分即可, 然后通过该 ICE 接口就可访问原数据库; 新应用就更方便了, 直接用欲实现的语言实现功能接口, 一次实现, 不限语言和平台的限制, 被任意客户端新业务接口调用。因此, 采用 ICE 中间件实现数据与业务之间的松耦合, 数据与业务分离, 互不影响; 只改变中间层的实现即可自适应数据与业务的变化, 不仅适应老系统的应用变化, 还可任意适应新业务功能的灵活扩充更新。

图 4 为通过综合库 Oracle 中间件接口访问测震 Mysql 数据库实例界面应用。



图 4 测震数据整合至 Oracle 库实例

Fig. 4 Instances of seismic data integrated to Orade library

5 结束语

笔者在开发层面采用透明网关技术实现异构数据库互连, 应用层面采取跨平台、跨语言的 ICE

中间件技术, 以达到很好整合异构数据服务之目的, 该方法的特点小结如下:

(1) 可访问不同异构数据库。使得应用系统避免因要访问不同的非 Oracle 数据库而引起的针对性量身定制开发, 减少开发成本, 并且增强了应用系统的可移植性。

(2) 数据无需重复存储。应用系统可以直接访问非 Oracle 数据库中的数据, 避免了大量数据在不同位置间的上传下载, 减少了数据重复, 节省磁盘空间。

(3) SQL 同时查询多个数据库。Oracle 数据库能接受查询多个不同数据库中数据的 SQL 语句, 并使用异构服务模块将相应的 SQL 语句直接发送给其它的 Oracle 数据库或非 Oracle 数据库, 合并结果后返回给客户端。

(4) 引入 ICE 中间件使业务系统与数据库相互独立, 降低了应用系统与数据库的耦合度, 在不影响现有应用系统情况下, 易扩展新业务应用。

(5) 应用开发灵活性高, 客户端应用可以是 ICE 支持的任意开发语言, 而无需更改服务器端任何代码, 皆可通过接入接口操作访问不同数据库。

本文撰写过程中得到高景春研究员、缪发军、刘胜国高级工程师、陈智勇工程师等的大力支持和热心帮助指导, 在此表示衷心感谢!

参考文献:

- 蓝永健. 2008. 利用 Oracle 透明网关技术进行系统整合的研究[J]. 广东教育学院学报, (5): 92-93.
- 刘裕, 吴坚. 2004. 中间件技术与 ICE[J]. 微机发展, (10): 37-39.
- 任连仲. 2007. 系统整合的两个关键[J]. 软件世界, (6): 76-77.
- 苏淑文, 翁敬农. 2004. Oracle 系统异构环境下的信息集成[J]. 微机发展, (10): 1-2.
- 汤晓燕, 徐竟. 2009. ICE 中间件的研究与应用[J]. 常熟理工学院学报(自然科学), (10): 116-117.
- Henning M, Spruiell M. Distributed programming with ICE [EB/OL]. (2009-02-20) [2012-09-02]. <http://www.zeroc.com/download/Ice/3.3/Ice-3.3.0.pdf>.
- Henning M, Spruiell M. ICE 分布式程序设计[EB/OL]. (2009-01-10) [2012-09-02]. <http://www.zeroc.com/download/Ice/1.3/Ice-1.3.0-C.pdf>.
- Oracle 公司网站. Oracle 数据库 10g 信息集成 Oracle 白皮书[J/OL]. (2008-04-07) [2012-09-02]. <http://www.oracle.com/global/cn/documentation/10g/database/InformationIntegration10g.pdf>.

New Method of Heterogeneous Seismic Database Integration Services

LIU Jian¹, LI Sheng-le¹, DAI miao²

(1. Key Laboratory of Earthquake Geodesy, Institute of Seismology, CEA, Wuhan 430071, Hubei, China)

(2. Earthquake Administration of Hubei Province, Wuhan, 430071, Hubei, China)

Abstract

The new method of heterogeneous seismic database integration services which takes the transparent gateway technology for heterogeneous database connectivity in data layer take the cross-platform, cross-language ICE middleware technology in application layer, could realize and provide the secondary development of interfaces related to various business, and the users can easily and quickly operated the interface business data through this interface, which will not only shield the difference between the heterogeneous database service, but also greatly enhance the efficiency of the integration data services.

Key words: heterogeneous database; seismic data integration; transparent gateway; ICE middleware