

# 清华大学图书馆自动化信息服务系统

安树兰 (清华大学)

## 一、序言

清华大学图书馆是中国高等学校大型图书馆之一,由于经费等各种原因,全面开始实施图书馆自动化计划始于1991年,经过三年时间已经建成以一台小型机为支持环境的馆藏目录查询系统及以两台服务器为支持环境的光盘信息检索系统。馆藏目录数据库内联机可查馆藏数据包括1976年以来我馆全部中文图书目录,西文各种连续出版图书目录,1986年以来全部西文专著目录,1966年以来全部期刊馆藏,共计书目数据达200,000多万条,馆藏700,000多万册,馆内各种业务工作,诸如采购、编目、期刊管理及装订、流通等等全部采用计算机在线处理以外,还在全校配有各种检索终端,远程联机工作站达两百多台,并为其它学校设立了远程通过电话网联机的检索工作站,每天服务达15小时,联机业务处理次数逾万次。光盘网络系统配有17台只读光盘驱动器,并配有一个大型新刊篇名查询库,全校近千台在线工作站均可联机查询,每天服务15小时,检索次数近万次。馆内还建有多媒体实验室,同时可提供给30位读者利用计算机进行自我学习。

## 二、高起点的设计目标与服务目标

清华大学图书馆的自动化建设基本分两个阶段完成,尽管起步较晚,在全面分析了技术支持的可能条件之后,确立了高起点的设计目标,即充分以网络技术与数据库及相关技术为依托,建立一个多种服务方式,多种服务内容的计算机管理系统。第一阶段完成的内容是馆藏目录管理和检索,包括馆舍内业务工作自动化及馆内目录公共检索服务。第二阶段完成的内容是目录服务的上网及远程检索,光盘网络建设。我们在工作中,一直密切注视国外同期技术条件及水平。首先抓住网络环境建设,分阶段分期投资完成各项工作,以达到三年基本完成的目

标。设计的基本思路是:

第一,以网络为支持基础,实现全校的在线服务。

第二,充分利用设备的各种资源条件,实现多途经工作方式,为读者多途径、多方式联机创造条件。

第三,速度要快,特别是要以最快的速度与读者见面,使读者直接受益。

第四,一边建设、研制、试验,一边服务,同时再抓规模扩大、发展与完善。

具体作法是:

1.基本建设中抓住网络建设,1991年就在全馆建成了速度为10兆的以太网,并根据馆舍条件设计了全馆网络系统的拓扑结构,与校园网的接口方式及电话线路。由于网络技术的结构方式比较成熟,所以该项工作是骨干架一次到位,各局部分期投资完成。

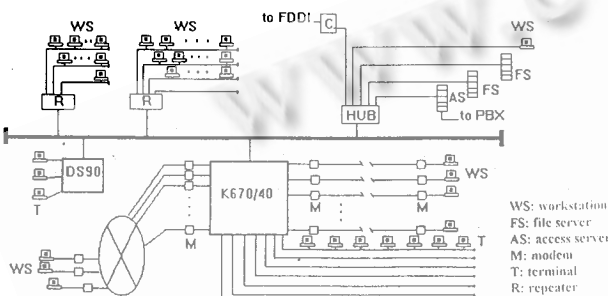
2.机器选型上注意的问题是:基本性能、安全性及开放程度与互联可能性综合考虑。因每台机器均有优势及劣势,所以在工作的开始就要充分注意。富士通的小型机安全性、可靠性很好,硬件故障率低,但开放及互联性不适于国内环境,我们就充分利用其优点。先搞数据库建设,联机业务处理自动化,并同时抓紧开放与互联工作的开发,使其能顺利上网服务。

3.实施过程中,突出速度要快,几项基本建设并头开展,条件充分的先上,先开展服务,条件差的则较晚完成,开始就进行数据库数据组织设计,数据录入,并很快开通流通及公共检索,为其它各项工作创造条件,同时也推动了这些工作。我馆小型机1991年9月中旬安装、调试,1991年11月开始装入第一批数据,1992年3月份开始公共检索试运行,1992年9月份开始流通管理,10月份开始编目与采购及期刊管理。光盘系统于1993年7月份开始考虑所需条件及技术问题,12月份资金到位,1994年3月份与读者见面运行。各种软件只有与读者见面,真正运行时,才会发现问题,进行完善。

这些作法带来明显的好处,其一是充分发挥投资效益,三年中我们累计投资约 130 万元,资金投入后基本没有设备搁置,而是快速投入运转。其二是以读者服务为推动力,推动了馆内业务的改革,不仅推动了图书馆业务流程、组织、工作方法的改进,特别是全面推动了读者教育与服务工作,大大加强了图书馆的作用。其三是便于在服务中检测与修正设计目标,因为我国图书馆自动化工作没有经验,原来的设计目标不一定是完全适合的。在开放服务后,经过自动化工作人员、业务人员、读者三方多次需求循环改善,才完善了最终目标。

### 三、系统配置与互联设计

清华大学图书馆系统环境配置如下图所示:



主要设备配置:

- K670/40 (富士通公司)专用操作系统及RDB,支持汉字处理;TCP/IP的文件服务;
- DS/90 (富士通公司)UNIX操作系统;支持TCP/IP文件服务及VT100仿真;
- FS1 (COMPAQ公司)PROSIGNIA服务器(486/33);支持NETWARE服务。
- FS2 (神通公司)4280G服务器(486/33);支持NETWARE服务。

联网说明:

- 根据我馆馆舍建筑条件,在主干网上分别联入一台 HUB 及两台 REPEATER。以保证分布在四个集中的工作站(公共检索服务大厅,多媒体实验室,读者教育实验室及计算机房)及各零散工作站的上网。

- 校园内所有联网工作站可以 NOVELL 工作站方式检索文献数据,以 VT100 仿真方式检索馆藏目录(或以客户机方式直接接受服务器服务)。

- 校园所有非联网工作站可通过 MODEM 直接联

入专用 NAS 服务器或 DS90,进行所需各种检索。

### 四、书目数据库的基本设计原则

馆藏书目数据库的建设工作量很大,往往影响自动化工作的顺利进行,我们为了保证尽快发挥机器效益,尽早为读者服务,采取了一些具体的办法,加快了工作。

1.数据的繁简程度以三个指标为主,一是馆藏规模,二是读者需求,三是机器容量,根据我馆的藏书规模及以科技图书为主的特点,我们认为基本著录项目以二级著录为宜,不必追求为全而全。而建库方法也是已有的数据项目先输入,不完全的以后再补充,比如西文期刊库,凡是数据不详的均只先输入本刊题名、中图公司刊号、ISSN、变刊前后刊名项,其它项目逐步补入,这样的好处是与读者见面快,又不影响使用。

2.输入方法多种途径,不同的书目库用不同的实施方法,西文图书因已有光盘支持,所以采用联机查重,脱机处理及批式转录方法,中文图书追溯则采用先输入书目信息,然后集中加馆藏方法,这些方法均是根据其手工目录及操作条件现状,图书分布地点不同而分别制订的。

3.充分利用计算机的特点,对检索点要作多作全,充分考虑读者使用方便。比如对题名检索点,除正题名外,还作了丛书名,部分副题名,分卷题名,合订题名等,并均作了题名关键词,而读者界面均为题名及题名关键词,因为读者分不清题名的性质,对于涉及多学科图书作了多分类检索点。对责任者,作了除第一责任者外,还提供其它各责任者及其笔名等的检索点。对期刊除本刊名外,还将变刊前名及变刊后名均作为检索点提供查询等等。多作检索点的方法正是充分发挥计算机优势,帮助读者查全的好方法。

4.抓好对检索点内容的规范处理。检索点的内容处理不规范,读者检索将十分困难,特别是中文项目中混有非汉字字符时更是这样。为此我们进行了相应规定,如非汉字字符一律大写,罗马字符一律以拉丁字符代替,汉字字符与非汉字字符间一律以一个空格相间,数字一律以阿拉伯字符输入等等。这些规定不仅对著录适用,同时交给读者以免漏检。

### 五、图书馆的网络化、虚拟化研究与建设

随着计算机与网络技术的发展,尤其是随着 Internet 的发展,通过网络能获取的图书情报资源大大增加,图书

馆的服务对象、方式、内容都较传统意义上的图书馆有了一些明显的变化,网络化、虚拟化的趋势越来越明显。清华大学图书馆长期以来一直跟踪国际先进图书情报服务的发展,并对此进行了研究与建设。

1.对国际上为解决不同机型、不同软件平台、不同数据库间的信息检索的一致性而制订的网络信息检索协议标准,包括美国国家标准 ANSI/NISO Z39.50 和国际标准化组织标准 ISO 10162,ISO 10163,做了比较细致的研究,并参照 Z39.50 编制了实用软件,即将投入运行。

2.广泛开展对外服务,通过 RS232 电话拨号和 Telnet 远程登录两种方式向用户提供清华大学馆藏资源检索。其中,前者已向国内十多家图书馆开放,后者则面对全世界能通过 Telnet 远程登录访问清华大学图书馆的网络用户服务。

3.对 Internet 的资源进行了广泛、深入的调查,把其中与图书情报密切相关的资源加以收集、整理,建立了超文本 WWW 界面的清华大学图书馆的 Home Page,并向用户提供简单、方便、有效的途径来访问清华大学图书馆及 Internet 上的许多资源,该服务向全世界的 Internet 用户开放,主要包括:

(1)国内(清华,北大,中科院)及国外上十所大学的图

书馆馆藏及其他书目信息检索。

(2)全球上千个图书馆 GOPJER 服务器。

(3)全球数千种免费电子期刊(按学科字母排序)。

(4)覆盖国内绝大多数学科的 WAIS 数据库资源(按学科排序)。

(5)覆盖国内绝大多数学科的 GOPHER 资源(可按关键字或学科查询访问)。

(6)覆盖国内绝大多数学科的 WWW 资源(可按关键字或学科查询访问)。

(7)查询 Internet 上的人事、组织机构信息(如查询 E-MAIL 地址等)。

(8)查询 Internet 上的免费软件地址(anonymous FTP Sites 查询)。

访问方法:使用 WWW(包括 Mosaic,Netscape,Lynx 等),地址为:

<http://www.cernet.edu.cn:80/tsinghua/library/index.html>

通过这个超文本的界面,读者就可以方便地获取国内外 Internet 网上的资源而不用区分,也不用了解这些资源究竟在网上哪个地方。在此基础上,初步形成清华大学虚拟图书馆服务的基本框架。