

文章编号:1001-9081(2010)03-0765-03

基于统计学模型的 VOD 负载均衡设计

白雪, 彭德巍

(武汉理工大学 计算机科学与技术学院, 武汉 430063)

(snowhitedelweiss@gmail.com)

摘要:现有的负载均衡算法大都基于 Web 服务,并不适用于视频点播(VOD)。而基于 VOD 的负载均衡算法没有考虑用户操作以及一个节目不同片段的访问概率差异等因素对系统负载的影响。因此,设计一种基于统计学模型的负载均衡设计方法,从用户的随机性操作出发,统计分析其规律,根据不同节目不同片段的访问概率对片源进行分块备份存储。通过仿真实验证实了算法较其他算法的优越性,提高了各节点服务器资源利用率,达到了负载均衡的目的。

关键词:负载均衡;视频点播;统计模型;用户操作

中图分类号: TP393 **文献标志码:** A

Load balancing design based on statistical model for VOD

BAI Xue, PENG De-wei

(College of Computer Science and Technology, Wuhan University of Technology, Wuhan Hubei 430063, China)

Abstract: Most load balancing algorithms are based on Web service and unfit for Video on Demand (VOD) system. They all neglect the effect of user's behavior and the differences among different segments in one program. Therefore, the load balancing algorithm based on statistical models was proposed. It commenced from random user behavior, analyzed the implied rules, divided the programs into pieces, copied and stored them, according to different order probability of diverse segments in diverse programs. The results of simulation testify its superiority to other algorithms. It improves the utilization rate of server resource and achieves load balance.

Key words: load balance; Video on Demand (VOD); statistical model; user operation

0 引言

在存储和宽带技术高速发展的今天,视频点播(Video on Demand, VOD)系统已成为炙手可热的服务,人们对 VOD 服务的需求在不断增长,使 VOD 的规模逐渐壮大。作为一个大规模 VOD 系统,其负载均衡已经成为衡量系统性能的一项重要指标,可以说体系结构及文件存储结构等设计都是在为系统的负载均衡服务。单服务器早已无法承载,然而用户需求的指数型增长速度远比服务器数量的添加速度要快,况且大规模 VOD 的服务器是一笔不菲的开销,因此从策略上优化负载均衡,必然成为首选之路,否则,再高级的服务器和宽带等硬件配备也终将成为 VOD 服务性能扩展的瓶颈。

VOD 服务不同于 Web 服务的一大特性在于,长事务的处理占有相当大的比例,因此要求 VOD 系统具有极强的负载均衡能力,才能保证 VOD 的服务质量(Quality of Service, QoS)。典型的负载均衡算法中,Round-Robin 算法和 Weighted Round-Robin 算法是静态方法,在实时服务之前就对用户请求的调度分配情况未卜先知,Source/Destination Hashing 算法是一种静态映射方法,这三种方法都没有考虑服务器的实时情况,负载均衡应是一种动态的均衡。Weighted Least Connection 算法是动态调度算法,但只靠当前连接数来衡量服务器的性能,产生的误差无法忽略。文献[1]提出一种最强能力优先调度算法(Greatest Capability First Scheduling, GCF),以节点状态、处理性能及当前用户请求数作为衡量服务器实时负载权值的标准,以片源流行度作为存储参照,但不够全面,会给负载带来

很多的不确定性。文献[2]提出一种 VOD 服务器集群文件复制与分发的算法,然而对整个片源的存储及备份粒度过大,冗余度过高,未考虑到一部片源中各片段访问力度的差异,而这种差异是造成了片源所在节点 I/O 瓶颈的主要原因之一。

本文提出了一种基于统计学模型的负载均衡设计方法,从用户的操作及需求出发,对用户并发请求的随机性进行统计分析,研究提取其规律,并对片源按照各片段的访问力度差异进行分块存储、备份及实时更新,使系统一定程度上克服网络带宽及存储服务器的 I/O 瓶颈,对多样化的用户操作及系统规模具有一定的自适应性。

1 VOD 中潜在的统计学模型

在 VOD 系统中,用户的操作是随机的,然而可以利用统计学知识对用户操作的随机性进行统计分析,建立相应的统计模型。

1.1 片源点播统计模型

用户的点播行为是一个随机事件,点播率是一个随机变量,且各用户的访问相对独立,对于精彩的节目而言,点播率相对较高。最经典的 Pareto Principle 即 80-20 规则,也就是说 80% 的点播集中于 20% 的节目。进一步的实验及研究表明,用户点播满足 Zipf-like 分布。设系统有 n 个片源,点播率按由大到小排列,则第 k 部片源的点播率 P_k 满足:

$$P_k = \frac{1}{k^{1-\theta}} \bigg/ \sum_{i=1}^n \frac{1}{i^{1-\theta}} \quad (1)$$

收稿日期:2009-09-08;修回日期:2009-10-31。

作者简介:白雪(1987-),女,黑龙江绥化人,硕士研究生,主要研究方向:并行分布式处理、高性能计算、计算机网络;彭德巍(1976-),男,湖北潜江人,副教授,博士,主要研究方向:移动计算、分布式处理、软件形式化理论。

其中修正参数 $\theta = 0.271$ 。

1.2 用户退出统计模型

用户点播一部节目后,通常顺序观看,逐渐产生对节目的认识和评价,之后决定继续看或者退出点播其他节目。文献[3]中表明,用户在一部节目的观看过程中的退出时间具有一定规律,如图1所示。

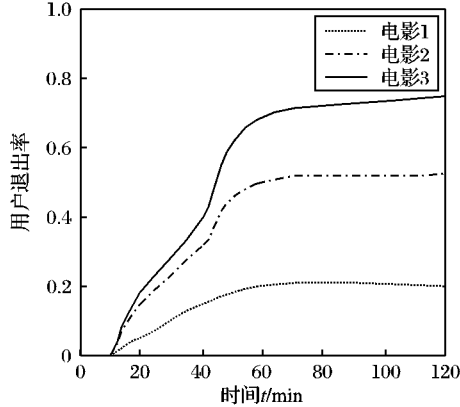


图1 用户退出概率分布

文献[4]根据统计学模型将离散函数转换为时间 t 为自变量的连续函数 $P_E(t)$:

$$P_E(t) = \begin{cases} 0, & 0 < t \leq T_b \\ K(t - T_b), & T_b < t \leq T_l \\ K(T_l - T_b), & t > T_l \end{cases} \quad (2)$$

根据以上用户退出的概率密度函数,可以得出:用户在小于 T_b 的时间内基本不会退出;在大于 T_b 小于 T_l 的时间段内,退出的用户数量逐渐增加;直至 T_l 时刻,几乎没有用户再退出。对于不同的影片, K 值则不同,在 T_l 时刻的退出率也将不同。

1.3 用户在线时长统计模型

经大量统计实验,得知用户在线时间接近于对数正态分布,即随机变量在线时长 t 的概率 $f(t; \mu, \sigma)$ 满足:

$$f(t; \mu, \sigma) = \frac{1}{t\sigma\sqrt{2\pi}} e^{-(\ln t - \mu)^2 / 2\sigma^2}$$

其中: μ, σ 是变量对数的均值与标准差。

说明如果用户在线时间已经很长,那么下一刻他继续收看的概率就越大。在统计学中, $t+1$ 时刻继续在线的概率是在已知 t 时刻在线时间很长的条件下,属条件概率分布。

2 基于统计模型的负载均衡设计方法

2.1 片源分块方法

根据用户退出的统计模型可以看出,用户对同一片源的不同片段观看的概率不同,即同一片源不同片段的点播率不同。基于此,如果存储单位为一整个片源,由于一个片源不同片段的点播率有明显差异,如图1中的电影3,从时间 T_l 至节目结束这一片段的点播率最低,而此时已有之前持续观看此节目用户的80%退出转而观看其他节目,导致负载的不均衡。因此,降低文件的存储粒度,从而减小文件的分配粒度,对片源进行分块存储,这对用户来说是完全透明的。采用文献[5]中设计的分块方法。

首先,定义 $Intensity$ 为相应片段的处理强度:

$$Intensity = \int_{t_b}^{t_e} P_k(1 - P_E(t)) dt \quad (3)$$

其中: P_k 为第 k 部片源的点播率; $P_E(t)$ 为用户在 t 时刻的退出概率; t_b, t_e 分别为相应片段的起止时间。为了使系统运行起来片段的分配粒度均衡,使各片段的处理强度相当。设一片源时长 t_d ,将片源分为 h 段的时间点为 t_1, t_2, \dots, t_{h-1} ,则应使

$$\int_0^{t_1} P_k(1 - P_E(t)) dt = \int_{t_1}^{t_2} P_k(1 - P_E(t)) dt = \dots = \int_{t_{h-1}}^{t_d} P_k(1 - P_E(t)) dt$$

显然,不是每个片源的最后一段都能偶然地和前 $(h-1)$ 个片段的处理强度相同,如果最后一段的时长小于阈值 Δ ,那么将其与倒数第二个片段合并。各片段的时长会随设置的处理强度的增大而增大,所设处理强度越大,每个片源分的段数越少。

2.2 片源片段备份存储方法

在分布式服务器集群体系结构中,中心服务器存有所有片源,是信息源。根据对用户点播特性的统计显示,用户的点播行为呈现极强的局部性。在整个大规模 VOD 系统中只存储一份显然不可能满足系统各个地区用户对它的点播请求,然而对所有节目都进行备份不切实际。所以将片源点播的分布性和存储的冗余度结合考虑,前面1.1节介绍过,80%的点播集中于20%的节目上。对于这20%的热门影片,为了加快 VOD 系统各边缘服务器用户的访问速度,提高系统的负载性能,对部分片源采取备份策略。至于对多少片源进行备份,根据 VOD 系统投入的存储服务器容量以及用户数量规模等实际情况而定。

为了提高系统负载性能同时避免造成存储空间的浪费,应使各片段的点播率相对均衡。设要备份的片源有 n 个,用 $Film = \{F_1, F_2, \dots, F_n\}$ 表示,第 i 个片源 F_i 要备份的片段有 m_i 个,用 $F_i = \{F_{i,1}, F_{i,2}, \dots, F_{i,m_i}\}$ 表示,则第 i 个片源的第 j 个片段的点播概率 $P(F_{i,j}) = P_i(1 - \int_{t_{i,j-1}}^{t_{i,j}} P_E(t) dt)$,其中 P_i 表示第 i 部节目被点播的概率, $t_{i,j-1}, t_{i,j}$ 表示第 i 个片源第 j 个片段的起止时间。设第 i 个片源第 j 个片段的备份数量为 $C_{i,j}$ (不大于存储服务器数量),为了使各片段点播相对平衡,根据统计学样本方差的原理提出如下衡量模型:

$$V = \frac{1}{sum - 1} \sum_{i=1}^n \sum_{j=1}^{m_i} \left(P(F_{i,j}) / C_{i,j} - \overline{(P/C)} \right)^2 \quad (4)$$

其中: $sum = \sum_{i=1}^n \sum_{j=1}^{m_i} C_{i,j}$ 表示所有片段备份的数量之和;

$\overline{(P/C)}$ 表示 $P(F_{i,j}) / C_{i,j}$ 的均值,即 $\frac{1}{sum} \sum_{i=1}^n \sum_{j=1}^{m_i} P(F_{i,j}) / C_{i,j}$ 。式(4)可以反应样本数据的波动性即稳定性,也就是说,反应的是各片段点播率的均衡程度,所以 V 的值越小,说明该样本各数据与其平均值相差越小,样本总体的稳定性越好,各片段的点播率与备份数量成比例,相对均衡度高。存储冗余备份的峰值则根据实际系统所拥有的存储量而定。存储策略如下:

1) 设系统有 n 台存储服务器,计算各服务器的负载值 $Load = \{L_1, L_2, \dots, L_n\}$ 并按负载值由小到大排列, $L_1 \leq L_2 \leq \dots \leq L_n$;

2) 进行初始备份,将所有片源片段按其点播率 $P(F_{i,j})$ 由大到小顺序排列,每 n 个片段为一组,每组按点播率由大到小顺序存储到负载值由小到大排列的对应的 n 个服务器中,即片段序列中的第 $(i+k \times n)$ 个存入服务器序列中的第 i 台(其中 k 是正整数);

3) 将需要备份的片段按点播率由大到小排列,用 S_i 表示第 i 台服务器,给出以下服务器序列: $S_1, S_2, \dots, S_n, S_1, S_2, \dots, S_n, \dots, S_n$, 将各备份按点播率的优先级顺序存储到各服务器节点。

2.3 片源替换方法

每隔一定时间,需对片源的存储进行调整,因为节目的点播概率是会随时间而变化的,一部节目的点播率不是固定不变的,文献[6]对节目的点播率随时间而变化的特性做了统

计分析。每个节目都有它的生命期:高峰期、中间期及低谷期。所以要对节目存储进行实时的调整替换,增加处于高峰期热播节目的备份数量,删除处于低谷期节目的冗余备份,充分利用有限的存储空间。替换策略如下:

1) 已知一部节目的第 j 个片段正在播放,那么可以根据 1.3 节的统计模型来预测用户观看第 $j+1$ 个片段的概率。设当前时刻 t 播放的片段集合用 F 表示,如果该片段不是影片的最后一块,计算用户访问下一块的概率,用集合 $P(F)$ 表示。

2) 根据片源点播统计学模型计算空闲片段的点播概率,用集合 $P(F')$ 表示。

3) 在此策略中加入用户对影片的评价值,用 e_i 表示对第 i 部片源的评价值(峰值为 E),反馈到服务器中进行统计,用 $t-1$ 表示 t 的上一个周期, o_i 表示第 i 部片源的点播数量, p 表示步骤 1) 和 2) 中计算的结果,定义式(5)作为下一周期片段的综合访问权值:

$$d_i = p \times (e_i/E) \times o_i \quad (5)$$

计算式(5),并由大到小排列, $d_1 \geq d_2 \geq \dots \geq d_m$ 。

4) 将 d_i 值未排在前 1/5 但存在冗余的节目片段删除,排在前 1/5 的节目片段按照其综合访问权值的优先级,依次在未存储该片段的服务器中优先选择负载权值小的服务器进行存储,如果该服务器的存储量达到预定阈值 Δ ,则选择负载值次小的服务器。

3 仿真测试

在仿真实验中,用一台 PC 机运行程序模拟用户的请求到达、点播及退出操作,用户请求到达率服从泊松分布^[4]。依此生成用户请求模型,用满足 Zipf-like 分布的随机访问序列对片源进行点播,各用户的在线时长服从对数正态分布,程序根据所服从的相应统计模型模拟用户操作。实验中 θ 取 0.271。根据式(3)将片源分段,按照式(4)进行备份存储,负载均衡程序会根据本文中的替换策略进行定期的热门片源替换。而且在系统运行过程中,负载程序会定时检查各节点服务器的可达性,避免将请求分配给宕掉的节点,若某节点不可达,则将其负载权值记为 0,以供维护。经过对本设计方法的实验与其他方法进行比较,结果如图 2~3 所示。

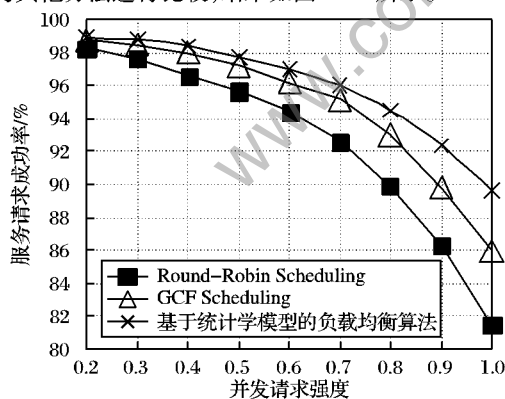


图2 三种算法的服务成功率比较

图2 描述了几种不同的负载均衡算法下,系统处理用户请求的成功率,所谓相对应的失败可能是由于用户等待时间过长或网络问题导致该请求服务失败。实验数据显示,基于统计学模型的负载均衡方法处理用户请求的能力较其他几种方法明显较好,即使用户请求数量过多时,系统仍能保持一定的成功率,增强了系统的可用性和健壮性,为系统扩展规模打下基础。

图3 描述了几种算法下系统的响应时间随用户并发强度的变化情况。算法一开始响应时间相差无几,随着并发用户数量的增加,开始体现出基于统计学模型的负载均衡算法的

优越性。GCF 虽然也是动态算法,然而只考虑节点服务器的实时负载权值的大小,没有考虑同一片源不同片段访问力度的差异性,也没有进行实时替换更新,响应速度不如本文算法快。基于统计学模型的算法可以在并发请求强度很大的情况下依然保持可接受的响应速度。

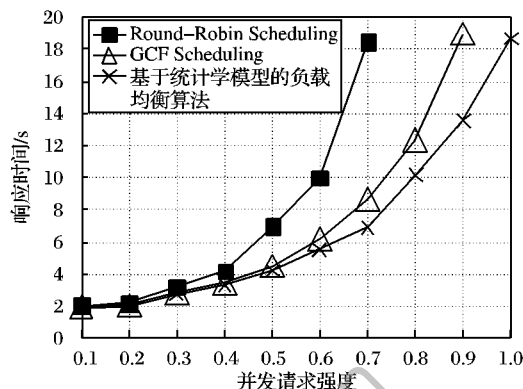


图3 三种算法的响应时间比较

4 结语

本文提出了基于统计学模型的负载均衡方法,从用户的操作规律着手,用户操作具有典型的随机特性。文献[4]表明暂停、快进、快退等 VCR 操作在点播观看过程中属于小概率事件,因此它对 VOD 系统的影响可以忽略。对用户的点播、退出操作和在线时长进行统计分析,提取统计模型,在此基础上对片源进行分块存储,根据各片段的点播率计算各片段合适的备份数量,以服务器各方面的综合权值作为衡量服务器性能差异的标准,将备份率高的片段存储在负载小的节点服务器中,并定期对存储节点中的节目根据点播率高低及用户兴趣进行预测更新,从而实现节点间的动态负载均衡。经过对算法进行仿真比较之后,证实本文提出的算法具有一定的优越性,其处理请求的成功率较其他算法高,即使在用户请求数量过大时仍能保持一定的响应速度和成功率,从而提高系统的稳定性及健壮性,并易于扩展,具有一定的自适应性。然而,替换方法需要实时获得用户在线时长信息,这势必会增大系统的管理开销,需要在性能的提升与开销的增加之间寻求一种折中方案。

参考文献:

- [1] 王永亮, 刘峰, 张春. VOD 系统的负载均衡存储策略及调度算法研究[J]. 电视技术, 2004(11): 40-42.
- [2] 李旸, 吴新亚. VOD 集群中基于 Zipf 定律的负载均衡算法[J]. 计算机工程与设计, 2009, 30(12): 2945-2947.
- [3] 张永坤, 金海. 一种基于进程剩余运行时间总和的集群动态负载均衡算法[J]. 计算机工程与科学, 2005, 27(5): 63-65.
- [4] 舒万能, 郑世珏. 一种基于线性变换遗传算法的 VOD 集群负载均衡方法[J]. 武汉大学学报, 2006, 31(9): 839-841.
- [5] 林光国, 戴琼海, 丁嵘. 基于用户行为统计的流媒体集群负载均衡算法[J]. 清华大学学报, 2005, 45(4): 525-528.
- [6] 李勇, 陈福接. 大规模 VOD 及其 Cache 机制[J]. 计算机工程与科学, 1999, 21(5): 52-55.
- [7] 胡涛. 大规模 VOD 系统及其负载均衡设计[D]. 合肥: 中国科学技术大学, 2008.
- [8] 罗建光, 赵黎, 杨士强. 基于用户行为分析的应用层组播树生成算法[J]. 计算机研究与发展, 2006, 43(9): 1557-1563.
- [9] 胡涛, 许胤龙. 分布存储 VOD 系统的负载均衡设计及其仿真[J]. 计算机仿真, 2009, 26(4): 186-190.
- [10] 李旸, 吴新亚. VOD 集群中基于 Zipf 定律的负载均衡算法[J]. 计算机工程与设计, 2009, 30(12): 2945-2947.