

九十年代数据库技术综述

——记第十三届全国数据库会议

张海晴 (北京软件中心 100039)

一、引言

1996年初,在黑龙江大学举办了第十三届全国数据库会议。会议共收到论文240多篇,录取了100篇。这些论文反映了目前我国数据库的研究进展情况,包含了我国数据库界的最新研究成果;涉及到当前数据库的各个研究领域,包括关系数据库理论和实现技术、并行数据库系统、分布式数据库系统、多数据库系统集成技术、多媒体数据库系统和图形图象数据库系统、面向对象数据库系统、演绎数据库系统、主动数据库系统、时态数据库系统、主存数据库系统、空间数据库系统、工程数据库系统、数据库应用系统及其开发工具和实现技术等。因此,从此次会议可归纳出目前我国数据库技术的进展状况。

二、九十年代数据库技术的发展趋势

1. 基本特点

数据库是计算机数据管理技术,具备数据结构化、低冗余度、高独立性、易扩充性、数据使用的非过程性以及数据共享等特点,能够高效地管理、维护海量信息。

E. Fcode 关系理论的提出,以及其成为核心技术及概念的发展,导致从结构的角度看,当今数据库技术最基本的特征是关系型数据库的存在。

另外,从数据库的产品来看,其高新技术含量越来越高,换代周期越来越短,产品功能向全方位发展,产品的开放性更好,同时由于竞争剧烈,价格愈来愈低。

2. 一些新兴理论和技术

(1). 数据仓库(data warehouse)。又称信息仓库(information warehouse),是信息工程领域的新概念,也是近年迅速发展起来的一种数据库技术。

数据仓库的主要设计思想是将分析决策所需的大量数据从传统的操作环境中分离出来,把分散的、难以访问的操作数据转换成集中统一、随时可用的信息而建立一种数据库存储环境。

与传统的操作环境相比,数据仓库环境为信息分析应用准备的数据具备以下四个特点:

·面向主题(subject-oriented)。数据仓库环境要求将数据按照主题来组织。主题域在数据仓库中体现为一系列相关联的表,共同的键值将同一主题域内所有数据联系在一起。

·集约化(integrated)。不同来源的多种数据一旦进入数据仓库,就必须按照统一的结构与编码规则重新组合。

·非挥发性(non-volatile)。数据仓库中的数据主要是服务于信息分析和决策,需要长期积累而成。通常被大量加载和存取,而非一般意义上的修改。

·时间变量(time-variant)。数据仓库中的数据值始终包括有时间元素。数据内容为一系列过去变化的快照(snapshot)。数据保存的时间通常是五至十年,具有做历史比较和趋势分析预测的长期数据基础。

(2)演绎数据库系统。演绎数据库是关系数据库的推广和发展,也是数据库技术发展的最新重要方向之一。在这一领域已取得的理论和技术成果正在被新一代数据库系统采用。

具体而论,演绎数据库是基于谓词逻辑的数据库,是一组规则的集合,规则的语法可以对应到一阶逻辑语法。

(3)多媒体数据库。九十年代是多媒体的时代。多媒体技术和传统的数据库技术相结合,导致了多媒体数据库的兴起。它使人们能够更方便地在集成环境下使用各种现代化的信息传导工具和信息管理工具,把本来独立的媒体获取、播放工具和信息管理工具有机结合为一个整体。因此可以说,多媒体数据库是当今数据库技术领域颇具魅力和发展潜力的一项热门技术。本届会议共有六篇文章对此做出了反响,分别就展示模式、时态参数、层次结构等多个领域展开论述。

(4)图形数据库系统。图形数据库是用数据库存储

与管理图形,并对图形进行查询、修改、编辑等操作,是计算机图形学和数据库技术相结合的产物,其应用非常广泛。

会议中“图形数据库系统 Gsybase 的设计与实现”一文,阐述了以现有的关系型数据库管理系统 sybase 为基础,设计和实现图形数据库系统 Gsybase 的思想和方法。

(5)多数据库系统集成技术。该技术是针对网上多种不同数据库访问而发展的技术。随着信息共享的要求,多数据库系统集成技术受到了人们更多的瞩目。其中的热点技术之一就是联邦数据库。所谓联邦数据库是由多个异构、分布的成员数据库系统组成,即把已运行的分布式异构数据库集成起来,使用户感觉是在同一种数据模型上访问各种类型的数据库。会上有两篇文章对联邦数据库均做了有一定理论和应用价值的探讨。

三、数据库热点及国内情况

目前,数据库热点集中在以下几个方面。

1. 关系数据库系统

有人说,数据库技术最有意义的成就是关系型数据库的发展。的确,关系数据库技术发展到今天,其理论与实现技术日趋完善,其应用已渗透到社会生活的各个领域,产生了巨大的社会效益和经济效益。所以,作为第四代数据库技术,目前关系数据库系统在应用上仍占主导地位。

这一专题下的“数据库加密技术研究”一文,从保证数据库数据存储安全的角度,提出了数据库数据加密的思想。介绍了密码系统的基本结构,详细讨论了数据库加密范围、数据库密码体制、加密算法、数据库加密对数据库管理系统功能的影响,并给出了一个 ORACLE 数据库加密系统的实例。

2. 面向对象数据库系统

近年来,由于需要数据库具备结构模拟与行为模拟的特征性,面向对象数据库(OODB)已成为数据库发展的重要方向之一。它将数据与操作方式统一为整体的一个对象。将数据及其过程一起封装。总之,OODB 在数据库域中引入面向对象的数据模型和方法,在建模能力和运行性能上都显示出比关系数据库更大的优越性和潜力,能够很好地支持多种非常规数据库应用。

在这次会议上,面向对象数据库系统成为讨论的重要专题之一。大致集中在数据模型、面向对象数据库系统原型、性能提高、程序设计和实现技术、索引技术、对

FoxPro for Windows 数据库系统面向对象机制的分析等等。

北工大蒋代梅等提出了一种以分子对象模式为基础的 OODB 三级模式结构。其意义在于:基于此模式的 OODB 管理系统可以实现将数据库技术与电子表格结合起来,以实现面向程序员或最终用户的超级报表系统。

会上还介绍了复旦大学开发的面向对象数据库系统 FOODB 的客户/服务器体系结构及其实现方法。FOODB 的客户端的一个处理用户操作的集成环境,服务器接受和处理面向对象级的命令,客户与服务器的通信方式是通过 Socket 和 RPC 实现的。即服务器接受客户端的命令之后,经过权限检查和事务管理,再把面向对象的命令转换成 Sybase 的 SQL 命令,交给 SQL Server 执行,最后把结果通过网络通信返回给客户端。此系统运行在 SUN 工作站上,使用的操作系统是 Unix 在 SUN 上的版本 Solaris。使用的开发软件有 C 和 C++ 语言,以及为图形用户界面而使用的 OpenWindows 环境下的 Open Look 及 X 协议。系统提供了一个包括应用编程、图形界面、交互命令界面和推理查询的集成环境。

此次会上介绍的另外一个面向对象的数据库系统原型,是上海交大的基于图形用户界面(GUI)的面向对象可视化数据系统(VOODBS)。其意义在于:将面向对象技术和以 GUI 为基础的可视化技术综合到关系数据库系统中,并充分发挥这三者的优势,以实现一功能强大的新型数据库系统。

总之,尽管面向对象概念在各个领域中尚无统一的标准,但我们完全有理由相信面向对象技术在未来将得到更广泛的发展和應用。

3. 分布式数据库系统

由于微机和计算机局域网、高速信息网的发展,分布式数据库系统在信息技术领域中的重要性日渐显著。本次会上多篇论文涉及这一领域,其中多是集中于分布式数据库系统原型的设计与实现方面。

“远程数据库服务器 WRDBS 的设计与实现”一文,介绍了一个基于 Client/Server 结构的远程数据库服务器 WRDBS,它为 PC 机访问远程数据库系统提供一个统一开放的 API 平台,透明访问异构的远程数据库。WRDBS 系统由客户机与服务器两个部分组成,服务器完成数据库操作的事务管理,客户机提供数据库的开放

接口及与服务器的透明联接。文章并介绍了 WRDBS 在 Windows 环境下的实现。

“主动型分布式数据库管理系统中规则的管理与触发”一文,则介绍了一个基于客户/服务器模型的分布式关系数据库管理系统 ADDB 的管理方法及其触发控制机制。此系统被引入了主动功能,实现了分布式环境下的规则定义、规则编译、规则管理、事件监视器及触发控制等功能。因而可为用户提供更方便灵活、更安全可靠地使用该系统的环境。

至于国内数据库系统开发应用的发展情况,可由以下几个方面看出:

(1)基于数据库管理系统的开发工具的研制。国际上目前最为流行的三大数据库管理系统:ORACLE、SYBASE 和 INFORMIX 都拥有各自的外围开发工具。目前应用开发工具正逐渐地独立于依赖数据存储及管理的后端数据库。

那么国内这方面的情况如何?本次会议上有多篇论文分别就联机应用开发工具、关系数据库辅助设计工具、FOXPRO 程序设计支援工具、分布式数据库辅助设计工具等多种数据库应用开发工具论述了各自的设计与实现方法。

(2)基于 UNIX 的关系数据库管理系统的研制。这一领域应以关系数据库管理系统——COBASE 比较有代表性。该系统主要特征如下:

- 具有 SQL 语言,提供 SQL 嵌入 C 语言的编程使用方式和交互式 SQL 使用方式。

- 提供完整的应用开发工具集,包括联机应用开发工具、报表工具、菜单工具、图示工具和 ISQL 工具等。其工具集具有支持快速原型方法开发、触发器语言、支持复杂报表的定义和自动生成等特色。

- 是多用户数据库管理系统。

- 支持 Client/Server 结构,可在多种平台上安装和运行。

- 具备较强的汉字处理能力,支持汉字数据和汉字标识符。

作为通用的多平台上的关系数据库管理系统,COBASE 将数据库管理技术、语言处理优化技术和工具实现技术结合起来,因而目前我国这一领域内具备一定的代表性。

(3)在决策支持系统方面的应用。决策支持系统

(DSS),是综合利用计算机科学、决策分析、人工智能、信息论、控制论、系统工程等学科,将数据库技术与建模仿真方法结合起来,面向决策需要而设计的人机交互系统,支持半结构化和非结构化决策问题。

DSS 在管理决策领域发挥着重要作用。在其模型的建立和管理上,如引入面向对象思想,并将其与关系数据库技术相结合,将会提高模型描述的精确性、模型的可扩充性以及用户界面的灵活性。估计未来几年,数据库技术将被更加完美地结合到这一领域内。

四、存在问题及对策

1. 存在问题

各种新型数据库在理论和概念上于各个应用领域仍未达成统一、形成标准。

真正意义上的多媒体操作系统和多媒体数据库管理系统;面对日新月异的大量涌现的多媒体硬件,多媒体系统的体系结构也并不十分清晰。回顾八十年代初的一些论文,我们今天的工作在概念、方法和技术上并未有本质的进展和突破。因此在多媒体领域尤其在软件领域应进行更系统和更深入的研究,以期和当今的多媒体时代相适应。

至今,面向对象概念在各个应用领域还没有一个统一的标准。应进一步考察、剖析已有的面向对象程序语言、知识表示语言、语义数据模型等,形成一个标准的面向对象数据模型,以使面向对象的新一代数据库系统得以完全实现。

在产品支持方面,市场上虽然有一些产品出现,而且在个别领域得到应用,但仍未形成统一的标准,加之应用支撑软件较少,因此这类产品的市场占有率不高。

确实,比如从多媒体产品来看,当前市场上的许多多媒体产品的技术方向多是主要针对实时处理和信息压缩两项,而并未真正解决多媒体数据的存取和管理,要解决这个问题,必要提出相关的一套完整的理论。

同时从主观认识上讲,虽然人们普遍认为这些新技术新产品前景乐观,但观念的转变、技术和资金的投资转向尚需要一个过程。

2. 对策

应加强有主观计划的基础研究指导,使相关的理论更加成熟,更具指导意义。同时应在成熟理论的指导下,加强商品化系统的开发,形成主流产品,在实践中进一步完善理论。