

MS SQL Server 6.5 复制技术研究及应用

尹晓勇 于洁 杜开雪 (军械工程学院计算机教研室 050003)

摘要:本文论述了数据库复制技术的基本概念,重点介绍了MS SQL Server 6.5数据复制的体系结构和实现方法,并结合一个典型的例子说明数据复制技术在分布式环境下的应用及其应注意的问题。

关键字:数据库 复制 数据分布

一、复制的概念

1. 复制的定义

复制是一种新的数据库技术,它用于将某个服务器上的某些表中的数据复制和分发到远程或本地的数据库服务器上,保证数据的同步更新,从而引用复制的数据和本地表在本端完成分布式查询操作,减少网络的开销。

2. 复制的模式

复制一般可有两种模式:紧凑一致性(tight consistency)模式和松散一致性(loose consistency)模式。

(1)紧凑一致性模式。紧凑一致性模式保证所有拷贝与原拷贝内容在任何时间都要绝对一致。它通过两阶段提交(two-phase commit)协议协调并同步所有结点上的数据。由于它的同步要求很高,因而只适用于一个高速局域网来同步复制,这样通常会导致一个可用性很差的网络系统。

(2)松散一致性模式。松散一致性模式允许原数据改变时与该数据的复制拷贝更新时,它们两者之间有一个时间延迟,且并不保证所有拷贝与原拷贝内容一致。但它的优点是对于网络的要求并不高,对于局域网、广域网、快速或慢速通信线路都具有较好的数据库可用性。

松散一致性模式的复制方法有三种:转储、表级快照和联机异步复制。

①转储和重载。转储和重载就是从主数据库服务器上,将数据转储到卡片、磁带或磁盘级设备中,然后通过各种传输方式装入其他结点的数据库服务器中进行处理。这种方法的不及时使得它通常用于数据备份和恢复。

②表快照。给主表生成只读快照,分散存放于别的结点,利用快速刷新机制进行传播。但由于它只是基于表级的复制,而没有维护多个快照表之间以及事务的完整性。

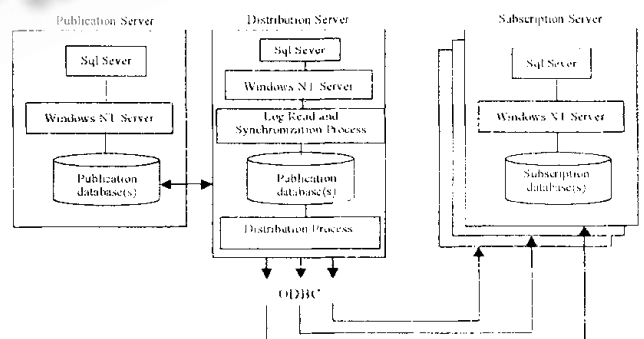
③联机异步方式。联机异步方式保存表的多个拷贝,使多个数据库密切同步。它也可以在多个数据库上间歇地更新数据,并且可在系统和网络发生故障之后自动将等待的信息前滚到远程服务器上,维护事务顺序及其完整性,重新使数据同步。

二、MS SQL Server 6.5 的复制技术

在MS SQL Server 6.5中提供了方便易用的分布式处理功能,主要的技术包括触发程序、远程存储过程、两阶段提交、开放数据服务、批拷贝程序和复制技术等,在此主要讨论其复制技术。

1. 复制数据系统结构

MS SQL Server 6.5的复制技术是基于事务日志的。它复制事务而不是整个表,即将有复制标志的事务从源数据库的事务日志中读出,并送到目的数据库。它的复制数据系统的结构如下图:



在复制的过程中出版、发行、订阅三个服务器担当了重要的角色:

(1) 出版服务器产生用于复制的数据。出版服务器

维护出版数据库,从数据库中产生用于复制的出版数据,并将对出版数据所作的全部修改情况的拷贝发送到分发服务器。

(2)发行服务器包含发行数据库。发行服务器接收出版数据的所有修改情况,并将其存储于发行数据库中。并且,在适当的时候将它们送入订阅服务器。发行服务器可以是充当出版服务器的同一台计算机,也可以是不同的计算机。

(3)订阅服务器维护目的数据库,它接收和维护出版数据的拷贝。

复制的原理如下:在出版服务器中包含一个出版数据库。出版数据库是复制数据的源数据库,它含有用于复制的表。日志阅读(Log reader)过程监视参与复制的各数据库的事务日志。当一个事务是有关于复制数据库的表操作时,出版服务器就为这个事务打上复制标志。日志阅读器发现有复制标志的事务时,它就将其送入发行数据库。发行数据库置于发行服务器中,是一种存储/转发数据库,它寄存着所有准备发送到订阅服务器的事务。发行数据库接收出版服务器的事务,寄存它们直到分发过程将其传送到订阅服务器。同步过程预处理(Synchronization processes)包含出版表模式和数据的初始同步文件,并在发行服务器的SQL Server工作目录中寄存该文件,以及在发行数据库中记录同步作业。分发过程(distribution process)将存储在发行数据库表中的事务和初始同步作业传送到订阅服务器。订阅数据库存放目的数据,订阅服务器负责维护目的数据库,供客户使用。

2. 复制的形式

订阅服务器接收复制数据时采用两种形式,即推订阅和拉订阅。

(1)推订阅。当管理焦点置于出版服务器时执行推订阅。即出版服务器将要复制的数据向一个或多个订阅服务器推出的一种复制方式。它的优点是简化和集中了订阅管理,不必为每个订户设立管理。

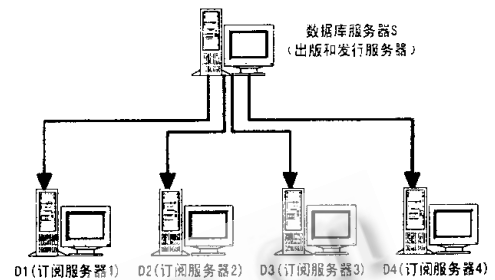
(2)拉订阅。当管理焦点置于订阅服务器时执行拉订阅。即订阅是通过将复制的数据从出版服务器中拉到订阅服务器中。拉订阅的优点是复制比较灵活、自主。

三、应用举例

下面通过一个简单的例子来说明复制技术的应用。

1. 系统设计

(1)系统结构。本应用系统采用如下结构。



在服务器S上有一个数据库EMPL-RECO,它包含三个表empl、marks和reco。我们将这三个表完整地分发到其他的数据库服务器上。在S服务器数据修改的情况下,同步更新其他的数据库服务器数据。为此我们将S做为出版和发行服务器,其他服务器均为订阅服务器。

(2)数据分布实现。为简单起见,我们采用推订阅的方式。由于各数据库服务器订阅过程是相同的,我们就以数据库服务器D1为例说明数据的订阅过程。

①首先在S上安装一个名为distribution的发行数据库。其数据设备为distdev,日志设备为distlog。

②然后在S上设置出版选项,将S设置为出版服务器,把EMPL-RECO设置为出版数据库,重新命名为REPL。

③将D1设置为订阅服务器。把inaire数据库设置为订阅数据库。

④将empl、marks和reco三个表设置成出版表,全部出版到D1上的目的数据库inaire中。

⑤在订阅服务器D1上订阅出版表。

至此,该自动复制过程就被建立了。如果出版物repl的缺省同步设置未被改变,则同步在五分钟内发生。同步将拷贝出版表的模式和数据到目的数据库。

2. 需要注意的问题

在数据分布应用中,复制时需特别注意以下几个问题:

(1)要为每个出版表至少定义一个主关键字。

(2)多数情况下,数据只能在出版服务器上被修改。最好将订阅服务器上的目的表上的所有复制数据作为只读数据维护。

(3)尽可能一次出版一个表。在多个出版物中用同样的表会引起多线索分发事件从而增加了网络流量。

(4)只出版所需的数据。

(5)在初始同步前尽可能多地建立订阅。

四、结束语

MS SQL Server 6.5 的复制服务器用来解决网络上相同数据多份拷贝及同步更新这一分布事务处理中的关键难题,在处理分布更新方面与传统的两阶段提交相比,明显提高了效率和可用性,为分布式联机事务处理提供了新的技术和方法。

除 MS SQL Server 6.5 以外,当前流行的许多关系数据库管理系统都提供了数据复制功能,以此支持分布式计算。如 Oracle 7 提供了全面的同步复制与异步复制的功能,包括由只读的快照(snapshot)和 Oracle 7 Symmetric

Replication 提供的基本的复制功能,以及异步的分布式复制技术;SYBASE Replication Server 的异步复制系统,采用存储之前滚的交付机制保证数据的完整和同步。各种数据库系统的复制技术的出现必将为分布式事务处理的迅速发展提供了强大的动力。

参考文献

- [1] 李昭原.数据库技术新进展.清华大学出版社,1997
- [2] 萨师煊,王珊.数据库系统概论.第二版.北京:高等教育出版社,1991
- [3] Robert D. Schneider.规划与建立高性能 SQL Server 6.5 数据库.机械工业出版社,1997

(来稿时间:1998年9月)