

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

胡文彦 叶德建

(复旦大学软件学院, 上海 201203)

摘要 将应用层组播技术应用在高清流媒体直播系统中,设计并实现了一个支持大并发、高性能、小延时的高清流媒体直播原型系统,同时创造性地提出了分布式时移技术并集成在系统中,使用户能随时收看直播或延时直播的节目。该系统在2006年夏世界杯期间在复旦大学校园网内进行了长达一个多月的大规模测试,最高在线人数460。测试数据表明该系统很好地解决了服务器出口带宽瓶颈以及普通P2P技术延时较长等缺点,为国家科技部重点项目高清电视的直播系统提供了一条可行的技术方案。

关键词 应用层组播 流媒体直播系统 分布式时移

中图分类号: TP301.6 **文献标识码**: A **文章编号**: 1006-8961(2007)10-1720-05

Implementation and Testing of ALM Based HDTV Live Prototype System

HU Wen-yan, YE De-jian

(Software School, Fudan University, Shanghai, 201203)

Abstract In this paper, we applied ALM(application layer multicast) into HDTV live system, design and implement a multimedia live prototype system with large concurrency, high efficiency and low delay. We also created a distributed time-shifting algorithm, which was integrated in the system in order to allow users to watch both live and delayed TV programs. This prototype system was put into large-scale testing for more than one month in Fudan Campus during World Cup 2006. The result shows that it is a good solution to output bandwidth bottleneck at server side and long delay of normal P2P technology at the client. The maximum amount of online users was over 460. It also provides a good solution to HDTV live system, which is the focus of National Science and Technology Department.

Keywords ALM, multimedia live system, distributed time-shifting

1 引言

随着多媒体技术的发展和宽带网络的普及,流媒体直播系统已受到越来越多用户的青睐。目前,互联网上的各种基于P2P(Pear-to-Pear)流媒体直播系统可谓百家争鸣,百花齐放^[1-3],在工业界运营的这些直播系统的共同点是运用P2P技术,避免了服

务器端网络带宽有限造成的输出瓶颈;同时摒弃了单源父节点的简单传输模式,而选择多个不同的有效源节点,并且大大地增加客户端缓存来确保播放的稳定,对P2P流媒体技术的研究是当今学术界的热点之一^[4-6]。

与此同时,众多住宅小区网络结构的调整和优化使得家庭用户的接入带宽大大改善,而数字电视机顶盒的兴起,更是把高清数字电视推向了热点的

基金项目:国家自然科学基金项目(60503044)

收稿日期:2007-07-15;改回日期:2007-08-01

第一作者简介:胡文彦(1982~),男,复旦大学计算机软件与理论专业硕士研究生。主要研究领域为网络多媒体。E-mail: 042053019@fudan.edu.cn

最前沿。但是我们必须看到,能承载高清电视视频流传输的网络环境和公网的环境有着很大的不同,其网络带宽相当高,能支持高清电视每秒大于 2MB (H. 264) 的传输码率;并且网络应用较为单一,没有公网上各式各样传输流的干扰。在这种情况下,公网上通用的多源父节点、大规模客户端缓存等方法已不再适用。而基于网状拓扑的 P2P 技术的弱点是它需要客户端的长时间缓存来支持数据同步,势必造成播放延时,这是直播系统中用户所无法忍受的。因此,需要针对高清电视的需求来设计一个全新的大并发、无延时、高性能的直播系统。综合以上各种因素的分析,基于网状拓扑的应用层组播技术是高清流媒体直播系统较好的选择。

2 基于应用层组播技术的高清流媒体服务系统的设计和实现

2.1 支持直播服务的流媒体服务系统

在过去几年的工作中,我们设计开发并投入运营了支持直播和时移服务的流媒体服务系统^[7],该

系统在上海网络电视的大宁模式中取得了很大的成功,图 1 给出了通过单播链路提供直播、点播和时移服务的流媒体服务系统的体系结构。但是,该系统仍然存在着支持用户数量有限,专用高清电视网络带宽没有充分利用等缺点,并不能解决服务器端网络出口带宽有限这一瓶颈。在现有的服务器系统基础上,引入应用层组播技术,通过对系统的重新设计和实现,可以比较好地解决这一矛盾。

2.2 基于应用层组播的流媒体直播服务系统

把基于应用层组播的流媒体直播服务系统简称为 ALMS (Application-layer-multicast based live media service) 系统,ALMS 系统为每个直播频道都分别建立一棵组播树,将该频道所有用户都组织到这棵树中。由发布该频道媒体数据的播放服务器充当这棵组播树的根节点。每个用户在接收直播、无缓冲地播放音视频数据的同时也会根据系统的指示将数据转发给收看同一频道的其他用户。将流媒体直播编码服务器、ALM 服务器和播放服务器统称为 ALMS 的服务器端,ALMS 系统的体系结构如图 2 所示。

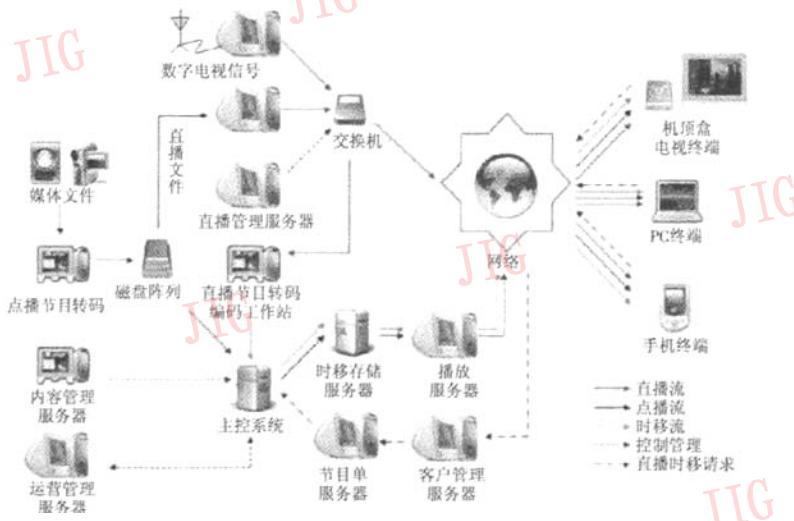


图 1 可提供直播服务的流媒体服务系统

Fig. 1 Live stream media server system

图 3 具体描述了 ALMS 系统的服务器端架构。播放服务器负责从直播编码器获取某一特定频道的数据,并对该频道媒体流的发布过程进行控制。系统初始化时,播放服务器向 ALM 服务器注册频道信息,然后它作为该频道组播树的根节点,根据 ALM 服务器的命令添加或删除用户节点,并将该频道的

媒体流转发给自己的每个子节点。

同时,图 3 给出了 ALM 服务器的各个组成模块以及它们之间的关系。ALM 服务器接收用户请求,维护用户信息,根据用户所要求的频道信息建立应用层组播树,并将用户在组播树中的位置、它们的子节点和父节点信息等数据发送给用户。

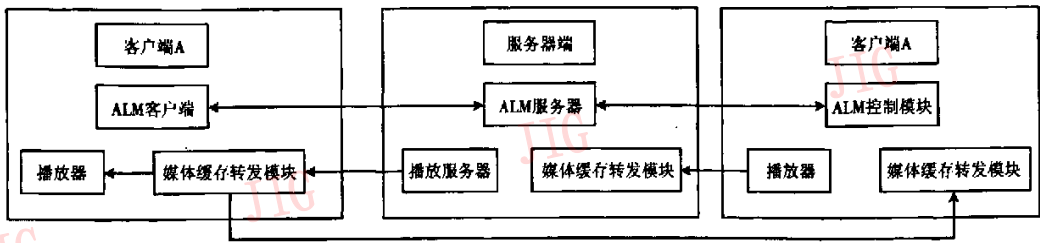


图 2 基于应用层组播的流媒体直播服务系统架构

Fig. 2 Architecture of ALM based stream media system

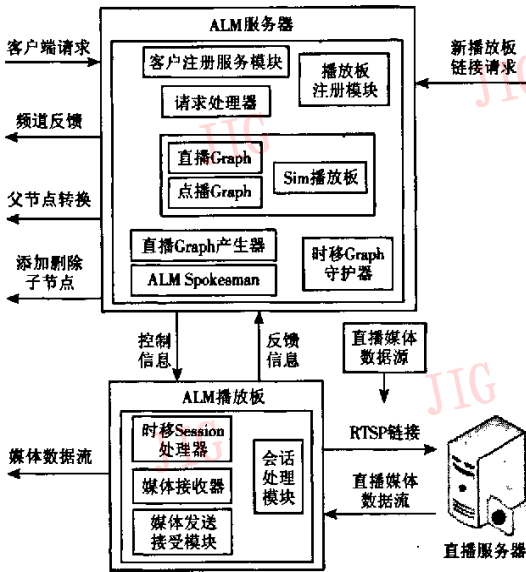


图 3 基于应用层组播的流媒体直播服务系统服务器端架构

Fig. 3 Architecture of ALM based stream media server

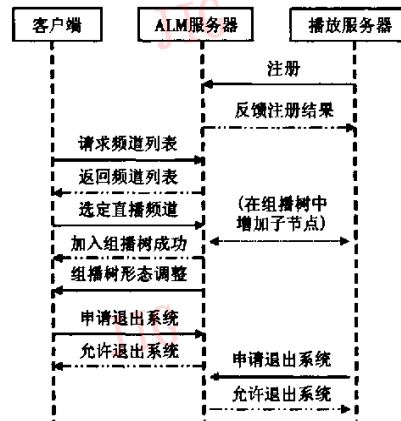


图 4 客户端、ALM服务器、播放服务器之间的交互过程

Fig. 4 Communication process among client, ALM server and playback server

2.3 基于应用层组播的流媒体直播服务系统的交互过程

在基于应用层组播的流媒体直播服务系统中，客户端软件在界面线程运行的同时，采用 TCP 链接和系统特定的通信协议与服务器端保持控制命令的通信，同时对自己的媒体接收端口进行监听和播放，收取媒体数据并转发给子节点。图 4 给出了客户端、ALM 服务器、播放服务器之间的交互过程。

2.4 分布式流媒体直播时移服务系统

时移服务系统建立在直播服务系统基础上。时移服务系统通过应用层组播方式为用户提供直播节目，将每个频道一定时间段内已经完成直播的媒体内容分散存储在该频道的各个用户节点上，同时在 ALM 服务器上建立用户节点存储的媒体内容索引。

为了保证能够存储所有的数据内容，时移服务系统仍然设立时移存储服务器，存储可供时移的所有媒体内容。时移服务系统中承担了媒体内容存储职能的用户被称为时移代理。

同时在 ALM 服务器中增加了时移控制模块。负责执行分布式存储算法，并对内容在组播节点上的分布进行控制并建立索引，处理用户的所有时移申请。作为时移代理的客户端程序增加了“本地存储/读取”功能，它监听 ALM 服务器发来的数据包，按照其要求在本地存储部分媒体数据；当 ALM 服务器指示由该时移代理对其他用户的时移申请进行响应时，它读取本地存储的对应数据并通过单播链路发送给有需求的用户，以减轻时移存储服务器对存储介质进行操作的压力和播放服务器通过单播支持时移操作的压力。

2.5 原型系统组播树组建算法

原型系统采用如下的函数考察当前组播树内节

点 h 作为新节点 h_{new} 的父节点时的性能:

$$g_1(h) = \|IP(h) - IP(h_{new})\| \quad (1)$$

$$g_2(h) = D(h)^a \quad (2)$$

$$g_3(h) = [d_{max}(h) - d_T(h) - 1]^b \quad (3)$$

$$g_4(h) = CPU(h) \quad (4)$$

$$g_5(h) = PL(h)^c \quad (5)$$

$$g(h) = \sum_i^5 w_i \cdot g_i(h) \quad (6)$$

其中, $IP(h)$ 和 $IP(h_{new})$ 分别为 h 和 h_{new} 的 IP 地址, $D(h)$ 为 h 在当前组播树中所处的层次, $d_{max}(h)$ 为 h 的度数上限, $d_T(h)$ 为 h 在当前组播树中的度数, $CPU(h)$ 为 h 当前的 CPU 占用率, $PL(h)$ 为从根节点到 h 之间链路的总丢包率。取使式(6)取值最小的那个节点作为 h_{new} 的父节点。该方法体现了对用户负载状况、组播树整体开销的关注。

2.6 原型系统组播树更新调整机制

系统客户端会周期性地(原型系统中为 10s)将所有的实时客户端信息反馈给服务器端,包括总丢包率、父子节点丢包率、CPU 占用率、内存使用量和所在组播树层次等。原型系统以此设立基于媒体服务质量和用户负载的组播树维护机制,通过考察用户接收到的媒体服务质量和用户负载,对组播树进行动态更新调整。

在原型系统记录的用户性能参数中,总丢包率是反映用户享受到的服务质量的重要参数。基于媒体服务质量的组播树更新调整机制的执行流程是:

(1)分析每个用户的性能参数,筛选出在刚结束的统计周期内总丢包率达到阈值(1%)的用户。

(2)若有用户连续若干个统计周期内总丢包率都达到这一阈值,就对树中从根节点到该用户的路径进行检查,考察每一跳的单链路丢包率。

(3)将从根节点到该用户路径上丢包率最大的那一跳的单链路中父节点作为调整对象,将它从当前组播树中删除,根据式(1)在被删除节点的子节点中选择一个代替该节点。

(4)被删除的节点重新执行加入组播树的操作。

基于用户负载的组播树更新调整机制的执行流程与基于媒体服务质量的机制类似。

通过组播树更新调整机制的处理,被删除的节点重新加入组播树后将作为叶子节点,不会承担为其他用户提供服务的职能,从而避免了对其

他用户的服务质量的影响。组播树在调整过程中只会影响被删除节点的后代,对组播树其他部分没有影响。

3 基于应用层组播的高清流媒体直播原型系统的运行测试

原型系统每运行一段时间后,都会对之前的运行数据进行检查,对系统中存在的问题进行修正,因此原型系统 33 天的持续运行时间被划分为 3 个连续时间段,其中最长的一个连续运行时间为 18 天 6 小时 50 分钟。

图 5 给出了这段时间内原型系统服务的用户数量。从图中可以看出,非比赛时段原型系统的用户总数都徘徊在比较低的水平上。在比赛时段,原型系统的用户数量多次超过 200 人的规模。如果通过单播方式提供服务,任何一场比赛的用户数量都不会超过 20 个,远远无法与原型系统相比。

图 6 则给出了 2006 年 7 月 10 日凌晨 0:50 开始的世界杯决赛在 5 个小时内 CCTV-1 频道的用

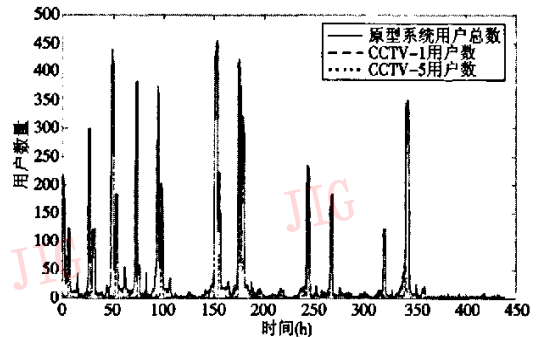


图 5 原型系统的用户数量

Fig.5 Number of users in prototype

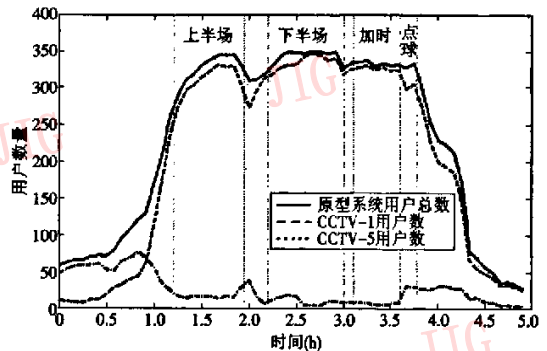


图 6 通过 CCTV-1 收看世界杯决赛的用户数量

Fig.6 User count of WC final match on CCTV-1

户数量变化。从图中可以看出,在比赛进行过程中,用户数量始终维持在较高的水平;当比赛进入中场休息时,用户数量会出现短暂的低谷过程。这表明原型系统能够容纳足够的用户,也能对用户数量在短时间内的剧烈变化进行良好的响应。

4 结 论

通过设计开发并运行测试了基于应用层组播技术的支持大并发、高性能、无延时的高清流媒体直播原型系统,证明层组播的技术应用在流媒体领域完全是可行的,并且能较好地解决服务器端网络出口带宽瓶颈以及直播节目延时的问题。分布式时移服务系统的设计和实现也在很大程度上减轻时移存储服务器对存储介质进行操作的压力和播放服务器通过单播支持时移操作的压力。

在今后的工作中,我们会继续改进该系统,使之真正成为可以投入运营的产品。

参考文献 (References)

- 1 pplive[OL]. <http://www.pplive.com>.
- 2 ppstream[OL]. <http://www.ppstream.com>.
- 3 coolstreaming[OL]. <http://www.coolstreaming.us>.
- 4 Hei X, Liang C, Liang J, et al. A Measurement Study of a Large-Scale P2P IPTV System[OL]. <http://cia.poly.edu/~ross/papers/P2PliveStreamingMeasurement.pdf>.
- 5 Wang Yan-li, Xian Ji-qing, Bai Jie. P2P based stream media techniques[J]. Computer Applications, 2005, 25(6): 1267 - 1270. [王艳丽, 鲜继清, 白洁. 基于 P2P 的流媒体技术[J]. 计算机应用, 2005, 25(6): 1267 - 1270.]
- 6 Wang Zhao. P2P system ALM based application in stream media[J]. Computer and Modernization, 2005, (2): 35 - 37. [王钊. 基于 P2P 系统应用层组播在流媒体中的应用[J]. 计算机与现代化, 2005, (2): 35 - 37.]
- 7 Xu Liang-ji, Ye De-jian. Design of stream media live (time-shift) system[J]. Journal of Image and Graphics, 2006, 11(1): 1516 - 1519. [徐良基, 叶德建. 流媒体直播(时移)系统的设计[J]. 中国图象图形学报, 2006, 11(1): 1516 - 1519.]