

# 嵌入式 P2P VOD 系统上面向内容的动态缓存策略

王 攀 黄 浩 谢长生

(华中科技大学计算机学院 武汉 430074)

**摘 要** 随着 P2P 技术的发展和 IPTV 的逐步推广,两者的结合应用越来越广泛。P2P 流媒体的实质是基于分享的精神,充分利用每个客户的上传带宽,以减少流媒体服务器端的负载。关于 VOD 点播中的 P2P 技术,现有的大部分研究都集中在节点的选择策略、P2P 网络的拓扑结构以及各个客户端的上传带宽和上传质量等。这些研究显著地改善了 P2P 点播系统的性能。采用基于嵌入式系统的面向内容的 P2P 点播缓冲机制,可以最大限度地减少嵌入式系统弱点(存储容量非常有限)的影响,并基于内容热度做出智能的缓冲机制。使用这种新的机制,嵌入式 P2P VOD 系统可以得到一个较小的网络组建时间和更高的数据共享率。

**关键词** P2P VOD 技术,面向内容,动态缓存,嵌入式系统

**中图分类号** TP399 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.10.030

## Dynamic Buffer Mechanism of P2P VOD on Embedded System

WANG Pan HUANG Hao XIE Chang-sheng

(School of Computer Science and Technology, Huazhong University of Science & Technology, Wuhan 430074, China)

**Abstract** P2P network is widely used in IPTV systems now. The essence of P2P streaming is making use of the upload bandwidth of each peer to reduce the load of the streaming servers. Regard to VOD P2P network, most of the existing researches are focused on peer selection strategies, such as the peer topology, peer bandwidth contribution, balance of peer churn and stream quality. These researches do make a significant improvement to the P2P live streaming and the P2P VOD systems. In this paper, we designed a new content-oriented buffering mechanism for P2P VOD on embedded system. It can minimize the weakness of embedded system (few memory and storage), and make an intelligent buffer mechanism at the basis of content popularity. Using this new mechanism, P2P VOD system can get a lower built-up time and a higher data sharing rate.

**Keywords** P2P VOD system, Content-oriented, Dynamic buffer, Embedded system

## 1 引言

随着宽带网络的不断发展, P2P 技术也得到了长足的发展,以 P2P 为基础的系统已被广泛地引入互联网来提供 IPTV 服务,其中包括直播(Live Streaming)服务和视频点播(VOD)服务<sup>[1-3]</sup>。基于 P2P 的视频点播系统的基本思想是充分地利用客户端带宽资源,通过资源共享来提供更好的播放效果,并减轻流媒体服务器的负担。特别是用户数量达到一定程度后,传统的单播模式的 VOD 系统将受到服务器的性能和网络带宽这两个瓶颈因素的影响而面临瘫痪,而采用 P2P 技术后,服务器的资源消耗和带宽占用将显著降低。

常见的 VOD 系统的 P2P 研究主要集中在节点选择策略、调度策略和缓存机制。但是这些研究大部分都基于 PC,并未很好地考虑到低端嵌入式系统的资源劣势。以 PC 为例,目前主流的 PC 都配置有至少 2GB 内存和 500GB 的硬

盘,在实际的 VOD 点播中可以充分利用内置的硬盘来存储资源。一般支持 P2P 的 VOD 播放器在 PC 上播放完一个节目后,整个节目文件就会被缓存到硬盘,为 P2P 系统提供了一个新的完整节目源。然而在中低端的嵌入式系统中,资源相当有限,由于系统限制,为了不缩短 NAND Flash 的寿命,系统运行的临时数据基本不会保存在 Flash 中,而是全部存储于内存中。当系统启动后,能用于 VOD 缓存的空闲系统资源不会超过 100MB,因此无法将完整的 VOD 节目或者大段的节目内容存储在客户端,而只能是选择性地存储部分数据。

在本文中,我们正是基于这种资源有限的嵌入式系统,设计出了一种动态缓存机制,采用面向内容的方法来动态调整缓存,以便减小 VOD 网络的组建时间和提高资源的共享效率。本文先介绍 P2P VOD 系统相关研究的现状,然后描述面向内容的动态缓存机制和实现原理,最后通过实际系统的

到稿日期:2013-06-15 返修日期:2013-08-10 本文受中国国家重点基础研究发展计划(2011CB302303),中国国家自然科学基金会(60933002),国家高技术研究发展计划(863 计划)资助项目(2013AA013203)资助。

王 攀(1985-),男,博士生,主要研究方向为嵌入式、数字家庭多媒体与网络,E-mail:techust@gmail.com;黄 浩(1965-),男,博士,副教授,主要研究方向为计算机系统结构、网络与多媒体、智慧城市;谢长生(1957-),男,博士,教授,主要研究方向为计算机系统结构、存储系统、网络存储与云存储系统。

测试来验证这一机制的有效性。

## 2 相关研究现状

当前已经有许多 P2P 服务器基于集中索引机制来建立,如 Napster。Napster 是一个音乐文件交换系统。此系统由终端用户和一个集中维护所有索引的索引服务器构成。集中索引服务器的维护难度相对较低,只需在收到客户端发送的索引信息后进行自我更新即可,同时它也能向客户端提供复杂的查询。这种机制面临的主要问题是,当中央索引服务器崩溃后,整个 P2P 网络将崩溃。网络的可靠性和安全性不能得到保证。分布式的非结构化索引机制已被用于 P2P VOD 系统,比如 Gnustream<sup>[4]</sup>,它可以称为第一个基于 P2P 的网络 VOD 系统,其使用 Gnutella<sup>[5]</sup> 来进行 P2P 网络的创建和维护。虽然它可以从多个客户终端获取数据来提高带宽利用率,但是其可扩展性不好。

Hefeeda 等人提出了一种基于视频文件分段的缓存机制<sup>[6]</sup>。该系统采用树形的拓扑结构,通过控制全局缓存的媒体信息来指导客户端进行缓存控制,其缺点是维护成本高,可靠性低。Cui 等人提出了基于会话的分段缓存模型<sup>[7]</sup>。这个模型在用户数量很大、观看同一个节目的用户很多时是有效的。它可以提高热度很高的节目的播放质量,但是对于那些热度不高,观看人数很少的节目,这种机制的效果很差。这种机制需要强大的服务器支持,具有可扩展性不好的缺点。Guo 等人提出了一种混合式 P2P 缓存管理机制——PROP<sup>[8]</sup>。PROP 有效地解决了缓存问题的可扩展性和可靠性。其缺点是,一个 peer 节点想要请求某一个节目片段时,需要向系统广播请求信息;当一个 peer 节点进行时间段跳转或者退出播放时,它会清空自己的缓冲区,同时需要发送一个更新索引列表的信息给索引服务器,这增加了网络的工作量。

## 3 动态缓存机制

嵌入式系统和普通的 PC 之间最大的区别是硬件性能和存储容量,特别是中低端的嵌入式平台。嵌入式系统一般没有足够的内存和存储空间来缓存用户观看的所有内容,只能够缓存很短的一个片段,而且不能长时间保存。

在普通基于 PC 的 VOD 视频点播系统中,当用户选择一个节目并开始播放时,从一开始这个节目内容就会被缓存在内存或硬盘,播放结束时,整个节目内容将被下载并存储在硬盘中。最终此内容是保留在硬盘中还是被删除取决于用户的设置。

但在嵌入式系统中,通常只有 256MB DDR 内存和 256MB NAND Flash 存储器,甚至更少。当系统启动后,由于系统自身的占用,空闲的硬件资源不多,没有足够的存储空间来存储整个节目内容。视频节目的播放流程是这样的:下载一个小片段,其在播放完成后被删除,播放一个小片段的同时,下载另一个小的片段,如图 1 所示。

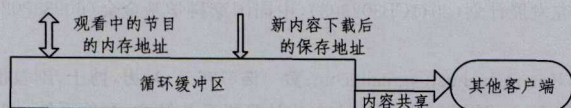


图 1 常用 VOD 缓存机制

当 P2P 网络部署到 VOD 视频点播系统中后,所有的客户端都需要存储播放过的内容,并按照需求分发到其他客户端。对于可用资源有限的嵌入式系统,通常只有一小块固定大小的内存(50MB 到 100MB)用于 VOD 节目内容片段的缓存。当用户开始观看节目后,此循环缓冲区将被节目内容填充,随着观看的进行,它会不断地刷新,以保持缓冲区中缓存的节目是当前播放时刻的前后一段时间的内容。当用户停止观看节目时,这个缓冲区的内容将被删除,空间将被释放。

### 3.1 面向内容的动态缓存机制的原理

VOD 系统中的内容有不同的分类,随着时间的推移,这些分类会被更新,每个分类中的内容也将被更新。各种分类和每个不同节目受欢迎的程度也就是热度是不同的,这意味着观看不同的节目内容的观众数量是大不相同的。在嵌入式的 P2P VOD 系统中,这一因素一般不会被考虑到,所有的节目都是同等对待。

由于嵌入式系统的缓冲区刷新速度比 PC 上的快很多,采用固定缓冲区模式的 VOD 系统的 P2P 效果不是很好。利用客户端之间互相传输数据来降低服务器端负载这一 P2P 的核心价值无法得到最大化的体现。为此我们设计了一个新的动态缓存机制来进行优化。

这种新机制以具体的节目内容为导向,客户端按照不同内容的热度来动态调整 P2P 缓冲,使嵌入式 P2P VOD 网络的性能得到提升。不同节目的热度等级被分为从 L1 到 L10。L10 意味着此节目内容是此时 VOD 播系统中最热门的。节目的热度等级很容易从服务器头端得到。在用户开始播放节目时,此热度参数随同详细的节目信息发送到客户端。

在新的机制中,系统的缓冲区被分为两部分:基础缓冲区(B)和扩展缓冲区(E)。扩展缓冲区由 10 个小缓冲区(E1 至 E10)组成。基础缓冲区是一个被用于 VOD 节目内容存储的固定大小的内容空间。扩展缓冲区是一个动态的缓冲区。扩展缓冲区的分配和释放依赖于所播放的内容的热度。当客户端在第一次开机时,系统的空闲内存将按照图 2 中所示的结构来进行划分。

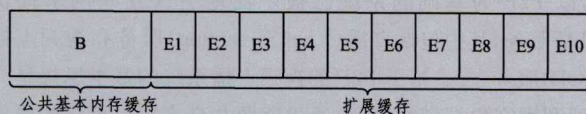


图 2 系统初始化时缓冲区的结构

如果用户开始观看热度为 L7 的节目 C1,编号为 E1—E7 的扩展缓冲区将分配给 C1,同时标记为被 C1 占用。这个时候,由于 E8—E10 这 3 个缓冲区为空,因此客户端还可以使用这 3 个缓冲区,但是不能将 E8—E10 标记为被 C1 占用,如图 3 所示。此客户端这时可以使用基本缓冲区 B 和指定的缓冲区 E1—E7 以及空闲缓冲区 E8—E10 来进行节目内容的下载、存储和分发,同时观看节目 C1。

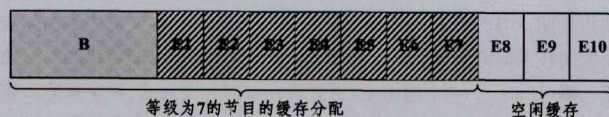


图 3 第一次观看一个节目时缓冲区的创建

### 3.2 动态缓存的更新机制

如果用户停止观看 C1,并开始观看一个新的节目 C2,其节目热度为 L2,缓冲区将被重新分配,如图 4 所示。当节目 C1 停止时,基础缓冲区 B 和 3 个未标记的缓冲区(E8-E10)将被立即释放并清空。然后,基本缓冲区被分配为 C2 占用。由于 C2 的热度只有 L2,因此只有 2 个扩展缓冲区被标记 C2。此时系统有 3 个未标记的缓冲区,其中 2 个将被分配给 C2,标记为 C2 占有,最后一个未分配的缓冲区将被暂时分配给 C2。重新分配后,基础缓冲区 B 和扩展缓冲区 E8-E10 将用于节目 C2 的缓存。

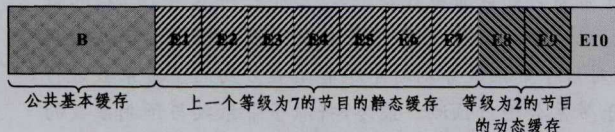


图 4 观看一个新的具有较低热度节目时的缓冲区重建

扩展缓冲区 E1-E7 将被保存在内存中,不会进行重新分配,数据不会清空也不会更新,它们将作为节目 C1 的一个 P2P 节目源,有助于此时 C1 这个热度很高的节目在 P2P 网络中的数据共享。这是这个面向内容的动态缓存机制的关键点。

如果用户停止观看 C2,并开始观看一个热度为 L4 的节目 C3,缓冲区将被再次重新分配,如图 5 所示。基本缓冲区和未标记的缓冲区(E10)将被立即释放。此时热度 L4 的节目 C3 需要 4 个扩展缓冲区,但是系统中只有一个未标记的缓冲区。在这个动态缓存机制中,缓冲热度相对较低的节目的缓冲区将优先释放,E8 和 E9 这两个标记为 C2 的将被首先释放并清空。但是 C3 还是需要一个扩展缓冲区缓冲。这种情况下,缓存有热度比 C3 高的节目的缓冲区中编号最后的一个缓冲区将被释放,E7 将被取消标记和清空。最后,基本的缓冲区,E7-E10 将被标记为被 C3 占用。在观看 C3 期间,E1-E6 的扩展缓冲区将保持不变,依然为 P2P 网络贡献一部分节目 C1 的内容。如果客户端开始观看一个热度高于 C1 的节目,在需要的情况下,被 C1 标记的缓冲区将被释放。

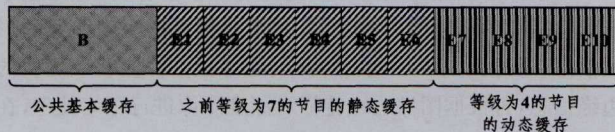


图 5 观看一个新的具有较高热度的节目时的缓冲区重建

为了方便对缓冲区进行操作,我们将缓冲区的数据结构定义如下:

```

typedef structure VOD_Buffer_S{
    Buffer_type {BASIC, EXTENDED};
    Buffer_index;
    Buffer_mark_flag {MARKED, UNMARKED};
    Buffer_empty_flag {OCCUPIED, EMPTY};
    Buffer_content_ID;
    Buffer_content_level;
    * memory_buffer;
} VOD_Buffer_t;
    
```

以此为基础即可设计出一个面向内容的动态缓存机制。

## 4 测试与结论

这个新的缓存机制设计完成后,我们在实际的系统中进行了验证。测试的嵌入式平台为基于 MTK8652 的 IPTV STB 平台。此平台配置有 256MB DDR 和 256MB NAND Flash 存储器。在系统正常开机,加载完应用程序后,系统可用内存接近 110MB。为此我们总共分配 100MB 的内存用于建立 VOD 的动态缓存机制。同时我们使用电脑模拟低配置的嵌入式平台,将可用内存限定为 100MB。通过 10 台嵌入式客户端以及 1000 个模拟客户端来对这个缓存机制进行测试。

测试中,我们定义热度的分级采用表 1 所列的原则。

表 1 节目内容热度分级原则

热度分级	用户收看比例	热度分级	用户收看比例
L1	<4%	L6	12%~14%
L2	4%~6%	L7	14%~16%
L3	6%~8%	L8	16%~18%
L4	8%~10%	L9	18%~20%
L5	10%~12%	L10	>20%

测试中,我们测试了在新的缓存机制下观看不同热度等级节目时,新加入的客户端的 P2P 网络建立时间与普通模式的对比,以及观看不同热度等级的节目时,P2P 网络数据共享率与普通模式的对比。这里我们定义的数据共享率是指观看某一个节目的所有用户下载该节目的数据大小与服务器实际发送所有这个节目的数据的大小的比值减 1。

图 6 所示为新加入客户端 P2P 网络建立时间的对比测试结果。从图中可以看出,对于低热度的节目,普通的缓存策略因为一视同仁的缓存机制,其 P2P 种子相对更多,建立时间短,但是随着节目热度的上升,新的动态缓存机制取得了更短的建立时间。

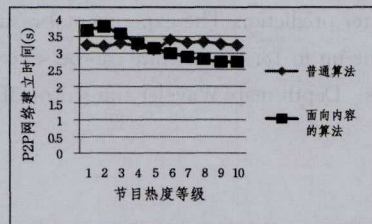


图 6 P2P 网络建立时间与节目热度关系的对比测试

图 7 为 P2P 网络数据共享率的对比测试结果。从图中可以看出,对于低热度的节目,由于用户数目少,P2P 数据共享率没有太大的差别。但是随着节目热度的升高,两种模式下的 P2P 数据共享率有提升,但是新机制下的共享率提升更快,而且绝对值也更大。

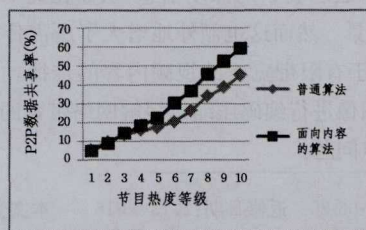


图 7 P2P 数据共享率与节目热度关系的对比测试

(下转第 163 页)

表 1 给出了不同特性序列在不同网络环境下的平均预测误差。其中 book 序列为运动缓慢序列, 相机间距 6.5cm; xmas 为中等运动剧烈序列, 相机间距 3mm; crowd 为运动剧烈序列, 相机间距 20cm; flamenco2 和 flamenco 为旋转运动序列, 相机间距 20cm。由表可知, 用 MSE 表征时, 左视点平均失真误差 4.11%, 右视点平均失真误差 4.21%; 用 PSNR 表征时, 左视点平均失真误差 0.82%, 右视点平均失真误差 1.04%, 证明本文所提模型适用于不同丢包率下不同视差大小和运动剧烈程度序列。

**结束语** 本文提出了一种递归式立体视频传输失真模型。模型能够精确预测终端的立体视频传输失真, 适用于不同运动剧烈程度和视差大小的视频序列, 且在不同的网络传输环境下均表现良好。模型可用于确定立体视频编码参数, 提高率失真优化性能, 优化立体视频传输机制(如拥塞控制), 指导信道的差错控制和终端错误隐藏技术, 为丢包网络环境下立体视频传输系统的进一步优化提供了有力保障。

### 参 考 文 献

[1] Zilly F, Kluger J, Berlin, et al. Production rules for stereo acquisition[J]. Proceedings of the IEEE, 2011, 99(4): 590-606

[2] Urey H, Erden E, Surman P. State of the art in stereoscopic and autostereoscopic displays[J]. Proceedings of the IEEE, 2011, 99(4): 540-555

[3] 同庆森, 李临生, 徐晓峰, 等. 视频跟踪算法研究综述[J]. 计算机科学, 2013, 40(26): 204-209

[4] Gurler, C G, Koc Univ E, et al. Flexible transport of 3-D video over networks[J]. Proceedings of the IEEE, 2011, 99(4): 694-707

[5] Thomas S, Tomas W, Stephan W. Optimized transmission of H. 26L/JVT coded video over packet-lossy network[C]//IEEE International Conference on Image Processing. NY, 2002: 173-176

[6] 唐振华, 刘文予, 赵爽, 等. 基于视频编码流的视频传输端到端失真度估算[J]. 华中科技大学学报: 自然科学版, 2008, 36(7): 25-29

[7] He Zhi-hai, Cai Jiang-fei, Chen Chang-wen. Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(6): 511-523

[8] Zhang Yuan, Gao Wen, Lu Yu, et al. Joint source-channel rate-distortion optimization for H. 264 video coding over error-prone networks[J]. IEEE Transactions on Multimedia, 2007, 9(3): 445-454

[9] Zhang Dong, Liang Jie, Singh I. Fast transmission distortion estimation and adaptive error protection for H. 264/AVC-based embedded video conferencing systems[J]. Signal Processing: Image Communication, 2013, 28(5): 417-429

[10] Tan A S, Aksay A, Akar G B, et al. Rate-distortion optimization for stereoscopic video streaming with unequal error protection [J]. EURASIP Journal on Applied Signal Processing, 2009, 2009: 7

[11] Xiang Xin-guan, Zhao De-bin, Ma Si-wei, et al. Rate-distortion optimization with inter-view refreshment for stereoscopic video coding over error-prone networks[C]//IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2009: 72570K-72570K-8

[12] Zhou Yuan, Hou Chun-ping, Pan R, et al. Distortion analysis and error concealment for multi-view video transmission[C]//2010 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting. Shanghai, 2010: 1-5

[13] Zhou Yuan, Hou Chun-ping, Xiang Wei. Modeling of transmission distortion for multi-view video in packet lossy networks [C]//Global Telecommunications Conference (GLOBECOM 2010). Maimi, 2010: 1-5

[14] Wang Yao, Wu Zhen-yu, Boyce J M. Modeling of transmission-loss-induced distortion in decoded video[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(6): 716-732

(上接第 133 页)

**结束语** 本文介绍了一种基于嵌入式平台的面向对象的 P2P VOD 缓存策略。由于嵌入式系统的性能和资源限制, 无法缓存大的 VOD 节目片段, PC 上的 VOD 系统中一些常用的 P2P 缓存策略效果不明显。本文提出的这个动态缓存机制结合了嵌入式平台资源有限的特点, 根据 VOD 系统中不同节目的热度来动态调整缓存结构, 最大限度地保障了 VOD 系统中热度高的节目资源尽可能多, 同时也让低热度的节目也有一个大小比较合理的节目缓冲区。该动态缓存机制较好地提高了 P2P VOD 系统的性能。

### 参 考 文 献

[1] Liu J, Rao S G, Li B, et al. Opportunities and challenges of peer-to-peer Internet video broadcast[J]. Proc. IEEE, 2007, 96(1): 11-24

[2] PPLive. Huazhong Univ. Sci. Technol., Wuhan, China [OL]. <http://www.pplive.com/>

[3] UUSEE. UUSEE, Inc., Beijing, China [OL]. <http://www.uusee.com/>

com/

[4] Jiang X, Dong Y, Xu D Y, et al. GnuStream: a P2P Media Streaming System Prototype[C]//Proceedings of International Conferences on Multimedia & Expo. (ICME'03). Maryland, USA, 2003

[5] Ripeanu M, Foster I. Mapping the Gnutella network: Properties of large-scale peer-to-peer systems and implications for system design[J]. IEEE Internet Computing Journal, 2002, 6(1): 50-57

[6] Hefeeda M M, Bhargava B K, Yau D K Y. A hybrid architecture for cost-effective on-demand media streaming [J]. Computer Networks, 2004, 44(3)

[7] Cui Y, Li B, Nahrstedt K. oStream: asynchronous streaming multicast in application-layer overlay networks[J]. IEEE Journal on Selected Areas in Communications, 2004, 22(1): 91-106

[8] Guo L, Chen S, Zhang X. Design and evaluation of a scalable and reliable P2P assisted proxy for on demand streaming media delivery[J]. IEEE Transactions on Knowledge and Data Engineering, 2006, 18(5): 669-682