

基于 DiffServ 的高效调度算法的研究*

曹结宝, 周井泉

(南京邮电大学 电子科学与工程学院, 江苏 南京 210003)

摘要: 区分服务(DiffServ)体系是未来 IP QoS 研究的主要发展方向,在区分服务的体系下,队列调度是实现 IP QoS 的核心技术。在深入研究区分服务体系下的基本分组调度算法优缺点的基础上,提出一种改进算法,以队列分组的延迟特性,保证实时业务的实时特性。对改进算法进行了仿真,在多约束下,对性能进行了评价。

关键词: 服务质量; 区分服务; 调度算法; 实时特性

中图分类号: TP393

文献标识码: A

文章编号: 1674-7720(2012)13-0007-03

Research of effective queue scheduling algorithm based on DiffServ

Cao Jiebao, Zhou Jingquan

(College of Electronic Science and Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: DiffServ system is the main development direction of the coming IP QoS research, in the DiffServ service system, the queue scheduling is the core technology of the IP QoS. Based on basic group scheduling algorithm of the DiffServ system, we put forward an improved algorithm to reduce the queue delay characteristics, ensure the real-time nature of real-time business. To simulate the improved algorithm, in many constraints of performance evaluation.

Key words: QoS; DiffServ; scheduling algorithm; real-time nature

随着 Internet 网络的迅猛发展, 促使了各种基于 Internet 应用的数量呈指数级增长, 业务种类同时也发生了很大的变化, 已经从单一数据向集成数据、视频语音、图像等多种业务转变, 例如: 语音 (VoIP)、VoD (Video on Demand) 以及视频会议等。这些新应用对网络的带宽、延迟、抖动等都有一定的要求, 而传统的尽力而为网络不能较好地满足这些新业务在带宽和时延等方面的具体需求。基于此, 本文提出使用 QoS 技术来保证网络的传输质量。QoS 技术并没有扩充网络的带宽, 只是根据业务的需求以及网络状况来管理带宽, 实现带宽的优化。IETF 在 1997 年 9 月制定了有关 QoS 定义和服务的一系列 RFC 标准, 并提出了区分服务 (DiffServ)^[1] 模型的解决方案。在此模型中, 它将 IP 服务类型 (ToS) 字段改名为 DS, 并用它承载 IP 包服务所要求的信息, 是严格意义上的第三层技术。区分业务主要通过两个机制来完成: DS 标记和一个包转发处理库集合的每跳行为 PHB (Per-Hop-Behavior)。通过对一个包 DS 字段的不同

标记以及基于 DS 字段的处理, 就能够产生一些不同的服务级别。IP 包头中的 DSCP (区分服务标记字段) 是 DS 区域的边缘节点和核心节点之间传递流汇聚信息的媒介, 是连接边界的传输分类和调节机制与内部 PHB 的桥梁。

网络 QoS 控制的实质在于资源的管理, 即控制缓冲队列、链路带宽等网络资源的分配与使用。队列管理在网络传输控制中发挥着相当大的作用, 是网络 QoS 控制的核心技术之一, 也是实现网络拥塞控制的重要手段, 是当前 QoS 研究的热点。本文采取 DWRR_PQ 的调度算法可以为 EF 业务提供低延迟、低抖动及对带宽的要求的服务, 将 EF 业务和其他业务进行很好的隔离, 避免 BE 业务可能出现的“饥饿”现象。

1 基于 DiffServ 的队列调度策略

目前最常使用的队列调度算法有 PQ、WRR、DWRR 等。PQ 算法的高优先级队列的分组总是优先得到调度, 低优先级队列的分组只能在高优先级队列分组被调度结束之后才会得到服务, 否则是得不到服务的。所以该

* 基金项目: 国家“863”计划基金资助项目 (2009AA012202)

算法能够给优先级高的业务数据报文提供低延迟和低抖动网络性能,而低优先级的报文可能出现“饥饿”现象。WRR算法:对于所有的业务流在排队等待调度的队列WRR是根据每个队列配置的权值与所有的业务流在排队等待调度的队列的权值总和的比来平等地分配带宽的,克服了PQ调度算法的低优先级队列可能出现长期的“饥饿”现象。同时WRR还可以提供类似公平队列算法的公平性、实现过程比较简单和算法的复杂度只有 $O(1)$,且适用在高速网络下。但是由于该算法轮询服务的特征,从而不能保证EF业务的低延迟、低抖动的性能,不能支持包长度的变化,一旦调度长度一样的数据包就会导致分组长的队列会得到更多的带宽资源,出现不公平的现象。DWRR算法:是基于WRR的一种算法,它分配给每一个队列的权值不是分组的个数,而是分组的比特数,当前轮询剩下的比特数会留到下一轮的轮询中去。基本的调度过程与WRR算法是一样的,但是它提供的带宽分配细度会更加公平,而它的复杂度与WRR一样。DWRR与WRR有同样的缺点,即对于在同一个队列的报文没有优先级的区分,对于这种“微公平”性没有很好的保证,并且对EF业务的低延迟和低抖动也没有很好的保证^[2]。

本文的调度算法主要由流量调节器和DWRR_PQ分组调度器两部分组成。为了避免EF业务流量霸占过多的带宽资源,采用了令牌桶算法作为流量调节器来控制EF业务的流量。同时为了满足EF业务的低延迟抖动的性能还是采用了优先级算法,能够为最高优先级的EF业务提供最佳的服务。但是由PQ算法的缺点可以看出,这样做有可能会使低优先级的业务长时间得不到服务,处于“饥饿”的状态^[3]。因此,分组调度器的算法是通过将DWRR算法和优先级算法相结合来实现的,区分服务模型是将大量的复杂性工作放在边缘路由器中来完成以达到伸缩性强的特点。边缘路由器虽然会降低流量的容量但是并不会减少流的数目,同时使服务基于汇聚流,从而使核心路由器的工作仅限于简单的业务判断并转发汇聚数据流。所以该调度策略适用于边缘路由器的队列调度,而核心路由器的调度策略就是该调度策略的简化,即不会包括令牌桶的流量调节器^[4]。当分组到达区分服务域时,边缘路由器首先会对该分组进行分类和调度。如果该分组被标记为EF业务,那么它就必须经过流量控制的令牌桶过滤,若令牌桶中有令牌则符合流量要求,就会通过PQ调度,被送到高优先级队列进行转发;否则该分组就会被降级处理重新标记,并且同时转到其他队列中去接受DWRR算法的调度。如果到达的分组数据包被标记为其他的非EF业务,那么它首先就要接受DWRR算法的调度,被DWRR调度后再会被发送到PQ算法中的优先级或低优先级队列中进行排队调度等待优先级调度器的调度。在上述算法原理分析

中,最高优先级的EF业务分组总是优先得到服务,但是由于它的流量会受到令牌桶算法的控制,并不会占用整个带宽的资源,因此低优先级业务的队列并不会出现“饥饿”状态。本文调度算法提高了网络的性能,满足了QoS对EF业务在低延迟、低抖动和公平性及对低优先级的业务的资源分配不合理现象^[5]的要求。

DWRR-PQ算法调度原理如图1所示。

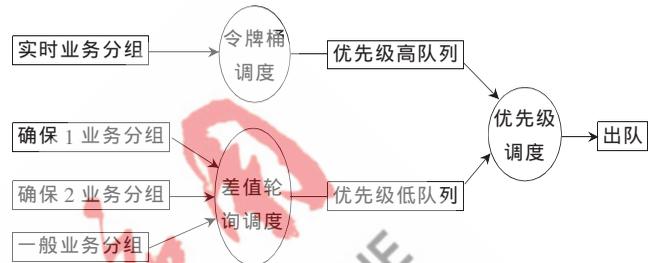


图1 DWRR_PQ算法调度原理图

2 仿真与数据分析

利用网络仿真软件OPNET实现了DWRR_PQ算法,网络拓扑图如图2所示,网络业务配置如表1所示。

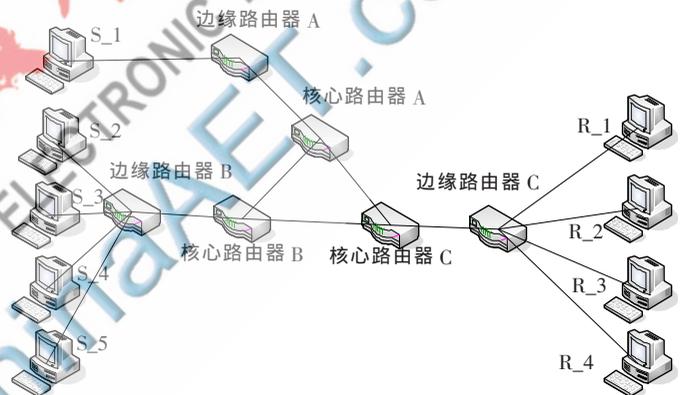


图2 网络拓扑图

表1 网络业务配置表

发送端	接收端	PHB 种类	业务类型
S_1	R_1	EF	视频业务
S_2	R_2	AF1	音频业务
S_3	R_3	AF2	低质量音频业务
S_4	R_4	BE	FTP 业务

2.1 调度算法对EF业务吞吐量的仿真

利用网络仿真软件OPNET实现了该算法。由图3可知,DWRR_PQ调度算法对EF业务的吞吐量与PQ算法非常接近,一般都是在500 packets处上下波动,波动范围较小。而在WRR算法下的EF业务吞吐量波动较大,范围在494 packet~502 packet之间。DWRR_PQ算法对EF业务的吞吐量的支持率比WRR要好。

2.2 调度算法对EF业务延迟抖动的仿真

在网络中传输EF业务时要求其具有低延迟和低抖动的性能。由图4所示可以看出,PQ调度和DWRR_PQ调度的端到端延迟都是在11.6 ms上下波动,而WRR

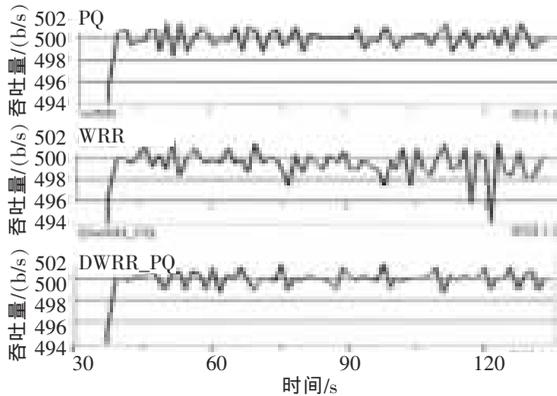


图3 调度算法对EF业务吞吐量仿真

的延迟比较高,在17 ms~18 ms之间波动。由这些数据可知,PQ和DWRR_PQ算法端到端的延迟比WRR算法低7 ms~8 ms左右,性能有很大的提高。同时由图4还可以看出三种调度算法下EF业务通信量的延迟抖动情况。PQ算法和DWRR_PQ算法下,EF业务通信量的延迟变化范围一般是在0.1 ms左右,而WRR算法的延迟的变化大约在0.2 ms。可见DWRR_PQ算法对EF业务的队列延迟和抖动有很好的保证。

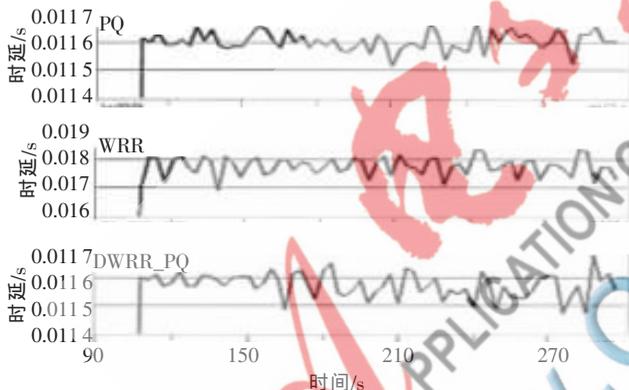


图4 EF通信量的端到端延迟比较

2.3 调度算法对BE业务支持的仿真

由图5可知,PQ调度算法没有数据(只有DWRR_PQ算法和WRR算法有数据),这是由PQ算法的特点决定的,当有BE业务流时,PQ调度器会将该业务分配给最低的优先级,再有高优先级时不会对低优先级的业务进行调度,使低优先级的业务处于“饥饿”状态,采集不到BE的数据,从而使PQ算法的接收方不会接收到BE的数据。同时从图中还可以看出,WRR算法对BE通信量的支持比DWRR_PQ方案还要好。

经过对仿真结果的分析,就EF业务队列的吞吐量、延迟和延迟抖动而言,DWRR_PQ算法的调度性能都要好于WRR算法;与PQ机制相比,DWRR_PQ降低了BE

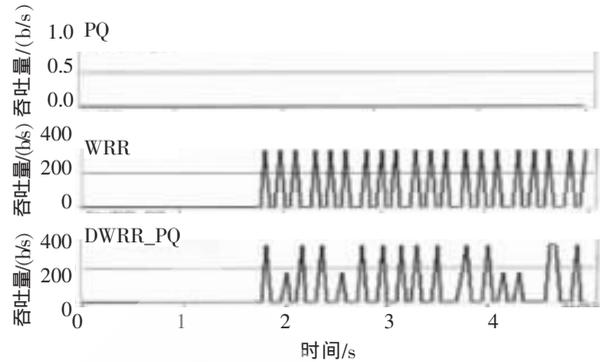


图5 BE通信量的比较

业务有可能处于“饥饿”状态的问题。

本文提出了一种基于区分服务(DiffServ)的拥塞管理的队列调度策略,满足了区分服务业务的EF业务流对网络的性能要求。仿真结果显示了本算法实现了为EF业务提供低延迟、低抖动及对带宽要求的服务,将EF业务和其他业务进行很好的隔离,避免了BE业务可能出现的“饥饿”现象,说明了本文提出的调度策略对区分服务体系的队列调度有较好的支持。

参考文献

- [1] 林闯. 计算机网络的服务质量(QoS)[M]. 北京: 科学出版社, 2004.
- [2] PITSILLIDES A, IAMBERT J. Adaptive congestion control in ATM-based networks: Quality of service and high utilization[J]. Computer, Communication, 1997(20):1239-1258.
- [3] PETERSON L L, DAVIE B S. Computer networks: a system approach(second edition)[M]. 北京: 机械工业出版社, 2002.
- [4] GUFFENS V, BASTIN G. Fluid flow network modeling for hop-by-hop feedback control of a multicast stream[C]. Benelux Meeting, 2003.
- [5] GOYAL P, VIN H M, CHEN H. Start-time fair queuing: a scheduling algorithm for integrated services[J]. IEEE/ACM Transactions on Networking, 1997,5(5):690-704.

(收稿日期:2012-03-09)

作者简介:

曹结宝,男,1986年生,硕士,主要研究方向:通信系统中的可靠性技术。

周井泉,男,1963年生,教授,博士,主要研究方向:通信网络中所涉及的优化、信息管理和控制问题。