

# 一个支持并行流的多媒体服务系统<sup>\*</sup>

吴 飞 陈福接 王 朴

(国防科技大学计算机系 长沙 410073)

**摘要** 分布式多媒体是分布式计算技术和多媒体技术相结合的新兴技术,其基本特征是实现异地多媒体信息共享.由于连续媒体数据与传统数据在特性上的巨大差异,使得传统网络服务器和相关的存储系统不再适宜支持异地连续媒体的服务,所以要求代之以新的存储系统.本文介绍了作者所开发的分布式多媒体并行服务系统.该系统通过引入系统资源秩概念,实现了多访问流的并行操作,从而弥补了现有服务系统在设计上的主要缺陷.另外,文中给出了一组视频访问流的测试数据,其结果显示系统对多媒体应用的良好支持.

**关键词** 分布式多媒体,连续媒体,视频服务,调度,并行访问.

**中图分类号** TP391,TP316

近年来,分布式技术、网络技术和多媒体技术的发展使得联机(on-line)访问异地巨量媒体数据(包括图象信息、声频信息、视频信息等)成为可能.然而,媒体信息不同于传统的科学处理的信息,有其与众不同的特征.<sup>[1]</sup>

• **连续性** 声频和视频类的信息都是以时间为轴的连续信息.通常,其采集设备(如摄影机、录音机等)都实时地产生连续媒体数据;播放时,也要求连续实时.因此网络传送媒体数据时,必须保持连续性和稳定性;相反,传统数据传送多呈现突发特性.所以,传送多媒体信息需要网络具有更高的频宽.

• **数据量大** 多媒体视频文件的数据量一般都远大于传统数据处理领域内文件的数据量.例如,当以 DVI 的标准采集视频时,即使采用了各种压缩技术,10m 的数据量也会高达 100MByte.可想而知,当采用更高视频采集标准时,长达数小时的数据量之巨大.

• **全局性** 该特性体现为某媒体信息段一经使用,很少存在缓存中重复使用;即使重复使用,也很少在缓存中保留该信息段.相反,传统的数据处理领域,数据使用具有局部性.因此,基于局部的可重复性的 Cache 技术对连续媒体处理无益.在这方面,Hanko 对工作站的使用状况进行了测试.<sup>[2]</sup>

综上所述,媒体信息的 3 大特性要求改变了在存储系统设计中的思路和方法.从解决巨量连续媒体的读、写入手,以求能从多方面——时、空、频宽——考虑资源管理而设计出满足

\* 本文研究得到并行与分布处理国家重点实验室资助.作者吴飞,1968年生,博士,主要研究领域为计算机组织与系统结构,分布式多媒体技术.王朴,1938年生,教授,主要研究领域为计算机组织与系统结构,人工智能,多媒体.陈福接,1935年生,教授,主要研究领域为超高性能计算机,多媒体计算机技术.

本文通讯联系人:吴飞,长沙 410073,国防科技大学计算机系

本文 1996-04-30 收到修改稿

分布式多媒体应用需要的并行存储系统.<sup>[3,4]</sup>

# 1 分布式服务系统组成

## 1.1 系统概述

通常,分布式多媒体服务系统主要有 2 类设计方案:①局限于视频服务要求,因而开发了只读的 VOD(video-on-demand)分布式视频服务系统.该系统适宜于通过“Broadcast”、“Batch”和“Multicast”技术实现对远程多用户的视频服务.它的好处主要是可利用现有的通讯线路,不需要太多附加设施.如文献[5]所论述的视频服务器即为此类.然而,该类设计方案不支持交互式的多访问流并行服务,也不能存储多媒体信息(尤其是视频信息).②局限于巨量媒体数据存储服务要求,开发了分布式的并行存储服务系统.由于现有网络条件的限制和设计上的原因,它对连续媒体服务的支持能力不强,并且缺乏多用户多访问流服务的支持.美军方采用的 ISS 系统中,就采用了如此的设计方案.<sup>[6]</sup>

为了更进一步改善存储系统对视频服务的支持,更为重要的是为了实现对多用户多访问流的并行服务,弥补现有设计的缺陷.我们设计和实现了 DMPSS 服务系统.从功能实现上看,该系统和传统的网络文件服务系统有许多相同和相似之处,尤其都有共同的目标——追求最大化的数据吞吐率.但是,前者更为强调对高速、稳定的访问流支持,强调对流同步的支持;而后者则更为强调对数据一致性和纠错能力的支持.

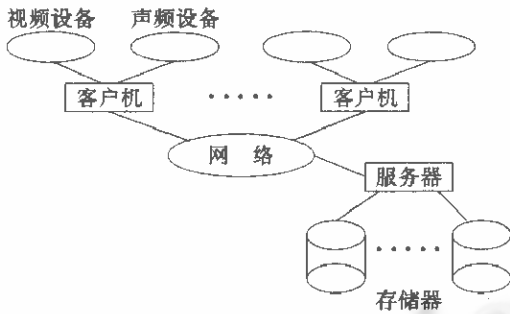


图1 分布式多媒体硬件构成系统

另外,从维护数据资源的一致性要求、减少不必要的网络传送要求以及利于管理的要求出发,我们选择了图 1 所示的分布式多媒体系统硬件结构.

## 1.2 网络服务

通讯网络连接服务器和终端设备(客户机)历来就是分布式系统的“瓶颈”,更不用说传输量巨大、频宽要求高的分布式多媒体系统了.表 1 列出了几种媒体数据所要求的网络服务质量(QoS).表中的数据是单路、无限时连续演播条件下的频宽要求;而当同时服务多路流时,根据频宽  $B$  与并行流数  $n$  的关系式:

$$B = f(n) \geq Qn;$$

以 5 路 HDTV 视频为例,即使采用了有损压缩技术,网络频宽至少也要 100Mbps.如考虑网络自身开销,那么要求将会更高.由此可见通讯网络在多媒体分布系统中的重要性,这也是现阶段网络视频服务大多采用 MPEG 视频标准的原因之所在.

# 2 并行存储系统设计

DMPSS 系统设计存在 2 条重要的背景要求:①DMPSS 系统是一个 Client/Server 的分布式系统.其工作方式是:只有当 Client 向 Server 发出请求后,Server 才服务,而 Server 不能主动为 Client 服务.另外一个隐含工作方式是:当某 Client 向 Server 发出请求后,在请求

表 1 各种媒体对象的 QOS 要求

| 媒体类型 | QOS 要求 | 范围                     | 对象类型        |
|------|--------|------------------------|-------------|
| 声 频  | 点到点延迟  | <400ms                 | 会议声频        |
|      | 会议延迟   | <800ms                 |             |
|      | 网络频宽   | $\geq 32\text{Kbps}$   | CD 质量       |
|      |        | $\geq 365\text{Kbps}$  |             |
| 视 频  | 点到点延迟  | <250ms                 | MPEG 标准     |
|      | 网络频宽   | $\geq 1.86\text{Mbps}$ |             |
|      |        | $\geq 140\text{Mbps}$  | TV, PCM 无压缩 |
|      |        | $\geq 20\text{Mbps}$   | HDTV 有损压缩   |
| 普通数据 | 点到点延迟  | <1s                    | 传统文件        |
|      | 网络频宽   | 0.2~10Mbps             |             |

没有得到服务(或撤销)之前,该 Client 不能向 Server 发出新的服务申请. ②该系统的现阶段应用开发主要是多媒体合著系统和会议系统. 为了很好地支持这两类应用开发,我们在并行存储系统设计过程中,充分考虑了应用领域的特性;采用了图 2 所示的并行存储系统设计方案.

### 2.1 多访问流的调度管理

通常,任何一个客户都希望自己的请求能尽快地被服务;在多用户的分布式环境下,假若没有恰当的调度机制来保证各个用户的请求被妥善地折中,那么不仅会降低系统服务性能,而且可能无法实现多访问流的操作;严重时,甚至造成系统的“崩溃”. 所以,我们的任务是在确保系统和用户要求的 QOS 不被降低的情况下,采用恰当的排队和调度策略,使系统服务性能趋近最佳,并且尽量降低用户的平均等待时间.

#### 2.1.1 系统访问模型

根据 DMPSS 系统设计的背景要求,可以得到图 3 的系统访问排队模型. 其中顾客总体数目  $m$  为 DMPSS 系统的客户机总体数目,而队列长  $n$  为并行存储系统中正在等待服务的客户数与正在被服务的客户数之和.

服务系统

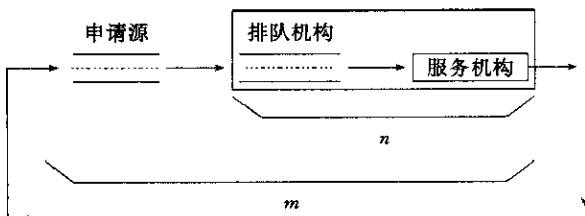


图3 系统工作的访问模型

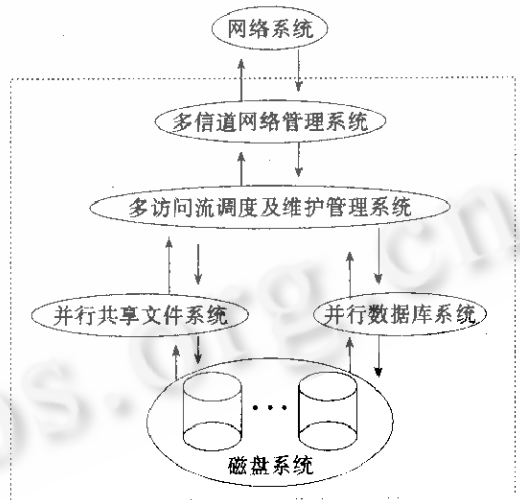


图2 多媒体并行存储功能结构图

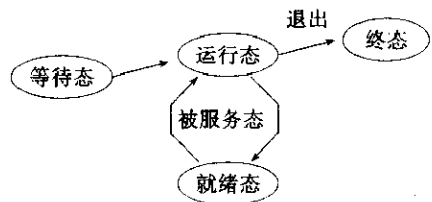


图4 多级动态优先级调度示意图

在 DMPSS 系统稳态工作时,通过取样分析发现:客户请求的到达时间(即发出服务请求的间隔时间)和服务时间(即客户请求所需的服务时间)均趋近于普阿松分布,因此我们令两者均服从普阿松分布.又由于服务台是单一的服务器,故依据排队论知识,可得出该系统访问模型是:  $M/M/1/m/m$  模型.<sup>[7]</sup>

### 2.1.2 动态优先级调度策略

调度策略的研究历来是系统设计的“重头戏”,对于分布多媒体的系统设计也是如此.<sup>[8,9]</sup>现有的调度策略主要有以下 2 类.

① 基于先来先服务的(FCFS)调度策略.该类策略实现简单,时间开销小.如文献[4]中的视频服务系统就选择了基于 FCFS 的调度策略.然而,FCFS 类策略无法体现请求的紧迫性,而且当先前请求的服务时间太长时,会出现严重滞后后继请求的现象.

② 基于优先级的调度策略.该类策略调度十分灵活,特支持多任务环境下的多用户并发工作,利于分布式多媒体服务系统使用.不过,如果设计不当,不仅无法避免本身固有的时间开销较大的缺陷,而且会导致某类请求被“饿死”的现象,甚至会出现系统“死锁”的情况.

为了支持多用户的多访问流调度,DMPSS 系统采用了“多级动态优先级调度”策略,所谓“多级动态优先级调度”是指如下的二级调度策略.(见图 4 所示的调度示意图).

第 1 级:对处于等待队列中的请求流实施调度.具体含义是:根据某种秩  $r_1(p)$  选择秩大的请求流,然后调度该请求流进入服务状态(即运行态).秩  $r_1(p)$  的实现机制是

$$\text{秩 } r_1(p) = \text{基本全局秩} + \text{时间秩} + \text{数据长度秩};$$

其中“秩  $r(p)$ ”定义为:某集合  $A$ ,为确定  $A$  中各元素  $p_i$  的排序关系;故对  $A$  中各元素实施一法则  $r$ ,从而得到一组新值  $r(p_i)$ ,再根据  $r(p_i)$  而确定出各元素  $p_i$  的排序关系.则称  $r(p_i)$  为秩.

第 2 级:对处于服务圈中的请求流实施调度,具体含义是:根据某种秩  $r_2(p)$  选择秩大的请求流,然后重新调度该请求流进入运行状态,并将原运行状态的请求流转换到就绪态;如此反复,直至该请求流被服务完成(或者被强迫中止)而退出到终态.秩  $r_2(p)$  的执行机制是

$$\text{秩 } r_2(p) = \text{秩 } r_1(p, t_0) + \text{系统资源秩};$$

其中秩  $r_1(p, t_0)$  是指当请求流从等待态调度到运行态时所具有的秩值,而  $t_0$  是调入时刻.

### 2.1.3 系统资源秩

由于要处理的对象是具有全新特性的连续媒体信息,因此决定了我们在调度中引入系统资源秩.通过该秩的引入实现了如下的控制策略.

(a) 当响应读请求流时,并行存储系统及时提供用于连续播放的数据块;另一方面,当读取了用于连续播放的足够数据时,系统要暂停供给媒体数据,以防止产生过量的数据而过分消耗(或耗尽)系统资源.

(b) 当响应写请求流时,并行存储系统及时存放生成的连续媒体数据,防止为了保持信息不丢失而过分消耗系统资源;另一方面,只有当生成了足够的媒体数据后,系统才进行存储,以防止服务器空等媒体数据的到达,从而避免系统资源的空耗.

从概念上讲,某请求流的“系统资源秩”反映了保持该请求流连续演播的最小系统资源开销与实际占用的系统资源比值的变化.也就是说当比值过小会导致系统资源秩迅速增大,从而调度该请求流到运行态;当比值过大时,系统资源秩则会迅速下降,从而强迫调度该请

求流到就绪态,这里所谈到的“最小系统资源开销”、“实际占用的系统资源”以及确定“比值的”变化”都是我们所研究的重要内容,本文只简单阐述单路视频演播流所涉及的“最小系统资源开销”,详细内容可参阅我们有关论文和报告。

在单路演播流的条件下,该值的大小与客户端的 QOS、网络的 QOS 和服务器的 QOS 相关,以服务器响应某客户的视频请求为例,令 0 时刻为该客户发出请求时刻,经过  $t_1$  时间后服务器调度该请求而开始服务响应,再经过  $t_2 - t_1$  时间客户端“能够”开始播放该视频,如果令 X 轴方向为时间 T 轴,取 Y 轴为数据块量 Q 轴,则可得图 5 的演播时间与数据量的 T-Q 图。

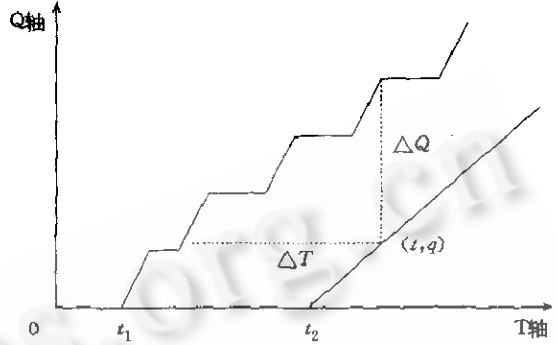


图5 演播时间与数据量的T-Q图

从 T-Q 图中,我们可以得到 2 个反映系统性能和资源使用状况的参数  $\Delta Q$  和  $\Delta T$ 。参数  $\Delta Q$  的含义是:在  $t$  时刻系统已经读出但还没有被播放的数据量,显然这些数据是被存放在系统的缓冲区内,因此,从理论上讲,只要系统保证提供至少  $\Delta Q$  大小的缓冲区,就可以实现连续播放媒体信息,也就可以实现带有系统资源秩的动态优先级调度,从而实现多访问流的操作,而参数  $\Delta T$  表明如此的关系:演播的媒体数据必须至少提前  $\Delta T$  时间被读入系统缓冲区内,值  $t_1 + \Delta T$  表示最早许可演播时间。

当具体地采用如下的符号时,我们得到了针对单路视频演播的公式(1)。

$$\Delta T \geq \begin{cases} \frac{W}{nB'} & (B' > W) \\ \frac{W}{nB'} + (q - \frac{q}{W}) & (B' \leq W) \end{cases} \quad (1)$$

$$\Delta Q \geq \begin{cases} \frac{W}{n} + \frac{B'}{k} & (B' > W) \\ \frac{W}{n} + (q - \frac{B'q}{W}) + \frac{B'}{k} & (B' \leq W) \end{cases} \quad (1)$$

其中  $\frac{1}{B'} = \frac{1}{S} + \frac{1}{B} + \frac{1}{V}$ ;  $q$  为演播数据量,  $B$ : 系统的传输网络的频宽;  $S$ : 数据从服务器缓冲区上网的速度;  $V$ : 数据从网络到客户机缓冲区的速度;  $W$ : 客户机上的连续媒体播放的速度;  $n$ : 单位时间内演播的视频帧帧数;  $k$ : 单位时间内网络传输的数据包数目。

表 2 是针对几种不同标准的媒体,在确定 QOS 要求下计算出的最小启动时间和最少缓存占用量。(为更清晰地说明,特取频宽  $B = 20\text{Mbps}$ ,  $k = 20$ ;  $V = 100\text{Mbps}$ ;  $S = 50\text{Mbps}$ ; 并假定各请求的平均等待时间  $T_w = t_1 = 0$ 。)

表 2 媒体演播所要的最小资源占用表

| 媒体类型        | 演播 QOS 要求<br>速度及帧数 | 演播时间(s) | 理论最小缓冲 $\Delta Q$ | 理论最小延迟 $\Delta t$ |
|-------------|--------------------|---------|-------------------|-------------------|
| MEPG-2 视频   | 4Mbps, 24          | $t$     | 792Kb             | 13.3ms            |
| NTSC 视频(压缩) | 27Mbps, 30         | $t$     | (1.5+14.5)Mb      | (72+1160t)ms      |

限于篇幅,本节内容的详细讨论可参阅我们有关论文和报告.

## 2.2 并行访问系统

DMPSS 系统的并行访问系统由共享文件系统和并行数据库系统组成见图 2. 客户机的请求通过调度系统调度后,最终由这两者访问磁盘系统而完成服务.

### 2.2.1 共享文件系统

该系统的目标存在 3 条:①实现向上兼容,增加系统的可移植性;②实现巨量媒体数据的访存,以支持会议系统的视频演播需要;③实现数据文件的联机(on-line)共享使用,以支持应用系统——多媒体合著系统 CSCW 的开发. 在最终设计时,我们通过管理“全局共享资源表”(GSRT)来实现上述目标. 而且,整个 DMPSS 的数据一致性维护也是由它完成的.

### 2.2.2 并行数据库系统

为了提高访问速度,尤其为了实现对结构复杂的“超媒体”数据的查询操作,我们引入了数据库系统. 为了实现对多访问流功能的支持,为了保证与动态优先级调度相匹配,我们采用了并行数据库服务系统,通过各个访问流上的服务线程实现并行访问.

## 2.3 同步维护和管理

从 DMPSS 结构上看,服务器的同步问题是属于源同步类问题. 由于同步问题是多媒体技术的 3 大经典问题之一,有待研究的内容很多,因此要完善地解决很难. 在现阶段我们研究的重点不在此处,对解决该类问题只采用了一些简单的方法.

### 2.3.1 保持流的连续性

这主要指保证流内媒体信息同步. 此工作是靠前面谈到的多访问流调度系统来完成.

### 2.3.2 保持流间的同步

这主要指不同媒体之间的同步. 如视频流与音频流的同步,音频流与文本的同步等.<sup>[10]</sup> 本文中,这项工作是通过向媒体流添加附加同步数据块来完成的.

## 3 相关应用和测试

为实现完整的分布式多媒体系统,课题组的其他成员从事于 2 类应用开发:①支持多点合著创作工作的多媒体合著系统;②支持多点视频会议的多媒体视频会议系统.<sup>[11]</sup> 如今,这 2 个系统通过支持服务器提供的接口协议而实现了访问操作连接. 所采用的接口协议标准可参阅系统的研究报告.

为测试用户请求的平均服务处理时间,评价系统的工作性能. 我们测试了一组符合 M/M/1/4/4 模型的访问流. 其中请求的平均发出间隔时间  $1/\lambda=12s$ ; 请求的平均服务时间  $1/\mu=15s$ . 测试结果见表 3.

在测试过程中,我们发现:当不考虑操作系统对缓冲区的管理开销,信号量的互斥开销,以及线程调度的上下文转换“context switching”开销的情况下,系统的性能上限最受流调度策略、网络协议和 I/O 调度策略 3 大因素影响.

## 4 将来的工作

为了适应将来的分布式多媒体技术发展的需要,我们计划在下一阶段增强 DMPSS 系

统的操作能力,这些目标包括完善基于系统资源秩的动态优先级调度策略;完善和加强视频服务功能,尤其是针对 VCR 控制的功能;根据连续媒体的结构信息和内在同步信息,增加对各种媒体流的流内和流间控制操作;改进磁盘访问系统,结合并行数据库系统完善多级的磁盘并行访问系统;根据对连续媒体采用的压缩标准,在存储体上实现布局优化。

表 3 系统模拟测试结果表

| 请求流       |    | S1                          | S2      | S3      | S4     | S5      |
|-----------|----|-----------------------------|---------|---------|--------|---------|
| 操作类型      |    | R(读)                        | R       | W(写)    | R      | R       |
| 客户        | 速率 | 2.0Mbps                     | 1.5Mbps | 2.0Mbps | 20Mbps | 1.5Mbps |
| QOS       | 帧数 | 20                          | 20      | 24      | 30     | 20      |
| 发生时刻(s)   |    | 0                           | 16.9    | 23.1    | 41.4   | 45.9    |
| 所需服务时间(s) |    | 22.2                        | 6.9     | 20.0    | 6.86   | 19.04   |
| 启动时刻(s)   |    | 0.005                       | 16.910  | 23.110  | 42.0   | 48.865  |
| 系统        | 理论 | 432Kb                       | 432Kb   | 980Kb   | 7540Kb | 950Kb   |
| 需求        | 实设 | 500Kb                       | 500Kb   | 1000Kb  | 7800Kb | 1000Kb  |
| 网络 QOS    |    | 实用频宽 B=20Mbps, 定长数据包 32K 发送 |         |         |        |         |

## 参考文献

- 1 吴飞等. 基于 Client/Server 结构的多媒体视频类服务器的研究与设计. 南京大学学报(自然科学), 1995, 31(10).
- 2 Hanko J G, Kuerner E M *et al.* Workstation support for time-critical applications. In: Proc. 2th International Workshop on Network and Operating Systems Support for Digital Audio and Video, Heidelberg, Germany, Nov. 1991. 4~9.
- 3 Lougher P, Shepherd D. The design of a storage server for continuous media. The Computer Journal (Special Issue on for Multimedia), Feb. 1993, 1:33~42.
- 4 Rangan P V, Vin H M *et al.* Designing an on-demand multimedia service. IEEE Communication Magazine, July 1992, 30:56~65.
- 5 Dan A, Sitaram D *et al.* Scheduling policies for an on-demand video server with batching. In: Proc. ACM Multimedia'94. ACM, Oct. 1994. 15~23.
- 6 Tierney B, Johnston W *et al.* Using high speed networks to enable distributed parallel image server systems. In Proc. Supercomputing'94, Nov. 1994. LBL~35437.
- 7 甘应爱等. 运筹学. 北京:清华大学出版社, 1990.
- 8 Peha J M, Tobagi F A. Evaluating scheduling algorithms for traffic with heterogeneous performance objectives. IEEE GLOBECOM, 1990. 21~27.
- 9 Ramakrishnan K K *et al.* Operating systems support for a video-on-dmand file service. In: Proc. 4th International Workshop on Network and Operating Systems Support for Digital Audio and Video. Lancaster, UK, Nov. 1993.
- 10 Leydekkers P, Teunissen B. Synchronization of multimedia data streams in open distributed environments. In: Proc. 2th International Workshop on Network and Operating Systems Support for Digital Audio and Video, Heidelberg, Germany, Nov. 1991.
- 11 朱海滨等. 基于 Client/Server 结构的多媒体合著及会议系统的设计. 第二届全国智能接口与应用研究学术会议, 山东, 1995.

## A MULTIMEDIA SERVICE SYSTEM SUPPORTING PARALLEL STREAMS

WU Fei CHENG Fujie WANG Pu

*(Department of Computer Science National University of Defence Technology Changsha 410073)*

**Abstract** Distributed multimedia, whose aim is to support non-local accesses to multimedia information, is a new research domain. Because of the significant differences the characteristics of continuous media data from ones of conventional data, conventional file servers and associated storage systems are not well suitable for non-local multimedia applications; instead, new service systems capable of doing so. In the paper, authors introduce a new distributed multimedia parallel service system designed and built. By the concept “system resource order”, the mutual parallel operations of multi-access stream are realized. So, the main defects in the design of existing systems are well solved. Additionally, a group of video access streams is simulated in the paper, and the conclusion is that the system is well suitable for continuous media applications.

**Key words** Distributed multimedia, continuous media, video service, schedule, parallel access.

**Class number** TP391, T316