

基于 Chord 的流媒体直播系统设计

张国东¹, 汪国安²

- (1. 河南大学 计算机与信息工程学院, 河南 开封 475004;
2. 河南大学 网络信息中心, 河南 开封 475001)

摘要:针对当前实时流媒体系统的维护开销较大和节点缓存资源块发布不完整等问题,提出了一种基于 Chord 的结构化的流媒体直播系统 RR-Media. 该系统中的节点分为超级节点和普通节点,所有路由信息都在超级节点上完成;任何节点都可以发布资源,节点之间通过定期交互来声明自己的存在. 仿真结果表明 RR-Media 具有较低的网络延迟、非常小的维护开销和高效的节点缓存资源发布完整性.

关键词:流媒体;P2P;RR-Media;Chord

中图分类号:TP391 **文献标志码:**A

Design of live streaming media system based on Chord

ZHANG Guo-dong¹, WANG Guo-an²

- (1. College of Comp. and Infor. Eng., He'nan Univ., Kaifeng 475004, China;
2. Network Center, He'nan Univ., Kaifeng 475001, China)

Abstract: Aiming at present the real-time streaming media system's problems caused by high maintenance overhead and the cache resource block of node released of integrity, a structuring live streaming model RR-Media based on Chord was proposed. Nodes were divided into super nodes and ordinary nodes in this system, all the routing information is being done on the super node. Any node can publish resources, and declare their own existence through regular interactions. Simulations result showed that RR-Media drastically has lower network latency and little maintenance overhead and efficient integrity of node resource publishing.

Key words: streaming media; P2P; RR-Media; Chord

0 引言

时移流媒体允许用户在观看视频直播过程中可以进行暂停、快进、回播等操作,但对带宽、存储以及其他资源的要求非常高. 而传统的网络服务是基于 C/S 模式的,在该模式中,服务器存在单点故

障和不可扩展等瓶颈问题. 为了解决这些问题,基于 P2P 的时移流媒体系统应运而生. 该系统通过充分利用普通客户端的剩余资源,减轻服务器的压力,大大提高了系统的可伸缩性和实际传输质量.

在 P2Cast^[1]和 P2VoD^[2]中,观看者只能从头观看视频. oStream^[3]提供观看者在视频的任意位置进

收稿日期:2011-10-26

作者简介:张国东(1984—),男,河南省许昌市人,河南大学硕士研究生,主要研究方向为计算机网络技术.

通信作者:汪国安(1957—),男,河南省新蔡县人,河南大学教授,主要研究方向为计算机网络及应用.

入,但是节点插入系统会造成新节点与子节点的冲突. P2TSS^[4-5] 系统有效地解决了时移问题和回放问题并减轻了服务器的压力,但是在 P2TSS 中,节点仅仅缓存下一个流媒体视频块,所有参与的节点都不存储初始播放时的第 1 个视频块,也不能对首个视频块回放. DRPSS^[6] 系统有效地解决了首个视频块存储的问题并降低了回放延迟,但 DRPSS 系统必须存储所有桥节点和本层所有非桥节点的信息,从而带来了高维护的问题,而且 DRPSS 系统没有发布节点缓存的全部视频块.

本文拟设计一个新的实时流媒体系统 RR-Media,使其能够有效地完成实时流媒体回放,以期大幅度降低终端用户观看视频的延迟,改善直播系统的实时性,有效降低系统的控制开销.

1 基于 Chord 的覆盖网络构建方法

RR-Media 是基于 Chord 的一种结构化分布式查询系统,该系统使能力度高的节点贡献更多的资源. 根据能力的强弱节点被划分成超级节点 (RP) 和普通节点 (NRP),所有的 RP 形成 1 个路由环 (routing-ring),路由环用于路由查询请求,NRP 不参与路由信息. 在 RR-Media 中,所有请求信息均在路由环上路由,最后跳到目的地. 当 1 个新的节点加入系统时,首先作为 NRP,经过一段时间 (例如 0.2 h),根据节点能力度决定是否成为 RP. RR-Media 不存在集中控制,因而可以对其进行扩展.

1.1 RR-Media 的结构

RR-Media 是一种分布式 P2P 结构,采用 Chord 协议在节点之间交换信息,构建了无结构的环状覆盖网络. 为了形成覆盖网,新节点 n 在加入系统时,首先通过物理位置映射 (Hash 函数) 获取其空间 ID (space ID) 即 SI 值,然后通过请求视频块获取其时间 ID (time ID) 即 TI 值.

1) NRP. 每个节点在 RR-Chord 环上按照节点 SI 值进行排序,每个 NRP 直接与 RP 连接,NRP 的路由表包含它的路由前驱、路由后继和路由节点.

2) RP. RP 维护它在路由环上的立即前驱与立即后继. 同时,每个 RP 还维护 1 个路由表和 1 个非路由表,RP 路由表中每一项依据 PNS (proximity neighbor selection)^[7] 指向 1 个 RP,非路由表记录从本身到其路由后继的所有的 NRP.

1.2 路由表的设计

图 1 是具有 6 b ID 空间的 RR-Media 结构图,其中,N2 是 1 个路由节点,拥有 2 个表,非路由表记录了从本身到路由后继的所有的 NRP,即表 1;路由表中每一项依据 PNS 指向 1 个 RP,即表 2. 路由表中的每一项记作 (i,P) ,其中 $0 \leq i < m, 0 \leq P < k$. 1 个路由节点 n 的路由表项 (i,P) 包含 1 个 RP 的 ID 和 IP,其 ID 在区间 $[n + 2^i + P \times 2^i/k, n + 2^i + (P + 1) \times 2^i/k]$ 内. 本文设定 $k = 2$.

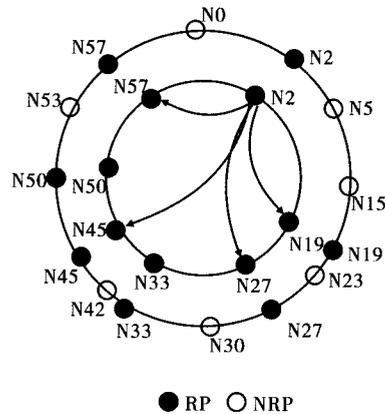


图 1 具有 6 b ID 空间的 RR-Media 系统结构图
表 1 N2 的非路由表

SI	TI
N5	3
N15	4

表 2 N2 的路由表

Entry	Interval	SI	TI
(0.0)	[3,4)	N19	3
(1.0)	[4,5)	N19	3
(1.1)	[5,6)	N19	3
(2.0)	[6,8)	N19	3
(2.1)	[8,10)	N19	3
(3.0)	[10,16)	N19	3
(3.1)	[16,18)	N19	3
(4.0)	[18,26)	N19	3
(4.1)	[26,34)	N27	4
(5.0)	[34,50)	N45	6
(5.1)	[50,66)	N57	9

2 相关技术

2.1 节点的维护

在流媒体系统中,在节点加入时需要构建一个

路由表信息,而当节点离开之后,需要通知相关节点随时更新其路由表信息.在 P2TSS 中,节点注册数目的增加导致更改关键字节点数目的增多;而在 DRPSS 系统中,桥节点保存了所有桥节点和本层节点的信息,因此在节点离开和加入时都需要通知所有桥节点,这样给系统带来了大量的维护开销.

在 RR-Media 中,节点之间通过定期交互来声明自己的存在.分别对 RP 和 NRP 讨论其维护开销:
 1) NRP 在加入和离开时只需通知 3 个节点,即节点的前驱、节点的后继以及 RP.若 NRP 非正常退出,由于节点之间定期进行交互,如果一定时间之内无交互信息,那么前驱和后继节点直接可以更新路由表信息,因此大大节约了 NRP 的维护开销;
 2) RP 在退出之后,由于 RP 并不是记录所有的 RP 信息,因此仅仅通知路由环上在它之前的 RP,并把非路由表信息发送给上一个 RP,因此,随着节点的增多,RP 的路由维护比桥节点路由维护少.

2.2 节点资源的发布

在 RR-Media 系统中,任何节点都可以发布资源.首先节点在本机上发布资源,更新本机的资源列表,同时节点向服务器发送资源 ID 以及资源的描述.服务器对得到的资源进行处理:1)如果在已有的列表中不存在,则添加至资源列表中;2)如果存在相同的资源则丢弃,服务器更新资源列表并记录更新时间,同时更新资源列表的文件.

在 P2TSS 和 DRPSS 中,每个节点在实时流缓冲时,仅仅发布本身在初始化时得到的视频块的 ID,而节点在观看实时流的同时也在缓存以后的视频块,但节点并没有把此视频块信息发布出去,这就造成了资源的浪费,甚至可能造成此视频块的丢失.

在 RR-Media 中,为了避免资源的浪费,节点在发布初始资源的同时也发布已经缓存的其他资源块.由于节点之间通过在一定时间内进行交互来保证节点的存在,因此 NRP 只需在交互的时候把信息表发送给其他节点就可以保持信息表的新鲜度.另外,为了降低资源延迟,在每次交互时同时发布节点最近视频块请求 ID.

3 仿真与分析

本文在 openchord 上建立 RR-Media 系统,节点数由 100 变化到 1 000,节点会保存在系统中直到视频播放过程结束.当节点加入系统时,将从起始播

放点复制流媒体到缓冲区,再进行播放.

3.1 查询跳数和维护开销

查询开销评估主要基于查询跳数和查询延迟指标.在 RR-Media 中,由于 NRP 不参与路由信息,只影响它的路由前驱的 NRP 表,因此仅采用 RP 来进行查询开销的评估,如图 2 和图 3 所示.

从图 2 可以看出 DRPSS 在其节点较少时具有最好的查询性能,但随着节点的增多,查询性能降低,而 RR-Media 的查询性能逐渐与 DRPSS 持平.从图 3 可以看出,随着本层节点的增多,DRPSS 中的桥节点随时要对路由信息进行更新,因此维护开销呈直线增长,当节点增加到一定程度时,将导致系统无法运转,因此 DRPSS 可扩展性较差.而 RR-Media 仅仅在路由环上进行,因此维护开销最低.

3.2 节点缓存资源命中率

图 4 为节点缓存资源命中率.当系统将节点数量增加到 1 000 时,由于 P2TSS 和 DRPSS 在节点

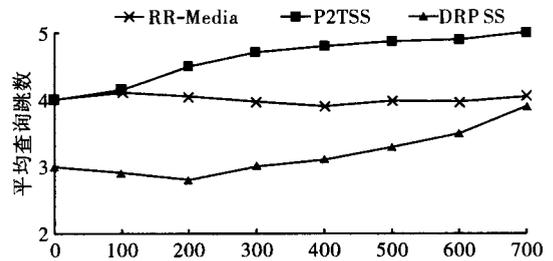


图 2 平均查询跳数

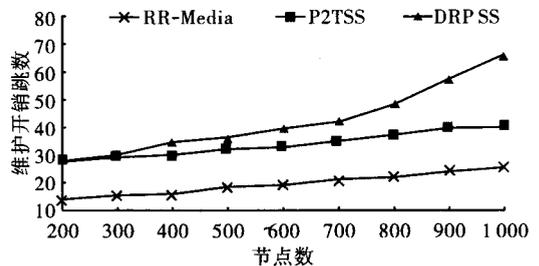


图 3 维护开销

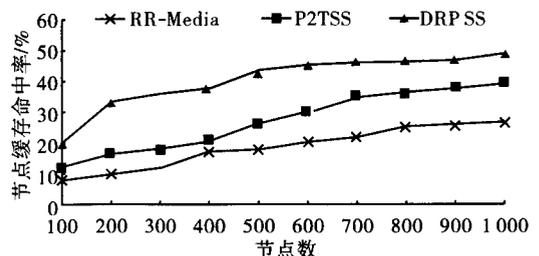


图 4 节点缓存命中率

资源发布时并没有把节点所缓存的视频块信息全部发布,因此随着时间的推移,存在节点中的缓存视频块并不能完全被其他节点找到,从而造成了资源的浪费,导致节点用户观看视频时的延迟较高.而 RR-Media 由于在节点发布资源的同时也发布初始视频块和缓存中的视频块,因此节点能更快更精准地找到所需要的视频块.

4 结论

本文设计了一个基于 Chord 的 P2P 流媒体资源定位结构的实时流媒体系统 RR-Media. 性能分析表明,RR-Media 大幅度降低了终端用户观看视频的延迟,改善了直播系统的实时性,降低了系统的控制开销,提高了节点缓存资源发布的完整性.因此 RR-Media 可大幅度提升视频的流畅性和完整性.

参考文献:

- [1] Guo Y, Suh K, Kurose J, et al. P2Cast: peer-to-peer patching scheme for VoD service[C]//Proc of the 12th World Wide Web Conf, Budapest: ACM Press, 2003:301-309.
- [2] Do T, Hua K A, Tantaoui M. P2VoD: providing fault tol-

erant video-on-demand streaming in peer-to-peer environment[C]//IEEE ICC, Paris: IEEE, 2004:20-24.

- [3] Cui Y, Li B, Nahrstedt K. oStream: synchronous streaming multicast in application-layer overlay networks[J]. IEEE J on Selected Areas in Com, 2004, 22(1):91.
- [4] Zhao B, Huang L, Stribling J, et al. Tapestry: a resilient global-scale overlay for service deployment[J]. IEEE Journal on Selected Areas in Communications, 2004, 22(1):41.
- [5] Deshpande S, Noh J. P2TSS: Time-shifted and live streaming of video in peer-to-peer systems[C]//IEEE Int Conf on Multimedia and Expo, Hanover: IEEE, 2008:301.
- [6] Deshpande S, Noh J. Pseudo-DHT: Distributed Search Algorithm for P2P Video Streaming[C]//Tenth IEEE Int Symposium on Multimedia, California: Emerald, 2008: 348-395.
- [7] Gummadi K P, Gummadi R, Gribble S, et al. The impact of DHT routing geometry on resilience and proximity[C]//Proc of the 2003 ACM Sigcomm on Comp Com. Karlsruhe: Association for Computing Machinery, 2003: 381-394.

(上接第 48 页)

- [2] 谢壮宁. 群体高层建筑风荷载干扰研究[D]. 上海: 同济大学, 2003.
- [3] 赵青春, 彭兴黔, 周显鹏, 等. 低矮双坡屋面房屋风干扰效应的数值模拟分析[J]. 福州大学学报: 自然科学版, 2008, 36(6):863.
- [4] Kumar A. Wind interference amongst low-rise buildings [R]. India: Dept of Civil Eng of Univ of Roorkee, 1994.
- [5] 张瑜明, 焦峰, 张珂. 低矮建筑群风荷载作用下干扰效应的数值模拟[J]. 河南科技大学学报: 自然科学版, 2009, 30(2):57.
- [6] 焦炳峰. 低层四坡屋面房屋风特性及相互干扰的数值模拟[D]. 杭州: 浙江大学, 2005.
- [7] 杨玉珍. 中国民居建筑的奇葩——福建土楼[J]. 百姓生活, 2010(5):62.
- [8] Mochida A, Lun I Y F. Prediction of wind environment and thermal comfort at pedestrian level in urban area

[J]. J of Wind Eng and Ind Aerodynamics, 2008, 96(10-11):1498.

- [9] 刘正先, 苗永森. 有曲率影响的湍流流场中几种 $k-\zeta$ 模型的数值[J]. 天津大学学报: 自然科学与工程技术版, 2000, 33(1):97.
- [10] Yoshihide Tominagaa, Akashi Mochidab, Ryuichiro Yoshie, et al. AIJ guidelines for practical applications of CFD to pedestrian wind environment around buildings[J]. J of Wind Eng and Ind Aerodynamics, 2008, 96(10-11):1749.
- [11] 曾锴, 汪丛军, 黄本才, 等. 计算风工程中几个关键影响因素的分析与建议[J]. 空气动力学学报, 2007, 25(4):504.
- [12] 郑朝荣, 张耀春. 不同截面尺寸双柱体静力干扰效应的数值模拟[J]. 武汉理工大学学报: 交通科学与工程版, 2007, 29(8):88.