

# DiffServ 队列调度算法研究

江文静,蔡祥宝

(南京邮电大学 电子科学与工程学院,江苏 南京 210003)

**摘要:**随着网络技术的发展,用户对网络的服务质量要求越来越高。传统的 IP 网络是针对非实时数据传输而设计,提供“尽力而为”的服务,难以满足不同业务的服务质量要求。研究表明区分服务(DiffServ)模型具有简单、有效、扩展性强等特点。它对业务流进行分类、整形使得不同类型业务数据在网络中得到不同服务。队列调度算法是实现区分服务的核心机制之一。文中分析了 DiffServ 模型下常见的 WRR 和 DWRR 调度算法的优缺点,在 DWRR 算法的基础上提出一种改进的队列调度算法(VDWR),同时将 PRI(优先调度)算法与 VDWR 算法相结合,很好地满足了业务的时延特性。用 NS-2 网络仿真器进行性能评价,结果表明了该算法的有效性和可行性,可以为不同业务提供更好的服务质量保证。

**关键词:**区分服务;队列调度;时延;服务质量

**中图分类号:**TP393

**文献标识码:**A

**文章编号:**1673-629X(2015)04-0085-04

**doi:**10.3969/j.issn.1673-629X.2015.04.020

## Research on Queue Scheduling Algorithm Based on DiffServ

JIANG Wen-jing, CAI Xiang-bao

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

**Abstract:** With the development of network technology, the quality service requirement of network is increasingly improved. Since the traditional IP network is designed for the real-time data transmission and providing “Best effort” service which is difficult to satisfy the QoS requirement. Many studies have shown that Differentiated Service (DiffServ) model provides IP networks with QoS processing, due to its simplicity and scalability. The queue scheduling algorithm is one of the core mechanisms for realizing the DiffServ. It analyzes the advantage and disadvantage between common WRR (Weighted Round Robin) and DWRR (Deficit Weighted Round Robin) scheduling algorithm under the DiffServ model, and based on DWRR algorithm, put forward an improved VDWR (Variable Deficit Weighted Round Robin) queue scheduling algorithm. At the same time, the PRI combined with VDWR algorithm well satisfies the time delay characteristic of the business. With NS-2 network simulator for performance evaluation, the results show the feasibility and effectiveness of the algorithm, and it can provide better QoS guarantee for different business.

**Key words:** DiffServ; queue scheduling; delay; QoS

## 0 引言

随着 IP 技术和各种 Internet 的迅速发展,IP 网络中承载的信息由单纯的文本信息演变到丰富的多媒体业务,出现了诸如 IP 电话、视频通讯会议、远程视频教育等多媒体实时业务。网络业务的不断增多,传统的 IP 网络缺乏服务质量保证(QoS)的缺点也日益突出,因此,IP QoS 成为 IP 技术发展的关键技术<sup>[1]</sup>。服务质量(QoS)就是在网络带宽一定的前提下,通过一定的策略和机制以保证不同业务的需求,具体可以量化为

带宽、抖动、延迟、吞吐量、丢包率等性能指标<sup>[2]</sup>。IETF 制定的 DiffServ 模型就是解决 QoS 的主要技术之一,由于其具有简单性及良好的扩展性,已经成为解决网络服务质量的方案<sup>[3]</sup>。

网络 QoS 的关键在于资源的控制管理,即管理缓冲队列与链路带宽等网络资源的分配和使用<sup>[4]</sup>。队列调度算法是实现区分服务(DiffServ)的核心机制之一,队列调度对链路带宽进行控制管理,使得 IP 报文在路由器中按照一定的策略暂时缓存在等待队列中,

收稿日期:2014-06-13

修回日期:2014-09-18

网络出版时间:2015-02-23

基金项目:国家自然科学基金(11304159);教育部博士点基金(20133223120006);南京邮电大学引进人才项目基金(NY213023)

作者简介:江文静(1987-),女,硕士,研究方向为光通信与光信息处理技术;蔡祥宝,教授,硕士生导师,研究方向为凝聚态物理、非线性光学、光通信与光信息处理。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20150223.1252.047.html>

然后再按照一定的规则选择分组发送。使输入业务能够按照预定方式共享输出链路的带宽<sup>[5]</sup>。

目前有很多队列调度算法,如:FIFO(First In First Out),CQ(Custom Queuing),PQ(Priority Queuing),WFQ(Weighted Fair Queuing),WRR(Weighted Round-Robin),DWRR(Deficit Weighted queuing Round-Robin)。轮询类算法公平性显然不如WFQ,但算法的简单性却在硬件实现上有很大优势,轮询类算法有WRR和DWRR<sup>[6]</sup>。

WRR算法可以确保低优先级业务始终得到一定的服务资源,但由于轮询算法本身的缺点,在分组大小不一的情况下,WRR会对分组较小的队列不公平。为了克服分组变长而带来的不公平,P. Goyal提出了DWRR算法,此算法给每一个队列分配的权值不是基于分组的个数,而是基于比特数<sup>[7]</sup>。这种算法给宽带分配提供了更精确的保证。在DWRR算法中,设置了如下参数:权值 $W$ 和差值计数器DC(每次允许调度器发送的队列总字节数),算法如下<sup>[8]</sup>:

(1)调度器服务当前非空队列,如果此队列的头部等待发送分组长度大于DC值,调度器将移向下一队列而DC值累积到下次轮询;

(2)如果队列头部分组长度小于或者等于DC值,调度器将发送此分组且差值计数器DC值消耗(DC值减去此时分组的长度)。反复发送分组直到等待发送

分组长度大于DC值,此时重复步骤(1);

(3)如果输出分组队列为空而DC值仍有剩余时设置DC值为零,调度器将移向下一队列。

DWRR算法对单个队列连续传输多个分组直到差值器DC的值小于队列分组的字节数或队列为空,比较容易引起较大的传输时延抖动,很难支持实时应用<sup>[9]</sup>。DWRR算法中队列的权值分配和调度次序是固定不变的,容易导致不必要的包丢失<sup>[10]</sup>。

若权值能根据负载情况在一定范围内动态变化,便在一定程度上解决此问题<sup>[11]</sup>。文中在DWRR算法基础上,提出了一种改进型算法VDWRR,结合PRI算法的优缺点设计出新的调度策略。

## 1 改进型队列调度算法VDWRR描述

VDWRR算法思想:为每个队列设置两个参数,分别是延迟指数 $D_i$ 和权值 $W_i$ 。延迟指数 $D_i$ 是队列 $i$ 的长度 $L_i$ 与分配权值 $W_i$ 的比值,即 $D_i = L_i / W_i$ 。每经过一段时间就让调度器重新计算每个队列的 $D_i$ 值,将所有的 $D_i$ 值按升序排列,调度器根据 $D_i$ 值升序排列方式依次服务队列。由于各队列的延迟指数 $D_i$ 是动态变化的,故调度器服务队列的次序也是动态变化的,弥补了DWRR算法的不足。VDWRR算法的示意图如图1所示。

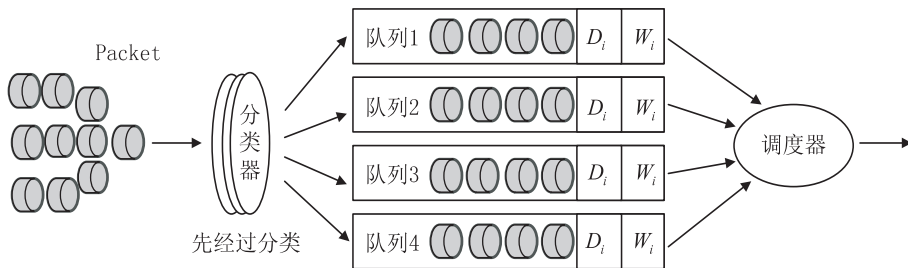


图1 VDWRR算法示意图

数据包经过分类器,被送往不同的队列进行等待,为每个队列设置两个参数 $W_i$ 、 $D_i$ ,假定在某一时刻调度次序是 $1, 2, \dots, N$ ,经过了 $\Delta T$ 时间,调度器会重新计算队列的 $D_i$ 值,根据 $D_i$ 值重新调整队列的服务次序。这样能减少转发数据包的时延抖动。

## 2 PRI+VDWRR算法概述

### 2.1 PRI+VDWRR算法思想

在DiffServ服务模型中,EF(Expedited Forwarding)PHB(每跳行为)提供的是“三低一保证”的虚拟专线服务<sup>[12]</sup>,AF(Assured Forwarding)PHB提供的是最小保证宽带的确保服务,BE(Best-Effort)PHB提供缺省的“尽力而为”服务。在所有的聚流中,EF流应该具有最高的优先级,AF流次之,而BE流则只需要“尽力而

为”的服务<sup>[13]</sup>。

VDWRR是基于DWRR的一种改进型算法,也属于轮询类的算法,不能满足EF流低时延、低抖动的要求,所以选用PRI调度算法。PRI算法属于“饥饿型”算法,当网络拥塞时会导致低优先级的流得不到服务。为了防止恶意的EF流肆意地侵占其他的流,可以通过设置EF流的上限(最高速率、最大突发量等),超过上限的分组一律丢弃。为弥补PRI算法的缺点,文中通过令牌桶实现对EF流流量的限制。对于BE流只需要“尽力而为”的服务,所以不适合使用轮询类的算法,在这里也采用PRI调度。而AF流之间则没有明显的优先级界限,需要宽带公平分配,所以选用前面提到的VDWRR。此调度策略(PRI+VDWRR算法)结合了两个算法的优点,不仅保证高优先级业务提供了实

时服务保证,而且由于高优先级业务流量已经受到限制,故不会占用太多的带宽,同时又保证低优先级队列不会饥饿。采用改进型的 VDWRR 算法保证了 EF 流业务的延时特性,又确保了低权值业务的带宽,实现了带宽分配的相对公平。调度算法整体框图如图 2 所示。

2.2 PRI+VDWRR 调度策略分析

(1)当分组到达时,根据各 IP 分组首部的 DSCP 值自动匹配映射表。如是 EF 业务,则进入 EF 缓冲队

- 列,否则进行下一次业务匹配;
- (2)当 EF、AF 和 BE 业务队列同时存在,EF 优先,再 AF,最后 BE,发送字节按权重分配;
- (3)当 EF 队列为空时,AF 队列由 VDWRR 调度,然后再送往 PRI 调度器调度;
- (4)VDWRR 算法根据权值  $W_i$  和延迟指数  $D_i$  动态调整队列调度次序,以减少 AF 业务的抖动时延;
- (5)进入下一队列,重复以上步骤。

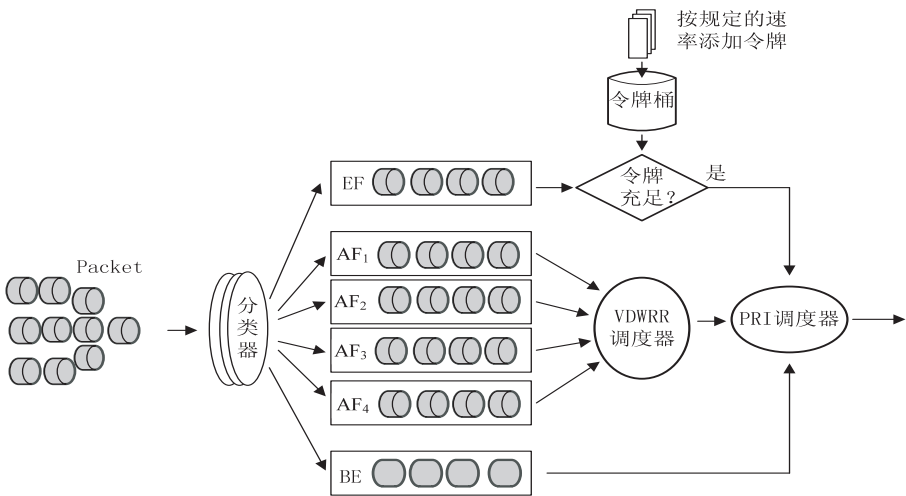


图 2 调度策略整体图

3 仿真实验结果

3.1 实验拓扑结构及主要参数

采用网络仿真器 NS-2<sup>[14]</sup>来验证,并且与 WRR 算法进行比较,由于 NS-2 网络仿真器的区分服务模块中并没有 VDWRR 算法,将改进的算法模块编译到 NS-2 中,实验结构如图 3 所示。

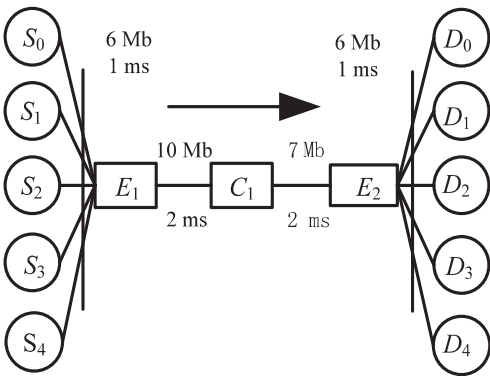


图 3 PRI+VDWRR 算法仿真拓扑结构

图中  $S_0 \sim S_4$  为发送端,  $D_0 \sim D_4$  为对应的接收端;  $E_1$ 、 $C_1$ 、 $E_2$  构成一个 DiffServ 网络区, 其中  $E_1$ 、 $E_2$  为 DiffServ 边缘路由器,  $C_1$  为 DiffServ 核心路由器<sup>[15]</sup>。瓶颈链路位于边缘路由器  $E_1$  和核心路由器  $C_1$  之间。边缘节点  $E_1$  和  $E_2$  需要具有分类、整形和标记的作用, 将各输入 DSCP 值相同的流汇聚,  $S_0$  至  $D_0$  的信息流标记

为 EF 流;将  $S_i$  至  $D_i$  的信息流标记为  $AF_i(i = 1, 2, 3)$  流, 将  $S_4$  至  $D_4$  的流标记为缺省的 BE 流。各链路的带宽和时延在图中已标出, 各源端的流类型和速率如表 1 所示。最小带宽单位为 SLA(Mbps)。

表 1 各流参数设置

DSCP	EF	AF <sub>1</sub>	AF <sub>2</sub>	AF <sub>3</sub>	BE
源端	$S_0$	$S_1$	$S_2$	$S_3$	$S_4$
最小带宽	2.0	1.6	1.2	0.8	0.1

3.2 实验结果及其分析

根据以上的仿真参数,各流的仿真结果如图 4~6 所示。

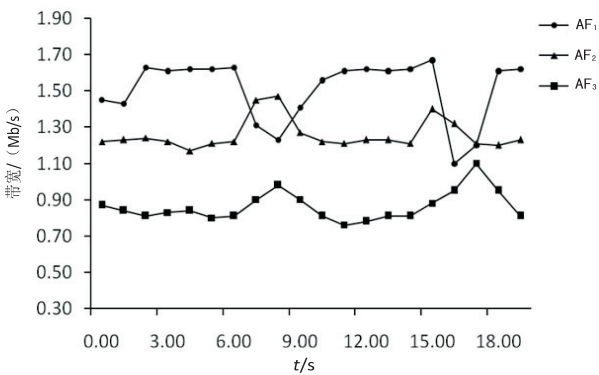


图 4 运用 WRR 算法各流在  $E_1 - E_2$  链路的实际带宽曲线

由图 4 和图 5 可以看出,VDWRR 算法为大小不

同的包分配带宽较 WRR 更公平。VDWRR 为不同的流提供最低的带宽保证,带宽曲线很平稳,所以不会出现较大的时延抖动。

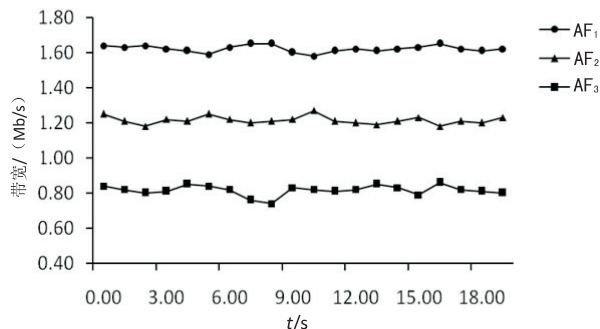


图 5 运用 VDWRR 算法各流在  $E_1 - E_2$  链路的实际带宽曲线

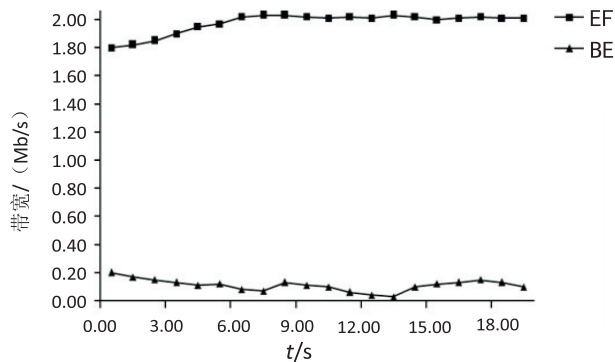


图 6 BE 流和 EF 流的带宽

由图 6 可以看出 PRI+VDWRR 调度算法结合了两个算法的优点,不仅保证高优先级业务得到实时服务,而且由于 EF 实时业务流量已经受到限制不会占用太多带宽,又保证低优先级队列不会饥饿,弥补了 PRI 算法的缺陷。

## 4 结束语

文中在 DWRR 队列调度算法的基础上提出了改进型的算法 VDWRR,该算法根据延迟指数的动态变化保证了业务低时延,提高了网络的服务质量。由于轮询类的算法 WRR 和 DWRR 对实时业务支持方面的性能较差,目前的路由大多数是将 EF 实时业务独立进入低延迟队列,提供绝对的优先级。而其他则采用轮询调度算法,这样增加了算法的复杂性,从而算法实现的难度也相应加大。因此下一步的工作是如何对现

有算法进行优化,以满足不同业务的需求,最终提高网络的服务质量。

## 参考文献:

- [1] 林 闯. 多媒体信息网络 QoS 的控制[J]. 软件学报, 1999, 10(10): 1016-1024.
- [2] 林 闯. 计算机网络的服务质量(QoS)[M]. 北京: 清华大学出版社, 2004.
- [3] Blake S, Blacke D, Carlson M, et al. An architecture for different services[S]. RFC 2475, 1998.
- [4] 赵 晨, 周润芳, 许小刚, 等. 集成服务网络中 QoS 路由的讨论[J]. 计算机工程与应用, 2003, 39(9): 176-178.
- [5] 张继军, 高 鹏. 基于分组网络的服务质量保证[M]. 北京: 北京邮电大学出版社, 2004.
- [6] Parekh A K, Gallagher R G. A generalized processor sharing approach to flow control in integrated services networks; the multiple node case[J]. IEEE/ACM Trans on Networking, 1994, 2(2): 137-150.
- [7] Ghyal P, Start H M. Start-time fair queuing: a scheduling algorithm for integrated services packet switching networks[J]. IEEE/ACM Trans on Networking, 1997, 5(5): 690-703.
- [8] 伍金富, 周井泉. 基于区分服务的队列调度算法研究[J]. 计算机技术与发展, 2011, 21(1): 140-142.
- [9] Dekeris B, Budnikas A. Analysis of QoS assurance using Weighted Fair Queuing (WFQ) scheduling discipling with Low Latency Queue (LLQ)[C]//Proc of 28th international conference on technology interfaces. [s. l.]: [s. n.], 2006: 507-512.
- [10] Li Miaoyan, Song Bo. Design and implementation of a new queue scheduling algorithm in DiffServ networks[C]//Proc of 2nd international conference on advanced computer control. Shenyang: IEEE, 2010: 117-122.
- [11] 李双庆, 许 峰. 支持区分服务的自适应队列调度管理[J]. 计算机应用, 2007, 27(B06): 170-171.
- [12] Shreedhar M, Varghese G. Effieient fair queueing using deficit round robin[J]. IEEE/ACM Transactions on Networking, 1996, 4(3): 375-385.
- [13] 潘迟龙, 张宏宏. 基于 WRR 的改进型队列调度算法[J]. 电信快报, 2012(1): 34-37.
- [14] 柯志亨. NS2 教学手册[DB/OL]. 2006. <http://140.116.72.80/~smallko/ns22/ns2.htm>.
- [15] 徐雷鸣, 庞 博, 赵 耀. NS 与网络模拟[M]. 北京: 人民邮电出版社, 2003.

DiffServ队列调度算法研究

作者：[江文静](#)，[蔡祥宝](#)，[JIANG Wen-jing](#)，[CAI Xiang-bao](#)

作者单位：[南京邮电大学 电子科学与工程学院](#)，[江苏 南京](#)，[210003](#)

刊名：[计算机技术与发展](#)

英文刊名：[Computer Technology and Development](#)

年，卷(期)：[2015\(4\)](#)

引用本文格式：[江文静](#)，[蔡祥宝](#)，[JIANG Wen-jing](#)，[CAI Xiang-bao](#) [DiffServ队列调度算法研究](#)[期刊论文]-[计算机技术与发展](#) 2015(4)