

# SQL Server 2000 分布式数据库系统的构建

葛芝宾

(盐城师范学院 信息科学与技术学院, 江苏 盐城 224002)

[摘要] 通过对 SQL Server 2000 分布式功能和机制的分析研究, 给出实现 SQL Server 2000 分布式数据库系统的逻辑架构. 介绍了 SQL Server 2000 服务器注册和构建链接服务器的步骤, 讨论了 SQL Server 2000 数据库分布存储的数据分区概念、分区设计和实现, 以及分布式分区视图的创建及其在分布式查询中的应用.

[关键词] SQL Server 分布式数据库, 系统构建

[中图分类号] TP 311.138 [文献标识码] A [文章编号] 1672-1292(2008)04-0199-04

## Construction of Distributed Database System Based on SQL Server 2000

Ge Zhibin

(Department of Information Science and Technology, Yancheng Teachers College Yancheng 224002, China)

**Abstract** Having analysed and researched on the distributed database management and the mechanism of SQL Server 2000, this paper shows the logic construction of distributed database architecture based on SQL Server 2000 and introduces in details the steps of registration and configuring linked server. It discusses the conceptions, plans and implementation of data partition in SQL Server 2000 database distributed storage. It also investigates the creation and application in distributed query of distributed partitioned view. This paper is worthy for the configuration, application and development of SQL Server 2000 distributed database system.

**Key words** SQL Server distributed database system construction

MS SQL Server 是中小型数据库管理系统的代表, 是国内相当流行的数据库管理系统和开发平台. 随着企业经营的发展 (如连锁店、分厂等) 和管理模式的改变 (如职能部门的垂直管理), 使得业务数据的采集和处理分布在不同的地理位置, 而作为全局性应用又需要将其在逻辑上作为整体对待和处理 (查询和统计等). 若能在原有的 SQL Server 系统上升级数据库应用系统, 则能节约成本, 是较好的选择方案. 本文就 SQL Server 分布式数据库系统的机制及其构建进行了探讨.

### 1 SQL Server 2000 的分布式机制

分布式数据库是利用计算机网络将地理位置分散而管理和控制又需要不同程度集中的多个逻辑单位 (通常是集中式数据库系统) 联接起来, 共同组成一个统一的数据库系统. 该系统负责实现分布式数据库的建立、查询、更新、复制和维护等功能<sup>[1]</sup>. SQL Server 2000 提供了实现以上分布式数据库系统主要功能的机制. 本文以下的讨论均以 SQL Server 2000 企业版为基础.

在许多关于 SQL Server 2000 的书籍中, 对 SQL Server 2000 的分布式数据库系统有关的概念和功能的介绍较分散, 缺少逻辑上的整体联系. 本文试图理清其逻辑关系, 从构建 SQL Server 2000 分布式数据库系统的总体需求出发, 研究其功能架构, 并介绍构建 SQL Server 2000 分布式数据库系统的步骤. 首先, 每个

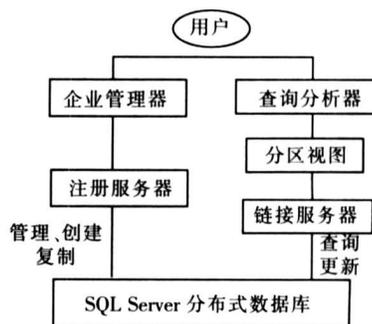


图 1 SQL Server 分布式系统逻辑架构

Fig.1 Architecture of distributed database system based on SQL server

SQL Server 2000 是作为独立数据库服务器存在,其分布式管理控制(包括对数据库的建立、管理和复制等)是通过注册服务器来实现的,而分布式查询是通过作用在链接服务器上的分布式分区查询来实现的,分布式更新则是通过建立可更新分布式分区视图来实现的.其逻辑架构如图 1 所示.

## 2 注册服务器<sup>[2]</sup>

SQL Server 2000 是分布式网络数据库系统,在数据库系统的管理工具和后台数据库服务器之间采取的是 Client/Server 机制,来实现管理工具和后台数据库的分离.数据库服务器的注册,就是将分布在 Intranet 或 Internet 上的数据库服务器,纳入到集成的管理工具的管理范围的途径.数据库服务器的注册流程如下:

### (1) 配置被管理的数据库服务器的服务器网络实用工具.

在被管理的数据库服务器的 SQL Server 2000 程序组里选择“服务器网络实用工具”,在“常规”标签页的“启用的协议”列表框里选择“TCP/IP”,单击“属性”配置数据库服务器的 TCP/IP 协议.在“网络协议默认值设置”会话窗中显示“默认端口”的值是 1433,除非必要且具备相应知识,一般可不改变该端口的默认值.对其它各项一般也无需修改.

### (2) 配置进行管理的数据库服务器的客户机网络实用工具.

在进行管理的数据库服务器上的 SQL Server 2000 程序组里选择“客户机网络实用工具”,在“常规”标签页的“按顺序启用协议”列表框里选择 TCP/IP,点击“属性”按钮,出现“网络协议默认值安装”会话窗,检查所显示的端口默认值是否与(1)中所设置的一致,若在(1)中对默认值进行了修改,则此处必须使之一致.

在“别名”标签页中,若“服务器别名配置”列表中还不存在所需注册的服务器,则点击“添加”按钮,在出现的“添加网络库”窗口的左栏中点选“TCP/IP”,在“服务器别名”输入框中输入对被注册服务器的命名;在“连接参数”栏的“服务器名称”输入框中,输入被注册服务器的 IP 地址(一般地,独立服务器均有其固定 IP 地址),“确定”回到“SQL Server 网络实用工具”窗口.其它内容无需设置,“确定”完成网络实用工具的配置.

### (3) 在配置进行管理的数据库服务器上完成注册.

在“企业管理器”中的服务器注册过程如下:

① 右击一个服务器或服务器组,然后单击“新建 SQL Server 注册”命令.如果上次使用注册服务器向导时,选择了“我希望今后在执行该任务时不使用向导”复选框,则 SQL Server 企业管理器将显示“已注册的 SQL Server 属性”对话框.否则,将启动“注册 SQL Server 向导”.

② “注册 SQL Server 向导”中,从“可用的服务器”列表中选择所要注册的服务器(若客户机网络实用工具配置完成则列表中必有),点击“添加”按钮;点“下一步”进入“选择身份验证模式”.在“使用 Windows 身份验证”和“使用 SQL Server 身份验证”中点选一项.

如果选择“使用 Windows 身份验证”,则需要注册 SQL Server 服务器中建立与本服务器上 Windows 启动帐户相同的登录帐户.如果选择“使用 SQL Server 身份验证”,则进入“选择连接选项”,需要在“用我的 SQL Server 帐户信息自动登录”下输入“登录名”和“密码”.

单击“下一步”进入“选择 SQL Server 组”,可以将被注册服务器加入指定的组.点“完成”按钮后进入系统的连接注册过程.若网络良好且登录名和密码均正确,那么就会显示“注册成功”信息并处于连接状态.关闭窗口即可在企业管理器的相应 SQL Server 组中出现新注册成功的服务器.已注册的服务器将一直保持注册状态,当系统重新启动后,欲访问被注册服务器,则需要“连接”它.

只要在被注册服务器上为登录帐户设置足够的权限,就可以对被注册服务器进行管理、维护、数据库操作以及“复制”的配置.但这并不影响被注册服务器的独立性.

## 3 链接服务器<sup>[3]</sup>

通过服务器的注册,可以将分布在不同地理位置的 SQL Server 2000 服务器通过网络连成一个整体.但在 SQL Server 2000 的查询分析器中,SQL 的 select insert update 等语句却不能对注册服务器上的数据库进行相关操作.这些 SQL 语句对非本地数据库的操作是作用于“链接服务器”的.链接服务器是 SQL Server 2000 的虚拟服务器,它通过 OLE DB 或 ODBC 对异构数据库或其它 SQL Server 2000 服务器上的数

数据库进行操作,其工作原理如图 2所示。

链接服务器的创建既可以通过调用系统过程 `sp_addlinkedserver` 也可以通过企业管理器进行创建. 现假定将网络上其它 SQL Server 独立服务器创建为链接服务器, 介绍在企业管理器下的创建步骤。

(1) 如果欲链接的服务器还没有用“客户机网络实用工具”进行设置, 则按前述的方法进行设置。

(2) 在“企业管理器”中展开本地 SQL Server 实例的“安全性”, 右击“链接服务器”选择“新建链接服务器”, 在“常规”标签页的“服务器类型”中选择“其它数据源”, 在“提供程序名称”输入框的下拉列表中选择“Microsoft OLE DB Provider for SQL Server”, 在“链接服务器”输入框中输入链接服务器名字 (作为使用该链接服务器的名字——命名需符合规范), 在“数据源”输入框中输入所要链接的 SQL Server 服务器的网络名 (通常为其 Windows 系统的网络标识), 在“目录”输入框中可以指定所要访问的数据库名 (该项为可选项)。

接下来在“安全性”标签页中建立登录帐户间的映射关系. 在“本地服务器登录到远程服务器登录映射”栏的“本地登录”中填入本地服务器的登录帐户 (一般为启动 SQL Server 的帐户), 并勾选“模拟” (也允许填写本地非启动帐户, 此时需填写远程登录帐户和密码, 不勾选“模拟”); 选中“用此安全上下文进行”, 在“远程登录”输入框中输入登录帐户名, 在“使用密码”输入框中输入密码. 其它标签页的内容无须设置, “确定”即完成该链接服务器的创建. 在企业管理器目录树中的链接服务器下就可看到创建的链接服务器。

若采用层次结构的分布式系统, 那么, 通常需要选择一台 (或几台) SQL Server 服务器作为主控服务器, 而将其它 SQL Server 服务器 (称为成员服务器) 配置为该主控服务器的链接服务器, 作为分布式数据处理 (查询、更新等) 的环境. SQL Server 2000 可以将 `server1` 和 `server2` 互为被注册服务器、互为链接服务器, 因而可以构建网状结构的分布式数据库系统. 需要说明的是, 目前国内有些 SQL Server 2000 企业版不能实现这种网状结构。

关于创建异构数据源的链接服务器和使用系统过程创建链接服务器的内容, 在本文引用的文献中有较为详细的介绍。

## 4 数据分区和分布式分区视图

作为分布式数据库系统, 数据的分布存储是主要组成之一. 分布存储通常将数据进行水平分割和垂直分割, 然后将分割后的各个片段分布存储在各成员服务器上, 并且还能作为逻辑整体进行查询和更新. SQL Server 2000 支持水平分割, 即所谓数据分区. 例如, 在连锁店的分布式数据库系统中, 可以把销售、客户、库存等信息分别存储在各个区域的销售店的 SQL Server 数据库的表中, 这样既适合于各店的业务处理, 也可以在逻辑全局上进行查询和必要的处理。

对数据的分区需要应用设计者进行必要的规划, 首先需要确定进行分区的数据表, 然后对每个数据表划分行的片段 (部分行), 每个片段作为某个成员服务器中存储的数据表的内容 (该表称为成员表); 通常, 为了标识各个成员表的行集范围, 需要确定一个“分区列”, 以该列的取值范围 (区间) 代表上述的行片段. “分区列”需要包含在主键 (Primary Key) 中 (主键约束), 并对该列定义 CHECK 约束, CHECK 约束中只能使用以下运算符: `BETWEEN`、`AND`、`OR`、`<`、`<=`、`>`、`>=`、`=`. 此外, 分区列不允许为空. 每个成员表的数据范围都在为分区列指定的 CHECK 约束中定义, 它既便于 SQL Server 通过对成员表使用 `Union All` 的查询优化, 并生成智能的查询计划以便有效地访问成员表中的数据; 也有利于应用程序根据该列的取值识别数据的归宿. 由这样的成员表构成的数据库称为成员数据库, 然后在各成员服务器中创建这些成员数据库和 (具有相同结构的) 成员表, 就实现了数据的分布存储. 这既能适应成员服务器作为独立服务器对数据进行处理, 也能通过 SQL Server 2000 提供的四部分名称及分布式分区视图, 对同结构的成员表在逻辑上

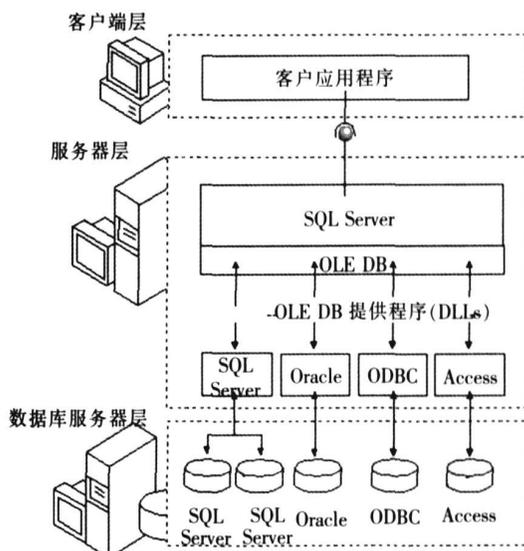


图 2 链接服务器的工作原理

Fig.2 Principles of linked server

作为整体进行查询和更新。

分区视图是通过使用 Union All 合并并在多个成员表上的查询结果集所定义的视图。分区视图需要满足以下条件:

(1) 成员表中的所有列应在视图定义的列列表中被选定。

(2) 每个 select list 中的同一序号位置上的列应属于同一类型, 包括排序规则。列至少属于隐性可转换类型 (在 Union 的通常情况下)。

(3) 成员表可以是本地表, 也可以是通过由四部分构成的名称或基于 OPENDATA SOURCE 或 OPENROWSET 的名称引用的其它 SQL Server 中的表; 在用 Union ALL 语句组合的表集中, 同一个表不能出现两次; 成员表不能在表中的计算列上创建索引; 成员表在数目相同的列上应具有所有 PRIMARY KEY 约束。这样创建的分区视图可以作为查询的只读副本。

引用多台服务器上的成员表所创建的分区视图称为分布式分区视图。所谓四部分名称, 其语法是:

< linked\_server\_name> . < catalog> . < schema> . < object\_name>

当链接服务器指代 SQL Server 实例时, catalog 指代数据库, schema 指代所有者 ID。使用由四个部分构成的名称在 T-SQL 的 select insert update delete 语句中引用已配置链接服务器中的对象。例如, 欲在 (本地) 服务器 server1 的 sales 数据库中创建包含 server1, server2, server3 上 sales 数据库的 kucun (库存) 表的分布式分区视图, 可用如下的 SQL 语句:

```
create view kucunv as
select* from sales dbo kucun
--本地服务器名可省略
union all
select* from server2 sales dbo kucun
union all
select* from server3 sales dbo kucun
go
```

执行上述语句即可在本地服务器的 sales 数据库中创建指定名的视图。分布式分区视图可为分布式查询提供行集。

在 T-SQL 语句 select insert update 和 delete 中, 通过使用四部分名称可以实现分布式查询和更新。此外, SQL Server 2000 还提供了 OpenQuery, OpenRowSet 和 OpenDataSource 函数以支持分布式查询。

## 5 结束语

本文仅就将分布在网络上不同地理位置的 SQL Server 服务器构成分布式数据库系统进行了探讨, 对诸如复制 (出版和订阅)、数据转换 (DTS) 等内容没有涉及, 这些内容在文献 [2-4] 中有较为详细的介绍。笔者在开发连锁商店管理系统等实践中, 已实现本文所述的内容 (SQL Server 2000 和 Windows 2000 Server 环境), 为了在 Internet 下获得稳定的网络环境, 我们采用 VPN 连接。尽管 SQL Server 2000 分布式数据库系统的部分功能还不尽完善, 但已能满足中小型数据库分布式应用系统的需求。

### [参考文献] (References)

- [1] 邵佩英. 分布式数据库系统及其应用 [M]. 北京: 科学出版社, 2000.  
Shao Peiyng. Distributed Database System and its Application [M]. Beijing: Science Press, 2001 (in Chinese)
- [2] 赵松涛. 中文版 SQL Server 2000 应用及实例集锦 [M]. 北京: 人民邮电出版社, 2002.  
Zhao Songtao. Applications and Samples of SQL Server 2000 in Chinese [M]. Beijing: Posts & Telecom Press, 2002 (in Chinese)
- [3] Shapiro Jeffrey R. Microsoft SQL Server 2000 参考大全 [M]. 周之, 黄玫, 译. 北京: 清华大学出版社, 2002.  
Shapiro Jeffrey R. SQL Server 2000 The Complete Reference [M]. Zhou Zhi Huang Mei Translated. Beijing: Tsinghua University Press, 2002 (in Chinese)
- [4] 闪四清. Microsoft SQL Server 2000 实用教程 [M]. 北京: 人民邮电出版社, 2000.  
Shan Siqing. A Practical Course of Microsoft SQL Server 2000 [M]. Beijing: Posts & Telecom Press, 2002 (in Chinese)

[责任编辑: 严海琳]