

文章编号: 1005-8893 (2004) 04-0056-04

基于数据库的复制技术与应用

江士方¹, 徐守坤²

(1. 江苏工业学院 图书馆, 江苏 常州 213016; 2. 江苏工业学院 计算机科学与工程系)

摘要: 基于数据库的复制技术结构简单、机制实用可靠、性能高。介绍了数据库复制技术的原理及特点以及数据库复制的类型和应用情况分析。结合 SQL Server 2000 实际工程实例, 对于复制技术在管理信息系统、热电厂监测、资讯数据服务等行业中的应用进行了理论和实践上的探讨, 并给出了实施具体应用时的规划与实施方法。由于应用背景具有很强的代表性, 因此对中小型企业集团实施和应用信息集成技术, 具有一定的指导意义。

关键词: 分布式数据库; 复制技术; 数据一致性; SQL Server

中图分类号: TP 311.13

文献标识码: A

网络技术和通讯技术的发展, 敦促人们把地理上分散的数据进行统一组织、协同处理、达到随时随地可查询数据。在分布式数据库中采用两阶段提交协议, 虽然很好地保证了分布式事务的一致性, 保证了分布式数据库的一致性, 但这种方式存在两个严重的缺陷: 一是与单节点事务相比响应太慢, 二是参与分布事务的某个节点有故障或网络不通时, 则会使其它节点也无法继续处理数据, 从而降低了系统的可靠性。数据库复制技术就是为克服上述缺点而提出的满足分布式应用的一项举足轻重的技术之一。目前像 Oracle、Sybase、SQL Server 等数据库产品都具有数据库复制功能。本文就在 SQL Server 2000 分布式环境下的数据库复制问题进行讨论。

1 数据库复制技术的原理与特点

1.1 数据库复制技术的原理

数据复制主要通过复制服务器 (Replication-Server) 来完成。在 SQL Server 2000 中, 复制系统由复制服务器、复制代理、复制服务器系统数据库和复制队列组成。复制服务器有其自己的一套系统表、系统过程、账号和命令语言, 系统表存放账

号、复制定义、路由、订阅等信息, 复制服务器根据系统表中设定的信息进行数据复制。系统表存放在系统数据库 RSSD (ReplicationServer System Database) 中, 而 RSSD 库则由数据库服务器进行管理。复制代理负责在数据库服务器和复制服务器之间传送数据更新事务。当一个数据库与复制服务器连接后, 其相应的复制代理 (复制代理类型取决于复制系统中的数据库服务器, 在 SQL Server 中为 LTM 即 LogTransferManager, 日志传输管理进程) 就监视着该数据库的事务日志, 并将符合复制定义和订阅条件的数据更新事务提交给复制服务器, 由复制服务器根据路由和订阅信息将事务复制到目标复制服务器中。目标复制服务器的复制代理将事务传送给目标数据库服务器执行, 从而完成相应的数据更新。因此数据库复制技术的原理就是由复制服务器中的复制代理实时监视数据库管理系统的事务日志, 捕捉针对数据库的操作及数据的变化, 并以事务为单位由复制服务器根据用户预先的定义向远程数据库管理系统复制数据。SQL Server 复制服务器的一大重要特性是在网络或某一节点发生故障时, 会将待复制的事务可靠地存储在队列中, 并在故障恢复后自动将队列复制到远地, 恢复数据一致性而无需人工干预。

收稿日期: 2004-05-06

作者简介: 江士方 (1962-), 男, 江苏江阴人, 工程师, 主要研究方向: 数据库的开发、计算机网络的应用。

1.2 数据库复制技术的特点

数据库复制技术结构简单、机制实用可靠、性能高。其主要特点如下：①高可用性；②一致的信息传输；③高性能；④高度场地自治。

2 数据库复制的类型及应用情况分析

SQL Server 2000 提供 3 种复制类型：快照复制、事务复制和合并复制^[1,2]。

2.1 快照复制

快照复制是指完全按照数据和数据库对象出现时的状态对其进行复制和分发的过程。快照复制不要求对更改进行连续的监视，因为对发布数据所做的更改不会增量地传播到订阅的服务器。订阅服务器用数据集的完全刷新而不是单独的事务来进行更新。因为快照复制是一次复制整个数据集，所以将数据修改传播到订阅服务器的时间较长。复制快照发布的频率通常低于其它发布类型。快照复制一般应用于下列情况：①数据主要是静态数据，不经常更改；②一个时期内允许有已过时的数据副本；③复制小批量数据；④站点经常不在线，并且可接受高滞后时间。

2.2 事务复制

事务复制是将数据的初始快照传播到订阅服务器，此后，当发布服务器上发生数据修改时，捕获个别的事务并传播到订阅服务器。

SQL Server 2000 监视 Insert、Update 和 Delete 语句，以及对存储过程执行和索引视图的更改（SQL Server 7 不提供）。SQL Server 2000 存储影响已复制的对象的事务，然后连续或按一定的调度间隔将这些更改传播到订阅服务器、保留事务边界。例如，如果在事务中更新 100 行，则带有所有 100 个数据修改的整个事务将被接受和传播到订阅服务器。当所有改动传播完毕后，全部订阅服务器的值将与发布服务器的值相同。事务复制提供的选项允许筛选已发布数据，允许订阅服务器中的用户修改复制的数据并将这些更改传播到发布服务器或其他订阅服务器，并允许在发布时间转换数据。事务复制通常应用于下列情况：①希望将数据修改传播到订阅服务器，通常在更新发生的几秒内进行；②订阅服务器通常连接到发布服务器；③应用服务器不能承受订阅服务器接收更改时的高延迟，因此

不适宜远程的数据订阅。

2.3 合并复制

合并复制使订阅服务器自主工作，并且过一段时间后将多个订阅服务器的数据修改合并为一个统一的结果。合并复制要求首先在订阅服务器上应用初始快照，然后 SQL Server 2000 在发布服务器和订阅服务器上跟踪已发布数据的决选项，可在配置合并复制时加以定义。当冲突发生时，合并代理程序唤醒调用冲突解决程序，决定接受和向其他站点传播哪些数据。合并复制通常应用于下列情况：①多个订阅服务器需要在不同时刻更改数据并将这些更改传播到发布服务器和其他订阅服务器；②订阅服务器需要接收数据、脱机更改数据、然后将更改同步到服务器和其他订阅服务器。

3 数据库复制技术的实施与应用

3.1 数据库复制技术的实施^[3,4]

数据库复制技术的实施一般需要以下几个步骤：①服务器配置；②生成和应用初始快照；③同步和传播数据更改。

对于快照复制，同步处理数据意味着在订阅服务器上重新应用快照，以便订阅服务器上的架构和数据与发布数据库保持一致。对于事务复制，对数据进行同步处理意味着数据插入、更新、删除等修改在发布服务器和订阅服务器之间分发。对于合并复制，同步处理数据意味着合并多个站点进行的数据修改、检测和解决冲突（如果有的话），并将数据最终汇聚为所有站点上相同的数据值。

3.2 数据库复制技术的应用领域

下面我们通过几个实际应用项目实例，对于复制技术在一些行业中的应用进行理论和实践上的探讨。

3.2.1 复制技术在公司人事管理系统中的应用

将一个大公司的各应用系统都采用 SQL Server 2000 作为后台数据库，每个管理系统后台均运行一个独立的数据库。由于人员基本信息在多个管理系统都有，如果要更改，必须进入多个系统，十分不便，因此实现人员基本信息在各个管理系统中的共享是管理系统的共同需求。

公司人事管理系统由人力资源部进行日常管理，人力资源部定期（每月 1 次）在公司局域网网

站上发布在职员工的基本信息,供大家查阅、复制使用,对新增人员则必须手持人力资源部的表格到各个管理系统的管理员处进行登记,才可被授权访问使用各系统提供的功能。离职和退休员工也必须手持人力资源部的表格到各个管理系统的管理员处进行系统销号。而对于职工部门变更、职务变迁,则需要各个系统管理员与局域网公布的人员信息进行核对,再做授权变更。

经过分析,确定采用数据复制技术来解决这个问题。目标是实现人事管理系统人员信息的实时更新和发布,各管理系统订阅到更新的数据后,根据业务逻辑和本系统的数据结构进行数据处理,更新本系统的人员信息,进行授权变更。由于公司的管理系统较多,为了方便共享数据的扩充,建立了中间数据库,专门用于共享数据的存储。如人员的基本信息、项目的基本信息等,便于系统之间交换和传递。发布数据库发布的数据表,中间数据库进行订阅,然后中间数据库再将数据表发布到各个管理系统。实现时选择的是事务复制方式。通过事务复制,先将人事数据库的人员表以快照的方式发布到各个管理系统,然后当人力资源部维护人员获取信息后,系统捕获人员信息的变化,传播到中间数据库,中间数据库依次传播到各个管理系统。这样的连续传播也只是几秒钟,大大提高了信息的及时性。

部署方案简要把它分成以下几个过程:①配置前的准备工作;②配置发布服务器;③创建和管理发布;④编写 Trigger 程序,实现发布数据表和业务数据表之间的数据互动。

总体上说,这种方案实施简单,不用重新开发系统,开发工作量小,应是企业数据集成的首选方案。

3.2.2 复制技术在监测系统中的应用

在一套热网监测系统中,热网用户端用单片机技术对各个被测站点的各个被测量进行实时测量、记录和处理,并通过电台向上位机远程实时发送相应数据。上位机通过电台接收数据,将数据存入数据库,热电厂内各个部门的客户机可以通过局域网浏览查询自己所需要的数据。

在提供的 3 种复制类型中,快照复制的每一次复制都会把所有数据都传送到订阅服务器,而不是仅仅传送被修改过的数据;而使用合并复制,将自动跟踪源数据库和所有的目标数据库,源数据库和所有的目标数据中的数据变化都将引起数据的同步

过程。也就是说出版服务器和订阅服务器都有权启动复制过程。在热电厂监测系统中既希望有较高的复制效率,又要限制订阅服务器对数据库的修改权限,所以经过比较,事务复制方式是最佳选择。事务复制利用事务日志来捕获出版物中数据所发生的变化,对数据的所有操作都按时间发生的先后顺序存储在分发数据库中,然后再将这些变化传送到订阅服务器,以相同的顺序将它们应用到目标数据库中^[5]。

事务复制包括快照进程、日志阅读进程和分发进程。利用复制技术,将数据采集服务器(即出版服务器)上的采样数据,根据不同部门的需要,复制和分发到本地的数据库服务器(即订阅服务器)上,保证数据的同步更新。另外复制数据库为只读数据库,不允许本部门内部工作人员修改采样数据,无需立锁,更无需管理锁,大大提高了数据的安全性和可用性;而本部门的另一些非采样数据不再与其他部门共享,这样既降低风险,又简化了管理的复杂性,符合提高效率的原则。采样数据在各个服务器上存放,查询和统计可以同时进行,从而提高了速度,源服务器的负荷大大减小,不再成为系统的瓶颈,因此监测系统的整体性能得到改观,整个系统可以高速、可靠地运行。

3.2.3 复制技术在资讯数据服务业中的应用

有一家数据供应商(可以是证券,地产方面等),它集中了大量的人力每时每刻去编辑处理原始繁杂的信息,并加工出有用及准确的信息提供给需要该信息的客户,它的客户在地理上分布很广,全国各地乃至全世界各地,所谓时间乃金钱,那么它就需要即时地将其数据传递给分布于各地的客户。在现有网络普遍的通讯条件下,采用 SQL 复制技术不失为一个很好的解决方案。在数据资讯服务上, Publisher 即是数据服务供应商, Subscriber 即是客户。服务商可以根据客户的要求量身定制出客户的数据,或者实现定制出相对标准的满足大部分需要的数据来。在不考虑内部处理路程的情况下,部署方案我们简要把它分成以下几个过程:

(1) 硬件上的部署:一条数据通讯线路和一台服务器,条件许可的情况下,还可以采用备份线路和备份服务器。针对数据的发布量,考虑发布与分发是否在同一服务器上。

(2) 软件上的部署:SQL 数据库服务器。对于处于公网上的服务器,还可以安装防病毒,防火墙等相对来说必要的工具。对于出版服务器来讲,

为了能保证通过 Internet 完成与订阅服务器的相互通讯, 需要采用 TCP/IP 协议。

(3) SQL 复制需要考虑的细节: 对于数据供应来讲, 数据只需发布。因此可以采用快照和事务结合的方式。如果数据量小, 比如只有几兆, 则可以采用快照复制。而对于需要流式发布的数据, 可采用事务复制。复制的速度控制根据客户接收的线路好坏进行调整。

(4) 客户端的要求: 数据复制对客户的环境要求不高, 只要能够访问到发布服务器即可, 当然如果纯粹是从 Internet 来强制订阅的客户来讲, 需要有一个固定的 IP 地址。接收软件只要是常用的数据库管理系统即可。

(5) 控制与管理: SQL Server 2000 提供了图形化的复制配置与管理特性, 并能实现到 ODBC、Oracle、Sybase、SQL Server、Microsoft Access 等常用数据库的发布, 其他不支持的也只需安装相对应的客户端即可。对客户的监控, 全部采用非匿名订阅就可方便地监控到客户接收情况, 并设置过期时间。

实践表明, SQL Server 2000 复制技术全面考虑了各种需求, 相对多数采用 FTP 下载数据的接收数据方式来讲, 其优点有: ①客户避免了在接收数据时感染病毒; ②由于 SQL 本身具有的通讯传

输方式, 减少了自行开发传输工具的成本; ③由于数据库数据本身的特殊格式, 使得数据供应商在 Internet 上传输数据时无须考虑加密。

4 结束语

数据库复制技术为分布式事务处理的迅速发展提供了强大的动力。SQL Server 2000 复制是将数据信息的副本在单位多部门之间分布的流行方法, 通过在许多不同的服务器之间分布负载, 就能够快速、可靠地在应用领域中传播数据。如何更好地规划、建立和管理复制, 增强复制性能和安全特性, 是我们今后进一步研究的任务。

参考文献:

- [1] 廖疆星, 张艳叉, 肖今秀. SQL Server 2000 数据库实用教程 [M]. 北京: 冶金工业出版社, 2003. 54-57.
- [2] 微软公司. SQL Server 7.0 系统管理 [M]. 北京: 北京希望电子出版社, 1999. 200-228.
- [3] 郭思媚. 数据复制技术在管理信息系统中的应用 [J]. 计算机辅助工程, 2000, (3): 75-78.
- [4] 田茵, 张云苑. 数据复制技术及实现方法 [J]. 实验技术与管理, 2003, 20 (4): 63-65.
- [5] 徐洁, 汪乐宇. SQL Server 7.0 复制技术在监测系统中的应用 [J]. 电测与仪表, 2001, 38 (3): 46-48.

Base on Database Replication Technology and its Application

JIANG Shi-fang¹, XU Shou-kun²

(1. Library, Jiangsu Polytechnic University, Changzhou 213016, China; 2. Department of Computer Science and Technology, Jiangsu Polytechnic University)

Abstract: Database replication is an exciting technology. Its structure is simple, the mechanism is practical and reliable while the performance is high. In many real projects, it plays a very important role. This paper introduces the principle and characteristics of database replication and probes into several application examples based on SQL Server 2000 theoretically and practically in certain profession. For example, the management information system, the measurement system of thermoelectricity factory and the consulting data service and so on. The plan and implementation methods of database replication presented with a demonstration. The works have more significance to implement the information integration of small group company due to the representative of the background application.

Key words: distributed database; replication technology; data consistency; SQL Server