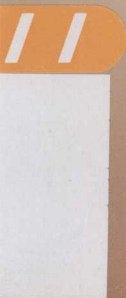




集群式流媒体服务系统

柴云鹏 著



煤炭工业出版社

013057718

TP37
201

集群式流媒体服务系统

柴云鹏 著



煤炭工业出版社

· 北 京 ·



北航

C1668158

TP37
201

图书在版编目 (CIP) 数据

集群式流媒体服务系统/柴云鹏著. --北京:煤炭工业出版社, 2013

ISBN 978-7-5020-4196-0

I. ①集… II. ①柴… III. ①多媒体技术 IV. ①TP37

中国版本图书馆 CIP 数据核字 (2013) 第 048221 号

煤炭工业出版社 出版
(北京市朝阳区芍药居 35 号 100029)

网址: www.cciph.com.cn
北京房山宏伟印刷厂 印刷
新华书店北京发行所 发行

*

开本 850mm × 1168mm $1/_{32}$ 印张 6
字数 142 千字 印数 1—800
2013 年 4 月第 1 版 2013 年 4 月第 1 次印刷
社内编号 7019 定价 30.00 元

版权所有 违者必究

本书如有缺页、倒页、脱页等质量问题, 本社负责调换

013027718

内 容 提 要

本书以集群式流媒体服务系统为中心，在探讨其体系结构、性能模型和仿真系统的基础上，重点介绍了集群式流媒体服务系统如何进行效率优化、性能优化和节能优化。

本书可供从事流媒体服务系统研究的科研人员、提供流媒体服务的企业管理人员及高级工程技术人员、高等院校计算机相关专业师生等参考。



计算机工业工程

第二版

前言

随着网络技术的进步，人们上网越来越方便，带宽成本也越来越低，因此从早期的 P2P (Peer-to-Peer) 视频直播到 YouTube、优酷网、土豆网和 CNTV 等在线视频服务，流媒体这个传说中的互联网“杀手级”应用终于真正流行起来。用户数的增加、服务质量的保障和高清视频的需求，使后端的流媒体服务系统承担着越来越大的压力。虽然 P2P 技术和 CDN (Content Delivery Network, 内容分发网络) 技术可以在一定程度上缓解核心流媒体服务器的压力，但面临日益增长的服务压力，流媒体服务的“强结点”仍是必需的。

多媒体领域的一些学者早已关注到了这个问题，例如清华大学的钟玉琢老师、杨士强老师和孙立峰老师等。流媒体服务系统其实是一个计算机体系结构和多媒体的交叉领域，而传统体系结构研究更关注通用系统的体系结构，对于专用领域关注较少。作为一名计算机体系结构的专业人员，我在李三立院士的鼓励 and 指导下，开始了集群式流媒体服务系统的研究工作，力求能提供一种高性能、低功耗、低成本的基于“强结点”的流媒体服务系统解决方案。

虽然集群技术已经是一项较为成熟的技术，经常用在大型数据中心和高性能计算机上，但是流媒体毕竟是一个数据和 I/O 操作都非常密集的应用，并且对服务质量的要求也非常高，集群技术并不是直接拿来就能造出很好的流媒体服务器，还需要针对流媒体应用的特征展开深入的研究工作。

在开展研究工作的过程中，我发现一个好的流媒体服务系统

的仿真工具是进行研究的必要条件，因为虽然基于实际系统的实验很重要，但仿真工具更为高效、快捷，并支持更大的参数配置范围。因此，在阅读了100篇以上的学术论文和技术规范的基础上，我首先建立了流媒体服务系统的性能模型，然后自己设计和开发了一套流媒体服务系统的仿真工具。相信这部分内容对打算进行流媒体服务系统研究或者设计的人员能有一定的帮助。

在此基础上，我开始探索集群式流媒体服务系统如何在效率、性能和节能方面进行优化。除了最容易理解也最受关注的性能优化外，效率优化和节能优化实际上也非常重要。效率优化指如何根据性能需求给出集群的最佳配置，这部分内容可供流媒体服务提供者参考；而高能耗问题在大型服务系统中日益凸显出来，除了带来高额的电费开销外，对空间、散热系统等都有很大影响，相信在今后几年会逐渐得到国内服务提供商的高度重视。

在进行本项研究期间，我所在的实验室正在承担一个教育部精品课程流媒体服务的项目，本书中介绍的研究内容有一部分应用在了这个项目中，并取得了很好的效果。

感谢清华大学的李三立、钟玉琢、史元春、杨士强、孙立峰和都志辉等老师对我的指导和鼓励！

感谢何川、顾雷、杜彬、王晓英、王坤、沈时军、张伟、张萌、肖鑫、罗建光、郝松、程志力、张天乐和叶银等同学对我的支持和帮助！

最后感谢我的家人对我这几年全身心投入工作的理解和支持！

著 者

2013年2月

目次

1 概述	1
1.1 发展集群式流媒体服务系统的必要性	1
1.1.1 流媒体应用的发展	1
1.1.2 P2P 和 CDN 等新技术能否取代流媒体服务系统	4
1.1.3 大容量流媒体服务系统面临的主要挑战	7
1.1.4 大容量的流媒体服务系统的解决方案	11
1.2 集群式流媒体服务系统的优势	13
1.3 集群式流媒体服务系统的性能优化的研究框架	17
1.3.1 集群式流媒体服务系统的基础研究	19
1.3.2 集群式流媒体服务系统的效率优化	20
1.3.3 扩展集群式流媒体服务系统的性能	21
1.3.4 集群式流媒体服务系统的节能优化	22
2 集群式流媒体服务器的体系结构和评价模型	23
2.1 引言	23
2.2 已有流媒体服务系统的体系结构	24
2.3 流媒体应用的特征及其对流媒体服务器 体系结构的影响	28
2.3.1 流媒体应用中资源消耗的特征	28
2.3.2 流媒体应用中用户行为的特征	32
2.4 集群式流媒体服务器 CSMS 的结构	34
2.4.1 CSMS 的逻辑结构	34

2.4.2	CSMS 的体系结构	34
2.4.3	CSMS 的简化结构和数据流向	37
2.5	集群式流媒体服务器的服务容量的评价模型	39
2.5.1	建立 CSMS 服务容量评价模型的重要性	39
2.5.2	CSMS 服务容量的评价模型	41
2.6	集群式流媒体服务器的服务容量的评测方法	51
2.6.1	绝对性能指标	52
2.6.2	效率指标	54
	本章小结	55
3	集群式流媒体服务器的仿真系统	57
3.1	引言	57
3.2	已有的流媒体服务器仿真系统	58
3.3	集群式流媒体服务器仿真系统 CSMSSim 的设计	59
3.3.1	CSMSSim 的体系结构	60
3.3.2	CSMSSim 的仿真模型	61
3.3.3	CSMSSim 的仿真功能	64
3.3.4	CSMSSim 的工作流程	64
3.4	CSMS 试验床构建及 CSMSSim 参数获取	68
3.4.1	CSMS 试验床构建	68
3.4.2	通过压力测试获取 CSMSSim 参数	70
3.5	集群式流媒体服务器仿真系统的应用	74
3.5.1	磁盘数据分布算法测评	75
3.5.2	内存缓存替换算法测评	77
	本章小结	79
4	集群式流媒体服务器的效率优化	81
4.1	引言	81

4.2	CSMS 的面向流媒体用户行为的非对称 体系结构 SUBOA	82
4.3	CSMS 的资源配置优化	87
4.3.1	CSMS 的资源配置优化的可行性	88
4.3.2	CSMS 资源配置优化问题的解决方案—— 分步优化	88
4.3.3	CSMS 的分步资源配置优化算法	90
4.4	CSMS 分步资源配置优化算法的具体案例分析	94
4.4.1	确定 A_{AM}^{sch} 、 N_{AM} 和 N_{AM}^{oNIC}	94
4.4.2	缩小 A_{AM}^{cache} 和 MS_{AM} 的范围	97
4.4.3	缩小 A_{SM}^{cache} 和 MS_{SM} 的范围	102
4.4.4	缩小 A_{SM}^{alloc} 和 N_{SM} 的范围	103
4.4.5	确定 Set_A 和 Set_N 中的所有参数	105
	本章小结	111
5	集群式流媒体服务器的性能优化	112
5.1	引言	112
5.2	集群式边缘流媒体服务器 CESS 的体系结构	114
5.2.1	CESS 的外部网络环境	115
5.2.2	CESS 的内部结构及工作流程	116
5.3	集群式边缘流媒体服务器的负载均衡问题	117
5.3.1	流媒体边缘缓存服务器负载分析	117
5.3.2	流媒体边缘缓存服务器负载平衡解决方案	118
5.4	支持负载均衡的多层缓存替换算法 MCLBS	119
5.4.1	缓存增益模型	119
5.4.2	MCLBS 缓存替换算法	120
5.5	MCLBS 算法的性能评测	121
5.5.1	实验平台	121
5.5.2	实验结果及分析	121

本章小结	124
6 集群式流媒体存储系统的节能优化	126
6.1 引言	126
6.2 集群式流媒体存储系统节能问题的 分析和模型化	128
6.2.1 节能的形式化	129
6.2.2 服务质量的形式化	130
6.2.3 流媒体系统中节能问题的分析	131
6.3 负载自适应的磁盘分组和数据交换的 节能模型	137
6.3.1 LADGDE 在线算法	140
6.3.2 LADGDE 不在线算法	145
6.3.3 节能数据分块策略	150
6.4 LADGDE 算法的性能评测	152
6.4.1 仿真平台	152
6.4.2 基于参数化用户记录的实验结果	154
6.4.3 基于实际系统用户记录的实验结果	157
本章小结	166
7 结论	167
7.1 总结	167
7.2 需进一步开展的工作	169
参考文献	171

1 概 述

1.1 发展集群式流媒体服务系统的必要性

1.1.1 流媒体应用的发展

近年来,流媒体应用发展迅速,在音视频点播、远程视频教育、视频会议、军用监控和卫星观测等众多领域都得到了广泛的应用,尤其在 IPTV 和 3G 手机领域形成了巨大的产业和市场。流媒体应用中规模最大的三类用户分别是:互联网用户、家庭电视用户和移动电话用户。下面分别介绍国内和国外这三类流媒体用户的情况。

1. 互联网

在互联网方面,截至 2012 年 6 月底,中国网民数量达到 5.38 亿人,稳居世界第一位,其中手机网民规模已达 2.88 亿。

表 1-1 展示了各种网络应用使用率的排名情况。网络应用使用率指的是在所有网民中使用这种应用形式的用户所占的比例。在各类应用中,属于流媒体应用的网络音乐(音频直播和点播)和网络视频(视频直播和点播)分别排名第三和第六。网络音乐的使用率达到 76.4%,是中国网民的重要娱乐方式,也是促进新增网民上网的重要驱动力之一,对中国互联网的推广功不可没。网络视频也是中国网民通过互联网娱乐的重要方式之一,其使用率为 65.1%,用户量已经达到 3.25 亿人。

同时,互联网上也涌现了一批非常成功的流媒体服务产品。在 P2P 音视频直播方面,最成功的有 pplive 和 ppstream 等,目前, pplive 日常在线用户 120 万,软件的月覆盖用户 2000 万。近几年, pplive、ppstream 和 Joost 等产品也为用户提供了 P2P 的视

表1-1 网络应用使用率排名

排 名	网 络 应 用 名 称	使用率/%
1	即时通信	82.8
2	搜索引擎	79.7
3	网络音乐	76.4
4	网络新闻	73.0
5	博客/个人空间	65.7
6	网络视频	65.1
7	网络游戏	61.6
8	微博	50.9
9	电子邮件	48.1

频点播服务。在视频点播网站方面,全球最大的视频网站 YouTube 的 Alexa 世界排名第 3 位,中国最大的两个视频网站——优酷网和土豆网的 Alexa 世界排名分别达到第 53 位和第 59 位。

2. 数字电视和 IPTV

家庭电视用户收看流媒体音视频内容主要可以通过两个途径,一是数字电视,二是 IPTV。

在数字电视市场上,截至 2008 年 11 月中国数字电视用户已经超过 4800 万户,比 2007 年底增长了 2000 万户。而数字电视在法国的渗透率达到了 66%,在英国达到 86%。英国是欧洲数字电视发展最快、渗透率最强的国家。

另外, IPTV 也取得了迅速的发展。微软公司宣布:不到一年的时间里,采用微软 Mediaroom IPTV 平台和多媒体软件平台的客户几乎翻了一番,达到 200 万个用户,并支持全球将近 400 万个机顶盒的业务。香港电讯管理局最新公布的年报显示,香港的宽频电视服务(IPTV)住宅用户达 110 万户,普及率为 48%,居全球首位。西欧市场在 2007 年的 IPTV 订户超过了 690 万户,

全球 IPTV 订户达到了 1234 万户。据赛迪顾问研究数据, 2008 年第三季度中国 IPTV 用户数已达到 205.4 万户, 2011 年中国 IPTV 用户数已超过 1350 万户。

3. 移动电话

手机上网的用户群已经初具规模, 目前中国 5.38 亿网民中, 使用手机上网的网民数已经达到 3.88 亿人, 占全部网民数的 72.12%。而随着 3G 的迅速发展和“千元智能机”的出现, 大幅降低了移动智能终端的使用门槛, 手机上网的用户数量, 以及服务的内容和质量都得到了迅速的提高。中国电信 3G 用户数约 526.2 万户, 居国内三家 3G 运营商之首。另外, 根据 Nielson 公司的统计, 美国用户在 2007 年和 2008 年这三类用户的规模和月平均观看时间分别见表 1-2 和表 1-3。

表 1-2 美国月平均用户数

方 式	平均用户/千人		增长百分比/%
	2008 年 5 月	2007 年 5 月	
在家看电视	282348	278400	1.4
观看时移电视	63265	46632	35.7
使用互联网	161690	154134	5
在互联网上观看视频	118673	—	—
使用移动电话	217076	—	—
使用移动电话观看视频	4400	—	—

表 1-3 美国用户月平均观看时间

方 式	平均观看时间		增长百分比/%
	2008 年 5 月	2007 年 5 月	
在家看电视	127 h 15 min	121 h 48 min	4
观看时移电视	5 h 50 min	3 h 44 min	56

表 1-3 (续)

方 式	平均观看时间		增长百分比/%
	2008 年 5 月	2007 年 5 月	
使用互联网	26 h 26 min	24 h 16 min	9
在互联网上观看视频	2 h 19 min	—	—
使用移动电话观看视频	3h 15 min	—	—

其中时移电视 (Time-shifted TV) 是 IPTV 的一种, 是在交互式宽频网络上实施的一种崭新的电视节目服务形式。传统电视的特点是固定时间、固定频道、单向广播; 而时移电视的特点是用户可在任意时间收看任意频道中的任意节目或片段, 可以像影碟机、录像机一样对收看的电视节目实行暂停、快进、快退等功能操作。

目前虽然观看传统电视的用户还非常多, 但是互联网上的流媒体用户已经有很大的规模。随着技术的进步和成本的下降, 会有越来越多的人通过数字电视、IPTV 和移动电话等形式观看流媒体内容, 所以流媒体用户的增长将会非常迅速, 用户对服务质量的要求也会越来越高。

1.1.2 P2P 和 CDN 等新技术能否取代流媒体服务系统

随着流媒体用户数量的增加和用户对服务质量要求的不断提高, 后端的流媒体服务系统必然将承受前所未有的巨大压力。然而, P2P (Peer-to-Peer) 和内容分发网络 CDN (Content Delivery Network) 等一些新技术的出现对于缓解后端流媒体服务系统的压力起到了一定的作用, 但这些新技术的广泛应用, 能否彻底取代高性能的流媒体服务系统呢?

实际上, 这些新技术虽然在一定程度上能够缓解流媒体服务系统面临的压力, 或者有助于流媒体服务系统内部结点之间的负载均衡, 但是它们并不能取代流媒体服务系统; 相反, P2P 技术

和 CDN 技术正需要高性能的流媒体服务系统为其提供支持。

1. P2P 技术

在目前的网络大环境中, P2P 技术很难完全通过用户之间的共享自给自足, 这是由以下两个方面的原因造成的:

首先, P2P 需要服务系统提供大量的上行带宽。一般来说, 在互联网中有上行带宽和下行带宽的概念, 这两个概念都是针对互联网的终端用户而言的。上行带宽指用户将内容上传到互联网上所消耗的带宽, 而下行带宽指用户从互联网上下载内容所消耗的带宽。在互联网上, 非对称数字用户线路 ADSL (Asymmetric Digital Subscriber Line) 得到了广泛的应用, 例如在世界上互联网用户最多的中美两国, ADSL 所占的比例就分别达到 73% 和 50% 左右。而 ADSL 的特点是用户端的下行带宽远远大于上行带宽, 例如目前常见的 ADSL 设备支持上行 640 kbps、下行 8 Mbps 的数据传输率。因此, P2P 系统一般不能从广大普通用户处得到足够的上行带宽, 这样用户总的下行带宽和总的上行带宽之间的缺口就迫切需要高性能的流媒体服务系统来弥补, 这样才能维持上行带宽和下行带宽之间的整体平衡。

其次, P2P 系统中存在着用户共享效率和骨干网带宽消耗之间的矛盾。P2P 技术通过用户之间的互相传播在很大程度上减小了服务器带宽的消耗, 对流媒体应用的推广起到了巨大的作用。但是, 这种数据传输方式会造成很多数据在骨干网上的重复传播。例如在图 1-1 所示的情况中, 分别处于互联网不同边缘的三个用户 A、B 和 C 如果进行通信, 那么需要通过骨干网络路由器。假如三个用户属于同一个 P2P 网络, 用户 B 和 C 需要从用户 A 处得到同一个数据包 D, 那么同一份数据包 D 会经过骨干网络路由器传输两次, 这种重复传输如果在 P2P 系统中不做限制, 那么会频繁发生, 从而耗费掉大量宝贵的骨干网络带宽。实际上, 一方面可以利用 CDN 技术在互联网边缘设置缓存服务器缓存热门内容; 另一方面, 通过对用户间的 P2P 行为进行限制,

尽量只在临近用户之间进行数据共享,可以很大程度上避免骨干网络上的数据大量重复传输的现象。但是用户共享范围的缩小也就意味着用户间资源共享率的降低,那么相对来说,用户群体能提供的上行带宽更小了,也就是服务器面临的带宽压力也就增大了。另外,近年来流媒体点播应用迅速发展,正在逐渐超过甚至取代流媒体直播应用,点播应用中同时观看相同内容的用户比直播要少很多,因此用户共享效率要下降很多,这也大大加剧了P2P系统对服务器带宽的依赖性。

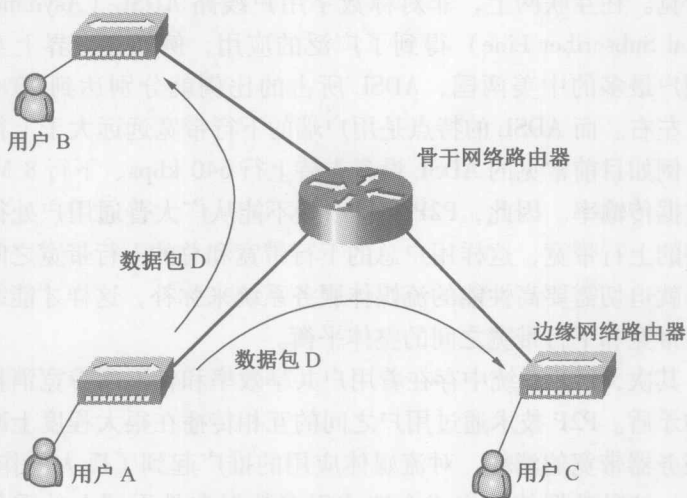


图 1-1 P2P 行为对骨干网带宽浪费情况示意图

通过以上的分析可知,P2P 技术的确能够通过用户之间的共享行为节省大量的服务器带宽,但是仅仅靠 P2P 技术并不能完全通过利用用户的带宽而消除对服务器的依赖;相反,仍然必须有大量的服务器带宽的支持才能维持整个 P2P 流媒体服务系统的正常运转和流媒体用户获得的服务质量。所以,P2P 技术并不能彻底取代流媒体服务系统。

2. 内容分发网络技术

CDN 技术指的是通过在现有的互联网中增加一层新的网络架构,即通过设置大量互联网边缘缓存服务器,将网站的内容发布到最接近用户的网络边缘,使用户可以就近取得所需的内容,解决互联网的网络拥挤问题,提高服务系统的服务能力和服务质量。由于大多数用户请求都可以通过边缘缓存服务器得到满足,因此 CDN 技术可以明显地减小骨干网络的流量,并且可以大大减少应用的核心服务器的压力。但是 CDN 技术实际是一种在服务器间分配负载的技术,它仍然需要强大的基础服务器来提供大量的计算资源、网络带宽资源、缓存资源和存储资源等,无论是在核心服务器上,还是在边缘服务器上。随着流媒体应用中用户数量的增长、流媒体视频内容的质量和数量的迅速提高,高性能的基础服务器的需求将会越来越强烈。

总的来说,P2P 和 CDN 技术能够对流媒体服务系统起到很好的辅助作用,例如 P2P 能够通过用户间的共享减小流媒体服务系统的负载,而 CDN 能够通过核心服务器和边缘缓存服务器的方式在流媒体服务系统内部做负载均衡,同时也有助于改善用户获得的服务质量。但是,P2P 和 CDN 等新技术的出现和发展,并不能消除对高性能的流媒体服务系统的需求,还是需要构建高性能、高扩展性和高效率的大容量流媒体服务系统。

1.1.3 大容量流媒体服务系统面临的主要挑战

随着流媒体用户规模的增长,以及用户对服务质量要求的逐渐提高,迫切需要大容量的流媒体服务系统来满足用户的需求。而且相对于传统服务,流媒体用户的行为特征更加复杂和多变,如流媒体应用中用户对系统的平均占用时间非常长,而且对服务稳定性的要求非常高;流媒体用户的同时在线人数一般呈现出规律性的剧烈变化,一天之内的最多和最少在线人数相差很大;流媒体用户的访问兴趣一般都集中于少量内容,但是用户的访问热点却一直在实时的转移,并不固定;而且绝大多数流媒体用户会