

# 基于带宽预测的 VCP 拥塞控制改进算法<sup>①</sup>

何建新, 周翠红

(湖南城市学院 信息科学与工程学院, 益阳 413000)

**摘要:** 反馈信息不足导致 VCP 无法根据网络拥塞程度进行相应窗口调整, 在网络突发业务流情况下存在慢收敛问题. 针对高带宽延时网络特性, 本文提出一种基于带宽预测的 VCP 拥塞控制改进算法 VCP-BE. 该算法结合端到端可用带宽预测和 2 位显式拥塞反馈信息提高拥塞反馈精度, 为源端拥塞窗口调节提供更精细的拥塞控制信息. 仿真结果表明, VCP-BE 能够有效适应高带宽时延网络环境, 具有比 VCP 及 MLCP 更快的公平收敛速度.

**关键词:** 带宽预测; 高带宽时延网络; 拥塞控制; VCP

## Improved VCP Congestion Control Algorithm Based On Bandwidth Estimation

HE Jian-Xin, ZHOU Cui-Hong

(College of Information Science and Engineering, Hunan City University, Yiyang 413000, China)

**Abstract:** VCP can't adjust the congestion window according to the congestion degree because of insufficient feedback information, and shows slow convergence speed with the burst traffic flows. This thesis presents a improved VCP congestion control algorithm named VCP-BE in high bandwidth delay product networks, which uses the end-to-end bandwidth estimation algorithm and two existing ECN bits to improve the accuracy of congestion information feedback, provides more precise congestion control information for the source congestion window adjustment. Simulations show that the improved VCP-BE can adapt to high bandwidth delay networks, has faster fair convergence speed than VCP and MLCP.

**Key words:** bandwidth estimation; high bandwidth delay product networks; congestion control; VCP

下一代互联网的发展趋势表明, 高带宽时延网络将大规模应用于互联网主干并成为未来的主流网络. 伴随着网络带宽和往返时延不断增加, 传统 TCP 的 AIMD 窗口管理机制的低效性导致网络带宽利用率偏低, 不同流的 RTT 差异导致吞吐量不公平性加剧, TCP 流的频繁抖动增加了网络不稳定性, TCP 协议本身成为了限制高带宽时延网络性能增长的瓶颈, 传统 TCP 拥塞控制机制在高带宽延时网络环境正面临严重挑战.

针对高带宽时延网络特性, 国内外学者相继提出一系列算法或改进协议. 根据调节拥塞窗口所采用反馈信息不同, 分为基于丢包或延时反馈的端到端控制协议和基于中间结点显式反馈协议两大类. 丢包反馈是一种被动拥塞控制机制. 显式反馈协议根据来自中

间节点的反馈信息能够更准确、及时地将拥塞信息传送给源端, 从而避免端到端拥塞控制协议丢包、延时反馈所产生的反映滞后和拥塞误判. YongXia 等人提出 MLCP<sup>[1]</sup>通过增加反馈信息至 7 位达到接近最优性能, 但额外的反馈开销增加了 MLCP 在 Internet 的部署难度. Qazi 等提出 VCP<sup>[2]</sup>通过计算链路负载指数, 使用 2 位编码表示链路拥塞程度并通过 ACK 返回发送端, 发送端据此反馈信息调整发送速率. 文献[3]研究表明 VCP 可以实现高带宽利用率、较低丢包率和业务流之间的合理公平性, 但反馈信息不足导致 VCP 为避免突发现象而对拥塞窗口增长幅度进行限制, 使其无法根据具体拥塞程度进行相应窗口调整, 存在慢收敛问题. 文献[4]同时改进 VCP 的效率收敛速度和公平收敛速

<sup>①</sup> 基金项目:湖南省科技计划项目(2014FJ4252)

收稿时间:2016-01-28;收到修改稿时间:2016-04-11 [doi:10.15888/j.cnki.csa.005387]

度,但却额外增加了显式反馈开销.因此,如何合理设置反馈信息位数保证显式反馈协议实用性,同时提高反馈精度,在兼顾高效性和公平性同时保证协议较快的收敛速度是提升 VCP 性能的关键.陈洁等人通过对各种拥塞控制协议在交叉网络环境下的性能测试、分析比较,发现不属于高速网络的 Westwood 和 Hybla 在高带宽时延网络中反而有较好表现,并分析指出采用带宽预测策略降低反馈开销是使其性能提升的关键原因<sup>[5]</sup>.为提升高带宽时延网络环境下 VCP 的公平收敛速度,在不扩展反馈信息长度的同时弥补 VCP 拥塞反馈信息不足的缺陷,本文在对 VCP 算法原理分析基础上,设计一种基于带宽预测的 VCP 拥塞控制改进协议 VCP-BE(Variable Structure Congestion Protocol-Bandwidth Estimation).

### 1 VCP算法基本原理

VCP采用一种不同于TCP的基于负载因子的拥塞控制机制.路由器每间隔固定周期 $t_p$ 计算网络负载指数 $f$ ,如公式(1)所示,其中 $y$ 是链路吞吐量, $q$ 为目标队列长度, $k$ 为队列收敛速度(取默认值0.5), $C$ 为链路出口带宽, $r$ 为链路出口带宽目标利用率(取默认值0.98),因Internet数据流的RTT值一般小于200ms,故周期 $t_p$ 取常数200ms.

$$f = \frac{y + kq}{rCt_p} \quad (1)$$

VCP根据负载指数 $f$ 的取值范围将网络拥塞程度划分为低负载,高负载和过载3种不同级别,并将其量化编码为2位二进制信息分别填入IP包首部中的ECN位,TCP源端根据ACK携带回的不同状态信息,分别采用不同拥塞窗口调节算法:当网络处于低负载状态,源端采用MI(Multi Increase)策略增加发送速率提高网络利用率;当网络处于高负载状态,源端采取AI(Additive Increase)策略逐步增加发送速率;当网络处于过载状态,源端采用MD(Multi Decrease)策略快速减小发送速率.VCP的ECN标记及源端拥塞窗口调节算法如表1所示,其中 $rtt=t_p$ , $\epsilon=0.0625$ , $\alpha=1$ , $\beta=0.875$ .

表1 VCP的ECN标记及源端拥塞窗口调节算法

网络状态	负载因子	ECN位	VCP拥塞窗口调节算法
低负载	$0 \leq f < 0.8$	01	MI: $cwnd(t+rtt) = cwnd(t) * (1 + \epsilon)$
高负载	$0.8 \leq f < 1$	10	AI: $cwnd(t+rtt) = cwnd(t) + \alpha$
过载	$f \geq 1$	11	MD: $cwnd(t+rtt) = cwnd(t) * \beta$

尽管VCP仅用2位反馈就获得了较好的网络性能,但反馈信息不足导致VCP仍存在公平收敛速度缓慢,低负载阶段拥塞窗口增长缓慢等问题.VCP根据当前RTT值按(2)式修正调节因子 $\epsilon$ ,按(3)式修正调节因子 $\alpha$ ,以消除RTT差异所带来的不公平性<sup>[6]</sup>.

$$\epsilon_s = (1 + \epsilon)^{rtt/t_p} - 1 \quad (2)$$

$$\alpha_s = \alpha * (rtt / t_p)^2 \quad (3)$$

### 2 VCP改进协议设计

拥塞反馈信息不足是造成VCP提速保守的根本原因.为了真实、有效地反映网络拥塞,提高端到端拥塞反馈信息的准确性和反馈精度,在保持2位ECN信息不变情况下,VCP-BE源端通过RTT变化来估测端到端可用带宽,利用可用带宽预测值以及2位ECN反馈信息为源端拥塞窗口控制算法提供更精细的拥塞控制信息.

当网络处于低负载状态时,VCP-BE主要关注算法的高效性,为实现拥塞窗口的快速调整,源端通过带宽预测可允许的发送速率,使用SI(Sudden Increase)算法将拥塞窗口设置为当前预测的可用带宽;当网络处于高利用率状态时,VCP-BE重点关注算法的公平收敛性,将高负载状态划分为远离过载和接近过载两种不同状态.当远离过载时,为加快收敛速度,源端采用较快的AI窗口增长机制;在接近过载时,为保持网络稳定性,采用比较平滑的II(Inversely proportional Increase)速率变化机制;当网络处于过载状态时,VCP-BE主要关注协议对拥塞的响应性,源端仍采用MD算法来减小拥塞窗口.VCP-BE中间节点ECN标记策略及源端拥塞窗口调节算法如表2所示.VCP-BE算法主要涉及以下三方面内容.

表2 ECN标记策略及拥塞窗口调节算法

网络状态	负载因子	ECN位	VCP-BE拥塞窗口调节算法
低负载	$0 \leq f < 0.8$	00	SI: $cwnd = \text{MAX}[cwnd_{i-1} * (1 + \epsilon), (RTT_{min} * BW_n / seg\_size)]$
远离过载	$0.8 \leq f < 0.9$	01	AI: $cwnd = cwnd_{i-1} + a / cwnd_{i-1}$
接近过载	$0.9 \leq f < 1$	10	II: $cwnd = cwnd_{i-1} + a / (cwnd_{i-1})^{0.5}$
过载	$f \geq 1$	11	MD: $cwnd(t+rtt) = cwnd(t) * \beta$

#### 2.1 可用带宽预测 ABE

VCP-BE通过在发送端对端到端可用带宽进行预测来调整拥塞窗口,以便更主动的利用网络资源,预防拥塞.连续ACK包收到时间间隔反映了网络资源的

利用情况, ACK 包到达发送端越快, 说明网络可用资源越充足.

当网络处于低利用率状态时, VCP-BE 主要关注带宽预测的响应性; 当网络负载较大时, 则重点关注带宽预测的准确性. 为确保带宽预测的响应性和准确性, 针对显式反馈的不同负载程度, VCP-BE 利用 ACK 的返回速率, 采用文献[7]如公式(4)所示的网络可用带宽预测算法 ABE(Available Bandwidth Estimation).

$$BW_n = \frac{\theta * RTT * BW_{n-1} + L_n}{\theta * RTT + (t_n - t_{n-1})} \quad (4)$$

其中  $BW_n$  表示  $t_n$  时刻第  $n$  个 ACK 到达源端时的预测带宽,  $t_n$  是接受到第  $n$  个 ACKs 的时刻,  $L_n$  是被确认分组  $n$  的大小,  $\theta$  为时间间隔调节因子, 是影响带宽预测准确性和响应性的重要因素.

## 2.2 低负载阶段 SI 算法

VCP-BE 低负载阶段主要目标是网络快速收敛于高利用率. 在连接初始阶段, 拥塞窗口较小, 分组发送间隔较大, 预测带宽将小于当前可用带宽, 此时 VCP-BE 采用 MI 乘性增长方式快速增加拥塞窗口; 随着发送速率逐步增大, 带宽预测值将逐步收敛于准确值, VCP-BE 根据现阶段带宽预测值将发送速率调整到网络可允许的程度, 这样既保证了拥塞窗口的快速增长, 又不至于使窗口增长过于激进. 低负载阶段 SI 拥塞窗口控制算法如公式(5)所示. 其中  $seg\_size$  为分组大小,  $RTT_{min}$  是最小 RTT 值,  $cwnd_{n-1}$  为 VCP 采用 MI 乘性增加拥塞窗口算法,  $\varepsilon$  取默认值 0.0625.

$$cwnd_n = MAX[cwnd_{n-1} * (1 + \varepsilon), (\frac{RTT_{min} * BW_n}{seg\_size})] \quad (5)$$

低负载阶段为使带宽预测值较快地跟踪当前可用带宽, 需采用较小的计算时间间隔; 高负载阶段为避免带宽预测偏高加速网络拥塞, 需采用较长的计算时间间隔. 为达到上述控制目标, VCP-BE 低负载阶段时间间隔调节因子  $\theta$  采用当前带宽预测值和已完成吞吐量比值来表示, 如公式(6)所示.

$$\theta = \frac{BW_{n-1}}{(cwnd_{n-1} * seg\_size) / RTT} \quad (6)$$

传输起始阶段, 较小的带宽预测值使窗口采用 MI 乘性增长, 吞吐量大于预测带宽, 此时  $\theta$  小于 1 使计算时间间隔减小, 从而加快带宽预测响应性; 随着预测带宽和吞吐量越来越接近,  $\theta$  逼近 1, 计算时间间隔将越来越大, 使带宽预测值平滑接近高负载.

## 2.3 高负载阶段算法

VCP-BE 高负载阶段主要目标是快速收敛于公平共享带宽. 当网络处于稳定状态时, VCP-BE 使用 AI-II-MD 算法. 远离过载时, VCP-BE 采用 AI 算法并设置较大增长因子  $\alpha$  以确保拥塞窗口较快增长; 当接近过载时, VCP-BE 采用与窗口平方根成反比例增长的 II 算法以增加速率变化平滑性, 从而减小网络振荡. 为解决带宽较大时 II 算法较慢的增长幅度依然导致协议在接近过载状态持续占用带宽较长时间, 从而影响公平收敛速度问题, VCP-BE 通过计算当前带宽占用比率(即当前发送速率与可用带宽预测值比值), 并依据当前带宽占用比率及 ECN 位将高负载阶段分为远离过载与接近过载两种状态, 并分别采用 AI 与 II 两种不同拥塞窗口控制算法. 当网络过载后, 采用 MD 算法将窗口减小到原窗口的  $\beta$  倍( $\beta$  取 0.875), 保证窗口退避后不再进入 MI 阶段. VCP-BE 高负载阶段算法实现如下:

$R_i = (cwnd_{i-1} * seg\_size) / RTT_{min}$  //源端当前发送速率

$K_i = R_i / BW_{i-1}$  //当前带宽占用比率

if ( $K_i < L_{thresh}$  && ECN=(01)<sub>2</sub>) //远离过载

$cwnd_i = cwnd_{i-1} + \alpha / cwnd_{i-1}$  //AI 控制算法

else if ( $K_i >= L_{thresh}$  && ECN=(10)<sub>2</sub>) //接近过载

$cwnd_i = cwnd_{i-1} + \alpha / (cwnd_{i-1})^{3/2}$  //II 控制算法

else if ( $K_i >= L_{thresh}$  && ECN=(10)<sub>2</sub>) //过载

$cwnd_i = cwnd_{i-1} * \beta$  //MD 控制算法

其中  $R_i$  为第  $i$  条流的发送速率,  $K_i$  为当前带宽占用比率,  $L_{thresh}$  为转换阈值(一般取经验值 0.95),  $\alpha$  为窗口加性增长因子, VCP-BE 将  $\alpha$  设置为 5 以加快协议收敛速度, 而 VCP 由于不能对高负载网络状态做出更精细的判断,  $\alpha$  取保守值 1. 为增加高负载阶段带宽预测准确性, 时间间隔调整因子  $\theta$  取固定值 1, 为保证不同 RTT 流之间的公平性因子,  $\varepsilon$  和  $\alpha$  仍按(2)式和(3)式根据当前 RTT 进行修正.

## 3 VCP-BE算法仿真及性能分析

将 VCP-BE 加入 NS2 核心组件, 测试不同瓶颈带宽、不同往返时延下协议公平收敛时间, 以评估带宽大小、往返时延对公平收敛速度的影响. 采用单瓶颈哑铃型实验拓扑, 缓存大小为带宽延时乘积. 实验中引入两条 TCP 流, flow2 在 flow1 达到稳定状态后开始传输, 瓶颈带宽从 100Mbps 变化到 1Gbps, RTT 从

20ms 变化至 200ms,  $L_{thresh}=0.95$ ,  $\alpha=5$ ,  $\beta=0.875$ , 其他参数均按 NS2 默认值设置. 不同瓶颈带宽公平收敛时间如图 3 所示.

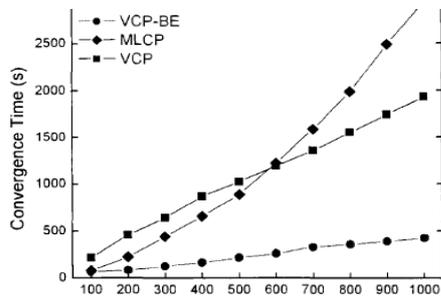


图3 不同瓶颈带宽的公平收敛时间

高带宽网络特性使协议的公平收敛时间随着带宽的增大而加长. 低负载阶段 VCP-BE 采用可用带宽预测源端可允许的发送速率, 快速收敛于高利用率, 高负载阶段采用 AI-II-MD 方式有效利用带宽, VCP-BE 的公平收敛时间远小于 MLCP 和 VCP. MLCP 虽在带宽小于 600Mbps 时具有比 VCP 更短的收敛时间, 但随着带宽的增大, 其收敛时间远大于 VCP 和 VCP-BE.

不同 RTT 对协议公平收敛时间影响如图 4 所示. 在不同 RTT 下 VCP-BE 的收敛时间都远小于 VCP, 且在 RTT 较小时明显优于 MLCP. 由于 VCP-BE 及 VCP 采用了固定的与当前 RTT 无关的  $tp$  值来调节拥塞窗口, 随着 RTT 增长收敛时间都增长缓慢. 针对显式反馈的不同负载程度, VCP-BE 采用了不同时间间隔调节方法预测可用带宽, 为源端拥塞窗口控制提供了合适的网络状态信息, RTT 对 VCP-BE 公平收敛时间的影响最

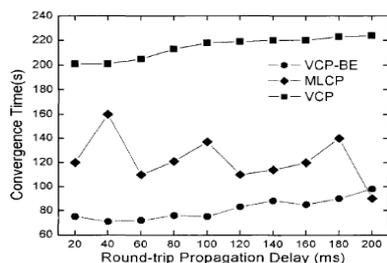


图4 不同 RTT 的公平收敛时间

小. MLCP 跟随当前 RTT 变化调整拥塞窗口增长幅度, 在不同 RTT 下收敛时间在一定范围内波动.

#### 4 结论及下一步工作

本文提出的基于带宽预测的 VCP 改进协议 VCP-BE 获得了比 VCP、MLCP 更好的性能, 能够真实有效地反映网络拥塞, 显著提高算法在高带宽长时延网络环境的公平收敛速度. VCP-BE 通过 TCP 源端进行可用带宽预测, 由于远离网络拥塞发生地, 其带宽预测的准确性受不同网络状态影响较大. 另外, 改进算法中  $\alpha$ 、 $\beta$ 、 $L_{thresh}$  等参数设置一般采用经验值, 算法所使用的静态参数不能随网络状态而动态调整. 因此, 如何对参数进行设置并对参数进行优化选择, 以动态适应不同网络环境, 将是在此基础上值得进一步研究的内容.

#### 参考文献

- 1 Qazi IA, Znati T. On the design of load factor based congestion control protocols for next-generation networks. IEEE INFOCOM. New York, USA. IEEE. 2008. 96-100.
- 2 Xia Y, Subramanian L, et al. One more bit is enough. ACM SIGCOMM Computer Communication Review, 2005, 35 (4): 37-48.
- 3 尹凤杰, 王金科, 岳承君. 高速网络中基于 VCP 拥塞控制算法的改进. 辽宁大学学报, 2010, 37(3): 193-197.
- 4 毛鹏轩, 肖扬. 一种提高 VCP 协议收敛速率的自适应算法. 北京交通大学学报, 2012, 36(2): 30-34.
- 5 陈洁. 高带宽时延网络的拥塞控制协议研究[硕士学位论文]. 长沙: 中南大学, 2011.
- 6 张慧翔, 戴冠中, 等. VCP 协议公平性算法的仿真分析与改进研究. 系统工程与电子技术, 2009, 31(11): 2773-2777.
- 7 Clark DD, Fang W. Explicit allocation of best effort Packet delivery service. IEEE/ACM Trans. on Networking, 2008, 6(4): 362-373.