

# 流媒体分发系统关键技术综述

杨 戈<sup>1,2</sup>, 廖建新<sup>1</sup>, 朱晓民<sup>1</sup>, 樊秀梅<sup>3</sup>

(1 北京邮电大学网络与交换技术国家重点实验室, 北京 100876;

2 北京大学深圳研究生院集成微系统科学与工程与应用重点实验室, 广东深圳 518055;

3 北京理工大学计算机科学技术学院智能信息技术北京市重点实验室, 北京 100081)

**摘 要:** 流媒体将是未来通信中的杀手业务. 本文讨论了流媒体分发系统的关键技术, 阐述了基于 CDN (Content Distributed Network) 和基于 P2P (Peer to Peer) 的流媒体的研究现状, 针对基于 CDN 的流媒体, 研究了流媒体调度算法, 代理服务器缓存算法, 基于 CDN 的交互式操作. 针对基于 P2P 的流媒体, 研究了数据分配算法, 激励机制, 流媒体对象的放置, 应用层组播, 基于 P2P 的交互式操作. 指出了流媒体的未来研究方向.

**关键词:** 流媒体; 对等网; 内容分发网

中图分类号: TP311 文献标识码: A 文章编号: 0372-2112 (2009) 02-0137-09

## Survey of Key Technologies of the Distribution System for Streaming Media

YANG Ge<sup>1,2</sup>, LIAO Jianxin<sup>1</sup>, ZHU Xiaomin<sup>1</sup>, FAN Xiumei<sup>3</sup>

(1 State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2 Key Laboratory of Integrated Microsystems, Shenzhen Graduate School of Peking University, Shenzhen, Guangdong 518055, China;

3 Beijing Key Laboratory of Intelligent Information, School of Computer Science, Beijing Institute of Technology, Beijing 100081, China)

**Abstract:** Streaming media is a killer application in the future communication system. The key technologies of the distributed system for streaming media are discussed. The research state based on CDN (Content Distributed Network) and P2P (Peer to Peer) for streaming media is presented. For streaming media based on CDN scheduling algorithms, proxy caching algorithms and interactive operations are discussed. For streaming media based on P2P data assignment algorithms, incentive mechanism, placement of media objects, multicast technology on the application layer and interactive operations are discussed. Finally, some issues to be further studied are pointed to.

**Key words:** streaming media; peer to peer network; content distributed network.

### 1 引言

随着 IP 业务的快速发展, 运营商和业务提供商面临着新的商业机会, 各运营商希望通过新业务的开发把网络能力转化为竞争能力, 为用户提供极具吸引力的业务, 同时满足运营商自身的盈利需求, 其中流媒体正受到越来越多的重视, 更有人预言, 流媒体将是通信中的杀手业务. 流媒体是指在网络中使用流式传输技术传输的连续时基媒体, 例如音频、视频、动画或其他多媒体文件. 流媒体的主要特点是以流的形式进行多媒体数据的传输, 即把连续的影音信息经压缩处理后放到网络服务器上, 让浏览者不需等到下载完成, 就可一边下载一边

观赏. 鉴于流媒体技术在应用中的重要地位, 对其在网络中分发策略的研究是非常有必要的, 也是非常紧迫的.

目前流媒体传送策略主要有两种: 基于 CDN (Content Distributed Network) 和基于 P2P (Peer to Peer).

### 2 基于 CDN 的流媒体系统

在基于 CDN 的流媒体系统中, 流媒体服务器和流媒体代理服务器是提供流服务的关键平台, 是流媒体系统的核心设备. 流媒体服务器一般处于 IP 核心网中, 用于存放流媒体文件, 响应用户请求并向终端发送流媒体数据. 流媒体代理服务器, 位于网络的边缘, 靠近用户,

使客户能从位于本地的缓存代理服务器上获取流媒体内容,从而提高用户访问的性能,并减轻骨干网络流量,同时也增加了系统容量,如图 1.

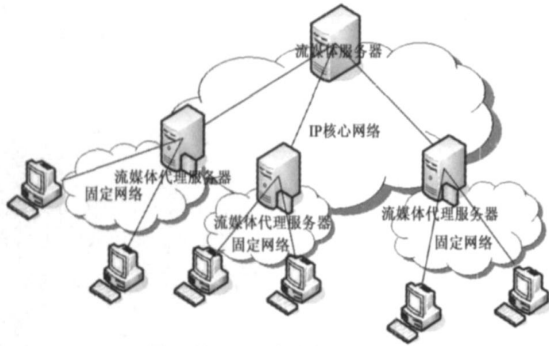


图1 基于CDN的流媒体系统

代理服务器的角色是:从流媒体服务器角度来说,代理服务器是终端,从用户角度来说,代理服务器是服务器。流媒体代理服务器一端支持用户,一端连接流媒体服务器。从流媒体代理服务器到客户端是最短的网络路径,这意味着能减少网络故障,缓解带宽瓶颈。当一部分流媒体已经缓存于代理服务器时,这部分流媒体可以直接从代理服务器以组播方式发到用户,不用再从远端的流媒体服务器提取。因此为了减轻流媒体源服务器和终端的负载,节省主干网络的带宽,应在代理缓存中实现如下情形:对于特别流行的媒体对象应该尽量缓存其所有的数据,使得代理可以尽量服务于请求的用户,对于比较流行的媒体对象应该缓存大部分数据,这样,源服务器只需要提供剩余部分就可以满足客户要求,对于最不流行的对象,只要缓存媒体对象开始部分(前缀部分),使得流媒体在代理中缓存的数据量尽可能与流媒体流行度成正比关系。

### 2.1.1 流媒体调度算法

在大规模的流媒体系统中,用户的点播往往集中于少数热门节目,这就使得合并用户服务、共享服务器和网络带宽等资源成为可能,于是流调度技术应运而生,它属于应用层面的技术。

典型的流调度算法可分为两类:静态调度算法和动态调度算法。静态调度算法采用服务器推模式,服务器不考虑用户动态行为而调度流,预留信道周期性广播视频文件,适用于分送最流行的媒体对象,动态调度算法采用客户端拉模式,媒体流的调度由用户请求驱动,服务器根据用户请求的情况动态选择相应的调度方案来做出响应,使不同的用户尽可能共享同一个数据流,从而降低服务器带宽资源消耗。

静态调度算法主要包括:金字塔算法<sup>[1]</sup>,摩天大楼算法<sup>[2]</sup>等,金字塔算法将节目划分为长度逐渐递增的若干片断,然后利用组播通道播放不同片断,为支持用户连续收看媒体流,金字塔算法要求用户在任意时刻

必须从两个组播通道中接收数据。摩天大楼算法不采用倍数增加的方式切分节目片断,而设计了特定数列,再按照数列中的比例切分流媒体对象。以上算法不能根据用户请求到达情况做出动态调整,占有较多缓存。

动态调度算法主要包括:batching 算法<sup>[3]</sup>、patching 算法<sup>[4,5]</sup>、Piggybacking 算法<sup>[6]</sup>、OBP<sup>[7]</sup>、OBP + Prefix caching 算法<sup>[8]</sup>、OBP+ prefix&patch caching<sup>[9,10]</sup>等。Batching 算法将不同用户的请求绑定于一个组播流中,具有很好的带宽效率,但用户会有较大的响应延迟;patching 算法中,客户端接收正在组播的节目,同时用一个单播接收已播放的节目前缀,其有效地降低了服务器需要传输完整的流的个数,进一步节省了传输带宽,但随着请求到达率的增大所需的补丁通道个数迅速增大,导致较大的服务器负担;OBP 算法结合了Batching 算法和 Patching 算法的优点,虽然降低了服务器网络带宽消耗,但在实践中依然会遇到很大困难,因为它依赖一个完全具有网络层组播能力的网络,而且要求客户端具备同时接收多个流的能力,客户端的启动延迟也较大,当网络规模较大时,缺乏足够的灵活性和可扩展性。Piggybacking 算法改变相邻视频流的播放速率,让后面的流赶上前面的流,然后进行合并,其提高了系统资源的利用率,但调节用户的播放速率影响了用户对媒体节目的收看效果。OBP + Prefix caching 算法通过在代理服务器中提前缓存流媒体对象的前缀部分<sup>[10]</sup>,可以降低或消除客户端的启动延迟。以上算法在骨干网络带宽消耗,服务器负载方面都获得较好性能,但没有考虑占流媒体对象大部分的后缀的缓存策略,对后缀的有效缓存不仅可以进一步降低骨干网络带宽消耗和服务器的负载,而且能有效提供用户的QOS。OBP+ prefix&patch caching 算法将补丁数据也进行分段缓存,让不同批处理区间到达的客户可以共享一部分补丁块,从而获得更好的性能,但当客户对/热门0对象的访问请求强度很高时候,这些算法仍然需要消耗较高的骨干网络带宽。文献<sup>[11]</sup>提出的P<sup>3</sup>S<sup>2</sup>A(Proxy2assisted Patch Prefetching and Service Scheduling Algorithm)调度算法,根据当前客户请求到达的分布状况,代理服务器为后续到达的客户请求进行补丁预取和缓存,特别,对于那些较流行的媒体对象,客户请求到达率很高时,效果更加明显,而在代理服务器缓存空间的消耗方面和OBP + prefix&patch caching 算法基本相同,但它没有区分不同流行度的流媒体对象,对每个流媒体对象都进行缓存,对那些很少被访问的媒体对象进行全部或者部分缓存都将造成代理服务器缓存效率的下降,P<sup>3</sup>S<sup>2</sup>A 算法将客户对媒体对象的访问等同起来,只是考虑客户对媒体对象的访问频率,当客户对媒体对象的访问时间不同时,没有区别对待,使得代理服务器为每个流媒体对象

分配的最大缓存空间都是静态变量,而不能根据流媒体流行度的变化动态调整。

## 212 代理服务器缓存算法

近年来,随着流媒体技术在 Internet 环境中的高速发展,对代理服务器的研究也逐步深入,其中代理服务器缓存分配算法和缓存替换算法是管理缓存的主要手段,也是决定代理服务器性能的核心因素。缓存分配算法决定哪些媒体对象的哪些部分被缓存到代理服务器中,当代理服务器没有足够空间存储新的媒体对象时,缓存替换算法决定哪些缓存单元以及多少缓存单元被清除,以释放足够的空间。

一个理想的流媒体代理服务器缓存系统应具有以下特性:

(1) 快捷性: 缓存系统应该能够有效地降低客户的访问延迟,这是缓存系统的根本目的;

(2) 鲁棒性: 鲁棒性意味着可用性,客户希望流媒体服务随时可用;

(3) 透明性: 缓存系统对客户应是透明的,客户得到的结果仅仅是快速的响应和良好的可用性;

(4) 高效性: 缓存系统给网络带来的开销越小越好;

(5) 稳定性: 缓存系统采用的方案不应给网络带来不稳定;

(6) 简单性: 简单的方案更加容易实现且易被普遍接受,一个理想的缓存方案配置起来应简单易行。

目前,已经有学者对缓存分配算法进行了一些研究。Sen 等<sup>[12]</sup>提出,缓存流媒体的前缀可以更有效地减少流媒体播放的延时和抖动。Zhang 等<sup>[13]</sup>提出了视频分段传输(staging)算法,通过在代理服务器缓存中存储数据率变化较大的流媒体数据,以减少流媒体服务器与代理服务器间的带宽变化。Ramesh 等<sup>[14]</sup>在网络支持多播的条件下提出了使得流媒体服务器与代理服务器间的传输成本最小的缓存分配算法。Kangasharju 等<sup>[15]</sup>针对分层可扩展性编码的流媒体,以层为缓存粒度,用最优方法求解缓存分配的问题,在缓存大小有限以及服务器与代理间带宽有限的条件下最大化服务提供者的收益。雷正雄等<sup>[16]</sup>定义了平均传输成本和平均播放启动延迟,综合考虑了这 2 个参数,提出了基于收益的缓存替换算法,但它假设单位长度的数据消耗的网络成本是相同的,缺乏一般性。杨戈等<sup>[17]</sup>根据移动流媒体对象段的流行度提出了新的缓存算法,使移动流媒体对象在代理服务器中缓存的数据量和其流行度成正比,并且根据客户平均访问时间动态决定该对象缓存窗口大小,但是,当替换相应媒体对象所带来的代理服务器的空闲空间小于将要缓存的媒体对象段时,新的段没有被缓存进代理服务器,原来文件又被替换掉,这

时代理服务器的整体效用将下降。

传统的缓存替换算法,典型的有 LFU(least frequency used)<sup>[18]</sup>, LRU(least recently used)<sup>[19]</sup>, LRU2K<sup>[20]</sup>等,这些算法以访问频率或最近访问时间判断数据的冷热程度,从而决定缓存的替换。LFU 算法总是替换出使用频率最小的对象,认为对象的使用频率越高,未来的使用价值越大,但 LFU 算法存在缓存污染的问题(过去曾被多次使用的对象即使不再被使用,仍占据缓存空间)<sup>[21]</sup>。LRU 算法总是替换最长时间没被使用的对象,但 LRU 算法存在长环模式问题(由于缓存大小小于对象的重用模式长度,存在对象刚被替换出缓存又被请求使用的情况)<sup>[21]</sup>。而 LRU2K 算法将访问频率和访问的最近性综合到效用函数的设计中,虽然取得了较好性能,但也同时存在 LRU 算法和 LFU 算法的缺点。

文献[22]提出 A<sup>2</sup>LS 算法,尽可能延迟媒体对象的分段,分段长度是由客户访问行为决定,另外根据文中提出的效用函数,媒体对象段能够被自适应地接纳和替换,但是文献[22]提出的效用函数没有考虑媒体对象段与媒体对象初始位置的距离,因为媒体对象最后部分通常具有最少价值,段缓存的优先权随着该段与媒体对象初始位置的距离成单递减<sup>[6,23]</sup>,距离媒体对象开始位置越近的段,被访问的频率越高。另外 A<sup>2</sup>LS 算法对于第一次被访问的媒体对象,都是全部缓存到代理中,这样造成一定盲目性,有可能频繁缓存或者替换某些媒体对象。还有 A<sup>2</sup>LS 算法虽然尽可能延迟分段,但一旦分段,其每个段长度就固定,不能根据客户访问行为进行改变,很有可能造成该分段长度不适应客户的请求。

Hyper Proxy 算法<sup>[24]</sup>尽可能延迟媒体对象的分段,分段长度是由客户访问行为决定,以主动预取的方式减少代理抖动,向客户提供较高的流媒体服务质量,另外根据文中提出的效用函数,媒体对象段能够被自适应地接纳和替换,但是文献[24]和文献[22]有共同的缺点。

文献[25]提出了基于焦点分段的流媒体代理缓存策略,它考虑到不同流媒体对象在流行性方面的差异和同一对象不同段上访问频率的差异,以段的流行度作为调度的依据,融合了部分前缀缓存的方法,避免了自适应滞后分段缓存算法在延时启动率方面的不足,但在准入和逐出法则中,只是尽量缓存被访问频率高的段,较少考虑段被访问的时间也是影响段的焦点程度的重要因素。

## 213 基于 CDN 的交互式操作

流媒体技术的出现,促进了视频点播业务在 Internet 上的快速发展,流媒体代理服务器缓存算法、流媒体动态调度算法的研究,进一步增加了视频点播类业务的性能,用户可以根据自身爱好选取节目,还可在节目

播放过程中与进行实时交互。

VCR 操作是视频点播系统中客户与服务器交互的重要手段. 文献[26]建立了 VCR 操作的数学模型: 当一个用户向系统请求某个媒体对象时, 用户交互行为的状态转移如图 2 所示, 其中  $p_i$  ( $i=0, 1, \dots, 9$ ) 定义为状态转移概率,  $d_j$  ( $j=0, 1, \dots, 8$ ) 定义为在特定状态的滞留时间. 假定在整个视频点播期间, 用户停留在每个状态的时间满足指数分布. 另外引入了三类参数: (1) VCR 请求发生的频率, 通过调节转移概率  $p_i$  实现; (2) VCR 操作的持续时间, 通过调节滞留时间  $d_i$  达到; (3) VCR 操作的偏向性, 即 VCR 操作中, 快进和快退哪个占主导, 可以通过调节  $p_2, p_3$  的比例来实现.

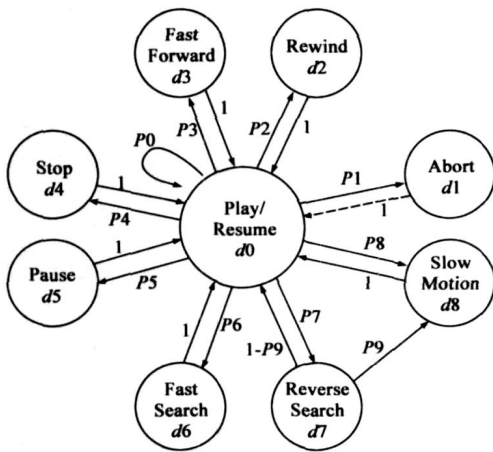


图2 用户VCR操作的状态转移图

文献[27]分析了 Monash 大学视频点播系统(McIV2 ER 系统), 提出了面向交互式视频点播系统的客户端缓存策略, 节省了网络流量, 研究了交互式视频点播系统中用户点播视频的统计: 包括视频再利用的范围; 用户拖动视频滑动块跳跃的距离; 跳跃之间的时间间隔, 它认为这个时间间隔服从对数正态分布, 而没有采用通常认为的指数分布. 文献[27]假定缓存足够大以致可以缓存全部通过网络传输的媒体对象, 然而, 对于大的媒体对象, 这个策略将不太可行. 文献[28]提出了主动式缓存管理技术, 在广播 VoD (video on demand) 系统中实现了交互式功能, 用户可以根据播放点在缓存中的位置, 从广播信道中有选择的预取段, 缓存的内容将相应调整, 以保证播放点的相对位置保持在缓存区的中间, 但是它是广播式, 即使没有客户请求媒体对象, 服务器仍然保持广播媒体对象, 造成资源的浪费.

Popwise 算法<sup>[29]</sup> 缓存最流行媒体对象的最流行片段, 为了兼容非交互式场景, Popwise 算法也包含了 Pre2 fix 算法<sup>[8-10]</sup>, 使得 Popwise 算法在不同的用户请求模式和交互强度下都可以提供较低的用户响应时延和链路占用带宽, 但是该算法以瞬时受访频率作为整个媒体

对象的流行度, 媒体对象流行度的判断不够准确<sup>[22]</sup>, 以累计受访次数作为媒体片段的流行度, 将每次访问同一个媒体片段等同起来, 但每次访问同一个媒体片段的时间不一定是相同的.

文献[30]针对单一代理服务器的流媒体系统, 考虑了 VCR 操作的请求到达时刻代理服务器中缓存的分布情况和请求的数据在数据流中的位置, 尽可能地选择合适的缓存为请求提供服务.

### 3 基于 P2P 的流媒体系统

P2P 的基本思想是充分利用网络上分布在不同地理位置上的客户端资源, 采用分布式计算模式来为网络上的用户提供各种服务, P2P 网络中没有集中的服务器, 网络的每个节点既可以作为客户接收其他节点的服务, 也可以作为服务器向其他节点提供服务. 基于 P2P 的流媒体系统也是借助这种思想进行流媒体内容的分发, 其目标是充分利用多客户机的空闲资源, 构建一个成本低, 扩展性好, 并有一定 QoS 保证的流媒体分发系统. 但是在 3G 网络中实现基于 P2P 的分发策略是比较困难的, 一般情况下, 无线信道传输可靠性差, 误码率高, 而且无线信道的传输带宽也明显小于 Internet 的带宽, 3G 终端的性能也明显小于 Internet 上的主机性能, 尤其是有限的终端缓存. 同时新的流媒体文件必须要分布到足够多的节点以后才能正常提供服务, 其中的管理也是非常复杂的.

#### 3.1 数据分配算法

在 P2P 网络中, 请求节点请求的媒体对象通常由不同的供应节点提供, 数据分配算法就是决定哪些供应节点提供被请求的媒体对象的哪些部分. 文献[31]提出了 OTS<sub>p2p</sub> 算法, 媒体对象的数据由不同的供应节点提供, 根据节点输出带宽的不同, 将节点分成不同的类, 第  $n$  类节点的输出带宽是  $R_0/2^n$ ,  $R_0$  是媒体对象的播放速率, 请求媒体对象的节点运行 OTS<sub>p2p</sub> 算法, 将媒体对象的数据分配给不同的供应节点, 获得最优的分配结果, 取得最小的缓冲延迟, 但是它只对每类节点提供的输出速率进行了假设, 要求每类节点的输出带宽是  $R_0/2^n$ , 而且要求被请求的媒体对象的段的数目是  $2^n$ , 缺乏一般性, 而且该算法是静态分配算法, 不能根据网络的环境进行变化, 当系统出现拓扑变化和带宽波动等情况时, 需要启动算法重新进行分配. 文献[32]提出了 MBDA<sub>p2p</sub> 算法, 它考虑了已分配资源块, 待分配的资源块以及各资源块产生的缓冲延迟, 实现了动态分配策略, 能够适应不稳定的网络环境, 但它只能在媒体对象的提供节点在提供媒体对象时不离开的情况下取得很好的性能, 而 P2P 网络的节点通常是不稳定的, 媒体对象

的提供节点有可能中途退出系统, Algorithm 1 算法<sup>[31]</sup> 是将被请求的媒体对象的段按照段的顺序依次分配给提供节点, 但应用于 P2P 流媒体传输时具有较大的缓冲延迟<sup>[31, 32]</sup>. 文献[33]提出了 2 种算法: 基于多目标优化的近似算法, 基于分枝定界的精确算法, 利用计算二分图中的最大流值来确定分枝上界和被裁减的分枝, 在节点之间对数据层进行优化分配, 减少了对源服务器带宽的占用, 但它假设了提供节点的上行带宽足够大, 而实际中节点的上行带宽资源是非常宝贵和有限制的.

文献[34]提出了一种基于分布式文件系统的调度机制, 实现了点播, 组播, 文件下载, 具有就近调度、站内调度客户节点、站间服务节点自调度的特点, 有效降低了主干网上的数据调度.

### 312 激励机制

P2P 网络中, 需要提供节点尽量贡献自己的流媒体对象, 提高网络的资源利用率, 快速增长网络的服务能力, 但是, 实际网络中众多节点更愿意接收其他节点提供的服务, 而不愿意向其他节点提供服务<sup>[35]</sup>, 目前通常采用激励机制, 使得向其他节点提供服务越多的节点获得的服务也越多, 使得网络中的节点在消费流媒体对象时, 能够主动承担生产者的角色. 有 2 种激励机制:

(1) 通过货币支付方式, 贡献资源的节点收取货币, 享受资源的节点支付货币<sup>[36]</sup>. (2) 通过区分服务方式, 贡献资源越多的节点将获得更好的服务质量<sup>[37]</sup>. 文献[38]提出了基于博弈论的激励机制, 微支付和共享酬劳, 但它要求中心节点行使仲裁. 文献[39]向具有不同贡献值和连接类型的节点提供了区分化的服务, 提出了基于激励的资源分配, 在 P2P 环境下, 保持累积贡献值和社会效用, 如果全部请求节点有相同的贡献值, 则最大化社会效用, 该机制取得了更高的累积效用和公平性, 一个节点提供的服务越多, 它获得的贡献值越多, 当贡献值大的节点和其他节点竞争服务时, 这个节点获得的效用越大.

文献[40]通过激励机制, 以提高社会福利为目的, 建立了一个文件共享模型.

文献[41]提出了面向 P2P 网络的成本模型, 当且仅当某个节点获得的服务比它给予其它节点的服务多时, 这个节点才加入系统中, 当系统的节点数目小于一个阈值时, 系统将不稳定, 当系统的节点数目大于这个阈值时, 系统将稳定而且可以自我发展.

文献[42]提出了一个面向对等网络流媒体的服务差异化的节点选择机制, 提供了灵活选择节点, 提高了流媒体质量, 这个激励机制遵从等级顺序竞赛的原理, 只是考虑竞赛者相对的性能, 竞赛的优胜者将获得更高的奖励, 但是它假设流媒体系统的用户和文件共享

系统的用户有相同的行为.

文献[43]针对自私节点, 提出了 P2P 流媒体系统中一种积分检测相结合的激励机制, 节点通过自身或子节点共享数据流获得贡献积分, 贡献积分越高, 节点获得数据流的优先权就越高; 为了贡献积分最大化, 节点有选择地对它的子节点进行检测来促进子节点数据的共享, 从而提高整个系统的性能.

### 313 流媒体对象的放置

流媒体对象的放置目前主要有: 基于分条 (striping)<sup>[44~46]</sup>和复制 (replicating)<sup>[47~49]</sup> 2 种方式, 分条指将一个完整的媒体对象分成一个个段, 按照一定的原则存储到磁盘上, 它能提高存储的利用率, 但是只是适用于同构的存储方式. 复制指在其他服务器上完全拷贝该媒体对象, 它以牺牲存储能力来获取系统的扩展性.

文献[44, 45]只是考虑了文件之间的访问倾向性, 没有考虑文件内部的访问倾向性, 虽然解决了前者造成的系统热点瓶颈问题, 但没有考虑后者对流媒体服务器产生的负面影响, 容易导致服务器系统负载不均, 降低系统性能. 文献[46]在研究了文件内部访问倾向性特征的基础上, 根据是否增加多余数据分别提出了 SPS 方法 (symmetrical pair scheme) 和 SRPS 方法 (symmetrical replication pair scheme), SPS 方法使用不同文件构造对称因子, 适用于没有冗余数据的流媒体服务器, SRPS 方法使用数据复制构造对称因子, 具有系统开销小, 简便实用, 性能优异的特点, 适用于大规模的流媒体服务器, 大幅度提供了系统的负载均衡能力, 但是 SPS 方法和 SRPS 方法只是考虑文件内部的访问倾向性, 没有考虑文件之间的访问倾向性, 很明显后者同样对前者有很大的影响.

文献[47]提出了启发式内容复制与替换算法, 在不同应用组间复制节目内容以提高系统的服务能力, 在应用组内复制节目内容能够增加服务的健壮性, 但它并没有针对视频服务的特点对应用组内节点的管理、内容存储策略、系统性能、负荷均衡等问题深入讨论.

文献[48]提出了一个解决媒体对象放置的最佳解, 综合考虑了传输成本和编码成本, 降低了访问媒体对象的成本. 文献[49]提出了在自治系统中点到点的媒体放置算法, 充分考虑了网络带宽和节点能力的限制以及媒体的访问频率.

### 314 应用层组播

应用层组播就是一种由主机终端在应用层实现组播通信的组播方案, 参与组播通信的节点自组织地配置进叠加网络, 端系统在组播网络中执行应用层组播路由由算法来确定一条虚拟逻辑组播路径, 数据包沿着这条逻辑链路转发, 逻辑路径其实是 2 节点间传统网络

物理链路组成的通信路径,这条实际的物理路径对端系统而言是透明的.文献[50]提出了分布式组播路由算法,实现了在一定延迟限制下,产生非最优网络成本的路由树,能够灵活适应组员的变化.文献[51]提出了正交(orthogonal)的遗传组播路由算法,在一定延迟下,使得源节点到目的节点的成本最小.文献[52]将应用层组播抽象成对边和点都带权的有向图求解,提出了2个启发式近似算法:基于度的算法和基于最大延迟路径的算法.文献[53]提出了构建具体分层排列图结构的应用层组播系统,参与的主机组合成多个排列图,每个排列图中嵌入多个独立组播树,排列图间形成树状结构,对应组播树相互衔接并保证独立性,流媒体数据采用多描述编码,每个描述数据通过一个独立的组播树传输,使得单个主机故障至多影响一个数据,提供了系统流媒体服务的稳定性.文献[54]使用应用层组播技术,设计和实现了流媒体直播原型系统.

### 3.1.5 基于 P2P 的交互式操作

在 P2P 网络中用户同样也需要进行交互.文献[55]提出了基于 P2P 和移动代理技术的交互式通信的体系结构,这个系统不但能提供文件数据,而且能够把代理发送到请求节点,完成数据的处理,同时把处理结果返回给提供节点,但它实现起来比较复杂.文献[56]针对交互式数字电视系统提出了一个模型,将多媒体文件表示成结构化文件,终端用户可以很容易地增强媒体内容,而且能够在 P2P 网络中和特定节点共享媒体对象增强层的内容.文献[57]在没有服务器的 P2P 网络中提出了一个交互协作的媒体框架,但是它没有考虑节点能够缓存的媒体对象的个数,而一般节点都是有有限的缓存空间.文献[58]提出了基于 P2P 的视频点播系统的设计和实现,尤其是发现机制.

## 4 总结

本文针对基于 CDN 和基于 P2P 的 2 种流媒体分发系统,讨论了流媒体分发的关键技术,今后需要考虑:

(1) 如何实现流媒体的传送策略,尽量避免 CDN 和 P2P 的缺点.基于 CDN 的传送策略对源服务器和代理服务器,以及它们之间的带宽都提出了较高的要求,基于 P2P 的传送策略对于各个节点之间的管理也是其得以实施的难点之一.

(2) 在 P2P 网络中,服务节点的选择问题.当用户请求服务时,选择哪一个或哪些服务节点提供服务;当正在提供服务的节点失效时如何保证流媒体服务的连续.这时考虑选择另外一个适合的服务节点来承担失效的服务节点的传输任务,需要综合考虑被选择节点的资源(如节点之间的带宽,节点的计算能力等).

(3) 在 P2P 网络中,对多个服务节点的传输调度.

即确定每个候选服务节点向请求节点提供流媒体中的哪些部分,既要满足流媒体适时的要求,又不能超过每个服务节点的资源限制,尽可能减少本次服务对整个网络资源的消耗.

(4) 在 P2P 网络中,如何评估共享资源的真实性和可靠性.网络中每个节点的安全性和可靠性是不同的,甚至会有欺诈的节点,需要进行区别对待,这可以通过数字证书对节点进行认证等方法来解决.

(5) 目前针对 P2P 网络中节点的行为和兴趣还没有统一的认识,而它又是很多问题的基础,比如建立适当的激励机制等问题,这就需要以后对这个问题深入的研究,进行统计建模.

总之,向用户提供快捷、稳定、高质量、低成本的流媒体是流媒体分发系统追求的目标,也是流媒体业务成为下一代网络的杀手业务的关键,为此需要不断地深入研究,其中基于 P2P 的流媒体分发系统将成为今后流媒体发展的热点之一.

### 参考文献:

- [1] S Viswanathan, et al. Pyramid broadcasting for video on demand service[A]. IEEE International Conference on Multimedia Computing and Networking[C]. San Jose, CA, 1995. 24 (17): 66- 77.
- [2] A Hua Kien, et al. Skyscraper broadcasting: A new broadcasting scheme for metropolitan video on demand system[A]. Proc ACM SIGCOMM Conference[C]. Cannes, France, Sept. 1997. 89- 100.
- [3] A Dan, et al. Scheduling policies for an on demand video server with batching[A]. Proc ACM multimedia[C]. San Francisco, California. 1994. 15- 24.
- [4] Y Cai, et al. Optimizing patching performance[A] IEEE International Conference on Multimedia Computing and Networking [C]. San Jose, California. 1999.
- [5] Kien A Hua, Ying Cai, Simon Sheu. Pathing: a multicast technique for true video on demand services[A]. Proceedings of ACM Multimedia[C]. Bristol UK. 1998. 191- 200.
- [6] L Golubchik, et al. Adaptive piggybacking: a novel technique for data sharing in video on demand storage servers[J]. ACM Multimedia Systems, 1996: 140- 155.
- [7] P P White, J Crowcroft. Optimized batch patching with classes of service[J]. ACM of the Communications, 2000. 30(4): 21- 28.
- [8] Chitra Venkatramani, Olivier Verscheure, Pascal Frossard, Kang Won Lee. Optimal proxy management for multimedia streaming in content distribution networks[A]. Proceedings of the 12th international workshop on Network and operating systems support for digital audio and video[C]. Miami, Flor2

- da, USA. May. 2002. 147- 154.
- [9] Olivier Verscheure, Chitra Venkatramani, Pascal Frossard and Lisa Amini. Joint server scheduling and proxy caching for video delivery[J]. Computer Communications. 2002. 25 (4) : 413- 423.
- [10] Frossard P, Verscheure O. Batch patch caching for streaming media[J]. IEEE Communications Letters, April 2002, 6(4) : 159- 161.
- [11] 覃少华, 李子木, 蔡青松, 胡建平. 基于代理缓存的流媒体动态调度算法研究[J]. 计算机学报, 2005, 28(2) : 185 - 194.  
Qin Shaohua, Li Zimu, Cai Qingsong, Hu Jianping. Study on dynamic scheduling algorithms for streaming media based on proxy caching[J]. Chinese Journal of Computers, 2005. 28 (2): 185-194. (in Chinese)
- [12] S Sen, J Rexford, D Towsly. Proxy prefix caching for multimedia streams[A]. Proc IEEE Infocom. 99[C], Amherst, MA, USA. 1999. 1310- 1319.
- [13] Zhen Zhang, Y Wang, D H C Du, and D Su. Video staging: a proxy server based approach to end-to-end video delivery over wide area networks[J]. IEEE ACM Trans. Networking, Aug. 2000, 8: 429- 442.
- [14] S Ramesh, I Rhee, and K Guo. Multicast with cache (Mcache): an adaptive zero delay video on demand service[J]. IEEE Transactions on Circuits and Systems for Video Technology, March 2001, 11(3): 440- 456.
- [15] J Kangasharju, F Hartanto, M Reisslein, and K W Ross. Distributing layered encoded video through caches[J]. IEEE Transactions on Computers, June 2002, 51 (6) : 622 - 636.
- [16] 雷正雄, 廖建新, 朱晓民. 基于补丁流传输机制的移动流媒体系统的缓存替换算法[J]. 高技术通讯. 2006. 16(7) : 671- 675.  
LEI Zhengxiong, LIAO Jianxin, ZHU Xiaomin. Cache replacement algorithm for patching stream transmission scheme based mobile streaming media system[J]. Chinese High Technology Letters, 2006, 16(7): 671-675. (in Chinese)
- [17] 杨戈, 廖建新, 朱晓民, 杨波, 黄海. 基于段流行度的移动流媒体代理服务器缓存算法[J]. 通信学报, 2007, 28(2) : 33- 39.  
YANG Ge, LIAO Jianxin, ZHU Xiaomin, YANG Bo, HUANG Hai. Proxy caching algorithm based on segment popularity for mobile streaming media[J]. Journal on Communications. 2007. 28(2) : 33- 39. (in Chinese)
- [18] Robinson J T, Devarakonda M V. Data cache management using frequency based replacement[A]. Proceedings of SIGMETRIC on Measuring and Modeling of Computer Systems[C]. Boulder, Colorado, USA, 1990. 134- 142.
- [19] Alghazo J, Akaaboune A, Botros N. SF2LRU cache replacement algorithm[R]. Records of the 2004 International Workshop on Technology, Design and Testing. Carbondale, IL, USA. Aug 2004: 19- 24.
- [20] Seung W S, Ki Y K, Jong S J. LRU based small latency first replacement (SLFR) algorithm for the proxy cache[A]. Proceedings of IEEE International Conference on Web Intelligence[C]. Daejeon, South Korea, 2003. 499- 502.
- [21] 肖明忠, 李晓明, 刘翰宇等. 基于流媒体文件字节有用性的代理服务器缓存替代策略[J]. 计算机学报, 2004, 27(12): 1633- 1641.  
XIAO Mingzhong, LI Xiaoming, LIU Hanyu et al. Proxy cache replacement strategies based on bytes benefit of streaming media file[J]. Chinese journal of computers. 2004, 27 (12): 1633- 1641. (in Chinese)
- [22] Chen S Q, Shen B, Wee S, Zhang X D. Adaptive and lazy segmentation based proxy caching for streaming media delivery[A]. Proceedings of The 13th International Workshop on Network and Operating Systems Support for Digital Audio and Video[C]. Monterey, CA, USA. Jun2003. 22- 31.
- [23] Kunlun Wu, Philip S, R Joel L Wolf. Segmentation of multimedia streams for proxy caching[J]. IEEE Transactions on Multimedia, October 2004, 6(5) : 770- 780.
- [24] Songqing Chen, Bo Shen, Susie Wee, Xiaodong Zhang. Designs of high quality streaming proxy systems[A]. In proc. IEEE INFOCOM. 04[C]. HongKong, vol. 3, March 2004. 1512- 1521.
- [25] 张东辉, 李文辉, 李鸣. 基于焦点分段的流媒体代理缓存策略[J]. 吉林大学学报(工学版), 2007. 9, 37(5): 1137- 1141.  
ZHANG Donghui, LI Weihui, LI Ming. Popular segment based proxy caching for streaming media over internet[J]. Journal of Jilin University(Engineering and Technology Edition), 2007. 9, 37(5) : 1137- 1141. (in Chinese)
- [26] AbramProfeta E L, Shin K G. Providing unrestricted VCR functions in multicast video on demand servers[A]. Proceedings of IEEE international conference on multimedia computing and systems[C]. Austin, TX 1998. 66- 75
- [27] Branch P, Egan G, Tonkin B. A client caching scheme for interactive video on demand[A]. IEEE of Int'l Conf. Networks (ICON. 99)[C], Brisbane, Australia, 1999. 391- 397.
- [28] Zongming Fei, Ammar M H, Kamel I, Mukherjee S. An active buffer management technique for providing interactive functions in broadcast video on demand systems[J]. IEEE transaction on multimedia, Oct. 2005, 7(5) : 942- 950.
- [29] 刘威, ChunTung Chou, 程文青, 杜旭. 交互式流媒体代理缓存[J]. 计算机研究与发展, 2006, 43(4) : 594- 600.  
LIU Wei, ChunTung Chou, CHEN Wenqing, DU Xu. Proxy Caching for Interactive Streaming Media[J]. Journal of Computer Research and Development, 2006, 43(4): 594- 600. (in Chinese)

- [30] 罗玮华, 杨坚, 郑, 胡晗. 一种支持视频 VCR 操作的共享动态缓存算法 [J]. 计算机应用, 2008, 28(6): 1624-1626.  
LUO Wenhua, YANG Jian, ZHENG Quan, HU Han. Shared running cache algorithm with supporting video VCR operation [J]. Journal of Computer Applications, 2008, 28(6): 1624-1626. (in Chinese)
- [31] Xu D, Hefeeda M, Hambrusch S, et al. On peer-to-peer media streaming [A]. In IEEE Conference on Distributed Computing Systems [C]. West Lafayette, IN, USA. 2002. 363- 371.
- [32] 杨薇薇, 黄年松. 一种 P2P 流媒体数据传输任务分派算法 [J]. 华中科技大学学报(自然科学版). 2005, 33(5): 26 - 28.  
YANG Weiwei, HUANG Niansong. Data assignment of P2P media streaming [J]. Journal of Huazhong University of Science and Technology (Nature Science), 2005, 33(5): 26- 28. (in Chinese)
- [33] 刘亚杰, 龚文华. P2P 分层流媒体中源服务器参与的数据层分配算法 [J]. 计算机研究与发展, 2005, 42(9): 1472 - 1477.  
LIU Yajie, DOU Wenhua. Layer allocation algorithms in layered peer-to-peer streaming with source servers participation [J]. Journal of Computer Research and Development, 2005, 42(9): 1472- 1477. (in Chinese)
- [34] 闫巧玲, 刘心松, 曹铮, 张汀汀. 一种基于分布式并行系统的流媒体数据分发系统计算机应用 [J]. 计算机应用, 2008, 28(3): 826- 828.  
YAN Qiaoling, LIU Xinsong, CAO Zheng, ZHANG Tingting. Design and implementation of a distributed parallel system based on streaming media distributing system [J]. Journal of Computer Applications, 2008, 28(3): 826- 828. (in Chinese)
- [35] Hughes D, Coulson G, Walkerdine J. Free riding on Gnutella revisited: the bell tolls? [J]. IEEE distributed systems online. June 2005. 6(6): 1- 18.
- [36] Krit Wongrujira, Aruna Seneviratne. Monetary incentive with reputation for virtual marketplace based P2P [A]. Proceedings of the 2005 ACM conference on Emerging network experiment and technology [C]. Toulouse, France. Oct. 2005: 135- 145.
- [37] PRichard T B Ma, Sam C M Lee, John C S Lui, David K Y Yau. Incentive and service differentiation in P2P networks: a game theoretic approach [J]. IEEE/ACM Transactions on Networking (TON). Oct. 2006. 14 (5): 978-991.
- [38] Golle P, Leyton-Brown K, Mironov I. Incentives for sharing in peer-to-peer networks [A]. Proc. Of second workshop on electronic commerce (WELCOM. 01) [C], Heidelberg, Germany, Nov, 2001. 264-267.
- [39] Ma R T B, Lee S C M, Lui J C S, Yau D K Y. An incentive mechanism for P2P networks [A]. Proceedings of the 24th international conference on distributed computing systems [C]. China. 2004: 516- 523.
- [40] Courcoubetis C, Weber R. Incentives for large peer-to-peer systems [J]. IEEE Journal on Selected Areas in Communications. May 2006. 24 (5): 1034- 1050.
- [41] Guangzhong Sun, Guoliang Chen, Junmin Wu. Incentive for participating in a hybrid peer-to-peer system [A]. Proceedings of the Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT. 05) [C]. Dec. 2005. 533- 536.
- [42] Habib, A, Chuang J. Service differentiated peer selection: an incentive mechanism for peer-to-peer media [J]. IEEE Transactions on multimedia. June 2006. 8 (3): 610-621.
- [43] 路卫娜, 杨寿保, 郭磊涛. P2P 流媒体系统中积分检测相结合的激励机制 [J]. 中国科学院研究生院学报. 2008, 25 (1): 61- 68.  
LU Weina, YANG Shoubao, GUO Leitao. A score and monitoring based incentive mechanism in P2P media streaming system [J]. Journal of The Graduate School of The Chinese Academy of Sciences. 2008, 25 (1): 61-68. (in Chinese)
- [44] Gafsi, Biersack. Data striping and reliability aspects in distributed video servers [J]. Cluster Computing, 1999. 2(1): 75 - 91.
- [45] Shenoy J P, Vin H M. Efficient striping techniques for variable bit rate continuous media file servers [J]. Performance Evaluation, 1999, 38(9): 175- 199.
- [46] 吴松, 金海, 邹德清. 一种流媒体文件的分块放置方法 [J]. 计算机学报. 2006, 29(3): 500-507.  
WU Song, JIN Hai, ZOU Deqing. A data placement strategy for parallel streaming servers [J]. Chinese Journal of Computers. 2006, 29(3): 500-507. (in Chinese)
- [47] Xiang Z, Zhang Q, Zhu W et al. Peer-to-peer based multimedia distribution service [J]. IEEE Transactions on Multimedia. 2004, 6(2): 343- 355.
- [48] Keqiu Li, Hong Shen, Chin F Y L, Liusheng Huang. Multimedia object placement for hybrid transparent data replication [A]. IEEE Global telecommunications conference [C]. Ishikawa, Japan. vol. 2 Dec. 2005. 631- 635.
- [49] Kan Hung Wan, Loeser C. An overlay network architecture for data placement strategies in a P2P streaming network [A]. The 18th international conference on advanced information networking and applications [C]. Germany. vol. 1, 2004. 119 - 125.
- [50] Xiaohua Jia. A distributed algorithm of delay-bounded multicast routing for multimedia applications in wide area networks [J]. IEEE transactions on networking. Dec. 1998, 6(6): 828- 837.
- [51] Qingfu Zhang, Yiu-Wing Leung. An orthogonal genetic algo-

- rithm for multimedia multicast routing[J]. IEEE transactions on evolutionary computation. Apr, 1999. 3(1): 53- 62.
- [52] 曹佳, 鲁士文. 应用层组播的最小延迟生成树算法[J]. 软件学报, 2005, 16(10): 1766- 1773.  
CAO Jia, LU Shiwen. A minimum delay spanning tree algorithm for the application layer multicast[J]. Journal of Software, 2005, 16(10): 1766- 1773. (in Chinese)
- [53] 田瑞雄, 向哲, 李星. 基于分层排列图结构的流媒体应用层组播系统[J]. 清华大学学报(自然科学版). 2004, 44(4): 493- 497.  
TIAN Ruixiong, XIANG Zhe, LI Xing. HAC based application layer multicast system for streaming media[J]. Journal of Tsinghua University( Science and Technology) 2004, 44(4): 493- 497. (in Chinese)
- [54] 胡文彦, 叶德建. 基于应用层组播的高清流媒体直播原型系统的实现和测试[J]. 中国图象图形学报, 2007, 12(10): 1720-1724.  
HU Wenyan, YE Dejian. Implementation and Testing of ALM Based HDTV Live Prototype System[J]. Journal of Image and Graphics, 2007, 12(10): 1720-1724. (in Chinese)
- [55] Deqiang He, Xiangyang Lu, Xiaoyang Yao. P2P and mobile agent based interactive communication in the digital passenger information system[A]. The Sixth International Conference on Intelligent Systems Design and Applications( ISDA. 06) [C]. Jinan. vol. 2, Oct. 2006. 573- 580.
- [56] Pablo Cesar, D C A Bulterman, A J Jansen. Benefits of structured multimedia documents in IDTV: the end user enrichment system[A]. Proceedings of the 2006 ACM symposium on document engineering[C]. Amsterdam, the Netherlands. Oct. 2006. 176- 178.
- [57] Araie H, Hirokawa M, Nishide R, Ueshima S. Proposing an interactive collaboration media in P2P settings[A]. The Fourth International Conference on Creating, Connecting and Collaborating through Computing(C5. 06) [C]. Berkeley, CA. Jan. 2006. 103- 111.
- [58] 刘宇杰, 秦肖臻, 彭力, 杨文俊, 汪秉文. 一种基于 P2P 的视频点播系统设计[J]. 计算机技术与发展, 2007, 17(1): 193- 198.

LIU Yu jie, QIN Xiaozhen, PENG Li, YANG Wenjun, WANG Bingwen. A design of video on demand system based on P2P technology[J]. Computer Technology and Development. 2007, 17(1): 193- 198. (in Chinese)

#### 作者简介:



杨 戈 男, 1974 年 11 月出生, 博士, 主要研究方向为下一代网络、流媒体。(本文通讯作者)

E-mail: g20044444@163.com



廖建新 男, 1965 年 2 月出生, 博士, 北京邮电大学网络与交换技术国家重点实验室教授, 博士生导师, 主要研究方向为下一代网络、智能网。



朱晓民 男, 1974 年 7 月出生, 博士, 北京邮电大学网络与交换技术国家重点实验室副研究员, 主要研究方向为智能网、下一代网络、3G 核心网。



樊秀梅 女, 博士, 北京理工大学副教授, 教育部 2007 年新世纪优秀人才入选者。主要研究领域为计算机网络 QoS 控制、无线 Ad hoc 网络路由技术、容迟网络、网络性能评价。