

网边缘可控的 IP QoS 体系结构及其算法

樊秀梅¹, 林 闯¹, 王忠民²

(1. 清华大学计算机科学与技术系, 北京 100084; 2. 北京科技大学计算机系, 北京 100083)

摘 要: IP QoS 要求是综合宽带网发展的重要瓶颈, 而现有 IP QoS 体系不能很好地满足用户需求, 因此, 本研究将对 IP QoS 体系做一个新探索, 依托当前网络环境, 着眼于下一代网络发展, 研究一种网边缘控制的自组织 IP QoS 体系结构及其算法, 本体系中采用探测分组、染色分组和信息分组三类分组, 并通过智能控制系统来控制三类分组的比例关系, 有效地提供用户所需求的 QoS. 本体系将克服现有机制的流状态保持, 并通过用户的自组织行为来加强网络的自适应性与可扩展性, 更好适应网络发展与应用需求.

关键词: IP QoS 体系结构; 自组织行为; 网边缘控制; 自适应性

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2002) 12A-2027-05

IP QoS Architecture and Algorithm Based Network Edge Controlling

FAN Xiu-mei¹, LIN Chuang¹, WANG Zhong-min²

(1. Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China;

2. Dept. of com. Science, uni. of Science and Technology Beijing, Beijing 100083, China)

Abstract: IP QoS is the important bottleneck of Integrated Broad Band Network development, but the active IP QoS architecture can't satisfy the users requirements commendably. So, our research would be a new exploration about IP QoS architecture. Relying on current network condition and next generation network development demand, we researched a sort of Self-organizing IP QoS architecture based on network edge controlling and provided its routing realization algorithms. In this architecture, artificial intelligence will control the proportion of explorer packets, chromosome packets and information packets, and provide effectively QoS for users. This architecture will overcome flow state keeping in the active mechanism and enhance network self-adapted and scalable through the users self-organizing behaving. This will accommodate the network development and application requirements much better.

Key words: IP QoS architecture; self-organizing behaving; network edge controlling; self-adapted

1 引言

迄今为止, Internet 已经发展成了一个连通世界各个角落的大型互连网络. 目前, 传统的 Internet 体系结构提供一个非常简单的端到端传输服务, 它是基于“best effort”服务模式的, 其网络结构、通信协议或质量要求均以数据通信的要求为前提, 传输的质量得不到保证. 在这种模式中, 网络提供的最高保证是可靠的数据传输, 它使用 TCP/IP 等协议族. 这对传统的数据应用, 像 FTP (File Transfer)、HTTP (Hyper Text Transfer Protocol)、DNS (Domain Name Service) 和 Telnet 等较为适用. 这些业务从本质来说有很大的弹性, 它们可以容忍较大的传输时延和一定程度的数据丢失, 在网络发生拥塞的情况下可以延迟发送, 数据传输的正确性比及时性更为重要. 现行的 Internet 体系结构是上世纪 70 年代确定的. 在 Internet 的发展初期面对的主要需求是网络互连、网络的健壮性、网络设备的异

构性、分布式的管理、低成本和容易上网, 因此, 当时确定的 Internet 网络体系结构是无连接的分组交换结构, 高层的功能放在网络边缘, 采用路由机制和尽力服务 (Best Effort) 原则. 但是, 30 年来, 随着网络规模、用户数量的迅猛增长, 以及多媒体应用的不断发展, 近年来发展的几种新型分布式应用, 像视频会议、视频点播 (VOD) 和虚拟现实 (Virtual Reality) 等多媒体实时应用和综合分布式控制, 已从单纯的数据向语音、图像、动画、视音频等多媒体信息扩展, 它们对从网络收到的服务质量 (QoS: Quality of Service) 保证是较为敏感的, 迫切要求 Internet 能够提供应用所需求的服务质量. 特别是, 它们对传统方式下确保正确的、公平的传输业务, 而延误了时延是不可能接受的. Internet 的巨大发展和各种以前未想到的应用已对网络体系结构提出了许多新的需求, 所有这些新应用要求都迫使我们不得不重新审视 Internet 体系结构, 甚至考虑要进行革命性的变革. 但网络体系结构的改变不仅要考虑技术上的

收稿日期: 2002-05-10; 修回日期: 2002-09-08

基金项目: 国家重点基础研究发展规划 (973 计划) 项目 (No. G1999032707); 国家高技术研究发展计划 (863 计划) 课题 (No. 2001AA112080, No. 2001NA411040); 国家自然科学基金项目 (No. 90104002, 60173012)

需求,还要考虑许多非技术的因素,包括企业模式、法律法规等等,因此有关基于网络边缘的研究^[11]一直是一个热点研究问题。

本文的研究将是对 IP QoS 体系的一个新探索.它依托当前网络环境,着眼于下一代网络发展,研究一种网边缘控制的自组织 IP QoS 体系结构及其具体的实现算法.本体系结构将克服现有 IP QoS 机制的流状态保持,并通过用户的自组织行为来加强网络的自适应性与可扩展性,更好地适应网络发展与应用的需求。

2 相关的研究成果

在应用的推动下,国内外同行对 IP QoS 问题进行了大量的研究与论证,已取得了一些基本的阶段性成果,目前主要的有 IntServ、DiffServ、SCORE 等 IP QoS 体系结构。

2.1 InteServ(Integrated Service)是 IETF 1994 年提出的具有 QoS 保证的网络体系结构(RFC 1633^[11]).它是在传统 IP 网络“best-effort”服务的基础上定义了一组服务扩充,通过 RSVP 提供端到端的 QoS 保证.其主要的不足有:

①其工作方式是基于流的,需要复杂的资源预留、接纳控制、QoS 路由和调度机制.在大规模网络中,链路的状态是不确定的,带宽的有效预留是一件非常困难的工作。

②RSVP 的有效实施必须依赖于分组所经过路径上的路由器.但是许多现有路由器和交换机不支持 RSVP,无法实施真正意义上的资源预留。

③高速网络核心路由器常常被迫去维护、调度数以千计的连接;随着连接数量的增加,连接建立和连接释放阶段的额外开销也将会增大.每个连接的状态信息需要保存在路径通过的每个节点上,即需要保存每流状态,其可扩展性较差。

④提供较好 QoS 保证的 IntServ^[8]具有某种面向连接的特性,而 IP 网络本身不具有面向连接特性的,故在具体实施中,如何与现有网络互连也是一个困难。

2.2 对 IntServ 在发展中遇到的困难,IETF 在 RFC2475^[2]中提出了 **DiffServ**(Differentiated Service)体系结构. DiffServ 比 IntServ 简化了信令,将网络上传输的业务分成不同的业务流,每种业务类定义一组单向的分组传输特征,由区分服务码(DSCP)标识. DiffServ 通过汇聚(aggregate)和 PHB(Per Hop Behavior)方式提供一定程度上的 QoS 保证. DiffServ 最大的特点是实现简单、扩展性好,但也存在如下一些不足:

①由于 DiffServ 结构中网络和端系统之间缺乏信号通信,不能提供端到端的 QoS 保证,而端到端的 QoS 保证一直是 IP QoS 研究的基本目标。

②DiffServ^[9,10]只能使类型不同的业务享受不同的处理速度、平均带宽和平均丢失率.仅是一种统计上的有限处理,所以仍不能很好地保证业务流的 QoS。

2.3 SCORE(Stateless Core)^[3]是近来研究者提出的一种核心无状态结构,该结构在网络边缘节点为每个流保留状态信息,并由边缘网络节点将部分状态信息记录在其每个数据包的头部.该体系结构的主要技术是 DPS(Dynamic Packet State),此信息随着数据包在网络节点间的转发而不断更新。

核心网络节点根据包头的状态信息进行处理,从而达到与传统的需要保留每个流状态信息的方法具有相同性能.其明显的不足在于:

①对现有的网络协议进行了一些改动,在分组头上加入了一些流状态信息。

②为了保持同步,加入一些空分组,浪费了带宽。

③需区分边缘节点与核心节点,并在不同的节点进行不同的处理,增加了复杂性与不可操作性。

3 网边缘可控的 IP QoS 体系结构

本文的研究主要是针对国内外 IP QoS 体系结构的研究现状与当前世界网络应用飞速发展的需求而提出的,是依托当前网络环境、着眼于下一代网络的一种 IP QoS 体系结构研究.该研究将网络分为边缘与核心,网络边缘将对反馈的分组进行一些智能化的处理,而在网络核心只根据传输分组头上的信息与路由表进行信息传递,是一种网边缘可控的 IP QoS 体系结构.本文提出的体系结构将利用用户的自组织行为来影响、控制网络提供的服务质量,并使得网络具有良好的扩展性、自适应性与自组织性,是对现有 IP QoS 体系的一种新探索。

3.1 网边缘可控的 IP QoS 体系结构的基本思想

在用户端,用户提出一个 QoS 需求,然后在发送信息(information)分组的同时发送一定比例的探测(explorer)分组.探测分组到达目的地时,目的地根据用户的 QoS 需求产生一个或几个染色(chromosome)分组.染色分组到达边缘路由器时,边缘路由器根据一定的规则修改染色分组所携带的信息,然后让其原路返回用户端,并在返回途中调整经过路由器的行为规范.而信息分组则按照到达路由器上的行为准则进行信息传输.这样,在网边缘可控的 IP QoS 体系中,边缘可根据用户的要求与行为来调整网络对他的服务提供.同时,在边缘路由器要记录每个流所得到的服务质量,并依此计费.这样可以避免用户盲目扩大自己的要求,避免了恶意用户的恶意攻击;而在核心路由器不保持任何流状态,具有较强的可扩展性,网边缘可控的 IP QoS 体系结构有效地提高了端到端的服务提供(QoS)。

3.2 具体实施方法与采用的技术

本文提出的网边缘可控的 IP QoS 体系结构中有探测分组、染色分组和信息分组三类分组类型.其具体的分组比例关系的有效控制是通过一个智能控制系统来进行的,这是我们正在进行的工作.不同分组的分组结构描述具体表示如表 1、2、3 所示。

3.3 网边缘可控的 IP QoS 体系结构的特点

(1)该体系中的 QoS 控制决断将主要由边缘路由器做出,核心路由器的任务将简化为通报网络信息和协调用户决断这两个较为简单的功能,而大量复杂的、较为慢速的计算工作将被推到网络边缘的各个终端去完成.这样,既充分利用了未来网络中边缘的剩余处理能力来构成强大的网络计算实体,又大大减少了网络本身实施同样计算规模所必须花费的成本。

(2)由网络边缘做出控制决断将更有效地把用户和网络

表 1 探测分组的分组结构描述:

0	8	16	31
分组类型(2)	分组大小	分 组 大 小	
分 组 时 戳	QoS 描 述		
源 IP 地 址			
目 的 地 IP 地 址			
记 录 分 组 跳 数	预 留		
预 留	记录分组在路径上的延迟		
记录目前分组所在节点的 IP 地址			
分组经过的第一个节点 IP 地址			
分组经过的第二个节点 IP 地址			
.....			

表 2 染色分组的分组结构描述如下:

0	8	16	31
分组类型(3)	分组大小	分 组 大 小	
分 组 时 戳	QoS 描 述		
源 IP 地 址			
目 的 地 IP 地 址			
记 录 分 组 跳 数	预 留		
预 留	记录分组在路径上的延迟		
记录目前分组所在节点的 IP 地址			
分组经过的第一个节点 IP 地址			
分组经过的第二个节点 IP 地址			
.....			

表 3 信息分组的分组结构描述如下

0	8	16	31
分组类型(1)	分组大小	分 组 标 号	
分 组 时 戳	QoS 参 量 说 明		
存活时间	协议	报头校验和	
源 IP 地 址			
目 的 地 IP 地 址			
信息数据			
.....			

提供的服务质量结合起来,最大限度地满足用户的各种应用需求。

(3)该体系是一种互动式自组织结构,每个网络边缘的工作不仅有利于自己,更多的是使其它边缘受惠.经过一段时间的自组织、互动行为,网络对每个用户将提供他需要的最佳服务,亦能充分利用网络资源。

(4)本体系保持了核心无状态的特性,即在核心路由器中无需为每个流保持流状态,这与传统的 IP 风格相吻合。

(5)具有较强的可扩展性。

综上所述,本研究适应当前计算机网络发展的需要,提供一种简单、有效、经济、可靠的 IP QoS 解决方案,既可以充实这一研究领域,又可以解决目前 Internet 上存在的问题,同时也保护了目前网络的资金投入,具有有效的经济效益和社会效

益。

4 具体的路由实现算法

已有的 QoS 路由^[5,6,7]研究已有很多.在本文中,我们进行了一种新的探索,提出一种网边缘可控的 IP QoS 路由算法。

4.1 提出了一种网边缘可控的分布式、可扩展路由与 QoS 路由算法

终端发送数据时,同时向同一目的地发送一定比例的路由探测分组.路由探测分组随机地选择到某一子网的出行端口,并在传输过程中,记录所经过的节点信息,将其标记在分组头上.当路由探测分组到达目的地时,目的地比较这些探测分组头上所记录的信息,根据自己的 QoS 需求选出最优的一个或几个路由探测分组,相应地产生路由染色分组.路由染色分组根据分组头上的历史记录返回源端.路由染色分组在到达边缘路由器时,边缘路由器根据现在的网络状态与网络信息修改路由染色分组,后将其返回.在返回途中,路由染色分组更新经过节点的路由表.我们定义路由表中的路由表项值称为信息素.假定开始状态时,路由表中对应每端口的信息素是一相同值,分组根据信息素的大小(或加上其它什么参数)依概率选择它的传输路径.依此过程,经过一定的时间,网络可找到满足用户需求的最优路径选择。

我们将网边缘可控的可扩展路由算法简称为 TDF(Terminal-Edge Decide Feedback),将网边缘可控的可扩展 QoS 路由算法简称为 TDFQ(Terminal-Edge Decide Feedback QoS).TDF 和 TDFQ 算法的路由过程总的来说是一个四步行为,即路由探测行为、终端决断行为、路由表更新行为与数据转发行为.这个四步行为并不是依序进行的,它们可能同时进行或交叉进行。

本文中基于边缘的分组交换网络中有路由探测分组、路由染色分组与信息分组三类分组,分组结构同上所述.具体的分组行为如下描述:

路由探测分组在网络传输时进行两种行为:记录信息行为和选择路径行为。

①**记录信息行为:**当路由探测分组在网络中传输时,将经过的每一个节点的有用信息记录在分组头上,以利于边缘路由器据此对网络到用户自身路由的选择和调整。

②**选择路径行为:**为了发现新的路由,路由探测分组的路径选择有一定的随机性,即随机选择路由表里对应每个子网的不同端口。

路由染色分组的行包括三部分:产生路由染色分组、路由染色分组返回源端和信息素调整行为。

①**产生路由染色分组:**当路由探测分组到达接收端时,接收端缓存并比较某一固定时间段内,来自于同一源的路由探测分组所记录的信息,选出接收端认为适合自己要求的、性能好的路由探测分组,以此产生路由染色分组。

②**路由染色分组返回:**边缘路由器根据路由染色分组头上的历史记录与网络的现有状态与网络信息修改路由染色分组,然后路由染色分组按照其来时经过的路由器反向返回源端。

③信息素调整行为: 当一个反向返回的路由染色分组经过某一节点时, 该节点应对输入链路的相应端口的已有信息素进行增强更新. 而信息素是会随着时间的推移而挥发的, 所以, 也要对对应于同一子网的其它端口的信息素进行减少更新. 如果值小于零, 则表示为零; 如果值大于或等于零, 则表示为实际值.

信息素根据如下公式被更新:

$$\tau_{ij}(t+n) = \rho \cdot \tau_{ij}(t) + \Delta\tau_{ij}^k \quad (1)$$

$$\tau_{ij}(t+n) = (1-\rho) \cdot \tau_{ij}(t) \quad (2)$$

这里, $\Delta\tau_{ij}^k = \sum_{k=1}^m \Delta\tau_{ij}^k$, $\Delta\tau_{ij}^k$ 是在时间段 $[t, t+n]$ 之间, 第 k 个分组放在边 (i, j) 上的信息素量. ρ 是坚持率 (有效率), 这样 $(1-\rho)$ 就是挥发率.

信息分组的传输是根据路由表里对应每个子网不同端口的信息素来选择路径的. 在本文的分组交换网络中, 信息分组选择信息素最大的端口输出数据. 当信息分组到达终端时, 就被终端用户接收, 不再产生任何行为.

4.2 仿真结果及分析

本文的仿真是在一个具有 30 个节点和 110 条链路的拓扑结构中进行的. 下面的所有仿真都是在相同的仿真环境和相同的流量场景条件下进行的, 在理论和实际中都具有可比性. TDF 的详细描述见参考文献 [4].

4.2.1 分组输入速率变化时 TDFQ 与 TDF、OSPF 的分组平均时延比较

为了验证 TDFQ 提供 QoS 的能力, 我们对不同的情况都做了仿真. 仿真结果显示, TDFQ 能较好地满足用户的时延需求, 是一种有效的 QoS 路由方案. 仿真结果如图 1 所示.

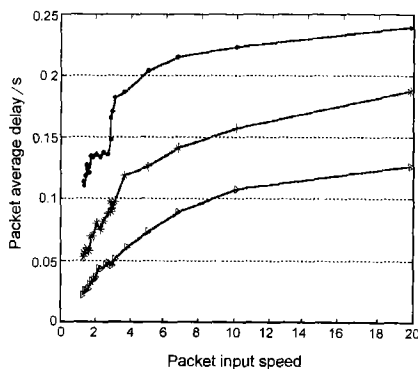


图 1 分组平均输入速率变化较大时的分组平均时延比较

从图 1 可看到, 随着分组输入速率的增长, 三种算法下的分组平均时延都在增大, 这是因为分组输入速率的增加, 就意味着网络平均负载的增大, 每个节点要处理的分组数也相应增大, 故分组平均时延会增加. 从图中还可看到 TDFQ 的分组平均时延是最低的, 这是因为 TDFQ 中的路由选择采用了基于用户时延优化的反馈机制, 它能很快地改变路由到当前对用户 QoS 有效的路径上.

4.2.2 TDFQ 与 TDF、OSPF 的分组丢失率比较

从图 2 中可以看到, TDFQ 与 TDF 的分组丢失率相差不

大, 且二者没有什么明显的优劣之分. 这是因为仿真时, TDFQ 中的分组反馈只强调了分组的时延最小, 而没有对丢失率进行着重改善.

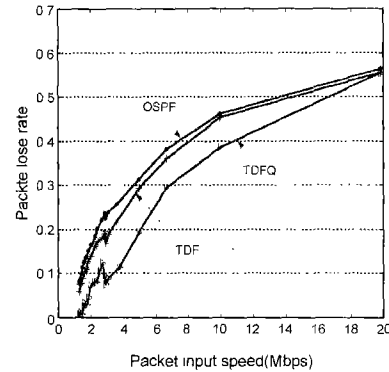


图 2 分组平均输入速率变化较大时的分组丢失率比较

4.2.3 TDFQ 与 TDF、OSPF 在网络链路动态变化时的性能比较

网络结构是动态变化的, 一个好的路由结构体系应很好地考虑到这种动态性, 使其具有实用性. 本文的结构和算法是具有这种动态适用性的. 在仿真期间 (60s), 设定网络中分组的平均输入速率约为 10.7Mbps, 节点以随机的间隔向随机选择的其它节点发送随机长度的数据包. 时间间隔服从平均分布, 其统计平均为 1000ms, 包长服从平均分布, 其统计平均为 800Byte. 在 9.0 秒始, 让链路 74、75、76、77、80、81、82、83 断掉, 这时节点 21 只剩下一个相连的节点 17, 图 3 的仿真结果就是反映如此情况下的分组平均时延对动态变化的自适应性.

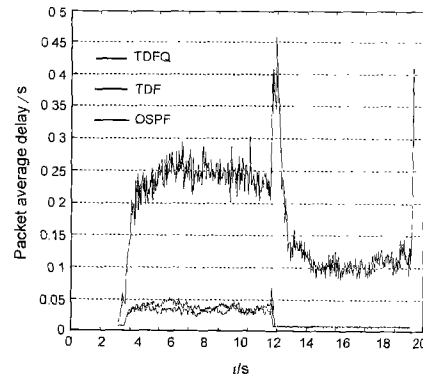


图 3 网络链路动态变化时, TDFQ 和 TDF 的分组平均时延比较

从图 3 可看到, 在网络因链路断掉或其它情况变动下, TDFQ 能很快地自适应这种变动, 分组的平均时延还是低于 TDF 和 OSPF 情况的. 这说明, 在网络动态变化时, TDFQ 也能提供较好的 QoS 保证.

5 小结

随着多媒体应用的不断发展, 网络用户开始提出诸如网络安全性、QoS 保证等各种要求. 为了适应这种不断增长的新需求, 本文提出了网边缘可控的 IP QoS 体系结构, 并提出该

体系结构下的一种基于边缘的分布式、可扩展 QoS 路由算法,并通过仿真说明该算法的有效性与实用性.与传统路由算法不同的是,TDFQ 路由算法不是建立在对全网拓扑完全了解的基础之上,路由形成是通过大量独立运行的主体互相协作而自组织涌现出来,是一种分布式运算方式.TDFQ 路由算法具有良好的动态性、可扩展性和鲁棒性,这使网络运营商大大节约了资本投入,也使用户有了更广泛的选择自由,更加适应新应用的需求.

参考文献:

- [1] RFC1633, Integrated Services in the Internet Architecture: an Overview [S].
- [2] RFC2475, An Architecture for Differentiated Services[S].
- [3] Ion Stoica, Hui Zhang. Providing guaranteed services without per flow management[A]. Proc. ACM SIGCOMM '99 [C]. 1999. Cambridge, MA, USA, 1999. 81 - 94.
- [4] 樊秀梅. 基于用户的分布式可扩展路由及 QoS 路由研究[D]. 北京:北方交通大学研究生院,2001 年.
- [5] RFC 2386, A Framework for QoS-based Routing in the Internet[S].
- [6] Shigang Chen, Klara Nahrstedt. An overview of quality of service routing for next-generation high-speed networks: problems and solutions [J]. IEEE Network, <http://www.nevada.edu/~chen/November/December 1998>.
- [7] Xipeng Xiao, Lionel M Ni. Internet QoS: A big picture[J]. IEEE Network, March 1999. <http://www.cse.msu.edu/~xiaoxpe/download.html>.
- [8] David D Clark, Scott Shenker, Lixia Zhang. Supporting real-time applications in an integrated services packet network[J]. proc. of ACM SIGCOMM. August 1992. 14 - 26.
- [9] Constantinos Dovrolis, Dimitrios Stiliadis, Parameswaran Ramanathan. Proportional differentiated services: Delay differentiation and packet

scheduling[A]. Proc. ACM SIGCOMM '99 [C]. Cambridge, MA, USA, 1999. 109 - 120.

- [10] Differentiated Services (diffserv) [DB / OL]. <http://www.ietf.org/html.charters/diffserv-charter.html>, Mar. 2002.
- [11] Rice University INCITE (InterNet Control and Inference Tools at the Edge) Project [DB / OL]. <http://incite.rice.edu>, 2002.

作者简介:



樊秀梅 女, 1967 年 7 月生于山西省山阴县, 博士, 副教授, 主要研究领域为计算机网络, 系统性能评价等. E-mail: xmfan@csnet1.cs.tsinghua.edu.cn.



林 闯 1948 年 7 月生于辽宁省, 博士, 清华大学计算机系教授, 博士生导师, 计算机网络研究所所长, 《计算机学报》编委, 国家自然科学基金重大项目“网络和信息安全”的科学指导专家, 中科院网络中心和北京科技大学兼职教授. 主要研究领域为计算机网络, 系统性能评价, 随机 Petri 网, 逻辑推演和推理系统. 曾在美国 Purdue 大学、美国 Texas 大学奥斯汀分校、香港科技大学、香港浸会大学和美国 Florida 大学长期做研究工作. 已在国内外一级期刊上和 IEEE Computer Society 的学术年会上发表论文 90 多篇, 并在国内出版专著 2 本.

王忠民 男, 1966 年 10 月出生于河北省新乐县, 博士, 北京科技大学信息工程学院计算机系副教授, 主要研究领域为智能管理、软件工程、数据库及电子商务.