

CB HSTCP: Ú Î © Ž Ĩ ¥ Ü TCP Ø E

1 1 1 2
~ v Ð 9 Ø Ð ý, „ ~ s Ü 310027; 2 s Ü È 0 S / v Ð 9 Ø Ð ý, „ ~ s Ü 310027)

L ? C HSTCP i ø × ¥ RIT (Round Trip Time), Ü Ÿ, 6 “ ? C š Ö Hi v ¥
Ü C` É) s , i s é > Ü 5.4 CB HSTCP Ø E, “ CW Ä „ Block
„ † © é > q. 1™ Ø E Ÿ?, 4 B ñ M Ü Ÿ S . Y V s i P “ ns?
£ HSTCP Ÿ? ¥ \$, 3 % RIT , Ü Ÿ Ü 5 a† © é > q.] H µ

Ü Î © Ž; TCP; HSTCP; š Ö e Ä

TP393 Ó D S M': A Ó c l | : 0372 2112 (2005)

~ : ° “ ~ 1/Ⅲ “ ~ fl · fl Ž ž! “ “ # S “

SU Fan jun¹, PAN Xue zeng¹, CAI Lian

(1 College of Computer Science, Zhejiang University, Hangzhou 310027)

(2 College of Computer Science, Hangzhou Dianzi University, Hangzhou 311126)

) * 20/4% experiments, it is found that HSTCP has
sion event. All the RIT unfairness in HSTCP
des CW and algorithm to alleviate RIT
proposed algorithm. The criterion is pro
fair problem
, “ - & ' f “ , TCP; i

ý
" © Ž Ü E Ä ; 8 / Ø E ¥ ? Z , C Ÿ ¥ {
1 ¥ Ü Î © Ž , v 7 P 3 1 ` Ü v “ ¥ Ü Ÿ
ö î 1 ? . B t < “ Ä S Ð x T a ù ñ • „ L H _ ©
V e “ Ü Î © Ž v ù ñ © Ä Ĩ ' a Ä r . { Ü { z L Ä
am ^ Ä . 6 “ Ä Z “ © ì ñ ½ (data intensive
gr application Ü Î © Ž [2] 9 ü 3 1 Ü Î © Ž . { v

Ü Ä { x p, “ - Inter
1 100bps Ž
T
© Z
AIMD (additive increase multiplicative decrease) { { ,
9 F α , ' š Ö ? 3 H , 3 g h l , f ü • Ä v š
3 g š Ö ? 3 a ë - α , K Ä | b = { z ¥ ? ĩ , ý
5 M' É B H W = æ “ q © [4] . L ç RIT 1 100ms,
v 1 1500byte, 4 ¥ { z 1 10Gbps, 5 3 1 ¥ 3 g v l

3 1 Ø «
Hu 7 ñ
ç õ α “ , ö
Ä ^ © ¿ š Ö e Ä
TCP (STCP) [7] , XCP [8] a
Ä . f Ä K Ä “ Z Ÿ .
“ Z Ÿ a ¶ z Ÿ
a , BIC ? Ä Ĩ q ¥ 9 É , Ä H “ STCP v . HSTCP “
STCP ^ S TCP \$ É) ¥ ¿
v ¿ B ç ¥ • ' a , P 3 g 9 F y 0 a (a
î 1 3 g v l w ¥ f “ . HSTCP Z Ä “ -
“ .

' P " RED ^ © H, HSTCP μ z ¥ Ÿ ?^[10], ÇE ^ DT (Drop Tail) ^ © @ ÷ L C e †, © Ž Ĩ P " ¥ ÷ F < W. ' P " DT(Drop Tail) ^ © H, ÓD[9] · HSTCP i ø x ¥ RTT, Ü Ÿ · RTT, Ü Ÿ · é s } B { z ¥ " @ ¥ { Î q 1, Ü · ÓD[4] · HSTCP i @ ÷ H Y ¥ " l ? „ HSTCP " Z Ÿ î / î ¥ v é >, ÇE Å μ é > ¥ Ø É) W % s .

' Ó Y V L £ L [Ù 5. Ú a • " ~ s RTT, Ü Ÿ ¥ ð y, i ð HSTCP ¥ ? Å Ø 3 d HSTCP š Ö ? 3 H ¥ é > Ø. N \$ 4 HSTCP ¥ ÷ Ê Ø E CBHSTCP (Constant window and Block pacing HSTCP). ¼ Ø E " CW Å „ Block pacing Å · CW Ø E Y V @ F Ü y 0 h " RTT 3 g 9 F ¥ • Y, h ^ RTT μ s { Ÿ ¥ ø x, Ü Ÿ, P RTT,] ¥ @ - W ¥ { μ s û RTT ' ¥ Q 1 S ¶ =, V 7 | x ç ¥ Ü Ÿ. block pacing Å V [h @ ÷ HSTCP ¥ " Z Ÿ ¥ 3 g 9 F „ v š Ö 3 g H Y l ? î „ ¥ é > " „ † @ é > q. 1 TM Ø E ¥ Ÿ ?, ' Ó » 6 † s Î 4 B ñ M Ü Ÿ S " ÷ ß N 4 Ø E ¥ Ÿ ?, i Ú a " ns2^[11] É) E L £ 4 Ø E ¥ Ÿ ?.

' Ó Å † s ½ † Á / : Ó Î » 2 « e 1 ° TCP „ HSTCP: » 3 « " L · RTT, Ü Ÿ „ é > Ù 5, " ~ s RTT, Ü Ÿ, i ð ? Å Ø s é > Ù 5; » 4 « W % í • ÷ É Z Å; » 5 « s Ø E ¥ l ù Ÿ „ x ç Ÿ; » 6 « É) L s . K a ^ 9 2 „ Z -.

2 M1 ù î : TCP „ HSTCP

' - TCP μ „ ñ ' , Ĩ TCP Reno < " K < W „ Ó ë 1 S TCP. S TCP „ HSTCP P " Á î 3 g Å „ ž ¥ a V Å Ÿ 4 V L " . {^[12] m l A U " ? Å H î • " ¥ Á î 3 g Å. 3 g ¥ v l ^ ¶ l î < l ¥ Ĩ b W v l „ ? Å š Ö 3 g v l ¥ K l ' . V " 3 g V U „ ? Å ¥ " „ © ÷ 3 g v l „ X Ü ?

Å ÇE Å μ \$ ' a ¥ " ¥ ¥ μ ' . B š ? Å € ¥ " l ž ACK i \$ ' a, 5 3 g L ö _ · M î .

TCP ¥ š Ö e Å Å " ¶ î a š Ö E ç ay Î ë - ©. š Ö E ç " (x M ") ¥ š Ö 3 g M Ä ? p Å / :

$$l ž • ¥ ACK: w w + a(w) / w \tag{1}$$

$$š Ö ? 3: w w - b(w) w \tag{2}$$

Ĩ w V U š Ö 3 g v l ([K v " 1 † Ê. 1 Z L í • , ' Ó [1 † Ê). f Ö š Ö e Å Å 9 \$ ë 1 AIMD (Additive Increase and Multiplicative Decrease) Å. V T(1) # TCP ¥ Å V [w :

$$w(t+1) = w(t) + a(w) \Delta t / RTT \tag{3}$$

9 ü ^ a š Ö 3 g Ä ñ RTT H W = ¥ 9 1 a(w). ž © ¥ f f :

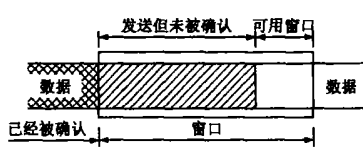


图1 滑动窗口机制

S TCP, a(w) = 1, b(w) = 0.5. ô f Ö Å, S TCP ¥ Q < f " 1 w = 1. 2 / √ p, ó Ü (š Ö 3 g „ x ç y é > q ¥ 1 " .

S TCP 4 [Ú Î © Ž Ĩ s æ " © Ž { z, 1 N Floyd ©: ÓD[4] Ĩ 4 HSTCP Z Å, ' ± X ^ © ÷ TCP ¥ Q < f " , i Y V a(w) „ b(w) ¥ © ÷ Ÿ L C. ! ç H - W d - W d l - P „ L - P + ñ • " , 1 £ „ TCP ¥ Y „ Ÿ „ ¶ z Ÿ, ' w L - W H, a(w) = 1, b(w) = 0.5; ' v ÷ L - W H 5

$$b(w) = (b(H - W) - 0.5) / (\log H - W - \log L - W) (\log w - \log L - W) + 0.5 \tag{4}$$

$$a(w) = w^2 p(w) 2 b(w) / (2 - b(w)) \tag{5}$$

Ĩ, H - W d - W d l - P d - P ¥ w i ' s Y 1 83, 000 a 8 a 10⁻⁷ a 0⁻³, 6 " p(w) = 0.078 / (w^{1.2}), i ! b(H - W) = 0.1. P " w i ' , V [9 Ø ,] w < V 1 HSTCP Ĩ ¥ a(w) ¥ a(w) „ b(w). L = P " Ĩ, „ b(w) | ' a(w) „ b(w) S î á y V (Å V 1), Y V f " Ÿ ` Ü.

w	a(w)	b(w)
38	1	0.50
118	2	0.44
221	3	0.41
347	4	0.38
8405	71	0.10

Ð HSTCP È » ¥ ^ STCP^[7]. STCP 3 g l ÷ • " L - wnd H, P "] S TCP M] ¥ š Ö 3 g ÷ • Ø E. ' š Ö 3 g v ÷ L - wnd H, | a(w) = 0.01 w, b(w) = 0.125. f Ö Z Å L C Ÿ ÷ F e †.

3 RTT, Ü Ÿ „ é > s

3 1 L % œ ê „ ¥ Å 1 s HSTCP ¥ Ÿ ?, • " ns2^[12] (ñ ' 2.26) É) E L . ns2 ^ ' - © Ž ù î 5 x Ĩ < 1 P " ¥ Ö Ü Y q © Ž E .

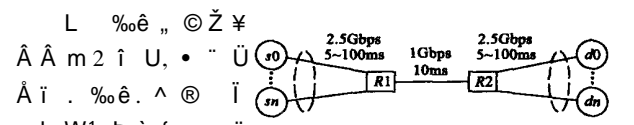


图2 实验拓扑和配置

L % œ ê „ © Ž ¥ Å Å m 2 î U, • " Ü Ĩ Ĩ . % œ ê . ^ © Ĩ b W 1 B à { z „ ü p ð v l, v l ! 1 1000byte, TCP P " SACK ñ ' , TCP ¥ š Ö 3 g K v ' ! Å 1 100000, V 7 „ v š Ö 3 g ? Å " . 6 " ! ç ¶ l μ @ ¥ Ĩ b W.

1 Ð V ? E C L ì â,] H h " phase effect^[13] Ù 5, P " web < @ [# B t ž _ „ Q _ ù TCP @ 6 " Í > 3 ~ 5 ñ S TCP @ ð ¶ S „ 2 • H W Ù ! ç. f t ë 1 E ¥ ß < , Ä Q E û > . d 9 " ¥ s | š Ö E ç " ' x M " (stable state) ¥ " . E ö 1 P " DT(drop tail) 5 Ø ! { .

3 2 Ú Î © Ž Ĩ ¥ RTT, Ü Ÿ „ s 3 2 1 RTT, Ü Ÿ Y V L ? C, • " DT 5 Ø ! { H, Ú Î © Ž Ĩ ¥ RTT, Ü Ÿ d È ø x. P " RED H, f Ö C ' i, ü A. î [' Ó L ö 1 Ç E P " DT(drop tail) ^ © ¥ f f :

$\beta < @f f / , > 5 \tilde{n} \text{HSTCP} @ m3 \cdot 1 \cdot 5$
 $\tilde{n} \text{HSTCP} @ \check{s} \check{O} 3 g \check{U} \check{H} \check{W} \check{M} \check{A} f f , \check{V} 2 \check{I} \check{o} ,] @$
 $\check{\neq} \text{RIT} \check{\cdot} \check{,} \check{,}] \text{H} \check{Y} \check{\neq} \check{P} \check{a} \{ z \} \mu q .$

$V m3 \check{,} \check{V} 2 \check{V} [A , \text{RIT} @ 1 \check{,} \check{2} \check{\neq} 3 g 9 F y ,$
 $\{ z \} \mu q \check{U} ; 7 \check{E} \text{RIT} @ 3 \check{\neq} \check{a} \check{\neq} 5 \check{U} \check{H} \check{W} 9 F K \check{O} 3 g /$
 $\check{t} \check{z} \check{''} @ \check{\neq} \check{\cdot} , \check{P} \check{a} \{ z \} \mu q 5 \div @ \check{\cdot} \check{a} \check{u} \text{HSTCP} \check{\cdot} \text{L} i$
 $\text{RIT} , \check{U} \check{Y} . 7 \check{O} \check{U} \check{H} \check{W} 9 F , f \check{O} , \check{U} \check{Y} \check{Y} \check{u}$
 $A , \check{o} x .$

$P \check{''} \text{STCP} \text{H} , \text{RIT} , \check{U} \check{C} \check{\cdot} \div F \check{o} x . \text{RIT} \check{''} 120 \text{ms} a$
 $160 \text{ms} \check{,} 240 \text{ms} \check{\neq} \check{U} \check{I} @ > 200 \text{s} \check{a} s \check{Y}] \check{''} \check{P} \check{a} \{ z$
 $\check{''} 91\% \check{a} . 73\% \check{a} . 77\% .$

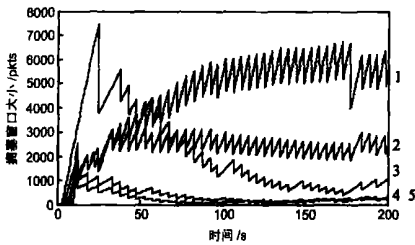


图3 拥塞窗口随时间的变化规律
 $V 2 \text{RIT} , \check{U} \check{Y} \check{L} \check{''} 2 T$

HW Wi (s)	P à { z æ " q (%)				
	@1 RIT= 60ms	@2 RIT= 80ms	@3 RIT= 140ms	@4 RIT= 200ms	@5 RIT= 240ms
50~ 100	38.98	20.96	18.24	2.38	1.06
100~ 150	48.10	21.35	14.10	1.77	0.84
150~ 200	52.97	21.39	11.56	1.52	0.80

$3 2 2 \check{z} \check{\sim} \check{\neq} s \check{''} 1 L \check{z} \check{O} 3 , P \check{''} \check{O} \check{D} [14] \check{I} \check{\neq}$
 $\check{\sim} s \check{s} \check{O} 3 g \check{M} \check{A} , N \$ \check{V} [\check{''} , \check{^} w \text{RIT} ,$
 $\check{U} \check{Y} .$

$\check{3} 4 \check{\sim} y \check{e} > \check{t}] \check{,} \check{e} > \check{\neq} L \check{z} / .] \check{,} \check{e} > \check{\cdot} ,]$
 $@] \text{H} \check{U} \check{a} \check{s} \check{O} . \check{\cdot} \check{''} \text{DT} \check{^} @ \text{H} ,] \check{,} \check{e} > f f 1 \check{o}$
 $R . - \check{U} \check{I} @ \check{Z} \check{I} ,] \check{,} \check{e} > d \check{E} \check{u} A , f \check{A} \check{V} m3 \check{V}$
 $A .$

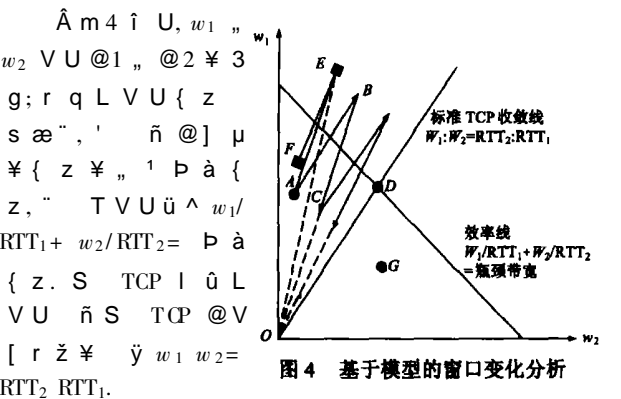


图4 基于模型的窗口变化分析

$L \check{z} \text{RIT} 1 < \text{RIT} 2 . \tilde{n} \text{H} \check{Y} , w_1 > w_2 (m \check{I} A \check{A} \check{V} \check{U}) . w_i$
 $\check{,} \text{RIT}_i \check{V} \check{U} @ i \check{\neq} \check{s} \check{O} 3 g v l \check{,} \text{RIT} \check{\cdot} .$

$n 5 s S \text{TCP} @ \check{\neq} 3 g \check{M} \check{A} ? p . S \text{TCP} @ \check{\neq}$
 $a(w) = 1 , b(w) = 0.5 . \check{A} \check{A} \check{r} q L / , \check{V} \check{U} \check{y} \check{^} \mu : \check{A} \{$
 $z , y N @ 1 \check{,} @ 2 9 F \check{s} \check{O} 3 g . \check{o} > 2 \check{''} \check{\neq} T (3) .$

$5 \Delta t \text{HW} = 3 g \check{\neq} 9 F \check{\cdot} \check{''} 1 a(w) \Delta t / \text{RIT} , y N] \check{''} \check{\neq}$
 $\text{H} \check{W} \check{w} \check{i} = , \tilde{n} @ \check{\neq} 3 g 9 F \check{\cdot} \check{\neq} 1 \check{\cdot} \check{^} \Delta w_1 / \Delta w_2 =$
 $\text{RIT}_2 / \text{RIT}_1 , \hat{i} [\check{M} \check{A} ? p \check{''} \text{A} \check{B} \text{L} , \text{A} \check{B} \check{U} > \check{z} \text{S} \text{TCP} \check{I} \check{u}$
 $\text{L} (\text{O} \check{D} \text{L}) . \check{B} \check{A} \check{r} q \text{L} , y N \check{N} \check{''} \check{O} \check{Z} \{ z , \hat{i} [\check{o} ?$
 $3 @ \check{Z} \check{s} \check{O} . \check{s} \check{O} ? 3 \check{a} , 3 g h l , \check{V} \check{B} \check{z} \text{C} \check{A} . \check{o} 3 g$
 $\check{M} \check{A} ? p (\check{\cdot} n > 2 \check{''} \check{\neq} T (2) , w_1 (t + 1) / w_2 (t + 1)$
 $= w_1 (t) / w_2 (t) , y N \text{B} \check{C} \check{\neq} \check{u} \check{E} \text{L} \check{Y} \check{V} \check{o} \check{A} . \check{z} v f \check{O} ? p \text{M}$
 $\check{A} , K \check{O} \check{o} l \check{u} \check{z} \text{D} \check{A} , \check{\cdot} \{ z s \check{a} \check{e} \check{''} ,] \text{H} 3 g 1 \check{\cdot} \check{''}$
 $\text{RIT}_2 \text{RIT}_1 .$

$\text{HSTCP} @ \check{\neq} \check{M} \check{A} ? p 5 \mu \hat{i} ,] . \text{HSTCP} \check{I} a(w) \check{U} 3 g$
 $9 F 7 9 F , b(w) \check{U} 3 g 9 F 7 h l . \check{\cdot} \check{\cdot} w_1 > w_2 \text{H} ,$
 $a(w_1) > a(w_2) , b(w_1) < b(w_2) . y N \Delta w_1 / \Delta w_2 > \text{RIT}_2 / \text{RIT}_1 ,$
 $\check{z} v \text{A} \text{E} \text{L} 9 F 3 g (| q v \check{z} \text{O} \check{D} \text{L}) . \check{s} \check{O} ? 3 \text{H} , 5 w_1 (t$
 $+ 1) / w_2 (t + 1) > w_1 (t) / w_2 (t) , \check{V} 7 \check{z} v \text{E} \text{F} \text{L} h \check{''} 3 g v l$
 $(\check{u} \check{E} \text{L} , \check{U} \check{V} \check{o} \check{A} , \text{F} \check{A} \check{E} \check{O} \check{D} \check{A}) , 9 \check{u} \check{^} \check{a} 3 g \mu \check{\cdot} 9$
 $F . y N , \check{U} \check{H} \check{W} 9 F 3 g \check{\neq} \mu \check{o} \check{Y} v . \check{V} m3 \check{I} \check{V}$
 $\check{u} \text{A} \check{A} , @ 1 \check{\neq} 3 g , 9 F , 7 @ 3 \check{\neq} \check{a} \check{\neq} \check{\neq} 3 g / \check{t} \check{z} \check{''}$
 $\check{I} \check{\neq} \check{\cdot} .$

$7 \text{S} \text{H} w_1 < w_2 (\check{A} \check{G} \check{V} \check{U}) \check{\neq} s \check{V} \check{n} \check{D} \check{e} \check{E} \check{''} .$

$? \check{A} \hat{i} q \check{''} v (t) = w (t) / \text{RIT} , y N \text{RIT} , \check{U} \check{Y} \check{D} 3 g$
 $\check{M} \check{A} ? p \text{B} \check{''} . 9 \check{u} \check{^} \check{a} \check{U} \check{H} \check{W} 9 F , \text{RIT} , \check{U} \check{Y} \check{Y} \check{o}$
 $x .$

$[\check{\neq} s \check{^} > \check{t}] \check{,} \check{e} > f f / \check{E} > \check{\neq} . \check{V} m3 \check{V}$
 $[A , \check{t} f f / 9 \check{V} [\text{E} \check{z}] \check{,} \check{e} > . \check{I} \{ z @ @ \check{z} 3$
 $g l , ? \check{A} \hat{i} q @ , \check{V} 7 \check{s} \check{O} ? 3 \check{A} q \check{z} \check{U} \{ z @ f \check{O} f$
 $f \text{B} \check{z} \check{n} \check{h}] \check{,} \check{e} > , \check{t} @ \text{RIT} , \check{U} \check{C} \check{\cdot} . 1$
 $\check{A} m3 \check{I} , 102 \check{o} \text{P} . \text{H} \check{U} \{ z @ 1 \check{,} \check{2} \check{U} \check{z} \check{s} \check{O} , 7 @ 3$
 $5 \check{A} \mu , \text{E} \check{z} \check{e} > . y N @ 3 \check{\neq} 3 g \check{\neq} [9 F . \check{A} \text{T} \check{^} @$
 $\text{P} \check{''} \text{R} \text{E} \check{D} \check{A} , 5 \check{V} [\text{E} \check{z}] \check{,} \check{e} > , y N \text{E} \check{z} \check{o}$
 $x \text{RIT} , \check{U} \check{Y} .$

$\hat{i} [\text{HSTCP} \check{O} \text{E} \check{\cdot} \check{\neq} \check{\neq} \check{A} [\# \check{\cdot} \check{''} \text{D} \check{I} \check{''} \check{A} \text{H} \check{o}$
 $x \check{\neq}] \check{,} \check{e} > \text{C} \check{\cdot} \check{^} / \hat{i} \check{o} x \text{RIT} , \check{U} \check{Y} \check{\neq} \check{o} y .$

$3 3 \check{U} \check{I} @ \check{Z} \check{I} \check{\neq} \check{e} > \# s$

$\check{V} 3 \check{e} \text{L} \check{I} \check{e} > \check{U} 5 \check{\neq} \mu 1 \check{''} . \check{V} \check{V} 3$
 $\check{\neq} \text{L} \check{''} \check{Y} \text{A} , \check{A} \check{Q} \check{s} \check{O} \text{H} \check{\neq} \check{e} > \check{''} \check{,} \check{''} 9 F y 0 a(w)$
 $\mu \check{a} \text{M} 1 \check{''} .$

V3 HSTCP 50~ 200s Hé f f

@	RIT (ms)	y ^ æ " q (%)	š Ö Q"	Ü (Ä Q é "	š Ö H 3 g v l	a(w)
1	60	46.68	33	31.73	6225~ 6753	22~ 24
2	80	21.23	33	16.55	2508~ 2971	13~ 15
3	140	14.63	27	5.78	620~ 1061	6~ 8
4	200	1.89	25	3.00	202~ 327	2~ 4
5	240	0.90	24	2.38	148~ 313	2~ 4

$f \check{O} \check{C} \check{\cdot} \check{,} \text{HSTCP} ? \check{A} \check{''} \check{\neq} \check{A} \hat{i} 3 g \check{A} \mu 1 \check{''} . \check{\cdot}$
 $\check{O} \text{L} \check{z} \check{p} \check{I} \mu @ \check{\neq} \check{I} b \check{W} , 9 \check{u} \check{^} \check{a} 3 g v l @ ? \check{A}$
 $\check{\neq} \check{s} \check{O} 3 g v l \check{Y} \% \check{z} . \text{L} ! \tilde{n} \text{H} \check{Y} ? \check{A} \check{\neq} \check{s} \check{O} 3 g$
 $v l \check{''} 1 w , ? \check{A} \check{o} \check{U} \check{V} \check{''} 3 g \hat{i} , \check{\neq} ? \check{A} \check{€} , \check{V} 7 \text{P}$
 $? \check{A} 7 \check{O} \check{A} \$ \check{\cdot} \check{a} \check{\neq} \check{''} \tilde{n} \check{''} \check{r} \check{z} \check{w} \check{V} \check{''} 3 g \check{z} \text{M} \check{''} 0 .$

6 “ BZë, Äñ? Ä €¥” ž r m l a, m l ö
 QCEBñ ACK . ? Ä l ž Bñ ACK a, Ä T ' a ñ
 ” \$ ž ' ? Ä (ô ¥½ |), L ž v T(1) ÷ •
 š Ö3 g, 6 “ 3 g _ . Ä î, V “ 3 g M¹ l, V 7 V [?
 Ä Bñ • ¥” . f Ö Ä 9 \$ ë¹ ACK clock. Y V f Ö
 Ä, ? Ä V [ô © Ž { z Ø « ? Ä î q [15]. CE^ Ä T
 Q _ y ^ i © Ž š Ö, 5 m l ? Ä ¥ ACK V ? ö \$ é
 , V 7 ö C Bñ ACK a V ñ ” ¥ f f, ? Ä
 ¥ V “ 3 g L ö v ž l, / î ? Ä HW = | ñ ?
 Ä €, TM î l ? . © ž HSTCP ? Ä V ñ Ĩ š Ö3 g “ v
 (Ä @ l r ž 6753), V 7, ^ TM î v ¥ ” l ? . Ä T ^ ©
 Ĩ Ĩ b W, , 5 ” ¥ l ? ö / î v ¥ é > . á Ĩ
 ë f Ö l ?¹ H Y = ¥ l ? .

' P Ä µ H Y = ¥ ” l ? , HSTCP ¥ ” Z Ÿ 9 ö /
 î B Q š Ö H ñ é > . T(3) · HSTCP Bñ RTT
 H W = š Ö3 g 9 F a(w). L ç ñ H Y, D T ^ © Ä µ
 Ĩ b W, a ü f ñ H Y ^ © ¥ Ä E g Ĩ q, , g Ĩ q r ž
 Ü TM Ä T ? Ä ? Ä 9 F ? Ä î q, 5 ö C é . © ž
 HSTCP ¥ 3 g ô T(1) ^ , 9 F ¥, y N ö µ ” \$
 é . ô TCP ¥ Ä, Ä T š Ö ? 3 , ? Ä ³ 1 Ü V B
 ñ RTT H W C E V [_ © š Ö [12] (Y V 3 ñ x - ¥ ACK) .
 Bñ RTT W i = , ? Ä ¥ Ĩ q 9 F¹ a(w), y N f t
 û ö \$ é .

8 , HSTCP ¥ ” Z Ÿ , v š Ö3 g ¥ H Y = ” l
 ? ^ / î š Ö ? 3 H v é > ¥ ð y .

4 ž É Ø E: CB HSTCP

ô ë ¥ s , 4 ž É Z Ä CB HSTCP (Constant
 window and Block pacing HSTCP), ô 1 “ ñ † s: CW Ä
 „ Block pacing Ä, s Y ” ž ž Ÿ Ü Ÿ , h é > q .

4.1 CW Ä

¾ Ä ö 1 “ ž ž Ÿ HSTCP ¥ RTT , Ü Ÿ . ô T
 (3), B ç ¥ H W W i Δ t = , HSTCP @ ¥ š Ö3 g 9 F¹
 (a(w) Δ t / RTT) . y N CW Ä ö 1 ± X ü ^ ô RTT ,
] , Y V @ F Ü y 0 , h ” RTT 3 g 9 F Ĩ q ¥ • Y , P
 ,] RTT ¥ @ ¥ š Ö3 g t Ĩ ž M © .

| Ü y 0 η = c RTT, c ¥ c l Ð CR [16] Ø E Ĩ ¥ E
 », 1 Ä c = 10, RTT = 100ms = 0.1s, 5 η = 1. Y V f Ö Z T,
 ,] RTT ¥ @ x M Ÿ ö m] ¥ š Ö3 g 9 F. Δ t H
 W = 3 g ¥ 9 F¹ M¹ (c a(w) Δ t), E ç ,] @ ¥
 RTT ,] 3 g M Ä ¥ • Y .

Ä / è a ü . L ç ñ H Y t, @ 1 ¥ RTT ' 1 100ms,
 @ 2 ¥ RTT ' 200ms, š Ö3 g v l û ^ w . @ F Ü y 0 - ,
 5 @ 1 t + 100ms a , 3 g v l ' 1 w + a(w); t + 200ms
 a , 3 g v l ' 1 w + a(w) + a(w + a(w)) w + 2 a(w), 7
 @ 2 t + 200ms, 3 g v l ' 1 w + a(w). y N 3 g v l i
 µ s . @ F Ü y 0 a , 1 Ä c = 10, 5 @ 1 ¥ η = 1, @ 2 ¥ η
 = 2, y N 200ms a , € ¥ 3 g û 1 w + 2 a(w). c | 6
 “ ¥ ' 9 V [m M] ¥ 2 .

© ž ? Ä î q v(t) = w(t) / RTT, 3 g M] ¥ f f / ,
 RTT s Y¹ RTT₁ RTT₂ ¥ ñ @ M] ¥ H W = Ü ({ 1
 ' 1 RTT₂ RTT₁, ' Ø , ž É Ø E ¥ RTT , Ü Ÿ¹ RTT
 ¥ Q 1 .

4.2 Block pacing Ä

Ä - î • , HSTCP š Ö ? 3 H i v ¥ é > . @ F
 Ü y 0 a , É RTT @ f Ö C ` ÷ F ø x . 1 Ä RTT¹
 300ms, c = 10, 5 Ä ñ RTT = 3 g ¥ 9 F¹ 1 3 a(w) . 1 N
 4 block pacing Ä .

B Z ë • ” pacing Ø E,] , H Y = ¥ ” l ? . “
 Interval V U - W ¥ W i , 5

Interval = RTT / w (6)

V 7 3 g ¥ ? Ä V [e Ä 3 g ? Ä , s ACK ž r
 o î „ é > ¥ • Y .

6 “ • ” block Ä, 9

ü ^ | Bñ š Ö Ü ù = ¥ 3 g
 9 F s î ñ † s (block): B
 † s ^ ” Z 9 É u (Ä m 5 Ĩ
 AB), 9 ü ^ ž v CW Ø E 9 F
 š Ö3 g; 6 “ B † s ^ Ĩ
 9 É u (Ä m 5 Ĩ BC), ' ? Ä
 ' © Ž " Ĩ š Ö H, | 3
 g 9 F y 0 a(w) ! Ä¹ l, V
 7 h CBHSTCP ¥ ” Z Ÿ î { Ÿ ¥ é > ” .

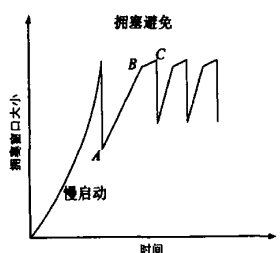


图5 block-pacing原理

š Ö ¥ ç © ö 1 ^ ô RTT • . RTT ' ¥ M Ä V [
 “ ž ç © © Ž š Ö Ÿ [17, 18]. “ R c } V ' - ¥ RTT '
 RTT_min V U RTT ¥ K l ' , RTT_max V U RTT ¥ K v ' . RTT_min ^
 © © Ž Ÿ ^ ¥ . l ü p î % ç , y N (R c - RTT_min) ü V U
 © ž ^ © Ĩ b W 9 F î „ ¥ ü p 9 F ' , A - (RTT_max
 - RTT_min) V U š Ö ? 3 H 9 F ' ¥ K v ' . y N P “ Ä /
 T Ÿ ç © š Ö:

Rc = RTT_min + β (RTT_max - RTT_min) (7)

Ĩ β ^ B ñ • ” . Ä m 5 î U, β v 5 B Ä m í C Ä.
 © ž - ” Ĩ š Ö H L @ block Ø E, y N ³ 1 É B ñ m í ž
 l ¥ β ' . l n ž RTT 9 i , ^ d É Ü ' , ' Ó ! Ä¹ 0.8.
 Ø E ¥ Ä } ' í • n m 6 (a(w) , b(w) ¥ W % 9 Ø V
 [• l » 2 «) . HSTCP m î ” ¥ ž É V n Ó D [19].
 m î ” 3 g 9 F “ y, y N @ F Ü y 0 ¥ Z E , ? “
 m î ” . ' Ó ¥ Ø E ö 1 Ø š Ö E ç ” , 7 0 3
 g v ž L - W a L @ V 7 £ TCP Y , Ÿ .

Initial values:

- RTT_min = 10000;
- RTT_max = 0;
- β = 0.8;

On receiving a new ACK in congestion avoidance state:

if (w - L - W) // L - W c l n » 2 «

a(w) = 1;
 else
 a(w) = increment(); // ž ž ñ (6) 9 Ø a(w)

```

increment= c* RTT* a(w)/cwnd; // cwnd ^ š Ö 3 g v l
if(increment > 1)
    increment= 1; // E ç 9 F y ç ð ð î
cwnd= cwnd+ increment;
interval= RTT/cwnd; // pacing Ĩ - W ¥ H W W ĩ
if(RTT < RTT_min)
    RTT_min= RTT;
if(RTT > RTT_max)
    RTT_max= RTT;
diff= RTT_max- RTT_min;
if((RTT - RTT_min+ ß* diff)
    a(w)= 1;

```

On congestion happening:

cwnd= cwnd* (1- b(w));

On sending date:

```

output(interval); // H W W ĩ interval ^ a ? Ä "
m6 CBHSTCP Ø E ¥ Ä } '

```

5 ĩ ũ Ÿ, x ç Ÿ^ a ü

› †] „ é > ¥ f f / , ô » 3.2 « 4 ¥ ~ É
› s V © , ' P , i RTT µ s , ' 3 g µ Y 1 v H ,
] ¥ HSTCP @ 9 4 [ĩ ũ . Œ ^ L = © Ž Ĩ ,] „ é >
t f f / V [E ç (» 3.2 « X Ũ É » ^ a ü) , y N
HSTCP K Ô ô ĩ ũ . Œ HSTCP 3 g µ s v H ĩ ũ ð ¥ Ũ
5. Ó D [20] 4 y ĩ ĩ ũ Z E , V [T ' ' Ó ũ ĩ ¥ €
ç É ¥ CBHSTCP Ø E ¥ ĩ ĩ ç Y V @ F Ü y 0 ,
,] RTT ¥ @ B Š ĩ ũ ž 3 g M] ¥ Ÿ , L ö µ M] ¥
a(w) „ b(w) ^ , V 7 E ç » 3.2 « ĩ s ¥ HSTCP ĩ
a(w) „ b(w) 3 g M Ä ¥ • Y , K Ô ĩ x ç ¥ Ü Ÿ .

6 E s) ,

6.1 ß N · S

E ¥ µ 1 ! Ä] » 3.1 « . ns2 ĩ L C CBHSTCP
Ø E , i É > E L Z Ä É > s ß N . ö 1 ¥ ß N ·
S ^ ý ^ æ " q „ Ü Ÿ . 1 ç Ü Ÿ ¥ s , 4 • ¥ M
Ü Ÿ S . ç ĩ M Ü Ÿ S ' : (1) £ S TCP @
¥ { z ^ 3 p , P S TCP @ µ @ ¥ { z ? Z b W (2) £
Ũ ĩ @ - W ¥ Ü Ÿ .

ĩ @ (1) 5^ a ^ 1] H 9 ĩ @ TCP ¶ z Ÿ , 9 ũ ^ a V
[„ S TCP ¶ z M) . f Ö ¶ z Ÿ 9 8 C] S TCP ¥
Y , Ÿ . S TCP @ ç ' & ¥ Œ J , a † < " ç Ũ é > q ¥ ĩ
â / . Ä T S TCP @ ¥ é > q ĩ ç B ç ^ , 5^ a ^ 1 X Ũ µ
@ ¥ { z ? Z b W , 9 V^ a ^ 1 Ũ ĩ TCP @ µ S TCP
¥ ¶ z Ÿ . ' Ó É ç ß N S ' HSTCP ĩ ç ĩ ¥ L - W ĩ
< ¥ é > q L - P : 10^-3 .

ç (2) , 5 P " Ü . " FI^[15] Ÿ ß N Ũ ĩ @ - W Ü
Ÿ . T ' :

$$FI = \left[\sum_{i=1}^n x_i \right]^2 \left[\sum_{i=1}^n x_i^2 \right]^{-1} x_i \geq 0$$

ĩ x ĩ ^ 1 ö ð @ ĩ] " ¥ { z 1 è .
¾ x p ö 1 " ç v " . { , y N H ü „ H ü ô ĩ ,
^ ß N „ TM ¥ ö 1 . S .

6.2 CW Ø E ¥ r T

f † s L ö 1 ' ^ a ü @ F Ü y 0 ¥ r T .
ß @ f f / , › RTT ' 80ms a 40ms a 200ms (s Y S : ' 1
1 a a) ¥ Ø ñ Ũ ĩ @ • " c = 10 . V m 7 V n , 4 ¥ Z Ä V
[P ,] RTT ¥ @ ð M] ¥ 3 g ^ , V 7 E ç P " HSTCP Ø E ¥ Ũ 5 (n m 2) , ð x ç ¥ Ü Ÿ . © N^ a ü
CW Ä V [µ r ç Ÿ HSTCP ¥ Ü Ÿ .

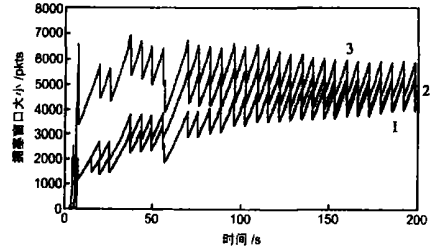


图7 采用CW算法后的拥塞窗口变化规律

6.3 Block pacing Ø E ¥ r T

' £ block pacing / Œ ¥ r T , É > L , 1 CB
HSTCP a W HSTCP (Ç Ç @ F Ü y 0 Œ Ä µ P " block pac
ing) . s Y › Ø ñ CW HSTCP @ „ CB HSTCP @ • " c =
10 . V V 4 ĩ A , CB HSTCP P " block pacing Ø E^ a é >
q h . f ö 1 ^ @ ç 4 ¥ Ø E µ r ' E ç 3.3 « ĩ 4
ž ¥ H Y = ¥ " ĩ ? ĩ V 4 Block pacing Ø E ¥ r T

š Ö ? 3 - † @ CB
HSTCP ¥ " Z Ÿ . 7 O block
pacing Ä š Ö ? 3 - L
@ ĩ , • Y CW Ä ĩ
ĩ ¥ Ü Ÿ „ x ç Ÿ , [/
¥ L ^ T £ L f B Ä .

RTT (ms)	é > q (10^-4)	
	CW HSTCP	CB HSTCP
80	1.2	0.08
140	1.8	0.10
240	2.9	0.16

6.4 Ü Ÿ s

f † s L 1 • " ,] Ø E H ¥ Ü Ÿ . Ä F L
ĩ ũ µ Ø ñ @ RTT s Y ' 80ms a 40ms a 200ms . s Y P " CB
HSTCP (c = 10) , HSTCP , STCP „ S TCP Ø E , Ü V 9 Ø ð
• " ,] Ø E H ¥ Ü y 0 . Ä m 8 V n , ç É Ø E v v ç Ÿ
Ü Ÿ .

ĩ @ ¥ ñ " 9 F ž 5 ñ , RTT s Y ' 60 a 80 a 160 a
200ms . E^ a 9 Ø ð , ý ^ æ " q s Y ' 24. 56% a 20.
23% a 15. 12% a 3. 43% a 1. 56% , ß @] " { z 8. 76% , 9
ý ^ æ " q 93. 66% . Ũ ĩ @ - W Ü . " 1 0. 93. y N , CB

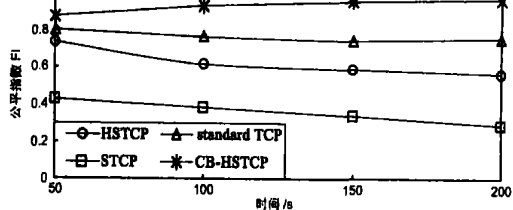


图8 不同算法的公平性

HSTCP Ø E μ z ¥ Ü Ÿ.

V 5 • " c ¥ • Y „ TCP ¶ z Ÿ

6 5 • " c ¥ Ê | „ TCP ¶ z Ÿ

c ' ¥ Ê | Ÿ ? ¥ • Y 9 1

v. V 5 1] " ¥ Æ H q / , • " c | ,

] ' H ¥ L ² T. V V 5 V n, c 1 5

H, 5 Ú Í @ î] " ¥ { z „ 9 ¥ ý ^

æ" q ú μ í / †, f a ü c V @ ö K

Å CB HSTCP ¥ " Z ? ĩ . c 1 20 H, 5 CB HSTCP 3 g " Z ? ĩ 9 <, OE ^ β @] " { z h , 7 O é > q 9 F, 9 ü ^ a TCP ¶ z Ÿ, @ 9 ¥ Ÿ f, Ê 4 c 1 10 1 † a, f a " é " Ú Í © Ž ũ p S ¶.

V V 5 V [A , β @ ú ^ S TCP @ é > q ú 10⁻³ [/ , î [4 ¥ Ø E μ TCP ¶ z Ÿ.

7 ² , „ / B „ ¥ ù î

CB HSTCP OE HSTCP Ĩ ¥ ø x RTT , Ü Ÿ Ü 5, Y V @ F Ü y 0 é Ÿ HSTCP Ü Ÿ, i P " block pacing / OE ç H W = ¥ " | ? „ HSTCP ¥ " Z Ÿ ĩ / ĩ ¥ B Q š Ö Ĩ ¥ ñ é > . Ø , s „ E L V ü CB HSTCP £ HSTCP δ μ a Ä Ä Ú ý ^ æ" q ð CP ¶ z Ÿ ¥ \$, 4 Ú Ú Í @ - W ¥ Ü Ÿ a † ® é > q. ¾ Z Ä L @ e †, Ç Ç ³ 1 © é ? Ä , μ z ¥ " Z Ÿ, V [" z 1 " Ú { z © - † ¥ © Ž Ĩ . Ú © Ž { z , 9 F, ¾ Ø E μ < - ¥ < " - β. N " Î ³ 1 É B „ ù î 1 ù ö î Ä (AQM) 4 Ü HSTCP Ÿ ? .

• I Ó D:

[1] I Foster, C Kesselman, S Tuecke. The anatomy of the grid: enabling scalable virtual organizations[J]. International Journal on Supercomputer Applications, 2001, 15(3): 200– 222.

[2] B Phillips. Have storage area networks come of age[J]. Computer, 1998, 31(7): 10– 12.

[3] V Jacobson, R Braden, D Borman. TCP extensions for high performance [S]. RFC1323, 1992.

[4] S Floyd. HighSpeed TCP for large congestion windows[S]. RFC3649, 2003.

[5] J Senke, J Makhavi, M Mathis. Automatic TCP buffer tuning[J]. Computer Communications Review, 1998, 28(4): 315– 323.

[6] H Sivakumar, S Bailey, R L Grossman. Pockets: the case for application level network striping for data intensive applications using high speed wide area networks[A]. Proceedings of Supercomputing [C]. Dallas: ACM Press and IEEE Computer Society Press, 2000. 63– 64.

[7] T Kelly. Scalable TCP: improving performance in high speed wide area networks[J]. ACM Computer Communications Review, 2003, 33(2): 83– 91.

[8] D Katabi, M Handley, C Rohrs. Congestion control for high bandwidth delay product networks[J]. Computer Communications Review, 2002, 32(4): 89– 102.

[9] L Xu, K Harfoush, I Rhee. Binary increase congestion control(BIC) for fast long distance networks[A]. Proceedings of INFOCOM [C]. Hong Kong: IEEE, 2004. 2514– 2524.

[10] F Souza, D Agarwa. A highSpeed TCP study: characteristics and de-

C	P à { z æ" q (%)					Ü . "	β @ ¥ é > q
	CB HSTCP RTT= 80ms	CB HSTCP RTT= 140ms	CB HSTCP RTT= 200ms	β @	9 „		
10	33.21	26.25	18.26	16.65	94.37	0.95	6 10 ⁻⁵ ~ 7 10 ⁻⁴
5	28.73	16.16	14.53	25.86	85.28	0.91	5 10 ⁻⁵ ~ 3 10 ⁻⁴
20	43.67	25.81	19.56	5.67	94.71	0.89	3 10 ⁻⁴ ~ 1 10 ⁻³

plyment issues[DB/OL]. <http://www.icir.org/floyd/hstcp.html>.

[11] Network simulation ns2[CP/OL]. <http://www.isi.edu/nsnam/ns>.

[12] W R Stevens. TCP/IP illustrated: the protocols[M]. Volume 1. Addison Wesley, 1994.

[13] S Floyd, V Jacobson. On traffic phase effects in packet switched gateways[J]. Internetworking: Research and Experience, 1992, 3(3): 115– 156.

[14] D Chiu, R Jain. Analysis of the increase and decrease algorithms for congestion avoidance in computer networks[J]. Journal of Computer Networks and ISDN, 1989, 17(1): 1– 14.

[15] V Jacobson. Congestion avoidance and control[J]. Computer Communication Review, 1988, 18(4): 314– 329.

[16] S Floyd. Connections with multiple congested gateways in packet switched networks Part 1: one way traffic[J]. Computer Communications Review. October, 1991, 21(5): 30– 47.

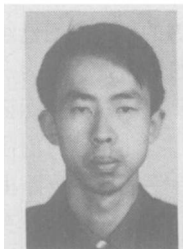
[17] L Brakno, L Peterson. TCP Vegas: End to end congestion avoidance on a global Internet[J]. IEEE Journal on Selected Areas in Communication, 1995, 13(8): 1465– 1480.

[18] A Kuzmanovic, E Knightly: TCP LP: a distributed algorithm for low priority data transfer[A]. Proceedings of IEEE INFOCOM [C]. San Francisco: IEEE, 2003. 1691– 1701.

[19] S Floyd. Limited slow start for TCP with large congestion windows[S]. RFC3742, 2004.

[20] M Nabeshima, K Yata. Improving the convergence time of highSpeed TCP[A]. IEEE International Conference on Networks[C]. Singapore: IEEE, 2004. 19– 23.

T € e 0 :



É O+ 3, 1976 M 7 3 é ö ê ü ½ „ ~ v ð 9 Ø Ð ý p V 3, ö 1 ù î Z _ : © Ž x p æ © Ž Qos. E-mail: sufanjun@163.com.



OE Ø 9 3, 1942 M 11 3 é „ ~ Ñ 4 „ ~ v ð 9 Ø Ð ý • q, p V 3 • =, ö 1 ù î Z _ : © Ž " o ½ † / OE / B } © Ž Y β Ð s f) Ø / OE