

# 面向 Internet 的流媒体组播技术<sup>\*</sup>

叶保留 吴敏强 顾铁成 陈道蓄

(南京大学计算机软件新技术国家重点实验室 南京210093)

## Streaming Media Multicast over the Internet

YE Bao-Liu WU Min-Qiang GU Tie-Cheng CHEN Dao-Xu

(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

**Abstract** The multicast of Streaming Media over the Internet is an important component of many current and future Internet applications. The heterogeneity and lacking of QoS guarantee make the application of this technique a challenge. In this paper, we analyze the difficulties in streaming media multicast design, survey technologies that have been proposed for it. We also discuss several congestion control and error control mechanisms. Finally we present some remained issues for further research.

**Keywords** Multicast, Streamed media, Fairness, TCP-friendly, Congestion control, Error control

## 1 概述

多媒体和现代网络技术的发展使流媒体应用(如视频会议、远程教育、视频点播等)成为下一代 Internet 服务的一个重要分支。流媒体传输持续时间长,并且用户对流媒体对象的访问往往呈现在特定黄金时段内,请求集中于少数热门对象,采用简单的多单播实现途径,一方面使得共享链路上可能存在多个复制数据报文,造成系统资源浪费并很快耗尽网络资源,另一方面需要在发送者和接收者集之间建立应用级的同步机制。组播(multicast)是一种可控制的单点对多点的数据传输技术,数据源发送一个数据拷贝,报文在网络传输过程中可根据需要自动复制,避免了单播方式下为每个接收者建立通信连接,并单独发送数据信息的缺点。组播能节约网络资源,减少网络管理要求,是下一代 Internet 应用的一项关键技术。

近年来,基于 Internet 的视频组播技术研究工作包括以下两个方面<sup>[1]</sup>:

(1)将对服务质量(QoS)的支持集成到 Internet,为视频服务提供资源预留并为数据传输延迟、延迟抖动和丢包率确立上界。

(2)采用闭环策略,使用端到端的自适应速率控制技术,根据可用网络资源动态调整视频流特征。

第一种方法以网络为中心,在网络层考虑视频流的 QoS 特性,需要中间节点如路由器、交换机等支持,因此技术实现涉及到对网络基础设施的更新和扩充,第二种方法以终端系统为基础,试图为流媒体传输提供端到端的 QoS 保证,相关技术旨在通过应用层设计对视频传输进行控制,该途径更易在短期内与现有 Internet 兼容。本文结合第二种途径对相关技术及实现进行了总结。

## 2 主要问题

面向 Internet 的流媒体组播的研究焦点是如何满足数字

视频信息组播的实时性需求,该焦点要求网络解决两个问题:

(1)提供支持数据传输的组播机制;(2)为视频信息传输提供实时服务。然而,传统 Internet 至少在两个方面制约着流媒体应用:(1)尽力而为(best-effort)服务特性不能提供 QoS 保证,满足流媒体应用的时序性要求;(2)目前,大多数 Internet 应用都采用 TCP 作为传输协议,TCP 的高可靠性是通过接收方为每个报文反馈 ACK 实现的,效率低下,且在组播中易产生反馈风暴(ACK Explosion);同时,TCP 基于窗口的速率调整策略 AIMD(1,1/2)是一个振幅较大的周期性波动曲线,直接应用到流媒体中易产生抖动。

从传输角度讲,流媒体与普通文件在 QoS 方面的主要差异有:(1)实时性。流媒体有较高的实时要求,否则会产生抖动,影响用户可感知 QoS,而文件传输在时间限制方面相对较弱;(2)可靠性。文件传输要求有严格的差错恢复机制来保证高可靠性,而流媒体应用能容忍一定程度的差错,是半可靠的(semi-reliable)。(3)顺序性。流媒体的编码格式(如 MPEG 中的 I、P、B 帧)存在前后相关性,因此,出于解码需要,报文传输必须保证一定的顺序性,而一般文件传输没有时序要求。

基于 Internet 的视频组播需要处理和协调好以下几个问题:

(1)流媒体自身特征与 Internet 的服务特性。流媒体有大容量、高实时性、半可靠性、强时序性、长持续时间等特点,数据传输有较高的带宽需求,必须围绕用户可感知 QoS(如延时、抖动、画面质量)改善和提高现有 Internet 尽力而为的服务特性。

(2)网络的异构性与流媒体组播实现。网络上不同链路、节点及终端设备在带宽容量、处理能力等方面都不尽相同,而流媒体组播是一种典型的“单点对多点”的数据传输模式,研究的一个重要问题是如何解决异构性,实现组成员间公平(intra-session fairness)。

(3)与 TCP 流的友好性(TCP-Friendly)。根据 MCI 的统

<sup>\*</sup> 本文得到国家863高技术项目基金(项目编号:2001AA113050)和江苏省高技术产业化推进项目“基于多服务器群的大规模多媒体信息服务平台的研究和开发”资助。叶保留 博士研究生,主要研究领域为分布式计算与并行处理。吴敏强 硕士研究生。顾铁成 博士研究生。陈道蓄 教授,博士生导师。

计,现有 Internet 上有90%的应用使用 TCP,而可靠组播的实现不能直接采用 TCP,因此相关拥塞控制机制要保证组播流与 TCP 流公平共享瓶颈链路,避免 TCP 流饿死,达到会话间公平(inter-session fairness)。

(4)系统结构的可伸缩性和自适应性。伸缩性在许多系统设计强调的一个基本目标,对多媒体而言伸缩性主要体现在规模的变化,要通过一定的自适应机制弱化和消除规模变化对多媒体组播性能的影响。

### 3 组播流策略

Internet 的异构性是制约高性能组播应用的一个重要因素,组播策略是流控制的基础,直接影响着成员间的公平,文[3]给出了成员间公平性的评价标准。已有视频流控制策略研究工作大致可分成三类<sup>[1]</sup>。

**单个流策略<sup>[2~7]</sup>** 发送源只用一个视频流为所有组成员服务,接收方将自身状况和统计到的网络信息反馈给发送方供速率调整。该方法最直接,但在大规模应用中存在两个缺点:容易产生“反馈风暴(feedback implosion)”而且不能很好解决异构性问题。T. Jiang 等在文[3]引入接收者公平(Receiver Fairness)和接收者间公平(Inter-Receiver Fairness)函数决定发送速率, D. Sisalem 在文[4]采用 LBA(Loss Based Adjustment)机制选择速率,并在文[5]中采用基于等式的调节策略扩展了该机制, I. Busse 等人在文[6]中提出根据拥塞状态用户和正常负荷状态用户分别所占的百分比来决定发送速率, Ch. Bouras 等在文[7]中采用 RTP/RTCP 协议,根据 RTCP 所反馈的网络信息及在保证用户间公平性的前提下,尽量提供与 TCP 流友好的发送速率。

**复制流策略<sup>[8,9]</sup>** 这种方法为解决异构问题对单个流技术做了简单扩展,发送源通过不同的组播地址为同一个视频服务请求开辟多个具有不同质量的组播流,接收者选择一个最能与自身接收能力相匹配的组播流,同时每个流可根据相应接收者的反馈,在一定范围内调整流速率。该方法的缺点是接收者能力会随着网络状况等的变化而变化,因此会在不同流之间跳跃。此外复制流技术所产生的冗余信息增加了网络负担。文[3]将接收者分成若干个子组,各子组接收特定层。Y. Cheung 等在文[8]中通过 DSG(Destination Set Grouping)协议将用户分组并处理、维护服务公平性。

**基于分层的流策略<sup>[17~20]</sup>** 该方法将视频信息分成一个基本层(base layer)和若干个增强层(enhanced layer),并分别在不同的组播地址空间组播。其中基本层提供了最基本的服务质量,可独立完成解码,而增强层提高了服务质量,但解码依赖于基本层及其他相关层。接收者至少必需加入基本层,并根据自身条件尽可能同时接收多个增强层。这种方法有效地解决了用户接入方式的异构性问题,但单个组成员的状态变化与 LAN 中的其他组成员相关,组管理及流间同步控制复杂。该方法有以下几个特点:(1)方案实现与具体的视频压缩技术相关联;(2)能有效解决异构性,但由于依赖于具体的视频编码技术,因此只能提供粗粒度的服务质量。文[10]给出了一种层次式编码方案。RLM<sup>[11]</sup>是一种由接收方引起的基于分层的全分布速率控制算法,接收者通过“加入试验(join-experiments)”来发现自己的接收能力,并通过“共享学习(shared learning)”机制来提高系统伸缩性。文[12]提出了一种带重传的层次式速率控制 HRC(Hierarchical Rate Control)组播方案—LVMR, LVMR 通过重传改善了服务质

量, HRC 将控制信息分布于发送者、接收者及代理(agents)之间, LVMR 通过对知识库的智能分区(intelligent partitioning)避免了 RLM 的缺点,并允许多个试验同时进行。文[13]给出了一种基于分层的 TCP 友好控制算法,但不适合于流媒体应用。

### 4 公平模型与拥塞控制

组播流控制策略着眼于从会话内部解决协议内公平性问题, TCP 是一种保守型协议,高可靠性以牺牲效率为前提,不适合流媒体传输。非面向连接协议 UDP 能为数据传输提供高效服务,但带有一定的侵占性,不能对网络状况做出反应,直接将 UDP 应用到 Internet 中的流媒体传输容易伤害 TCP 流,甚至造成网络瘫痪。因此,流传输协议设计的关键是使流媒体的实时性和 TCP 公平性同时得到满足,协议要能根据客户反馈信息(丢包率、抖动、延迟),自动调节组播流的发送速率,从而有效避免和处理拥塞问题。

协议间公平的评价标准有 TCP 友好公平(TCP-Friendly fairness)、最小公平性(essential fairness)、fi(fairness index)公平性。TCP 友好公平要求拥塞控制机制中吞吐量和丢包率满足  $T \propto 1/\sqrt{P}$ ; 最小公平性要求在瓶颈带宽下,组播流和 TCP 流的平均吞吐量(分别设为  $T, T'$ )满足  $aT' < T < bT'$  (其中  $a, b$  是关于组成员个数的函数); fi 公平性假设有  $N$  个组播流和 TCP 流共享瓶颈带宽,其中  $T_0, T_1, \dots, T_n$  分别为平均吞吐量,则 fi 可定义为  $fi = (\sum_{i=0}^n T_i)^2 / N \sum_{i=0}^n T_i^2$ , fi 取值介于  $1/N$  到 1 之间,如果  $fi = 1$ , 则拥塞控制算法是绝对公平的(TCP 友好的)。相关工作<sup>[13~19]</sup>主要从以下两个角度研究了公平拥塞控制算法。

**基于窗口的拥塞控制算法** 该类算法模拟了 TCP 滑动窗口机制,算法在保证 TCP 友好的前提下避免单个报文丢失引起拥塞窗口减半。基于窗口的拥塞控制算法为拥塞窗口大小的增减定义了不同的规则,无拥塞时接收窗口在每个 RTT 内递增,出现拥塞时接收窗口立即减少。

窗口调整幅度是这类拥塞控制算法的关键, S. Floyd 等在文[17]中给出与 AIMD(1, 1/2)兼容的确定型响应函数 AIMD( $a, b$ )中的  $a, b$  关系表达式,但相关实验结果与理论分析具有一定的偏差。文[18]提出了一种二项式拥塞算法,如公式(1):

$$\begin{aligned} \text{Increase: } w_l + R &\leftarrow w_l + a/w_l^b, a > 0, \\ \text{Decrease: } w_l + \delta &\leftarrow w_l - \beta w_l, 0 < \beta < 1. \end{aligned} \quad (1)$$

算法使增量与当前窗口的  $k$  次幂成反比,而减量与当前窗口的  $l$  次幂成反比,相关工作证明当  $l+k=1$  时可满足 TCP 友好。

J. Golestani<sup>[19]</sup>给出了一种更为通用的接收窗口计算公式,见公式(2):

$$w_i = \begin{cases} w_i + A_i(w_i, \tau_i) \\ w_i - B_i(w_i, \tau_i) \end{cases} \quad (2)$$

算法中  $w_i$  为第  $i$  个接收者窗口,  $\tau_i$  为速率计算的时间间隔,算法一般要求满足条件:  $A_i(w_i, \tau_i) \geq 0, B_i(w_i, \tau_i) \geq 0, A_i(w_i, \tau_i)/B_i(w_i, \tau_i)$  为关于  $w_i$  的递增函数。

根据具体的组播流策略,算法又可分为单窗口单速率和多窗口多速率两类,设计的一个要点是防止“归零(Drop-to-Zero)”现象,一般要求将拥塞控制和差错恢复相结合,以便适应大型异构环境的需要。

这类算法的特点是响应速度快,但波动也更大,很难适应

连续媒体流的应用。

基于速率的控制算法 也称为 TCP 友好速率控制算法。根据对 TCP 拥塞控制机制的分析,在稳定状态下的拥塞控制算法可由与报文大小( $s$ )、往返时间(RTT)、丢包率( $P$ )相关的响应函数来表示。若丢包率较低且不考虑超时(TO)因数,该响应函数可用式(3)来表示:

$$T = \frac{C * s}{RTT * \sqrt{P}} \quad (C \text{ 介于 } 0.9 \sim 1.3 \text{ 之间}) \quad (3)$$

在考虑超时因素且丢包率相对较高时,响应函数可由公式(4)表示:

$$T = \begin{cases} \frac{s}{RTT * \sqrt{P} * [\sqrt{2b/3 + T * \sqrt{27b/8} * \sqrt{P} * P * (1 + 32P^2)}]}, & P \leq \frac{8}{27b} \\ \frac{s}{RTT * \sqrt{P} * [\sqrt{2b/3 + T * P * (1 + 32P^2)}]}, & P > \frac{8}{27b} \end{cases} \quad (4)$$

该公式要求超时(timeout)是 RTT 的整数倍,其中常量  $b$  指收到的 ACK 报文数,一般取值为 2。

基于速率的拥塞控制算法相对简单,易于实现,但也存在几个难点:(1)需要每个接收者都测量、计算并反馈 RTT 及丢包率,计算有效性难以验证,易产生 ACK 风暴;(2)速率调节策略相对困难,容易产生归零现象;(3)由于超时现象的存在,算法对拥塞响应比较慢;(4)很难验证算法的公平性。

## 5 差错控制

Internet 上由于多种业务的存在,包丢失现象不可避免,流媒体传输能容忍一定程度的丢包,但高丢包率直接影响到用户可感知的服务质量,因此必须通过差错控制提高服务质量。Internet 上视频信息的差错控制机制包括 4 类:前向纠错 FEC (Forward Error Correction) 技术,重传 (Retransmission-based) 技术,错误恢复 (error resilience) 和错误隐藏 (error concealment) 技术。FEC、重传方法以及错误恢复需要发送和接收两端的参与,而错误隐藏技术只在接收方执行,错误恢复和错误隐藏技术通过压缩途径,减少丢包对服务质量的影响。流媒体组播常采用前两种及其混合途径来处理差错<sup>[1]</sup>。

### 5.1 FEC 技术

FEC 通过对视频信息增加冗余,出现丢包时可根据冗余信息重建原信息。该方法的优点是具有较低延迟,无需反馈,可避免反馈风暴,但附加信息增加了网络负担,并且在突发丢包超过 FEC 编码恢复能力时会影响系统性能。从编码角度看,FEC 可分为 3 类<sup>[20]</sup>:(1)信道编码(channel coding)。该方法把视频流分成大小为  $k$  的数据段,信道编码器在每个数据段中增加数据位使得段长度达到  $n$  ( $n > k$ ),接收方在收到每段中任意  $K$  ( $K > k$ ) 位信息就能恢复原报文。信道编码使数据恢复完全在客户端执行,适用于任意大小的组播。此外,在带宽或处理能力受限时允许网络和接收者丢弃部分报文,而且对异构性有较强的适应性。该方法的缺点是由于在编码和播放时都要受“等待  $k$  长度”的限制,因此增加了延迟。根据压缩视频流的编码特征,文[21]提出了一种基于优先级的 FEC (Priority Encoding Transmission) 技术,结合 MPEG 中 I、P、B 帧的优先关系对不同信息设置不同的差错保护级别。文[22]提出了一种适合于分层组播的基于接收方的层次式前向纠错技术 HFEC (Hierarchical FEC),HFEC 的关键特征是把根据各个视频层生成的 FEC 冗余信息组织成相应的 FEC 流,接收

者根据自己的接收情况预定相应的 FEC 流。(2)基于源编码的 FEC (source coding-based FEC)。和信道编码一样,SFEC 也通过增加冗余信息来恢复丢失,但两者的冗余信息增加方法不同。SFEC 冗余信息的附加是对原视频的增强压缩,而信道编码则根据块编码添加冗余。出现丢包时,信道编码能提供完全恢复而 SFEC 只能提供质量受损的恢复,但与信道编码技术相比,SFEC 有较低延迟。(3)源/信道组合编码 (joint source/channel coding)。信道编码侧重编码的鲁棒性,而 SFEC 则强调编码效率,组合编码的目标是根据速率变形原理,对给定的总速率和丢包率,优化源编码速率,使得视频信息的变形度最小。组合编码过程包括在 SFEC 和信道编码之间寻找优化的速率分配、根据目标速率设计 SFEC 方案、设计/选择与丢包特征相匹配的信道编码等三个阶段。Davis 和 Danskin 在文[23]提出了一个基于源/信道组合编码技术的 Internet 图片传输方案。

### 5.2 重传技术

流媒体信息重传有着严格的时间限制,只有在播放点之前能到达的重传报文才有效。传统的重传技术如 ARQ 不适合流媒体的差错控制,必须采用受延迟约束的重传技术。在视频组播中,时间限制要求重传尽量控制在位置最近的组播成员中。此外,为适应大规模应用,重传请求所产生的反馈风暴也应得到有效控制。典型的重传机制通过将接收者组织成某种层次型逻辑树或本地组来限制重传请求的个数和范围。IP 组播对网络拓扑信息的缺乏使得系统不能为如何将接收者组织成修复树(组)提供显式途径。一些研究提出手工建立修复树和配置网络,这种方法的缺点是要求预先知道接收者集、网络结构和所有参与网络的管理策略。文[22]通过在树的每层静态配置 DRs (Designated Receivers) 来协助重传丢失报文,同时借助客户缓存提高按时接收重传报文的可能性。此外,该算法还通过抑制不能按时恢复的重传请求来提高带宽效率。与文[24]不同,STORM<sup>[25]</sup>可为每个接收者在任何时候根据可靠程度和延迟要求动态选择最佳 DR,提高了即时重传的可能性。

## 6 相关协议

### 6.1 组播协议

组通信是组播实现的基础,组播协议要能适应网络结构的变化,维护动态组关系。对 Internet 而言,组管理协议功能包括提供在 LAN 内跟踪组成员关系、准确传输组播信息的局部机制和 LAN 之间路由报文的全局机制。

RFC 1112<sup>[2]</sup>描述了组播对 IPv4 协议栈各层的影响,给出了组播模型及实现机制,并在 IPv4 中通过集成 IGMP (Internet Group Management Protocol) 提供 LAN 内的组管理功能。IGMP v. 1 中定义了组成员关系查询、报告两种消息的处理方法,IGMP v. 2 为降低离开延迟重新定义了消息格式,扩充了消息类型,但没有从根本上消除离开延迟问题。B. Cain 等人从安全性角度提出了 IGMP v. 3 草案,增加了源过滤 (source filtering) 功能,使得系统能接收指定源发送的数据包和排除特定的源地址,同时可限制报文发送到特定的组播地址。

组播路由协议 (DVMRP, MOSPF, CBT, PIM-DM, PIM-SM) 提供了广域环境下域内组播报文的转发机制,这些路由协议虽分别针对不同成员分布特点结合协议依赖性、路由生成方式、路由模型等几个因素提出了不同的路由算

法,但都是采用嫁接(grafting)和剪枝(pruning)技术提供有效组播路由。

现有组播路由协议很难跨越非组播路由器提供域间(Inter-Domain)路由功能。虽然 PIM-SM 的加入消息能到达非组播路由器,但在域间使用 PIM-SM 还存在以下问题:不同域的实体之间一般不存在信任关系,因此很难交换路由信息。针对上述问题的一个过渡方案是结合使用 MBGP、PIM-SM 和 MSDP。MBGP 允许为不同协议维护多个路由表,使得路由器能分别为单播路由器和组播路由器构造路由表,通过 MBGP, PIM-SM 可绕过单播路由器发送加入消息,MSDP 为每个域建立组共享树,不同域间的 RP 通过 MSDP 实现互连,解决了 ISP 间的独立性问题。但这种方案由于要求域中的每个 RP 都要维持所有发送源信息,因此伸缩性较差。

BGMP 的目标是提供一个解决域间组播的长期方案,草案的基本思想是使用一个根节点在域间建立双向共享树,并允许接收域根据需要建立特定源域间分枝, BGMP 的基本设想是域间的依赖性可以通过严格的地址分配方案来消除。

## 6.2 RTP/RTCP

RTP/RTCP 是 Internet 上实时传输的应用层协议,最初设计主要针对视频会议的数据传输,底层协议可以是 TCP 或 UDP, RTP 支持单播和组播, RTP 主要负责会话的建立、管理,提供对负载信息的标识、序列化。RTP 支持流媒体信息的分层传输,协议中 mixer 利用 SSRC 能实现对多个流的合并和同步控制, translator 保证了流媒体数据能穿越防火墙,提供可靠服务。RTP 是一个可扩充协议框架,用户可根据特定的应用需求,定义自己的数据类型。

RTCP (Real-time Transport Control Protocol) 是与 RTP 集成的传输控制协议, RTCP 负责对网络信息如丢包率、抖动延迟等统计,并将相关数据反馈给组成员和发送方。RTCP 的报告分为接收方报告 RR (Receiver Report) 和发送方报告 SR (sender Report), 服务器可根据反馈动态改变传输速率及载荷类型。为避免反馈报文对网络资源的影响,协议规定 RTCP 报文只能占用网络带宽的 5%。

同一个会话中的 RTP、RTCP 连接使用同一个 IP 地址和相邻的两个端口号 (RTP 的端口号为偶数, RTCP 的端口号为随后的奇数)。会话过程中, RTP 报文中的同步源 SSRC 可以改变,而 RTCP 中的规范名字 CNAME 是 RTP 源的永久性标志,是唯一且不可改变的,会话通过 CNAME 跟踪参加者。RTP/RTCP 保证了流媒体传输的实时性、有序性,负责监视传输服务质量,但不为服务质量提供保证,用户可根据反馈信息和协议机制,在应用层为流媒体的有效传输引入端到端的流量控制,根据当前网络状况,动态调节数据发送速率。

**展望** Internet 为应用研究提供了新的舞台,同时也带来了新的挑战。本文综述了 Internet 上流媒体组播实现的基本思想和协议设计,相关工作为流媒体传输提供了应用基础。然而, Internet 的自身特点和流媒体应用的特性使得流媒体在 Internet 上的大规模应用还存在一些有待解决的问题,有关流媒体组播的进一步研究可包括以下几个方面:

(1) 服务集成。Internet 是一个多业务网络,各种业务流有着不同的服务特性,流媒体组播协议机制要能与多业务流共存和共享网络资源,现有工作大多分别从端到端和网络层开展研究,进一步研究可结合 QoS 从服务集成的角度考虑有关工作的融合。

(2) 异构性与公平性问题。网络元素的异构性导致了用户

接收、处理能力的异构性,基于分层的组播策略只能为流媒体传输提供粗粒度控制,结合编码技术对会话成员公平性的形式化描述将有助于视频组播协议的设计和性能评价。同时,有关异构流的公平性研究还有待于验证。

(3) 安全问题。流媒体组播的安全性研究包括组播机制和信息传递两个方面。组播是一个开放环境,为防止恶意用户的侵入和参与,必须提供针对发送者和组成员的身份鉴别技术。与流媒体特别是视频信息版权相关的传输控制也是与流媒体传输相关的一个热点。安全性是流媒体组播得以大规模应用的前提。

(4) 价格机制。带宽是目前 Internet 上的宝贵资源,然而许多研究工作的基本假设都是用户希望尽可能多地获取资源,提高服务质量。进一步工作可考虑结合价格、服务质量来考虑用户行为,考察相关的资源分配策略和协议设计。

## 参考文献

- 1 Li X, Ammar M, Paul S. Video Multicast over the Internet. *IEEE Network Magazine*, April 1999
- 2 Deering S. Host Extensions for IP Multicasting. RFC 1112 (1989)
- 3 Jiang T, Ammar W, Zegura E M. Inter-Receiver Fairness: A Novel Performance Measure for Multicast ABR Sessions. *SIGMETRICS*, 1998
- 4 Sisalem D. Fairness of Adaptive Multimedia Applications. In: *Proc. IEEE SUPERCOMM*, 1998
- 5 Sisalem D, Wolisz A. LDA+ TCP-Friendly Adaptation: A Measurement and Comparison Study. In: *the 10<sup>th</sup> Intl. Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV' 2000)*, Chapel Hill, NC, USA, 2000. 25~28
- 6 Busse I, Deffner B, Schulzrinne H. Dynamic QoS Control of Multimedia Applications Based on RTP. *Computer Communications*, Jan. 1996
- 7 Bouras C, Gkamas A. A Mechanism for Multicast Multimedia Data with Adaptive QoS Characteristics. In: *Proc. PROMS 2001*
- 8 Cheung Y, Ammar M, Li X. On the Use of Destination Set Grouping to Improve Fairness in Multicast Video Distribution. In: *Proc. IEEE INFOCOM*, 1996
- 9 Jiang T, Zegura E W, Ammar M. Inter-receiver Fair Multicast Communication over the Internet. In: *Proc. NOSSDAV*, 1999
- 10 Shacham N. Multipoint Communication by Hierarchically Encoded Data. In: *Proc. IEEE INFOCOM*, 1992
- 11 McCanne S, Jacobson V, Vetterli M. Receiver-driven Layered Multicast. In: *Proc. ACM SIGCOMM*, 1996
- 12 Li X, Paul S, Ammar M. Layered Video Multicast with Retransmissions (LVMR): Evaluation of Hierarchical Rate Control. In: *Proc. IEEE INFOCOM*, 1998
- 13 Vicisano L, Crowcroft J, Rizzo L. TCP-like Congestion Control for Layered Multicast Data Transfer. In: *Proc. IEEE INFOCOM*, 1998
- 14 Whetten B, Conlan J. A Rate Based Congestion Control Scheme for Reliable Multicast. *Technical White Paper*, Nov. 1998
- 15 Widmer J, Handley M. Extending Equation-based Congestion Control to Multicast Applications. In: *Proc. ACM SIGCOMM 2001*, San Diego, CA, 2001. 275~286
- 16 Widmer J, Denda R, Mauve M. A Survey on TCP-Friendly Congestion Control (extended version): [Technical Report TR-01-002]. Department for Mathematics and Computer Science, University of Mannheim, 2001

- 17 Floyd S, Handley M, Padhye J. A Comparison of Equation-Based and AIMD Congestion Control. URL: <http://www.aciri.org/tfrc/>.
- 18 Bansal D, Balakrishna H. Binomial Congestion Control Algorithms. In: Proc. IEEE INFOCOM, April 2001
- 19 Golestani J, Sabnani K. Fundamental Observations on Multicast Congestion Control in the Internet. In: Proc. IEEE INFOCOM, 1999
- 20 Wu D, Hou Y T, Zhang Y. Transporting Real-Time Video over the Internet: Challenges and Approaches. In: Proc. of the IEEE. VOL. 88, NO. 12, Dec. 2000
- 21 Alanese A, Blomer J, Edmonds J, Luby M. Priority Encoding Transmission: [Tech. Rep. RT-94-039]. Department of Computer Science, University of California, Berkeley, 1994

- 22 Tan W, Zakhor A. Multicast Transmission of Scalable Video Using Receiver-driven Hierarchical FEC. In: Proc. Packet Video, New York, April 1999
- 23 Davis G, Danskin J. Joint Source and Channel Coding for Internet Image Transmission. In: Proc. SPIE Conf. Wavelet Applications of Digital Image Processing XIX, Denver, CO, Aug. 1996
- 24 Li X, Paul S, Pancha P, Armmar M. Layered Video Multicast with Retransmissions (LVMR): Evaluation of Error Recovery Schemes. In: Proc. IEEE INFOCOM, 1998
- 25 Xu X R, Myers A C, Zhang H, Yavatkar R. Resilient Multicast Support for Continuous-media Applications. In: Proc. IEEE Int. Workshop Network and Operating System Support for Digital Audio and Video (NOSSDAV'97), May 1997, 183~194

(上接第9页)

### 3.3 元数据组织与表示

P2P 网络面向的是异构网络与操作系统, 这样就需要在这些系统之间交换数据资源, 但是因为这些系统的数据表示并不都是完全相同的, 这样就需要一个能够在多个系统之间确定一个通用的元数据表示方案。关于元数据的组织包括数据资源的表示、消息通信协议等, 很多系统都支持 SOAP 或者 XML-RPC 等协议。

### 3.4 匿名性的支持

对匿名性进行支持是很多 P2P 系统所要实现的一个重要功能, 所谓的匿名性是指对系统中信息资源的操作者的操作行为的隐私进行保护, 具体包括对一个信息的作者、发布者、阅读者、信息提供服务器、存储位置、检索的用户行为的隐私性进行支持。在 Freenet、Free Haven 等系统都对这项功能进行了支持。

### 3.5 P2P 网络的支撑技术

Internet 技术的发展使得连入互连网络中的设备不再局限于计算机, 在 P2P 的计算环境中要求任何设备都可以在任何地点很容易地加入到这个环境中, 所谓的计算设备既包括有线设备也包括无线设备, 这样就需要很多很多网络传输的支撑技术来支持各种不同设备连入整个 P2P 网络。Bluetooth、Jini 都是目前所研究和采用的 P2P 网络支撑技术。

### 3.6 P2P 网络的安全问题

安全问题是一直伴随着互联网发展的重要问题, 安全问题包括很多相关的问题, 比如应该防止他人控制整个系统, 增加恶意信息等, 同时系统应能够保证系统中信息资源的正确性。在 P2P 系统中系统安全同样面临着巨大的挑战。P2P 系统需要在没有中心节点的情况下, 提供身份的认证、授权以及数据信息的安全存储、数字签名、加密、安全传输等工具, 同时 P2P 系统要有能力抵抗过量存储负载、Dos 攻击等攻击行为。

**总结** P2P 技术的兴起使得人们从信息资源共享、普及计算、协同工作、实时通信技术、信息检索技术、广域网络存储系统等领域来应用该技术, 同时为了能够充分地发挥其技术优势需要解决与该技术相关的一些典型的问题, 文中对这些应用领域以及这些典型的问题分别进行了讨论。

## 参考文献

- 1 Napster. <http://www.Napster.com/>
- 2 Clarke I, et al. Freenet: A Distributed Anonymous Information Storage and Retrieval System. ICSI Workshop on Design Issues in Anonymity and Unobservability, July 2000
- 3 Gnutella. <http://gnutella.wego.com/>
- 4 Free Haven. <http://www.freehaven.com/>
- 5 Ohaha. <http://www.ohaha.com/>
- 6 SETI@Home. <http://setiathome.ssl.berkeley.edu/>
- 7 Foster I, Kesselman C. The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann Publishers, Inc. 1999
- 8 <http://www.research.ibm.com/autonomic/>
- 9 Groove. <http://www.groove.net/>
- 10 Jabber. <http://jabber.org/>
- 11 WebGather. <http://e.pku.edu.cn/>
- 12 google. <http://www.google.com/>
- 13 JXTA Search: Distributed Search for Distributed Networks. <http://www.jxta.org/>
- 14 Kubiawicz J, et al. OceanStore: An architecture for global-scale persistent storage. In: Proc. of ACM ASPLOS. ACM, Nov. 2000
- 15 JXTA. <http://www.jxta.org/>
- 16 Ratnasamy S, Francis P, Handley M, Karp R, Schenker S. A scalable content-addressable network. Submission to ACM SIGCOMM, 2001
- 17 Druschel P, Rowstron A. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. Submission to ACM SIGCOMM, 2001
- 18 Stoica I, et al. Chord: A scalable peer-to-peer lookup service for internet applications. Submission to ACM SIGCOMM, 2001
- 19 Zhao B Y, Kubiawicz J, Joseph D. Tapestry: An infrastructure for fault-resilient wide-area location and routing. Submitted for publication, 2001
- 20 LI J, et al. A scalable location service for geographic ad hoc routing. In: Proc. of the 6th ACM Intl. Conf. on Mobile Computing and Networking (Boston, Massachusetts, August 2000), pp. 120~130
- 21 Bolosky W, Douceur J, Ely D, Theimer M. Feasibility of a serverless distributed file system deployed on an existing set of desktop pcs. In: Proc. of Sigmetrics, June 2000