

基于SQL Server的数据同步方案的研究及应用

陈 潇

(江苏联合职业技术学院徐州财经分院, 江苏 徐州 221008)

摘 要: 随着企业对信息系统的依赖性日益增加,数据库作为信息系统的核心承担着重要作用。为了缓解服务器访问压力,提升读写效率,很多数据库管理员选择搭建服务器集群。通过数据主从复制技术可以实现集群内部服务器之间数据的同步,实现负载均衡。本文对SQL Server数据库发布/订阅模式下的三种同步方案进行研究,并通过事务发布方式实现数据的同步。

关键词: SQL Server; 数据同步; 主从复制

中图分类号: TP392 **文献标识码:** A

Research and Application of Data Synchronization Scheme Based on SQL Server

CHEN Xiao

(Xuzhou Finance and Economics Branch, Jiangsu Union Technical Institute, Xuzhou 221008, China)

Abstract: With the increasing reliance on information systems in the enterprises, database, as the core of information systems, is undertaking the extremely vital role. To alleviate the pressure of read and write access and enhance the effect of application responses, many database administrators tend to build server clusters to accomplish data synchronization between the servers and load balancing by using the technology of data master-slave replication. This paper studies three replication schemes of SQL Server database publish/subscribe model, and accomplishes data synchronization through publish/subscribe model.

Keywords: SQL Server; data synchronization; master-slave replication

1 引言(Introduction)

随着企业对信息系统的依赖性日益增加,数据库作为信息系统的核心承担着重要作用。为了保证系统的高可用性和稳定性,很多企业扩展单台SQL Server服务器,构建服务器集群。服务器之间数据同步是保证集群数据一致性的基础,各服务器间通过相应的策略同步数据^[1]。在业务复杂的系统中,主数据库主要负责写数据,从数据库主要负责读数据,这样能有效地减轻单台服务器读写压力。数据同步还应用于以下场合:

(1)数据汇总。如某个公司在不同地区设置销售分公司,利用数据同步可以将各地区销售分表进行汇总,形成公司销售总表。另一种情况是对水平切分(如按顺序切分、Hash散列切分)的表进行数据统一,切分后可以利用数据同步进行汇总。

(2)业务拆分。根据不同的业务进行划分。如企业可以划分为供内网办公用服务器,外网客户访问服务器,备份服务

器等。利用同步方案可以保证各服务器的数据一致性。

(3)故障转移。建立数据同步机制后,如果主服务器出现故障则可以随时切换其他服务器保证业务正常运行。

2 发布/订阅模式(Publish/subscribe model)

SQL Server提供的多种数据同步的模式包括:镜像方式、log shipping、发布/订阅和故障集群转移等。本文主要介绍通过发布/订阅方式实现数据库主从同步。

发布/订阅方式是常见的一种主从同步模式,主服务器进行消息发布,从服务器进行订阅,当主服务器中数据发生变化时,就将消息发送至订阅服务器,订阅服务器接收消息后就开始从发布服务器向订阅服务器进行数据传递,进而在两者之间实现数据同步。SQL Server中的发布模式有:快照发布、事务发布、合并发布。

2.1 快照发布

快照发布是完全按照数据和数据库对象出现时的状态来

复制和分发它们的过程^[2]。首先在某一时刻将发布数据库进行快照后产生镜像文件, 然后在约定的时刻或时间间隔内由分发代理程序将文件复制到订阅数据库, 如图1所示。订阅服务器复制的内容是发布数据库某一时刻的状态。

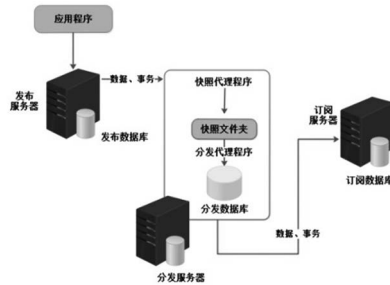


图1 快照发布

Fig.1 Snapshot publishing

快照方式不实时进行数据同步, 而是每隔一个时段或在某一个时间点进行数据库复制, 同步方式为单向同步。优点是数据复制方式稳定, 不必监视对数据的更新。由于是将数据库复制到订阅服务器, 所以在大数据量的同步场景下存在以下缺点: 占用网络带宽较大、数据同步周期较长^[3]。快照方式适用于以下场景:

(1)主库数据量小、更新不频繁。例如, 学生管理系统中的学生表在每学期开学后都要进行一次更新, 因为更新周期相对固定, 所以可以在主库学生表更新数据后从库进行快照复制。

(2)主从库接受同步延迟。快照方式由于复制数据量大导致延迟较高, 可以利用读写较少的时间(如深夜)进行快照复制, 将影响范围缩减至最低。

2.2 事务发布

事务发布是将发布服务器的初始快照将复制到订阅服务器。订阅服务器接收到初始数据后, 代理服务会连续监测数据变化, 当发布服务器上发生数据修改时, 会将这些新增的事务传递给日志代理, 然后日志代理通知订阅服务器, 订阅服务器进行新增事务的执行, 如图2所示。

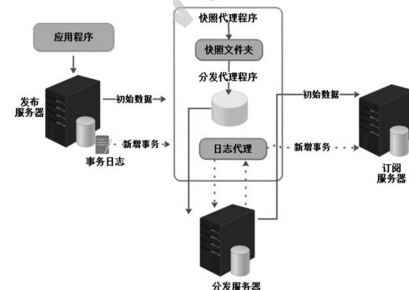


图2 事务发布

Fig.2 Transaction publishing

主从服务器之间通过新增事务复制的方式进行主从同步, 发布数据库上增删改等操作的事务日志会在短时间内复制到订阅数据库。相对快照复制, 事务复制优点是复制内容小、数据传递快, 所以从库更新延迟小, 对大数据量的同步

有可靠保证。缺点是如果主库进行读写操作频繁, 则同步过程中会对服务器性能造成影响。由于发布和订阅服务之间是实时进行事务传递, 所以要保证两者之间相互连接的网络稳定可靠^[4]。这种同步方式适用于以下应用场景:

(1)要求主从库同步延迟小。因为复制的内容是数据库更改的增量, 且事务日志内容较小, 所以可以在较短的时间实现同步。

(2)主库频繁进行DML操作(增加、删除、修改)。如果采用快照方式同步, 则同步时延过高, 不能做出及时响应。如铁路票务系统数据库中的订票信息表频繁进行订票和改签操作(增加、修改), 如同步不及时则会造成主从库查询余票结果不一致的情况。

(3)数据同步方式为单向。同步的方向是由发布同步至订阅服务器, 订阅服务器上的数据库为只读状态。但如果订阅数据库上产生更改, 则不会将事务回传至发布服务器。

2.3 合并发布

快照发布和事务发布数据同步的方式都是单向, 即从库复制主库数据。合并发布数据同步是双向的, 订阅数据库复制到发布数据库的快照后, 可以各自进行读写操作^[5]。在各服务器上产生的数据更改操作会被SQL Server监视, 订阅服务器和发布服务器连接时, 交换自上次发布快照以来各自发生的更改。这种复制方式的优点是允许在主从服务器上分别进行读写操作, 从库独立性较高。缺点是在读写量极大的情况下数据合并更新可能会产生冲突。合并发布方式适用于以下场景:

(1)接受一定的同步延迟。集群订阅服务之间的同步不是实时进行, 某个订阅服务器发生更改后, 会脱机进行数据更改, 并将更改同步到发布服务器和其他订阅服务器, 所以会造成一定的时间延迟。

(2)应用程序访问最终更改结果。例如订单管理系统数据库, 如果同步前订阅数据库中的订单表中数据进行了多次更改, 则与发布数据库同步时只同步最后一次的更改结果, 而不会反映中间的修改过程^[6]。

(3)可在主从服务器上分别进行读写操作。不同于前两种同步方式的单向性, 合并发布方式中发布订阅数据库都可以进行读写操作。例如电子商务系统后台数据库可以分布在不同服务器上, 各服务器都可接受前台应用程序的I/O操作。

3 使用事务发布方式实现数据同步(Data synchronization through transaction publishing)

SQL Server中进行数据同步需要满足以下条件: 被同步的数据表需要有主键, 如果没有主键则会增加一个int类型的自增列id(增量为1); SQL Server代理服务必须启动; 在进行SQL Server注册时, 发布服务器和订阅服务器必须使用计算机名称, 不能使用IP地址或别名注册。

本文将采用事务发布模式实现主从服务器上对应的数据库E_business_DB(电子商务数据库)数据同步。

3.1 主库建立发布

(1)E_business_DB是主数据库, E_business_DB_R是被同步的数据库既从库, 初始化状态下主库和从库中的userInfo表(用户信息表)数据一致。

(2)将主库E_business_DB进行发布, 并设置发布数据库并选择发布类型为“事务发布”, 这样发布服务器将事务流式地传送到订阅服务器。

(3)根据需要进行发布数据库对象, 如图3所示。



图3 选择发布对象

Fig.3 Selecting publishing objects

(4)在快照代理设置内容中选择“立即创建快照”, 也可以在“计划在下时间运行快照代理”中设置运行时间。

(5)指定运行时的数据库账户并进行代理安全性设置后, 数据库即可发布成功。

快照发布的流程和事务发布类似, 区别主要在主从数据库同步的及时性, 快照发布在规定时间内或重启代理服务时会进行数据同步^[7]。

3.2 从库进行订阅

(1)在从库E_business_DB_R上建立订阅, 选择主数据库创建的发布, 设置在分发服务器上运行所有代理。

(2)选择订阅数据库E_business_DB后, 设置运行分发代理的计算机账户和连接的订阅服务器账户, 如图4所示。



图4 设置账户

Fig.4 Setting up an account

(3)同步计划中选择代理计划选项为“连续运行”; 在初始化订阅页面中设置初始化时间, 确认订阅信息后即可创建订阅。订阅成功后会在主库中显示数据库订阅信息。

3.3 主从库数据同步

在主库的userInfo(用户信息)表中增加2条记录, 使记录数为12, 如图5所示。查询从库中的userInfo表, 发现表中数据

行数已经和主库表一致, 如图6所示。说明主库与从库已经完成数据同步。



图5 主库更新数据

Fig.5 Updating Master Database

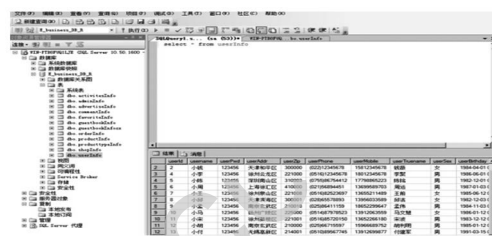


图6 从库完成同步

Fig.6 Synchronizing Slave Database

4 结论(Conclusion)

通过主从复制技术可以实现集群内服务器间的数据同步。数据库管理员在选择同步方案时, 需要综合考虑数据量、同步延迟、网络可靠性等影响因素^[8]。结合具体业务的应用场景选择合适方案, 从而缓解单台服务器读写压力, 实现集群间负载均衡。

参考文献(References)

- [1] 夏明慧. 分布式数据库数据同步的应用分析[J]. 智慧城市, 2016(03):82.
- [2] 季刚. 基于存储过程的主从服务器数据库同步的实现[J]. 计算技术与自动化, 2012, 31(3):114-116.
- [3] 景慎艳. 分布式数据库同步技术及其应用[J]. 计算机与自动化技术, 2008, 34(12):48-50.
- [4] 刘娟娟, 刘帅. 数据库同步技术的研究与实现[J]. 软件工程, 2017, 20(1):1-4.
- [5] 肖刚. 异构数据库更新同步研究与实现[J]. 数字通信世界, 2017(07):47.
- [6] 李伟洪. 服务器负载均衡的实现与分析[J]. 网络安全技术与应用, 2019(07):12-14.
- [7] 李宁. 一种异构关系型数据库间的数据同步方案[J]. 福建电脑, 2018(10):114-117.
- [8] 刘维志. 数据库应用的高可用性及其实现技术研究[J]. 信息系统工程, 2013(4):97;119.

作者简介:

陈 潇(1983-), 男, 硕士, 讲师. 研究领域: 数据库管理与应用, 管理信息系统开发。