

Ad Hoc 网中的移动数据库应用

范 俊, 李晓宇

(郑州大学信息工程学院, 郑州 450001)

摘 要: 传统的移动数据库模型应用到 Ad Hoc 网中, 会带来通信代价增大等问题。为此, 对传统的移动数据库模型进行改进, 加入本地服务器作为中介, 形成由移动主机、本地服务器和主服务器 3 类结点构成的移动数据库模型, 并提出 2 个算法用来解决模型中本地服务器与主服务器上的数据同步和事务重做问题, 从而使移动主机能够高效正确地访问数据库。实验结果证明, 该移动数据库模型具有较好的稳定性。

关键词: Ad Hoc 网; 移动数据库模型; 本地服务器; 主服务器; 过时数据; 事务重做

Mobile Database Application in Ad Hoc Network

FAN Jun, LI Xiao-yu

(School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

【Abstract】 As a result that the traditional mobile database model can not adapt to the actual situation of Ad Hoc network, this paper improves the traditional mobile database model with adding a local server as the agent to adapt to the situation of Ad Hoc network, forming the mobile database model which comprises mobile computer, local server and master server. Furthermore, two algorithms which are used to solve problems of transactions redoing and data synchronization between local server and master server are proposed, so that mobile computer can access database efficiently and correctly. Experimental results show that mobile database model can gain good stability.

【Key words】 Ad Hoc network; mobile database model; local server; master server; outdated data; transaction redoing

DOI: 10.3969/j.issn.1000-3428.2012.17.014

1 概述

移动数据库是支持移动计算的分布式数据库, 通常应用在诸如掌上电脑、PDA、车载设备、移动电话等嵌入式设备。一个理想的移动数据库系统要实现可用性、移动性、可串行性、收敛性这 4 个目标^[1]。

传统的移动数据库模型主要包括服务器、移动支持结点以及移动主机 3 类结点, 其中服务器和移动支持结点都存在于固定网络中, 称为可信部分, 移动主机则处在移动网络中, 称为移动部分^[2]。传统的移动数据库模型具有很低的事务夭折率、数据操作的正确性、低通信代价和高并行性、自治性等优点。但是, 传统的移动数据库模型服务的网络是拥有大量的基站、移动交换中心等基础通信设备, 并且多是围绕城市这样的环境而构建, 因而不用过多地去考虑环境等因素的影响。而 Ad Hoc 网作为一种多跳的、无中心的、自组织无线网络, 多是因为环境等因素不利于铺设基础设施或者基础设施损坏等原因才使用的。在现实中, Ad Hoc 网的一类典型应用是侦查小分队到野外执行任务、灾后救助、矿场勘测等。在 Ad Hoc 网中, 考虑到移动主机的能源和存储能力有限、带宽受限以及频繁

的与主服务器发生断接等原因, 将传统的移动数据库模型应用到 Ad Hoc 网中, 会带来很多问题, 如: 通信代价增大, 事务冲突数量明显增加, 错误的事务操作增多, 事务延迟时间变长, 事务夭折率升高, 甚至会因为与主服务器长时间断接而导致整个网络系统瘫痪等。为了将移动数据库技术很好地应用到 Ad Hoc 网中, 需对传统的移动数据库模型进行改进, 以适应 Ad Hoc 网络。为此, 本文提出一种改进的移动数据库模型, 并给出检测本地服务器上正在执行的事务是否使用到过时数据的算法和对使用到过时数据的事务决定是否进行事务重做的算法, 解决模型中的本地服务器和主服务器上的数据同步和事务重做问题。

2 Ad Hoc 网中的移动数据库模型

2.1 改进的移动数据库模型

根据 Ad Hoc 网的特性及它应用的实际环境等因素, 构造出适用于 Ad Hoc 网的移动数据库模型^[2-3]。如图 1 所示, 模型包括 3 个主要元素:

(1) 主服务器(MS): 一般是固定网络的固定结点, 带有自己的数据库, 数据库中保存着客户要用到的全部数据, 只和本地服务器联系。

基金项目: 国家自然科学基金资助项目(61073023); 河南省教育厅自然科学基金资助项目(2010B520025)

作者简介: 范 俊(1986—), 男, 硕士研究生、CCF 会员, 主研方向: 移动数据库技术; 李晓宇, 副教授

收稿日期: 2011-10-20 **修回日期:** 2011-12-29 **E-mail:** fanjun1986120@163.com

(2)本地服务器(LS): 从移动主机中选出来的, 属于 Ad Hoc 网中的移动结点, 存储主服务器的部分数据副本, 作为主服务器和移动主机之间的中介, 负责处理移动主机的事务请求。另外, 由于自身本就是移动节点, 因此性能会受到环境、自身配置和电源等因素的影响。

(3)移动主机(MC): 处理能力、存储能力、能源都是受限的, 带有缓存, 可以储存一些常用的数据^[2]。由于处在 Ad Hoc 网中, 不时会发生网络断接, 可靠性比较低, 网络延迟时间比较长。

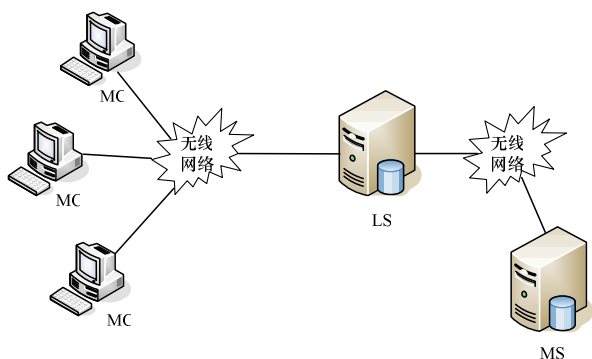


图1 移动数据库模型

2.2 移动数据库模型的工作原理和特点

本地服务器和主服务器数据库中的数据采用加锁机制, 保证在加锁期间只能有一批读事务或一个写事务在使用某一个数据。此外, 移动主机和本地服务器之间的关系仍遵守传统的移动数据库模型, 移动主机带有缓存, 缓存数据库服务器上的部分数据, 减少客户访问数据库服务器的频率, 从而提高系统的性能^[4]。

当移动主机需要访问数据库时:

(1)如果本地服务器的数据库中有事务所申请的数据, 则申请访问数据。

(2)如果本地服务器的数据库中没有移动主机将要访问的数据, 则本地服务器访问主服务器, 下载该数据到本地服务器的数据库中, 然后供移动主机访问。如果在本地服务器上修改了某些数据, 则将修改后的数据上传给主服务器, 更新主服务器上的数据。

模型的核心思想是使用本地服务器作为主服务器和移动主机之间的中介, 在本地的 Ad Hoc 网中充当主服务器的作用, 工作原理如图 2 所示。

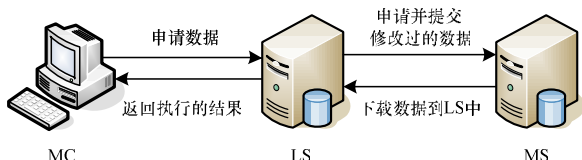


图2 移动数据库模型工作原理

移动数据库模型具有以下特点:

(1)本地服务器上保存主服务器上的部分数据副本, 它负责从主服务器获取数据和向主服务器提交数据更新, 并且所有的移动主机想要申请数据时, 均访问本地服务器上的数据库, 不直接和主服务器联系。

(2)考虑 Ad Hoc 网的特性, 移动主机发起的所有事务都在本地服务器上执行, 所有的移动主机只和本地服务器通信。这样可避免移动客户端无序地和主服务器进行联系, 减少上行链路不必要的带宽浪费, 节约移动主机的能源, 而且能更快地处理事务, 尽早将执行结果反馈给客户。

(3)采用即时更新的策略^[5]: 当主服务器上的数据更新之后, 随即发送数据更新给本地服务器; 本地服务器的数据发生更新后也会立即上传给本地服务器。可以尽量降低移动主机读到过时数据的概率。

(4)由于移动主机只和本地服务器通信, 当与主服务器断接时, 移动主机的工作不受任何影响。

但是, 由于移动主机只访问本地服务器上的数据副本, 而后者并不能够保证始终与主服务器数据一致, 这样移动主机发起的事务可能会读到过时的数据。如何判断读到的数据是否过时, 如果过时, 是否需要重做事务, 如果需要, 如何重做事务等, 将在下文讨论。

3 过时数据检测与事务重做算法

3.1 检测数据过时

本地服务器和主服务器上的数据存储形式都是由数据在主服务器上的更新时间 and 数据内容两部分构成。如何判断本地服务器上的事务使用的数据过时, 假定本地服务器上的数据 d 的主服务器更新时间是 T_{mu} , 本地服务器上的使用数据 d 的事务开始时间设为 T_{ts} , 判断本地服务器上的事务使用数据过时的方法如下^[6]:

(1)若 $T_{ts} < T_{mu}$, 则事务使用的数据 d 不过时。

(2)若 $T_{ts} \geq T_{mu}$, 则事务使用的数据 d 过时。

3.2 事务重做

定义 1 事务重做是指将事务对数据库的所有已完成的操作全部撤销, 重新开启一个新事务执行相同的操作。

当事务使用到的数据过时, 就要决定事务是否需要重做。针对读写事务的执行过程不同, 并且考虑到数据的重要程度, 采用不同的算法来处理事务重做问题^[2,7-8]。

3.2.1 读事务

考虑到当用户读到一些并不重要的数据时, 这些数据如果已经过时, 但对用户没多大影响, 没有必要进行事务重做, 所以将读事务分成高、中、低 3 个等级。

读事务的重做算法:

(1)当主服务器的数据更新到达本地服务器, 若 $T_{ts} < T_{mu}$, 则转向步骤(3)。

(2)若 $T_{ts} \geq T_{mu}$, 则检查事务的等级:

1)如果是低级, 就不需要重做, 转向步骤(3)。

2)如果是中级, 则事务继续执行, 同时本地服务器给移动主机发送一个数据过时提示, 请(移动主机)用户选择是否需要重做: 如果用户决定不重做, 则转向步骤(3); 如果用户决定重做, 则撤销事务, 重新开始并转向步骤(1); 如果没有收到反馈, 则默认为不必重做, 转向步骤(3)。

3)如果是高级, 事务撤销, 重新开始并转向步骤(1)。

(3)读事务结束。

3.2.2 写事务

对于写事务, 由于要将修改的数据更新到数据库中, 为了保证主数据库中存放着正确和最新版本的数据, 写事务一旦用到过时数据就要重做事务。整个执行过程可以分成2个阶段, 第1阶段是本地服务器的数据更新, 第2阶段是主服务器的数据更新。

写事务的重做算法:

(1)在第1阶段, 当 $T_{ls} < T_{mu}$ 时, 撤销事务并发起一个新事务做同样的工作; 否则, 接着执行步骤(2)。

(2)在第2阶段, 在更新主服务器数据之前, 申请对主服务器上相对应的数据加写锁。

1)若申请加锁不成功, 则说明数据正在被使用, 撤销此事务, 并发起一个新事务做同样的工作, 转至步骤(1)。

2)若申请加锁成功, 提交本地数据更新给主服务器, 进行相应的数据更新, 接着执行步骤(3)。

(3)数据在主服务器上更新完毕, 写事务结束。

4 实验结果

定义2 事务平均长度是指事务独立运行时(其他事务都不存在)所需要的时间。

定义3 事务完成时间是指事务从申请数据开始到事务正常结束所用的时间。

定义4 事务重做次数是指事务使用到过时数据后进行事务重做的次数。

实验使用的电脑配置 CPU 为奔腾双核 E5300, 主频是 2.60 GHz, 内存是 2 GB, 主板是惠普 OB10H, 操作系统是 Windows XP, 实验所用的开发环境是 J2ME。

通过实验检验本文移动数据库模型的性能和稳定性。考虑到像侦查小分队, 矿藏勘察大概能并发执行的事务数目一般为几个到数十个, 移动终端上执行单个事务的平均时间在数十到数百毫秒的量级。在给定事务个数分别为 10、20、30、40、50、60、70、90 时, 事务平均长度分别是 250 ms、680 ms、960 ms、1 175 ms, 实验得出事务平均完成时间和事务重做次数的变化情况如图 3~图 6 所示。

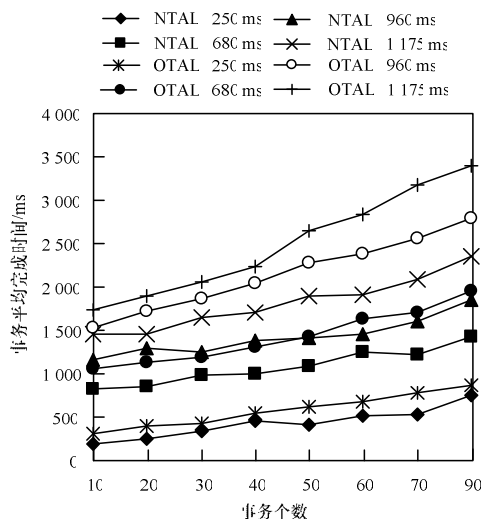


图3 事务平均完成时间与事务个数变化情况1

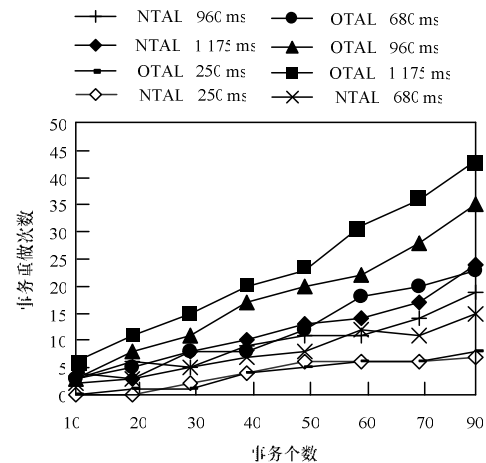


图4 事务重做次数与事务个数变化情况1

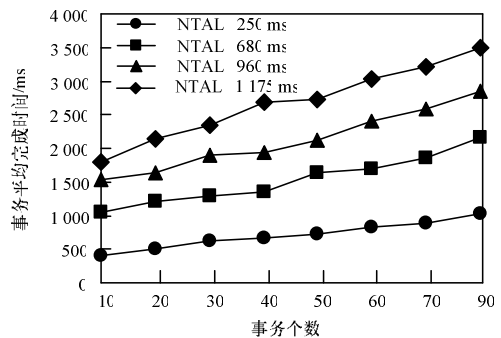


图5 事务平均完成时间与事务个数变化情况2

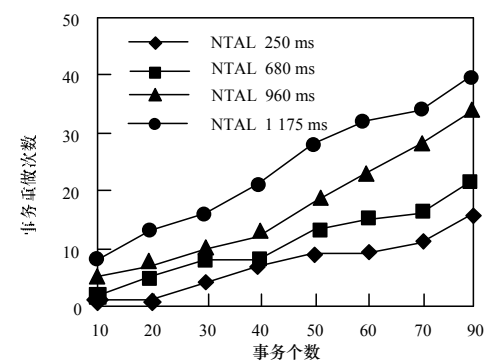


图6 事务重做次数与事务个数变化情况2

其中, NTAL 为改进的移动数据库模型下的事务平均长度; OTAL 为传统移动数据库模型下的事务平均长度。

如图3、图4所示, 在相同的运行环境下, 对传统的移动数据库模型和改进的移动数据库模型的事务平均完成时间和事务重做次数进行了比较。实验得出, 改进的移动数据库模型在不同的事务平均长度和事务个数下, 事务平均完成时间和事务重做次数都要比传统移动数据库模型下要少, 并且随着事务平均长度和事务个数的增加, 这种趋势将越来越明显。充分说明改进的移动数据库模型优于传统的移动数据库模型。

如图5、图6所示, 对改进的移动数据库模型中移动主机和主服务器频繁断接的情况进行了针对性的实验。从图中可以得出, 随着事务个数的增多和事务平均长度的增加, 事务的平均完成时间和事务重做次数都相应地随之增

(下转第55页)