

## 移动网格中负载自平衡的集群式边缘流媒体服务器

柴云鹏 顾 雷 李三立

(清华大学计算机科学与技术系 北京 100084)

(chaiyunpeng@china.com)

### Cluster-Based Edge Streaming Server with Adaptive Load Balance in Mobile Grid

Chai Yunpeng, Gu Lei, and Li Sanli

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

**Abstract** With the fast development of 3G network, and the gradual popularity of Wi-Fi which can cover the whole city area, the requirement of multimedia service in mobile grid has been growing significantly in recent years. Because the streaming media service on the Internet is the most important supplier for mobile multimedia users in mobile grid, the edge streaming servers which locate between the boundary of wireless network and the Internet, can act as bridges and buffers to achieve the notable effect on reducing the load of the Internet and improving the QOS of mobile multimedia service. Accordingly, a new design of cluster-based edge streaming server (CESS) is introduced; in the meanwhile, aiming at solving the most important problems of CESS, namely loading balance, some analysis and experiments are carried out. On the basis of such work, a new kind of caching replacement algorithm called MCLBS is brought out to make CESS an adaptive loading balance system. Finally, the experiment results and related analysis show that in comparison with the traditional cache replacement algorithm, the MCLBS algorithm which is more suitable for cluster server architecture makes the cache-hit-rate increase significantly and greatly reduces remote server's bandwidth requirements in the same environment.

**Key words** streaming media; mobile grid; cluster; load balance; cache replacement

**摘 要** 迅速发展的 3G 网络和覆盖范围可达整个城域的 Wi-Fi 网络的逐渐普及,使移动网络环境中对流媒体服务的需求迅速增长. 由于移动用户主要依赖 Internet 上的流媒体服务资源提供服务,所以位于无线网络和 Internet 交界处的边缘流媒体服务器可以充当桥梁和缓冲区,对于降低网络负载和提高服务质量有显著的效果,因而提出一种新的集群架构的边缘流媒体服务器(CESS)的设计,而且针对集群服务器中最重要的负载均衡问题做了分析和测试,提出了一种新的 MCLBS 缓存替换算法来使 CESS 能够达到自适应的负载均衡. 最后,实验测试和结果分析证明,相对于传统的缓冲替换算法, MCLBS 算法更适合集群式的服务器体系结构,在同样的条件下,缓存命中率明显提高,大大降低远程服务器的带宽.

**关键词** 流媒体;移动网格;集群;负载平衡;缓存替换

中图法分类号 TP302;TP37

迅速发展的 3G<sup>[1-2]</sup>和覆盖范围可达整个城域的 Wi-Fi<sup>[3]</sup>使移动网络环境中对流媒体服务的需求迅速增长,流媒体服务已经成为移动网格研究中目前

以及今后一段时间内的重要课题. 由于移动终端设备受到体积、机器性能以及耗电等问题的限制,其提供服务能力有限,所以移动用户获取流媒体服务主

要通过访问 Internet 中的多媒体资源,这需要在无线网络和 Internet 的交界边缘上建立有效的传输机制和服务。在 Internet 上,位于网络边缘的代理服务器被证明可有效降低 Internet 网络流量并减轻网络负载,因为用户的点播规律满足 Zipf 分布<sup>[4]</sup>,热点现象非常突出,所以在网络上距离终端用户很近的代理服务器的缓存能满足大多数用户的请求,有效地降低远程流媒体服务器的负载,并提高服务质量。

移动网格所基于的无线网络一般都采用蜂窝结构,即整个蜂窝无线系统由若干无线蜂窝组成,如图 1 所示。一般每个蜂窝跨越 10~15km。无线蜂窝的中心是一个无线基站,每个无线蜂窝范围内的用户利用无线信道与基站通信,并通过基站连接远程的 Internet 服务资源,一般在基站内通过设置代理服务器来提高整个系统的性能<sup>[5]</sup>。但是每个代理服务器只负责一个蜂窝内的用户,而不同蜂窝区域内同时使用流媒体服务的用户数不同,导致不同的代理服务器的负载可能有巨大的差异,一方面,负载重的代理服务器不能满足蜂窝区域内所有用户的请求;另一方面,负载轻的代理服务器却有大量的服务资源处于闲置状态,造成巨大的浪费。更重要的是,蜂窝区域的负载水平并不是固定的,对于同一个蜂窝区域,一段时间负载非常重,而另一段时间负载又很轻,比如写字楼集中的区域可能工作时段负载非常重,而休息时段负载非常轻。

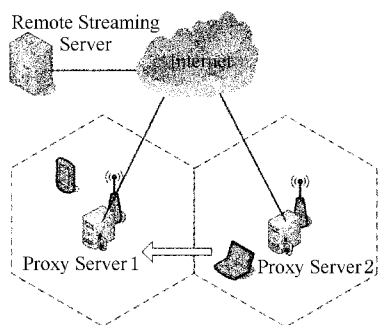


Fig. 1 The bee hive architecture in mobile grid.

图 1 移动网络中的蜂窝结构

另外,移动网格用户会在不同的蜂窝之间穿梭移动,经过蜂窝边界时,涉及到负责该用户的代理服务器的转移。这时新的代理服务器需要重新从远程代理服务载入该用户需要的媒体内容,但是临时载入需要花费较长时间,不能保证用户连续观看流媒体内容的质量,这就是“移交”问题,是移动网格流媒体应用的一个重要问题。

测试数据表明,用户的点播规律基本满足 Zipf 分布,所以不同的蜂窝区域内用户点播的热点相差不多,每个独立的流媒体服务器缓冲的内容中有很多是相同的,从整体来看,对于代理服务器的存储资源是一种浪费,如果代理服务器之间能够协作,同样的资源就能覆盖更多的用户请求,大大提高用户请求的缓存命中率。

独立的代理服务器系统存在上述的 3 个问题。而集群结构是目前大型服务器领域的主流,而且近年来也有一些对集群式的代理服务器进行的研究<sup>[6-7]</sup>,但一般是从利用集群扩充独立代理服务器能力的角度出发。本文提出一种新的移动网格环境中的集群式边缘流媒体服务器(cluster-based edge streaming server,CESS)的架构,位于传统 Internet 和无线网络的交界边缘,可以为很多蜂窝区域内的移动用户提供流媒体服务,也同时在多个蜂窝区域之间形成缓存资源的共享,而且提出了 CESS 中的一种 MCLBS(multilevel cache and load balance supported)缓存替换算法,来使 CESS 达到自适应的负载均衡。

## 1 CESS 的体系结构

### 1.1 CESS 的外部网络环境

在 Internet 上提供流媒体服务的集群式服务器分为两类:一是集群架构核心流媒体服务器(cluster-based core streaming server,CCSS),是提供流媒体服务的核心服务器,储存着所有的流媒体资源,但是在网络上距离绝大多数终端用户较远,也称为远程流媒体服务器;第 2 类是集群架构边缘流媒体服务器(cluster-based edge streaming server,CESS),它一般位于网络边缘,覆盖一定范围的区域,距离终端用户较近,提供缓冲代理服务,加快用户访问流媒体资源。

CESS 设计的目的是使一个服务能力强大的集群式服务器作为一个比较大的区域的边缘服务器,其距离终端的用户即无线蜂窝基站的无线网关的网络距离只有一跳或者很少的几跳,它们之间的访问速度非常快,可以忽略它们在网络传输上的延迟。所以 CESS 在网络中的布置如图 2 所示,从覆盖若干个无线蜂窝区域代理服务器的路由器上架专线到 CESS 上。这样的路由器级别越高覆盖的无线蜂窝区域范围越大,但是同时付出的代价是终端到 CESS 的网上距离会略微增加;如果路由器的级别比较低,则覆盖的蜂窝范围较小,但是终端到 CESS 的延迟就会很小。可以根据实际各个蜂窝的平均负

载情况,以及实际的需求来进行调节 CESS 覆盖的区域范围。

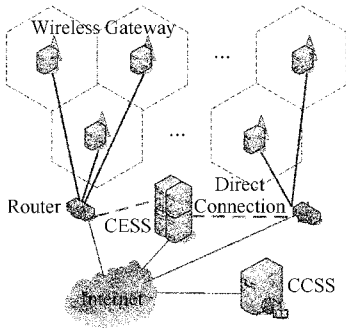


Fig. 2 Network environment of CESS.

图2 CESS 外部网络环境

### 1.2 CESS 的内部结构及工作流程

如图3所示,CESS内部的结点可以分为3类:管理结点、代理结点和资源结点。管理结点是对其其他各个结点进行管理以及使其他结点协作的结点;代理结点有外部网络连接,直接连接外部路由器引进的专线,负责一定区域的流媒体代理服务;资源结点为了维护各个代理结点的负载均衡,通过提供自己的存储资源,以及从远程服务器处进行媒体预取工作来调节各个代理结点的负载,使其趋于平衡。代理结点以及资源结点的数量和配置都可以灵活配置,来适应不同的实际需求。这3类结点通过内部高速网络连接,可以在此基础上,实现彼此之间的缓存资源共享以及协作。当代理结点收到用户的媒体请求时,依次查询如下位置是否有请求部分数据的缓存:1)代理结点本地内存缓存;2)代理结点本地磁盘缓存;3)资源结点和其他负载低于一定阈值的代理结点的缓存。如果以上都不命中,则访问远程的CCSS获取需要的数据。

总之,作为集群式服务器,CESS中负责不同区域的代理结点和资源结点之间的缓存是通过内部高速互连网络共享的,这样可以避免不同的结点都去

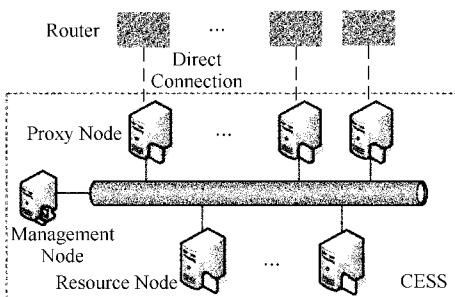


Fig. 3 CESS architecture.

图3 CESS 的体系结构

缓存同样的热门节目浪费存储资源,而且可以作出全局的缓存的优化分布,提高CESS整体的缓存效率。另外,移动用户的“移交”问题在CESS中也能很好地解决,由于代理结点之间的缓存共享,用户的移交可以做到无缝切换,原来的代理结点进行媒体预取得到的内容也并不浪费,可以为新的代理结点所利用。

## 2 CESS 的负载均衡问题

负载均衡一直是网格系统中的重要问题,在多个服务提供点之间达到负载的平衡可以大大提高整个系统的效率。同样,负载均衡在CESS中也很重要。但是在移动网格环境中,移动用户是各个区域之间自由移动,每个无线区域内的实际负载是动态变化的,不同区域之间可能相差悬殊。而CESS中每个代理结点负责固定范围的无线区域,所以在CESS体系结构的限制下,分配新的用户请求给低负载结点这种集群系统中传统的负载均衡方案并不适用,CESS只有采用通过内部资源优化配置来进行自适应负载均衡的解决方案。

### 2.1 流媒体代理服务器负载分析

不同的应用方向对于负载的定义实际上各不相同,往往是应用中的瓶颈所在,比如传统的计算型集群系统的负载就基本上等价于CPU的利用率。但是流媒体服务集群系统中,系统的瓶颈在于存储访问和网络传输的延迟,每次用户请求的处理时间绝大多数都被上述的延迟占用。所以在CESS系统中,我们可以用CESS单位时间内所有用户请求的实际处理时间总和 $T_{total}$ 作为反映结点负载的指标。而缓存的不同位置需要的处理时间是不同的,所以结点负载可以用式(1)表示:

$$L_{node} \propto T_{total} = \sum_{all\_reqs} T_p = \sum_{lmem\_hit} T_{lm} + \sum_{ldisk\_hit} T_{ld} + \sum_{nmem\_hit} T_{nm} + \sum_{ndisk\_hit} T_{nd} + \sum_{remote} T_r, \quad (1)$$

其中, $T_{lm}$ 是访问一次本地内存中缓存的处理时间, $T_{ld}$ 是访问一次本地磁盘中缓存的处理时间, $T_{nm}$ 是通过内部高速网络访问一次其他结点内存中缓存的处理时间, $T_{nd}$ 是通过内部高速网络访问一次其他结点磁盘中缓存的处理时间, $T_r$ 是访问一次远程服务器获取相关数据的处理时间。

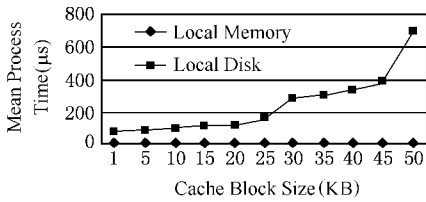
为了对结点的负载进行分析,需要了解式(1)中各部分的处理时间 $T_{lm}$ , $T_{nm}$ , $T_{ld}$ , $T_{nd}$ 的具体关系,

所以我们构建了如下的测试环境:8 结点集群服务器,结点具体参数如表 1 所示.实验主要测试以上 4 种不同的方式,读取不同大小数据块时实际的处理时间.为了保证数据尽量准确,所有的数据均为 1000 次测量数据的平均值.实验数据如图 4 所示.

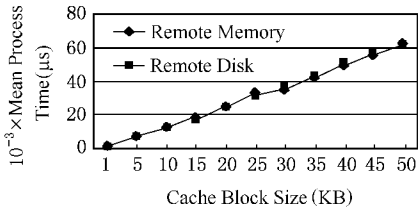
Table 1 Parameters of Cluster Nodes

表 1 测试集群结点参数列表

Parameter	Value
Processor	AMD Athlon 1900 + × 2
Memory(MB)	512
Disk	60GB, IDE, 7200RPM
Inner Network(Mbps)	100



(a)



(b)

Fig. 4 Processing time experiment result. (a) Local memory and Disk and (b) Remote memory and disk.

图 4 处理时间测试结果. (a) 本地内存和磁盘 (b) 远程内存和磁盘

## 2.2 流媒体代理服务器负载均衡解决方案

实验数据表明,  $T_{lm} < T_{ld} \ll T_{nm} \approx T_{nd}$ , 其中  $T_{nm}$  和  $T_{nd}$  相差无几. 而  $T_r$  的值在不同的实际网络情况下差距很大, 但一般都比以上 4 个大很多. 所以我们可以把 CESS 中缓存的存储资源进行分级: 本地内存缓存、本地磁盘缓存和其他结点缓存依次分为高中低级, 分别用  $A, B, C$  表示. 由式(1)可知, 提高命中率、命中高级别的缓存有助于缩短结点对用户的处理时间, 从而降低结点的实际负载.

所以, 解决负载均衡的方案有两个方面: 一是缓存在 CESS 各个结点中的分布, 要偏向对负载重的结点进行优化的方向来分布, 这主要靠缓存替换算法来具体实现; 二是进行媒体资源预取, 在资源紧张的情况下, 优先响应负载重的结点的预取请求, 通过这两种方法做到 CESS 的负载均衡的自适应. 媒体

资源预取并不是本文讨论的重点. 本文下面重点介绍缓存替换算法的设计和实现.

## 3 CESS 的缓存算法

### 3.1 缓存增益模型

相对于用户直接连接远程服务器获取流媒体服务, CESS 的存在具有很多优点, 其中最主要的是降低用户启动延迟和节约远程服务器的带宽资源. 而且在 CESS 内部合理的缓存策略还能平衡 CESS 内各代理结点的负载, 打破瓶颈, 提高整体性能. 所以, 通过建立缓存增益模型<sup>[8]</sup>可以量化以上指标, 准确判断缓存操作对系统整体性能带来的提高.

降低启动延迟增益  $G_{Delay}(C_i)$  如式(2)所示. 用户到 CESS 的网络距离很近, 可以忽略不计, 式中  $D(C_i)$  表示 CESS 到远程服务器的网络延迟;  $L_j$  表示代理结点  $j$  的负载, 负载重的结点的  $L_j$  值较大;  $f_j(C_i)$  表示代理结点  $j$  在过去一段时间内对数据块  $C_i$  的访问次数, 数据块  $C_i$  包括缓存数据块, 也包括 CESS 中得到但是没有进行缓存的数据块;  $Rank(C_i)$  是根据前面提出的缓存分级而设置的权值,  $Rank_A(C_i) > Rank_B(C_i) > Rank_C(C_i)$ ;  $f_j(C_i) \times L_j$  是根据代理结点负载加权过的访问频率, 负载重的结点的权重增加, 这样负载重结点所需的缓存数据块的增益值会相对增大, 这样缓存替换和缓存调度时会作出优先于负载重结点的决策, 有助于负载均衡;  $f_j(C_i) \times Rank(C_i)$  是根据缓存分级加权过的访问频率, 处于高级别的缓存数据块为提高系统性能贡献更大, 所以这样可以提高其增益, 前缀缓存是媒体文件的开头部分, 用户连接上之后首先必须获取的数据, 只有前缀缓存对减小启动延迟有作用.

节约远程服务器带宽增益  $G_{Network}(C_i)$  如式(3)所示. 其中  $S(C_i)$  表示缓存数据块  $C_i$  的大小:

$$G_{Delay}(C_i) = \begin{cases} D(C_i) \times \sum_{j=1}^m [L_j \times f_j(C_i) \times Rank(C_i)], & C_i \in Prefix\ Cache, \\ 0, & C_i \notin Prefix\ Cache. \end{cases} \quad (2)$$

$$G_{Network}(C_i) = D(C_i) \times S(C_i) \times$$

$$\sum_{j=1}^m [L_j \times f_j(C_i) \times Rank(C_i)]. \quad (3)$$

综合以上两种增益的定义, 缓存数据块  $C_i$  在 CESS 中总的系统增益如式(4)所示,  $\lambda_D$  和  $\lambda_N$  分别表示二者的权重, 可以根据不同的实际情况配置:

$$G(C_i) = \lambda_D G_{\text{Delay}}(C_i) \times \lambda_N G_{\text{Network}}(C_i). \quad (4)$$

缓存策略的目的实际上是实时地形成缓存数据在

存储系统上的一个最优分布,即求  $\text{Max} \sum_{i=0}^n G(C_i)$ :

$$\text{s.t.} \quad \sum_{i=0}^n S(C_i) \leq S_{\text{All}}, \quad (5)$$

其中  $S_{\text{All}}$  是 CESS 所有缓存空间的大小. 实际上, 上述问题可以转化为背包问题, 也就是 NP 完全问题<sup>[7]</sup>. 所以该问题不可能求出最优解, 我们希望能够通过设计合适的算法得到次优解. 定义单位存储空间带来的增益为增益密度, 表示为  $\rho(C_i) = G(C_i)/S(C_i)$ , 优先缓存增益密度大的数据块能取得更接近式(5)的结果<sup>[5]</sup>. 基于这个思路, 我们设计了下面的基于缓存分级的负载平衡缓存算法, 可以简称为 MCLBS (multilevel cache and load balance supported) 缓存算法.

### 3.2 MCLBS 缓存替换算法

MCLBS 缓存替换算法是实时性的. 每次缓存替换并不是寻找整个 CESS 缓存系统中增益密度最小的数据块直接进行替换, 而是选择一个能使增益增量达到最大值的位置进行替换, 如果不能使增益增量大于 0, 则不发生替换. 其算法描述如下所示.

*MCLBS\_Cache\_Replacement* ( ):

$\text{Set}_1 = \{ \{ \rho_{A1}, \rho_{A2}, \dots, \rho_{Am} \}, \{ \rho_{B1}, \rho_{B2}, \dots, \rho_{Bm} \}, \rho_C \}$  = 各个代理节点的 A, B 两级缓存, 以及所有资源节点的 C 级缓存中各自最小的增益密度;

$\text{Set}_2 = \{ \{ \rho_{\text{New}_A1}, \rho_{\text{New}_A2}, \dots, \rho_{\text{New}_Am} \}, \{ \rho_{\text{New}_B1}, \rho_{\text{New}_B2}, \dots, \rho_{\text{New}_Bm} \}, \rho_{\text{New}_C} \}$  = 假设分别用新数据块  $C_{\text{New}}$  替换以上各个数据块时的增益密度;

while(1):

( $\rho_{\text{Max}}, \text{Pos}$ ) =  $\text{Set}_2$  中最大的增益密度及对应位置

if ( $\rho_{\text{Max}} > \rho_{\text{Pos}}$  in  $\text{Set}_1$ ):

if ( $\rho_{\text{Pos}}$  in  $\text{Set}_1$  不是  $\text{Set}_1$  中最小的):

$\text{Pos}2 = \text{Set}_1$  中增益密度最小的缓存块的位置;

缓存替换:  $C_{\text{Pos}} \rightarrow C_{\text{Pos}2}$ ;

缓存替换:  $C_{\text{New}} \rightarrow C_{\text{Pos}}$ ;

break;

else

从  $\text{Set}_2$  中删除  $\rho_{\text{Max}}$ ;

## 4 实验结果及分析

### 4.1 实验平台

本实验仍然采用前面所述的 8 结点集群系统作为实验平台, 其中 3 个结点用来模拟无线网关, 3 个结点用来当做 CESS 代理结点, 2 个结点当做 CESS 资源结点, 结构如图 5 所示. 每个结点模拟无线网关与一个 CESS 代理结点直接相连, 向其发送模拟的用户请求. 用户请求媒体资源时遵循 Zipf 分布; 同时为了观察算法在负载不均衡的情况下的表现, 3 个模拟无线网关的同时在线人数比例基本上满足 4:2:1; 另外, 为了使模拟更接近真实情况, 已进入的用户有一定的概率离开, 而且离开的概率与已观看时间成反比.

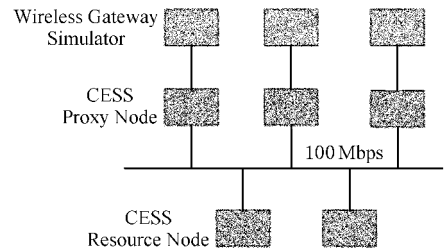


Fig. 5 Experiment platform of CESS.

图 5 CESS 实验平台

### 4.2 实验结果及分析

基于上述实验平台, 我们分别用 LFU<sup>[9]</sup> 和 MCLBS 两种缓存替换算法进行对比实验. 实验中, 我们选择 50KB 作为缓存块的大小. 测量当 CESS 系统总缓存空间与影片库总空间为不同比值情况下的缓存命中率、远程服务器带宽占用、CESS 各代理结点的负载的平均值和方差. 缓存命中率是反映缓存系统效率的常规而有效的测量标准; 远程服务器带宽占用也能反映算法对于节约 Internet 带宽以及缓存的效率所起的作用; 实验中的负载指的是结点单位时间对所有用户请求的处理时间, 负载平均值能反映算法对降低服务器工作压力方面的作用; 而从结点负载方差能比较科学地了解整个系统的负载均衡情况. 为了使算法具有可比性, 在 LFU 算法中, 图 7 中所示的 CESS 资源结点的缓存资源平均分配给 3 个 CESS 代理结点. 实验结果如图 6~9 所示.

从实验数据中可以看出, MCLBS 缓存替换算法使 CESS 充分发挥了集群系统中结点间的高速互连网络和存储资源共享的优点, 在同样的条件下, 缓

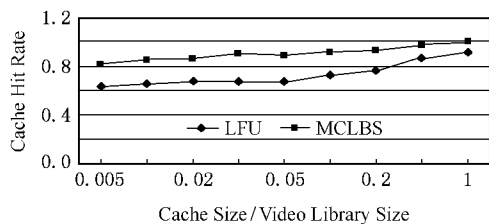


Fig. 6 Comparison of cache hit rate.

图 6 缓存命中率对比

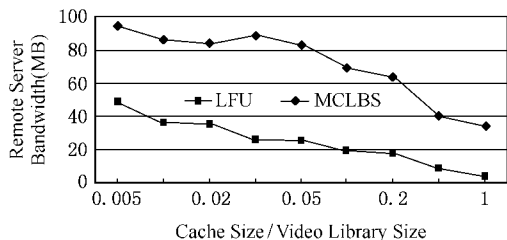


Fig. 7 Comparison of remote server bandwidth.

图 7 远程服务器带宽占用对比

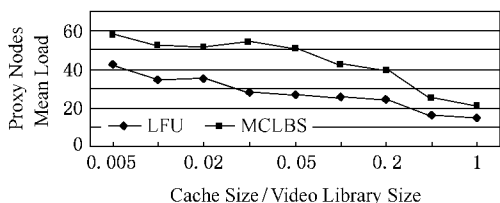


Fig. 8 Comparison of mean load of proxy nodes.

图 8 代理结点平均负载对比

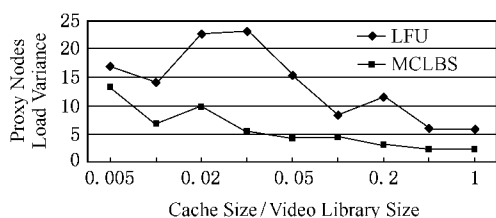


Fig. 9 Comparison of variance of proxy nodes.

图 9 代理结点负载方差对比

存命中率明显提高,远程服务器带宽大大降低。同时由于对远程服务器访问减少,缓存命中率的提高,使得结点的平均负载也显著降低。同时,除去随机产生的用户请求和模拟的用户中途退出造成的不稳定因素,结点负载的方差整体上是随着缓存总空间/影片库空间的增大而呈下降趋势的,而且 MCLBS 比 LFU 也有明显的下降,这是因为一方面负载整体水平的降低导致方差的减小;更重要的是 MCLBS 算法中考虑结点负载的增益模型使高负载结点要求的缓存块以更大的概率留下,而不是被替换掉,所以

在缓存有限的情况下,优先保证高负载结点的缓存命中率,从而降低其负载,各结点负载趋于平衡。

但是实验中也发现 MCLBS 的问题:由于缓存增益模型的计算远比 LFU 等算法复杂,所以计算上占用的 CPU 更多,为了保证系统的实时性,在硬件上对 CPU 的要求比较高。

## 5 结束语

本文提出了一种新的在移动网格环境中的集群式边缘流媒体服务器 CESS 的设计,并在其基础上,通过 MCLBS 缓存替换算法使其做到自适应性的负载均衡。CESS 具有如下优点:①负载均衡,解决了负载动态变化的问题;②不同区域的代理之间形成协作,通过共享有效充分利用资源;③很好地解决“移交”问题,在集群内部调度解决,不存在额外的开销;④扩展性强,可根据实际需求灵活调整规模。

## 参 考 文 献

- [1] P Haavisto, R Castagno, H Honko. Multimedia standardization for 3G system [C]. The IEEE 5th Int'l Conf on Signal Processing Proceedings, Beijing, 2000
- [2] Minoru Etoh, Takeshi Yoshimura. Wireless video applications in 3G and beyond [J]. IEEE Wireless Communications, 2005, 12(4): 66-72
- [3] Erina Ferro, Francesco Potorti. Bluetooth and Wi-Fi wireless protocols: A survey and a comparison [J]. IEEE Wireless Communications, 2005, 12(1): 12-26
- [4] Cai Qingsong, Li Zimu, Hu Jianping. Study on characterizing streaming media and user access to media on the Internet [J]. Journal of Beijing University of Aeronautics and Astronautics, 2005, 31(1): 25-30 (in Chinese)  
(蔡青松, 李子木, 胡建平. Internet 上的流媒体特性及用户访问行为研究 [J]. 北京航空航天大学学报, 2005, 31(1): 25-30)
- [5] Zhong Yuzhuo, Xiang Zhe, Shen Hong. Streaming Media and Video Server [M]. Beijing: Tsinghua University Press, 2003 (in Chinese)  
(钟玉琢, 向哲, 沈洪. 流媒体和视频服务器 [M]. 北京: 清华大学出版社, 2003)
- [6] Song Wu, Hai Jin, Jie Chu, et al. A novel cache scheme for cluster-based streaming proxy server [C]. The 25th IEEE Int'l Conf on Distributed Computing Systems, Ohio, USA, 2005
- [7] Yang Guo, Zihui Ge, Bhuvan Urganekar, et al. Dynamic cache reconfiguration strategies for a cluster-based streaming proxy [J]. Computer Communications, 2006, 29(10): 1710-1721

- [8] Qian Zhang, Zhe Xiang, Wenwu Zhu, *et al.* Cost-based cache replacement and server selection for multimedia proxy across wireless internet[J]. IEEE Trans on Multimedia, 2004, 6(4): 587-598
- [9] E J O'Neil, P E O'Neil, G Weikum. The LRU-k page replacement algorithm for database disk buffering[J]. ACM SIGMOD Record, 1993, 22(2): 297-306



**Chai Yunpeng**, born in 1983. Received his B. A. 's degree in computer science and technology from Tsinghua University, Beijing, China, in 2004. Since 2004, he has been Ph. D. candidate in computer architecture in Tsinghua University. His current research interests include simulation, performance analysis and performance improvement of clustered streaming server.

柴云鹏, 1983年生, 博士研究生, 主要研究方向为集群式流媒体服务器的仿真、性能分析和性能优化。



**Gu Lei**, born in 1979. He has been Ph. D. candidate in Tsinghua University since 2002. His main research interests include resource management in grid computing.

顾雷, 1979年生, 博士研究生, 主要研究方向为网格计算中的资源管理。



**Li Sanli**, born in 1935. Professor Ph. D. supervision at Tsinghua University and member of the Chinese Academy of Engineering. His main research interests include grid computing and micro-architecture.

李三立, 1935年生, 教授, 博士生导师, 中国工程院院士, 主要研究方向为网格计算和微体系结构。

## Research Background

Cache server between the Internet and mobile users plays an important role in reducing the servers' network transmission load on the Internet and improving the QOS of mobile service, especially for multimedia application which consumes large amounts of network bandwidth. With the fast evolution of 3G and Wi-Fi network, which can cover city area, the requirement of multimedia service in mobile grid grows significantly in recent years. But these large numbers of independent caching servers can not cooperate with each other in the current solutions, resulting in low utilization of resources. In this paper, a new design of cluster-based edge streaming server (CESS) is introduced, making adjacent cache servers organized as a cluster so as to improve the resource utilization. In the meanwhile, aiming at balancing the load of CESS, some analysis and experiments have been carried out, and a new kind of caching replacement algorithm called MCLBS, which is shown to be a more effective cache replacement algorithm than the traditional ones in the cluster environments, is put forward to make CESS an adaptive load balance system. Finally, the simulation experiment results show that, compared with the traditional cache replacement algorithm, MCLBS is more suitable for cluster server architecture, making the cache-hit-rate increased significantly and the remote server's bandwidth requirements reduced greatly in the same environment.