

# 并发多媒体负载访问存储系统的模型分析

李 中 王 刚 刘 璟

(南开大学 信息技术科学学院, 天津 300071)

**摘 要:** 多媒体应用要求存储系统提供满足服务质量 QoS(Quality of Service)需求的数据访问服务,反应时间超过截止期限的访问请求必须小于一定的比例.在开放环境下,并发多媒体负载中访问请求的随机性增加了实现这一目标的难度.并发多媒体应用的访问请求经汇聚后到达存储系统的过程服从泊松分布,存储系统的服务时间服从指数分布;并发多媒体负载访问存储系统的过程用 M/M/1 排队论模型描述.通过对该模型的分析推导出构成并发负载的多媒体应用数量  $N$  与负载中访问请求的响应时间分布之间的关系.仿真试验的结果证明该模型对于保证并发多媒体应用的 QoS 是非常有效的.

**关 键 词:** 多媒体负载; 存储系统; 服务质量; 排队模型

**中图分类号:** TP 302

**文献标识码:** A

**文章编号:** 1001-5965(2004)11-1043-05

## Analysis of model of concurrent multimedia workloads retrieving data form storage systems

Li Zhong Wang Gang Liu Jing

(Collage of Information Technology Science, Nankai University, Tianjin 300071, China)

**Abstract:** The multimedia applications need the guarantee of the QoS(quality of service) when retrieving data form the storage systems, the fraction of requests whose response time exceeds a specific delay limit must below a certain proportion. In the concurrent multimedia workloads, many applications access a storage system at the same time, Which makes it difficult to satisfy the QoS requirements. The aggregate requests of concurrent multimedia applications Poisson arrive the storage system, the storage system service time is exponential distribution, such stochastic process was abstracted as an M/M/1 queuing model. Based on the queuing model, the relationship between the number of concurrent multimedia applications and the statistical distribution of the requests response time was determined, the relationship can be used to guarantee the QoS of concurrent multimedia applications. The experimental results validate the novel model is accurate.

**Key words:** multimedia workload; storage system; QoS(quality of service); queuing model

随着网络技术和多媒体技术的发展,多媒体应用对存储设备上多媒体数据的访问模式发生了变化:由单一用户的本地访问发展为开放网络上大量用户的并发访问,形成了并发多媒体负载.多媒体应用不仅要求存储系统提供持续的大吞吐量数据访问服务,而且还要求所提供的服务保证特

定的服务质量需求 QoS(Quality of Service),同时多媒体系统要求存储系统支持尽可能多的并发应用,因此并发多媒体负载对存储系统产生了空前的压力<sup>[1]</sup>.

为了提供有效的多媒体服务,必须研究并发多媒体负载访问存储系统的过程,其中的关键是

收稿日期: 2004-06-25

基金项目: 国家自然科学基金资助项目(60273031); 高等学校博士点专项科研基金资助项目(200220055021)

作者简介: 李 中 (1972-), 男, 黑龙江哈尔滨人, 博士生, createform@163.com.

建立相应的访问服务模型. 目前, 大部分的相关研究侧重于单一多媒体应用负载的情况, 少数对并发多媒体负载的研究也没有考虑到多媒体应用的具体 QoS 需求. 本文根据多媒体应用的特点建立了并发多媒体负载访问存储系统的排队论模型, 该模型不仅真实地抽象了并发多媒体应用访问存储系统的过程, 而且为研究存储系统保证多媒体应用的 QoS 需求提供了有效的工具.

### 1 存储系统的访问服务时间

多媒体应用一方面要求高速、低延迟的数据访问, 另一方面要求海量的数据存储空间. 在目前的各种存储介质中, 只有磁盘能够提供性能、空间和价格之间的完美平衡. 由多个磁盘组成的磁盘阵列能够大幅度提高数据访问的吞吐量和数据的可用性, 因此磁盘阵列是多媒体数据的主流存储系统.

磁盘为一个数据访问请求提供服务的服务时间  $T_{serv}$  主要由 2 部分组成: 定位延迟和数据传输时间. 对于随机的磁盘数据访问, 它们都是随机变量<sup>[2]</sup>. 因此  $T_{serv}$  是一个随机变量, 但可以通过分析和测试统计确定它的平均值  $E(T_{serv})$ <sup>[3]</sup>.

在磁盘阵列中, 数据通过条理化技术被均匀地散布在所有的磁盘上. 当存在大量的数据访问请求, 且访问请求的到达速率足够大时, 磁盘阵列中的磁盘实现对访问请求的并行处理, 且每个磁盘处理的访问请求的数量近似相等, 此时包含  $m$  个磁盘的磁盘阵列的访问服务时间为  $T_{serv}/m$ <sup>[4]</sup>.

由于在磁盘和磁盘阵列中还存在 Cache 机制、预读功能、优化调度等影响访问服务时间的因素, 进一步增加了存储系统访问服务时间的不可预测性和随机性.

### 2 多媒体应用访问的工作模式

多媒体应用访问存储系统是一个连续的周期性过程<sup>[5]</sup>. 为满足多媒体连续播放的需求, 多媒体应用按一定的周期  $T$  向存储系统发出数据访问请求, 每次从存储设备上读取一定大小的数据块, 形成一个访问请求序列. 每个访问请求都有一个完成截止期限  $d$ , 存储系统必须保证在  $d$  到达之前, 完成该周期的访问操作, 即访问请求的响应时间小于  $d$ , 以便多媒体应用能够及时处理本周期的数据, 比如通过网络发送到客户端进行播放, 如图 1 所示.

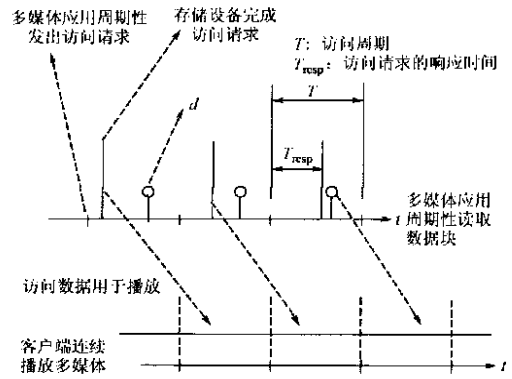


图 1 多媒体应用访问存储系统的工作模式

同时, 由于多媒体应用自身的特点 (如播放缓冲机制), 通常并不严格要求所有访问请求的响应时间都必须小于  $d$ , 允许小于一定比例上限  $P$  (如 1%) 的访问请求超过  $d$  完成. 因此, 多媒体应用对存储系统的 QoS 需求, 可以由 2 个参数确定: 访问请求的完成截止期限  $d$  和比例上限  $P$ .

多个多媒体应用同时访问存储系统上的数据, 形成并发多媒体负载. 在本文的研究侧重于并发多媒体负载中的多媒体应用为同一类型的情况, 它们具有相同的数据带宽  $B$  和 QoS 需求, 混和类型负载的情况, 可以由此推导. 在并发多媒体负载中, 相同类型的多媒体应用的访问请求序列汇聚成一个访问请求流 (stream), 属于不同应用的访问请求根据到达的时间顺序依次接受存储系统的服务. 因此并发多媒体负载可以表示为  $S(B, T, N, d, P)$ , 其中  $N$  为并发多媒体应用数量.

### 3 并发多媒体负载访问的模型

在开放环境下, 一个多媒体应用将从随机的时刻开始访问存储系统, 这一过程与电话总机接收到呼入的过程相似. 因此并发多媒体负载中的多媒体应用开始第一次访问存储系统是一个泊松过程. 虽然单一多媒体应用周期性向存储系统发出访问请求, 但当负载中并发多媒体应用的数量较多时, 可以认为访问请求是相互独立地到达存储系统的, 它们到达的时间间隔服从参数为  $\mu$  的指数分布,  $\mu$  是访问请求到达存储系统的平均速率. 对于并发多媒体负载  $S(B, T, N, d, P)$ ,

$$\mu = \frac{N}{T} \tag{1}$$

另一方面假设存储系统中磁盘的访问服务时间  $T_{serv}$  也是独立同分布的, 且服从参数为  $\mu$  的指数分布,  $\mu$  为磁盘处理访问请求的平均速率. 虽然很

多关于时间和空间局部性的研究质疑  $T_{serv}$  是否服从指数分布. 但在研究中, 通过将数据在磁盘阵列上的条纹化散布, 合理选择条纹单元的大小, 以及多媒体应用访问数据对象的随机性, 该假设是可以接受的.

$$\mu = \frac{1}{E(T_{serv})} \quad (2)$$

在对并发多媒体负载的研究中, 更值得关注的是重负载的情况, 此时  $S$  中的  $N$  数量较大时, 使  $N$  足够大, 可以将存储系统中的所有磁盘“填满”, 保证存储系统中所有的磁盘同时并发处理访问请求. 当存储系统为  $m$  个磁盘构成的磁盘阵列时, 其处理访问请求的平均速率为  $m\mu$ .

在并发多媒体负载访问存储系统的过程中,  $S$  中的访问请求以平均速率  $\lambda$  到达存储系统, 存储系统以平均速率  $m\mu$  对访问请求进行服务, 访问请求到达的时间间隔和存储系统的访问时间都是服从指数分布的随机变量. 当一个访问请求到达时, 如果存储系统空闲, 它将立刻得到服务; 否则这个访问请求将加入一个等待队列的尾部并等待, 直到得到服务为止. 上述过程显然是一个随机服务系统, 可以用  $M/M/1$  排队模型描述, 如图 2 所示.

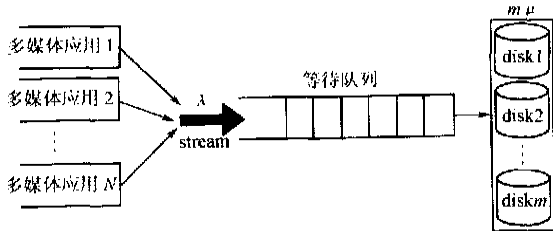


图 2 并发多媒体负载访问存储系统的模型

等待队列的最大长度由多媒体系统的数据缓冲区资源确定, 在开放系统中缓冲区容量几乎是无限的, 因此等待队列的长度可以认为是无限. 等待队列中等待服务的访问请求按照 FCFS 的规则接收服务, 从一个访问请求到达存储系统直到它完成服务离开系统之间的时间就是访问请求的响应时间; 响应时间由 2 部分组成: 在等待队列中的等待时间和实际的服务时间.

#### 4 分析并发多媒体应用的 QoS

对并发多媒体负载访问存储系统的排队模型进行分析, 可以研究存储系统如何保证并发多媒体应用的 QoS 需求.

存储系统为并发多媒体负载  $S$  中每一个应

用分配服务资源. 当  $S$  中的  $N$  较少时, 存储系统的负载较轻, 为应用分配的资源能够保证每个访问请求在  $d$  到达之前完成. 随着  $S$  中的  $N$  的增加, 存储系统的负载逐渐加重, 服务资源被逐渐消耗; 当  $N$  超过某一阈值时, 响应时间超过  $d$  的访问请求的比例将超过  $P$ ; 由于这些响应时间大于  $d$  的访问请求在  $S$  中随机分布, 因此存储系统将无法保证所有多媒体应用的 QoS 需求.

显然为了保证一个多媒体应用的 QoS 需求, 存储系统必须保证该应用的访问请求中响应时间超过  $d$  的比例小于  $P$ ; 而当并发的多媒体应用汇聚为  $S$  时, 要保证  $S$  中所有多媒体应用的 QoS 需求, 只需保证  $S$  中的所有访问请求的响应时间超过  $d$  的比例小于  $P$  即可. 根据排队理论,  $S$  中访问请求响应时间的概率分布由访问请求的到达速率

和平均服务时间  $1/m\mu$  决定, 因此对于一个  $S(B, T, N, d, P)$  只要推导出其访问请求的响应时间大于  $d$  的统计概率与并发多媒体应用数量  $N$  之间的关系, 就能通过控制  $N$  大小来保证并发多媒体应用的 QoS 需求.

图 2 中  $M/M/1$  排队模型的稳定状态概率  $P_n$  定义为排队系统中访问请求的数量恰为  $n$  的概率, 通过对稳定状态流量等式的递归计算可以求得

$$P_n = (1 - \rho)^{\rho^n} \quad n = 0, 1, 2, \dots \quad (3)$$

其中

$$\rho = \frac{\lambda}{m\mu}$$

$L^a$  为当一个访问请求到达存储系统时系统中已经存在的访问请求的数量,  $D_k$  表示其中第  $k$  个访问请求所需的服务时间. 此时正在接受服务的访问请求所需要的剩余服务时间并非完全的服务时间, 但由于服务时间为指数分布, 具有无记忆性, 剩余服务时间的分布与服务时间的分布相同, 因此  $D_k$  服从平均值为  $1/m\mu$  的独立指数分布. 访问请求的响应时间  $R$  为

$$R = \sum_{k=1}^{L^a+1} D_k \quad (4)$$

由于  $L^a$  与  $D_k$  相互独立, 采用基于  $L^a$  的条件概率, 得到访问请求的响应时间  $R$  大于  $d$  的统计概率

$$P(R > d) = P\left\{ \sum_{k=1}^{L^a+1} D_k > d \right\} = \sum_{n=1}^{L^a+1} P\left\{ \sum_{k=1}^n D_k > d \right\} P(L^a = n) \quad (5)$$

根据 PASTA 定理, 一个访问请求到达存储系

统时发现系统中恰存在  $n$  个访问请求的概率等于稳定状态概率  $P_n$

$$P(L^a = n) = P_n = (1 - \rho)^n \quad (6)$$

将(6)带入(5),并根据  $D_k$  实际是 Erlang -  $(n + 1)$  分布<sup>[6]</sup>,得到

$$P(R > d) = \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{(m\mu d)^k}{k!} e^{-m\mu d} (1 - \rho)^n =$$

$$\sum_{k=0}^{\infty} \frac{(m\mu d)^k}{k!} e^{-m\mu d} (1 - \rho)^n =$$

$$\sum_{k=0}^{\infty} \frac{(m\mu d)^k}{k!} e^{-m\mu d} = e^{-m\mu(1-\rho)d} \quad (7)$$

将式(1)~式(3)带入式(7)得到  $S$  中访问请求的响应时间大于截止期限  $d$  的统计概率与并发多媒体应用数量  $N$  的关系

$$P_N(R > d) = e^{\left(\frac{N}{T} - \frac{m}{E(T_{serv})}\right)d} \quad (8)$$

为了保证  $S$  中的多媒体应用的 QoS 需求:响应时间大于  $d$  的访问请求的比例小  $P$ , 必须保证  $S$  中的并发多媒体应用的数量  $N$  满足

$$P_N(R > d) < P \quad (9)$$

满足公式(9)的最大  $N$  就是存储系统能够支持的最大并发多媒体应用数量。

### 5 仿真试验

为了验证上述模型的精确性,采用仿真试验对模型确定的最大并发多媒体应用数量进行验证。仿真试验采用 DiskSim 3.0 仿真软件,该仿真软件是目前精度最高的存储系统仿真软件,其仿真结果具有极高的可信度<sup>[7]</sup>。

存储系统是一个由 8 个 HP\_C2490A 磁盘构成的磁盘阵列,阵列配置为 RAID0。数据通过条纹化均匀散布到每个磁盘上,并充满全部的存储空间,条纹单元的大小设置为与一次访问请求所需读取的数据量相等。

多个多媒体应用进程并发访问存储空间,形成并发多媒体负载。进程的初始访问地址在存储空间内均匀随机分布,通过对齐机制,保证每个访问请求仅从一个磁盘上读取数据。这些进程的启动时间间隔为指数分布。

试验采用了 2 种类型的多媒体应用,一类为窄带多媒体应用,数据带宽为 64 kb/s;另一类为宽带多媒体应用,数据带宽为 1024 kb/s。2 种情况下访问周期都设定为 1 s,窄带应用的每个访问请求读取 64 kb 数据,宽带应用每次读取 1024 kb 数

据。两类多媒体应用的 QoS 需求为  $d = 500$  ms,  $P = 1\%$ 。

利用公式(8)和(9),计算磁盘阵列允许的最大并发多媒体应用数量,其中  $E(T_{serv})$  由厂商提供的详细磁盘参数计算得到,并通过实际测试进行验证。表 1 为计算结果:

表 1 采用模型确定的最大并发应用数量

数据带宽/(kb s <sup>-1</sup> )	$E(T_{serv})$ /ms	基于模型的最大并发数量
64	19.550492	399
1024	49.046864	153

用表 1 中最大并发多媒体应用数量进行试验,得到访问请求的响应时间分布曲线,如图 3、图 4 所示。

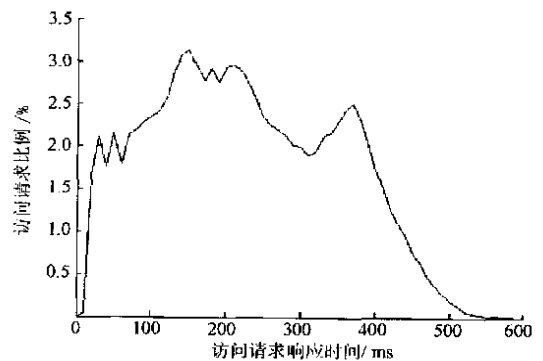


图 3  $B = 64$  kb/s,  $N = 399$  时访问请求响应时间分布曲线

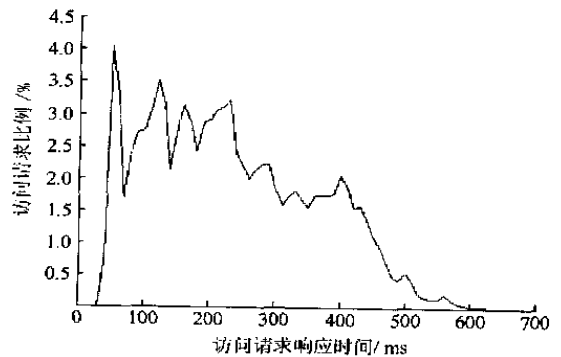


图 4  $B = 1024$  kb/s,  $N = 153$  时访问请求响应时间分布曲线

在  $B = 64$  kb/s 情况下,模型确定的最大并发多媒体应用数量为  $N = 399$ ,此时响应时间超过  $d$  的访问请求占 0.33%;  $B = 1024$  kb/s 情况下,模型确定的最大并发数量为  $N = 153$ ,响应时间超过  $d$  的访问请求占 1.69%。在进一步试验中,逐渐增加  $S$  中  $N$  的大小,得到响应时间超过  $d$  的访问请求所占比例随  $N$  变化的曲线。

从图 5 和图 6 的曲线可以发现满足多媒体应用 QoS 需求 ( $d = 500$  ms,  $P = 1\%$ ) 的实际最大并发

数量为 :  $B = 64 \text{ kb/s}$  时  $N = 401$ ,  $B = 1\,024 \text{ kb/s}$  时  $N = 151$ .

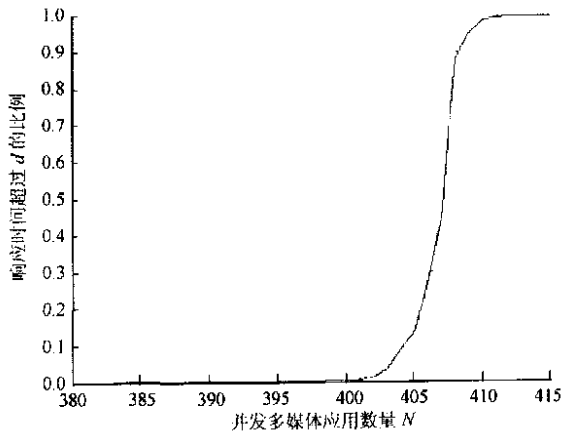


图 5  $B = 64 \text{ kb/s}$  时响应时间超过  $d$  的访问请求所占比例随  $N$  变化的曲线

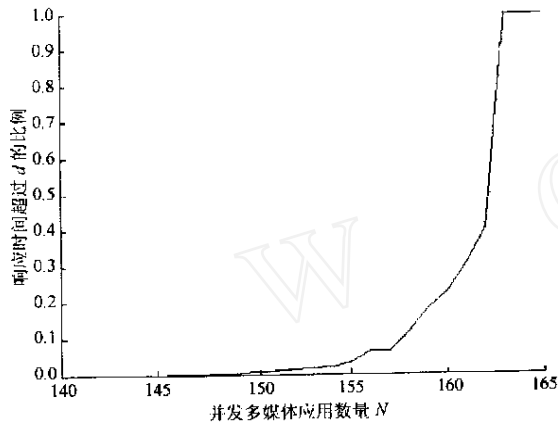


图 6  $B = 1\,024 \text{ kb/s}$  时响应时间超过  $d$  的访问请求所占比例随  $N$  变化的曲线

通过比较可以看到采用模型确定的最大并发多媒体应用数量与实际结果非常接近,该模型的精确度是值得信赖的.

## 6 结 束 语

通过对存储系统访问服务时间和多媒体应用工作模式的研究,建立了并发多媒体负载访问存储系统的分析模型.该模型不仅真实地抽象了并发多媒体应用访问存储系统的过程,而且为研究存储系统保证多媒体应用的 QoS 需求提供了有效的工具.试验结果证明该模型具有非常高的精确性.

### 参考文献 (References)

- [1] Schmitt J, Wolf L. Quality of service-an overview [R]. TR-KOM-1997-01, 1997
- [2] Ruenmmler C, Wilkes J. An introduction to disk drive modeling [J]. Computer, 1994, 27(3): 17~28
- [3] Shriver E. Performance modeling for realistic storage devices [D]. New York: Department of Computer Science, New York University, 1997
- [4] Patterson D, Gibson G, Katz R. A case for redundant arrays of inexpensive disks (RAID) [A]. In: Proc of 1988 ACM SIGMOD Int 1 Conf on Management of Data [C]. New York: ACM Press, 1988. 109~116
- [5] Leung Yiu Wing, Chan T K C. Design of an interactive video-on-demand system[J]. Multimedia, IEEE Transactions, 2003, 5(1): 130~140
- [6] Cooper R B. Introduction to queuing theory [M]. Second Edition. New York: Elsevier North Holland Inc, 1981
- [7] Bucy J S, Ganger G R. The disksim simulation environment version 3.0 reference manual [R]. CMU-CS-03-102, 2003