一种增强的速率比例调度器

王重钢, 隆克平, 龚向阳, 程时端
(北京邮电大学交换技术与通信网国家实验室, 北京 100876)

摘 要: 作为分组交换网络中提供服务质量保证的一种重要机制, 队列调度算法近年来得到了较为广泛的研究. 本文提出了增强的速率比例调度器模型(Enhanced RPS). ERPS 模型给出了设计系统虚时间函数的具体方法和条件, 并计算出了其相应的公平性指数, 最后给出了 1 个基于 ERPS 的分组级的队列调度算法.

关键词: 加权公平排队; 速率比例调度器; 服务质量; 分组交换网络
中图分类号: TN393.03 文献标识码: A 文章编号: 0372-2112(2001)07-0912-04

An Enhanced Rate-Proportional Server

WANG Chong-gang, LONG Ke-ping, GONG Xiang-yang, CHENG Shi-duan
(National Laboratory of Switching Technology & Telecommunication Networks, BUPT, Beijing 100876, China)

Abstract: As an important mechanism to provide guaranteed QoS in packet-switched networks, queuing scheduling algorithms have been widely researched in recent years. We present an Enhanced Rate-Proportional Server (ERPS), which gives the method and condition of designing system virtual-time function and compute their fairness index. We also give an example of packet- level queuing scheduling algorithm based on ERPS.

Key words: weighted fairness queuing; rate-proportional server; quality of service; packet-switched networks

1 引言

目前, Internet 网络中的服务质量(QoS)保证问题正成为研究的热点. 一方面是由于目前 Internet 的工作方式仍然是传统的尽力而为(Best-Effort)的分组转发方式; 另一方面, 现在和将来的各种应用可能有着不同的服务质量需求, 具体表现为在带宽、时延、时延抖动、分组丢失率等方面有不同的要求. 因此网络本身必须具有给不同的应用提供相应服务质量的能力. 而队列调度算法正是提供服务质量保证的重要机制之一.

近年来, 基于 GPS(General Processor Sharing)[1]模型的分组公平排队(PFQ: Packet Fair Queuing)调度算法得到了较为广泛的研究[2~6]. 所有的 PFQ 算法皆基于: (1)系统维持一个全局函数, 称为系统虚时间函数[3](或 system potential function[7]), 用以记录调度器已提供的服务量; 调度器利用系统虚时间函数为系统中的每个分组计算其相应的开始时间标签和完成时间标签, 请见式(1). (2)执行一定的分组选择策略, 有最小完成时间标签优先(SFF[3])、最小开始时间标签优先(SSF[8])、最小合格完成时间标签优先(SEFF[4]).

$$S_i^k = \max(F_i^{k-1}, P(a_i^k))$$
$$F_i^k = S_i^k + (l_i^k / r_i) \tag{1}$$

式中: 设 p_i^k 表示连接(或会话) i 的第 k 个分组; a_i^k 表示 p_i^k 的到达时间; S_i^k 表示 p_i^k 的服务开始时间标签; F_i^k 表示 p_i^k 的服务结束时间标签; l_i^k 表示 p_i^k 的分组长度; r_i 表示连接 i 的预约速率, $P(t)$ 表示在 t 时刻的系统虚时间函数值.

我们可以把每个队列(或连接)的头部分组的虚时间标签(包括开始时间标签和结束时间标签)称为此队列的虚时间标签(包括开始时间标签和结束时间标签). 调度算法按照一定的分组选择策略依次发送各分组, 但总是虚时间标签值小的分组首先得到发送机会. 所以如果某两个队列的虚时间标签的差值很大, 那么在一段较长的时间里, 只有虚时间标签小的那个队列首先得到发送分组的机会, 而虚时间标签值较大的队列则只有等待, 这种情况会一直持续到二者虚时间标签值相等或其差值小于零为止. 可见, 如果队列之间的虚时间标签差值很大或没有上限, 则调度算法会带来各队列之间接受服务的不公平性.

速率比例调度器(RPS: Rate-Proportional Server)[7]是一种较好的综合流体调度模型, GPS 模型是其特例之一. 分组级的 RPS(PRPS)调度算法也几乎包括了目前已经存在的所有 PFQ 调度算法(SCFQ[3]例外). 利用 RPS 模型, 通过选择不同的系统虚时间函数(但 RPS 并没有告诉我们如何选择), 可以方便地设计出符合特定需求的分组级的队列调度算法.

本文提出了增强的速率比例调度器模型(Enhanced RPS). ERPS 模型给出了设计系统虚时间函数的具体方法和条件, 并进行了较为详细的分析和证明. 这对设计分组级的队列调度

收稿日期: 2000-07-10; 修回日期: 2000-11-27
基金项目: 国家自然科学基金(No. 69972008)

算法具有重要的指导意义. 最后, 我们还给出了一个基于 ERPS 的分组级的队列调度算法.

2　速率比例调度器(RPS)

2.1　调度算法的分组级模型和流体流模型

一个调度算法(设为 S)的流体流(fluid-flow)模型(记为 f-S)是一种理想的模型, 在此模型下系统可以在任意瞬时时刻同时为多个连接服务. 而实际上, 信息是以分组的形式进行转发的, 在某一时刻, 系统只能为一个连接进行服务, 这便是调度算法的分组级模型(记为 p-S). 后面提到调度算法时, 如无特别申明, 皆指其分组级模型. 一个调度算法的这两个模型是一一对应的. 在研究一个调度算法的性能时, 我们常借助于流体流模型; 即先研究流体流模型下的性能, 再转化成相应分组级模型中的性能即可.

2.2　速率比例调度器(RPS)

RPS[7]模型给我们提供了一种可用来方便地设计满足特定需求的分组级队列调度算法的方法. 而且分组级的 RPS (PRPS) 调度算法具有同 WFQ 一样的时延特性, 且不依赖于系统虚时间函数的具体形式[7]. RPS 主要涉及到 RPS 的判别规则、各连接虚时间函数的计算, 具体如下.

2.2.1　RPS 的判别规则　RPS 属于流体流模型. 一个流体流调度算法当且仅当其系统虚时间函数 $P(t)$ 满足下面两个条件时, 才属于 RPS 类调度算法.

(1) 在系统活动期(即系统中有分组等待调度)中的任意时间间隔$(t_1, t_2]$里, 满足

$$P(t_2) - P(t_1) \geq t_2 - t_1 \tag{2}$$

(2) 在任意时刻 t, 系统虚时间函数不大于任何活动连接(即有分组堆积的连接)的虚时间函数值,

$$P(t) \leq \min_{i \in B(t)} \{P_i(t)\} \tag{3}$$

$B(t)$ 表示在时刻 t 有分组堆积连接的集合; $P_i(t)$ 为连接 i 在时 t 刻的虚时间函数值.

2.2.2　各连接虚时间函数 $P_i(t)$ 的计算　在 RPS 中, $P_i(t)$ 的计算遵循下面三条规则.

(a) 如果连接 i 无分组堆积, 则 $P_i(t)$ 不变.

(b) 如果连接 i 在时刻 τ 才开始突然有分组堆积, 则

$$P_i(\tau) = \max(P_i(\tau -), P(\tau -)) \tag{4}$$

(c) 对任何的 $t > \tau$, 且连接 i 在$(\tau, t]$期间内保持有分组堆积, 则

$$P_i(t) = P_i(\tau) + [W_i(\tau, t)/r_i] \tag{5}$$

这里 $W_i(\tau, t)$ 表示连接 i 在$(\tau, t]$期间获得的服务量.

系统根据 $P_i(t)$ 调节每个活动连接接受服务的机会, 因为只有具有最小 $P_i(t)$ 值的连接才能接受服务. 在 RPS 里, 每个正接受服务的连接都能即刻得到与其预约速率成正比的服务量, 所有正接受服务的连接的虚时间函数值以相同的速率(大等于 1)增加, 这提供了公平调度的保证.

3　ERPS 模型

RPS 模型告诉我们, 通过选择不同的系统虚时间函数可以设计出满足特定需求的分组级队列调度算法. 但是 RPS 模型中并没有给出如何去构造满足特定需求的函数. 这里, 给出 RPS 模型的增强模型: ERPS. 下面首先介绍 ERPS 模型中对队列的虚时间函数和系统虚时间函数的计算.

3.1　队列虚时间函数的计算

在 ERPS 模型里, 对各队列虚时间函数的计算遵循同 RPS[7] 一样的规则(请见前面第 2 部分).

3.2　系统虚时间函数的计算

ERPS 采用文[6]中相同的方法来计算系统虚时间函数, 具体如下:

(1) 如果系统处于空闲期, 则 $P(t) = 0$.

(2) 在系统忙期内, 系统虚时间函数为关于时间 t 的一个线性分段且递增的函数, 并把各分段处的非连续点 $\tau_0, \tau_1, \ldots, \tau_k(\tau_0 < \tau_1 < \ldots < \tau_k)$ 称为校正时刻. 在 ERPS 里, 校正时刻为每发送完一个分组的时刻.

(3) 在两个连续的校正时刻之间, $P(t)$ 随 t 线性增加, 且斜率为 1, 即

$$P(t) = P(\tau_j) + t - \tau_j \quad (\tau_j \leq t < \tau_{j+1}) \tag{6}$$

(4) 在某个校正时刻 τ_j, 根据下式对系统虚时间函数进行校正

$$P(\tau_j) = \max(P(\tau_j -), S^p(\tau_j)) \tag{7}$$

式中: $P(\tau_j -) = P(\tau_{j-1}) + \tau_j - \tau_{j-1}$; $S^p(t)$ 称为校正函数, 它是不同 RPS 算法之间的差别所在.

(5) 每当系统变为空闲, 则置 $P(t)$ 等于零.

在 ERPS 中, 通过理论分析给出了 $S^p(t)$ 函数的上限(记为 $S^U(t)$)和满足一定公平性要求时的下限(记为 $S^L(t)$), 并得出了其相应的公平性指数(见定理 1、2、3). 这样在设计具体的分组级调度算法时, 只需在 $S^U(t)$ 和 $S^L(t)$ 之间选择一定的 $S^p(t)$ 函数即可, 而且可以根据定理 3 方便地求出算法的公平性指数, 这对调度算法的设计具有较为重要的指导意义.

需要说明的是, 公平性指数的计算不仅在于可以定量地评价算法的公平性, 而且对于调度算法的具体实现也具有重要的意义. 设 ΔT 为系统中任何两个队列之间时间标签的最大差值, 则在算法的具体实现时, 可根据 ΔT 的大小选取不同的字长, 来表示各队列的时间标签且在排序时不影响它们之间真实的大小关系. 如果公平性指数越小, 则 ΔT 也越小, 从而可以选择较小的字长来表示各队列的时间标签(如果采用无符号整形数来表示时间标签, 可以采用比特的字长, 只要 $2^{N+1} \geq \Delta T$ 即可).

引理 1　设有调度算法 S(其虚时间函数的计算根据 3.1 和 3.2 节的定义), 只要 $S^P(t) \leq F_{服务}(t)$, 则此调度算法的流体流模型(f-S) 一定属于 RPS 类调度算法. 这里 $F_{服务}(t)$ 表示在 t 时刻正在接受服务的分组的完成时间标签.

证明　为了证明 f-S 属于 RPS, 需要证明式(2)和(3)成立. 根据式(7)的说明, 式(2)显然成立. 下面将通过数学归纳法证明式(3)也成立.

第 1 步　证明在 $\tau_0 \leq t \leq \tau_1$ 时, 式(3)成立. 其中 τ_0 是系统忙期的开始时刻, τ_1 为第一个校正时刻. 由于 $P_j(\tau_0) = P(\tau_0) = 0$, $\forall j \in B(\tau_0)$, 利用文[6]中的引理 1, 可以得到:

$$P_j(t) \geq P(t) \quad \forall j \in B(t), \tau_0 \leq t < \tau_1$$

当 $t = \tau_1$ 时, 记 $s(\tau_1)$ 为在 τ_1 时刻刚发送完其第一个分

组的连接, $F_{s(\tau_1)}$ 为此分组的完成时间标签(等于连接 $s(\tau_1)$ 在 τ_1 时刻的虚时间函数值). 由于连接 $s(\tau_1)$ 在 τ_1 时刻正接受服务, 所以 $F_{s(\tau_1)}$ 是所有活动连接中的最小值, 有

$$F_{s(\tau_1)} \le P_j(\tau_1) \qquad \forall j \in B(\tau_1 -) \setminus \{s(\tau_1)\}$$

于是有 $S^p(\tau_1) \le F_{s(\tau_1)} \le P_j(\tau_1)$

此外, 从式(5)和(6)可以看出, 在[0, τ_1)期间 $P_i(t)$ 的增长速度等于 $P(t)$ 的增长速度, 所以下式成立

$$F_{s(\tau_1)} \ge \tau_1 = P(\tau_1 -)$$

于是得到 $P(\tau_1) \le P_j(\tau_1) \qquad \forall j \in B(\tau_1)$

对第一步的证明完成.

第2步 我们将证明如果式(3)在 τ_{n-1} 时刻成立, 则一直到 τ_n 时刻也成立(包括 τ_n), 也就是要证明:

$$P_j(\tau_{n-1}) \ge P(\tau_{n-1}) \qquad \forall j \in B(\tau_{n-1}) \Rightarrow$$
$$P_j(t) \ge P(t) \qquad \forall j \in B(t) \quad \tau_{n-1} < t \le \tau_n$$

文[6]中的引理1说明了在 $\tau_{n-1} < t < \tau_n$ 时上式成立, 只需要证明: $P_j(\tau_n) \ge P(\tau_n) \quad \forall j \in B(\tau_n)$

与第1步类似, 存在

$$F_{s(\tau_n)} \le P_j(\tau_n) \qquad \forall j \in B(\tau_n)$$

于是 $S^p(\tau_n) \le F_{s(\tau_n)} \le P_j(\tau_n) \qquad \forall j \in B(\tau_n)$,

此外, 从式(5)和式(6)可以看出在[τ_{n-1}, τ_n)期间 $P_i(t)$ 的增长速度等于 $P(t)$ 的增长速度, 所以有下式成立:

$$F_{s(\tau_n)} \ge \tau_n - \tau_{n-1} + P(\tau_{n-1}) = P(\tau_n -)$$

于是便有: $P(\tau_n) \le P_j(\tau_n) \qquad \forall j \in B(\tau_n)$

以上两步说明了引理1成立.

引理2 如果调度算法 S(其虚时间函数的计算根据2.2和2.3节的定义)满足条件 $S^P(t) > F_{服务}(t)$, 则此调度算法的流体流模型(FS)一定不属于RPS类算法.

证明 假设在[τ_{n-1}, τ_n)期间, 系统对连接 i 进行服务. 则在 τ_n 时刻, 由式(7)可知:

$$P(\tau_n) = \max(P(\tau_n -), S^p(\tau_n)) \ge S^p(\tau_n) > F_{服务}(\tau_n)$$
$$= P_i(\tau_n) \Rightarrow P(\tau_n) > P_i(\tau_n)$$

即系统虚时间函数值大于连接 i 的虚时间函数值, 违反了RPS算法的第二个条件(即不等式(3))所以FS不可能是RPS类算法. 命题得证.

定理1 系统虚时间函数的校正函数存在一个上界 $S^U(t) = F_{服务}(t)$. 也就是说, 当且仅当校正函数 $S^P(t)$ 满足 $S^P(t) \le S^U(t)$ 时, 此调度算法的流体模型才属于RPS类调度算法(根据文[7]的结果, 其相应分组级的调度算法具有同WFQ一样的时延特性, 且与校正函数 $S^P(t)$ 的具体形式没有关系).

证明 由引理1和引理2可直接得到.

定理2 设有调度算法 S(其虚时间函数的计算根据3.1和3.2节的定义), 如果其校正函数满足条件 $S^P(t) = F_{服务}(t) - \Delta(t)$, 且 $0 \le \Delta t \le \Delta_1$, 其中 Δ_1 为大等于零的常数, 则此算法的公平性指数(采用文[3]和[7]中相同的定义)为:

$$F_{i,j}^{(S)} = \max(f_{i,j}, f_{j,i})$$

其中: $f_{i,j} = \Delta P + C_i + (L_{max}/r_j) + (L_i/r_i)$; $C_i = \min((N-1)(L_{max}/r_i), \max_{1 \le n \le N}(L_n/r_n))$; $\Delta P = \max(l_j/r_j + l_i/r_i - l_i/r, 2l_i/r_i + \Delta_1 - l_i/r)$; L_{max} 为系统中的最大分组长度; N 为系统中的最大连接数. 为了方便对下面定理3的描述, 可以把公平性指数 $F_{i,j}^{(s)}$ 表示为 ΔP 的函数, 记为 $F_{i,j}^{(S)} = f(\Delta P)$.

证明 由定理1, FS属于RPS类. 设 ΔP 为系统虚时间函数小于活动连接的虚时间函数值的最大差值, 即 $\Delta P = \sup_j (P_j(t) - P(t)), j \in B(t)$. 下面分两个步骤来证明.

第一步 证明 $\Delta P = \max(l_j/r_j + l_i/r_i - l_i/r, 2l_i/r_i + \Delta_1 - l_i/r)$

考察系统中任意一段时间[τ_k, τ_{k+1}), 设连接 i 在这期间发送了一个分组 P_i(假设紧接这之前连接 j 发送了另一个分组 P_j, 长度为 l_j), 其长度为 l_i. 根据算法对 $P_i(t)$ 和 $P(t)$ 的计算方法可知: $P_i(t)$ 在[τ_k, τ_{k+1})期间增加 l_i/r_i, $P(t)$ 在[τ_k, τ_{k+1})期间增加 l_i/r; 再加上此算法属于RPS类调度器, 所以 $P_i(t)$ 的增加速度大于 $P(t)$ 的增加速度. 于是可以得到:

$$P_i(t) - P(t) \le P_i(\tau_k) - P(\tau_k) + l_i/r_i - l_i/r, \tau_k \le t < \tau_{k+1} \qquad (8)$$

现在分别考虑下面两种情况:

情形1 连接 i 在时刻 τ_k 之前已经有分组堆积, 这样便有 $P_i(\tau_k) \le P_j(\tau_k) \Rightarrow P_i(\tau_k) - P_j(\tau_k) \le 0$.

根据此算法的校正机制和校正函数的定义, 有 $P(\tau_k) = \max(P_j(\tau_k) - \Delta(\tau_k), P(\tau_k -)) \ge P_j(\tau_k) - \Delta(\tau_k)$.

代入式(8)可得到:

$$P_i(t) - P(t) \le P_i(\tau_k) - P(\tau_k) + l_i/r_i - l_i/r \le$$
$$P_i(\tau_k) - P_j(\tau_k) + \Delta(\tau_k) + l_i/r_i - l_i/r \Rightarrow$$
$$P_i(t) - P(t) \le l_j/r_j + l_i/r_i - l_i/r \quad (\tau_k \le t < \tau_{k+1})$$

情形2 连接 i 在 τ_k 时刻才有第一个分组到来, 这样根据式(1)便有

$$P_i(\tau_k) = \max(P_i(\tau_k -), P(\tau_k)) + l_i/r_i \qquad (9)$$

情形2.1 如果 $P_i(\tau_k -) \le P(\tau_k)$, 则由式(9)可得: $P_i(\tau_k) = P(\tau_k) + l_i/r_i$, 代入式(8)可得:

$$P_i(t) - P(t) \le 2(l_i/r_i) - l_i/r \qquad (\tau_k \le t < \tau_{k+1})$$

情形2.2 如果 $P_i(\tau_k -) > P(\tau_k)$, 则由式(9)可得: $P_i(\tau_k) = P_i(\tau_k -) + l_i/r_i$.

另外, 由此算法的校正机制和校正函数可得: $P(\tau_k) \ge S^P(\tau_k) \ge P_i(\tau_k -) - \Delta(\tau_k -)$, 那么有:

$$P(\tau_k) \ge P_i(\tau_k) - l_i/r_i - \Delta(\tau_k -) \Rightarrow P_i(\tau_k) - P(\tau_k)$$
$$\le l_i/r_i + \Delta(\tau_k -)$$

代入式(8)可得:

$$P_i(t) - P(t) \le 2l_i/r_i + \Delta(\tau_k -) - l_i/r \le 2l_i/r_i + \Delta_1 - l_i/r$$
$$(\tau_k \le t < \tau_{k+1})$$

综合情形1和情形2的结果, 便可得到:

$$\Delta P = \max(l_j/r_j + l_i/r_i - l_i/r, 2l_i/r_i + \Delta_1 - l_i/r)$$

第二步 利用文[7]中的定理4再结合第一步的结论可以直接得到定理2. 综合上面第一步和第二步的结果, 定理2得证.

定理3 设有调度算法 S(其虚时间函数的计算根据3.1

和3.2 节的定义), 另外有 $S^{U}(t)=E_{服务}(t)$, $S^{L}(t)=F_{服务}(t)-\Delta(t)$, 且 $0\le\Delta(t)\le\Delta_1$(其中 $\Delta_1\geqslant0$ 的常数), 则有如下结论: (1)当且仅当校正函数 $S^{p}(t)$满足条件 $S^{p}(t)\le S^{U}(t)$时, 此调度算法 S 的流体模型(f-S) 属于 RPS 类调度算法, 从而此调度算法的分组级模型(p-S)具有同 WFQ 一样的时延特性. (2)当校正函数 $S^{P}(t)$满足条件 $S^{L}(t)\le S^{P}(t)\le S^{U}(t)$时, 此 f-S 属于RPS 类调度算法, 而且 p-S 的公平性有保障. 设 $F_{i,j}^{(S)}$ 为其公平性指数, 则有 $F_{i,j}^{1}\le F_{i,j}^{(s)}\le F_{i,j}^{2}$. 而且当 $S^{P}(t)=S^{L}(t)$时, $F_{i,j}^{(s)}=F_{i,j}^{2}$; 当 $S^{P}(t)=S^{U}(t)$ 时, $F_{i,j}^{(S)}=F_{i,j}^{1}$. 其中 $F_{i,j}^{1}=f(\Delta P1)$, $F_{i,j}^{2}=f(\Delta P2)$, 函数$f(\cdot)$的定义见定理2. 这里

$$\Delta P1=\max(l_j/r_j+l_i/r_i-l_i/r,2l_i/r_i-l_i/r)$$

$$\Delta P2=\max(l_j/r_j+l_i/r_i-l_i/r,2l_i/r_i+\Delta_1-l_i/r)$$

证明　由定理 1 和定理 2 可以直接得到.

4　一个基于 ERPS 模型的调度算法

在这部分, 将给出 1 个基于 ERPS 模型的分组级队列调度算法(P-ERPS), 它的系统虚时间校正函数位于 $S^{P}(t)$的上限和下限之间. 通过同目前已经存在的 PFQ 调度算法进行比较, 可以得到如下结论: 此调度算法具有可以同 WFQ 相比拟的公平性和时延特性, 而且复杂度近似为 $O(1)$, 是目前综合性能较好的队列调度算法之一.

实例: 调度算法 $S1$(在文[9]中有详细的描述), 其校正函数为 $S^{P}(\tau_j)=S_{服务}(\tau_j)-\max\limits_{i\in B(\tau_j)}(L_i/r_i)$, (其中, $S_{服务}(t)$为 t 时刻正接受服务的连接的开始时间标签, 即此连接的第一个分组的开始时间标签)

由定理 3 可得: f-S1(流体模型) 属于 RPS 类, 而且 p-$S1$ (分组级模型) 公平性指数为 $F_{i,j}^{(S1)}=f(\Delta P1)$, 其中:

$$\Delta P1=\max(l_j/r_j+l_i/r_i-l_i/r,2l_i/r_i+\Delta_1-l_i/r)$$

$$\Delta_1=\max(2L_i/r_i)(因子\ F_{服务}(t)=S_{服务}(t)+l_i/r_i)$$

注: 也可利用定理 2 的证明方法直接证明 $S1$ 算法的公平性.

在已提出的 PFQ 算法中, MD-SCFQ (Minimum-Delay SCFQ)[6]是目前综合性能较好的一种算法, 其系统虚时间函数的计算复杂度为 $O(1)$, 且公平性及时延特性皆与 WFQ 比较接近; 但是在计算其系统虚时间函数时, 需要执行一次除法和最多 N 次乘法及 N 次加法运算, 这给调度算法在高速路由器中的硬件实现带来了很大的困难. $S1$ 算法在计算系统虚时间函数时虽然涉及到比较操作, 但当业务流变化(到来和离去)较慢时, 其计算复杂度可近似为 $O(1)$, 另外 $S1$ 算法不需要任何除法和乘法运算, 带来的好处是更快的分组调度速度和更容易的硬件实现; $S1$ 算法是目前所有 PFQ 算法中综合性能较好的算法之一.

5　结论

本文提出了增强的速率比例调度器模型(Enhanced RPS). ERPS 模型给出了设计系统虚时间函数的具体方法和条件, 并通过较为严密的理论分析和证明, 得到了系统虚时间校正函数的上界和下界及其相应的时延特性和公平性指数. 在设计具体的分组级队列调度算法时, 我们可以在上界和下界范围内寻找合适的校正函数, 以在算法的公平性和复杂性方面取得一个较好的折衷. 最后, 我们还给出了 1 个具体的基于 EPRS 模型的分组级队列调度算法, 在目前已经存在的 PFQ 算法中, 此算法具有更好的综合性能, 运算简便且易于实现. 很适合应用到目前高速路由器的设计之中.

参考文献:

[1] A K Parekh, R G Gallager. A generalized processor sharing approach to flow control in integrated service networks: the single-node [J]. IEEE/ACM Trans. Networking, 1993, 1(3): 344-357.

[2] H Zhang. Service disciplines for guaranteed performance service in packet switching networks [J]. Proceedings of the IEEE, Oct 1995, 83(10).

[3] S J Golestani. A self-clocked fair queuing scheme for broadband applications [C]. IEEE INFOCOMM, 1994, 2: 636-645.

[4] J C R Bennett, H Zhang. WF^{2}Q: Worst-case fair weighted fair queuing [C]. IEEE INFOCOM'96, San Francisco, CA, 1996, 3: 120-128.

[5] J C R Bennett, H Zhang. Hierarchical packet fair queuing algorithms [J]. IEEE/ACM Trans. On Networking, Oct. 1997, 5: 675-689.

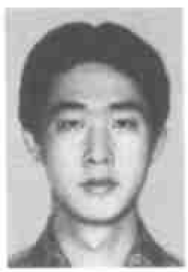
[6] F M Chiussi, A Francini. Minimum-delay self-clocked fair queuing algorithm for packet-switched networks [C]. IEEE INFOCOMM, 1998, 3: 1112-1121.

[7] D Stiliadis, A Varma. Rate-proportional servers: a design methodology for fair queuing algorithms [J]. IEEE/ACM Transactions on Networking, April 1998, 6(2): 164-173.

[8] P Goyal, H M Vin. Start-time fair queuing: a scheduling algorithm for integrated services packet switching networks [J]. IEEE/ACM Transactions on Networking, October 1997, 5(5): 690-703.

[9] W Chonggang, et al. A simple weighted fair queue scheduling algorithms [C]. IEEE ICC2001, Helsinki, Finland, June 2001.

作者简介:



王重钢　男. 1974年 5 月出生于四川, 北京邮电大学交换技术与通信网国家重点实验室博士研究生. 1996 年毕业于西北工业大学电子工程系, 1999 年 4 月在电子科技大学获得工学硕士学位, 1999 年 9 月考入北京邮电大学攻读博士研究生. 目前研究方向为网络流量控制和拥塞控制、队列管理和队列调度及无线网络服务质量.

隆克平　男. 1968 年 5 月出生于四川省通江县, 1998 年获电子科技大学博士学位, 1998 年 9 至 2000 年 8 月, 北京邮电大学通信网国家重点实验室博士后研究, 现为该室副教授、硕士生导师. 主持和承担过国家级、省部级及国际合作项目 8 项. 发表学术论文近 50 篇, 其中以第一作者近 30 篇. 主要研究方向: SDH/ATM 网络生存性、TCP/IP 协议改进机制及性能分析、增强 Internet 实时多媒体业务 QoS 保障的策略及其实现机制、IP/ATM 综合技术、移动 IP 技术及应用、路由器的队列调度和缓存管理策略及算法等.