

基于消费者-提供者协作管理模型的 QoS 管理^{*}

QoS Management Based on Resource Consumer-Provider Cooperative Model

蔡皖东

陈亚清

(西北工业大学 西安710072)(西安交通大学计算机信息技术研究所 710049)

Abstract The rapid advances in distributed multimedia systems pose higher requirements to the QoS (Quality of Service) management mechanisms. This paper presents a management method of network bandwidth based on resource consumer-provider cooperative model. On an endsystem, it supports QoS negotiation by defining an abstract interface, estimates the QoS of current network by the use of control mechanism of the RTP (Realtime Transport Protocol) and manages the network bandwidth adaptively by an algorithm of bandwidth adjustment. The management method can ensure steady the QoS of end-to-end.

Keywords Distributed multimedia systems, QoS, QoS management mechanisms, RTP

1 引言

在分布式多媒体系统中,常用服务质量(Quality of Service, QoS)来说明网络服务的“良好”程度。由于不同的应用对网络的性能要求不同,对网络系统所提供的服务质量的期望值也不同。这种期望值可以用一种统一的 QoS 概念来描述。在不同应用系统中, QoS 参数集的定义方法可能是不同的。例如,可以用一些相关的网络性能参数,如吞吐量,差错率,端到端延迟,延迟抖动等来定义网络系统的 QoS。为了在端系统上提供更加灵活和动态的 QoS 选择,必须通过一种 QoS 管理机制对 QoS 进行有效的管理,并且这种 QoS 管理机制应当支持 QoS 的可配置性、可协商性、可伸缩性以及自适应性等管理特性。

近年来,国外对分布式多媒体系统的 QoS 管理机制进行了深入研究,提出了一些管理模型和原型系统。文[1]提出一种资源管理协作模型对端系统上的 CPU 和网络 I/O 资源进行协作和自适应管理,但未涉及网络带宽管理问题。文[2]提出一种基于服务合同的使用者-提供者管理模型对网络带宽资源进行动态分配和自适应管理,但需要 ATM 网络的强力支持。文[3]提出一种协作多媒体环境通过端点间的相互协作实现对 QoS 进行控制和管理,但在端点间必须建立一个用于协商 QoS 的会话连接。

本文提出的基于消费者-提供者协作管理模型的

QoS 管理方案是根据以上研究的最新成果,从集成 QoS 管理的角度,对端系统上的 CPU、缓冲区、I/O 带宽及网络带宽等可用资源进行综合管理。这种集成化 QoS 管理环境由用户界面、QoS 管理器和这种资源控制器组成。用户界面为应用定义和协商 QoS 提供了可视化的接口, QoS 管理器根据应用的 QoS 需求对各个资源采取相应的管理策略,各种可用资源由各自的资源控制器进行具体的控制和管理。由于系统采用层次化和模块化设计,各种资源控制器相互独立,且通过标准接口与 QoS 管理器绑定,连接到集成 QoS 管理环境中。本文主要论述对分布式多媒体环境中的 QoS 产生重要影响的网络带宽管理问题,而其它资源的集成管理将另文讨论。

2 消费者-提供者协作管理模型

在分布式多媒体环境中,端系统上的可用资源有 CPU、缓冲区、I/O 带宽及网络带宽等,这些资源应当通过一种集成化的 QoS 管理机制实行统一的管理,提供端到端的、且具有协作性、动态性和自适应性的 QoS 管理功能。这些 QoS 管理特性应通过适当的管理模型或框架来描述和实现。因此,我们采用了一种消费者-提供者协作管理模型来实现系统可用资源的管理。这种管理模型基本原理是,一个应用(资源消费者)在初始启动时,首先以适当的方式提出 QoS 请求,系统(资源提供者)根据其可用资源容量计算和分配该应用所

^{*} 本文工作得到航空高校科研基金的资助。

需的资源。在该应用运行时,提供者和消费者将协作管理和动态监测该应用的资源需求和实际的 QoS。当系统负载发生变化而导致 QoS 改变时,提供者和消费者需要重新协商,使之在可用资源约束内自适应于该应用的 QoS 需求。这种消费者-提供者协作管理模型不

仅可提供上述的管理特性,而且还隐藏了 QoS 测量和协商细节,应用只需按易于理解的方式定义其 QoS 参数表示。图1显示了一种基于该管理模型的集成化 QoS 管理体系结构。

在端系统的可用资源中,网络带宽对分布式多媒

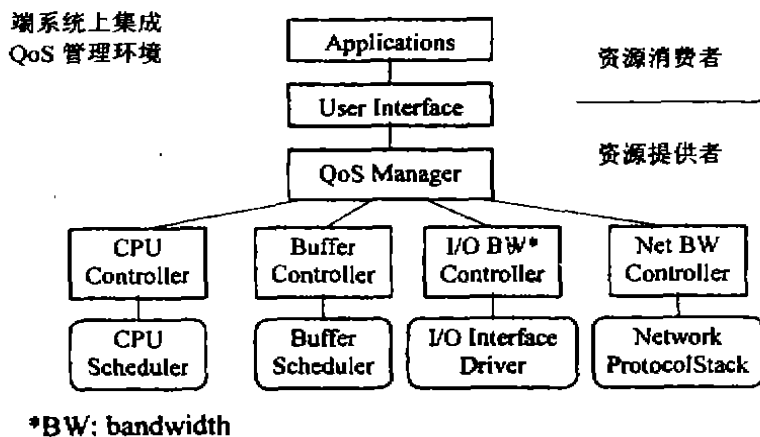


图1 端系统集成 QoS 管理体系结构

体系统的性能和 QoS 产生更直接的影响,必须进行有效管理。图2给出一种端系统上网络带宽管理结构,它是基于上述管理模型,并且利用 RTP 协议控制机制进行网络 QoS 动态检测,从而实现端系统上的 QoS 协商和网络带宽自适应管理的。

表1 QoS 参数例

Name	Type	Value
CurrentFrameRate	IntegerValue	30
FrameRateRange	IntegerRange	[1...30]
Cost	IntegerRange	[100(150)200]
Priority	OrderedSet	Low, medium, high

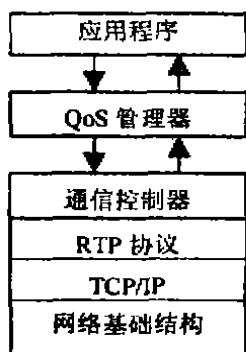


图2 端系统网络带宽管理结构

3 QoS 管理器

QoS 管理器是实现端系统 QoS 管理的核心部件,主要提供 QoS 的协商、映射及动态监控等管理功能。

(1)QoS 的协商。QoS 管理器通过定义一种易于

理解的 QoS 参数表示法与应用程序接口。在 QoS 参数集中,每个参数用一个三元组 {Name, Type, Value} 来说明,见表1。应用程序使用这些 QoS 参数向 QoS 管理器请求所需的 QoS 级。QoS 管理器根据当前网络资源状况来测算是否满足要求,以及所能达到的 QoS 级,并向应用程序给出相应的指示。

(2)QoS 的映射。QoS 管理器将应用程序的 QoS 请求映射成网络 QoS 级。每个网络 QoS 级可用下列参数描述:网络吞吐量、延迟、延迟抖动以及数据丢失率等。通过这种映射隐藏了与网络资源提供者进行协商的细节,使应用程序能以一种易于理解的方式进行 QoS 参数的定义和协商。为了支持这种映射机制,QoS 管理器必须能够测算出当前网络资源使用状况,以验证网络当前 QoS 是否能满足用户所请求的 QoS。如果不能满足,则 QoS 管理器需要与应用程序进行重新协商,使之在可用网络资源约束内确保其可预测的服务。QoS 管理器对网络状态的测算通过调用通信控制器的初始化功能实现的,该功能调用将返回所测量的网络当前 QoS 参数。

(3)QoS 的动态监控。在数据传输过程中,网络负载和可用资源是不断变化的,这将引起 QoS 的波动。为了适应网络这种动态特性,QoS 管理器将周期地轮询通信控制器来检测当前网络 QoS 状态,并根据该状态对网络性能参数进行自动调节,以保证应用所需的 QoS。例如,可通过改变发送端的数据速率(或称带宽)来调节对数据丢失率的要求。我们在 QoS 管理器上设计了一种网络带宽的自动调节算法。在本算法中,使用

分组丢失率 λ 作为网络状态指示器,这里设置了两个丢失率阈值 λ_0 和 λ_1 ,并以此来判定网络状态, $0 \leq \lambda \leq \lambda_0$,则表示网络处于轻载状态;若 $\lambda_0 < \lambda \leq \lambda_1$,则表示网络处于载荷状态;若 $\lambda > \lambda_1$,则表示网络处于过载状态,根据经验, λ_0 和 λ_1 取值为 2% 和 4%。并且定义了三种带宽控制操作:增加、保持和减少,其带宽自动调节算法如下:

①如果网络处于过载状态,则执行减少操作,即 $b_s \leftarrow \max\{b_s * \mu, b_{\min}\}$;

②如果网络处于轻载状态,则执行增加操作,即 $b_s \leftarrow \min\{b_s + \nu, b_{\max}\}$;

③如果网络处于载荷状态,则执行保持操作,即保持 b_s 不变。

其中, b_s 为检测到的网络带宽, b_s 为所允许的带宽, $b_{\min}$ 和 $b_{\max}$ 分别为用户可接受的最小带宽和最大带宽,且 $b_s, b_s \in [b_{\min}, b_{\max}]$, μ 和 ν 分别为减算子和加算子。

4 通信控制器

通信控制器是在 RTP 协议基础上构造的,因此,首先简要地介绍 RTP 协议。

4.1 RTP 协议

RTP(Realtime Transport Protocol)是由 IETF 开发的一种实时传输协议,它是一种端-端的应用层协议,可以在面向连接或无连接的低层协议上工作。通常和 UDP 协议一起使用,RTP 包括数据和控制两部分。RTP 数据分组用于传送媒体数据(如音频和视频),它由分组头和分组数据组成,分组头中包含了媒体类型、时间标签、序列号、同步源标识等字段,这些字段为接收端实施媒体流内和流间同步控制提供了必要的支持。

RTP 控制部分称为实时传输控制协议(RTCP),主要用于支持协议控制功能,RTCP 分组是一种轻载分组,每个 RTP 实现都要周期地将 RTCP 分组广播给参加 RTP 会话组中的各个成员(一对多点广播)。发送端可以使用 RTCP 分组向各个接收端报告所发送的分组与字节累加计数值以及媒体间同步相关信息,接收端可以用这些计数器来估算实际的数据速率。接收端可以使用 RTCP 分组向发送端报告当前所接收的最高序列号、丢失的分组数、延迟抖动以及发送者和接收者报告往返的时间标签。

可见,RTP 提供了多媒体流内和流间的同步机制和网络 QoS 状况的检测机制,端系统可以借助于这些机制实施同步控制和 QoS 管理。

4.2 通信控制器的功能

通信控制器通过一种标准的接口与 QoS 管理器交互。在该接口中定义了三种操作:

①初始化。QoS 管理器向通信控制器发出该命令,通信控制器执行初始化过程,并返回当前网络 QoS 状况。

②QoS 检测。QoS 管理器向通信控制器发出该命令,通信控制器通过 RTCP 检测当前网络 QoS 状况,并返回给 QoS 管理器。

③带宽控制。QoS 管理器向通信控制器发出该命令,通信控制器按该命令给出的带宽参数增加、减少或保持数据发送速率,对传送带宽实施控制。

通信控制器利用 RTP 协议控制机制来测量当前网络 QoS 状况,并使用分组丢失率作为其指示器。为了避免 QoS 波动,通信控制器采用一种低通滤波器来平滑分组丢失率。在点-多点通信的场合下,发送端将会接收到多个接收端发来的接收者报告。这时,通信控制器将会根据一种多数表决策算法来确定当前网络 QoS 状况,并报告给 QoS 管理器。

结论 我们在 LINUX 上实现了基于上述 QoS 管理模型的原型系统。实验表明,通过对 QoS 的调节和控制,可使端到端的多媒体演示 QoS 稳定在一定的水平级上,同时增强了端用户对 QoS 的调控能力。由于 RTP 协议只是为端点的 QoS 检测和媒体同步提供一种支持机制,而协议本身并不提供网络带宽保证和调节能力,为了使端系统上的 QoS 管理器在出现网络带宽资源严重不足时也能在一个较大的范围内对带宽资源进行分配和调节,一方面需要网络基础结构(如 ATM 网络)提供必要的支持,人们对此做了大量的研究和实验^[2];另一方面可采用资源保留协议(RSVP)提供端到端的实时传输服务。RTP 和 RSVP 各有长处和适用的场合,我们今后工作的目标是在 QoS 管理器中实现对多协议的支持功能。

参考文献

- 1 Lakshman K, et al. Integrated CPU and network-I/O QoS management in an endsystem. Computer Communication, 1998, 21: 325~333
- 2 Ott M, et al. An architecture for adaptive and its application to multimedia systems design. Computer Communication, 1998, 21: 334~349
- 3 Alfano M. Design and implementation of a cooperative multimedia environment with QoS control. Computer Communication, 1998, 21: 350~361
- 4 Busse I, et al. Dynamic QoS of Multimedia Applications based on RTP. Computer Communication, 1996, 19: 49~58