

医院信息系统数据库 性能调整

上海铁道大学医学院附属甘泉医院信息管理中心 郭旭升 金宇红

在介绍我院HIS系统的基础上,探讨如何结合医院特点,调整数据库使用性能。对SQL SERVER数据库采用反规范化、增加存储冗余数据等方法。结果较好地平衡和解决了医院信息系统中联机事物处理能力和决策支持能力之间的矛盾。

医院信息系统(HOSPITAL INFORMATION SYSTEM--HIS)是管理和利用医院信息的计算机网络系统,是一个大型的数据库应用系统。数据库管理软件(DBMS)是负责数据组织和管理的支撑软件,数据库负责按一定结构组织、集成数据,数据库的运行性能直接关系到 HIS 系统的性能,关系到应用者和决策层对系统的评价。医院信息系统是同时运行联机事物处理(OLTP)和决策支持(DSS)的计算机网络系统。如何在现有的硬件平台上对数据库性能进行微调,使联机事物处理性能和决策支持性能达到平衡和优化,需要在实践中紧密结合国内医院的运作特点进行研究和调整。

系统介绍

我院的医院信息系统门诊模块于1999年6月6日投入使用,病区模块于1999年12月31日正式使用。在新系统刚刚使用的一个月内,不可避免地出现了一些混乱,在各方配合和努力下,目前门诊模块已经非常稳定,病区模块业已趋于基本稳定。新系统取得了良好的使用效果和经济效益,解决了诸如单科工作量统计、药品控制、病人费用全部实现明细打印等问题,同时能为决策层实时提供以前需要延迟几天到几周方能得出的数据,极大减轻了信息科统计人员的工作量。

1. 系统简介

网络拓扑: 星形层次以太网

网络协议: TCP/IP

网络介质: 双绞五类线

主机服务器: HP LH3 512M内存, 9GB硬盘(4个)

工作站: DELL PC 内存 64M

网络操作系统: WINDOWS NT4.0 中文版

工作站操作系统: WINDOWS 95/98

数据库管理软件: SQL SERVER 6.5

系统结构: CLIENT/SERVER

前端开发工具: VB5.0 中文版

工作站点数: 52个

2. 相关表介绍

(1)tmz_mzbrjflb(门诊病人交费记录表,主要记载病人就诊信息)

主要字段

c_mzh: 门诊号

n_xuhao: 序号(某个病人某次看病唯一对应一个序号)

d_rq: 交费日期

(2)tmz_yjsfb(一级收费表,主要记载病人收费金额)

主要字段

n_xuhao: 序号(对应tmz_mzbrjflb的字段n_xuhao)

y_je: 金额

c_cfdh: 处方单号

(3)tmz_ejsfb(二级收费,主要记载费用明细)

主要字段

c_cfdh 处方单号(对应tmz_yjsfb的字段 c_cfdh)

c_mc 收费项目名称

y_dj 单价

n_sl 数量

想要查询某个病人的费用完整信息,可以通过以上三个表之间的连接查得,连接字段分别是 tmz_mzbrjflb 和 tmz_yjsfb 中的 n_xuhao 以及 tmz_yjsfb 和 tmz_ejsfb 中的 c_cfdh 如:

```
select * from tmz_mzbrjflb t1,tmz_yjsfb
t2,tmz_ejsfb t3
where t1.n_xuhao=t2.n_xuhao and t2.c_cfdh-
t3.c_cfdh
```

记录某个病人的收费信息同时要对以上三个表实行插入或更新。

以上三个表的设计符合关系数据库的规范化要求,通过字段和表的组合可以查询出符合我院目前要求的各种门诊收费信息。

同时,tmz_mzbrjflb、tmz_yjsfb和tmz_ejsfb(此三张表简称临时表)分别对应三个用于存储每张表的长期数据的永久表tmz_@_mzbrjflb、tmz_@_yjsfb和tmz_@_ejsfb,通过schedule中的存储过程,每天定时把三张表5天前的数据倒入到各自的永久表中,不至于临时表过度膨胀,以保持数据库的性能。遇到的实际问题如下:

①有关收费的报表查询速度慢。收费员在每天结束工作或次日上班时生成门诊收费的日报表,查询速度很慢,通常花费2-3个小时,并且在查询报表时导致门诊划价程序速度明显下降。病人看病时心情一般都较焦急和不愉快,在上午门诊病人集中时,划价速度的稍微下降就会导致病人不满或投诉,影响医院声誉。

门诊有关收费的月报表查询是在报表服务器上进行的,每月都要执行一数据库的transfer,得到全部月报表所花费的时间较长。

②病人查询发票时导致挂号、划价等程序速度减慢。门诊病人经常出现丢失发票的现象,为了报销看病费用,就要到数据库重新查询完整的发票信息,如果看病时间在5天以内,发票数据仍然在数据库临时表中,比较容易查询;但如果是5天前看的病,发票数据已经进入到永久表中。此时,虽然查询发票的速度能够令人接受,但却导致划价、挂号、病区等所有系统上其他应用程序速度变得极慢,有时几乎停止。

问题解决

医院信息系统是同时运行联机事务处理(OLTP)和决策支持(DSS)的计算机网络系统,以上第①、②个问题就是OLTP和DSS相互矛盾的体现。定义SQL SERVER的性能有三种方法,分别是查询的响应时间、吞吐量和并行处理能力。调整单项性能指标是相当简单的,获得单项性能指标的方法有:

- (1)反规范化(存储重复数据)——能加快检索速度
- (2)创建或删除索引
- (3)把表分割后放进多个数据库中
- (4)添加硬件或重新配置软件

解决某个单一的问题是相当直截了当的,然而对于多种性能指标,要找到一种合理的方法通常是复杂的平衡过程。比如,在删除用于更改查询性能的索引后,可以加

快吞吐速度,对表的修改要快得多,系统每秒钟能处理更多的事务,但同时意味着查询性能的损失。因此在调整数据库运行性能时,需要考虑一些通用的常见的折衷方案。比如反规范化往往需要更多的存储空间,这是因为要存储重复数据,而这不是免费的,比较其他解决性能的组合造价,硬盘的代价可能要便宜得多。

1. 原因分析

以上①、②问题的出现,原因是门诊收费信息的存储采用tmz_mzbrjflb、tmz_yjsfb和tmz_ejsfb三级表,这样设计虽然满足关系数据库的规范,减少了冗余数据,但报表查询和发票查询(5天前)时往往需要此三张表的连接,需要CPU资源用来执行连接逻辑和维护数据和引用完整性,对I/O和CPU的开销很大。观察报表查询和发票时的CPU资源占用率几乎达到了100%,此时意味着系统所有应用程序都无法正常工作。

对于第一个问题,试图通过在tmz_mzbrjflb、tmz_yjsfb和tmz_ejsfb上创建多个索引来解决,结果相关报表查询速度得到明显改善,但同时划价程序的运行速度却变慢了(表现为从保存收费信息到收费完成这一段时间变长)。这是因为在数据修改、插入或删除时需要同时更新索引,而这会增加额外的开销,这是必须同时支持OLTP和决策支持类型应用程序的数据库中存在的共同问题。

2. 处理方法

为了提高查询性能,又不对吞吐量产生显著影响,参照提高查询性能的常见方法中的存储冗余数据的方法,在数据库中增加了一张表tmz_mxb(收费明细表)。该表主要字段有门诊号、姓名、发票号、作废标志、收费日期以及门诊10种收费类别的各自金额字段,所含信息能够满足病人发票查询和大部分的报表查询。在划价收费程序保存收费信息于三张临时表时,把部分相关收费信息也写入到收费明细表中。

去掉tmz_yjsfb和tmz_ejsfb上的所有索引,仅在tmz_mzbrjflb的门诊号字段上建立索引,目的是提高系统吞吐量。

临时表tmz_mzbrjflb、tmz_yjsfb和tmz_ejsfb中存储门诊病人5天内的收费信息,而我院规定病人挂一次号于三日内有效。为了提高系统性能,使临时表更加精练,决定在临时表中仅存储3天内的数据。

通过以上数据库调整,较好地解决了两个问题,报表查询和发票查询可以在不影响其他程序的前提下以很快的速度完成,达到了调整性能的目的。■