

探索·争鸣

●万齐鸣

信息高速公路与未来图书馆

ABSTRACT It's necessary for the library to start at the present the building of electronic library information resources, and to complete the normalization and standardization of information itself, the interface of information processing and the structure of network system and the building of a nation-wide electron text data base. 1 ill.

SUBJECT TERMS Information super-highway—Influences
Future libraries—Constructions

CLASS NUMBER G 258. 94

1 信息高速公路的内涵

信息高速公路实际上是一个高速信息网络体系,主要可分为三个层次:信息源、通讯平台、应用信息系统。

信息源包括政治、军事、经济、科技、文化、教育、法律等社会信息资源,它们以有序或无序的形式存在。

通讯平台把有关的应用信息系统联接和综合起来,完成各种信息的处理、传输、交换和分配,最大限度地发挥信息资源共享和信息系统整体功能的效用。它又可分为三个层次:(1)物理传输网络。除现在已非常普及的802·3以太网,还有高速以太网、FDDI、高速ATM LAN、窄带和宽带ISDN、公用数据交换网(X·25)、帧中继网、数字数据网(DDN)、卫星网、高频微波无线网、SDH同步系列通讯网、有线广播电视网等;(2)分布式处理和管理软件平台,负责对各

种物理传输网络和分布式资源的管理与控制;(3)群体协同工作平台。它是在一个网络环境下,为用户提供最佳性能和最大系统利用率的面向应用的支撑工具与界面。

应用信息系统是在高速信息网络上运行的形形色色以最大限度共享信息资源为目的的分布式系统,如分布式军事指挥系统,分布式金融、证券系统,电子图书馆,电视会议系统,远程医序系统,远程教育系统等。

2 信息高速公路的关键技术

这里仅介绍一些与未来图书馆发展信息相关的主要技术。

2.1 超级并行计算机技术

有些应用系统由于信息资源量非常大,系统要求的实时性极强,一般的计算机无法承受信息处理或传输的负荷,这就需要由一

些超级并行计算机来承担中央处理机的任务。如电子图书馆,信息储藏量极大,用户浏览时实时性又极强,为减少硬件设备等资源浪费和减少重复劳动、提高效率,更主要的,电子图书馆的当前用户和潜在用户数量极其巨大,这些都对中央处理机的硬件性能提出了极高的要求。使用现有的一些小型或大型机,一旦同时使用系统的用户太多,中央处理机必将难以承受。有两种解决方案:一是采用巨型机,美国和日本都已研制出运行速度在千亿次以上的巨型计算机,中国也有了运行速度十亿次以上的“银河Ⅱ”,但其成本极高,商品化还有很大的困难;二是利用高性能计算机网络技术,将许多异构或同构多机系统互连起来,构成易于裁减的群机式巨型机。这种巨型机性能价格比高、坚固性好,随着网络技术的发展,处理机间延迟降低,带宽增加,使许多应用可以在并行系统中高效实现。

2. 2 高速传输技术

随着人们信息需求的增加,具有 T3/E3 传输速度(45Mb/s)的广域网 Internet 也变得日益不能满足多媒体信息实时传输的需要。比如一本 40 万字的图书,其电子文本约有 1Mb(纯文字,不包括图象),若包括插图,恐怕字节更多。电子文本传输速度要求极高。高速传输技术已成为制约信息高速公路主干网络建设的关键。当前,研究热点主要集中在两方面:异步交换模式(ATM)和数字同步系列(SDH)。两种技术都具有组网灵活、管理方便、传输速度高(可达 Gb/s 级或更高)等特点,被认为是未来高速信息公路主干网络的主要传输方式。市场目前已有这些技术的初期产品。

2. 3 网络互连技术及网络协议

高速信息网络必须和现有各种中、低速网络互连。目前各种连接中、低速网络的桥、路由器产品品种较多,且在世界范围形成产业,但高速网和中、低速网的互连方面还未

见相应产品。再者,信息网络的各种应用系统大都基于不同体系结构的协议,如何使得具有相同功能的应用能在不同体系结构的协议环境下互操作,也是必须解决的关键问题。TCP/IP 协议和 SPX/IX 协议等目前已获得广泛应用,成为事实上的信息网络通讯标准协议。但这些协议在处理速度、地址定义及路由算法等方面还存在着不足,无法满足高速实时传输的需要。只有解决这些问题,网络用户才有可能在各自的基于不同硬件配置、不同操作系统、不同网络体系结构、甚至不同的协议环境下与网络交换信息。

2. 4 多媒体分布式数据库技术

任何一种数据库都无法将非常巨大的信息量储存到一个单一的服务器中,且能保持良好的运行性能;即使能够储存起来并运行,但对网络来说,要承担大量的信息传输任务必然会导致公路的交通堵塞。这种情况下,客户机/服务器结构的分布式数据库越来越受到用户欢迎。它将大量的信息分布存储在多台服务器中,而且服务器只传送给用户的检索结果,而不是传输整个数据库,这大大减少了网络信息流量。目前客户机/服务器结构分布式数据库比较典型的有 ORACLE7、SYSDBASE、INFORMIX 等产品。但以面向对象数据库技术为主,能同时存储和处理多种媒体的多媒体数据库的研究尚未取得突破性进展。电子图书馆要让用户满意,有一个良好的用户界面,又不丢失信息,必须要应用到多媒体分布式数据库技术。

2. 5 分布式多媒体操作系统

当前流行的网络操作系统有两大类:以 UNIX 为代表的多用户操作系统和以 Netware 为代表的局域网操作系统,它们的支持以 TCP/IP 协议为基础的 Internet 和以 SPX/IPX 协议为基础的 NOVELL 网方面取得了巨大的成功。但是,无论是 Internet 广域网还是 NOVELL 局域网,都是以传输

字符信息为主的中、低速网络。如何能够支持多媒体信息传输,且提高信息传输的速度,需要进一步研究。

2.6 多媒体处理与输入/输出技术

用户在欣赏或阅览电子图书时,不仅希望能够看到文字信息,还希望能够看到图象,听到声音,这就要求电子图书馆系统应用多媒体处理和输入/输出技术。该技术主要包括:多媒体传输标准和协议;多媒体资源的管理技术和互操作技术;多媒体信息的压缩、恢复技术;多媒体信息的输入/输出技术,例如语言识别与语音合成技术;多媒体设备的开发技术等。

2.7 各种接口技术

信息源和信息种类极其繁多,获取和输出信息的设备种类也相当广泛,包括摄像机、扫描设备、键盘、显示器、打印机、电话、传真、光盘、录像磁带、电视、高清晰度大屏幕电视和微生物控制器件等。信息高速公路应为各类设备提供平等的逻辑接口,让它们能够顺利地集成到网络环境中。

3 走入千家万户的电子图书馆

3.1 信息高速公路对图书馆发展的贡献及实现电子图书馆的可行性和必要性

(1) 计算机存储器技术的发展使得将大量文献以电子文本的形式存储起来成为可能。目前海量存储器的研究已取得显著的进展:磁盘的存储容量越来越大,只读光盘存储器和可读写光盘存储器技术也逐渐成熟,光盘存储容量也不断提高,光盘盘片可随时拆卸。将文献用计算机存储起来,有助于解决图书馆空间日益紧张的危机,减少在图书文献维护方面的投入,可以把更多的资金用到更需要的地方。

(2) 通过高速网络传输技术,可将存储在服务器中的数据库信息通过信息高速公路传输到全世界的任何地方,极大地方便了

用户使用图书馆信息资源,最大限度地实现信息资源的共享和开发。

(3) 多媒体处理和输入/输出技术,大大改变了计算机与用户的交互界面,使得用户在使用电子图书馆时更方便、更直观和更有趣味性。用户愿意使用电子图书馆,提高了信息资源的使用效率,带来的经济效益和社会效益无法估量。

(4) 各种接口技术的研究和发展,使用户可以在不需要严格的硬件条件下就可与信息高速公路联网。用户可以通过微机、终端、电话、传真、打印机、家用电视等设备与网络上的电子图书馆系统联接而不需要太多的额外投资。方便了用户,势必促进电子图书馆用户群体的壮大。

(5) 超级并行计算机技术、多媒体分布式数据库技术、多媒体分布式操作系统及高速实时网络协议的研制和发展使得电子图书馆的实现成为可能。虽然尚有一些技术不够成熟,但它们必然会飞速发展和成熟起来,10年来计算机技术日新月异的进步足以证明。而且,计算机发展到今天,成本越来越低,价格逐渐下降,已开始向家庭普及。这样,作为电子图书馆的用户必将越来越多,信息资源的利用率也必将越来越高,这反过来又会促进电子图书馆的发展。

(6) 电子图书馆建立初期虽然需要投入大量资金、设备和人力,但不能因此就盲目地否定和排斥它,而不考虑其长远效益。图书馆将丰富的信息资源转换为电子文本的形式,用户通过信息高速公路加以利用并付出一定的服务费用。用户群极为庞大,即使信息服务价格低一点,服务费用收入也将极其可观。电子图书馆的经济状况应比传统图书馆要好得多。有了充足的资金,就可以进一步开发其它的信息资源,提高信息的使用价值,开拓范围更为广阔的信息服务,真正进入良性循环。从经济效益和社会效益看,建立电子图书馆也是必要的。

3. 2 电子图书馆的信息服务

电子图书馆主要面向家庭,其信息服务范围是极为广泛的,而且也必将加大服务深度。

3. 2. 1 电子图书的阅览和信息查询

这是电子图书馆最主要的信息服务内容。用户在自己的计算机或终端前发出检索或阅览图书的指令后大约几十秒钟,查询或检索的图书即显示在屏幕上。如果感到眼睛疲劳,可打开声音开关,由计算机发出优美的声音,就象有人在朗读书本内容一样。

3. 2. 2 研究项目专业信息的查询和检索

这类信息服务主要对象是科研人员。检索服务结果可以是书目信息,也可以是电子文本全文信息。服务可实行跟踪定价,视所取得效益来决定信息服务价格。

3. 2. 3 研究项目查新

主要服务对象也是科研人员。这类服务的社会效益相当可观,它避免了重复研究,避免了大量人力、物力、财力的浪费。确是物有所值。

3. 2. 4 定题信息服务

服务对象是科研人员或科研团体。电子图书馆定期为他们提供某一课题的最新研究资料。这类用户一般来说都比较固定,信息服务的价格也不宜过高。

3. 2. 5 预测与决策服务

利用电子图书馆丰富的信息资源,可以对某一专题、某一技术、某一市场的发展前景进行预测,向用户提出预测研究报告;也可为用户某一方面的决策提供辅助信息,进行决策服务。这些都属于信息服务的更深层次。

信息高速公路上运行的电子图书馆所能提供的信息服务并不仅仅只有上面提到的这些。电子图书馆丰富的信息资源,必能在信息高速公路上发挥它的最大效益。电子图书馆信息服务的收益将随着信息服务用户数量的增长以及信息服务的深入不断增

长。电子图书馆将不会再象传统图书馆一样,依靠国家事业经费拨款艰难度日。信息服务可观的收益,可用来进行新的信息资源建设和开发,拓展新的信息服务项目,提高图书馆员待遇,对图书馆员进行再教育。

3. 3 电子图书馆的运行模式

信息高速公路上,电子图书馆必将是一个极其复杂的分布式系统。这个系统必须具有良好的扩充能力,以便当用户群大量增长、系统响应速度下降时,能够扩充和派生出子系统。电子图书馆的运行模式将是一个分布式的可扩充的开放系统。笔者在考虑了电子图书馆的各种复杂因素后,提出了一种电子图书馆运行模式(如图),其实用性和有效性有待今后实践的检验。

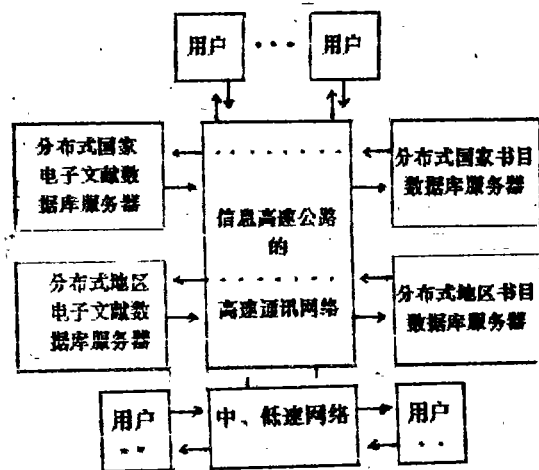


图 电子图书馆运行模式

此运行模式建有一个分布式的书目数据库服务器和一个分布式电子文献数据库服务器(它们实际上也是一个分布式网络)。数据库服务器采用客户机/服务器结构。书目数据库服务器主要存储书目数据和收藏有与书目对应的电子文献的数据库服务器的地址;电子文献数据库服务器则收藏电子文献信息。分布式数据库服务器网络特别大,服务器数量极多,每个服务器也许仅存储一个门类甚至一个小类的电子图书信息。

用户发出检索或阅览图书的指令后,指令信息通过高速通讯网络传到分布式书目

数据库服务器中央处理机,处理机到分布式书目数据库查询,将查询结果(主要是用户检索图书在分布式电子文本服务器的地址)和用户地址传送至对应的电子文献数据库服务器,从中检索出电子文献全文信息;按照用户地址,传送到用户的计算机。

当用户数量急剧增加,系统性能急剧下降时,必须对系统进行向下的扩充,在用户数量集中的地区建立地区分布式书目和电子文献数据库服务器网络,或在学科专业集中的地方建立学科分布式书目和电子文献数据库服务器网络。当国家之间要实现信息资源共享时,它也可向上扩充,国家书目数据库服务器成为全球分布式书目数据库服务器之一,国家电子文献数据库服务器也成为全球分布式电子文献服务器的成员(但首先需解决因信息格式不同而需进行的转换)。地区书目和电子文献服务器网络主要面向用户提供本国信息的服务,国家书目和电子文献服务器网络主要满足用户对国外的信息需求。

4 图书馆肩负的历史使命

建设电子图书馆是项规模巨大的工程。其中,电子信息资源的建设就是个突出问题。从长远的眼光看,图书馆有必要开始着手电子图书馆信息资源建设,而不能等到用户有强烈需求时才考虑;那时要把浩如烟海的文献一下子输入计算机,转换为电子文本信息,恐怕已来不及。

建立电子图书馆电子信息资源,摆在我们面前的当务之急是要完成信息本身、信息处理界面和网络体系结构的规范化、标准化。信息本身的标准化涉及到信息资源的规

则制订和编码,信息处理界面的规范化和标准化提供独立于设备和网络的统一应用界面,网络体系结构的规范化和标准化则提供开放的网络互连方法;目的都是为了信息资源的充分共享和信息的快速、准确传输和使用。与图书馆最为密切的是信息本身的规范化和标准化。

还需要建立一个国家性的书目数据库。要有一个有影响和有权威的国家图书馆牵头,联合各个地区的图书馆分工协作,每个图书馆负责一个门类或一个门类下的小类的书目数据库建立工作,最后将各个图书馆编制的书目数据库汇集成一个全国性书目数据库。

建立全国性的电子文本数据库也是图书馆面临的最艰巨的任务。当然,这也必须由许许多多的图书馆分工协作才能完成,任何一个图书馆都无法独自承担。由于信息量极大,传统的手工键盘输入速度太慢已无法承受,必须寻求更为快速的输入方式,印刷体汉字识别技术将得到更广泛的应用。目前已有这方面的技术和产品,如清华OCR印刷体汉字识别系统,但在性能和输入速度方面还有待进一步提高。

图书馆信息资源建设所需的投资、设备、人力都将是巨大的。除了政府资助、图书馆自身投资外,还可寻求一些公司、企业投资。本着谁投资谁受益的原则,一旦电子图书馆投入应用,投资者将得到丰厚的回报。

万齐鸣 武汉大学图书情报学院 93 级硕士研究生。通讯地址:武汉大学,邮码 430072。

(来稿时间:1995—10—04。编发者:刘喜中。)