

# 基于 IP 的视频点播系统中 QoS 管理的分析与设计<sup>\*</sup>

Analysis and Design of QoS Management for IP Based VOD Systems

荣波 孙景鳌 蔡安妮

(北京邮电大学 北京 100876)

**Abstract** VOD(video on demand) systems and other multimedia applications generally have high QoS requirements, but most of networks that now exist can't provide QoS guarantee. Due to this fact, distributed multimedia systems must have QoS management mechanism themselves. In this paper, we propose a QoS management scheme for VOD systems in IP environment. The architecture and implementation issues of the scheme are also discussed.

**Keywords** VOD systems, QoS management

在视频点播(VOD)系统中,多媒体应用一般对系统和网络都有较高的服务质量 QoS 要求,然而现有的各种系统和网络对 QoS 概念的支持还比较薄弱。多数网络虽然允许用户说明对服务提供方的 QoS 要求,但是通常并不保证一定满足用户的 QoS 要求。例如,在 IP 协议中,只允许 IP 用户通过 IP 分组头部中的服务类型域进行简单的定性 QoS 说明,不能进行严格的定量 QoS 说明,而且下层网络机制也不提供 QoS 保证。实际上,现有的多数网络系统(如 Internet)在设计时并没有考虑对定量 QoS 的支持。其次,多数现有网络中的 QoS 一般都是静态的,即通信双方一经商定 QoS 参数值,在通信过程中,就不能对 QoS 参数值进行动态“再协商”。例如,用户不能选择将己有的一条视频连接的质量从彩色降级到单色,以便打开一条新的音频连接。此外,服务提供方必须承诺在整个连接生命期内提供协商好的 QoS,如果不能继续维持协商一致的 QoS,服务提供方往往就单方面地关闭连接。因此,分布式多媒体系统及其应用要求系统自身具有强有力的 QoS 管理机制,以满足其 QoS 要求。在本文中,提出了一种适用于基于 IP 的视频点播系统的 QoS 管理方案,并讨论了具体实现方法。

## 1 视频点播系统中的 QoS 管理框架结构

为了保障多媒体业务的 QoS,曾经提出过几种 QoS 管理框架结构<sup>[1-3]</sup>,但这些框架结构大都是针对通用多媒体通信系统或视频会议系统的,没有针对视

频点播系统的具体特点。我们在 M. Fry 等人提出的一种用于视频会议系统的 QoS 管理框架结构<sup>[2]</sup>的基础上,经过适当地扩展和改进,得到了适用于视频点播系统的 QoS 管理框架结构,并将该结构应用于我们开发的基于 IP 的视频点播系统中。该 QoS 管理框架结构如图 1 所示,图中各组成部分介绍如下:

(1)多媒体应用:当多媒体应用进程开始运行时,它将申请用于多媒体通信的连接,该连接请求中包含有相应的 QoS 要求。例如,在视频点播系统中,点播者可以根据自己的意愿指定喜爱的服务级别,然后该服务级别将被翻译成相应的 QoS 要求,提交给 QoS 管理者。

(2)QoS 管理者:QoS 管理者接受多媒体应用进程提出的 QoS 请求,然后将该请求转换为一组 QoS 参数,根据这些参数与资源管理者进行协商,以获得相应资源。

(3)协议提供者:由于现有网络所提供的传输服务并不能完全满足多媒体应用的要求,所以需要由协议提供者来提供合适的、能够满足一定要求的通信协议栈。

(4)资源管理者:资源管理者接受 QoS 管理者提出的 QoS 要求,并为满足该 QoS 要求而预留资源。预留的资源包括主机的 CPU 时间片、内存、IO 资源等。网络资源也同样需要预留,虽然将来的网络会提供 QoS 保障,但是现存的实际使用中的大多数网络至今还没有提供预留网络资源的手段。如果知道了不同服

<sup>\*</sup>国家自然科学基金资助项目(69972009)。荣波 博士研究生,主要研究方向是多媒体信息网及其网络管理,孙景鳌 博士生导师,从事电视、多媒体通信和成像技术方面的教学和研究工作。蔡安妮 教授,博士生导师,从事多媒体通信、图像处理 and 识别方面的教学和研究工作。

务级别对应的不同网络带宽要求,那么在无法预留网络资源的情况下,可以根据当时的网络带宽情况适当降低服务级别,使在这个服务级别上的 QoS 要求能够得以保障,而且一旦将来网络通信情况转好,还可以恢复到原有的服务级别上来。

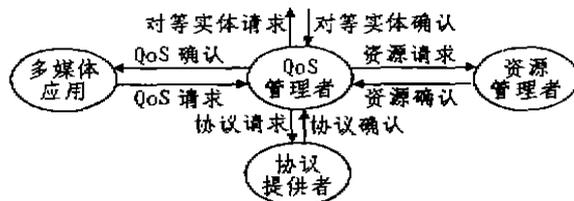


图1 用于视频点播系统的 QoS 管理框架结构

在视频点播系统中,无论是点播客户端还是视频服务器端都采用了上述 QoS 管理框架结构,除了本机资源的协商与预留之外,点播客户端与视频服务器端的 QoS 管理者之间也要进行协商,以最终确定能够提供的服务级别。另外,在整个连接的生命期内,QoS 管理者还应不断监视多媒体数据的传输质量,并采取适当措施来保证用户点播质量。

在基于 IP 的视频点播系统中,我们采用 RTP 协议作为多媒体数据传输的协议提供者,同时还采用该协议对网络中的数据传输质量进行监测,因此下面首先对 RTP 协议加以介绍。

## 2 RTP 协议

### 2.1 RTP 协议概述

RTP 协议<sup>[1]</sup>作为第一个为 ITU-T 所接受的 Internet 标准,已经成为了局域网环境下视听会议标准的一部分,并在媒体点播类应用,如视频点播(Video On Demand, VOD)中发挥着重要作用。RTP 协议包括实时传输协议(RTP)和实时传输控制协议(RTCP)两部分。

RTP 提供实时数据端对端的网络传输功能,其报文专门被设计用来承载连续媒体数据,但它不为实时业务保留资源,也不保证服务质量。RTCP 则只提供数据传输的基本控制和检测功能。RTP 协议独立于传输层和网络层,它没有具体规定底层网络所能提供的传输服务类型,通常情况下,它运行于 UDP 之上。在 RTP 协议中,RTP 通信服务的参与者之间的连接被称为 RTP 会话。对每一个参与者来说,会话由一对传输层地址(一个网络地址和两个端口号)组成,一个地址用作 RTP 报文通信,另一个用作 RTCP 报文通信。按协议的规定,RTP 将选择一个偶数位的端口号,而 RTCP 则选用高一位的端口号。在多媒体会话中,不同媒体类型的数据各自使用自己的 RTP 会话传输数据,

因而可以根据各个会话来分配网络资源。

在基于 IP 的视频点播系统中,我们采用 RTP 协议中的 RTCP 对网络的数据传输质量进行监测,因此,RTCP 在系统的 QoS 管理中占有较为重要的地位,下面加以详细介绍。

### 2.2 RTP 协议的控制功能—RTCP

RTCP 通过向会话内的所有参与者周期性地发送控制报文以实现其功能,其报文的传送机制与 RTP 数据报文一致,如两者的底层协议都采用 UDP。常用的 RTCP 报文有以下几类:①SR:源报告包,用于报告活动源的统计信息。②RR:接受者报告包,用于报告非活动源的统计信息。③SDS:源描述包,用于报告和站点相关的信息。④BYE:站点离开系统报告包。⑤APP:特殊应用包。借助于上述控制报文,RTCP 完成以下四项功能:

(1)数据传输质量的反馈。这项功能是 RTP 协议的重要组成部分,它大致上可以和其它传输协议的流量控制及阻塞控制功能相对应。这些反馈信息可以直接用作自适应编码的控制,也可以用来发现在数据分发时的错误。RTP 会话向所有的参与者发送 RTCP 报文的目的是为了那些发现在数据传输过程中出现问题的用户能够确定问题是仅和其自身有关还是和全体参与者有关。

(2)识别信息。对于每一个 RTP 源而言,RTCP 为其产生一个持久的传输层标识,称做 CNAME(Canonical NAME)。虽然在 RTP 报文中的 SSRC 标识符也可以区分不同的码流源,但是当 RTP 发现不同码流的 SSRC 标识符发生冲突或应用重新启动时,SSRC 值会被更改,为了接收者能够跟踪会话的每一个参与者就必须利用 CNAME。另外,借助 CNAME,接受者可以关联由同一个用户产生的、由不同 RTP 会话传输的相关码流(如音频码流和视频码流)的同步。CNAME 由 RTCP 中的源描述包携带。

(3)会议大小估计与控制报文发送速率的调整。通过接收来自于其它参与者的 RTCP 报文,一个 RTP 会话的用户能够独自了解会话参与者的数目。这个数目被用来计算控制报文的发送间隔,以达到 RTCP 所占用的通信带宽不超过会话带宽 5%的目的。

(4)基本的会话控制功能。这个功能是可选的,它主要应用于控制较为松散的会话,这类会话对用户的加入和离开没有严格的管理。

以上介绍的四项功能中,前三项是在多点通信环境下应用 RTP 协议所必需的,而在视频点播系统中,点播客户和视频服务器之间只是点对点的连接,因而针对每一个 RTP 会话主要使用了 RTCP 的第一项功能,其目的在于监测媒体数据的传输质量,并不断向

QoS 管理者进行汇报。

### 3 基于 IP 的视频点播系统中 QoS 的管理与保障

#### 3.1 视频点播系统中 QoS 管理的总体设计

在我们开发的基于 IP 的视频点播系统中,根据图 1 的 QoS 管理框架结构,引入了 QoS 管理机制。图 2 显示了该系统中 QoS 管理部分的系统结构框图,其中,点播客户端与视频服务器端分别拥有各自的 QoS 管理者;客户端高层、服务器端高层与各自的本地操作系统相配合,分别对点播客户端和视频服务器端的资源进行管理与控制;RTP 协议则在客户端和服务端充当协议提供者的角色。下面对系统实现中的几个关键问题加以介绍。

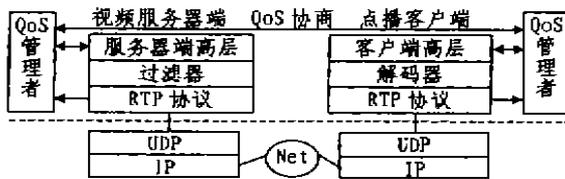


图 2 基于 IP 的视频点播系统中 QoS 管理部分的系统结构框图

#### 3.2 利用 RTCP 实现数据传输质量的监测

如前所述,在基于 IP 的视频点播系统中采用 RTP 协议做为多媒体数据传输的协议提供者,同时还采用该协议中的 RTCP 对数据传输质量进行监测。视频点播系统中的 RTP 会话是比较简单的客户机/服务器型(点对点)通信,因此,在视频服务器与点播客户之间只存在由服务器到客户机的 RTP 报文传输和双向的 RTCP 报文传输,如图 3 所示。在本系统中,点播客户端根据自己接收数据的实际情况利用“QoS 包”不断向视频服务器端汇报与通信质量相关的信息,QoS 包被设计为 8 字节长,属于 RTCP 的 APP 类,格式如图 4 所示。其中, Packets\_received 为自上次发送 QoS 包以来所接收到的数据包数; Packets\_expected 为自上次发送 QoS 包以来应该接收到的数据包数。视频服务器端根据 QoS 包中的信息就可以计算出数据包丢失率等参数。一般认为有两个原因引起数据包丢失,一是缓冲区溢出,二是比特错误。在大多数网络中,比特错误很少发生,因此可以认为数据包丢失主要是由网络

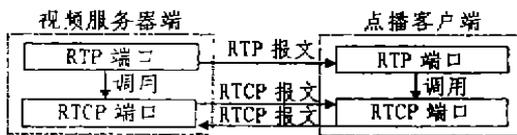


图 3 客户机/服务器结构下 RTP 会话的传输模式

拥塞带来的缓冲区溢出引起的,基于上述分析,可以使用数据包丢失率来作为网络拥塞程度的指示。

0			32		
F	QoS	Length=2	MBZ		
Packets_received			Packets_expected		

图 4 QoS 包格式

系统中网络拥塞程度的判决方法参考了 Ingo Buss 等人提出的方案<sup>[7]</sup>,如图 5 所示。视频服务器端根据点播客户端传来的反馈信息可以计算出数据包的丢失率  $\lambda$ ,为了防止 QoS 抖动,使用下述低通滤波器对丢失率  $\lambda$  进行平滑处理:

$$\lambda \leftarrow (1-a)\lambda + ab$$

其中,  $b$  为最新的丢失率值,  $0 \leq a \leq 1$ 。增加  $a$  会加强新值对最终计算结果的影响,减小  $a$  会使新计算值更接近于原有值。平滑后的丢失率  $\lambda$  再经过网络拥塞程度判决器就可得到最终的网络拥塞程度。网络拥塞程度判决器使用两个门限  $\lambda_c$ 、 $\lambda_u$  来决定最终的网络拥塞程度为网络拥塞、网络负载程度适中或网络空闲。门限  $\lambda$  的大小要适中,以使数据包丢失率限制在正常通信所能容忍的范围内, $\lambda_c$  的取值要小一些,以防止 QoS 抖动。在 100M 局域网环境下,取  $\lambda_c = 4\%$ 、 $\lambda_u = 2\%$ ,效果较好。

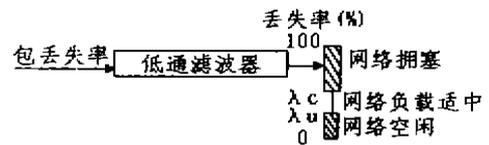


图 5 视频点播系统中网络拥塞程度的判决方案

#### 3.3 点播服务级别的协商与改变

每当用户点播视频节目时,点播客户端将向视频服务器申请建立相应的 RTP 会话,服务器端的 QoS 管理者将与点播客户端的 QoS 管理者进行协商,以根据当时的具体情况,确定最终能够提供的服务级别。在视频数据传输过程中,QoS 管理者通过 RTCP 不断监视多媒体数据的传输质量,当网络拥塞程度判决器判决当前网络状态为网络拥塞时,QoS 管理者就会认为网络当前不能满足用户所提出的 QoS 要求,两端的 QoS 管理者将重新协商,并采取相应措施,如降低服务级别等,以适应网络变化。一旦将来网络状态变为空闲,还可以恢复点播客户原有的服务级别。

当决定改变用户的服务级别时,视频服务器端的 QoS 管理者将通知服务器端高层,服务器端高层利用过滤器来控制由视频服务器发往点播客户端的视频数

(下转第 37 页)

表1 RED、CBT、D-CBT类的  
流平均吞吐率和公平性比较

	秒	TCP-Reno	标记UDP (MPEG-1流)	未标记UDP (CBR流)	公平性
RED	0~10	537	851	0	0.951
	10~20	293	720	4681	0.480
	20~30	161	601	4503	0.447
CBT	0~10	587	760	0	0.984
	10~20	615	739	624	0.993
	20~30	305	657	318	0.829
D-CBT	0~10	622	653	0	0.999
	10~20	659	660	637	0.999
	20~30	356	378	313	0.943

比较第6列的公平性值,我们可看到D-CBT能获得更好的公平性。

**结束语** 本文比较分析了RED、FRED、CBT、D-CBT四种不同的主动队列机制,重点分析了D-CBT, D-CBT将所有的业务分成三类:TCP、有流控的UDP流、没有拥塞控制的UDP流,对它们应用不同的队列策略,从而可限制每个UDP类最多使用它们应该得到的平均队列缓冲区空间。比较而言,D-CBT有如下优点:①D-CBT具有RED的拥塞避免功能;②D-CBT在增加较小负荷的情况下,能取得类似FRED的公平

特性;③D-CBT对每类进行惩罚,不是使用FRED执行严格的每个流的惩罚方法,因此D-CBT允许有流控机制的多媒体流能更公平地使用其输出带宽;④D-CBT的阈值动态更新能使所有流更公平地共享带宽。

我们下一步将改进D-CBT,研究它对区分服务的支持,并且结合端-端拥塞控制机制研究其公平性。

### 参考文献

- 1 Floyd S, Jacobson V. Link-sharing and Resource Management Models for Packet Networks. IEEE/ACM Trans. on Networking, 1995, 3(4): 7~9
- 2 Floyd S, Jacobson V. Random Early Detection Gateways for Congestion Avoidance. IEEE/ACM Trans on Networking, 1993
- 3 Lin D, Morris R. Dynamics of Random Early Detection. In: Proc. of SIFCOMM'97, Cannes, France, 1997
- 4 Parris M, Jeffay K, Smith F D. Lightweight Active Router-Queue Management for Multimedia Networking. Multimedia Computing and Networking, SPIE Proc. Series, Vol 3020, San Jose, CA, January 1999
- 5 Chung J, Claypool M. Dynamic CBT—Better Performing Active Queue Management for Multimedia Networking [Tech Report WPI-CS-TR-00-10]. Worcester Polytechnic Institute, April 2000

(上接第44页)

据格式,从而达到改变服务级别的目的。例如,对于采用分层编码的视频流而言,当QoS管理者决定降低服务级别时,可以命令过滤器过滤掉与图像细节相关的高层编码数据,只传送视频图像的基本信息,这样由视频服务器端发往点播客户端的数据流量就会减少,整个网络的拥塞程度将会有所改善。

由于在视频点播系统中采用了上述QoS管理方案,使得该系统能够较好地适应各种网络环境,并且在网络通信条件不断变化的情况下,仍然能够提供必要的QoS控制与保障。

**结束语** 由于目前大多数的网络系统(如Internet)尚不能对QoS要求提供充分的支持,因此采取适当的QoS管理和控制手段以保证点播质量对现有的视频点播系统是非常必要的,本文介绍的QoS管理方案在这方面具有较高的实用价值。随着通信与计算机网络技术的不断发展,未来的Internet网将对QoS提供良好的支持,在这种情况下只需引入QoS管理者与Internet网络设备之间的QoS申请、协商机制,本文介

绍的方案同样会取得良好的效果。

### 参考文献

- 1 Campbell A, et al. Integrated Quality of Service for Multimedia Communications. Proc. IEEE INFOCOM'93, San Francisco, USA, 1993, 3: 732~739
- 2 Fry M, Seneviratne A. A QoS Managed Video Conferencing Application. [Technical report]. CRC for Distributed Systems Technology, 1994
- 3 Neumar B. Modeling resources for integrated performance management. In: Proc of the Third Int Symposium on Integrated Network Management. San Francisco CA, April 1993. 109~121
- 4 Pacific G, Stadler R. An architecture for performance management of multimedia networks In: Proc. of the Fourth Int Symposium on Intergrated Network Management. Santa Barbara CA, May 1995. 174~186
- 5 Lazar A A. An architecture for real-time control of broadband networks. In: Proc of the IEEE Global Telecommunications Conf. Phoenix AZ, December 1991. 289~295
- 6 Schulzrinne H, et al. RTP: A Transport Protocol for Real-time Applications. IETF, Nov. 1995
- 7 Buss I, et al. Dynamic QoS control of multimedia applications based on RTP. Computer Communications, 1996, 19: 49~58