

# 基于自相似流量水平分级预测的网络队列调度算法

魏德宾<sup>1,2,3</sup>, 沈婷<sup>2,3</sup>, 杨力<sup>2,3</sup>, 戚耀文<sup>1</sup>

(1. 南京理工大学自动化学院, 江苏 南京 210094; 2. 大连大学信息工程学院, 辽宁 大连 116622;  
3. 大连大学通信与网络重点实验室, 辽宁 大连 116622)

**摘要:** 网络流量的自相似特性会导致网络中数据的突发状态持续, 为有效降低网络流量突发引起的队列排队时延和分组丢失率, 提高不同优先级业务的传输能力, 保障业务服务质量需求, 提出了一种基于网络流量自相似特性的队列调度算法——P-DWRR。该算法设计了基于自相似流量水平分级预测结果的动态权重分配方法及服务量子更新方法, 并根据业务优先级和队列等待时间确定队列的服务次序, 以减小数据分组排队时延, 降低分组丢失率。仿真结果表明, P-DWRR 算法在满足网络不同业务优先级要求的基础上, 降低了数据分组的排队时延、时延抖动和分组丢失率, 性能优于 DWRR 和 VDWR。

**关键词:** 网络流量; 自相似性; 分级预测; 队列调度; 轮询

**中图分类号:** TP393

**文献标识码:** A

**doi:**10.11959/j.issn.1000-436x.2020066

## Network queue scheduling algorithm based on self-similar traffic level grading prediction

WEI Debin<sup>1,2,3</sup>, SHEN Ting<sup>2,3</sup>, YANG Li<sup>2,3</sup>, QI Yaowen<sup>1</sup>

1. School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China  
2. School of Information Engineering, Dalian University, Dalian 116622, China  
3. Communication and Network Laboratory, Dalian University, Dalian 116622, China

**Abstract:** Self-similarity characteristic of network traffic will lead to the continuous burstness of data in the network. In order to effectively reduce the queue delay and packet loss rate caused by network traffic burst, improve the transmission capacity of different priority services, and guarantee the service quality requirements, a queue scheduling algorithm P-DWRR based on the self-similarity of network traffic was proposed. A dynamic weight allocation method and a service quantum update method based on the self-similar traffic level grading prediction results were designed, and the service order of the queue according was determined to the service priority and queue waiting time, so as to reduce the queuing delay and packet loss rate. The simulation results show that the P-DWRR algorithm can reduce the queuing delay, delay jitter and packet loss rate on the basis of satisfying the different service priority requirements of the network, and its performance is better than that of DWRR and VDWR.

**Key words:** network traffic, self-similarity, grading prediction, queue scheduling, round robin

### 1 引言

随着通信技术的发展, 网络中业务种类和业务量都在迅速增加, 良好的服务质量需要较高的链路

带宽来保证。虽然人们可以通过提高硬件性能来增加网络带宽, 但是仍然难以满足日益增长的用户需求和避免网络中某些路由或交换节点的拥塞。当拥塞发生时, 如果没有有效的队列管理和调度方法,

收稿日期: 2019-10-08; 修回日期: 2020-03-18

通信作者: 沈婷, shenting0207@126.com

基金项目: 国家自然科学基金资助项目 (No.61722105, No.61931004)

**Foundation Item:** The National Natural Science Foundation of China (No.61722105, No.61931004)

大量的数据分组由于不能及时转发，积压在路由器的缓冲区中。极端情况下会导致缓冲区溢出，丢失分组，网络将无法为业务提供保障。

队列调度是指路由器以数据流的相关信息为依据，按照某种规则从队列中选择待转发的分组，从而为数据流提供公平或有差别的服务。目前，队列调度算法主要有3种，分别是简单队列调度算法、基于时间戳的调度算法和基于轮询的调度算法。

简单队列调度算法包括先来先服务调度算法、随机调度算法和基于优先级的调度算法等。先来先服务调度算法按照分组到达顺序确定分组服务顺序，实现简单，管理方便，且最大时延可由队长决定，但其不能为高优先级的分组提供服务质量(QoS, quality of service)保障，无法支持区分服务。随机调度算法是在所有等待队列中随机选择转发分组，该方法可在某种程度上满足统计意义的性能保证，但不能满足确定性的时延保证。基于优先级的调度算法虽然能为高优先级分组提供QoS保障，但只有当高优先级队列都为空时，低优先级队列分组才会被调度，这会导致低优先级队列的“饿死”现象。

基于时间戳的调度算法通过对分组记录开始服务时间和结束服务时间进行排序，选择具有最小服务时间的分组进行调度。这类算法主要是对Parekh等<sup>[1]</sup>提出的最理想的队列调度算法模型——通用处理器共享(GPS, generalized processor sharing)的近似模拟，如加权公平队列(WFQ, weighted fair queuing)、最坏情况加权公平队列(W<sup>2</sup>FQ, worst-case weighted fair queuing)和开始时间公平队列(STFQ, start time fair queuing)等，具有良好的时延性能和公平性，但时间复杂度高，实现困难。

基于轮询的调度算法是指调度器轮询地对每个队列中的分组进行调度，一次调度发送一个分组，不考虑业务的优先级和处理能力，使不同队列平等地使用带宽。该算法实现简单、复杂度低，适合高速分组网络，但不能解决业务不同优先级的需求和变长分组带来的不公平性。因此，研究人员提出了一系列改进算法，如加权轮询(WRR, weighted round robin)、差额轮询(DRR, deficit round robin)、差额加权轮询(DWRR, deficit weighted round robin)<sup>[2]</sup>等。文献[3]依据各队列的平均分组到达率，调整各队列的调度权值，提出了PFWR (proportion fairness WRR)算法。文献[4]提出了逐次最小权值轮询(SMRR, successive minimal-weight round

robin)调度算法，它能保证在每个轮次中为每个活动数据流提供与本轮次中的最小权值相当的服务机会。文献[5]综合考虑网络中分组长度及队列权重，提出了一种改进型WRR(EWRR, enhanced WRR)算法。文献[6]针对传统WRR算法权值分配和调度次序固定不变无法适应网络负载加重带来的时延增大问题，提出了可变差额加权轮询(VDWRR, variable deficit WRR)调度算法。文献[7]针对DWRR在考虑分组截止时间时，对即将过期的分组简单转发或直接丢弃，而不是防止违反最后期限的问题，提出了一种破产差额加权轮询(I-DWRR, insolvency-deficit WRR)调度算法。文献[8]针对传统WRR带宽利用不充分的问题，提出了一种负差额加权轮询(N-DWRR, negative-deficit WRR)算法。文献[9]针对大型数据中心数千个虚拟机之间实现负载平衡的问题，提出了一种基于神经网络的动态负载平衡算法。

上述的调度方法虽然各有优势，但它们缺乏对网络流量特性的考虑。已有的研究表明，局域网、广域网、万维网等不同通信网络的实际网络流量都具有自相似性<sup>[10-12]</sup>。网络业务流自相似性的发现和推翻之前网络流量短相关的基础假设。由于自相似性网络流量的突发性更强，持续时间更长，需要更大的网络资源和带宽，容易导致网络路由或交换节点发生拥塞，这使网络流量的统计特征提取、排队性能分析和缓冲区设置等均有所变化，同时也给网络交换节点的队列管理和调度带来挑战。

通过文献检索发现，针对上述问题，国内外的研究(如文献[13-15])都是基于网络流量自相似性对主动队列管理算法的影响展开的，缺少基于网络流量自相似性的队列调度算法研究。为此，本文综合考虑网络流量自相似性对网络性能的影响和不同数据业务的传输需求，在传统差额加权轮询调度算法的基础上，根据自相似流量水平分级预测结果，动态分配权值和更新服务量子，并根据业务优先级和队列等待时间对队列进行排序，完成调度，从而达到减小队列时延、降低分组丢失率的目的。

## 2 DWRR 队列调度算法

队列管理与调度算法基本原理如图1所示。队列管理机制一般位于队列输入端，依靠网络节点主动感知缓冲区的占用率来管理缓存，在网络发生拥

塞时通过分组丢失管理队列长度。队列调度机制则在队列的输出端，按规则决定下一次要发送的分组，管理各流之间的带宽分配。

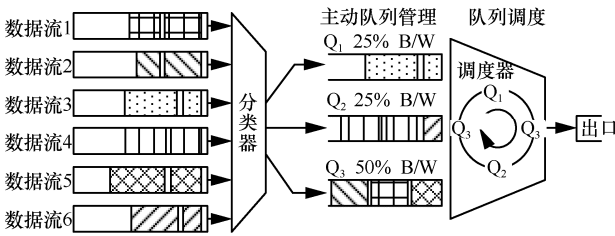


图 1 队列管理与调度算法基本原理

DWRR 算法为每一个队列分配的权值是基于字节数的，其中主要参数定义如下。

- 1) 权值。分配给各队列的输出端口带宽的比例。
- 2) 差值计数器 DC。在某一服务周期内，每个队列在每次进行调度服务时允许队列传输的字节数。
- 3) 服务量子  $q$ 。用字节数表示，正比于其队列权值。调度器对某队列调度服务结束后，在下次轮询到此队列时，将此队列差值计数器的值增加服务量子，为此队列调度服务做准备。

DWRR 算法针对网络分组大小可变的情况，调度器依次服务当前非空队列，基本调度过程如图 2 所示，其中阴影部分为已经调度出去的分组。如果此队列首部等待发送分组长度小于或等于 DC 值，则发送此分组，在差值计数器中减掉相应的字节数，并反复发送分组，直到此队列首部等待发送分组长度大于 DC 值，调度器将移向下一队列，此时剩下的 DC 值累积到下次轮询。如果此队列为空，DC 值仍有剩余，设置 DC 值为 0，调度器移向下一队列。

### 3 基于流量分级预测的 P-DWRR 算法

鉴于现有队列调度算法没有考虑网络流量的自相似特性，导致数据分组时延和时延抖动增大、分组丢失率增高的问题，本文提出 P-DWRR (prediction DWRR) 算法，即在原 DWRR 算法的基础上采用基于流量自相似特性的流量分级预测与队列优先级的权值设定：每隔一定的时间间隔  $\Delta t$ ，根据流量当前的水平等级预测下一个时间间隔的流量水平等级，进而依据流量水平等级动态地调整每个队列的权值和服务量子，再根据业务优先级和队列等待时间调整调度顺序。

#### 3.1 自相似过程

设  $X = \{X_i : i = 1, 2, \dots\}$  表示一个广义平稳离散

随机过程，其中， $X_i$  表示第  $i$  个时间间隔到达网络节点的数据分组数。 $X$  具有恒定均值  $\mu$  和有限方差  $\sigma^2$ ，且其自相关函数为  $r(k)$ 。

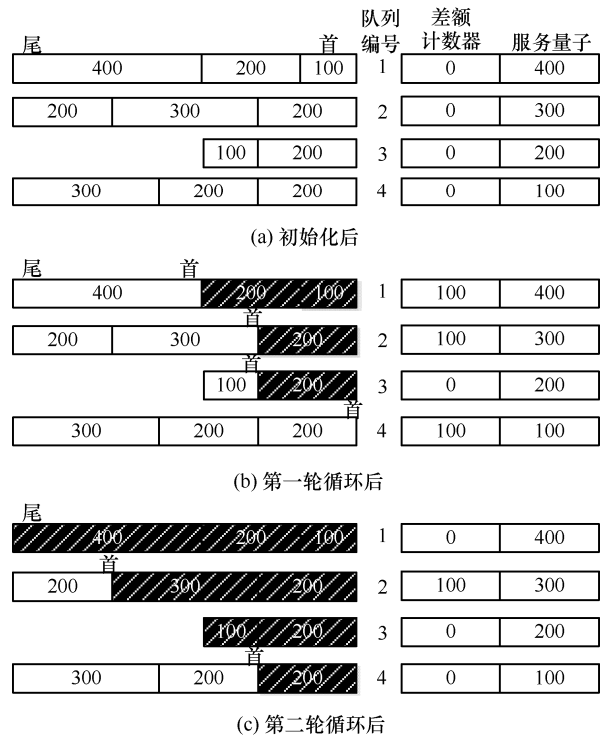


图 2 DWRR 算法的调度过程

随机过程  $X$  的  $m$  阶聚集过程  $X^{(m)} = \{X_i^{(m)}, i = 1, 2, \dots\}$  的定义为

$$X_i^{(m)} = \sum_{k=m(i-1)+1}^{mi} \frac{X_k}{m}, m = 1, 2, \dots \quad (1)$$

对每个  $m$ ， $X^{(m)}$  都定义了一个广义平稳随机过程，其方差和自相关函数分别为  $V^{(m)}$  和  $r^{(m)}(k)$ 。

如果随机过程  $X$  的自相关函数满足  $r(k) = \frac{[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}]}{2} \sim g(k), k \in Z_+$ ，则称  $X$

为严格二阶自相似过程，且具有 Hurst 参数  $H = 1 - \frac{\beta}{2}, 0 < \beta < 1$ 。如果随机过程  $X$  的自相关函数满足  $r(k) \sim ck^{-\beta}, k \rightarrow \infty$ ，其中  $c$  为正常数，则称  $X$  为长相关过程。如果随机过程  $X$  的  $m$  阶聚集过程  $X^{(m)}$  的自相关函数满足  $\lim_{m \rightarrow \infty} r^{(m)}(k) = g(k), k \in Z_+$ ，则称  $X$  为渐近二阶自相似过程。

当  $H \in (0.5, 1)$  时，随机过程具有自相似性，并且  $H$  值越大，自相似程度越高。文献[16]指出，当足够多的、服从重尾分布的 ON/OFF 过程叠加在一起时，

叠加后的过程具有自相似性，其 Hurst 参数为  $H = \frac{3-\alpha}{2}$ ，其中  $\alpha$  为重尾分布的形状参数。

### 3.2 流量水平分级

设  $\{X(t), t \in T\}$  是一个广义平稳随机过程， $x(t)$  是随机过程的一个样本函数。

取 2 个参数  $T_1, T_2 > 0$ ，在  $t$  时刻，可以使

$$a = \sum_{i \in [t-T_1, t)} x(i), b = \sum_{i \in [t, t+T_2)} x(i) \quad (2)$$

$$V_1 = \sum_{i \in [t-T_1, t)} X(i), V_2 = \sum_{i \in [t, t+T_2)} X(i) \quad (3)$$

其中， $a$  表示在最近的过去  $[t-T_1, t)$  上观察到的总流量， $b$  表示在最近的未来  $[t, t+T_2)$  上观察到的总流量， $V_1$  和  $V_2$  表示最近的过去和最近的未来的复合随机变量。

定义条件概率计算式为  $\Pr\{V_2 = b | V_1 = a\}$ ，表示在最近的过去的总流量为  $V_1 = a$  的条件下，最近的未来的总流量为  $V_2 = b$  的概率。

假设随机过程  $\{X(t), t \in T\}$  具有有限的均值和方差，分别为  $\mu_k = E(V_k)$  和  $\sigma_k^2 = \Lambda(V_k)$ ， $k=1,2$ ，为了描述流量水平的“高”和“低”，本文将  $V_k$  的变化范围分为以下 6 个级别

$$\begin{aligned} &(-\infty, \mu_k - 4\sigma_k), [\mu_k - 4\sigma_k, \mu_k - 2\sigma_k), \\ &[\mu_k - 2\sigma_k, \mu_k), [\mu_k, \mu_k + 2\sigma_k), \\ &[\mu_k + 2\sigma_k, \mu_k + 4\sigma_k), [\mu_k + 4\sigma_k, +\infty) \end{aligned} \quad (4)$$

定义 2 个新的随机变量  $L_1$  和  $L_2$ ，其中  $L_1$  为  $T_1$  时间段上的流量等级， $L_2$  为  $T_2$  时间段上的流量等级，则有

$$L_k = \begin{cases} 1, & V_k \in (-\infty, \mu_k - 4\sigma_k) \\ 2, & V_k \in [\mu_k - 4\sigma_k, \mu_k - 2\sigma_k) \\ 3, & V_k \in [\mu_k - 2\sigma_k, \mu_k) \\ 4, & V_k \in [\mu_k, \mu_k + 2\sigma_k) \\ 5, & V_k \in [\mu_k + 2\sigma_k, \mu_k + 4\sigma_k) \\ 6, & V_k \in [\mu_k + 4\sigma_k, +\infty) \end{cases} \quad (5)$$

其中， $k=1,2$ ， $L_k$  是  $V_k$  的函数，即  $L_k = L_k(V_k)$ 。因此，如果  $L_k \approx 1$ ，那么流量水平相对于平均值是“低”；如果  $L_k \approx 6$ ，那么流量水平相对于平均值是“高”。

### 3.3 流量水平条件转移概率估计

取长度为  $n$  s 的聚合流量序列  $X_i$ ，将其分为

$N = \frac{n}{T_1 + T_2}$  块，每个连续非重叠块的长度为  $T_1 + T_2$ ，并且对于第  $j=1,2,\dots,N$  个非重叠块，计算长度为  $T_1, T_2$  上的总流量，分别记为  $V_1, V_2$ 。令  $\ell, \ell' = 1,2,\dots,6$

分别为  $T_1, T_2$  上的流量等级， $h_\ell$  为满足  $L_1(V_1) = \ell$ （即  $T_1$  上的流量等级为  $\ell$ ）的总块数， $h_{\ell'}$  为当  $L_1(V_1) = \ell$  时，满足  $L_2(V_2) = \ell'$ （即  $T_1$  的流量等级为  $\ell$  条件下， $T_2$  的流量等级为  $\ell'$ ）的总块数，则流量水平条件转移概率计算式为  $\Pr\{L_2 = \ell' | L_1 = \ell\} = \frac{h_{\ell'}}{h_\ell}$ 。

### 3.4 P-DWRR 算法设计

P-DWRR 算法基本步骤介绍如下。

**Step1** 根据  $n$  个队列优先级初始化队列的权值  $w_{0i} (i=1,2,\dots,n)$ ，若最小权值不为 1，则需将其设为 1，其他权值同比例缩放，最后归一化为  $\bar{w}_{0i}$ 。

**Step2** 计算过去  $\Delta t$  时间段内队列  $i$  的流量水平的具体等级  $\ell$ ，根据条件转移概率计算式  $\Pr\{L_2 = \ell' | L_1 = \ell\}$ ，预测下一时间段队列  $i$  的流量等级为  $\bar{\ell}' = \sum_{\ell'=1}^6 \ell' \Pr\{L_2 = \ell' | L_1 = \ell\}$ 。

**Step3** 根据  $S = \bar{w}_{0i} \bar{\ell}' \frac{t_i}{\sum_{j=1}^n t_j}$  的值将队列降序排

列，其中  $t_i$  为队列  $i$  的等待时间，依次从高到低进行调度，以平衡队列优先级、数据突发程度和队间公平性。

**Step4** 判断队列  $i$  是否为空，若为空，则设  $DC[i] = 0$ ，队列  $i = i + 1$ ；若不为空，转到 Step5。

**Step5** 根据计算出的流量等级改变权值，计算更新后的  $w_i = w_{0i} + \Delta w_i$ ，并将结果写入队列的权值表中，其中  $\Delta w_i$  的计算过程介绍如下。

由式(5)可知，当流量等级小于 3 时，该  $\Delta t$  时间段内流量水平远低于其均值，可减少其预先分配的带宽，即  $\Delta w_i = \max i_{\Delta w_i} \frac{\bar{\ell}' - 3}{2}$ ；当流量等级处于 3~4 时，该  $\Delta t$  时间段内流量水平在均值上下浮动，令  $\Delta w_i = 0$ ，即流量等级处于 3~4 时，不调整其预先分配的带宽；当流量等级大于 4 时，该  $\Delta t$  时间段内流量水平高于其均值，且随着流量等级的增大，数据突发程度增高，可将其预先分配的带宽增大  $\max i_{\Delta w_i} \frac{\bar{\ell}' - 4}{2}$ ，其中  $\max i_{\Delta w_i}$  表示队列  $i$  权值增量的最大阈值。

总的来说，权值增量  $\Delta w_i$  的计算式为

$$\Delta w_i = \begin{cases} \max i_{\Delta w_i} \frac{\bar{\ell}' - 3}{2}, & \bar{\ell}' < 3 \\ 0, & 3 \leq \bar{\ell}' \leq 4 \\ \max i_{\Delta w_i} \frac{\bar{\ell}' - 4}{2}, & \bar{\ell}' > 4 \end{cases} \quad (6)$$

**Step6** 从队列的权值表读取队列  $i$  的新权值  $w_i$ , 归一化为  $\bar{w}_i$ 。

**Step7** 根据队列  $i$  的权值  $\bar{w}_i$ , 分配队列  $i$  一次可增加的服务量子  $q=C\bar{w}_i$ , 其中  $C$  为服务速率。

**Step8** 判断当前调度队列中的首部分组的字节数和差值计数器值的关系。

如果差值计数器的值大于队列首部分组的字节数, 则调度器允许从输出端口将该首部分组发送出去, 并且差值计数器的值减去队列首部分组的字节数  $P\_size$ 。调度器在发送完该分组后, 继续检测当前队列新的队列首部分组的字节数与差值计数器值的大小情况。如果该队列首部分组的字节数仍然小于差值计数器的值, 继续发送该队列首部分组, 并将差值计数器的值减去首部分组的字节数, 重复该过程, 直到当前队列为空, 或者当前队列首部分组的字节数大于差值计数器值。如果队列为空, 则转到 Step4。

如果差值计数器的值小于当前队列首部分组的字节数, 将拒绝对该队列进行调度服务, 差值计数器将该次未使用的额度保留, 并在下一次轮询到该队列时加入差值计数器中使用, 然后转到 Step4。

**Step9** 如果所有队列中都没有等待调度转发的数据分组存在, 则调度算法结束。

P-DWRR 算法中, 计算均值和方差的时间复杂度均为  $O(n)$ , 队列调度算法的时间复杂度为  $O(n^2)$ 。P-DWRR 算法流程如图 3 所示。

## 4 仿真校验

### 4.1 流量等级预测结果

为了得到自相似流量, 本文利用 100 个独立 Pareto 分布的 ON/OFF 源叠加模型来模拟网络自相似流, Pareto 分布的累积分布函数为

$$F(x) = \begin{cases} 1 - \left(\frac{\beta}{x}\right)^\alpha, & x \geq \beta \\ 0, & x < \beta \end{cases} \quad (7)$$

其中,  $\alpha$  值分别取 1.2、1.4 和 1.6, 由  $H = \frac{3-\alpha}{2}$  可得其对应的 Hurst 参数值分别为 0.9、0.8 和 0.7, 数据分组大小为 128 B, ON 持续时间均值为 50 ms, OFF 持续时间均值为 10 ms, 仿真时间长度为 10 000 s, 取  $T_1 = T_2 = 5$  s, 通过 3.2 节流量水平分级的条件转移概率计算方法, 可得表 1~表 3。

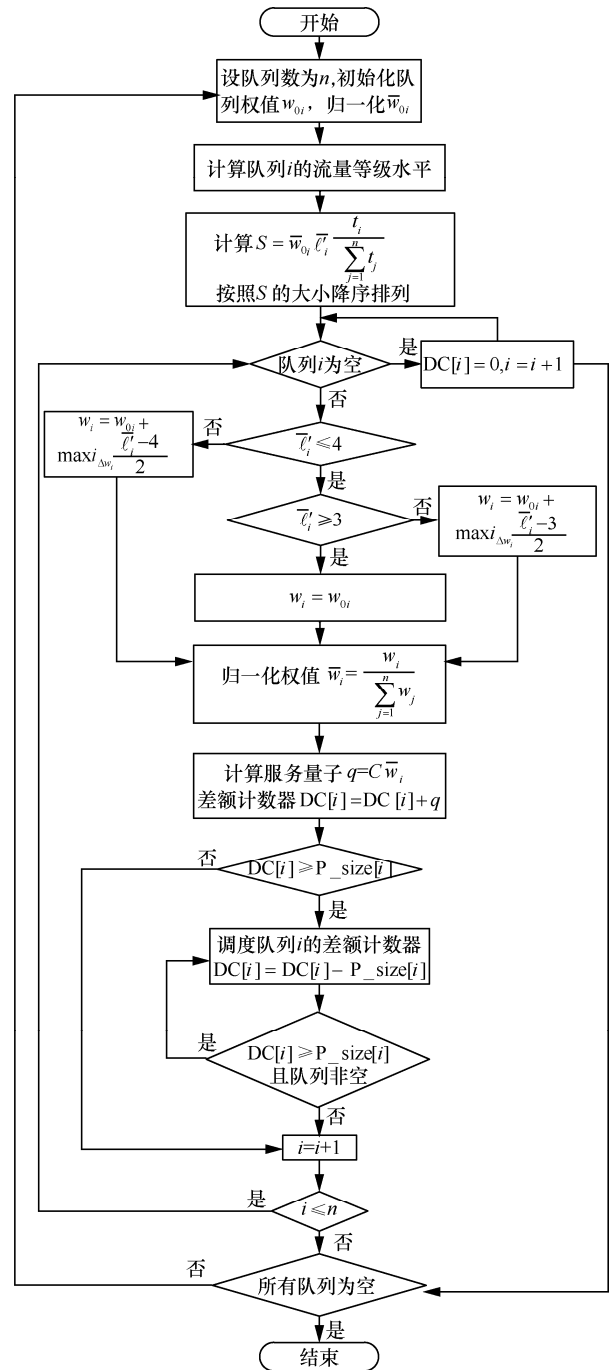


图 3 P-DWRR 算法流程

表 1  $\alpha = 1.2$  时流量水平条件转移概率

$L_2$	$L_1$						$\bar{c}'$
	1	2	3	4	5	6	
1	0.614	0.312	0.074	0	0	0	1.5
2	0.006	0.512	0.297	0.185	0	0	2.7
3	0.002	0.153	0.340	0.301	0.204	0	3.6
4	0	0.003	0.327	0.394	0.284	0.002	4.0
5	0	0.002	0.235	0.422	0.337	0.004	4.1
6	0	0	0.002	0.356	0.500	0.142	4.8

表2  $\alpha = 1.4$  时流量水平条件转移概率

$L_2$	$L_1$						$\bar{t}'$
	1	2	3	4	5	6	
1	0.553	0.263	0.122	0.062	0	0	1.6
2	0.048	0.462	0.285	0.174	0.031	0	2.7
3	0.007	0.196	0.338	0.291	0.164	0.004	3.4
4	0.004	0.075	0.285	0.357	0.271	0.008	3.8
5	0	0.023	0.260	0.386	0.325	0.006	4.0
6	0	0	0.013	0.387	0.487	0.113	4.7

表3  $\alpha = 1.6$  时流量水平条件转移概率

$L_2$	$L_1$						$\bar{t}'$
	1	2	3	4	5	6	
1	0.457	0.217	0.197	0.118	0.011	0	2.0
2	0.063	0.381	0.266	0.157	0.128	0.005	2.9
3	0.017	0.227	0.326	0.241	0.182	0.007	3.4
4	0.011	0.143	0.295	0.313	0.223	0.015	3.6
5	0.005	0.046	0.319	0.372	0.247	0.011	3.8
6	0	0.002	0.083	0.406	0.424	0.085	4.5

### 4.2 性能指标选取

#### 1) 分组丢失率

分组丢失率是指测试中丢失数据分组占所发送数据分组的比例，与数据分组长度和发送频率相关。在队列调度算法中，要求在缓冲区大小固定的情况下，尽可能降低分组丢失率。

#### 2) 时延

时延是指数据从网络的一端传送到另一端所需的时间，一般由发送时延、传播时延、排队时延和处理时延组成。在队列调度算法中，以排队时延作为衡量指标。

### 4.3 仿真结果及分析

仿真实验采用 Matlab 仿真软件进行，实验使用的仿真拓扑结构如图4所示。

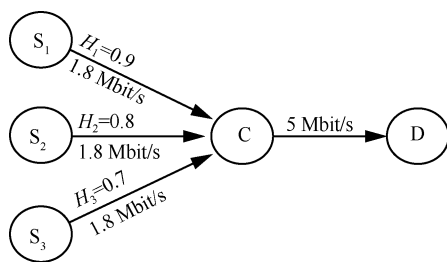


图4 仿真拓扑结构

图4中， $S_1$ 、 $S_2$ 和 $S_3$ 为3个源节点，分别采用100个独立Pareto分布的ON/OFF源叠加模型来模拟网络自相似流，具体参数与4.1节的设置相

同。其中Pareto分布的 $\alpha$ 值分别取1.2、1.4和1.6，对应的队列分别为队列1、队列2和队列3，Hurst参数值分别为0.9、0.8、0.7，3个队列的优先级按由高到低的顺序排列，初始权值设置为3:2:1。C为中间节点，D为目的节点，中间节点服务速率为5 Mbit/s，源节点向中间节点的发送速率均为1.8 Mbit/s， $\max i_{\Delta w_i} = 1$ 。

#### 1) 调度算法分组丢失率

当DWRR算法<sup>[2]</sup>、VDWRR算法<sup>[6]</sup>和P-DWRR算法在缓冲区长度从5个分组变化到100个分组时，分组丢失率曲线如图5所示。

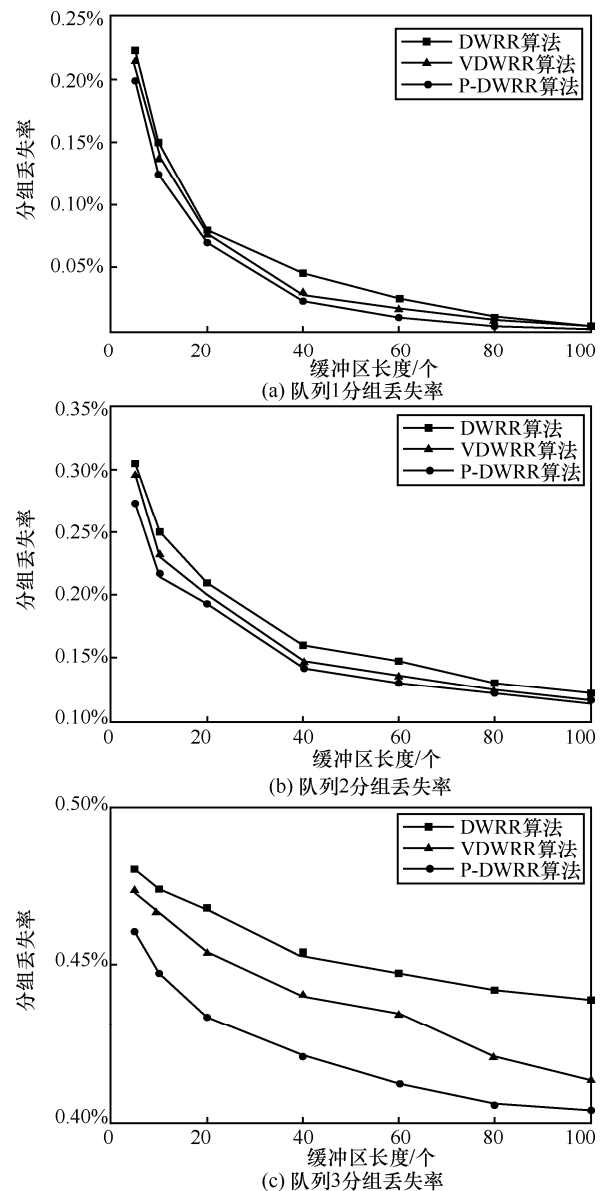


图5 3个队列的分组丢失率曲线

在队列分组丢失率方面，随着缓冲区长度的增加，3种算法的分组丢失率都呈现出减小趋势。

其中在同一个缓冲区长度下分组丢失率的不同主要受队列权值设定的影响,而在同一个队列的不同调度算法的分组丢失率比较中,P-DWRR 算法的分组丢失率是最低的。以队列 3 为例,在不同的缓冲区长度下,P-DWRR 算法相对于 DWRR 算法,平均分组丢失率降低了约 7.1%;相对于 VDWR 算法,平均分组丢失率降低了约 4.4%。这是由于本文的 P-DWRR 算法根据优先级和队列的流量水平预测结果动态地调整权值和调度顺序,把调度过程划分为多个调度周期,在一个调度周期内根据事先预测的队列流量突发程度,设置一个合适的权值比例进行调度,减少队列由于流量突发导致从缓存中溢出的数据分组数量。

2) 调度算法排队时延

DWRR 算法、VDWR 算法和 P-DWRR 算法的排队时延曲线如图 6 所示。为了更好地观察时延的变化情况,设定此处的缓冲区长度为 100 个分组,仿真时间为 1 000 s。

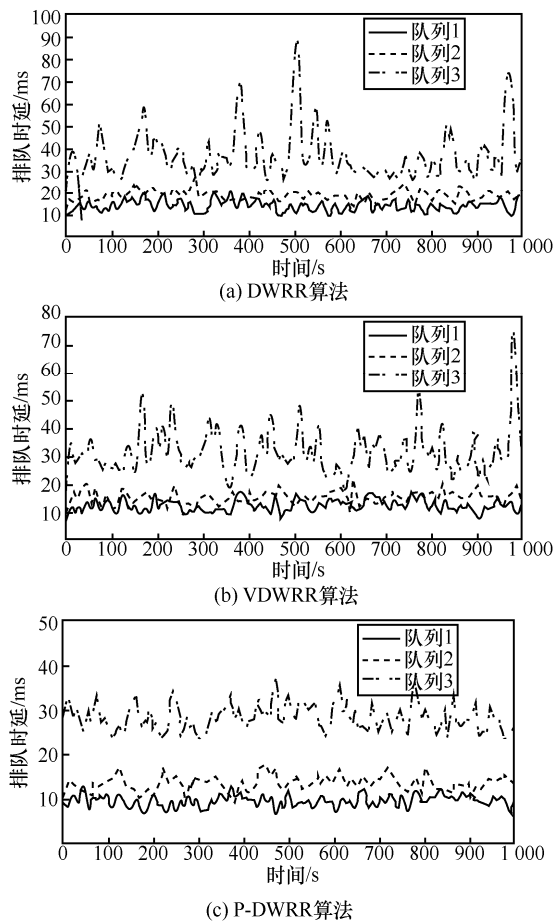


图 6 3 种算法的排队时延曲线

3 种算法平均排队时延的比较如表 4 所示。从表 4 中可以看出,本文的 P-DWRR 算法的平均排队时延是最低的。

表 4 3 种算法平均排队时延的比较

队列	DWRR 算法/ms	VDWR 算法/ms	P-DWRR 算法/ms
队列 1	13.5	12.5	9.8
队列 2	16.9	15.7	13.7
队列 3	38.2	33.2	28.8

以队列 1 为例, P-DWRR 算法比 DWRR 算法的排队时延降低了 27%左右,比 VDWR 算法的排队时延降低了 22%左右。这主要是因为不同 Hurst 参数下的流量突发程度不同,其流量水平在不断变化,相对于权值和调度顺序都不变的 DWRR 算法、调度顺序动态调整的 VDWR 算法而言,本文的 P-DWRR 算法权值的动态设置利用了预测的网络流量水平等级,更能贴合队列中实际的流量水平,同时因其调度顺序考虑了业务优先级和排队等待时间,使其排队时延更低。

3 种算法排队时延标准差的比较如表 5 所示。

从表 5 可以看出, P-DWRR 算法的队列时延变化更平稳。

表 5 3 种算法排队时延标准差的比较

队列	DWRR 算法	VDWR 算法	P-DWRR 算法
队列 1	8.4	7.8	5.8
队列 2	9.8	9.3	7.9
队列 3	19.7	16.3	13.6

5 结束语

本文通过分析网络流量自相似特性对队列调度算法影响以及不同业务的 QoS 需求,设计了队列调度算法 P-DWRR。队列调度算法在业务优先级决定队列初始权值的基础上,根据每个队列在一定时间间隔内流量等级的不同进行调整,实现权值的动态分配,并对队列的服务顺序进行调整。实验结果表明, P-DWRR 算法具有较好的时延和分组丢失性能,可满足网络不同业务类型的数据在不同自相似程度下的 QoS 要求。

参考文献:

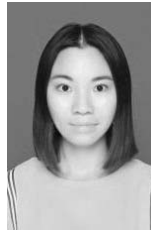
[1] PAREKH A K, GALLAGER R G. A generalized processor sharing approach to flow control in integrated services networks-the single

- node case[C]//Proceedings of INFOCOM'92. Piscataway: IEEE Press, 1992: 915-924.
- [2] SHREEDHAR M, VARGHESE G. Efficient fair queuing using deficit round robin[J]. IEEE/ACM Transactions on Networking, 1996, 4(3): 375-385.
- [3] 王胜灵, 侯义斌, 黄建辉, 等. PFWR: 能实现比例公平的增强型 WRR[J]. 计算机学报, 2008, 31(5): 775-782.  
WANG S L, HOU Y B, HUANG J H, et al. PFWR: an enhanced WRR scheduling realizing the proportion fairness principle[J]. Chinese Journal of Computers, 2008, 31(5): 775-782.
- [4] 刘桂开. 用逐次最小权值轮询算法实现公平和低时延分组调度[J]. 系统科学与数学, 2014, 34(9): 1080-1099.  
LIU G K. Fair and low-latency packet scheduling using successive minimal-weight round robin[J]. Journal of Systems Science and Mathematical Sciences, 2014, 34(9): 1080-1099.
- [5] 李秉权, 张松, 王兆伟, 等. WFQ 与 WRR 调度算法的性能分析与改进[J]. 北京理工大学学报, 2015, 35(3): 316-320.  
LI B Q, ZHAN H S, WANG Z W, et al. Performance analysis and improvement about WFQ and WRR scheduling algorithm[J]. Transactions of Beijing Institute of Technology, 2015, 35(3): 316-320.
- [6] 江文静, 蔡祥宝. DiffServ 队列调度算法研究[J]. 计算机技术与发展, 2015, 25(4): 85-88.  
JIANG W J, CAI X B. Research on queue scheduling algorithm based on diffserv[J]. Computer Technology and Development, 2015, 25(4): 85-88.
- [7] BOK P, KOHLS K, TUCHELMANN Y, et al. I-DWRR-an insolvency enabled scheduling scheme extending Deficit Weighted Round Robin[C]//The 3rd IEEE International Workshop on Management of Emerging Networks and Services. Piscataway: IEEE Press, 2011: 695-700.
- [8] OUNI R, BHAR J, TORKI K. A new scheduling protocol design based on deficit weighted round robin for QoS support in IP networks[J]. Journal of Circuits, Systems, and Computers, 2013, 22(3): 1-21.
- [9] LI C C, WANG K. An SLA-aware load balancing scheme for cloud datacenters[C]//2014 International Conference on Information Networking (ICOIN). Piscataway: IEEE Press, 2014: 58-63.
- [10] LELAND W E, TAQQU M S, WILLINGER W, et al. On the self-similar nature of Ethernet traffic (extended version)[J]. IEEE/ACM Transactions on Networking, 1994, 2(1): 1-15.
- [11] CROVELLA M E, BESTAVROS A. Self-similarity in world wide web traffic: evidence and possible causes[J]. IEEE/ACM Transactions on Networking, 1997, 5(6): 835-846.
- [12] WILLINGER W, TAQQU M S, SHERMAN R, et al. Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level[J]. IEEE/ACM Transactions on Networking, 1997, 5: 71-86.
- [13] 王晖, 季振洲, 孙彦东, 等. 基于时间槽的自相似流量随机早检测算法——SFRED[J]. 通信学报, 2010, 31(10): 115-120.  
WANG H, JI Z Z, SUN Y D, et al. Time slot-based RED algorithm on self-similar flows: SFRED[J]. Journal on Communications, 2010, 31(10): 115-120.
- [14] 别玉霞, 潘成胜, 刘海燕, 等. 基于 AOS 的 Hurst-优先级自适应 RED 与动态调度算法[J]. 通信学报, 2012, 33(10): 156-165.  
BIE Y X, PAN C S, LIU H Y, et al. Hurst-priority adaptive RED and dynamic scheduling algorithm based on AOS[J]. Journal on Communications, 2012, 33(10): 156-165.
- [15] DOMAŃSKI A, DOMAŃSKA J, CZACHÓRSKI T, et al. Self-similarity traffic and AQM mechanism based on non-integer order  $PI^{\alpha}D^{\beta}$  controller[J]. Computer Networks, 2017, 718: 336-350.
- [16] TAQQU M S, WILLINGER W, SHERMAN R. Proof of a fundamental result in self-similar traffic modeling[J]. ACM SIGCOMM Computer Communication Review, 1998, 27(2): 5-23.

## [作者简介]



魏德宾 (1978- ), 男, 辽宁大连人, 南京理工大学博士生、大连大学副教授, 主要研究方向为空间信息网络传输技术、流量工程和网络优化。



沈婷 (1995- ), 女, 安徽滁州人, 大连大学硕士生, 主要研究方向为空间信息网络传输技术。



杨力 (1982- ), 女, 黑龙江哈尔滨人, 博士, 大连大学教授, 主要研究方向为空间信息网络传输技术、无线通信网络协议理论与方法。



戚耀文 (1994- ), 男, 山东威海人, 南京理工大学博士生, 主要研究方向为空间信息网络传输技术、空间信息网络路由协议。