

IP 网络中支持 QoS 的接纳控制算法^{*})

黄新刚 向 勇 史美林

(清华大学计算机系网络研究所 北京 100084)

QoS-Supported Admission Control Algorithms in IP Network

HUANG Xin-Gang XIANG Yong SHI Mei-Lin

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

Abstract In the research of NGN, Network QoS control is one of currently research hotspots. Admission control is a significant method used to support QoS. This paper surveys the research of QoS-supported admission control algorithms in IP network. We categorize and compare some main admission control mechanisms in IP network, try to analyze the relationship among these mechanisms and propose our views on the future relevant research.

Keywords IP network, QoS, Admission control

1. 引言

传统的计算机网络,特别是大型的广域网(如因特网)只能提供无传输性能保证的“尽力而为”(Best effort)服务。近年来,网络应用的蓬勃发展对 IP 网络提出了新的传输要求。大量的语音和视频流等实时数据在 IP 网络上传输时,对带宽、延迟、丢失率等指标有较高的要求,传统“尽力而为”的传输模式不能满足这些要求,从而促使人们改进对传统 IP 网络的设计,使之能够提供一定的服务质量(QoS)保证。

为了在 IP 网络上提供传输服务质量的保证,人们在资源预留(Resource Reservation)、流量整形(Traffic Shaping)、分组调度(Packet Scheduling)、队列管理(Queue Management)等方面展开了广泛的研究。实际上,在网络带宽一定的情况下,为提供服务质量的保证,网络系统可承载的分组流的数量是有限的^[1]。因此,对于大规模的 IP 网络而言,要提供良好的服务质量,除了要考虑上述几个方面,还必须具备执行接纳控制(Admission Control)的能力。接纳控制使得网络能够根据当前的负载状态决定是否接纳新的分组流,通过控制同一时刻网络服务的流的数量来保证服务质量。如果网络不能控制同时服务的流的数量,就可能导致这些流的服务质量要求和超过网络能够承载的上限,从而影响这些流所获得的服务质量。因此,在当前通过 IP 网络传输实时分组流的要求越来越多的情况下,为传统 IP 网络引入接纳控制机制是非常必要的。

20 世纪 80 年代中后期以来,针对每个实时流提供服务质量保证的研究广泛开展。在 IETF 提出的集成服务(IntServ, Integrated Services)网络体系结构中^[2],接纳控制机制是提供确保型(guarantee)服务质量保证必不可少的组成部分。在这个体系结构中,一个分组流在进入网络之前首先要通过 RSVP^[3](资源预留协议, resource ReSerVation Protocol)这种信令机制向网络发出服务请求报文,然后由通向接收方的路径上的路由器根据该请求报文执行接纳控制。只有当该报文所经过的路径上的所有路由器都能满足该流对服务质量的要求时,这个流才能进入网络。虽然 IntServ/RSVP 能够提供最高级别的服务质量保证,但是由于它要求路由器维护每流的资源预留信息以及每流的状态信息,因此

IntServ/RSVP 的可扩展性比较差。

为了解决可扩展性的问题, IETF 提出了另一种服务体系——区分服务(DiffServ, Differentiated Services)^[4]。区分服务没有对每流执行接纳控制的机制,路由器也不需要维护每流信息,只要求路由器能够根据分组头部携带的 DS 域信息执行类似于优先级方式的调度和缓冲机制。区分服务体系中由于网络中继节点不维护每流状态信息,因而具有较好的可扩展性。但是区分服务体系可能因进入网络的流数量太多而使网络利用率降低。因此,区分服务并不能提供严格的服务质量保证。

在过去的十年中,人们在接纳控制领域做了大量的研究工作。研究主要目标是,找到既能提供较为精确的服务质量保证,又能保持较高的网络利用率,而且可扩展性也比较好的接纳控制算法。例如,为了提高基于 IntServ/RSVP 的接纳控制算法的网络利用率,文[5]提出了基于测量的接纳控制算法。为了满足网络管理员和运营商的管理需求,文[6]提出了基于策略的接纳控制机制。相对于 IntServ/RSVP 分布式的接纳控制机制,文[7]提出了基于带宽代理(BB, Bandwidth Broker)的接纳控制机制,它采用集中式的工作方式,在网络中使用一个节点来执行接纳控制的功能。文[8]提出了在网络出口(egress)路由器对网络负载进行主动检测的方式。为了避免使用复杂的信令机制和特殊的分组处理过程,文[9]提出了一种基于探测(probing)的端点接纳控制方式(Endpoint Admission Control),在一个流进入网络之前,发送方首先发送一个探测流到网络中,以测定网络的拥塞程度,并据此做出是否接纳该流的决定。为了避免探测方式造成流设置延迟的缺陷,文[10]提出了基于网络的端点接纳控制算法。

本文将对 IP 网络中支持 QoS 的接纳控制算法进行特征分析,比较它们的异同和相互关系。本文对几种主要的 IP 网络中支持 QoS 的接纳控制机制进行分析;对 IP 网络中支持 QoS 的接纳控制算法研究的现状和趋势进行分析;最后是本文的总结。

2. 几种主要的接纳控制算法

一般而言,接纳控制的工作方式如下:首先,在一个流进入网络之前,发送方向网络发出一个要求资源预留的请求报

^{*})本文受国家自然科学基金项目(批准号:60273010)资助。

文,该报文携带了该流对服务质量参数的具体要求,然后由通向接收方的路径上的路由器根据该报文携带的参数和当时的网络情况执行接纳控制,决定是否接纳该流。最后,将判断结果通知发送方。广义的接纳标准为,如果接纳这个新的流会对网络中已经接纳的流的服务质量造成损害,那么该流不被接纳;反之,该流可以接纳。

资源预留协议是传递 QoS 请求的信令机制,可协调数据收发和接纳控制。最常用的资源预留协议是 RSVP^[3],用于 IntServ 体系结构中传递接纳请求,在沿路的所有路由器上预留资源,从而保证分组流的服务质量。

RSVP 的工作过程包括以下 4 个步骤^[3]:(1)在发送方正式开始发送之前,首先发送一个 PATH 报文给接收方。该报文携带了该流的流量特性(traffic characteristics)和应用的服务要求(service requirements of application)等信息。在该报文经过每一个路由器时,路由器记录了该报文刚刚经过的上一路由器的地址,然后将该报文转发给下一个路由器;(2)当接收方收到此 PATH 报文后,根据该报文产生一个 RESV (reservation request)报文,RESV 报文中包含了接收方要求的服务质量参数信息,并沿 PATH 的到达路径返回给发送方;(3)当该路径上的每个支持 RSVP 的路由器接收到这个 RESV 报文时,它使用接纳控制(Admission Control)进程来鉴别这个请求。如果这个请求能够满足(该请求通过鉴别,并有足够的资源分配给它),该请求报文就被传递到靠近发送方的下一个路由器;如果这个请求未能满足,路由器就会返回一个错误信息给接收方,此次资源预留的请求失败;(4)当该路径上的最后一个路由器(最靠近发送方的路由器)接收到该 RESV 报文并满足该报文的请求后,它就发回一个确认信息给接收方。这样,一条满足实际数据传输 QoS 要求的链路就在发送方和接收方之间建立起来了。

IntServ/RSVP 能够提供确保型的 QoS 保证,但是它需要在沿路的每个路由器中维护每个流的状态信息并为每个流预留带宽资源,增大了路由器的工作负载。由于需要维护的流的数目太多,在广域网核心路由器上应用 IntServ/RSVP 几乎是不可能的。因此,RSVP 最大的缺陷就是复杂性过高,可扩展性比较差。为此人们也试图使用流聚(Aggregation)和层次化(Hierarchy)的方法来提高 IntServ/RSVP 的可扩展性^[11~13]。

需要说明的是,RSVP 并不是一种接纳控制算法,但作为一种传递 QoS 请求的信令机制,它是传统的接纳控制机制不可缺少的一个组成部分。下文的基于参数、基于测量、基于带宽代理和基于策略的接纳控制算法都需要在具备 RSVP 能力的网络中应用。

2.1 基于参数的接纳控制算法

此类算法常常基于预知或简化的流量模型。例如把所有流的到达都视为泊松过程并假设它们服从独立同分布。据此获得描述每个流的流量特性的参数,然后根据这些参数估计所需资源(如最坏情况下所需带宽的峰值)的总和。当新的流发出接纳请求时,就利用这些估值加上新的流的资源请求。如果所有流请求的资源之和不超过链路的能力,那么该流可以被接纳,反之该流不被接纳。

这类算法的典型代表是文[14]中提到的 Simple Sum 算法。其接纳条件的数学表达式为:

$$\nu + \gamma^{\alpha} < \mu$$

其中 ν 表示已经保留的带宽之和, μ 表示链路的带宽, α 表示请求接纳的流, γ^{α} 表示流请求的带宽。作为最简单的接纳控制算法,它被当前各路由器和交换机厂商普遍应用。

基于参数的接纳控制算法的主要缺点在于很难构造流量模型和网络利用率较低。首先,在 IP 网络中,大部分应用流都具有长程相关性(long-range dependence),并不符合泊松分布;而且,由于 IP 网络中各个流到达的随机性,很难描述多种流汇聚后的流量特征,因此难以构造汇聚后的流量模型。其次,由于该算法为每个流按照其所需的峰值带宽预留资源,当流的突发性(bursty)比较大时,此类算法的网络利用率比较低。

2.2 基于测量的接纳控制算法

为了解决基于参数的接纳控制算法的网络利用率问题,人们考虑使用实时测量网络负载的方法来获得网络的剩余能力,并根据剩余能力是否能够满足新流的要求来决定是否接纳新流。这就是基于测量的接纳控制算法(MBAC, Measurement-Based Admission Control)工作方式。目前已有大量文献^[5]提出了各种基于测量的接纳控制算法,这些算法无须预知流量模型,而是以网络负载的实时测量作为接纳控制的依据。

所有基于测量的接纳控制算法都包括两个部分,即测量当前网络负载情况的测量机制(measurement procedure)和接纳决策算法。下面我们这两方面来分析几种典型的 MBAC 算法。

(1) 接纳决策算法

测量和(Measured Sum)^[14]。经过测量得到已接纳流当前占用网络带宽的估值,如果这个估值加上新流的速率不超过链路带宽的倍值(比如 0.9 倍),那么该流可以被接纳,反之,该流不被接纳。其接纳决策条件的数学表达式为:

$$\hat{\nu} + \gamma^{\alpha} < \nu \mu$$

其中 ν 是用户指定的目标利用率, $\hat{\nu}$ 是测量得到的已接纳流的负载。引入 ν 的目的是控制延迟^[15]。对于一个简单的 M/M/1 队列,当系统达到完全的利用率时,延迟的变化会趋于无穷大(diverge)。这种情况下,基于测量的方法就会失效。因此需要找到一个目标利用率,并保证接纳控制算法能够达到的利用率不会超过这个值。

目标利用率的确定与通过链路的流量的特性有关。如果每个流的速率相比链路带宽比较小,而且突发时间不长,那么这个目标利用率可以设得高一些;反之应该设得低一些。

等价带宽(Equivalent Bandwidth)。文[16]提出使用等价带宽法测量已接纳流占用带宽。当一个流 α 发出接纳请求时,此算法根据该请求报文获得该流的峰值速率 p^{α} 。当下列不等式成立时,接纳该流。

$$\hat{C}_H + p^{\alpha} \leq \mu$$

其中: $\hat{C}_H(\hat{\nu}, \{p_i\}, \epsilon) = \hat{\nu} + \sqrt{\frac{1}{2}(\ln(1/\epsilon)) \sum_i (p_i)^2}$, ($i=1, \dots, n$)是根据 Hoeffding bounds 计算得到的已接纳的 n 个流的等价带宽(Equivalent Bandwidth)。 $\hat{\nu}$ 是已接纳流平均到达速率的测量值, p_i 是已接纳流 i 的峰值速率, ϵ 是到达速率超过链路能力的概率, μ 是链路带宽能力。

(2) 测量机制 基于测量的接纳控制算法所使用的测量机制有如下几种典型方法。

时间窗口(Time-Window)法 文[17]提出了一种可用于 Measured Sum 算法的测量机制——时间窗口。该算法将测量过程划分为若干个长度为 T 的时间窗口,在每个时间窗口内再划分出若干个采样周期 S ,在每个采样周期 S 内测量链路平均负载,并将时间窗口 T 内测得的最高平均负载值作为一个时间窗口内对链路负载的估值。当一个新的流被接纳,该估值就要加上该流占用的带宽值进行更新。如果在某个采样

周期 S 内测得的链路平均负载值高于链路负载估值,那么就要将估值更新为测量值。

在时间窗口方法中,链路负载估值对接纳控制的影响取决于时间窗口 T 和采样周期 S 的长度。采样周期 S 越短,时间 T 内测量得到的最大链路平均负载值就越高,接纳控制就越趋于保守;同样,如果 T 越长,那么对链路负载的估值保持时间越久,也导致保守的接纳控制。为获得一个更佳的结果,文[14]建议 $T/S \geq 10$ 。

指数加权移动平均(Exponential Weighted Moving Averaging)法。文[16]提出了一种可应用于 Equivalent Bandwidth 算法的测量机制——指数加权移动平均。该方案在每隔一个采样周期 S ,测一次流的平均分组到达速率 $\hat{\nu}^s$,然后使用下述公式计算本采样周期内分组平均到达速率的估值:

$$\hat{\nu}' = (1-w) \times \hat{\nu} + w \times \hat{\nu}^s$$

其中 $\hat{\nu}^s$ 是本采样周期内平均分组到达速率的测量值, $\hat{\nu}$ 是一个采样周期的平均分组到达速率的估值, w 是用户指定的加权系数。

在过去十年中,已提出了若干种基于测量的接纳控制算法。文[14,18]对这些算法进行了归纳和分析,其研究表明,对于同样的 QoS 要求,所有基于测量的接纳控制算法性能非常接近。同时文[18]也指出,由于此类算法的两个固有点:偏向于对链路资源要求比较小的和传输路径比较短的流,因此对现有此类算法可改进的余地有限;后续的研究工作可以考虑如何使基于测量的算法能够支持更多的策略(policy)。

2.3 基于带宽代理的接纳控制算法

上述两种接纳控制算法都是采用分布式的工作方式,需要在分组流沿路的每一个路由器中执行接纳控制的过程,并维护每流的信息。因此,此类算法的可扩展性比较差。从提高可扩展性的角度出发,文[7]提出了基于带宽代理(BB, Bandwidth Broker)^[19]的接纳控制算法。该算法不需要核心路由器维护每流的服务质量信息和状态信息,而是采用中央集中的方式来维护和管理这些信息。此算法能同时支持每流的和基于流量类(class-based)的确保型服务。

带宽代理是网络中路由器(或交换机)中的一个软件包,或者是一个网络中独立的服务器。带宽代理主要包括两个组成部分:接纳控制模块和路由模块。前者用于维护带宽代理所管辖的网络域的每条链路上的服务质量状态,负责执行接纳控制和资源预留的功能;后者负责向域中其他路由器获取网络拓扑结构信息,用于执行路径选择和设置。带宽代理还包含了一个存储所管辖网络拓扑结构、策略信息、每个链路和节点上的流以及服务质量状态信息的数据库。

基于带宽代理的接纳控制假设网络是基于虚拟时间参照系(VTRS, Virtual Time Reference System)的核心无状态的单个网络域。该算法的执行过程如下:

当一个需要 QoS 保证的流发出接纳请求时,首先到达入口路由器,然后由入口路由器发送一个该流的接纳请求报文给 BB。BB 首先根据事先制定的策略判断该流是否可以接纳,如果根据策略不能接纳该流,就立即拒绝。如果可以接纳,BB 根据其存储的网络拓扑信息为该流选择一条适当的路径,然后再判断该路径是否有足够资源可接纳该流。如果该流能够被接纳,BB 就发送一个肯定接纳的报文给入口路由器,在入口路由器处设置(或重配置)一个调节器(conditioner,用于在分组进入网络之前在其头部插入状态信息)。同时更新 BB 中的相关数据库。

基于带宽代理的接纳控制算法的接纳条件是,为该流选

择的路径上的最小剩余链路带宽大于或等于该流的速率。BB 利用存储在数据库中的每个路由器中所有流的状态信息来计算每个链路的可用带宽的。

如果有特性相同的流存在,BB 将把这些流作为一个服务类来处理,以便减少它所需要维护的流状态信息的数量,从而可以更快地进行计算和接纳数量更多的流。基于类提供接纳控制的具体接纳条件可参见文[7]。

这种算法的优点是不需要核心路由器来执行接纳控制,极大地降低了核心路由器的工作强度;相比 IntServ/RSVP,可扩展性较好,也可以提供确保型服务。它的缺点是 BB 需要管理整个网络,并存储网络中的所有流和路径的信息。对于大型网络来说,只使用一个 BB 是不可能的,因此,大型网络采用一种分布式的结构,即将整个网络划分为多个域,每个域设置一个 BB。但是需要在各个域的 BB 之间交换域状态信息,从而增加了实现的难度。关于上述分布式的带宽代理方案,也是当前研究的一个热点,仍有许多问题需要解决^[7]。

2.4 基于策略的接纳控制算法

上面的几种接纳控制算法仅根据发送方的请求和网络的可利用资源来执行接纳控制,而没有考虑到网络管理员和服务提供者希望对网络使用情况的控制策略。比如根据用户的身份、对带宽的要求、对安全的考虑、日期以及统计和计费等等因素来执行接纳控制。因此 IETF 在 IntServ/RSVP 的基础上提出一个基于策略的接纳控制算法(PBMC, Policy-Based Admission Control)^[6]。这种算法既要考虑传统的接纳控制标准,还要根据指定的策略标准来执行接纳控制。

PBMC 的体系结构主要有两个元素,策略执行点(PEP, Policy Enforcement Point)和策略决策点(PDP, Policy Decision Point)。

PEP 是运行在需要执行基于策略的接纳控制的每个网络节点上的程序,PDP 可以是远端的一个实体,比如一个策略服务器。PDP 根据事先制定的策略做出接纳决策,PEP 是接纳策略被实际执行的地方。PDP 可以利用其他机制或协议来实现诸如用户认证、统计、策略信息存储等功能。比如 PDP 可以使用基于 LDAP 的目录服务来存储和取得策略信息^[20]。

接纳控制的执行过程首先从 PEP 开始。PEP 首先收到一个要求进行策略决策的报文,然后 PEP 根据这个报文构造一个策略决策的请求报文,并发送给 PDP。该请求报文中包含了一个或多个策略元素以及执行接纳控制需要的信息(比如流的特性参数、对带宽的要求数量等)。PDP 根据这个请求进行策略决策,并执行传统的接纳条件判断,然后把结果(接纳或者拒绝)返回给 PEP。

IETF 的 RSVP 工作组已经提出了一种用于在 RSVP 网络中提供基于策略接纳控制能力的协议——通用开放策略服务协议(COPS, Common Open Policy Service Protocol)^[20]。这是一种典型的集中式的接纳控制机制,其优点是结构简单易于实现。但是其缺点也是明显的。首先,一旦中央策略决策点失效,会导致整个网络系统陷入瘫痪;其次是此方案的可扩展性也不好,一旦网络规模增大,中央决策点就会成为一个瓶颈。

2.5 端点接纳控制算法

自从 IntServ 提出以来,在接纳控制领域针对 IntServ 体系结构的可扩展性问题就展开了广泛的研究,希望通过减少或避免在网络核心路由器维护每流信息和资源预留信息,提高可扩展性。基于这个思路提出了只在网络边界节点(出/入口路由器或交换机)或者端系统(主机)处执行接纳控制的方法。

案,可以统称为端点接纳控制方案(EAC, Endpoint Admission Control)。虽然这些方案都是在端点处执行接纳控制,但具体的接纳控制机制差别也比较大,可以分为以下几种类型:

(1)基于探测(probing-based)的端点接纳控制 文[8,9,21~24]涉及的算法可以归入这一类型。它们的特点是只在端系统(主机或者入/出口路由器)处根据将要接纳的流的到达路径上的拥塞程度来执行接纳控制。在一个流建立之前,发送方首先向网络中发送一个具有和该流相同流量特性参数的探测流,然后在接收端测量这个探测流的分组丢失率、或者根据探测流分组携带的ECN^[25]拥塞标记,来推断网络的拥塞程度。如果探测流的丢失率(或者是被标记ECN的比例)低于指定的极限值,那么该流被接纳;反之,该流被拒绝。由接收端向发送端发送接纳结果,然后发送端根据此结果发送或者不发送正式的分组流。

在这种机制中,接纳控制由端系统来执行,而且不需要任何的信令机制,也不需要边界或核心路由器具备任何特定的功能(只需具备普通的分组丢弃和ECN标记功能即可)。由于此机制根据不是很精确的测量结果来执行接纳控制,因此该机制主要用于提供负载受控(controlled-load)的服务。

它的缺点是使用探测流会给网络增加额外的负载,而且探测过程通常长达数秒,造成初始化一个请求所需的时间延迟比较大,使得这一算法并不适合某些对延迟要求比较高的应用场合。此外,在路由器使用公平队列调度(Fair Queuing)和基于类(class-based)的调度算法的场合,此算法无法感知其它流量的存在。因此,基于探测的端点接纳控制并不适用于这些场合^[8]。

(2)基于网络(network-based)的端点接纳控制 为了克服基于探测的端点接纳控制方案中分组流的建立时间过长的缺点,文[26]提出了一种无需探测而基于网络的端点接纳控制方案。

该方案的原理是,在重配置(re-configuration)网络的时候(离线状态,比如网络的设计和配置阶段),通过预分配(pre-allocating)资源减少运行时执行资源分配的开销。具体做法是,在离线时根据资源的实际利用情况,建构网络资源的汇聚树(sink tree)^[27]信息。该汇聚树以出口路由器为根,入口路由器为叶子。因此从某个入口路由器到任意一个出口路由器,就只有唯一的一条路径。此外汇聚树还标识了每条路径的各段链路上可用资源的数量。在运行时,入口路由器可以根据到达流的目的地址,迅速找到出口路由器的汇聚树,由于通往出口路由器的路径是唯一的,路径上的可用资源也是预知的,就可以很快制定接纳控制的决策。

具体的接纳控制过程如下:当一个实时流(假定该流是由一个漏桶源发出来的)需要建立连接时,发送一个包含了目标IP地址,需要的带宽和对延迟的要求等信息的请求报文给入口路由器。入口路由器根据该报文在它的路由表中查找到相应的出口路由器和端口,根据这些信息,入口路由器可以决定使用哪一个汇聚树,并根据这个汇聚树上可用的带宽和最坏情况下的网络延迟来决定是否接纳该流。

一旦确定了一个流要经过的汇聚树,那么在入口路由器处就可以标记该流的每个分组,表明其所属的汇聚树。此后网络核心路由器就可以根据此标记进行转发,而不需执行复杂的路由协议。

相比基于探测的端点接纳控制方案,该方案具有以下优点:1)流的建立没有任何延迟;2)路由开销很小,只需在入口路由器处对分组进行标记;3)较高的接纳概率和网络利用率;

4)可提供严格的端到端分组延迟上界。

该算法的特点是在离线状态下,通过对网络资源的预分配,利用汇聚树建构和描述了网络的可用资源情况,使得在运行时,只需在入口路由器处执行接纳控制,并可以极大地减少运行时的计算负载。该方案的缺点是在网络中的每个路由器上必须支持汇聚树的计算,并需要有少量的信令消息在路由器之间交换资源共享信息,降低了该算法的可部署性(deploy-ability)。

(3)基于测量(measurement-based)的端点接纳控制 同样是为了避免流的建立时间并获得较好的可扩展性,文[10]提出了一种只在出口路由器(egress router)执行接纳控制的方案。

该方案将整个网络域看作一个“黑盒”系统,可以不需知道网络域内部使用的调度算法、网络核心负载,甚至也不必了解交叉流量(cross traffic)的情况,只需在出口路由器测量端到端路径的可用资源,并以此为依据决定是否接纳新的流。

具体的接纳控制过程如下:当一个流发生之前,首先使用RSVP的简化协议,发送一个资源预留请求报文给出口路由器,核心路由器只负责转发该报文。然后,出口路由器测量该路径上的可用资源是否满足该请求。如果满足,就接纳该流,反之拒绝该流。

这个方案的关键在于,如何在出口路由器测量端到端路径的可用资源。该方案应用了基于测量的包络线(envelope)理论^[22]来测量黑盒系统的到达(arrival)和服务(service)包络线(即到达入口路由器的流速率和经过出口路由器后的流速率的包络线),并根据这两个包络线推断出端到端路径的可用资源。

虽然基于测量的端点接纳控制只能提供负载受控型服务,但它的优点也很明显:1)不需核心路由器处理资源预留消息;2)不需要核心路由器执行复杂的调度算法;3)对于因路由波动引起的服务变化,通过使用服务包络线计入其影响,不需使用路由绑定(route pinning);4)本方案是基于类提供服务的,边界路由器也是针对流量汇聚执行接纳控制,不需维护每流信息。

上面我们已介绍了主要的接纳控制机制,比较了它们之间的联系和不同。我们认为,基于IP网可提供QoS保证的接纳控制算法的研究主要是围绕减少核心路由器所需维护的状态信息,提高算法的可扩展性这一目标开展的。在已经提出的算法中,基于参数的算法按照每流的峰值带宽预留资源,虽然能为每流提供确保型服务,但是该算法的网络利用率太低。基于测量的算法通过实时测量网络被占用的资源,从而得到网络剩余能力,并据此执行接纳控制,提高了算法的网络利用率。但是此类算法只能提供部分确保型的服务,而且仍然需要维护核心状态,没有解决可扩展性问题。与上述两种采用分布式方法来执行的接纳控制机制不同,基于带宽代理的算法采用中央集中式方式;它虽然满足了核心无状态的要求,但是仍然不能满足大规模网络的要求,还有许多需要解决的问题。基于策略的算法也是采用中央集中的方式,目的是给网络管理员和服务提供者提供对网络的更多的控制能力。最后一类基于端点的接纳控制机制,为了解决核心状态的问题,提高可扩展性,只在网络边界执行接纳控制,彻底保证了网络核心的无状态;这一思路对网络规模的适应性很好,可部署性也比较好,但是所有不需核心路由参与的接纳控制算法,只能提供部分确保型的服务。

3. 接纳控制算法的发展趋势

根据上述分析,我们认为现阶段接纳控制算法的主要研

究思路会从最初注重于接纳控制结果的精确性,转向以降低、避免核心状态,提高接纳控制算法的可扩展性。在这一领域的研究工作还包括以下方面:

(1)对已有算法进行分类、综合和性能评价。如文[5,14,18]对基于测量的这一类接纳控制算法进行总结与评价,并得出了一些指导性的结论。对于当前已提出的几种基于端点的算法之间的联系与区别,性能的比较和分析,还有待进一步研究。这方面研究的意义在于揭示所综合的各种算法之间的内在联系,为今后对这些算法的改进和新算法的提出提供了指导方向。

(2)对已有算法的改进。目前已经进行的工作主要集中在对基于测量的接纳控制算法上,但由于此类算法可改进的余地有限^[18],改进工作主要是在网络利用率一定的前提下,对其它指标的改进或基于策略参数的改进。如文[28]提出的基于测量的自适应的接纳控制算法,在系统资源利用率(或接纳能力)接近的情况下,与传统固定时间窗口的算法相比,能获得更好的分组丢失率。文[29]提出的基于ON/OFF马尔可夫数据流可根据流的相对价格(优先权大小)优化接纳控制的方案。

(3)随着近年来无线应用的快速发展,基于无线IP网络的接纳控制算法研究也是当前的一个热点。

通过以上综述和分析,我们认为,由于基于端点的接纳控制机制是核心无状态的,不需要核心路由器维护每流的状态信息和QoS信息,接纳控制只在网络边界执行,因此具有较高可扩展性,同时也能提供比较严格的(负载受控的)QoS保证,综合了IntServ和DiffServ的优点,可部署性也比较好。因此,在这一方向上的研究将是当前以及将来一段时间内的重点。

总结 在下一代Internet技术的研究中,如何保证IP网络服务质量(QoS)是其中的一个关键内容,而接纳控制机制是实现QoS控制不可或缺的一个关键技术。从20世纪90年代以来,在接纳控制领域进行了许多研究,本文比较全面地综述了基于IP网络用于提供QoS控制的接纳控制算法的研究现状,详细地描述各种主要的接纳控制算法的特征,并分析了每种算法的优缺点以及相关研究的发展趋势。

考虑到算法的精确性、可扩展性和可部署性等问题,上述接纳控制机制都还不是最佳的解决方案。如何找到一种能够灵活适应各种网络规模,能为不同要求的流提供多种确保类型的服务,并能与其它网络协议和算法(如组播协议、各种分组调度算法等)协调工作的接纳控制机制,仍然有待人们的努力。

参 考 文 献

- 汪芸,顾冠群. QoS接入控制必要性研究. 小型微型计算机系统, 1999, 20(9)
- Braden R, Clark D, Shenker S. Integrated Services in the Internet Architecture: An Overview. Internet RFC 1633, 1994
- Zhang L, et al. RSVP: A New Resource ReSerVation Protocol. IEEE Network, 1993, 7: 8~18
- Blake, et al. An Architecture for Differentiated Services. Internet RFC 2475, 1998
- Breslau L, Sugih J, Shenkers. Comments on the Performance of Measurement-Based Admission Control Algorithms. In: Proc. of IEEE Infocom 2000, March, 2000. 26~30
- Yavatkar R, Pendarakis D, Guerin R. A Framework for Policy-based Admission Control. Internet RFC 2753, 2000
- Zhang Z, et al. Decoupling QoS Control from Core Routers: a Novel Bandwidth Broker Architecture for Scalable Support of Guaranteed Services. In: Proc. of ACM SIGCOMM'00, Stockholm, Sweden, August 2000
- Breslau L, et al. End-point Admission Control: Architectural Issues and Performance. In: Proc. of ACM SIGCOMM'00, Stockholm, Sweden, Aug. 2000
- Cetinkaya C, Knightly E. Egress admission control. In: Proc. of IEEE INFOCOM 2000, Tel Aviv, Israel, Mar. 2000
- Centinkaya C, Knightly E. Scalable Services via Egress Admission Control. In: Proc. of IEEE INFOCOM'00, Tel Aviv, Israel, March 2000
- Berson S, Vincent S. Aggregation of internet integrated service state. Internet Draft 1997. draft-Berson-classy-approach-01.txt
- Pan P, Hahne E, Schulznhne H. BGRP: A framework for scalable resource reservation. Internet Draft 2000. draft-pan-bgrp-framework-00.txt
- Ferrari T, Almesberger W, Boudec J L. SRP: a scalable resource reservation protocol for the Internet. In: Proc. of IWQOS '98, Napa, CA, May 1998. 107~116
- Jamin S, Shenker S J, Danzig P B. Comparison of measurement-based admission control algorithms for controlled-load service. INFOCOM '97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Driving the Information Revolution, Proceedings IEEE, 1997, 3: 973~980
- Clark D D, Shenker S, Zhang L. Supporting real-time applications in an integrated services packet network: Architecture and mechanism. SIGCOMM Symposium on Communications Architectures and Protocols, 1992. 14~26
- Floyd S. Comments on Measurement-based Admissions Control for Controlled Load Services: [Technical report]. Lawrence Berkeley Laboratory, July 1996
- Jamin S, Danzig P B, Shenker S, Zhang L. A Measurement-based Admission Control Algorithm for Integrated Services Packet Networks (Extended Version). ACM/IEEE Transactions on Networking, Dec. 1996
- Breslau L, Jamin S, Shenker S. Measurement-based admission control: what is the research agenda? Quality of Service, 1999. IWQoS'99. 1999 Seventh International Workshop on, 1999
- Nichols K, Jacobson V, Zhang L. A two-bit differentiated services architecture for the Internet. Internet RFC 2638, July 1999
- Herzog S, Boyle J, et al. COPS usage for RSVP. Internet RFC 2749, 2000
- Bianchi G, Capone A, Petrioli C. Throughput analysis of end-to-end measurement-based admission control in IP. In: Proc. of IEEE INFOCOM 2000, Tel Aviv, Israel, Mar. 2000
- Elek V, Karlsson G, Ronngren R. Admission control based on end-to-end measurements. In: Proc. of IEEE INFOCOM 2000, Tel Aviv, Israel, Mar. 2000
- Gibbens R, Kelly F. Distributed connection acceptance control for a connectionless network. In: Proc. of ITC'99, Edinburgh, UK, June 1999
- Kelly F, Key P, Zachary S. Distributed admission control. IEEE Journal on Selected Areas in Communications, 2000
- Ramakrishnan K, Floyd S. A proposal to add explicit congestion notification (ECN) to IP. 1999. Internet RFC 2481
- Choi B K, Bettati R. Endpoint Admission Control: Network Based Approach. In: Proc. of the 21st Intl. Conf. on Distributed Computing Systems, Phoenix, AZ, April 2001
- Ferrari T, Almesberger W, Boudec J L. SRP: a scalable B. Choi, R. Bettati, Efficient Resource Management for Hard Real-Time Communication over Differentiated Services Architectures. In: The 7th IEEE Intl. Conf. on Real-Time Computing Systems and Applications, Cheju, Korea, Dec. 2000
- 马小骏, 顾冠群. 基于测量的接纳控制研究. 计算机学报, 2001, 24(1)
- 傅晓明, 张尧学. 一种用于提供QoS保证的准入控制方法. 电子学报, 2000, 28(10)