

# 视频点播系统两级存储结构的设计与实现

林 坤 辉

(厦门大学计算机科学系, 福建 厦门 361005)

**摘要:** 随着网络和多媒体技术的飞速发展, 对视频点播系统的需求急速增加, 但传统的视频点播系统造价昂贵, 不利于视频点播系统的推广应用. 本文论述了一种基于大容量 IDE 硬盘的视频点播系统两级存储结构的设计思想、硬盘配置方案、视频文件的存储方式、系统开发模式和实现方法, 为视频点播系统的设计提供了一种经济实用的新方案.

**关键词:** 视频点播; 存储设计; IDE 硬盘

**中图分类号:** TP 37

**文献标识码:** A

视频点播(VOD)系统处理的是多媒体信息. 多媒体信息既有传统的结构化文本和数值数据, 又有非结构化的文本、图形、图像、音频和视频. 视音频信息的特点之一是数据量大, 如每片 VCD 需要 1.2 GB 的存储量, 每片 DVD 需要 4.7 GB 的存储量, 当存储大量的节目时, 对存储器的容量要求很高; 视音频信息的特点之二是实时性强, 信息的再现有严格的时间间隔约束, 必需以均匀稳定的速率实时传输, 如 VCD/MPEG-1 每路视频流为 1.5 Mbps, DVD/MPEG-2 每路视频流为 3~8 Mbps, 特别当大量用户并发请求时, 对存储器的内部速率和 I/O 吞吐量要求很高<sup>[1]</sup>. 因此, 视音频数据的存储和获取模式直接影响视频点播系统的性能. 视频服务器的存储器必须容量大且有良好的物理结构和逻辑结构, 才能保证磁盘的有效利用和应用的快速存取.

## 1 传统的三级存储结构

传统的视频点播系统通常采用三级存储结构: 内存(RAM)作为第一级存储设备, SCSI 硬盘或 SCSI 磁盘阵列作为第二级存储设备, 磁带库或光盘库作为第三级存储设备. 由于视频流必须实时、连续、并发地传输, 需要使用高速的磁盘阵列来存放节目. 通常, 视频服务器中配置 SCSI 硬盘和 SCSI 磁盘阵列, 大多数专

用的视频服务器机型需采用 SCSI 硬盘和热插拔 SCSI 硬盘组成大容量存储器. SCSI 硬盘在采用阵列功能后可大幅度提高磁盘输出能力, 但 SCSI 硬盘存储成本高, 容量价格比低. 为了有效降低存储成本, 节省开支, 一般采用低速大容量低价位的存储体, 如磁带库、光盘库或光盘镜像塔, 来代替昂贵的 SCSI 硬盘. 光盘镜像塔实际上是采用硬盘作为主存储器, 通过 CD-ROM 将光盘镜像到硬盘上. 光盘镜像塔由于配置了 100~400 GB 的硬盘存储器, 所以, 光盘镜像塔可镜像 160~640 片 650 MB 的光盘片. 一般把点播频率高的节目存储在第二级存储设备中, 来响应实时的点播请求, 而将点播频率较低的节目存放在第三级存储设备中, 作为非在线节目<sup>[2]</sup>. 由于磁带库和光盘库的速度较慢、存取时延大, 为保证播放的质量, 当需要访问磁带库或光盘库上的节目时, 通过节目管理软件的调度算法, 再把它们上载到视频服务器上<sup>[3]</sup>. 由此引起视频点播系统的存储结构复杂, 存储管理与节目调度算法复杂, 软件设计复杂, 点播响应时间长. 采用这种三级存储结构的 VOD 系统, 其存储系统造价昂贵, 系统性能价格比低, 不利于视频点播系统的推广应用.

## 2 两级存储结构的设计

### 2.1 IDE 硬盘与其他存储介质的比较

近年来 IDE 硬盘的容量越来越大、内部速度和 I/O 速度越来越快、价格越来越低. IDE 硬盘与 SCSI 硬盘虽然结构不同, 但并没有本质的差别, 其核心是

收稿日期: 2001-11-22

作者简介: 林坤辉(1961-), 男, 副教授.

表1 IDE硬盘与SCSI硬盘对比表

Tab.1 The table of comparing IDE HD with SCSI HD

比较参数	DeskStar75 GXP 系列 IDE 硬盘	UltraStar36LP 系列 SCSI 硬盘
容量(GB)	75/60/45/30/20/15	36.9/18.3/9.1
转速	7 200 RPM	7 200 RPM
接口类型	Ultra ATA/100 IDE	Ultra160 SCSI
持续数据传输率	37 MB/sec	19.5 ~ 31.9 MB/sec
平均寻找时间	8.5 ms	6.8 ms
最大内部传输率	444 Mbits/sec	248 ~ 400 Mbits/sec
最大接口传输率	100 MB/sec	160 MB/sec
数据磁头数(物理)	10/8/6/4/3/2	10/5/3
数据碟数	5/4/3/2/2/1	5/3/2
每 GB 存储费用(人民币)	20 元左右	200 元左右

表2 IDE硬盘与其他存储介质的对比表

Tab.2 The table of comparing IDE HD with other storage medium

比较项目	IDE 硬盘	SCSI 硬盘	磁带库	光盘库	光盘镜像塔
相同盘体寻找延迟	8.5 ms	6.8 ms	几秒	几十毫秒	几毫秒
更换盘体寻找延迟	在线不需更换	在线不需更换	几秒	几秒	手动更换光盘
同时服务用户数	100/每硬盘	160/每硬盘	1/读取驱动器	10/CDROM	66
同时读取驱动器数	4~8	4~8	2~4	2~7	6
每个读驱动器非压缩持续传输率	20~37 MB/s	19.5~31.9 MB/s	6 MB/s	6 MB/s(40速)	20~30 MB/s
每 GB 存储费用	20 元左右	200 元左右	200~300 元	100~200 元	100 元左右

一致的,都是在磁介质上达到最高的存储密度.以同样是 7 200 转的 IBM 的 DeskStar75 GXP 系列 IDE 硬盘与 UltraStar36LP 系列 SCSI 硬盘为例作个比较<sup>[4]</sup>,如表 1 所示.

ATA/100 接口允许主机和硬盘之间以 100 MB/s 的数据传输率进行传输数据,这个接口已被正式确立为硬盘的下一代接口类型,ATA/100 (DMA100) 标准于 2000 年 6 月 2 日在美国正式确立.ATA/100 接口同样包含 CRC 特性,这能增加传输数据的完整性和可靠性,同时它能检测到数据传送中的错误.因此,具有 ATA/100 接口的 IDE 硬盘与具有 SCSI 接口的硬盘性能相近,但 IDE 硬盘每 GB 成本 20 元左右,SCSI 硬盘每 GB 成本要 200 元左右,由于存储成本占视频服务器造价的 50% 以上,因此,为取得最佳的容量价格比,必须采用价廉物美的大容量 IDE 硬盘,替代比较昂贵的 SCSI 硬盘.

大容量 IDE 硬盘与磁带库、光盘库等低速率的存储介质相比也具有寻找延迟低,传输速率快,每 GB 存储成本最低等优点.采用 IDE 硬盘或其他存储介质作为 VOD 系统的存储器,其主要参数对比如表 2 所示.

从上述比较可以看出,IDE 硬盘具有以下突出的优点:

- 1) 容量大,单硬盘容量可以达到 75 ~ 100 GB.
- 2) 速度快,接口传输率 100 MBps,最大内部传输率 55 MBps,持续的数据传输率 37 MBps.
- 3) 价格低,每 GB 价格仅 20 元左右.

IBM 的 DeskStar 系列 IDE 硬盘面向的是全球范围内的高级桌面电脑或影音应用,它具有更低的工作噪声、更少的能耗及更高的磁盘性能等特点,这使得它能满足高性能电脑应用对硬盘性能指标的要求,而且 IDE 硬盘的性能指标与 SCSI 硬盘相近,但价格仅是 SCSI 硬盘的十分之一,采用 IDE 硬盘替

代 SCSI 硬盘,可获得理想的性能价格比.同时,由于 IDE 硬盘容量大且价廉,没有必要再采用低速率的第三级存储器,因此可取消传统的第三级存储器,使 VOD 系统中多媒体数据的存储采用 RAM 内存—IDE 硬盘在线两级存储器结构.

## 2.2 第一级存储设计

第一级存储设备采用廉价的 RAM,建立多级缓存机制.视频点播的工作模式是当视频服务器响应客户的视频请求后,从存储系统读入一部分视频流数据到服务器缓存中,然后服务器缓存中的内容送入网络接口发送到客户端的缓存中,服务器必须保证在客户端缓存中的视频数据播放完之前,将所需的视频数据再传送到客户端的缓存中.RAM 是计算机中必不可少的部分,它可以作为视频和音频的稳定缓冲区.多级缓存如图 1 所示.

视频点播系统是有服务质量(QoS)要求的,比如视频流和音频流必须同步、响应要及时、延时要少、并以稳定的速率传输,而 TCP/IP 协议不能保证视频流和音频流的服务质量.因此,必须采用多媒体流控制、支持 802.1p 国际标准.由于 RAM 已十分廉价,可以配置足够容量的 RAM,建立两套缓冲区,分别作为读取缓冲和传输缓冲.使用两套缓冲区可以使对于每个用户的每个请求,数据读取和传输服务可以并行进行,同时采用多线程的视频流的调度算法.能有效防止视频数据从存储系统读取和在网络中传输所产生的延时抖动问题,提高视频服务质量,保证客户端可获得流畅的视频和音频信息.

## 2.3 第二级存储设计

第二级存储设备采用廉价的 IDE 硬盘,建立大容量存储器. IBM 75 GXP 系列玻璃硬盘特性如下<sup>[4]</sup>:单碟 15 GB、7 200 转/分、2 M cache、最大持续

速率 37 MB、支持 ATA 100/ATA 66/ATA 33.推荐采用具有 ATA/100 接口的 IBM 75 GB 或 Maxtor 100 GB 的 IDE 硬盘.通过以下几种方法可以有效扩展硬盘存储器容量:

1) IDE 硬盘一般均采用内置方式,但内置的硬盘个数受机箱和电源的限制.当采用特制的 300 W 以上电源时,有很多型号的服务器或机箱可以支持 8 个或 8 个以上的内置 IDE 硬盘.比如使用 Enable 的机箱可内置 8/12/16 个 IDE 硬盘<sup>[5]</sup>.

2) IDE 硬盘支持热插拔及 RAID 5,采用 IDE 阵列卡不仅能提高磁盘输出能力,还为机器增加 4 个 IDE 内部硬盘连接能力.在单台服务器上采用 Promise 公司的 Promise SuperTrack 66 IDE RAID 5 卡,可以联合多个 IDE 硬盘.还可以采用外置 IDE RAID 产品,通过 SCSI 卡或 LAN 连接到视频服务器<sup>[6]</sup>.

3) 专门配置一台有大容量 IDE 硬盘和网卡的 PC 机,在视频服务器上将该机的硬盘共享为本机的盘符,再将这些盘符加入存储设定.

4) 通过在视频服务器上加入 IEEE 1394 卡,外置 IEEE 1394 接口的硬盘即可扩展存储容量,IEEE 1394 接口的硬盘支持热插拔,且不受机箱电源限制.

5) 通过服务器集群,增加视频服务器,可线性增加存储容量.每台视频服务器可以配置 600 GB 的存储器,最大可输出 600 Mbps 的视频数据,通过多台视频服务器集群可达到在线数 TB 的存储容量.不同服务器除了 20%左右的存储容量存放点播频率较高的相同节目外,其余 80%左右的存储容量存放不同的节目<sup>[7]</sup>.因此每增加一台视频服务器,可增加 600 Mbps 的输出能力和 480 GB 的存储容量.

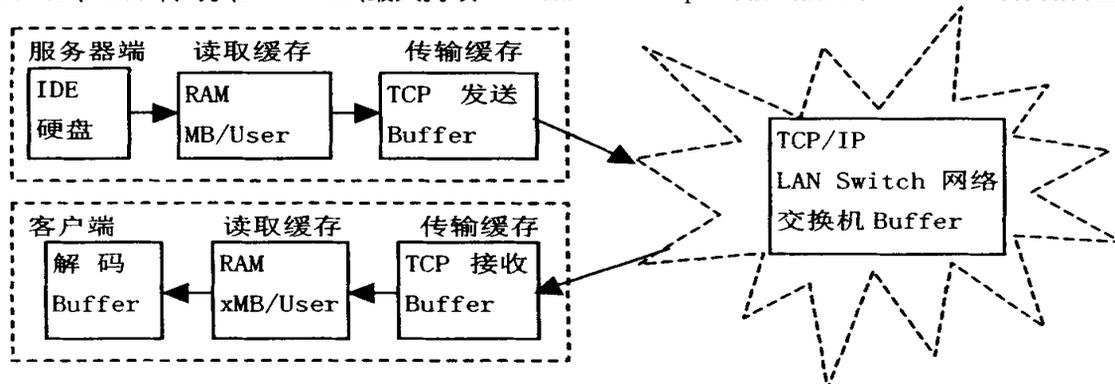


图 1 视频点播系统的多级缓存

Fig. 1 The multilevel buffer storage of VoD system

VOD 系统面对的是大量客户的点播请求,可以是不同的用户点播不同的节目,也可以是不同的用户点播相同的节目,这就要求有一种能适应这种需求的数据存储方式。而且,多媒体数据之间的时空相关性也要求各媒体数据的存储、检索和获取需要作特殊的处理。因此,一般 VOD 系统的视音频数据都采用专用的存储格式和专用的数据管理工具。当采用 IDE 硬盘方案时,每个硬盘可服务的用户数多,可满足一般应用的要求,因此视音频数据已经不再需要专用格式存储。视音频数据在 FAT 格式、FAT32 格式(每个分区管理 4 GB 数据)及 NTFS 格式中达到 I/O 最高性能,视音频数据可以采用标准通用的文件格式直接存储在视频服务器的硬盘上,从而大大简化了服务器对视音频数据的管理。

### 3 原型系统开发

VOD 系统开发必须解决两个最重要的问题:处理随机的并发请求问题和大量数据存储问题。对于处理请求问题,要求视频服务器具有大规模并行处理能力,要求网卡有足够的 I/O 速度,网络有足够的带宽;对于数据存储问题,要求有大容量的存储器,磁盘有足够的读吞吐量。下面是一个具有两级存储结构的 VOD 原型系统的开发实例。

1) 服务器配置。Web 服务器、数据库服务器、视频服务器由单一服务器担任。采用 Intel 架构的服务器已能满足系统对主机 I/O 速度的要求<sup>[8]</sup>,CPU 采用 P III 1 GHz 133 MHz FSB,内存 PC-133 RAM 512 MB,IDE RAID 卡为 Promise FastTrack 66,采用 IBM 75 GB IDE 硬盘,配置 4 个硬盘共 300 GB 作为存储器。配置 1 个 1 000M 网卡(3c985)。

2) 网络设计。网络结构分为四层,第一层:服务器;第二层:100 M 交换机,100 M/1 000 M 连接服务器;第三层:10 M 交换机,100 M 连接 100 M 交换机;第四层:10 M 交换机经 10M 共享 HUB 联 PC 机或直接联 PC 机。10 M 网络最多只能满足 6 个用户同时点播 VCD,100 M 网络最多只能满足 50~60 个用户点播同时 VCD。系统设计时,需要考虑网络其他应用所需要的网路流量,因此,一般情况下,10 M 网络设计时不应超过 4 个 VCD 并发用户,100 M 网络设计时不应超过 40 个 VCD 并发用户,1 000 M 网络设计时不应超过 400 个 VCD 并发用户,留一部分的带宽作为其他应用。对于点播 VCD 片,采用 10

M 交换到桌面即可。DVD 的平均码率在 6~8 Mbps,有时超高质量 DVD 的动态码率可达 10 Mbps。系统设计时,应考虑能支持超高质量的 DVD。对于普通的 DVD 片,10 M 交换到桌面就可以支持,对于超高质量 DVD 片,需要 100 M 共享或 100 M 交换到桌面才能支持。

3) 软件开发。服务器平台采用 Windows 2000 Server,使用 ASP 与 HTML、VBScript、JavaScript 混合编程开发 Web 应用程序<sup>[9~11]</sup>,使用 Microsoft Access 数据库存储用户点播信息和多媒体数据的文件名及路径信息,使用 VBScript 和 ADO 对象进行数据库应用程序开发,采用 C++ Builder 5.0 开发视频服务器端的点播 A/V 数据流的控制和管理软件<sup>[12~15]</sup>,采用 Visual C++ 6.0 开发一个传输及控制 A/V 数据流播放的浏览器插件。

4) 客户端环境。客户机平台采用 Windows 98,视频点播软件的用户界面采用纯网页架构,用户只要使用 Microsoft IE 浏览器就可以运行纯 Web 架构的视频点播软件。

### 4 系统性能分析

采用 IDE 硬盘作为大容量存储器,可获得最佳的性能价格比。以相同容量和读取速度的 7 200 rpm 的 IDE 硬盘与 SCSI 硬盘比较,价格相差几倍至数十倍。同样配置 600 GB 的存储器,IDE 方案将比 SCSI 方案节省几万甚至数十万元人民币的投资。单视频服务器按配置 600 GB 的存储器计算,可存储 500 多部 VCD 电影,或 1 000 张 VCD 片,或 12 000 首 MTV,或 120 000 首 MP3 歌曲。由于大量的多媒体数据直接存储在视频服务器的硬盘里,使得每台服务器可服务大量用户的并发访问。每个硬盘可同时服务 100 个 VCD 并发用户,每台服务器最多可配置 600~800 GB 的内置硬盘容量。通过测算,安装 1 000 M 网卡的服务器最大可以并发 400 个 VCD/MPEG1 用户,允许 200 个 VCD 并发用户同时点播同一节目,节目点播延迟时间小于 3 s,节目切换位置延迟时间小于 1 s。如果通过服务器集群,可以支持数千个用户进行 VCD 级别的点播。DVD 的平均码率在 6~8 Mbps 之间,系统能支持 DVD 的用户数按支持 VCD 用户数的 1/4~1/5 计算。

通过选用大容量快速廉价的 IDE 硬盘,甩掉了昂贵的 SCSI 硬盘和低速率的第三级存储器,实现

了 VOD 系统的两级存储结构,简化了 VOD 系统的存储设计、算法设计和软件设计,降低了系统开发难度,缩短了系统开发周期,通过对容量、带宽、成本的优化配置,使得普通的 PC 服务器就可以作为高性能视频服务器使用,大大降低了 VOD 系统的造价,为设计性能价格比高的视频点播系统提供了一种经济实用的解决方案。

### 参考文献:

- [1] 催杜武,李家和,张烨,等. 网络多媒体实用技术[M]. 北京:人民邮电出版社,2000.
- [2] 孟桂娥,杨宇航. 等级存储的分布式可扩展 VOD 系统的设计[J]. 电子技术,2000(8):15-18.
- [3] 傅锡天,孙景鳌,蔡安妮. 分布式 VOD 系统的分级存储调度[J]. 通信学报,2000(3):64-69.
- [4] IBM Corporation. Hard Disk Drives. <http://www.storage.ibm.com/hdd/>. 2001-08.
- [5] Enable Computer Technology Co. File Server Chassis. <http://www.enable.com.hk>. 2001-08.
- [6] Promise Technology, Inc. ATA RAID Cards. <http://www.promise.com/>. 2001-08.
- [7] 彭宇行,陈福接. 分布式 VOD 系统的视频数据存储[J]. 计算机学报,2000(6):671-673.
- [8] Intel Corporation. Chipsets. <http://www.intel.com/products/>. 2001-08.
- [9] Stephen Walther 著. 希望图书创作室译. Active Server Pages 2.0 Unleashed 揭密[M]. 北京:北京希望电子出版社,2000.
- [10] 任伟,郝水侠. Visual InterDev 6.0 开发指南[M]. 北京:北京航空航天大学出版社,1999.
- [11] 李世杰. Active Server Pages (ASP) 2.0 网页设计手册[M]. 北京:清华大学出版社,1999.
- [12] 刘衡竹,陈旭灿,陈福接. 视频服务器中视频流存储分配算法的研究[J]. 计算机学报,1998(4):289-295.
- [13] Kuo F, Effelsberg W, 著. 龙晓苑译. 多媒体通信协议与应用[M]. 北京:清华大学出版社,1998.
- [14] 裘晓峰,张春红. 宽带网络技术及其应用[M]. 北京:清华大学出版社,1997.
- [15] 网络编程系列编委会. C++ Builder 5 Web 编程[M]. 北京:中国人民大学出版社,2001.

## The Design and Implementation of Two Level Storage Architecture in VoD System

LIN Kun-hui

(Dept. of Compu. Sci., Xiamen Univ., Xiamen 361005, China)

**Abstract:** As the network and the multimedia techniques develop rapidly, the need for VOD system is increasing quickly. But the price of the traditional VOD system is very expensive. It is not beneficial to the spread application of VOD system. This article describes the design idea of two level storage architecture, the schema of hard disk configuration, video data storage method, system development model and implementation method of VOD system which is based on large-capacity IDE hard disk. It provides a kind of economic practical new scheme for the design of VOD system.

**Key words:** VOD; storage design; IDE hard disk