

# DiffServ 中基于优先级的队列调度算法

李娟, 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:**为了提高 DWRR 算法的时延性能及应对突发流的能力, 文中结合比例时延区分服务模型, 在 DWRR 算法的基础上提出了一种能够体现优先级的队列调度算法-PDDRR。该算法的控制目标是保证各业务流的平均排队时延在无突发业务流时维持在给定的比率, 而在应对突发流时忽略时延的比例公平原则, 根据网络实时负载动态地调整各队列的权值, 合理地分配资源。仿真结果表明, 算法不仅能实现不同优先级业务的时延区分, 而且在业务流突发时为各等级业务提供服务质量保证。

**关键词:**比例时延; 优先级; 业务流突发

**中图分类号:** TP301.6

**文献标识码:** A

**文章编号:** 1673-629X(2012)07-0091-04

## Queue Scheduling Algorithm Based on Priority in DiffServ

LI Juan, ZHOU Jing-quan

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

**Abstract:** In this paper, a queue scheduling algorithm based DWRR is proposed to improve the delay performance and the capability to handle burst traffic, which combines the proportional delay differentiated model. The algorithm aims to keep the average queueing delay of all the traffic in a given ratio, ignore the proportional fairness principle under bursty environment, and adjust all the weights according to the real-time load status to distribute resources dynamically. Simulation results demonstrate PDDRR not only merely realizes the delay differentiated between varied PRI services, but also ensures the quality of service under burst traffic environment.

**Key words:** proportional delay; priority; burst traffic

## 0 引言

随着 Internet 的迅速发展, 人们不仅需要现有的网络能够传输各式的数据还要求对业务提供一定的 QoS 保证。IETF 制定的区分服务 (DiffServ) 模型由于具有简单性及良好的扩展性已经成为解决 QoS 问题的首选方案, 而相对区分已经成为 DiffServ 模型发展的主流, 除了能够保证各业务服务质量之间的相对区分, 其网络管理部署等方面都较为简单, 比例区分服务模型就是其中的一种<sup>[1,2]</sup>。DiffServ 模型下中间节点如何在保证服务质量的前提下对分组进行调度已成为当前研究的热点。

目前已经提出了许多支持区分服务的分组调度算法, 如 WFQ 队列、WRR 队列、DWRR 队列、PAD 等, 其中 WRR 及 DWRR 是轮询类的算法, PAD 是针对比例

区分服务模型研究出的算法, 主要是保证任意两个业务流的平均时延控制在给定的比例<sup>[3]</sup>。DWRR 算法虽然操作简单, 但其时延性能及应对突发流的能力较差<sup>[4]</sup>, PAD 算法虽然能够实现业务流间的比例时延特性, 但在较短的时间间隔内却有不可预测的时延区分。所以文中将结合比例区分服务模型, 在 DWRR 算法的基础上提出了能够体现优先级的自适应队列调度算法-PDDRR, 该算法在保证轮询类算法简单性的同时提供比例时延特性, 且改善了应对突发流的能力。

## 1 比例区分服务模型

比例区分服务模型是一种典型的相对区分服务模型, 它能够在不同的时间段内保证各队列服务间的差别。若  $\bar{q}_i(t, t + \tau)$ ,  $\bar{q}_j(t, t + \tau)$  ( $\tau > 0$ ) 代表队列  $i$  和队列  $j$  的性能量值,  $s_i, s_j$  ( $i, j = 1, 2, \dots, N$ ) 表示各队列的 QoS 区分参数, 那么比例公平原则为每个队列作如下限制, 即:

$$\bar{q}_i(t, t + \tau) / \bar{q}_j(t, t + \tau) = s_i / s_j \quad i, j = 1, 2, \dots, N \quad (1)$$

收稿日期: 2011-12-13; 修回日期: 2012-03-20

基金项目: 江苏省普通高校研究生科研创新基金 (CXLX11\_0400)

作者简介: 李娟 (1988-), 女, 安徽淮北人, 硕士, 研究方向为网络的可靠性技术; 周井泉, 教授, 博士, 研究方向为通信网络的可靠性技术。

若上式中  $\bar{q}_i(t, t + \tau)$  用一个时延参数  $\bar{d}_i(t, t + \tau)$  代替,  $s_i$  用时延区分参数  $\delta_i$  替代, 便衍生出一种比例时延区分服务模型<sup>[5,6]</sup> (PDD, Proportional Delay Differentiation)。PAD (Proportional Average Delay) 是一个能够实现比例时延区分的调度算法, 该算法采用归一化的队列时延来描述 PDD 模型, 定义  $\bar{d}_i = \bar{d}_i / \delta_i$ , 每次选择具有最大  $\bar{d}_i$  的业务流进行调度, 由于没有考虑队列中积压分组的时延情况, 导致在较短的时间尺度内性能较差, 然而对于 PDD 模型, 一个好的调度算法应该保证在不同的时间间隔内提供可预测的时延区分<sup>[7]</sup>。

## 2 PDDRR 算法

### 2.1 PDDRR 基本原理

将缓存分为  $N$  个不同优先级的虚拟队列  $Q_1, Q_2, \dots, Q_N$ , 且优先级依次升高, 第  $i$  个队列的时延区分参数为  $\delta_i$  ( $0 < i < N + 1$ )。对于队列调度算法而言, 平均排队时延及平均分组丢失率是评价其性能的重要参数<sup>[8,9]</sup>, 所以文中的控制目标是在各队列的平均分组到达率  $\lambda_i$  ( $0 < i < N + 1$ ) 小于阈值的情况下保证队列间的平均排队时延的比例公平性, 而当业务流突发需要丢弃分组时将队列的平均丢包率为评价标准, 在这两种情况下将通过不同的策略动态的调整各队列轮询的权值  $\omega_i$  及定额值  $Q_i$  ( $0 < i < N + 1$ ), 确定好服务速率后再采用差额轮询对各队列进行调度。

### 2.2 PDDRR 算法结构

#### 2.2.1 比例时延公平原则

假设网络中提供  $N$  类业务, 且各业务类的平均分组到达率服从  $\lambda_i$  ( $i = 1, 2, \dots, N$ ) 的泊松分布, 那么根据 Little 公式  $\bar{L}_i = d_i \lambda_i$ <sup>[10]</sup>, 可得到第  $i$  类业务的平均分组排队时延  $d_i$ , 如式(2):

$$d_i = \bar{L}_i / \lambda_i, \quad i = 1, 2, \dots, N \quad (2)$$

式中  $\bar{L}_i$  为各队列的平均队列长度, 根据比例时延区分服务模型的基本要求, 各队列的平均排队时延应满足式(3):

$$d_1 : d_2 : \dots : d_N = \delta_1 : \delta_2 : \dots : \delta_N \quad (3)$$

其中,  $\delta_1, \delta_2, \dots, \delta_N$  为各类业务流的时延区分参数, 结合上面两式可得:

$$\bar{L}_1 : \bar{L}_2 : \dots : \bar{L}_N = \lambda_1 \delta_1 : \lambda_2 \delta_2 : \dots : \lambda_N \delta_N \quad (4)$$

将系统时间划分为大小为  $T$  的等距的时间段, 令  $L_{ij}$  为队列  $i$  在第  $j$  个时间段内的平均队列长度, 如果系统满足在每个时间段内队列长度成比例的原则, 即

$$L_{1j} : L_{2j} : \dots : L_{Nj} = \lambda_1 \delta_1 : \lambda_2 \delta_2 : \dots : \lambda_N \delta_N \quad j = 1, 2, \dots \quad (5)$$

若式(5)成立, 那么式(4)必然成立。又因为

$$L_{ij} = L_{i(j-1)} + \lambda_i T - Q_i, \quad i = 1, 2, \dots, N, j = 1, 2, \dots \quad (6)$$

上式中  $Q_i$  ( $i = 1, 2, \dots, N$ ) 为系统在时间段  $T$  内为队列  $i$  ( $i = 1, 2, \dots, N$ ) 服务的字节数, 能够反映差额轮询调度中各队列分配的量子值情况<sup>[11]</sup>, 所以文中将其归一化值视为各队列的服务权值  $\omega_i$  ( $i = 1, 2, \dots, N$ )。同时令  $B$  为系统在一次服务时间段  $T$  内最大可服务的字节数, 那么根据式(5)、式(6)及系统的服务能力可得到:

$$(L_{1(j-1)} + \lambda_1 T - Q_1) : \dots : (L_{N(j-1)} + \lambda_N T - Q_N) = \lambda_1 \delta_1 : \dots : \lambda_N \delta_N \quad j = 1, 2, \dots \quad (7)$$

$$Q_1 + Q_2 + \dots + Q_N = B \quad (8)$$

结合上面两式便可得到系统在时间间隔  $T$  内为队列  $i$  ( $i = 1, 2, \dots, N$ ) 服务的字节数  $Q_i$  ( $i = 1, 2, \dots, N$ ), 从而得到差额轮询调度中各队列的权值  $\omega_i$  ( $i = 1, 2, \dots, N$ )。通过上文的分析可知, 当监测到各业务流的平均分组到达率  $\lambda_i$  没有超过给定的门限值  $\lambda_{minhi}$ , 便可根据式(7)及(8)调整队列的服务权值  $\omega_i$  ( $i = 1, 2, \dots, N$ ), 从而保证队列间平均排队时延的比例公平性。

#### 2.2.2 业务流突发情况下权值的自适应调整

若网络环境中无业务量突发, PDDRR 算法则根据上节中提出的比例公平原则为各优先级队列调整服务权值, 而当网络中有突发流量时, 将会导致该队列的长度骤然变长, 此时网络便需要根据一定的原则丢弃分组, 若仍然按照比例时延公平策略不仅不能保证各服务队列平均时延的比例公平性, 还会产生较高的分组丢失率, 所以 PDDRR 算法中的自适应权重机制能够有效地解决突发流存在的问题。

假设系统检测到某优先级业务流  $i$  的分组到达率  $\lambda_i$  超过门限值  $\lambda_{minhi}$ , 此时自适应权值机制启动, 权值调整模块便根据测量器的测量结果进行相应的处理:

(a) 当  $\lambda_{minhi} < \lambda_i < \lambda_{maxhi}$  时,  $\Delta\omega_i$  则按照速率的大小进行线性的变化:

$$\beta = (\lambda_i - \lambda_{minhi}) / (\lambda_{maxhi} - \lambda_{minhi}) \quad (9)$$

$$\Delta\omega_i = \beta \cdot \Delta\omega_{maxi} \quad (10)$$

其中  $\beta$  为权值调节因子,  $\Delta\omega_i$  是队列  $i$  权值增量,  $\Delta\omega_{maxi}$  指队列  $i$  权值增量的最大值,  $\lambda_{maxhi}$  指队列  $i$  的平均分组到达率上限阈值。

(b) 当  $\lambda_i > \lambda_{maxhi}$  时, 令调节因子  $\beta = 1$ , 有  $\Delta\omega_i = \beta \cdot \Delta\omega_{maxi} = \Delta\omega_{maxi}$ ;

(c)  $\omega_i = I\omega_i + \Delta\omega_i$ ,  $I\omega_i$  为业务流突发前的权值, 确定各队列服务权值后, 依据式(11)便可得到在差额轮询调度中需要的量子值  $Q_i$ , 式中  $P_{mu}$  指传送单元的最大分组长度:

$$Q_i = \omega_i \cdot P_{mu} \quad (11)$$

为了保证其他优先级业务流的公平性, 本算法在

增加突发流权值的同时,会将所有队列的权值按照相同的比例 $\beta$ 增加,虽然 $\beta$ 相同,但各优先级队列的 $\Delta\omega_{\max i}$ 不同,并要求 $\Delta\omega_{\max 1} < \dots < \Delta\omega_{\max N-1} < \Delta\omega_{\max N}$ ,所以增加的权值 $\Delta\omega_i$ 也会有差异,在保证各优先级业务流相对公平性的基础上,也实现了区分服务。

### 3 性能仿真

#### 3.1 仿真拓扑结构

文中将基于不同的目的用 NS2 进行仿真实验<sup>[12]</sup>,网络拓扑结构如图1所示。图中 S3 到 S1 发送的业务流优先级依次降低,瓶颈链路位于节点 CR1 和 ER2 之间,带宽容量为 6Mbps,其它链路的带宽为 10Mbps,且所有链路的传输时延均为 2ms。实验中三个业务流的平均分组到达率阈值 $\lambda_{\min i}$ 及 $\lambda_{\max i}$ ( $i=1,2,3$ )均为 3.3Mbps 和 5Mbps,权值的最大调整范围分别为 $\Delta\omega_{\max 1}=0.5$ , $\Delta\omega_{\max 2}=0.8$ , $\Delta\omega_{\max 3}=1$ ,且三个业务流具有相同的分组大小分布(20%的分组长度 40bytes,30%的分组长度 550bytes,50%的分组长度 1500bytes)。

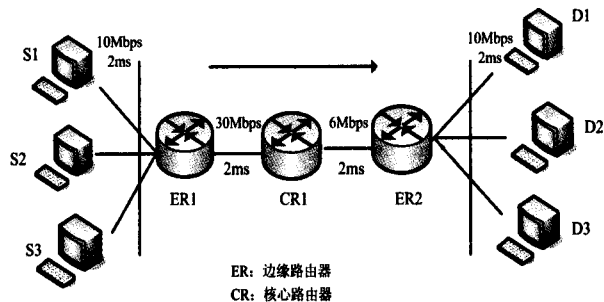


图1 仿真拓扑结构图

#### 3.2 实验过程及其结果分析

实验一:验证算法在合适的区分参数配置下各队列的平均排队时延的比例区分。三个业务流具有相同的平均分组到达率 2.2Mbps,时延区分参数为 $\delta_1=4$ , $\delta_2=2$ , $\delta_3=1$ ,为了保证实验的过程中无分组丢失,将各队列的缓冲大小设为 10Mb。为了更好地比较 PDDRR 算法的比例时延性能,在此网络环境下对 PAD 算法进行了同样的仿真,结果如图2所示。从图中可以

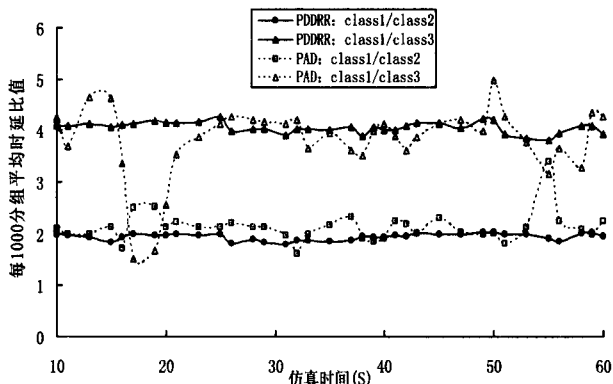


图2 PDDRR及PAD算法的平均时延比值

看出 PDDRR 算法能够在短时间间隔内实现业务流间的比例时延区分,虽然不能严格地满足 4:2:1 的关系,但较小的偏离影响并不大,而 PAD 算法在有的时间周期内 $\overline{d_1}/\overline{d_2}$ 比 $\overline{d_1}/\overline{d_3}$ 的比值要大,这说明流3比流2要经历更多的延时,这是一种不可预测的时延区分,是任何一个保证质量的网络不能接受的。

实验二:验证 PDDRR 算法在不同负载分布比例情况下对各业务流平均排队时延比值的影响。仿真基本条件如实验一,只是将三个业务流的负载分布分为 8 种比例情况,如图3横坐标所示,仍然对 PAD 算法进行了相同的测试。从图中可以看出 PDDRR 的相邻业务流间的平均时延比值能够在所有的负载分布情况下比 PAD 更靠近理想值( $\delta_1/\delta_2=\delta_2/\delta_3=2$ ),且每种情况下的时延比值差别很小。

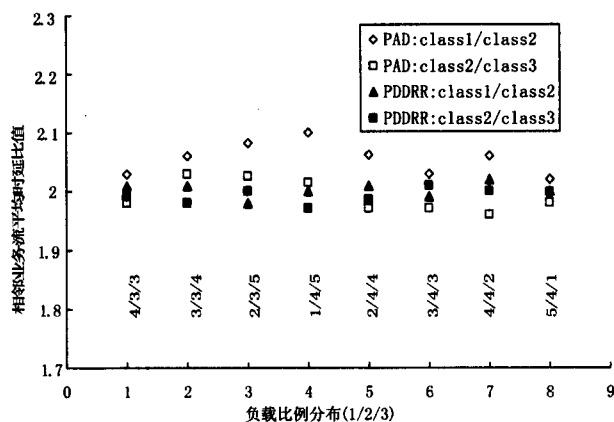


图3 不同负载分布下的平均排队时延比值

实验三:验证算法在不同的链路负载情况下各业务流的平均时延比值情况。基本实验条件如实验一,三个业务流的发送速率从 2.3Mbps 变化到 3.3Mbps,仿真结果如图4所示。从图中可以看出,随着各业务流平均分组到达率的增加,两个调度算法业务流间的平均时延比值都比较接近于期望值,所以算法能够保证不同超载程度下平均排队时延的比例公平性。

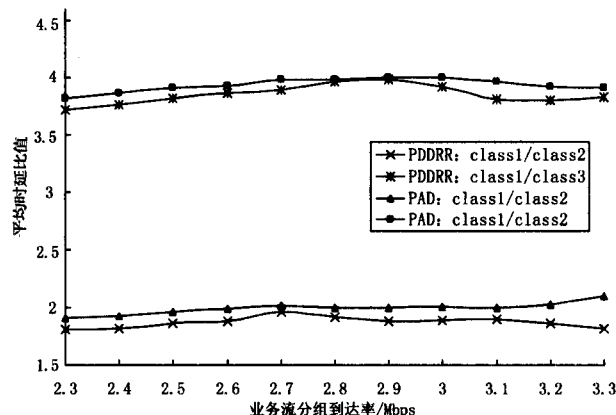


图4 不同链路负载下的平均排队时延比值

实验四:验证算法在业务流突发的情况下各队列的平均分组丢失率。基本条件如实验一,业务流1及

流 3 的平均分组到达率为 2.2Mbps, 业务流 2 的分组到达率由 2.2Mbps 增加到 4Mbps, 缓存空间设置为 1Mb, 三个队列的初始权值设为  $\omega_1 = 1, \omega_2 = 2, \omega_3 = 4$ 。PDDRR 及 DWRR 的实验结果如表 1 所示。从表中的数据可以看出 PDDRR 算法明显地改善了 DWRR 算法应对突发流的能力, 且低优先级业务的突发不会影响高等级业务的服务质量。

表 1 PDDRR 及 DWRR 丢包率的对比

Class	3	2	1
DWRR 丢包率	0.0568	0.3403	0.4391
PDDRR 丢包率	0.0399	0.2716	0.4187
降低百分比	29.75%	20.19%	4.65%

#### 4 结束语

文中提出的 PDDRR 算法主要基于轮询调度及比例时延区分服务模型, 对 DWRR 算法进行了很大的改进, 根据负载的状态动态地调整各队列的服务权值, 不仅提高了 DWRR 算法的时延性能, 同时也改善了应对突发流的能力。实验表明, 该算法在各队列平均分组到达率不超过给定阈值时保证队列的比例时延区分; 当存在业务突发时能够有效地减少突发流对高等级业务的影响, 提高了网络的整体性能。

#### 参考文献:

- [1] Dovrolis C, Stiliadis D, Ramanathan P. Proportional Differentiated Services: Delay Differentiation and Packet Scheduling [J]. IEEE/ACM Transactions on Networking, 2002, 10(1): 12-26.
  - [2] Ippoliti D, Zhou Xiaobo, Zhang Liqiang. Packet Scheduling with Buffer Management for Fair Bandwidth Sharing and Delay Differentiation [C]//International Conference on Computer Communications and Networks. [s. l.]: [s. n.], 2007: 569-574.
  - [3] Leung M K H, Lui J C S, Yau D K Y. Adaptive Proportional Delay Differentiated Services: Characterization and Performance Evaluation [J]. IEEE/ACM Transactions on Networking, 2001, 9(6): 801-807.
  - [4] Ji Liang, Arvanitis T N, Woolley S I. Fair Weighted Round Robin Scheduling Scheme for DiffServ Networks [J]. Electronics Letters, 2003, 39(3): 333-335.
  - [5] Dovrolis C, Ramanathan P. A Case for Relative Differentiated Services and the Proportional Differentiation Model [J]. IEEE Network, 1999, 13(5): 26-34.
  - [6] Lai Yuancheng. Packet Schedulers to Provide Proportional Delay Differentiation and Reduce Packet Queueing Delay Simultaneously [C]//IEEE International Conference on Communications. [s. l.]: [s. n.], 2004: 1968-1972.
  - [7] 孙 祺, 高振明, 李 骐. 比例区分服务模型下的绝对时延保证 [J]. 计算机应用, 2004, 24(9): 70-73.
  - [8] 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(S1): 170-171.
  - [9] 王胜灵, 侯义斌, 黄建辉. PFWR: 能实现比例公平的增强型 WRR [J]. 计算机学报, 2008, 31(5): 775-782.
  - [10] Wei Jianbin, Xu Chengzhong, Zhou Xiaobo. A Robust Packet Scheduling Algorithm for Proportional Delay Differentiation Services [C]//Global Telecommunication Conference. [s. l.]: [s. n.], 2004: 697-701.
  - [11] Venkataramanan V J, Lin Xiaojun. On Wireless Scheduling Algorithms for Minimizing the Queue-overflow Probability [J]. IEEE/ACM Transactions on Networking, 2010, 18(3): 788-801.
  - [12] 宋 玲, 刘勃兰. NS2 中添加路由协议的研究与实现 [J]. 通讯和计算机, 2006, 3(10): 33-37.
- 
- (上接第 90 页)
- [5] Schmid C, Mohr G, Bauckhage C. Evaluation of Interest Point Detectors [J]. International Journal of Computer Vision, 2000, 37(2): 151-172.
  - [6] Koenderink J J. The structure of images [J]. Biological Cybernetics, 1984, 50(5): 363-370.
  - [7] Zhang Bin, He Saixian. Improved Edge-detection Method Based on Canny Algorithm [J]. Infrared Technology, 2006, 28(3): 165-169.
  - [8] Kuglin C, Hines D. The phase correlation image alignment method [C]//Proc of IEEE Conference on Cybernetics and Society. New York: [s. n.], 1975.
  - [9] de Castro E, Morandi C. Registration of translated and rotated images using finite Fourier transforms [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1987, 9(5): 700-703.
  - [10] Reddy B S, Chatterji B N. An FFT-based technique for translation, rotation and scale-invariant image registration [J]. IEEE Transactions on Image Processing, 1996, 8(5): 1266-1271.
  - [11] Lowe D G. Distinctive Image Features from Scale-invariant Key Point [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
  - [12] Harris C, Stephens M. A combined corner and edge detector [C]//Proceedings of the 4th Alvey Vision Conference. [s. l.]: [s. n.], 1988.