

基于 P2P 的流媒体直播技术与展望

彭 凯¹ 武 娟² 杨宗凯¹ 唐 宏²

(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

摘 要 覆盖网络的组织与维护对于 P2P 流媒体直播系统的性能有着极其重要的影响。根据覆盖网络拓扑组织形式的不同,分别研究了基于单树结构、多树结构、随机拓扑结构的现有 P2P 流媒体直播技术方案,并结合节点的动态行为和传输延迟对方案的优缺点进行了分析比较。以此为基础概括总结有待解决的主要问题,展望了未来的研究方向,为基于 P2P 流媒体直播的进一步的研究和应用提供有益的参考。

关键词 对等网络,视频直播,流媒体,覆盖网

中图法分类号 TP393 **文献标识码** A

Research and Prospect on P2P Live Streaming System

PENG Kai¹ WU Juan² YANG Zong-kai¹ TANG Hong²

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

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

Abstract The performance of a P2P live streaming system is mainly determined by how to build up and maintain its overlay network. We surveyed existing P2P live streaming protocols by which the overlay networks are respectively based on single tree, multiple tree and stochastic topology. Considering the dynamic behaviors of peers and the latency of transmission, we compared the existing protocols and summarize some open issues for further studies and deployment of P2P live streaming.

Keywords Peer-to-Peer network, Live streaming, Video streaming, Overlay network

1 引言

随着因特网规模的不断扩大和网络用户的急剧增加,流媒体应用以其特有的娱乐性和交互性正在成为推动未来宽带应用的主动动力。基于客户机/服务器(Client/Server, C/S)模式的流媒体直播系统,由于服务器性能和网络带宽两个方面的基本限制,无法用于大规模用户网络环境。组播技术虽然能够有效解决流媒体直播系统中的服务器和网络负载问题,但由于其可扩展性、可靠传输、拥塞控制以及部署复杂性等方面的问题,目前还难以广泛部署。

将对等(Peer-to-Peer, P2P)网络技术引入流媒体直播系统,能够显著地减少服务器负载和网络带宽的消耗,具有组播技术的优点,但无需升级路由节点,部署简单方便。相关研究已提出多种 P2P 流媒体直播技术方案,如 PeerCast^[1], Zig-Zag^[16], Coopnet^[20], Splitstream^[25], CoolStreaming^[26] 等。这些方案分别按照单树、多树和网状结构组织节点拓扑结构。各种直播技术方案都是以解决 P2P 直播系统面临的节点动态性,启动时长等问题,减少节点动态行为对系统的影响,以保证视频播放的可用性和连续性为目标^[1,2]。

本文以 P2P 直播系统中的节点拓扑结构为线索,对基于

P2P 的直播的基本概念、已有技术方案及原理进行分析,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的参考。本文其他部分组织结构如下:第 2 节分析了基于 P2P 的流媒体直播系统结构;第 3 节讨论了基于单棵树直播方案;第 4 节讨论了基于多棵树直播方案;第 5 节讨论了基于随机拓扑的网状结构直播方案;第 6 节对各种方案的优缺点进行了综合比较,指出了流媒体直播中需要进一步研究解决的问题;最后总结全文。

2 P2P 流媒体直播简介

P2P 网络也称为对等网络,其中每个节点(称为 Peer)具有相同的地位,既是资源(服务和内容)提供者,又是资源(服务和内容)获取者。对等节点之间通过共享信息、存储和计算资源,无需集中式服务器,即可协同完成相应的任务或资源就可完成。

参与某一特定 P2P 会话(如观看同一视频直播、下载同一文件等)的 Peer 及其相互之间的服务连接关系,构成一个逻辑意义上的覆盖网络(Overlay Network)。覆盖网是一种计算机网络,它是建立在另一个网络之上的叠加网络。在叠加网络中的节点采用虚拟或逻辑链接,在唯一的路径上进行

到稿日期:2008-01-04 本文受中国博士后科学基金:可控 P2P 网络直播算法及其大规模仿真研究(42 批),国家自然科学基金:无线 mesh 网交互式流媒体分发研究(60773193)资助。

彭 凯(1976-),男,博士后,研究方向为现代信息网络技术,E-mail:pkhust@163.com;武 娟(1977-),女,硕士,研究方向为数据通信技术;杨宗凯(1963-),男,博士生导师,研究方向为现代信息网络技术;唐 宏(1974-),男,硕士,高级工程师,研究方向为数据网络与业务。

消息交互,该路径可能经过了很多物理链接。

流媒体直播对于网络数据传输有着较为严格的时限要求,而且持续时间较长。与以服务器为核心的 C/S 模式不同,P2P 网络中的对等节点具有很高的自治性和随意性^[3,4],可以随时加入和离开。因此,如何组织覆盖网络拓扑,在满足流媒体传输和播放要求的同时减少节点动态行为带来的影响,对于 P2P 流媒体直播系统的性能有着决定性的影响^[5,6,8]。

P2P 流媒体直播覆盖网络拓扑的组织主要有单树、多树和随机拓扑 3 种方式。后续几节对基于这 3 种覆盖网络的 P2P 流媒体直播技术方案分别进行介绍。

3 基于单树的 P2P 直播系统

基于单树的 P2P 流媒体直播系统中,所有参与同一流媒体直播会话的 Peer,构成一棵以媒体源服务器为根的树。

Deshpande 等最早在其原型系统 SpreadIt SI 中提出了一种组播树构造协议^[9]:新节点加入时,从源服务器也就是根节点出发沿组播树进行搜索,直到发现某个非饱和(Unsaturated)有剩余带宽资源的节点为止,并让新节点作为该节点的子节点加入系统;当节点主动离开系统时,首先向其父节点进行注销(Unsubscribe),并向其每个子节点发送一条重定向(Redirect)消息,以便让它们直接加入离开节点的父节点或从根节点重新开始加入搜索过程;对于节点的失效,在所有父子节点之间维持心跳(heart-beat)功能,如果在某段时间内收不到父节点的心跳信号,就认为父节点已经失效,并重新向根节点发起加入搜索过程。SpreadIt 的主要问题是一旦节点离开或失效,其子节点都将成“孤儿”节点,并在一般情况下需要从根节点开始重新发起搜索加入过程,从而导致子节点服务被中断的时间较长。

ESM: ESM(End System Multicast)^[10]是由美国 CMU 大学开发的一个基于对等网络的视频直播系统,可以直播一段实时视频或音频节目,应用于中小型规模的会议或者虚拟教室。ESM 采用 Narada^[10]协议构建一棵自组织的、分布式的、应用层上的多播树。它的多播树构建分两步:首先利用 Narada 构建一个丰富的连接网(mesh),由 Narada 协议保证图中任意一对节点之间最短路径时延,并且图中每个节点的邻居个数有上限;然后利用已有的连接图构建一个最短路径树。但是这样构建的多播树的质量取决于连接图信息的准确和丰富程度,连接图的质量直接影响系统的 QoS,另一方面构建和维护连接图的负载将会直接影响整个系统的扩展性能。

PeerCast^[11]:PeerCast 是美国 Stanford 大学 Peer-to-Peer 研究小组的研究成果。它采用一个简单的单根树结构,通过限定每个 peer 所能服务的节点个数来优化节点间的传输延迟。为了减少拓扑结构的动态变化对播放质量的影响,它采用了冗余连接的方式。拓扑的构建采用简单的重定向技术。由于节点放置的策略直接影响到系统的稳定性和可用性,目前 PeerCast 主要用于小规模因特网无线电广播,并已经成功地运行于已有的 Gnutella 网络上。

ALMI^[12]:ALMI 采用了集中式的方式构建多播树。在 ALMI 协议的系统中存在一个中心节点,该节点负责维护整个覆盖网络中的所有连接信息。当某个节点加入系统时,会

首先访问中心节点,以获得其最近的服务节点。但是要构建一棵出度有限的有向树,使其达到最小化的平均延迟,是一个 NP 完全问题,所以集中式选取父节点的方法很难适用于大规模视频流媒体服务系统。

Maryland 大学 Banerjee 等基于分层簇(hierarchical cluster)的思想提出了一种分布式自适应组播树构造协议 NICE^[13],它把参与活动的节点划分为多个簇,每个簇的大小为 $[k, 3k]$, k 作为系统参数指定。分层簇的形成如图 1 所示。新节点在加入时,从根节点开始搜索,直到在最底层的某个簇中找到合适位置;节点在离开或失效时,仅需要经过簇内协商即可为其子节点找到新的父节点;当节点的加入离开行为,使得簇的大小不在 $[k, 3k]$ 范围时,需要进行簇的拆分或合并。NICE 协议中对组播树的维护和管理是分布自治的,每个节点只需要维护其父子节点以及同一簇内其它节点的状态信息。这个系统的优点是较低的控制开销和数据传输过程中的低延迟。NICE 的主要问题在于它的协议比较复杂,需要考虑群的合并、分裂等操作,并且层次越高的节点其负载越重。如对最高层次的节点而言,其子节点的数目达到了 $O(k \log N)$ (N 为整个系统中 Peer 节点的数目)。当系统规模很大时,位于较高层次的节点就可能成为系统的性能瓶颈。一旦群的总领节点失效,群的重建开销很大。

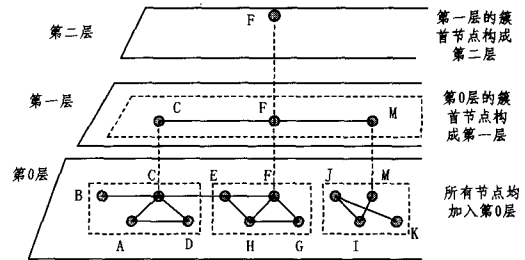


图 1 NICE 协议的分层簇结构

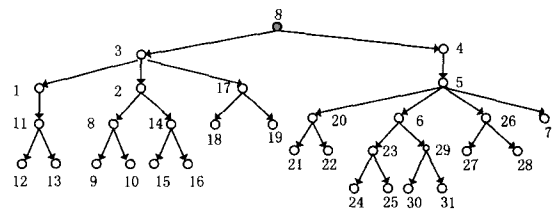


图 2 ZIGZAG 协议中组播树结构

针对 NICE 协议中层次越高的节点负载越重的问题,Tran 等提出了一种新的基于分层簇思想的组播树构建协议 ZigZag^[14],其组播树结构如图 2 所示。ZigZag 是一种基于对等网络的单源传输视频流媒体数据的技术,它构建了一棵以服务器为根的应用层的多播树,通过设定这棵树的节点个数和层次,达到减小端对端时延的目的。ZigZag 最大的贡献是在无法预知整棵树的变化情况下,通过一定的预先控制来维护树,在不影响现有节点、不给服务器带来任何负担的情况下达到局部容错的特点。ZIGZAG 组播协议中簇的层数即组播树的高度与 NICE 同为 $O(\log N)$,但每个节点所带子节点的数目最多为 $O(k^2)$,与节点在组播树中所处的位置层次及系统中总的节点数 N 无关。ZigZag 提供了完整的构造多播树的规则,保证树的深度维持在 $O(\log N)$ 。此外,ZigZag 还有如下的特点:节点数目 N 可以任意多;节点离开而引起的树

的修复限制在局部;控制协议的开销低;新节点加入快速,簇的维护成本低并独立于 N 。但是在 ZigZag 中,管理组织和多播树之间的映射,以及因为网络的动态变化而对这两种结构的维护造成了过多的资源耗费,而且 ZigZag 对节点的异构性处理也存在着缺陷。

在基于单棵树的直播方案中,由于节点的离开或失效都将导致其后继子节点传输服务的中断,而在此前提下为维护每个节点上视频播放的连续性,Guo 等提出了一种基于 Patching 流的数据分发模式^[16];源服务器节点除了用一条频道分发正常的直播数据外,还额外启用多条频道,每条频道用于传输一条 Shifting 流,这些 Shifting 流之间存在固定的发送启动延迟。当节点服务被中断时所产生的数据“空洞”被 Patching 流填充完毕后,随即释放 Patching 流所占用的数据频道,从而恢复仅从组播树上单个节点获取数据的状态。上述模式中的 Patching 流均直接从源服务器节点获取。为了减少对源服务器数据频道的占用,Guo 等还提出了另一种改进 Patching 流的数据分发模式^[17],它要求每个 Peer 节点缓存固定时长的其最近所播放过的数据。当节点的传输服务被中断时,让节点从其它的 Peer 节点获取 Patching 和正常流,省略了在服务器上额外保留数据频道的要求。在一定的播放启动延迟设置条件下,上述数据分发方式可让节点的播放保持连续性,但它要求服务被中断的节点在数据恢复的过程中同时获取 Patching 流和正常流,因此要求每个 Peer 节点必须至少拥有两倍于流速率的有效下行带宽资源。

为减少节点离开失效或链路故障对 Peer 节点播放质量的影响,Banerjee 等提出了一种联合随机转发(Randomized Forwarding)技术及 NAK 触发(Triggered NAK)机制的流媒体直播 SRMS 系统^[18,19]。在 SRMS 中,每个节点除了给其子节点转发数据外,还从系统中随机挑选出固定数目的 Peer 节点,并以较小的概率向这些节点转发数据;而对任一节点 X 而言,在给另一节点 Y 所转发的每个数据包中还须捎带(Piggy backs)一些数据元信息,用以表示节点 X 已正确收到了该数据包之前的那些数据包。在 SRMS 中,以多大的概率向固定数目的 Peer 节点转发数据是系统设计的一个关键。如果概率选择过大,则可能会导致节点接收到大量的重复报文,而造成带宽资源的浪费;如果概率选择过小,当节点离开失效或链路发生故障时,子节点可能需要经历较长的等待时间才能有可能会发送 NAK 报文,从而使得传输服务被中断的时间也较长。

由上述介绍可知,基于单树的 P2P 流媒体直播系统的主要优点包括:

- ①结构简单,数据流向单一;
- ②加入和退出的速度较快;
- ③信令简单,易于管理和控制。

相应的缺点则包括:

- ①树的深度难于控制,当树的深度达到一定层次之后,根节点到叶子节点的延迟难以估量;

- ②当面对大量在线节点频繁地加入、退出时,组播树难于平衡;

- ③针对带宽非对称节点,特别是当节点的上行带宽小于节目自身码流时,该节点难以向后继节点提供满意的服务;

- ④当组播树中的关键节点意外退出或者服务能力不能满

足时,后继节点的服务质量难以保证。

4 基于多树的 P2P 直播系统

基于多树的 P2P 直播系统中,数据流在源服务器端被分割成多条子流,每条子流用一棵单独的组播树进行传输,每个 Peer 节点可以根据自己的带宽资源情况加入到单棵或多棵组播树中。

针对基于单棵树的直播方案中节点离开或失效行为会中断后续子节点数据传输的问题,Padmanabhan 等首先提出了一种基于多棵树的直播方案 Coopnet^[20,21]。CoopNet 是 Microsoft Research 的第一个多树直播视频流媒体网络系统,是主要面向现场直播的流媒体的应用。它首先在源服务器端对视频节目进行 MDC^[22](Multiple Description Coding)编码,产生多条 MDC 子流,这些 MDC 子流在解码时不存在依赖关系;而对每一条 MDC 子流均采用一棵独立的组播树进行分发;新节点在加入系统时,可选择同时加入到多棵组播树;对任一节点而言,其接收到的 MDC 子流数量越多,则播放质量越好;当某棵组播树上的父节点离开或失效时,只会导致一条 MDC 子流的数据传输被中断,节点还可以从其它组播树上继续接收其它的 MDC 子流,在这种情形下除了画面质量会有所下降外,并不影响节点对节目数据的连续播放;对于多棵组播树的组织和维护,主要由中心索引服务器集中承担,节点在加入或重新加入时均向中心索引服务器提交请求,以获取多个或单个父节点。

微软研究院的 Castro 等认为,在基于单棵组播树的直播方案中存在如下两方面的矛盾:一方面组播树上的中间节点承担着较重的数据传输负载,另一方面还有大量叶节点的带宽资源得不到有效利用。针对此问题,他们提出了一种基于多棵组播树的直播模型 SplitStream^[25],它与 CoopNet 一样在源服务器端采用 MDC 编码,并让每棵组播树对应传输一条 MDC 子流;不同之处是在 SplitStream 中,每个 Peer 节点只在其所参与的一棵组播树中充当中间节点,而在剩余的其它组播树中只充当叶节点,这样就可以减小节点离开或失效对其它节点所造成的影响。它旨在提供纯粹的对等网络服务。在 SplitStream 中每棵组播树的构造均基于应用层组播系统 Scribe^[26,27],而后者建立在结构化 P2P 路由机制 Pastry 之上,并且采用了 DHT(Distributed Hash Table)技术^[28],使得它的查找速度得以优化。SplitStream 的关键思想是把内容分成 k 个带(stripe),并且每个带利用一棵独立的树来广播。节点根据它们愿意接收的带的个数,加入同样多的多播树中,每个节点还给定一个它们愿意转发的带的上限。问题转变为构造一个多播树的森林。在这个森林中一个在某棵树中是内部节点的节点在加入其它树中都是叶子节点,并且满足节点的出度限制(即转发的带的个数)。这可以保证转发负载散布到所有参与的节点。图 3 描绘了 Splitstream 中的带和多播森林。

设计 Splitstream 面临的关键挑战是以一种分布式的、可以扩展的、高效的并且自组织的方式构造一个由内部节点正交的多播树森林,一个节点最多在一棵树中担当内部节点,而在其它树中为叶子节点。Splitstream 利用 Pastry 的路由和 Scribe 的组成员管理来辅助内部节点正交树集的构建。SplitStream 的底层实现是基于分布式散列表(Distributed

Hash Table, DHT) 覆盖网络, 它依赖于 Pastry 实现和 Scribe 实现。

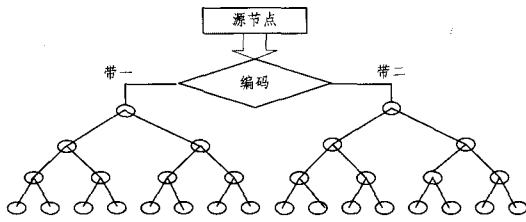


图3 Splitstream 中的多播森林

与基于单棵树的直播方案相比, 以上两种方案都建立在 Microsoft Research 提出的多表述编码 MDC 的基础上。多个多播树结构的最大好处在于它能提供良好的容错性, 这对于拓扑结构呈动态变化的对等网络系统来说是非常重要的。多树直播的出发点一方面在于解决单树状结构叶子节点只获得资源而不向外提供资源的浪费, 另一方面可以通过在覆盖网中部署 MDC 编码来实现节点的多个数据获取源, 从而解决单树结构带宽不对称的先天不足。通过上述分析, 可以分析出多树结构的优缺点。

多树结构的优点:

- ①提高了节点的带宽利用率;
- ②带宽不对称节点在播放码率大于上传码率的节目时可以降低其服务能力, 仅向其他节点提供一部分节目数据而不会影响其他节点的服务质量;
- ③MDC 编码的部署可以从一定程度上提升用户的主观观看质量, 面对网络拥塞或者抖动都较易于处理。

多树结构的缺点:

- ①组播森林大大增加了组播结构的复杂度;
- ②树的深度同样难于控制, 当树的深度达到一定层次之后, 组播森林中不同组播树的延迟可能会不均衡;
- ③当大量在线节点频繁地加入、退出时, 组播树难于平衡; 某棵组播树的不平衡甚至会导致整个组播森林的崩溃;
- ④组播森林中的关键节点意外退出或者服务能力不能满足时, 后继节点的服务质量同样难以保证;
- ⑤节点的加入、退出机制更加复杂, 加入时间更长;
- ⑥组播森林的维护开销过高, 且开销随着组播树的数目呈指数级递增。

多树结构以组播森林的方式解决了单树结构的不少问题, 但是由于其仍然基于树状拓扑的组播结构, 很多树状结构的基本问题仍难于解决。另外, MDC 编码理论上虽然已经成型, 但是实际应用上还有不少工程问题没有解决, 至今仍然不甚实用; 而且由于 SplitStream 的实现和部署过于复杂, 至今仍然没有一个可应用于验证和演示的系统实现。

5 基于随机拓扑的 P2P 直播系统

基于随机拓扑的 P2P 直播系统中, 新节点在加入时随机挑选多个 Peer 节点作为其邻居节点, 并通过运行相关的数据调度算法来从多个邻居节点同时调度获取数据, 其数据传输的模式也从传统的“推”(Push)模式转化为“推”和“拉”(Pull)相结合的模式。由于一个节点可同时从其它多个节点获取数据, 因此可在很大程度上降低 Peer 节点离开或失效行为对其它节点所造成的影响。

基于网状结构的视频流媒体网络模型 DoNet, 是近年来所提出的基于随机拓扑结构应用层的多播的典型例子。该方法的主要思想是利用轻量级的 Gossip 协议来构建一个应用层的覆盖多播网。在网中的每个节点都维护两张表, 其中一张是网络中的所有成员节点的列表, 另外一张是和本节点交换媒体数据的伙伴节点列表。由于该方法使用了轻量级的协议来维护覆盖网络并且使用了网状结构, 其可扩展性以及可靠性都得到了极大的提高。然而, 该方法使用了随机的调度和选择伙伴节点的方法, 而没有考虑上层覆盖网络和底层物理网络的拓扑匹配问题。DONet 的一个简单的网络拓扑图, 如图 4 所示。

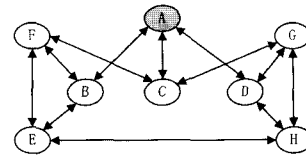


图4 DONet 网络拓扑图

DONet 的系统架构如图 5 所示, 它主要包含 3 个模块。(1)节点的管理 (Membership Manager): 负责维护系统中部分其它节点的视图; (2)伙伴管理 (Partnership Manager): 和其它节点建立协作关系; (3)数据调度: 负责动态地和其它节点, 交换数据。在 DONet 中, 每个节点既是数据的接收者, 也是数据的提供者。服务器是一个特殊的节点只作为数据的提供者, 称为源节点。

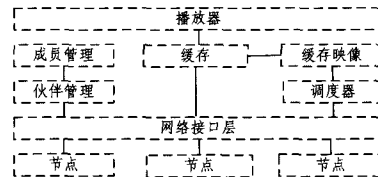


图5 DONet 系统架构图

节点管理 (Membership Manager): DONet 中每个节点有一个在整个系统中唯一的标识, 比如 IP 地址, 并且维护一个系统中其它节点的标识的缓存 mCache。当新节点加入时, 首先请求源节点, 源节点从它的 mCache 中随机地选择一个节点作为新节点的代理。新节点从代理节点获得初始的伙伴节点的列表。在 DONet 中, 消息的发送部署采用 SCAM (Scalable Gossip Membership Protocol) 协议。

伙伴管理 (Partnership Manager): 在 DONet 中, 节点的伙伴及伙伴之间数据的传输方向并不固定, 伙伴之间根据各自缓存的数据情况进行数据交换, 所以节点和伙伴需要相互知道所缓存的数据的内容。节点和伙伴通过不断交换缓存信息块来了解相互间的缓存情况。

数据调度 (Scheduler): 调度的目的就是如何从伙伴节点获取数据。在 DONet 中, 数据的调度采用最多供应优先的算法。在保证 Deadline 的情况下, 选择带宽最大的提供者。算法被周期性地执行更新调度。

Zhou 等人针对 VoD 应用需求中也提出了一种基于随机拓扑的系统 TAG^[27]。它与 DONet 的主要区别是新加入的节点在初始化其邻居节点列表时, 并不是从整个系统范围内随机挑选节点, 而是仅仅从那些缓冲了目标数据段的节点集合中随机挑选邻居节点。上述系统在数据获取方式上属于“推”

和“拉”相结合的模式。与基于多棵组播树的直播方案比较,基于随机拓扑的直播方案的最大优势是既不需要在 Peer 节点间维护复杂的网络逻辑拓扑结构,也不需要源服务器端把数据流切分为多条子流。

DONET 提供的是一个近似图状的非结构化组播模型,在这一模型中数据的流向不再单一而更加多变,需要引入专门的数据分块和调度机制,以适应这一结构。

基于随机拓扑的 P2P 直播系统的优点主要在于:

- ①结构上具有更好的可扩充性,可以支持更多的用户;
- ②较好地解决了异构节点和不对称节点带来的服务质量下降问题,更加适宜在 Internet 上使用;
- ③稳定性较好,节点的非正常退出对整个组播结构影响较小,组播结构也易于平衡;
- ④延迟相对较小,相对于同等用户数目的树状结构而言,延迟较低。

其缺点主要体现在如下几个方面:

- ①由于增加了数据缓冲,节点从加入到第一次播放的等待时间稍长;
- ②当在线用户较多时,实际拓扑情况难于测量和管理;
- ③物理网络拓扑适应能力与拓扑构建随机性匹配较差;
- ④数据调度算法效率不高,控制开销相对较高。

6 比较与分析

如上所述,已有的 P2P 流媒体直播技术方案根据覆盖网络拓扑组织方式的不同,分为单树、多树、随机拓扑 3 类,其主要特点和比较如表 1 所示。

表 1 P2P 流媒体直播技术方案比较

系统特点	组播树		随机拓扑	
	单棵树	多棵树		
网络特性	可扩展性	差	中	好
	网络稳定性	差	好	好
数据特性	视频编码	可用	需要	可用
	负载均衡	差	好	好
	通信开销	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	差
应用性	实现	简单	复杂	简单
	应用场景	小规模	大规模	大规模
	管理复杂度	低	高	中

针对 P2P 流媒体直播系统设计中的主要问题,现有研究已提出了多种解决方案。但对于大规模网络中的实际部署,P2P 流媒体直播仍然存在如下问题有待于进一步研究:

• Overlay 与物理网络的一致性

P2P 网络在 Internet 网络分层模型中属于应用层上的 Overlay。在 P2P 流媒体中,Peer 节点所组成的网络逻辑拓扑建立在底层的物理拓扑之上,并且在一般情况下这两种网络拓扑并不具有一致性;而这种不一致性会使得 Peer 节点经常从距其较远而不是较近的其它 Peer 节点获取媒体数据,导致播放延时的问题。因此,在构建 Peer 节点的网络逻辑拓扑时,如何对底层的物理网络逻辑拓扑进行快速有效的探测,并使前者在构造时尽量多地与后者保持一致,提高系统的总体性能,减少媒体播放延迟,是今后 P2P 流媒体直播中一个值得研究的方向。

得研究的方向。

• Small world 理论与 P2P 直播策略的结合的研究

非结构化 P2P 直播系统中节点发现技术一直采用随机转发的方式,与 DHT 的启发式发现算法相比,可靠性差,对网络资源的消耗较大。最新的研究从提高发现算法的可靠性和寻找随机图中的最短路径两个方面展开,也就是对重叠网络的重新认识。其中,Small world 特征和幂规律证明实际网络的拓扑结构既不是非结构化系统所认识的一个完全随机图,也不是 DHT 发现算法采用的确定性拓扑结构。大量研究证明了以 Gnutella 为代表的 P2P 网络符合 Small world 特征,也就是网络中存在大量高连通节点,部分节点之间存在“短链”现象。因此,P2P 发现算法中如何缩短路径长度的问题变成了如何找到这些“短链”的问题。P2P 媒体网络具有小世界现象的 W-S 模型特征,随着短链的增加,P2P 媒体网络的节点间的平均路径长度减小,聚合程度加大,节点的可达路由增多,网络处理负载的能力增强。在媒体资源的查找过程中,实现媒体资源平均传输时间的减少以及媒体资源搜索和查询效率的提高。如何产生和找到“短链”,是发现算法设计的一个新的思路。Small world 特征的引入会对 P2P 资源发现算法产生重大影响。

• 延时与抖动性研究

从现有几款 P2P 流媒体系统的运行效果来看,都存在播放启动延时过长,而且在播放过程中由于受 Peer 服务节点离开或失效的影响而导致播放质量出现一定程度的抖动。而在混合结构网络中,由于存在固定部署的缓存服务器比普通的主机节点具有更高的处理能力、更好的稳定性和安全性,缓存服务器可以同时服务于若干个 Peer 节点。因此,在 P2P 流媒体内容直播研究中如何通过缓存服务器的帮助,来减少 Peer 节点的播放启动延迟,并在播放过程中减少播放质量的抖动,同时尽量减少对网络带宽资源的分配占用,也是今后值得研究的一个方向。

• 激励机制

在 P2P 媒体环境中需要参与直播服务的 Peer 节点尽量贡献自己的资源,以提高网络上的资源利用率。但在实际应用中,许多用户更倾向于以“搭便车(free-riding)”的形式参与系统,即只愿意从其它节点获取服务而不愿意贡献出本身的空闲资源,为其它节点提供服务。这种行为并不能充分利用 Internet 上的空闲资源,也不利于充分发挥 P2P 模式的优点。因此,如何建立一套有效的激励机制,使得 Peer 节点在行使消费者角色的同时,主动承担为其它节点提供服务的职责,也是 P2P 流媒体直播中一项值得研究的课题。

• Peer 节点异构性的处理

对 P2P 网络中的 Peer 节点而言,它们所具备的网络带宽资源、主机处理能力实际上是有差异的。不同 Peer 节点对视频播放的质量等级有不同的需求,这种质量等级可体现在视频的分辨率、帧率等方面。如何满足 Peer 节点对目标节目异构性的质量需求,同时如何适应并利用 Peer 节点异构的资源能力,也是 P2P 流媒体直播技术研究中的挑战之一。

结束语 覆盖网络拓扑的组织对于 P2P 流媒体直播系统的性能有着决定性的影响。本文主要根据覆盖网络拓扑组织方式的不同,全面介绍了各种 P2P 流媒体直播技术方案,并结合流媒体直播应用的特点对其进行分析和比较,以此为

基础提出有待于进一步研究的主要问题和挑战。

参 考 文 献

- [1] Chawathe Y. Scattercast: An architecture for Internet broadcast distribution as an infrastructure service. Ph. D Thesis, University of California, 2000
- [2] Deering S. Multicast routing in internetworks and extended LANs // Proceedings of ACM SIGCOMM. Stanford; ACM Press, 55-64
- [3] Shi S, Turner J. Multicast routing and bandwidth dimensioning in overlay networks. IEEE Journal on Selected Areas in Communications, 2002, 20(8):1444-1455
- [4] Banerjee S, Kommareddy C, Kar K, et al. Construction of an efficient overlay multicast infrastructure for real-time applications // Proceedings IEEE INFOCOM. San Francisco, USA; IEEE Communication Society, 1521-1531
- [5] Banerjee S, Kommareddy C, Kar K, et al. OMNI: An Efficient Overlay Multicast Infrastructure for Real-time Applications. Accepted for publication in Special Issue of Computer Networks on Overlay Distribution Structures and their Applications
- [6] 吴家皋, 叶晓国, 姜爱全. 一种异构环境下覆盖多播网络路由算法. 软件学报, 2005, 16: 1112-1119
- [7] Xu D, Chai H-K, Rosenberg C, et al. Analysis of a Hybrid Architecture for Cost-Effective Streaming Media Distribution // Proceedings of SPIE/ACM Conference on Multimedia Computing and Networking (MMCN). Santa Clara, CA, January 2003
- [8] 刘亚杰. P2P 流媒体内容分发关键技术研究. 工学博士学位论文. 国防科技大学, 2005
- [9] Deshpande H, Bawa M, Garcia-Molina H. Streaming Live Media over a Peer-to-Peer Network. Technical Report. Stanford Database Group(2001-20), August 2001
- [10] Chu Y H, Rao S G, Zhang H. A Case for End System Multicast // Proc. of ACM SIGMETRICS. Jun. 2000
- [11] Bawa M, Deshpande H, Garcia-Molina H. Transience of Peers & Streaming Media. Technical Report. Stanford University, 2002
- [12] Pendarakis D, Shi S, Verma D, et al. ALMI: An Application Level Multicast Infrastructure // Proceedings of 3rd Usenix Symposium on Internet Technologies & Systems (USITS 2001). San Francisco CA, March 2001
- [13] Banerjee S, Bhattacharjee B, Kommareddy C. Scalable application layer mul-ticast // Proceedings of ACM SIGCOMM 2002. Pittsburgh, USA; ACM Press, 2002
- [14] Zigzag T D. An efficient peer-to-peer scheme for media streaming // Proc of IEEE INFOCOM03. San Francisco, USA, April 2003
- [15] Yang M, Fei Z. A Proactive Approach to Reconstructing Overlay Multicast Trees // Proceedings of IEEE INFOCOM. Hong Kong, IEEE Communications Society, March 2004
- [16] Guo M, Amman M. Scalable live video streaming to cooperative clients using time shifting and video patching // Proceedings of IEEE Infocom. Hong Kong, IEEE Communications Society, March 2004
- [17] Guo M, Ammar M, Zegura E. Cooperative Patching: A client based P2P Architecture for supporting continuous live video streaming // Proceedings of IEEE International Conference on Computers, Communication and Networks (IC3N). Chicago, October 2004
- [18] Banerjee S, Lee S, Braud R, et al. Scalable resilient Media Streaming // Proceedings of 14th ACM International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV). Kinsale, County Cork, Ireland. June 2004
- [19] Banerjee S, Lee S, Bhattacharjee B, et al. Resilient multicast using overlays // Proceedings of ACM Sigmetrics. San Diego, CA, USA, June 2003
- [20] Padmanabhan V N, Wang H J, Chou P A, et al. Distributing Streaming Media Content Using Cooperative Networking // Proceedings of International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV). Miami Beach, FL, USA, May 2002
- [21] Padmanabhan V N, Wang H J, Chou P A. Resilient Peer-to-Peer Streaming // Proceedings of IEEE International Conference on Network Protocols. Atlanta, GA, USA, November 2003
- [22] Goyal V K. Multiple Description Coding; Compression Meets the Network. IEEE Signal Processing Magazine, 2001, 18(5): 74-93
- [23] Chow P A, Wang H J, Padmanabhan V N. Layered multiple description coding // Proceedings of IEEE Packet Video Workshop. Names, France, April 2003
- [24] Padmanabhan V N, Wang H J, Chou P A. Supporting Heterogeneity and Congestion Control in Peer-to-Peer Multicast Streaming // Proceedings of Third International Workshop on Peer-to-Peer Systems (IPTPS). San Diego, CA, USA, February 2004
- [25] Castro M, Druschel P, Kennarrec A-M, et al. SplitStream: High-bandwidth multicast in a cooperative environment // Proceedings of 19th ACM Symposium on Operating Systems Principles (SOSP). Lake Bolton, New York, October 2003
- [26] Zhang X, Liu J, Li B, et al. DONet/CoolStreaming: A Data-driven Overlay Network for Live Media Streaming. Technical Report. June 2004
- [27] Zhou M, Liu J. Tree-assisted gossiping for overlay video distribution. Kluwer Multimedia Tools and Applications, 2005

欢迎订阅《计算机科学》杂志