第 10 卷第 10 期
1999 年 10 月

软件学报
JOURNAL OF SOFTWARE

Vol.10, No.10
Oct. 1999

一个基于差值有界的通信量调度算法*

孙利民 窦文华 龚正虎 周兴铭

(国防科学技术大学计算机系 长沙 410073)

摘要 调度算法决定不同应用的包在输出链路上的发送顺序,是网络提供 QoS 服务的关键技术.该文给出了系统虚时钟和连接虚时钟差值有界的调度算法框架 VDBF(virtual clock difference-bounded framework).在此基础上,采用周期修正系统虚时钟的方法和层次式排序结构,提出了简单公平的调度算法 DFQR(difference-based fair queueing by re-calibration).此调度算法在网络延迟、调度公平性和实现复杂度之间达到了很好的折衷,理论证明和模拟显示出其性能的优越性.

关键词 调度算法,虚时钟差值有界,周期修正,层次结构.

中图法分类号 TP393

调度算法决定不同应用的包在输出链路上的发送顺序,是网络提供 QoS 服务保证的关键技术之一[1].实用的调度算法必须具有低网络延迟、良好的调度公平性和低实现复杂度等优点,而目前提出的大量调度算法未能同时具有这些优点或在以上几方面之间达到很好的折衷[2].我们在本文提出的系统虚时钟和连接虚时钟差值有界调度算法框架 VDBF(virtual clock difference-bounded framework)的基础上,采用周期修正系统虚时钟的方法和层次式排序结构,提出了简单公平的调度算法 DFQR(difference-based fair queueing by re-calibration).它在网络延迟、调度公平性和实现复杂度之间达到了很好的折衷,具有较强的实用性.

1 差值有界的调度算法框架 VDBF

定义 1. 系统忙周期是指结点机一直处于工作状态的最大时间间隔.在系统忙周期内,结点机连续发送各连接的包.

定义 2. 连接累积周期是指连接在结点机中连续有包等待服务的最大时间间隔.处于累积周期的连接称为累积连接.

定义 3. 连接 i 在 $(t_1,t_2]$ 期间内获得的服务量 $W_i(t_1,t_2)$ 与预约带宽 ρ_i 之比,称为连接 i 在 $(t_1,t_2]$ 期间获得的规格化服务,用 $w_i(t_1,t_2)$ 表示.

我们在 VDBF 框架[3]中设置一个系统虚时钟 $P(t)$ 来描述系统状态,每个连接 i 设置一个连接虚时钟 $P_i(t)$ 来描述连接状态.其流模型描述如下.

(1) 设输出链路的带宽为 C,共享输出链路的连接个数为 N,连接 i 的预约带宽为 ρ_i,为保证网络的稳定性,存在关系 $\sum_{i=1}^{N}\rho_i \le C$.

(2) 每个连接的虚时钟描述在 t 时刻连接的状态,它具有以下特性:

(a) 当系统不处于忙周期时,连接 i 的虚时钟设置为零;连接 i 不处于累积周期时,连接 i 的虚时钟保持

---

* 本文研究得到国家自然科学基金资助.作者孙利民,1966 年生,博士后,主要研究领域为高速计算机网络,实时通信.窦文华,1946 年生,教授,博士生导师,主要研究领域为高性能网络技术,FMS.龚正虎,1945 年生,教授,博士生导师,主要研究领域为网络互连,QoS 技术.周兴铭,1938 年生,教授,博士生导师,中国科学院院士,主要研究领域为高性能机体系结构,高速计算机网络,分布式数据库.

本文通讯联系人:孙利民,北京 100080,中国科学院软件研究所

本文 1998-07-13 收到原稿,1998-10-16 收到修改稿

不变;

(b) 当连接 i 在 t_1 时刻进入累积周期时,它的虚时钟为 $P_i(t_1)=\max\{P(t_1-),P_i(t_1-)\}$;

(c) 在累积周期 $(t_1,t_2]$ 期间,连接 i 在 $t\in(t_1,t_2)$ 时的虚时钟为 $P_i(t_1)=P_i(t_1)+w_i(t_1,t)$.

(3) 系统的虚时钟描述在 t 时刻系统的状态,它具有以下特性:

(a) 当系统不处于忙周期时,系统虚时钟设置为零;在系统忙周期的任意区间 $(t_1,t_2]$,

$$P(t_2)-P(t_1)\ge t_2-t_1;$$

(b) 在任何时刻系统虚时钟都不大于任何处于累积周期的连接虚时钟,如果 $B(t)$ 表示 t 时刻处于累积周期的连接集合,那么

$$P(t)\le \min P_i(t),\quad i\in B(t);$$

(c) 在任何时刻,系统虚时钟与任何处于累积周期的连接虚时钟之差有界,即

$$P_i(t)-P(t)\le \Delta P,\quad i\in B(t),\qquad \Delta P\text{ 为常量}.$$

(4) 系统选择累积连接服务策略为:系统在 t 时刻只服务虚时钟最小的累积连接集合,该集合各连接获得的服务速率正比于它们的预约带宽.

流模型中多个连接以不同速率可同时获得服务,而包模型中任意时刻只能有一个包获得服务.在 VDBF 相应的包模型系统中,每个包带有虚拟完成时间,包的虚拟完成时间是包的最后一位离开流模型系统时连接的虚时钟值,系统选择虚拟完成时间最小的到达包服务.可以验证 WFQ[4]和 VC[5]调度算法符合 VDBF 框架.

由 VDBF 框架构造的包调度算法称为 PDFQ(packet_by_packet difference_based fair queueing)调度算法.PDFQ 调度算法具有如下特性.

性质 1. 在 PDFQ 服务器中,$L_{\max}$ 表示所有连接中最大包的长度,如果连接 i 的到达通信量符合漏桶模型 (σ_i,ρ_i),那么,连接 i 的包延迟满足 $D_i\le\sigma_i/\rho_i+\dfrac{L_{\max}}{C}$.

定义 4(调度算法公平性). 在系统忙周期内,对于任意时间间隔 (t_1,t_2) 内连续处于累积周期的任意两个连接 i 和 j,如果连接 i 和 j 在 (t_1,t_2) 期间获得的规格化服务之差有界,即 $\left|\dfrac{W_i(t_1,t_2)}{\rho_i}-\dfrac{W_j(t_1,t_2)}{\rho_j}\right|\le F$ (F 为常量),那么调度算法具有公平性,F 值越小公平性就越好,理想情况 F 等于零.

如果 L_i 表示连接 i 最大包长度,$\hat{W}_i(0,t)$ 表示连接 i 在 $(t_1,t_2]$ 期间内获得的包括非完整包服务量.PDFQ 算法的调度公平性如下.

性质 2. 在 PDFQ 服务器中,如果连接 i 和 j 在 τ 时刻后一直处于累积周期,那么对于 τ 时刻后的任意时间间隔 $(t_1,t_2]$,满足关系:

$$\left|\frac{\hat{W}_j^{\,P}(t_1,t_2)}{\rho_j}-\frac{\hat{W}_i^{\,P}(t_1,t_2)}{\rho_i}\right|\le \max\left(\Delta P+C_j+\frac{L_{\max}}{\rho_i}+\frac{L_j}{\rho_j},\Delta P+C_i+\frac{L_{\max}}{\rho_j}+\frac{L_i}{\rho_i}\right).$$

其中 $C_k=\min\left((N-1)\dfrac{L_{\max}}{\rho_k},\max\limits_{1\le n\le N}\left(\dfrac{L_n}{\rho_n}\right)\right)k=i,j$, ΔP 是系统虚时钟与连接虚时钟之差的最大值.

2 DFQR 调度算法

VDBF 框架规定了系统虚时钟和连接虚时钟的特性以及连接虚时钟的计算方法.符合 VDBF 框架的调度算法具有相同的低网络延迟,但调度公平性存在差异.根据 VDBF 框架具体构造调度算法的核心问题是系统虚时钟的计算,使用不同的系统虚时钟计算方法,算法的实现复杂度和调度公平性不同.调度算法复杂度主要包括包的虚拟完成时间计算以及包的排序和选择发送.系统虚时钟的计算决定虚拟完成时间计算复杂度.基于实现简单和调度公平考虑,我们采用如下周期修正系统虚时钟的方法.

2.1 系统虚时钟的周期性修正方法

设置 $\Delta D=\max\limits_{1\le i\le N}\{L_i/\rho_i\}$,$\Delta D$ 是包模型下连接虚时钟差值的最小上界.

如果系统虚时钟按真实时间增加,那么它的计算很简单,但存在系统虚时钟与连接虚时钟之差可大于任意值而造成调度公平性差的问题.为了计算系统虚时钟简单和调度公平性好,系统虚时钟通常按照真实时间增加,当系统虚时钟与连接虚时钟之差超过一定值时,及时修正系统虚时钟.在包模型下,采用独立式方法*计算系统虚时钟实现简单,系统虚时钟的修正只能在包到达或离开时发生.

假设系统虚时钟在系统忙周期开始时设置为零,而且按照真实时间增加,连接虚时钟的计算和选择连接的服务按照 VDBF 框架规定.在其相应的包模型下,每个包的虚拟完成时间等于在流模型下包的最后一位完成服务时连接的虚时钟值,结点机选择虚拟完成时间最小的包发送.在包模型下发送完成包时刻(记为 T_f 时刻),用 p_h 表示所有累积连接中未获得服务的虚拟完成时间最小的包,F_{p_h} 表示 p_h 的虚拟完成时间,存在下面的关系.

定理 1. 在包模型下,在发送完成包 T_f 时刻存在如下关系式:

$$P_i(T_f) \ge F_{p\_h} - \Delta D, \qquad \forall i \in B(T_f).$$

证明:反证法.假设累积连接 j 的虚时钟值不满足上述关系式,即

$$P_j(T_f) < F_{p\_h} - \Delta D, \qquad j \in B(T_f).$$

在 T_f 时刻,设连接 j 在结点机中未获得服务的最先到达的包为连接 j 的第 k 个包,那么,

$$F_{k,j} = P_j(T_f) + \frac{L_{k,j}}{\rho_j} \le P_j(T_f) + \Delta D,$$

因此 $F_{k,j} < F_{p\_h}$.　□

这与已知 p_h 是未获得服务的虚拟完成时间最小包矛盾,即证.

在包模型下,在发送完成包时刻 T_f,如果 $F_{p\_h} > P(T_f-) + \Delta D$ (此 T_f 时刻记为 T_R 时刻),由定理 1 可知,所有累积连接的虚时钟值都大于 $(F_{p\_h} - \Delta D)$,而系统虚时钟 $P(T_f-)$ 小于 $(F_{p\_h} - \Delta D)$,可将系统虚时钟修正为 $P(T_f) = F_{p\_h} - \Delta D$.周期修正系统虚时钟的方法具体描述为:

(1) 在系统忙周期开始,设置系统虚时钟为零;

(2) 在每个包发送完成时刻 T_f,判断结点机中未服务包的最小虚拟完成时间是否比系统虚时钟值大 ΔD,即 $F_{p\_h} > P(T_f-) + \Delta D$,若关系式成立,修正系统虚时钟为 $P(T_f) = F_{p\_h} - \Delta D$,否则不修正;

(3) 其他时刻 $P(t) = P(t_1) + t - t_1$, $(t_1, t]$ 属于系统忙周期.

定理 2. 在包模型下,采用周期修正的计算系统虚时钟方法,在发送完成包时刻 T_f,存在关系:

$$P_i(T_f) < P(T_f) + \Delta D, \qquad \forall i \in B(T_f).$$

证明:反证法.假设 $P_j(T_f) \ge P(T_f) + \Delta D$, $j \in B(T_f)$,在时刻 T_f,连接 j 必存在一个已经发送完成的包(设为第 k 个包),它的虚拟完成时间为 $P_j(T_f)$.考虑连接 j 第 k 个包在时刻 $t(t<T_f)$ 成为 p_h,此时,

$$F_{p\_h} = P_j(T_f) \ge P(T_f) + \Delta D.$$

由于 $P(t) < P(T_f)$ $(t<T_f)$,因此 $F_{p\_h} > P(t) + \Delta D$,在 t 时刻需要修正系统虚时钟为

$$P(t) = F_{p\_h} - \Delta D = P_j(T_f) - \Delta D, \qquad (t \text{ 时刻})$$

即

$$P_j(T_f) = P(t) + \Delta D,$$

所以, $P_j(T_f) < P(T_f) + \Delta D$, $P(t) < P(T_f)$.

这与已知假设相矛盾,即证.　□

由定理 1 和定理 2 可知:在包模型下,如果采用周期修正的计算系统虚时钟方法,那么在发送完成包时刻 T_f,所有累积连接的虚时钟值都小于 $(P(T_f) + \Delta D)$ 且都大于 $(F_{p\_h} - \Delta D)$,系统虚时钟与连接虚时钟之差小于

* 独立式(self-contained)方法是在包到达或包离开时,根据连接 i 的虚时钟和系统的虚时钟值计算新的系统虚时钟及连接 i 的虚时钟,不需要其他连接的虚时钟值,具有实现简单的优点.SCFQ(self-clocked fair query)调度算法[6]中采用此方法.

ΔD.

在相应流模型下的调度算法中,连接虚时钟最小的所有累积连接同时获得服务,所有累积连接的虚时钟值趋近相等.在包模型下,后到达的虚拟完成时间小的包不能剥夺正在发送的包.在流模型下,后到达(包)的虚时钟值小的累积连接可剥夺正在发送的连接.因此,对于任意给定时刻,流模型下最小的累积连接虚时钟不小于包模型下该连接的虚时钟值;流模型下最大的累积连接虚时钟不大于包模型下该连接的虚时钟值;流模型下累积连接虚时钟之间的差值比在包模型下小.因为包模型下 T_R 时刻,所有累积连接的虚时钟值都小于 $(P(T_R)+\Delta D)$ 值且都大于 $(F_{p\_h}-\Delta D)$ 值,所以,在流模型下 T_R 时刻,所有累积连接的虚时钟值也都小于 $(P(T_R)+\Delta D)$ 值且都大于 $(F_{p\_h}-\Delta D)$ 值,此时可将系统虚时钟修改为 $(F_{p\_h}-\Delta D)$.这样修改系统虚时钟,一方面符合 VDBF 框架的条件,另一方面保持两种模型下系统虚时钟相同.

定理 3. 对于流模式下的调度算法,如果按照 VDBF 框架计算连接虚时钟及选择连接服务,采用周期修正的系统虚时钟计算方法,那么:

1° 在系统忙周期的任何区间 $(t_1,t_2]$, $P(t_2)-P(t_1)\ge t_2-t_1$;

2° 在系统忙周期的任意时刻 t, $P(t)\le P_i(t)$, $\forall i\in B(t)$;

3° 在系统忙周期的任意时刻 t, $P_i(t)-P(t)\le \Delta D$, $\forall i\in B(t)$.

证明:1° 分两种情况:

(1) 在 $(t_1,t_2]$ 期间,没有修正系统虚时钟,那么 $P(t_2)=P(t_1)+t_2-t_1$,满足关系式.

(2) 在 $(t_1,t_2]$ 期间,如果在时刻 $\tau_1,\tau_2,...,\tau_m$ 修正了系统虚时钟,也就是在 $\tau_k(k=1,2,...,m)$ 时, $P(\tau_k)=F_{p\_h}-\Delta D>P(\tau_k-)$,那么 $P(t_2)=P(\tau_m)+t_2-\tau_m>P(\tau_m-)+t_2-\tau_m>P(\tau_{m-1}-)+t_2-\tau_{m-1}$,递推可得 $P(t_2)>P(\tau_1-)+t_2-\tau_1>P(t_1)+t_2-t_1$,满足关系式.

2° 反证法.假设在系统忙周期 t 时刻有 $\exists i\in B(t)$, $P(t)>P_i(t)$,考虑离 t 最近的时刻 $t-\Delta t$, $P(t-\Delta t)\le P_i(t-\Delta t)$,连接 i 在 $(t-\Delta t,t]$ 期间处于累积周期.分两种情况:

(1) 在 $(t-\Delta t,t]$ 期间没有修正系统虚时钟值.连接 i 在 $(t-\Delta t,t]$ 期间是虚时钟值最小的连接,它获得服务的速率不小于 ρ_i,因此, $P_i(t)\ge P_i(t-\Delta t)+\dfrac{\rho_i\cdot\Delta t}{\rho_i}\ge P_i(t-\Delta t)+\Delta t\ge P(t)$,这与假设相矛盾,所以 $P(t)<P_i(t)$.

(2) 在 t 时刻修正系统虚时钟,则存在关系式: $P(t-)<F_{p\_h}-\Delta D\le P_i(t-)$,系统虚时钟修正为 $P(t)=F_{p\_h}-\Delta D$,而 $P_i(t)=P_i(t-)$,所以 $P(t)\le P_i(t)$.

3° 由周期修正系统虚时钟方法可知:连接与系统的虚时钟之差最大值是在修正系统虚时钟的时刻之前.如果在时刻 t 修正系统虚时钟值,则存在关系式:

$$P(t-)<F_{p\_h}-\Delta D\le P_i(t-)<P(t-)+\Delta D,\qquad \forall i\in B(t-),$$

因而

$$P_i(t-)-P(t-)\le \Delta D,\qquad \forall i\in B(t-),$$

所以,对于系统忙周期的任意时刻 t, $P_i(t)-P(t)\le \Delta D$, $\forall i\in B(t)$.

推论. 在流模型下,如果采用周期修正系统虚时钟的方法,并按照 VDBF 框架计算连接虚时钟及选择连接服务,那么该调度算法符合 VDBF 框架,且 $\Delta P=\Delta D$.

2.2 层次式优化排序策略

对每个连接而言,按照包到达的先后顺序服务,只需连接在结点机中未服务的最先到达的包参加排序.我们采用如下层次式排序策略:每个连接设置一个等待队列,按照到达先后顺序存放属于该连接的包;系统设置一个公共等待队列,所有连接在结点机中未服务的最先到达的包存放到公共等待队列中,它们按照虚拟完成时间从小到大排序,这样,包的排序复杂度为 $O(\log_2 N)$.结点机选择公共等待队列中虚拟完成时间最小的包发送,如果发送包对应的连接队列中仍有包,计算它第 1 个包的虚拟完成时间后插入公共等待队列中.

2.3 DFQR调度算法

在具体给出算法之前,首先说明以下几点:*pub_queue* 表示按照虚拟完成时间从小到大排序包的公共等

待队列;

i_queue 表示连接 i 按照到达先后排序包的等待队列; $F_{p\_h}(t)$ 表示 t 时刻 *pub_queue* 中未服务包的最小虚拟完成时间; F_i 表示当前时刻连接 i 最近发送包的虚拟完成时间,系统忙周期开始设置为零.

系统虚时钟计算:

(1) 在系统忙周期开始时刻,设置系统虚时钟 $P(t)$ 为零;

(2) 在系统忙周期内除修正系统虚时钟 $P(t)$ 时刻外,系统虚时钟 $P(t)$ 按真实时间增加.

包离开算法: 假设连接 i 的第 k 个包 t 时刻发送完成.

(1) 修改系统虚时钟值:

```
if ( pub_queue 为空)
   then { P(t)=0;
          F_i = 0; ∀i∈C(t), C(t)是共享输出链路的连接集合
          exit() }
   else { if ( F_p_h > P(t)+ΔD )
           then  P(t) = F_p_h − ΔD
          endif }
   endif
```

(2) 设 *pub_queue* 中连接 j 第 m 个包的虚拟完成时间最小,选择该包发送;

(3) $F_j = F_{m,j}$

(4) if (*j_queue* 非空)

then {计算连接 j 第$(m+1)$包的虚拟完成时间 $F_{(m+1),j} = F_j + \dfrac{L_{(m+1),j}}{\rho_j}$;

根据虚拟完成时间将连接 j 第$(m+1)$个包插入 *pub_queue* 中}

endif

包到达算法: 假设连接 i 第$(k+1)$个包在 t 时刻到达.

if ((*i_queue* 为空) and (*pub_queue* 没有连接 i 的包))

then { $F_{(k+1),i} = \max\{F_i, P(t)\} + \dfrac{L_{(k+1),i}}{\rho_i}$;

根据虚拟完成时间将该包插入 *pub_queue* 中}

else 根据到达顺序将包存放在 *i_queue* 中

endif

3 与其他调度算法的比较与性能模拟

3.1 与其他调度算法的比较

从表 1 可以看到 DFQR 算法的延迟边界与 WFQ 相同,实现复杂度与 VC 相同,调度公平性也比 PFFQ[3] 调度算法好.

表 1　调度算法的延迟、公平性和实现复杂度比较

调度算法	延迟	公平性	复杂度
WFQ	$\sigma/\rho_i+L_{max}/C$	$F_{i,j}(0)$	$O(N)$
VC	$\sigma/\rho_i+(L_{max}/C)$	∞	$O(\log N)$
SCFQ	$\sigma/\rho_i+L_{max}/C(N-1)$	$L_i/\rho_i+L_j/\rho_j$	$O(\log N)$
DFQR	$\sigma/\rho_i+L_{max}/C$	$F_{i,j}(\Delta D)$	$O(\log N)$
PFFQ	$\sigma/\rho_i+L_{max}/C$	$F_{i,j}(2\cdot\Delta D)$	$O(\log N)$

表中,
$$F_{i,j}(\Delta P) = \max\{\Delta P + C_j + L_{max}/\rho_i + L_j/\rho_j, \Delta P + C_i + L_{max}/\rho_j + L_i/\rho_i\},$$

$$C_k = \min((N-1)^{L_{\max}}/_{C,\max_{1\le n\le N}}(^{L_n}/_{\rho_n})), k=i,j.$$

3.2 性能模拟公平性

延迟模拟模型为 8 个连接共享输出链路,连接的通信量符合 ON_OFF 模型并经过漏桶整形.表 2 给出调度算法在包长度为 1(模拟 ATM)时连接最大延迟统计结果.从表 2 看到:各连接的最大延迟小于理论上的延迟边界;占带宽 50%以上的连接 1 在 DFQR 调度算法中的最大延迟介于 SCFQ(PFFQ)与 WFQ(VC)之间;预约带宽是控制连接延迟的关键因素,通过稍微调整连接的预约带宽就可改变连接的延迟.

表 2 调度算法最大延迟统计结果(σ_i=2,$L_{\max}$=1)

连接	预约带宽	到达速率	VC 等理论值	SCFQ 理论值	VC	WFQ	DFQR	PFFQ	SCFQ
1	0.500 000	0.498	5	11	1.80	1.60	1.80	4.00	7.20
2	0.062 500	0.061	33	39	20.00	18.00	23.00	22.00	23.00
3	0.062 500	0.062	33	39	22.00	19.00	24.00	25.00	24.00
4	0.062 500	0.062	33	39	23.00	21.00	25.00	27.00	31.00
5	0.078 125	0.076	27	33	12.40	12.60	16.40	18.40	17.60
6	0.078 125	0.076	27	33	14.80	13.20	17.80	19.40	18.60
7	0.078 125	0.076	27	33	15.80	13.60	18.80	20.40	19.60
8	0.078 125	0.076	27	33	19.80	15.60	19.80	21.20	21.60

通过比较在同一段时间内连接获得的规格化服务之差评价调度算法的公平性,规格化服务之差最大值通常发生在有连接刚进入累积周期后的时间段.为了测试算法的公平性,我们将连接分为奇数和偶数两组,两组中对应连接的预约带宽相同,各连接的通信量较大以保证在任何时刻连接都处于累积周期.奇数组连接包在 t=0 开始到达,偶数组连接包在 t=1000 开始到达,统计每 100 个时间单位时各连接发送包的个数,连接 1 和连接 2 在时间间隔(700,1800]内的统计结果如图 1 所示.图中显示出 VC 调度的公平性差,其他调度算法的结果基本相近.近一步的实验说明 DFQR 算法的公平性介于 PFFQ 与 WFQ 之间.

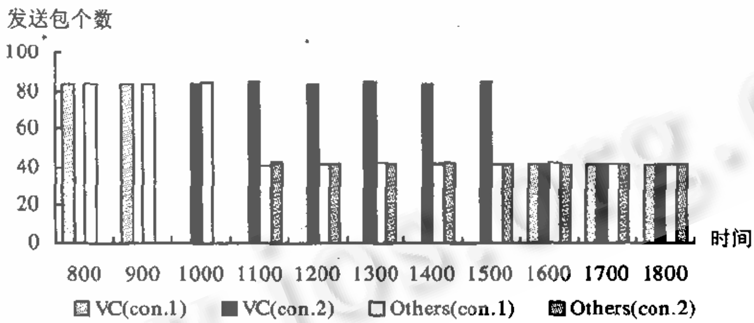

图 1 连接 1 和连接 2 在不同调度算法下发送包的个数(时间间隔:100s)

4 小 结

本文在差值有界的调度算法 VDBF 框架基础上,通过周期修正系统虚时钟,采用层次式优化排序结构,并提出了 DFQR 调度算法.理论分析和性能模拟显示,DFQR 调度算法具有低网络延迟、较好的调度公平性和实现简单的优点,在网络延迟、公平性和实现复杂度之间达到好的折衷.PFFQ 调度算法已在 ATM 交换机中用硬件实现.DFQR 调度算法比 PFFQ 算法公平性好,实现更为简单,这说明 DFQR 调度算法完全可以用硬件实现,用在高速的结点机中,是一个较为实用的高性能调度算法.

参考文献

1 Zhang Hui. Service disciplines for guaranteed performance service in packet-switching networks. Proceedings of IEEE, 1995,83(10):1373~1396

2 Stiliadis D, Varma A. Frame-based fair queueing: a new traffic scheduling algorithm for packet-switched networks. Technical Report, UCSC-CRL-95-39, 1995. http://www.cse.ucsc.edu/research/hanlab/publications/

3 孙利民.有界延迟实时服务网络的研究[博士学位论文].国防科学技术大学计算机系,1998
(Sun Li-min. Real-time networks with bound-delay service guarantees [Ph.D. Thesis]. National University of Defense Technology, 1998)
4 Parekh A, Gallager R. A generalized processor sharing approach to flow control in integrated services networks: the single-node case. IEEE/ACM Transactions on Networking, 1993,1(3):344~357
5 Figueira N R, Pasquale J. An upper bound on delay for the virtualclock service discipline. IEEE/ACM Transactions on Networking, 1995,3(4):399~408
6 Golestani S J. Network delay analysis of a class of fair queueing algorithms. IEEE Journal on Selected Areas in Communication, 1995,13(8):1057~1570

A Difference-based Traffic Scheduling Algorithm

SUN Li-min DOU Wen-hua GONG Zheng-hu ZHOU Xing-ming

(*Department of Computer Science National University of Defense Techonolgy Changsha* 410073)

Abstract Packet scheduling algorithm controls the order of packet service, which is crucial in the design of a QoS network. In this paper, the authors first give a virtualclock difference-bounded framework (VDBF), then present a new scheduling algorithm——DFQR (difference-based fair queueing by re-calibration), in which they use a hierarchical structure, recalibrate the system virtual clock periodically, compromise network delay with fairness and implementation complexity. Finally, they demonstrate the performance benefit of DFQR by analysis and simulation.

Key words Scheduling algorithm, virtualclock difference-bounded, recalibrate periodically, hierarchical structure.