

嵌入式数据库的现状和发展趋势^①

史恒亮^{1,2,3} 白光一^{1,3}

(1. 南京理工大学 计算机学院 江苏 南京 210094; 2. 河南科技大学 电信学院 河南 洛阳 471003;
3. 方舟信息技术(苏州)有限公司 江苏 苏州 215021)

摘要: 从嵌入式系统结构入手, 阐述了嵌入式数据库的发展现状, 并具体介绍了当前主流嵌入式数据库的共性、基本特点、关键技术、构件、分类, 并提出了发展趋势和进一步研究的技术突破点。

关键词: 嵌入式数据库; 复制; 缓存技术; 现状; 趋势

Current Situation and Development Trend of Embedded Database

SHI Heng-Liang^{1,2,3}, BAI Guang-Yi^{1,3}

(1. School of Computer Science, Nanjing University of Science and Technology, Nanjing 210094, China;

2. School of Electronic Information, Henan University of Science & Technology, Luoyang 471003, China;

3. Noah It Solution Com., LTD, Suzhou 215021, China)

Abstract: From the Embedded system's architecture, this paper illustrates the current situation about Embedded database, and clearly introduces main-stream embedded database's common features, basic characters, key technologies, components and its classification. It points out embedded database's development trend and its breakpoint technically.

Keywords: embedded database; replication; cache; current situation; development trend

随着计算机技术及相关技术的不断发展, 嵌入式产品由于其智能化、小型化、便携等特点, 已经普遍应用于我们的生活, 比如智能手机、车载导航设备、数字移动电视、机顶盒、智能控制设备等。这些产品通常是在计算资源、存储资源等条件受限的情况下处理大量的数据, 而传统的数据库一般都是运行在各种各样的服务器或大容量的存储器上, 显然把传统的数据库直接移植到嵌入式设备上是不可行的。那么如何在苛刻的条件下, 充分利用有限的资源便成为了人们研究的热点。嵌入式数据库也正是在此条件下, 和嵌入式操作系统、嵌入式应用软件相伴而生的。

1 嵌入式系统的体系结构

嵌入式数据库位于嵌入式系统的中间部分, 下面

有嵌入式操作系统层, 当前主流的嵌入式操作系统有 Windows CE, PalmOS, Linux, VXworks, Symbian OS 等平台, 硬件平台主要有 Dragonball, HandSpring, MIPS, Hitachi SH, Power PC 以及 ARM 处理器。嵌入式应用软件位于嵌入式数据库之上, 可以通过 API 函数接口, ODBC, JDBC 等与嵌入式数据库通信, 来实现数据的检索、删除、添加等操作, 更为重要的是要实现数据的完整性、一致性。

当前主流的嵌入式数据库, 具有支持多个平台、面向多种开发语言、具有多个灵活的接口等特点, 主流的嵌入式数据库有: Oracle 公司的 Berkeley DB、Sysbase Adaptive Server Anywhere、Linter 以及 SQLite 等, 国内的科研机构也积极开展这方面的研究并推出了自己的产品, 主要有方舟公司的 Noahbase,

① 基金项目: 国家部委基础研究项目(J1500C002)

收稿时间: 2009-06-02

人大金仓 kingbase 等。

在图 1 的体系结构中,只是描述了单一的以嵌入式设备为核心的体系架构,底层的数据采集程序把传感器集所测得的具有不同意义的数传送给嵌入式系统进行实时存储,以备数据中心进行数据的查询。嵌入式系统作为下位机负责数据的采集和处理,数据中心数据库作为上位机可以实现下位机和上位机的通信。很多情况下,这样的系统架构应用在移动通信场合。

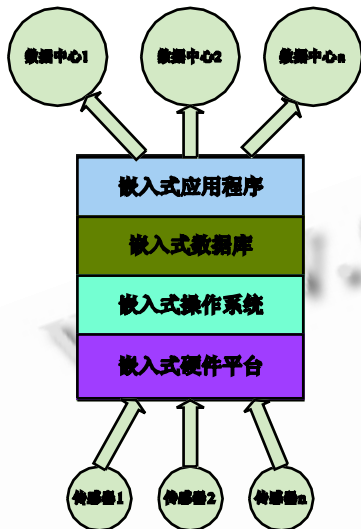


图 1 嵌入式系统体系架构

2 嵌入式数据库的基本特点

嵌入式数据库和企业级数据库的发展都是一个螺旋式上升的发展过程。但它们却有着本质的根本性的区别。外在形式的相似性,并不能代表二者的实现方式和运用方式的相似。

嵌入性: 嵌入式数据库不仅可以嵌入到其它的软件中,也可以嵌入到硬件设备中,因此无论是在网络上不同计算机之间还是在同一台计算机的不同进程之间,数据库操作并不要求进程间通信,而且其对所有数据的操作都使用 API 接口,不需要对某种查询语言进行解析,也不用生成解析计划^[1]。

实时性: 实时性和嵌入性是分不开的,只有具有了嵌入性的数据库才能够第一时间得到系统的资源,对系统的请求在第一时间做出响应。但是,并不是具有嵌入性就一定具有实时性^[2]。要想使嵌入式数据库具有很好的实时性,必须做很多额外的工作。

灵活性: 当前的嵌入式数据库产品大多具有很强的灵活性,支持多种开发平台,面向多种开发工具,

预留有灵活的开发接口。

移动性: 移动性是目前在国内提的比较多的一种说法,这和目前国内移动设备的大规模应用有关。可以这么说,具有嵌入性的数据库一定具有比较好的移动性,但是具有比较好的移动性的数据库,不一定具有嵌入性^[3]。

伸缩性: 伸缩性是嵌入式数据库的必备特点,没有伸缩性就无法在较小的空间上管理大规模存储空间上的数据。大部分嵌入式数据库可以提供线程/任务级别的引擎,使系统的伸缩空间大大扩大。

某种程度上说,嵌入式场合的数据比企业级应用的数据还要复杂,所以要支持各种类型数据(如多媒体数据和空间数据等)以外,还要支持各种数据结构,除了传统的关系型,还要能处理树状结构和网状结构。当然,嵌入式数据库肯定要具备企业级数据库所具有的一些特点,比如,一致性是数据库所必需的特性。通过事务、锁功能、日志记录、以及数据同步等多种技术保证数据库中各个表内的数据的一致性,同时也保证数据库和其它同步或镜像数据库内数据的一致性。

3 嵌入式数据库应用的关键技术

3.1 复制与缓存技术

嵌入式数据库一般用于移动存储设备,而移动设备的一个显著特点是移动终端之间以及与服务器之间的连接是弱连接,即低带宽、长延迟、不稳定和经常性的断开。复制性嵌入式数据库一般采用复制技术(上传、下载或混合)与服务器数据库进行映射,满足人们在不同时间和地点对数据的需求问题。现在普遍采用乐观复制方法(Optimistic replication)允许用户对本地缓存上的数据副本进行操作, Ding Zhi-ming 等提出了 TLRSP(Transaction-level Result-set Propagation)模型^[4]。在网络断开的情况下,嵌入式数据库在本地数据库子集进行数据操作和事务处理;待网络重新连接后再与数据库服务器或其它终端交换修改信息,并通过冲突检测和协调来恢复数据的一致性。由于嵌入式系统的零管理特点,也就是当系统出现故障的时候,可以在没有人员干预的条件下实现自动恢复,文献^[5]提出了基于功能替代的嵌入式系统恢复方法,该方法以替代作为并发控制和调度的基本单位,可以提高事务处理的成功率。

在嵌入式实时数据库中, Choi 等人提出了采用双

CPU 的并行处理结构, 一个 CPU 负责正常的事务处理, 另一个专门负责有关恢复处理, 如记录日志、做检验点以及故障后进行数据库恢复, 该方法是通过改善硬件的处理性能来换取系统整体性能的改善, 但该方法在资源条件苛刻的嵌入式系统中显然是不适用的。有人提出了基于影子的恢复技术, 该技术的优点是消除了日志开销, 恢复速度快, 缺点是在事务生命周期内数据库需维持其更新数据页的两个版本---当前页和影子页^[6-8], 同时需维护大量的页表指针。该方法是通过适当的增加存储数据量来提高系统性能。相比较而言, 后者在嵌入式系统, 以及嵌入式数据库恢复中应用的更多。为了保证嵌入式数据库中数据的正确性, 文献[9]引入时间有效性, 当数据库实时数据真实反映对应实体对象物理环境时, 这个数据在时间域内是新鲜的。

3.2 事务处理技术和移动查询优化

嵌入式数据库系统中的事务处理在前端可以简单化, 但在整个应用系统中要结合计算环境的特征进行事务处理控制。有效平衡用户事务和更新负载是提供满意服务的关键, 文献[10]提出了 Qos(Quality of Server)技术, 还提出了一种 QMF 方法去保证事务满足截止期的比率。文献[11]提出了用数据版本概念去解决事务间访问数据的冲突和加强并发控制。为了保证活动事务的顺利完成, 必须设计和实现新的事务管理策略和算法。

移动查询优化是指在传统分布式数据库查询优化技术的基础上, 利用多种方法消除带宽多样性、断接等因素所造成的影响, 使查询引擎能够根据当前可采用网络条件采取恰当的优化算法。文献[12]将移动 Agent 引入移动数据库的查询中, 对嵌入式移动数据库查询进行优化。文献[13]提出了 data windowing 方法, 该方法能够最大限度地使用缓存中的数据, 从而实现数据库查询的最优化。在移动计算环境中, 许多信息都是快速持续变化的, 如移动用户的位置等。

3.3 安全性

许多嵌入式产品是应用系统中的关键部分, 而其数据库部分涉及到个人的隐私性很高, 因此在防止碰撞、震动、磁场干扰、遗失、盗窃等对个人隐私数据安全的威胁上需要提供充分的安全保证性。主要措施有:

A)对终端进行认证, 防止非法终端的欺骗性侵入;

B)无线通信进行加密, 防止数据信息泄露;

C)对下载的数据副本加密存储, 以防止移动终端物理丢失后的数据泄密。

4 嵌入式数据库的分类

嵌入式数据库的分类方法很多, 可以按照嵌入的对象不同分为: 软件嵌入数据库、设备嵌入数据库、内存数据库。也有人将它们粗略的分为: 嵌入数据库、移动数据库、小型的 C/S 结构数据库等。目前更被大家所接受和认可的更偏向于按照下面的方式进行划分:

小型 C/S 数据库。这种数据库其实是企业级数据库的一个缩小版, 缩小以后可以在一些实时性要求不高的设备内运行。

面向软件嵌入数据库。它将数据库作为组件嵌入到其他的软件系统中。一般用在对数据库的安全性、稳定性和速度要求比较高的系统中。这种结构资源消耗低, 最终用户不用维护数据库, 甚至感受不到数据的存在。

面向设备嵌入数据库。它将关系型数据库嵌入到设备当中去, 作为设备数据处理的核心组件。这种场合要求数据库有很高的实时性和稳定性, 一般运行在实时性非常高的操作系统当中。

内存数据库。数据库直接在内存内运行, 数据处理更加高速, 不过安全性等方面需要额外的手段来保障。

5 发展趋势

近年来, 嵌入式数据库的研究取得不少进展, 以下几个趋势值得关注:

(1) 智能化和主动化 以往的嵌入式数据库往往是以存储为目的的被动型的数据库, 只能被动地接受操作系统和应用程序的调用来执行相应的动作。能够根据数据库中存储的情况和自身特点, 适当地做出优化来满足不同条件下的应用需求的智能化、主动型嵌入式数据库是以后研究的热点。

(2) 多媒体嵌入式数据库 随着高档电子消费品日益受到人们的青睐, 能够对视频、音频、文字、图像进行存储和快速检索的嵌入式数据库会有更大的市场, 并成为人们研究的热点。

(3) 时空数据库 导航设备、水文、地质、地形

地貌相关电子产品的快速发展,迫切需要能够同时处理时间和空间的数据库问世。

当然,除了上述所介绍的嵌入式数据库的发展方向以外,还有很多的方向,这里不逐一论述,比如说最近盛行的云嵌入式数据库等等。

那么关于嵌入式数据库具体技术的改进和升级可以从以下几点寻找突破点:

(1) 数据库内核 由于嵌入式系统可用资源的受限性,嵌入式数据库内核的大小也是一个关键的问题,那种以牺牲数据库的功能来换取较小的内核显然是与发展趋势相背离的。

(2) 数据库的可靠性 嵌入式数据库的应用越来越广泛,也越来越复杂,其功能已经从传统的添加、删除等操作向对声音、视频、三维或多维数据、智能控制等方向发展,因此可靠性的提高也是至关重要的一环。

(3) 数据库的可移植性 尽管目前的嵌入式数据库产品可以支持多种不同的操作系统,但在嵌入式系统开发过程中,开发人员还要充分考虑硬件平台、操作系统平台以及它们的接口定义,如果有能适用于绝大多数的操作系统平台,对硬件平台和操作系统平台透明的嵌入式数据库,势必可以大大地提高开发效率,降低开发成本。

6 结语

本文从嵌入式系统的体系结构入手,阐述了嵌入式数据库在系统中的作用和功能,并讨论了嵌入式数据库的基本特点、关键技术、分类、构件,并结合当前的应用现状,提出了嵌入式数据库的发展趋势以及技术突破点。

参考文献

- 1 <http://www.oracle.com/technology/products/Berkeley-db/index.html>.
- 2 李国徽,杨兵.移动环境下支持实时事务处理的数据预取.计算机学报,2008,(5):36-41.
- 3 <http://www.empress.com/>
- 4 Ding ZM, Meng XF, Wang S. A novel conflict detection and resolution strategy based on TLRSP in replicated mobile database systems. IEEE, 2001.
- 5 刘云生,夏家莉.嵌入式数据库系统的事务调度.软件学报,2002,13(5):23-27.
- 6 Song EM, Kim YK, Ryu C. No-Log recovery mechanism using stable memory for real-time main memory database systems. Proc. of the 6th Int'l Conf on Real-time Computing Systems and Applications. 1999.
- 7 Shu LC, Sun HM, Kuo TW. Shadowing-Based crash recovery schemes for real-time database systems. Proc of the 11th Euromicro Conf On Real-time system 1999. <http://ieeexplore.ieee.org/iel5/6302/16859/00777473.pdf>
- 8 Abraham S, Henry FK. Database System Concepts. 3rd ed., Beijing: China Machine Press/McGraw-Hill, 1999.
- 9 Stankovic JA, Son SH, Hansson J. Misconceptions about real-time database. Computer, 1999,32(6):25-31.
- 10 Kang KD, Son S H, Stankovic JA. A Qossensitive approach for timeliness and freshness guarantees in real-time databases. Proc. of the 14th Euromicro Conference on Real-time Systems. Washington, DC: IEEE Computer Society Press, 2002.
- 11 Bouazizi E, Duvallet C, Sadeg B. Management of QoS and data freshness in RTDBSs using feedback control scheduling and data versions. Proc. of the 8th IEEE International Symposium on Object-Oriented Real-time Distributed Computing. Washington, DC: IEEE Computer Society Press, 2005.
- 12 李荣鑫,饶泓,等.基于移动Agent的移动数据库查询优化.计算机与现代化,2003,(7):45-51.
- 13 Pisharath J, Choudhary A, Kandemir M. Data Windows: A Data-Centric Approach for Query Execution in Memory-Resident Databases. In IEEE Proc. Of the Design, Automation and Test in Europe Conference and Exhibition, 2004.