

文章编号: 1007-2861(2010)01-0091-06

P2P 流媒体直播系统的组播树构建及设计

谈广云, 许华虎, 高珏

(上海大学 计算机工程与科学学院, 上海 200072)

摘要: 研究了 P2P 流媒体直播系统的特性和体系结构. 对组播树的构建与设计进行了详细的说明, 提出一种单播树、多播树结合的组播树构建方案, 该方案结合了单组播树维护方便、多组播数节点负载均衡的优点. 对节点组织策略进行了分析, 对新加入节点利用及时服务策略进行了研究, 提出一种及时优化策略. 在新节点加入的时候就探测其服务能力, 调整组播树结构, 及时发挥新加入节点的优势, 提升系统服务能力.

关键词: P2P; 流媒体; 组播树

中图分类号: TP 393

文献标志码: A

Construction of Multicast Tree of P2P-Based Live Media Streaming and System Design

TAN Guang-yun, XU Hua-hu, GAO Jue

(School of Computer Engineering and Science, Shanghai University, Shanghai 200072, China)

Abstract: This paper studies characteristics and architecture of P2P streaming broadcast systems. Construction and design methods of multicast tree are first discussed. A scheme combining unicast and multicast trees is proposed, which combines advantages of easy maintenance of unicast tree and load balancing of multicast tree. Based on analyses of the strategies of node organization and timely service of new nodes, an optimized strategy for immediate service is introduced. When new nodes are added, the system will detect their abilities of providing services, and adjust the multicast tree structure to utilize the ability of the new nodes, thus enhance the system capacity.

Key words: P2P; streaming media; multicast tree

随着互联网技术的飞速发展,人们接收信息的方式发生了巨大的改变.由单纯的文字、音频信息,变为集声音、文字、视频于一体的多媒体信息.多媒体传输技术现已成为研究的热点,其中流媒体直播技术因内容丰富性、及时性,已成为多媒体传输的重要手段之一.根据国外相关调查表明,有31%的上

网用户经常使用流媒体直播.

传统的C/S模式,由于其服务规模小、服务器性能依赖性强以及需要大量带宽等缺陷,已淡出流媒体的应用领域.在P2P系统中,每个peer既享受服务,又提供服务,将服务器的负担分布在各个peer中,从而减轻了服务器的负载和网络带宽的占用,提

高了系统的扩展性.因此,P2P技术的应用已成为流媒体直播技术研究的趋势,并得到研究者的广泛重视.

当前,P2P流媒体直播系统已经在网络电视、视频会议、视频监控等方面得到了普遍应用,通过网络观看视频及电视节目已经变得越来越流行.P2P直播技术被大规模应用的同时,也面临着以下几个问题:

(1)如何进一步提高图像的质量,为用户带来更好的视觉感受.现行的直播系统在寻求低码率的同时,牺牲了图像的质量,所以寻找一种高效率又能保证图像质量的编解码方案,已经成为一个很紧迫的任务.

(2)如何构建高效的组播树,降低视频流在节点间的转发次数,提高整个直播系统的稳定性,这是P2P直播技术面临的最大挑战.现行的直播系统都不能在低延迟和画面的流畅性这两个方面同时做得很好.

1 P2P流媒体直播系统

1.1 P2P流媒体直播系统的特性

与P2P资源共享系统相比,P2P直播系统具有较高的实时性、稳定的传输速率、较高的码率等特性.在资源共享系统中,需要用几天的时间下载一部影片,而在直播系统中,只需一个小时或者几个小时下载这些数据,所以维持高速率的数据传输是直播系统的一个特性.低延迟是直播系统追求的一个目标,没有人愿意打开一个软件然后等上几分钟才看到节目.从直播技术诞生以来,低延迟一直是研究的热点.这其中涉及许多方面,其中最关键的是如何构建高效的组播树.高效的组播树要做到,节点的频繁加入和退出不会影响到系统的服务质量,树的高度要控制在最低的范围内,数据流的转发次数要控制在最少的范围内,节点的出度、入度要控制在最优的范围内.

1.2 P2P流媒体直播系统体系结构

P2P流媒体直播是指基于应用层组播且有同步时序要求的流媒体技术.其工作原理是:视频源服务器采集视频信号,对视频数据压缩、编码,然后将数据推送到由Peer节点组成的覆盖网络中,通过在线的Peer节点对媒体数据进行中继传输,Peer节点收到编码数据之后解码观看.P2P流媒体直播通常采用的拓扑有单组播树拓扑、多组播树拓扑和网状拓

扑.单组播树结构在初始的P2P流媒体直播系统中得到了大量应用,因为它拓扑结构简单,系统维护代价小.但是它过分依赖组播树的高层次节点的性能,某个节点的离开或者失效对其子孙节点有严重的影响,并且叶子节点的资源没有被利用,所以P2P流媒体直播技术逐渐向基于多组播树结构过渡.在多组播树结构中,一个节目源在多棵单组播树中转发,节点依据自身的带宽能力,从一棵或者多棵组播树中接收数据.多组播树方案解决了单组播数结构不够稳定,以及过分依赖单个节点的问题,但是其节点组织比较复杂,维护代价高.故当前已部署P2P直播系统都采用基于网状拓扑结构.在新节点加入时,管理服务器会从当前节点中挑选出几个Peer节点作为其邻居节点,然后该节点通过相关的调度算法从邻居节点获取数据.由于可以从多个节点获得数据,因此可以较好地解决节点扰动对系统造成的影响,系统的鲁棒性相对较高.其中选取邻居节点以及降低启动延时是网状拓扑结构的技术关键.

2 组播树的构建与设计

2.1 单组播树结构

单组播树结构中,最具代表性的就是NICE,Zigzag组播树构建方案.二者的思路都是分层和分群,成员只和少量固定的节点联系.NICE(见图1)的节点组织具有层次性,节点的维护代价小,节点加入和退出对管理系统的扰动较少,越高层次的节点对系统的服务质量和稳定性越重要.高层次节点的失效和退出对其子孙节点的影响很大.当系统规模很大时,高层次节点的服务能力成为系统的瓶颈.Zigzag(见图2)节点组织方式和NICE一样,不过它把数据转发和节点管理从同一个节点上分离出来,由一个节点专门负责节点Cluster的管理,从而解决了NICE系统瓶颈问题.改进后组播树中节点所带子节点数目是一定的,与参与组播树的节点数目无关.

2.2 多组播树结构

在多组播树结构中,最具代表性的就是CoopNet,SplitStream多组播树构建方案.在SplitStream方案中(见图3),视频流被分割成多个视频段,每个视频段在一个组播树中传输.每个节点只在一个组播树中转发数据,在别的组播树中作为叶子节点.这种方法平衡了系统负载,增强了系统稳定性,单个节点的离开或者失效对各节点影响有限.在

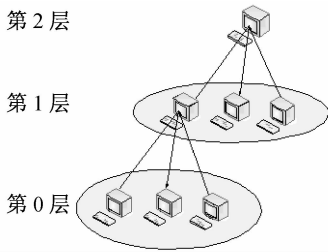


图1 NICE 数据转发模型

Fig.1 NICE transmitting-data model

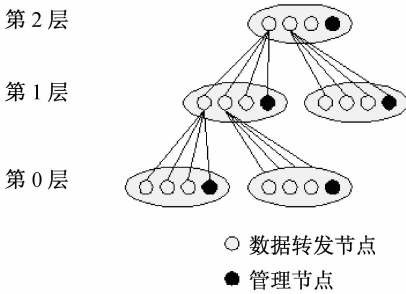


图2 Zigzag 数据转发模型

Fig.2 Zigzag transmitting-data model

CoopNet 方案中,通过 MDC 多描述编码,把视频流分成多个视频段,每个视频段在一棵组播树中传播.它和 SplitStream 的区别为此方案采用了服务器/客户端模式优先策略.当服务器没达到最大负载时,服务器直接向节点提供服务,否则由享有服务的节点提供服务.多组播树结构平衡了节点负载,增强了系统稳定性,但管理系统要维护多棵组播树,开销很大,同时多路传输时数据的同步很难保证.

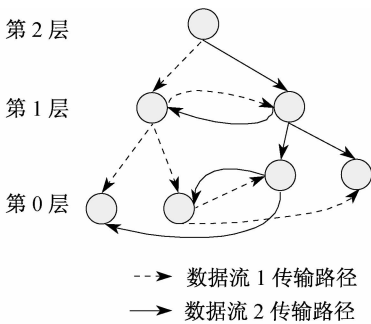


图3 SplitStream 数据转发模型

Fig.3 SplitStream transmitting-data model

2.3 组播树构建方案

单树结构的节点组织简单,维护开销小,但高层次节点退出或失效对子孙节点影响较大;多树结构平衡了各个节点的负载,单个节点的失效影响少,但

管理系统需要维护多棵树,系统开销较大.本研究提出一种结合单树和多树优点的组播树构建方案(见图4),整个结构可由A层区的单树结构加上B层区的多树结构组成.A层区的节点按照属性(ISP、城市、区)相似原则来划分层次,相同属性的集合处于同一层次.B层区的节点按照其服务能力(能提供的带宽、数据转发的延迟)划分各个层次,层次越高的节点,服务能力越强.系统的工作原理是:视频流由视频源服务器(SS)转发给ISP代表节点,ISP代表节点把数据转发到具有相同ISP属性的城市代表节点,城市代表节点把数据转发到具有相同城市属性的区代表节点.区代表节点根据相关算法把视频流分割成各个视频段,然后转发给B层区.B层区的Cluster为每个视频段建立组播树,构成多组播树网络.B层区各个Cluster中的节点具有ISP、城市、区3个相同的属性,各个节点在物理和逻辑路径上相近,数据转发的延迟较小.节点经过ISP、城市、区这3个属性分层,Cluster节点数目已经不是很多,构建的多树结构不会很复杂,管理开销不大.B层区中的多树结构是由Cluster中的区代表节点管理,不会增加管理源服务器(GS)的开销,在弥补了多树结构开销大的缺陷的同时,还降低了节点之间的数据转发延迟.

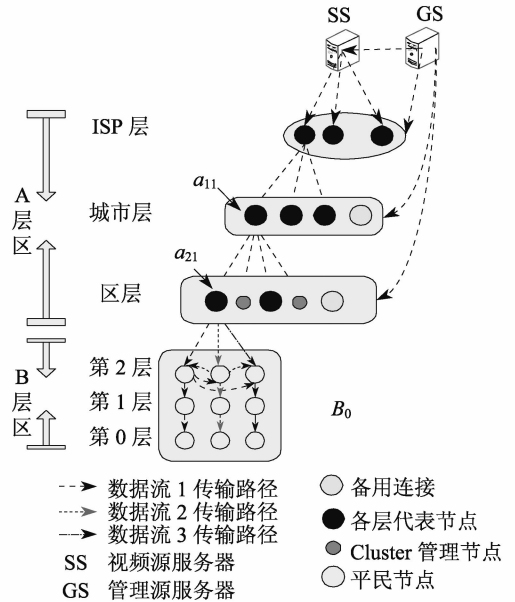


图4 组播树构建方案

Fig.4 Scheme of constructing multicast tree

2.4 节点组织策略的设计

2.4.1 按属性相似原则划分 Cluster

当新节点申请加入的时候,管理服务器会根据

该节点提供的物理位置信息和网络信息为该节点分配一个ID,该ID的组成如图5所示.

IP类型	公网IP	内网IP	ISP	城市	区	TCP端口	UDP端口
------	------	------	-----	----	---	-------	-------

图5 ID组成结构

Fig.5 Composition of ID

管理服务器根据ID,将该节点分配到与该节点有相同ISP、城市、区属性的Cluster中.管理服务器会从Cluster中选出一个区代表节点,作为该Cluster数据转发的支点.具有相同城市属性的区代表节点组成区层的Cluster,管理服务器会从区层各个Cluster中选出一个有优势的节点作为该Cluster的城市代表节点,作为具有相同城市属性的节点的数据转发的支点.同理,城市层的Cluster的组成方法与代表节点的选择方法和区层相同.

2.4.2 新节点的加入策略

现行的P2P直播系统对新加入节点的服务能力有一个反应延迟,即新节点在加入系统的一段时间内只接受服务,不提供服务,系统忽略了它的服务能力,如图6所示.

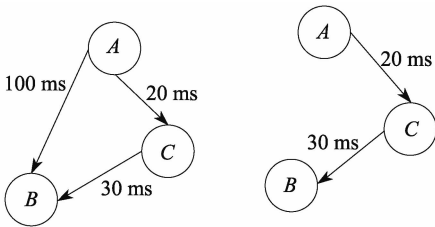


图6 A, B, C关系图

Fig.6 Relationship of A, B, and C

A→B的延迟为100ms,新加入节点C,A→C的延迟为20ms,C→B的延迟为30ms,B通过C获得服务比之前快50ms,但是系统并没有发现C的优势.因此,本研究提出一种及时优化策略,在新节点加入的时候就探测其服务能力,调整组播树结构,及时发挥新加入节点的优势.例如,新节点e申请加入,节点调度步骤如下:

- ① 管理系统根据属性相似原则把e划分到B₀中(见图4);
- ② e作为a₂₁(B₀区代表节点)的兄弟节点加入系统,e和a₂₁都接受a₁₁(城市代表节点)提供的服务;
- ③ e探测自身到B₀中所有节点的延迟D_i,求出

$$\text{平均延迟 } D_{e_0} = \sum_{i=1}^n D_i/n;$$

- ④ e探测自身到a₁₁的延迟D_{e1},然后计算出服务延迟D_e=D_{e0}+D_{e1};
- ⑤ 用D_e和D_{a21}作比较,如果D_e<D_{a21},并且D_e-D_{a21}<ξ(ξ为实验得出的值),转⑥,否则转⑨;
- ⑥ 用e节点代替a₂₁,作为区代表节点;
- ⑦ a₂₁作为平民节点,下放到B₀中;
- ⑧ 依据B₀中所有节点的服务能力重新划分层次,转⑩;
- ⑨ 把e作为平民节点,下放到B₀中,转⑧;
- ⑩ 结束.

在这种情况下,因为新加入的节点e有优势,能够立即为其他节点提供服务,从而使整个系统的服务能力得到提升.

2.4.3 试验数据分析

本研究用16台PC模拟要加入系统的节点.这16台PC处于同一个局域网中,用发包延时来模拟各个节点之间的通信延时.各个节点之间的发包延时时间如表1所示.

节点加入系统的顺序为N₉,N₂,N₅,N₆,N₄,N₃,N₇,N₁₁,N₁₃,N₁₅,N₁,N₁₂,N₁₄,N₈,N₁₀,N₁₆,每个节点的出度小于等于4.在未使用及时服务策略时,系统所构建的组播树如图7所示.

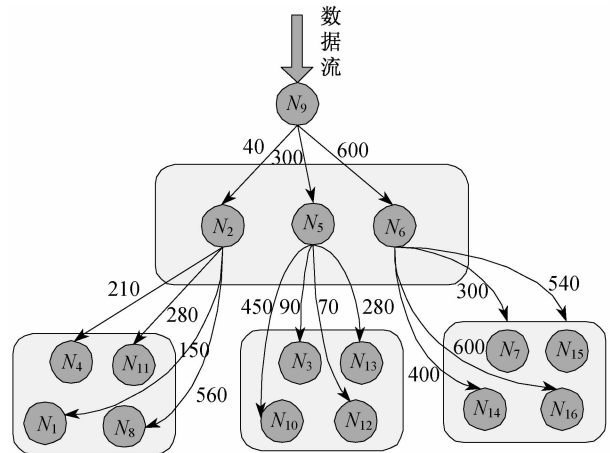


图7 未使用及时服务策略构建的组播树

Fig.7 Constructing multicast tree without providing services immediately strategy

16个节点的平均延时为

$$\bar{D} = \sum_{i=1}^{16} D_i/16 = (40 \times 5 + 300 \times 5 + 600 \times 5 + 210 + 280 + 150 + 560 +$$

$$450 + 90 + 70 + 280 + 400 + 600 + 300 + 540) / 16 = 439.$$

表 1 节点发包延时表

Table 1 Delay of nodes sending data packets

节点	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}	N_{13}	N_{14}	N_{15}	N_{16}
N_1	0	150	500	50	100	200	40	350	120	70	100	800	200	60	140	300
N_2	150	0	60	210	90	320	180	560	40	110	280	30	210	70	90	200
N_3	500	60	0	130	90	80	20	60	70	90	120	110	10	25	60	90
N_4	50	210	130	0	200	300	150	150	180	320	240	180	100	290	380	200
N_5	100	90	90	200	0	200	60	40	300	450	60	70	280	180	320	160
N_6	200	320	80	300	200	0	300	400	500	320	190	230	340	400	540	600
N_7	40	180	20	150	60	300	0	100	200	210	120	230	320	340	80	130
N_8	350	560	60	150	40	400	100	0	100	30	20	40	50	90	70	10
N_9	120	40	70	180	300	500	200	100	0	220	20	40	20	320	450	60
N_{10}	70	110	90	320	450	320	210	30	220	0	260	250	240	320	220	210
N_{11}	100	280	120	240	60	190	120	20	20	260	0	280	320	400	500	210
N_{12}	800	30	110	180	70	230	230	40	40	250	280	0	270	260	250	240
N_{13}	200	210	10	100	280	340	320	50	20	240	320	270	0	280	290	100
N_{14}	60	70	25	290	180	400	340	90	320	320	400	260	280	0	40	30
N_{15}	140	90	60	380	320	540	80	70	450	220	500	250	290	40	0	60
N_{16}	300	200	90	200	160	600	130	10	60	210	210	240	100	30	60	—

在使用及时服务策略情况下,系统所构建的组播树如图 8 所示.

由此可见,使用及时服务策略极大地缩短了节点延时,提高了系统的反应能力.

3 结束语

通过对现行的 P2P 直播系统的研究,提出了一种单组播树、多组播树结合的组播树构建方案.该方案结合了单组播树维护方便、多组播数节点负载均衡的优点,同时对新加入节点利用及时服务策略,发挥新节点的优势,提升系统服务能力.

参考文献:

[1] 高峻. P2P 流媒体直播系统中监控技术研究[D]. 武汉:华中科技大学,2006.

[2] 邹勇. 基于 P2P 技术的流媒体服务平台的研究[D]. 上海:上海交通大学,2006.

[3] 张朕. 基于对等网络架构的流媒体在线直播系统的设计和实现[D]. 上海:华东师范大学,2006.

[4] 杨莉丽. 应用层多播技术实现与性能分析[D]. 上海:南京理工大学,2006.

[5] 李倩. 基于 P2P 的流媒体直播系统关键技术研究[D]. 湖南:中南大学,2007.

[6] 刘亚杰. P2P 流媒体内容分发关键技术研究[D]. 长沙:国防科学技术大学,2006.

[7] 康芳,杨志义,王道彬. 一种 P2P 直播系统调度管理方

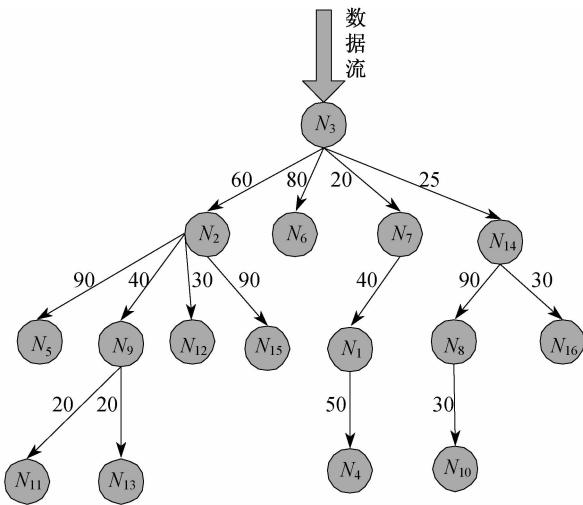


图 8 使用及时服务策略构建的组播树

Fig. 8 Constructing multicast tree with providing services immediately strategy

16 个节点的平均延时为

$$\bar{D} = \sum_{i=1}^{16} D_i / 16 = (60 \times 7 + 80 + 20 \times 3 + 25 \times 4 + 90 + 40 \times 3 + 30 + 90 + 40 \times 2 + 90 \times 2 + 30 + 20 + 20 + 50 + 30) / 16 = 88.$$

- 法的设计与实现[J]. 信息安全与通信保密, 2008(1): 46-47.
- [8] 韩俊伟, 王少锋. 基于 P2P 的流媒体直播系统研究与设计[J]. 计算机应用研究, 2006(6): 227-229.
- [9] JIN H, GUO X Z, ZHANG C, et al. A new layered P2P architecture with efficient resource location strategy [C]// Proceedings of IEEE Conference One-Commerce Technology for Dynamic E-Business (CEC-EAST' 04). 2004: 298-301.
- [10] PADMANABHAN V N, WANG H J, CHOU P A, et al. Distributing streaming media content using cooperative networking [C]// The 12th International Workshop on Network and Operating Systems Support for Digital Audio and Video, Miami, Florida, USA. New York: ACM Press, 2002: 177-186.
- [11] ORAM A. Peer-to-peer: harnessing the power of disruptive technologies [M]. Sebastopol: O' Reilly and Associates Inc, 2001: 1-452.

(编辑: 刘志强)

(上接第 90 页)

- [9] JELONEK J, KRAWIEC K, SLOWINSKI R. Rough set reduction of attributes and their domains for neural networks [J]. International Journal of Computational Intelligence, 1995, 11(2): 339-347.
- [10] 石峰, 娄臻亮, 张永清. 一种改进的粗糙集属性约简启发式算法[J]. 上海交通大学学报, 2002, 36(4): 478-481.
- [11] University of California. KDD Cup 1999 Data [CP/OL]. [2008-05-06]. <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>.
- [12] 张文修, 梁怡, 吴伟志. 信息信筒与知识发现[M]. 北京: 科学出版社, 2003.
- [13] CRISTIANINI N, TAYLOR J S. 支持向量机导论[M]. 北京: 电子工业出版社, 2004.
- [14] The University of Waikato. Weka Software [CP/OL]. [2008-05-09]. <http://www.cs.waikato.ac.nz/ml/weka/>.
- [15] BERTINO E, FERRARA E, QUICCIARINI A C. Trust2X: a peer to peer framework for trust negotiations [J]. IEEE Transactions on Knowledge and Data Engineering, 2004, 16(7): 827-842.

(编辑: 刘志强)