

基于 P2P 的流媒体点播技术与展望^{*}

彭 昭¹ 吕冠中² 梁 洁² 杨宗凯³

(华中科技大学电信系 武汉 430074)¹ (中国电信股份有限公司广州研究院 广州 510630)²

(华中师范大学信息技术系 武汉 430074)³

摘 要 视频点播数据传输的离散程度,给 P2P 流媒体点播技术带来了极大的挑战。分析了 P2P 流媒体点播技术的基本问题,从媒体资源定位和分发两个方面介绍各种已提出的协议和算法,结合 P2P 视频点播的特点和网络应用的实际需求分析其优点和不足,并以此为基础提出有待解决的主要问题,为进一步的研究和应用提供有益的参考。

关键词 对等网络,视频点播,流媒体,覆盖网络

Survey on Peer-to-Peer Video-on-demand Streaming Protocols

PENG Zhao¹ LU Guan-zhong² LIANG Jie² YANG Zong-kai³

(Department of Electronic and Information, Huazhong University of Science and Technology, Wuhan 430074, China)¹

(Guangzhou Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China)²

(Department of Information Technology, Huazhong Normal University, Wuhan 430074, China)³

Abstract The diversity of video-on-demand data transmission imposes great challenges for the development of Peer-to-Peer (P2P) Video-on-demand (VoD) streaming. Presented primary problems of developing P2P VoD streaming system. Then introduced existing protocols and algorithms of P2P VoD streaming with respect to locating supplying peers and content distribution, and summarized their merits and demerits according to the requirements in practical networks. Some open issues were finally presented for further studies and deployment of P2P VoD streaming.

Keywords Peer-to-Peer network, Video-on-demand, Video streaming, Overlay network

1 引言

传统的网络视频点播(Video on Demand, VoD)系统主要采用客户机/服务器(Client/Server, C/S)模式。由服务器为每个客户端单独传输一个视频流,随着用户异步请求数目的增加,基于 C/S 模式的网络视频点播系统会面临一系列根本性问题,包括服务器负载过重,带宽利用率较低,服务质量难以保证等。针对这些问题,已提出了基于 IP 多播技术的 Patching^[1,2], Periodic Broadcasting^[3,4] 等解决方案,同时针对异步点播服务和网络可扩展性,还提出了 Batch 协议簇^[5]、Patch 协议簇^[6]以及代理缓存和内容分发网络^[27](CDN, Content Distribution Network)等技术方案。但是,这些解决方案仍然存在维护复杂,服务质量不能保证,网络扩展性不足等方面问题。

将对等(Peer-to-Peer, P2P)技术引入网络视频点播系统,以有效解决 C/S 模式的上述问题,近年来获得了广泛关注。通过共享网络中节点的计算、带宽和存储等资源,P2P 模式彻底改变了视频点播资源分发和用户交互的方式,显著地减少服务器负载,提高网络资源利用率。已有多个研究机构对此展开研究,提出多种 P2P 视频点播协议和算法,如 P2Cast^[7], DirectStream^[8], P2VoD^[9], GnuStream^[10,11], Hybrid Overlay Network^[12]以及基于 DHT 流媒体系统^[13]等。这些协议和算法主要从媒体资源定位和分发两个方面,分别解决 P2P 点播

系统面临的节点动态性和带宽限制性问题,减少节点动态行为对系统的影响,保证视频播放的可用性和连续性。

本文从媒体资源定位和分发两个方面全面介绍 P2P 流媒体点播系统的基本概念、已有技术方案及原理,以此为基础总结和提出有待解决的主要问题,为进一步的研究和应用提供有益参考。

本文其余部分组织如下:第 2 节介绍 P2P 流媒体点播技术的基本概念和分类,提出 P2P 流媒体点播的基础问题;第 3 节和第 4 节分别介绍媒体资源定位和分发的主要技术方案以及相关 P2P 流媒体点播系统,并对相关的算法和协议进行比较,概括其适用场景和主要优缺点;第 5 节总结和提出 P2P 流媒体点播的研究问题;最后总结全文。

2 P2P 流媒体点播

2.1 P2P 流媒体点播技术特点

P2P 是一种节点对等的网络服务模型,每个节点(称为 Peer)具备相同的地位,既是服务的提供者又是服务的获取者。Peer 之间直接互相通信,共享和处理计算和信息资源,协同完成任务,无需依赖集中式的服务器。将 P2P 模式引入流媒体服务中,可以充分发挥每个 Peer 的作用,通过缓存部分视频信息为其它 Peer 提供服务,从而有效减少服务器负载和网络带宽的高消耗。

与传统的 C/S 模式相比,P2P 模式具有以下优点。

^{*}基金项目:国家自然科学基金(60602029),湖北省智能互联网技术重点实验室开放基金(HSIT-200603),国家发改委 CNGI 示范工程资助项目(下一代互联网示范工程 2006 年产业化及应用实验)资助。彭 昭 博士研究生,研究方向为 P2P 流媒体技术;吕冠中 硕士,工程师,研究方向为数据通信与应用;梁 洁 硕士,工程师,研究方向为数据通信与应用;杨宗凯 博士生导师,研究方向为现代信息网络技术。

1) 资源利用率高: 闲散计算机资源得到充分利用, 减少服务器容量和带宽的瓶颈。

2) 可扩展性高: 每个 Peer 都参与贡献资源, 节点数越多, 网络性能越好。

3) 稳定性高: 不依赖中心服务器, 只要有一定数量的 Peer, 整个网络就可以正常运转。

纯粹意义上的 P2P 网络中, 控制信息交互和数据传递均为分布式。而流媒体点播系统中的异步请求和频繁的用户交互行为, 使得纯 P2P 模式并不适合, 因此常见的 P2P 流媒体点播网络主要采用集中架构或混合架构^[14]。

1) 集中架构: 采用 C/S 模式管理协调控制信息, 即由集中的服务器管控所有节点信息。

2) 混合架构: 采用集中与分布相结合的方式^[15], 由超级节点集中地提供控制信息和数据传递服务, 其中超级节点与普通节点之间多采用 C/S 模式, 超级节点之间采用 P2P 模式。

2.2 P2P 流媒体点播基本问题

在因特网中采用 P2P 模式能够支持大规模的流媒体服务^[16]。但对于流媒体点播而言, 由于其固有的一些特点, 即节点服务能力有限且存在差异, 节点可以随时加入退出点播会话, 播放时间长, 需要较高带宽, 时限要求严格以及用户 VCR 行为等, 使得 P2P 流媒体点播系统的设计需要面临如下挑战:

1) 节点的动态性

Peer 可能因突发阻塞或离开失效^[17,24], 点播特有的 VCR 操作也导致 Peer 的不稳定性, 向其请求内容的 Peer (称为后继 Peer) 需要寻找新 Peer 取代。因此, P2P 点播系统应该具备足够的弹性, 以解决 Peer 失效问题。

2) 带宽的限制性和动态性

Peer 带宽容量有限, 通常只能支持一部分 Peer 的内容请求。在内容分发过程中, Peer 的可用带宽可能上下波动, 因此 P2P 流媒体点播系统应该能够自适应调整发送速率, 以保证各个 Peer 接收所点播的流媒体质量不受影响^[17]。

针对上述基本问题, P2P 流媒体点播系统的设计主要从媒体资源定位方案和媒体资源分发策略两个方面入手, 提高系统弹性和播放连续性。媒体资源定位是指 Peer 如何尽快找到提供媒体内容的 Peer, 而媒体资源分发主要涉及 Peer 的缓存策略和资源内容的分发调度。下面将从媒体资源定位方案和媒体资源分发策略两个方面(分类如图 1 所示)全面介绍已有的 P2P 流媒体点播。

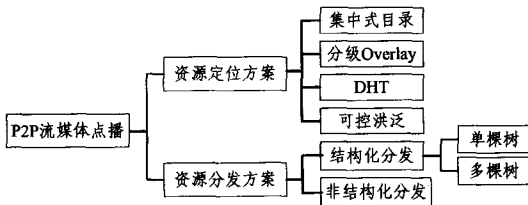


图 1 P2P 媒体点播定位和分发解决方案

3 媒体资源定位

P2P 流媒体系统中, 媒体资源定位是 Peer 加入传输会话和接收视频流的前提。而点播的异步请求和 VCR 操作给媒体资源定位提出了更高的要求和挑战, 不仅要求准确定位媒体资源的内容或片段, 同时要求快速的响应和较低的系统资

源消耗。目前, 常见的点播系统如 P2VoD^[9], 仅实现了异步请求和中断恢复, 属于准点播系统。一个真正 P2P 点播系统的媒体资源定位涉及 3 种应用场景:

1) Peer 加入。快速定位资源片段并在时限内播放视频。

2) Peer 中断。资源节点发生阻塞时, 快速有效地中断恢复, 包括资源节点的重定位。

3) VCR 操作。及时准确定位任意视频片段, 实现视频选择性播放。

这些应用场景涉及的目标参数包括: Peer 间的距离(跳数)、Peer 的可用带宽和资源、视频片段的信息、Peer 所在区域、Peer 优先级等。根据所选择的目标参数的不同可将算法分为如下几类:

1) 最小延迟算法。选择可用带宽大且跳数少的 Peer, 减少骨干网络带宽消耗, 将 P2P 媒体分发限制在一定的区域^[9]。

2) 最快匹配算法。选择可用带宽大的 Peer, 减少 Peer 的加入等待时间。

3) 最长空闲算法。选择服务时间短的 Peer, 保证网络服务的公平性。

4) 最佳优先算法。综合考虑 Peer 的地域、距离、类型和可共享资源, 选择优先级较高的 Peer, 构造稳定的分发网络。

这些算法与查找策略及网络结构构成不同的媒体资源定位方案: 集中式目录、分级覆盖 (Overlay) 网络、DHT 以及可控洪泛。

3.1 集中式目录

集中式目录方案是在集中式目录服务器上维护所有 Peer 的全局信息, 包括网络地址、可用带宽、视频段标识等。目录服务器运行相关查找算法, 向 Peer 提供节点路由和定位候选列表, Peer 根据某种算法选用其中一部分。集中式目录实现和部署较为简单, 相应的 P2P 点播系统包括 P2Cast 和 DirectStream^[7,8]。

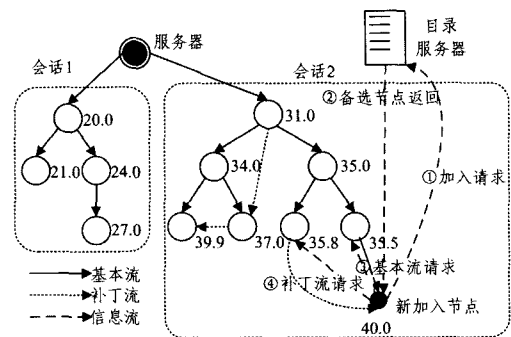


图 2 P2Cast 系统节点框架

P2Cast^[7] 是一个基于应用层多播树的 P2P 准点播系统 (不提供 VCR 操作)。每个新加入 Peer 需缓存时间长度为 T 的视频数据, 并与后续 T 时间内到达的 Peer 构成一棵应用层多播树, 每棵树占用服务器的一个数据频道并形成会话。为了保证不同时刻加入的 Peer 获取完整的视频流, P2Cast 采用补丁 (Patching) 的方式来取得加入时刻之前的媒体数据, 因此大部分 Peer 除了获取会话中传输的基本流 (Base Stream), 还需获取补丁流 (Patch Stream), 如图 2 所示。为了构造良好的基础树形结构, P2Cast 采用带宽优先原则和局部信息原则, 根据可用带宽和节点基本信息选择父节点, 同时采用 BF (Best Fit) 算法选择最佳带宽的补丁服务器。但该系统

中每个 Peer 需要较多的缓存空间缓存基本流,因此对节点带宽和存储资源要求较高。

DirectStream^[8]也是一个基于应用层多播树的 P2P 点播系统。系统中目录服务器协调管理全局网络,记录所有 Peer 的视频信息,支持任意视频位置的请求。每个 Peer 通过 QoS Parent 选择算法选择合适父节点并建立传输连接,同时缓存最近接收的媒体内容。QoS 算法使用距离带宽比(定义为 n_i/x_i ,其中 $0 \leq r < \infty$)作为选择依据,其中 n_i 为请求节点到备选父节点 c_i 的跳数, x_i 为其标准可用带宽。较高的带宽有助于负载均衡,较小的距离可以减少整个网络通信量。但本系统采用集中式目录方案依然存在系统扩展瓶颈。

集中式目录方案虽然实现简单,但存在如下问题:

- 1) 在节点数为 N 的系统中,目录服务器必须维护 $O(N)$ 条状态信息,所需的周期性更新^[17]使得 N 较大时服务器负载过大。
- 2) 集中的目录服务器容易导致单点失效,一旦服务器失效,整个系统将陷于瘫痪。

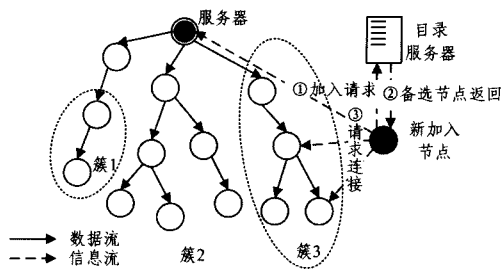


图3 DirectStream 系统框架

集中式目录方案主要适用于规模较小的网络,如公司或校园网,服务器能够很好满足节点的更新及交互请求。

3.2 分级 Overlay

分级 Overlay 方案是通过 Overlay 网络中各层节点共同维护全网拓扑和路由。节点请求资源时,联系 Overlay 网络中的根节点或数据源节点获取下一级节点列表 $\{P_1, P_2, \dots, P_k\}$,探测各节点并根据算法或规则选择节点 P_x ,联系 P_x 获取下一级节点列表并选择。以上过程反复迭代,直至找到目标节点^[17]。该方案将负载有效分担至全网,避免了集中式目录方案中单点失效和可扩展性问题。

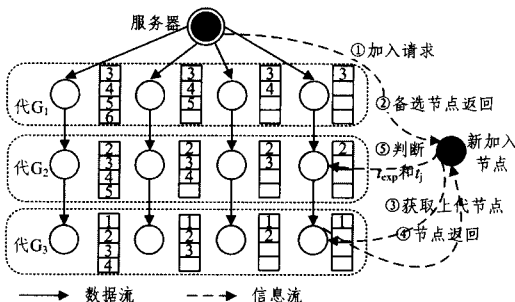


图4 P2VoD 系统节点框架

P2VoD^[9]是基于分层结构的 P2P 准点播系统。系统引入“代”(Generation)和不等长缓存的概念解决系统抖动问题,定义缓存数据块最小序号相同的 Peer 为同一“代”,且根据 FIFO 队列和加入时间确定缓存数据块的大小和数目。系统中 Peer 之间周期性交换控制信息,只在加入和离开时才与目录服务器交互,有效减少服务器负载。如图 4 所示,新加入

Peer 依据 t_{exp} (定义为节点初始视频段的生存时间)和 t_j (定义为节点的加入时间)判断加入最新一代还是重新创建一代。在父节点选择时采用三种算法:最长空闲选择算法、最快匹配选择算法、最小延迟选择算法。但 P2VoD 没有考虑 VCR 操作,每个会话结束时低代节点离开导致系统极不稳定,缓存利用率也比较低。

分级 Overlay 方案依赖于 Overlay 网络架构,存在如下缺陷:

- 1) 对于频繁的节点动态行为,网络架构维护复杂。
- 2) 定位与网络拓扑密切相关,查找性能需依赖完善的拓扑维护和控制协议。

因此分级 Overlay 方案适用于系统节点数量较少、对实时 VCR 操作要求不是太高的应用场景。

3.3 DHT

DHT 方案是通过分布式哈希表定位资源节点。该方案中,每个视频文件索引被表示成一个 (K, V) 对, K 是媒体资源信息, V 是存储媒体资源的 Peer 地址信息。所有 (K, V) 组成文件索引哈希表,并分布在所有参与节点上,每个节点维护部分哈希表信息。节点请求时只需输入 K 值,就能快速定位所有存储该视频的节点地址。一个大小为 N 的 DHT 网络中,平均的查询跳数是 $O(\log N)$ ^[18,19,21]。DHT 方案具有查询长度短、查询结果准确性高、消息通信量少等特点。

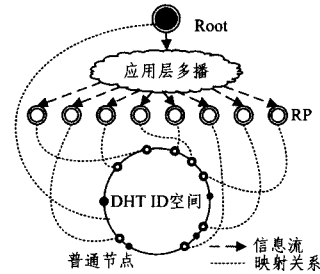


图5 基于 DHT 系统节点框架

表1 Location Table 实例

	S_1	S_2	S_3	S_4	S_5
Location 1	Peer013	Peer007	Peer007	Peer007	Peer007
Location 2	Peer10D	Peer023	Peer213	Peer09E	Peer068
Location 3	Peer023	Peer213	Peer09E	Peer342	
Location 4	Peer006			Peer02C	
Location 5	Peer057			Peer0A4	
Location 6				Peer213	
GRN	5	3	3	6	2
GRN^0	6	NextRP	Peer055	A_n	204

文献[13]提出一种基于 DHT 方案的视频点播系统。该系统采用类 Pastry 的 DHT 网络缓存媒体数据块的地理信息。网络由 3 个部分组成:RP(Rendezvous Peer)、根服务器(Root)和普通节点(Ordinary Peers),如图 5 所示。系统中 Peer 节点注册和查找基于 Location Table(分发给 RP 上),该表包含两个部分:一部分由 m 列 n 行组成,第 i 列 j 行的条目表示对应视频单元第 i 段第 j 个地址;另一部分对应视频分段 S_i ,其中 GRN_i 表示全局拷贝数目, GRN^0 表示视频单元内预期所有分段的拷贝数目, NEXT_RP 表示后继单元指针, A_n 表示总共访问相关单元的数目。系统节点的选择基于 ITLD(IP to Landmark Database),包括 inter-country 层、inter-ISP 层、WAN(Wide-Area Network)层、MAN(Metropolitan-Area Network)层和 LAN(Local-Area Network)层的网

络坐标信息,选择差异最小的节点作为父节点。该系统具备较好的扩展性和较少的延迟抖动,但局限在同种类型局域网环境中,且节点 VCR 操作和抖动增大分布式哈希表维护开销。

DHT 方案将路由请求均匀分布在网络中所有节点,通过 DHT 协议简化系统设计与实现,使得系统无需处理复杂节点管理就能实现媒体流有效定位。现有 DHT 结构有 chord^[19], CAN^[25], Pastry^[18] 等,但仍存在如下缺陷:

- 1) 哈希表维护机制复杂。
- 2) 节点频繁加入和退出所造成的抖动会增大系统维护开销。
- 3) 支持精确关键词匹配查询,无法支持内容/语义等复杂查询。

因此 DHT 方案适用于节点较为稳定的应用场景。

3.4 可控洪泛

可控洪泛是基于节点 Mesh 网的资源定位方案。该方案 Peer 通过向邻居节点(不包括请求报文的发送节点)广播查询请求来实现资源定位。采用 TTL (Time-to-Live) 限制网络中大量无用数据,每经过一次节点广播, TTL 减 1,直到为 0。在广播过程中,拥有媒体资源的节点回复请求节点,如在规定时间内没有节点回复则认为资源不存在。可控洪泛方案构建的网络模式消除了对拓扑结构的依赖,减少网络的维护开销,对热点资源有很好的搜索效率和质量。

GnuStream^[10,11] 是基于网状多播策略的 P2P 准点播系统。系统中每个 Peer 可以从多个 Peer 获取媒体资源。GnuStream 构建在 Gnutella^[20] 网络上,使用洪泛定位资源,并从返回列表中选择足以支持播放的部分节点建立连接,同时将其余节点作为中断或阻塞的备用节点。GnuStream 采用两种分配策略分发媒体流:

- 1) 均匀分配。将流负载均匀分配到每个发送节点,这种策略适用于节点相似的环境中(如内网)。
- 2) 比例分配。根据发送节点的能力分配流负载,这种策略有更大灵活性,适用于动态的异构环境中(如广域网)。

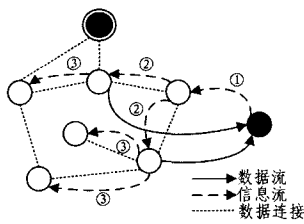


图6 GnuStream 系统节点框架

可控洪泛方案具有较高的自治性、负载均衡性和健壮性,但也存在如下缺陷:

- 1) 随着时间的增长,查询和同步更新消息呈指数级增长,增大网络负载,在低带宽节点处易形成瓶颈。
- 2) TTL 机制可能导致在有限节点跳数内无法找到媒体资源,特别是非热点资源。

因此可控洪泛方案适用于节点数量相对较少,热点资源较多的分布式查询场景。

3.5 比较

上述几种 P2P 点播系统资源定位方案,在适用场景、可扩展性、复杂度等方面均有所不同,相应的比较如表 2 所示,其中部分内容引自文献[17]。

表2 资源定位方案比较

特性	资源定位方案			
	集中式	分级覆盖	DHT	可控泛洪
网络拓扑	结构化或非结构化	结构化	结构化	非结构化
单点失效	存在	不存在	不存在	不存在
可扩展性	差	好	好	中
VCR 支持	好	差	中	中
系统弹性	中	差	好	好
服务器查询复杂度	O(N)	O(1)	O(1)	O(1)
节点查询复杂度	O(1)	O(logN)	O(logN)	O(1)
系统实现	简单	复杂	中等	中等
应用场景	节点少	VCR 少	精度高、稳定	节点少、纯分布

4 媒体资源分发

P2P 流媒体系统中,媒体资源被分解成多个数据块传输,并在接收端顺序重组播放。每个数据块都有对应的截止时间,只有在截止时间内到达才能保证媒体正常播放。如何提高分发速率,保证播放连续性,是媒体资源分发面临的主要问题,在异步请求和 VCR 操作的点播系统中尤为突出。根据分发网络拓扑的不同,现有媒体资源分发策略可分为两类:结构化媒体资源分发和混合式媒体资源分发。

4.1 结构化媒体资源分发

结构化媒体资源分发可采用单棵树和多棵树两种结构,其区别在于后者每个节点存在于多棵树中,如图 7 所示。树中根节点是数据源,提供媒体资源;中间节点和叶子节点都是 Peer,中间节点接收并转发媒体,而叶子节点只接收媒体。

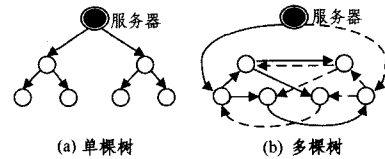


图7 结构化媒体资源分发

P2Cast, DirectStream 以及 P2VoD 均属于单棵树分发。根据维护方式的不同, P2Cast 和 DirectStream 属于集中式,而 P2VoD 属于分布式。基于单棵树分发维护简单,但也存在着一些缺陷:

- 1) 公平性。叶子节点不需要转发数据,只有中间节点承担网络负载,导致负载不均衡。
- 2) 稳定性。每个 Peer 都只有一个源节点,源节点的动态行为可能导致该 Peer 以及后继子节点的数据短缺。
- 3) 限制性。中间节点带宽限制会导致子节点数据阻塞,随着树的深度增加,这种限制性体现得更加明显。

单棵树分发策略的上述缺陷使得其只适合于小规模应用。多棵树分发策略^[22]能够在某种程度上克服上述缺陷。根据多棵树分发策略,在源服务器端对视频内容进行 MDC (Multiple Description Coding) 编码,产生多个可以单独解码的子流,并沿着不同的分发树传输。由于节点在不同树中角色不同,从而可以减少单棵树被破坏所带来的影响,如图 7 (b) 所示。该策略能够均衡网络负载,有效利用了所有节点的链路带宽,提高了系统分发的速率和吞吐量。但是,基于多棵树的分发策略增加了系统的维护负载,而且多棵树结构难以优化,通常只适用于 VCR 操作较少的异步点播应用。

4.2 混合式媒体资源分发

结构化媒体分发保证了数据传输的实时性,但节点带宽波动和失效都有可能导致缓存下溢^[23]。与结构化相对应的

是非结构化,即不依赖某种特定的结构化拓扑,Peer 之间构成随机拓扑结构,可以同时从多个源获取数据,降低了节点离开或失效带来的影响,如点播系统 GnuStream。但对于纯粹非结构化 P2P 点播而言,媒体资源离散性和用户异步请求会极大地增加交互信息量,造成大量数据冗余以及视频传输的非实时性。因此 P2P 点播系统中通常并不采用纯粹的非结构化分发策略,而是采用结构化和非结构化相结合的混合式分发方式,同时具有结构化网络视频传输延迟低和非结构化网络弹性好的优点。

文献[12,23]提出一种混合点播系统,即 HON(Hybrid Overlay Network),采用 Tree-Assisted Gossip 协议,构建和维护两个 Overlay 网络,分别为树型和随机型拓扑。HON 系统假设所有视频点播都是从头开始播放,以此为基础设计 Overlay 网络构造和数据传输算法,无缝集成两种不同分发策略。一般情况下,Peer 数据传输使用随机型网络,只有在获取数据超过时限的情况下,才基于树型网络从父节点获取。Peer 获取到新的媒体片段时,采用“推”模式将信息以 Gossip^[26](即随机选择邻居节点)方式推送给其他邻居节点;如果需要某一媒体片段数据,可直接向邻居 Peer 发送请求,采用“拉”模式获取,避免数据冗余。

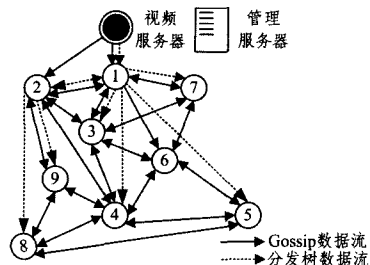


图 8 HON 模型

混合式媒体资源分发策略中树型和随机型拓扑的维护增加部分系统开销,但该策略在某种程度上保证低传输时延和较好的网络弹性,因此可适用于一定条件下的大规模点播应用。

4.3 比较

各种资源分发策略中,基于单棵树的分发能够优化控制延迟,但稳定性较差,而基于多棵树的混合式资源分发能够达到较高的传输速率,且不会受到节点带宽的限制^[17],但需要较大的系统开销维护多个 Overlay 网络。表 3 详细说明和比较了两种分发策略的特点。

表 3 资源分发策略比较

特点	结构化分发		混合式分发
	单棵树	多棵树	
可扩展性	差	中	好
网络稳定性	差	好	好
视频编码	可用	需要	可用
负载均衡	差	好	好
协议开销	小	中	大
多源下载	单源	多源	多源
传输速率	中	高	高
视频延迟	好	中	中
实现	简单	复杂	简单
应用场景	小规模	VCR 操作较少的异步点播	大规模

5 P2P 媒体点播研究问题

5.1 现有系统比较分析

对上面介绍的 P2P 流媒体点播系统,表 4 从资源定位查

找和资源分发两个方面进行了比较和小结。

由表 4 可知,已有 P2P 流媒体点播系统在资源定位查找方面主要考虑带宽、节点距离、视频内容、服务时间、缓存等因素。就资源分发方案而言,主要采用结构化拓扑。此外,现有系统大都还仅仅为准点播系统。

5.2 研究问题总结

将 P2P 技术引入流媒体点播系统,能够有效地减轻视频点播服务器的负载和带宽消耗,提高视频点播服务的可用性。如前所述,近年来关于 P2P 流媒体点播技术已有大量的研究成果,甚至原型系统。但是,对于 P2P 流媒体点播技术的实际应用,还存在如下问题有待于进一步研究。

1) 基于物理网络的 Overlay 网络拓扑优化

Peer 构成的 Overlay 网络属于应用层的逻辑网络。关于流媒体点播 Overlay 网络的构建,主要的研究目前侧重在 Overlay 网络层面考虑采用合适的 Peer 连接策略(如结构化或非结构化),以迅速和可靠地定位和获取所需媒体数据,并没有充分考虑所基于的实际物理网络拓扑和网络负载状况,甚至可能严重地背离。如何在已有媒体分发策略的基础上,根据物理网络拓扑和网络当前负载状态,生成更加合理的 Overlay 网络拓扑,还有待于进一步探索。

2) P2P 与 CDN 技术的结合

流媒体点播系统中的异步请求、用户 VCR 操作和数目庞大的媒体源,使得需要实时传输的媒体数据离散程度远基于文件下载和流媒体直播,采用 P2P 模式只有在较大用户数量的基础上才能获得性能的显著提升。CDN^[27]的效率虽然也取决于用户请求数据的集中度,但是其缓存服务器不同于普通的 Peer 用户,有着较高的链路带宽和较强的计算及存储能力。在流媒体点播系统中如何有效地结合 P2P 和 CDN 两种技术,提高点播服务的可用性,还需要在多个方面深入研究。

3) 系统适应性

P2P 流媒体点播网络中,Peer 的随机动态行为(如加入/离开、VCR 操作)以及意外的失效,会导致媒体数据传输阻塞,进而使得许多下游的 Peer 播放质量下降甚至中断。通常 Peer 缓存备用节点实现中断快速恢复,但也存在资源分配和全局优化的问题。P2P 流媒体点播系统的设计,应从资源定位和资源调度两个方面充分考虑上述问题,使得下游 Peer 所受的影响降到最低,以合理的开销提高整个系统的适应性。

4) 媒体资源定位

如前所述,媒体资源定位主要涉及 Peer 加入、Peer 中断恢复和 VCR 操作 3 种应用场景,可采用集中式、分布式、DHT、洪泛等方式。但是这些方式有着不同的限制:集中式定位快速,但服务器负载较大,容易单点失效;分布式负载分布均衡,但响应时延较长;DHT 查询准确快速,但不支持语义查询;洪泛系统适应性好,但冗余数据过多。因此,结合 P2P 流媒体点播的实际需求,设计性能全面的媒体资源定位算法,仍然是一个挑战性的问题。

5) 媒体资源调度

点播应用本身的数据请求分布离散程度,以及流媒体播放所需的实时性,使得 P2P 流媒体点播系统中的 Peer 资源分发和缓存调度策略对于系统性能有着决定性的影响。资源分发主要有结构化和混合式两种方式,缓存调度包括节点缓存分配和缓存数据块调度两个方面。设计合理的资源分发和缓存调度算法组合、源分发和缓存调度算法组合,并能适应于不

同的用户规模和媒体数据请求离散度,对于 P2P 流媒体点播而言仍然是一个有待解决的问题。

表 4 视频点播系统比较

系统	资源定位方案			资源分发方案		
	定位	查找算法	参数	任意位 置视频 请求	分发	数据源
P2Cast	集中式目录	Best Fit	带宽	不支持	结构化	单源
Direct-Stream	集中式目录	QoS Parent	节点距离,带宽	支持	结构化	单源
P2VoD	分级 Overlay	Round Robin, Smallest Delay, Smallest Distance	服务时间,节点距离,带宽	不支持	结构化	单源
DHT-Based	DHT	DHT ITLD	网络坐标	支持	结构化	单源
Gnutstream	可控洪泛	Flooding	带宽	不支持	非结构化	多源
HON	集中式目录	Condition-Based	服务时间,带宽,缓存大小	不支持	混合式	多源

6) 激励机制

与文件下载等应用不同,流媒体播放本身只需要媒体数据到达速率不低于编码速率即可,更高的到达速率意义并不明显。对于 P2P 流媒体点播系统中激励机制的效用和设计,目前还缺乏有效的研究成果。

7) 服务质量

对于 P2P 流媒体点播服务的实际部署而言,服务质量保证是一个重要的技术问题。简洁有效的服务性能理论分析模型、基于物理和 Overlay 两个网络层面考虑传输质量的 Peer 查找定位及调度缓存协议算法、分层媒体编码在复杂异构网络条件下的应用,在已有研究成果中还比较匮乏。

上述问题的研究成果,将有助于进一步认清 P2P 模式用于流媒体点播系统的优势和局限,为在实际网络中大规模部署 P2P 流媒体点播服务提供重要的技术依据和行之有效的实施策略。

结束语 流媒体点播应用中的异步请求、用户 VCR 操作和数目庞大的媒体源,使得需要实时传输的媒体数据非常离散,给 P2P 流媒体点播系统设计带来的困难远甚于其它的 P2P 应用(如 P2P 文件下载和流媒体直播)。近年来关于 P2P 流媒体点播技术已有大量的研究结果,包括较为完整的系统技术方案。本文分析了 P2P 流媒体点播技术的基本问题和挑战,从媒体资源定位和分发两个方面介绍各种协议算法的主要原理,结合 P2P 视频点播的特点和网络应用的实际需求分析其优点和不足,并以此为基础提出有待解决的主要问题,为进一步的研究和应用提供有益的参考。

参 考 文 献

[1] Hua K, Cai Y, Sheu S. Patching: A multicast technique for true video-on-demand service [A]//Proceedings of the ACM Multimedia'98. September 1998:35-43

[2] Gao L, Towsley D, Kurose J. Efficient schemes for broadcasting popular videos [A]// Proc. Inter. Workshop on Network and Operating System Support for Digital Audio and Video, July 1998

[3] Hu A. Video-on-demand broadcasting protocols: A comprehen-

sive study [A]//Proc. IEEE INFORCOM. April 2001

[4] Guo Y, Gao L, Towsley D, et al. Seamless workload adaptive broadcast [A]// Proc. of International Packet Video Workshop, April 2002

[5] Dan A, Sitaram D, Shahabuddin P. Dynamic Batching Policies for an On-demand Video Server [A]. ACM Multimedia System, 1996, 4(3):112-121

[6] Cai Y, Hua K A. Sharing multicast videos using patching streams [J]. Multimedia Tools and Application Journal, 2003, 21(2):125-146

[7] Guo Y, Suh K, Kurose J, et al. P2Cast: peer-to-peer patching scheme for VoD service [A]//Proc. WWW' 03. Budapest, Hungary, May 2003

[8] Guo Y, Suh K, Kurose J, et al. A Peer-to-Peer On-Demand Streaming Service and Its Performance Evaluation [A]// Proceedings of IEEE International Conference on Multimedia & Expo (ICME). Baltimore, MD, July 2003

[9] Do T T, Hua K A, Tantaoui M. P2VoD: Providing Fault Tolerant Video-on-Demand Streaming in Peer-to-Peer Environment [A]// Proceedings of IEEE International Conference on Communications (ICC). Paris, France, June 2004

[10] Jiang Xu-Xian, Dong Yu. Gnutstream: A mpeg video streaming system over peer-to-peer network [R]. Technical report. Project report for cs590n p2p networks and services. Fall, 2002

[11] Dong Y, Xu D, Jiang D, et al. Gnutstream: a p2p media streaming system prototype [A]// Proceedings of the International Conference on Multimedia and Expo (ICME). 2003, 2:325-328

[12] Zhou M, Liu J. A Hybrid Overlay Network for Video-on-Demand [A]//Proc. IEEE Int'l Conf. Comm. (ICC), IEEE Press, 2005:1309-1311

[13] Yin Zuoning, Jin Hai. DHT Based Collaborative Multimedia Streaming and Caching Service [A]//TENCON 2005 IEEE Region 10. Nov. 2005:1-6

[14] Backx P, Wauters T, Dhoedt B, et al. A comparison of peer-to-peer architectures [EB/OL]. Germany, 2002. <http://allserv.rug.ac.be/~pbackx/A%20comparison%20of%20peer-to-peer%20architectures.pdf>

[15] Yang B, Garcia-Moline H. Designing a Super-Peer Network [R]. Technical Report. Stanford University, February 2002

[16] Sripanidkulchai K, et al. The Feasibility of Supporting Large-scale Live Streaming Applications with Dynamic Application End-Points [A]//Proc. ACM SIGCOMM. ACM Press, 2004: 107-120

[17] Yiu W-Pk, Xing Jin, Chan S-H G. Challenges and Approaches in Large-scale P2P Media Streaming [J]. IEEE Multimedia, 2007, 14(2):50-59

[18] Rowstron A, Druschel P. Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems [A]// Proceedings of IFTP/ACM International Conference on Distributed Systems Platforms (Middleware 2001). Nov. 2001:329-350

[19] Stoica I, Morris R, Karger D, et al. Chord: A scalable peer-to-peer lookup service for internet applications [A]//Proceeding of ACM SIGCOMM, Aug. 2001:149-160

[20] Ripeanu M. Peer-to-peer architecture case study: Gnutella network [EB/OL]. Chicago: University of Chicago, 2001. <http://www.cs.uchicago.edu/matei/PAPERS/gnutella-rc.pdf>

(下转第 27 页)

Company(C_ID, C_Name, Location, Phone, Coal_ID)
 Coalition(Coal_ID, Coal_Name, Coal_Desc)
 Object(O_ID, O_Name, O_Desc, Con_ID)
 ObjectConstraints (O_ID, Con_ID, Desc)
 Constraints(Con_ID, R_ID, T_ID, C_ID, C_Type, Coal_ID, Permission, Start, End)
 UserRole(U_ID, R_ID, Desc, Status)
 UserTask(U_ID, T_ID, Start, End)

对于给定的 RFID 数据和关联信息,我们可以根据信息的性质、所属的企业、进行访问的用户角色、用于完成的任务和为完成该任务所动态组成的任务联盟等定义访问控制信息。企业在发布信息时,将相关的访问控制信息保存在供应链系统的关系数据库中,同时在系统中提供一个访问 RFID 及相关信息的 Web 服务,在用户请求访问该 RFID 数据及关联信息时,判定用户的访问请求是否得到系统的许可。从第 2 节的讨论可以看出,用户对信息的访问许可判断是通过一个由多个步骤组成的流程来完成的,可以在访问请求许可判断的 Web 服务实现中,引入工作流引擎,进行实现。

基于 Microsoft 公司的 SQL Server 关系型数据库管理系统、Visual Studio .NET 软件开发环境,和提供的工作流开发工具 Windows Workflow Foundations,我们实现了 Internet 环境下数字供应链中 RFID 数据及关联信息的复合访问控制的 Web 服务,用户在对 RFID 数据及相关信息进行访问时,首先要调用该 Web 服务,根据复合访问控制模型进行权限判断,以决定用户是否有访问该信息的权限。

目前,本文提出的 RFID 信息的复合访问控制模型已经在我们开发的 RFID 应用系统中中间件中实现和应用。在 RFID 应用系统中中间件中,有机集成了 RFID 标签信息和相关的企业产品多种实体信息,并基于本文提出的访问控制模型,可以将这些信息安全地授权给多个组织中的用户进行访问和使用。

结束语 供应链系统中多个企业之间关系的动态性、交叉性和多变性导致了供应链中 RFID 数据及关联的产品信息访问的安全控制机制的复杂性。本文在分析供应链中实体和实体之间、用户和实体之间的不同关系的基础上,给出了一种 RFID 数据及关联产品信息在供应链中的安全访问复合控制模型,综合使用基于角色的访问控制机制、基于任务的访问控制机制、基于联盟的访问控制机制和关系驱动的访问控制机制等,建立了一种 RFID 及关联信息在供应链中的安全访问控制模型,给出了一种对发布信息指定访问权限的规范化语言,给出了用户和信息访问权限之间的映射关系,讨论了访问许可的判断流程,以实现供应链中 RFID 及关联信息的安全

访问。

同时,我们提出了一种该模型的实现机制。并基于微软公司开发的软件开发工具、关系型数据库管理系统和工作流引擎,将该复合访问控制模型通过一个 Web 服务进行实现,可以通过 Internet 进行调用。供应链中的用户在访问 RFID 及相关信息时,必须首先调用该服务以判断用户是否有对该信息的访问许可。这种供应链中 RFID 信息的安全访问控制复合模型已经用于我们开发的 RFID 应用系统中中间件系统中。

参 考 文 献

- [1] Finkenzeller K. RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification[M]. 2nd Edition, John Wiley & Sons, 2003
- [2] Niederman F, Mathieu R G, Morley R, et al. Examining RFID Applications in Supply Chain Management[J]. Communications of ACM, 2007, 50(7): 92-101
- [3] 陈安, 刘鲁. 供应链管理问题的研究现状及挑战[J]. 系统工程学报, 2002, 15(2): 179-186
- [4] Ferralolo D F, Sandhu R, Gavrila, et al. Proposed NIST standard for role-based access control[J]. ACM Transactions on Information and System Security, 2001, 14(3): 224-274
- [5] 许峰, 赖海光, 黄皓, 等. 面向服务的角色访问控制技术[J]. 计算机学报, 2005, 28(4): 686-693
- [6] Thomas R K, Sandhu R S. Task-based Authorization Controls (TBAC): A Family of Models for Active and Enterprise-oriented Authorization Management[C] // Proceedings of the IFIP WG11. 3 Workshop on Database Security. California, USA. Aug. 1997: 13-19
- [7] Cohen E, Thomas R K, Winsborough, et al. Models for coalition-based access control (CBAC)[C] // Proceedings of the 7th ACM Symposium on Access Control Models and Technologies. Monterey, CA, 2002: 97-106
- [8] Kang H M, Park J S, Froscher J N. Access control mechanisms for inter-organizational workflow[C] // Proceedings of the 6th ACM Symposium on Access Control Models and Technologies. 2001: 66-74
- [9] 谭民, 刘禹, 曾隽芳, 等. RFID 技术系统工程及应用指南[M]. 机械工业出版社, 2007
- [10] Ahn G, Sandhu R. Role-based Authorization Constraints Specification[C]. ACM Transactions on Information and System Security, 2000, 3(4): 207-226

(上接第 14 页)

- [21] Plaxton C G, Rajaraman R, Richa A W. Accessing Nearby Copies of Replicated Objects in a Distributed Environment [J]. Theory of Computing Systems, 1999, 32: 241-280
- [22] Padmanabhan V N, Wang H J, Chou P A. Supporting Heterogeneity and Congestion Control in Peer-to-Peer Multicast Streaming [A] // Proc. 3rd Int'l Workshop Peer-to-Peer Systems 2004 (IPTPS). LNCS 3279. Springer, 2005: 54-63
- [23] Zhou M, Liu J. Tree-assisted Gossiping for Overlay Video Distribution [J]. Kluwer Multimedia Tools and Applications, 2005
- [24] Meddour D-E, Mushtaq M, Ahmed T. Open Issues in P2P Mul-

timedia Streaming [A] // MULTICOMM. 2006

- [25] Ratnasamy S, Francis P, Handley M, et al. A Scalable Content-addressable Network [A] // SIGCOMM '01. San Diego, CA, August 2001
- [26] Jenkins K, Hopkinson K, Birman K. A Gossip Protocol for Subgroup Multicast [A] // International Workshop on Applied Reliable Group Communication (WARGC 2001). Apr. 2001
- [27] Peng Gang. CDN: Content Distribution Network [R]. Research Proficiency Exam Report. Computer Science Department, SUNY at Stony Brook, NY, USA 2003, 2