

一种改善 RTT 公平性的 TCP 慢启动算法

王思思¹, 朱金鹏¹, 王建峰²

(1. 空军指挥学院, 北京 100097; 2. 空军工程大学 电讯工程学院, 陕西 西安 710077)

摘要:针对 Internet 拥塞控制策略在慢启动阶段中,具有不同 RTT 的 TCP 连接在竞争带宽时存在的不公平性,以及慢启动后期窗口增长过快等问题,提出一种改善 RTT 公平性的 TCP 慢启动新算法。新算法在慢启动初期窗口仍保持指数增长,慢启动后期引入与 RTT 成正比的窗口增长控制因子,使 RTT 较大的 TCP 连接窗口增加较快,反之增加较慢。性能分析和仿真结果表明,新算法在慢启动后期减缓了拥塞窗口的增长速度,削弱了 RTT 较小的 TCP 连接竞争带宽的侵略性,在一定程度上保证了不同 RTT 的 TCP 连接共享带宽的公平性,并且保持了标准慢启动算法的端到端语义,简单易实现,具有一定的理论和实际应用价值。

关键词:拥塞控制;TCP;慢启动;RTT 公平性

DOI:10.3969/j.issn.1009-3516.2009.03.019

中图分类号: TP393 **文献标识码:** A **文章编号:** 1009-3516(2009)03-0086-05

TCP 协议采取的是加性增加乘性减少(Additive Increase and Multiplicative Decrease, AIMD)的拥塞窗口(congestion window, cwnd)机制,其中慢启动策略的好坏对其性能有着非常明显的影响^[1]。在慢启动阶段,拥塞窗口 cwnd 的初始值设置为 1 个分组,源端每收到 1 个肯定应答 ACK, cwnd 的值就增加 1 个分组,这样, cwnd 的增长速度与 ACK 的到达速度呈指数关系。假设 $ssthresh = 2^N + 1$, 所有往返时间(Round-Trip Time, RTT)均相等,经过前 N 个 RTT, cwnd 从 1 增加到 $ssthresh/2$, 而第 $N+1$ 个 RTT, cwnd 将从 $ssthresh/2$ 增加到 $ssthresh$ 。可见,在慢启动后期,窗口增长速度过快,尤其最后一个 RTT 的突发流量将对网络的稳定性造成很大的影响,引起网络中多个分组丢失,形成全局同步网络性能下降,与慢启动的设计初衷(即逐渐探测网络带宽)有冲突。

此外,由于 ACK 的到达速度与 RTT 成反比,拥塞窗口 cwnd 的增长速度与 RTT 的长短呈负指数关系。因此,长时延连接与短时延连接相比, TCP 打开窗口的速率慢了许多。当具有不同 RTT 的 TCP 连接共享同一瓶颈链路时, RTT 较小的 TCP 流对网络资源的占用具有较大的侵略性, RTT 较大的 TCP 连接则只能获得它们应得资源的小部分。由此带来相同瓶颈链路上不同 RTT 的 TCP 流之间竞争带宽的不公平性。

针对标准慢启动算法存在的上述问题,研究者提出了平滑慢启动^[2],有限慢启动^[3],分阶段的慢启动改进算法 P-start^[4-5]以及基于余弦函数的新慢启动策略等方法^[6-7]。文献[8]提出使用估计的 $ssthresh$ 以及设定一个安全的 $ssthresh$,但由于网络动态变化,难以保证估计的门限阈值与网络可用带宽的完全匹配。文献[9]针对 GEO 卫星链路,提出在慢启动后期引入窗口增长控制因子,使 RTT 较大的 TCP 流窗口增长较快,反之增长较慢。文献[10]提出在慢启动阶段,每到达 1 个 ACK, cwnd 增加 $aRTT/cwnd$,但参数 a 为经验参数,难以准确确定。考虑到上述方法的局限性,提出一种能改善不同 RTT 的 TCP 连接竞争带宽不公平性并且容易实现的慢启动新算法,记为 SS+(New Slow Start Algorithm to Improve RTT Fairness)。

1 改善 RTT 公平性的慢启动算法

针对标准 TCP 拥塞控制策略的慢启动过程中不同 RTT 的 TCP 流在竞争带宽时的不公平性及慢启动后

* 收稿日期:2008-10-14

基金项目:国防预研基金资助项目(51422010108JB3202);陕西省自然科学基金资助项目(SJ08F14)

作者简介:王思思(1982-),女,工程师,主要从事宽带网络和装备维修保障研究. E-mail: wss8294@126.com

期拥塞窗口增长速度过快2个问题,提出一种改善不同 RTT 的 TCP 连接竞争带宽公平性的慢启动算法(SS+)。新算法设计目标如下:

1)改善相同瓶颈链路中不同 RTT 的 TCP 流之间竞争带宽的不公平性;

2)减缓慢启动后期拥塞窗口增长速度,使 cwnd 越接近 ssthresh 增长越趋于平缓,较平滑地过渡到拥塞避免阶段;

3)保持原有慢启动算法的端到端语义和简单易实现性,只在发送端作修改。

为满足以上设计目标,提出的新算法主要从以下2个方面考虑:①针对具有不同 RTT 的 TCP 连接竞争带宽的不公平性问题,设法让每个 TCP 连接打开窗口的速度逾期自身的 RTT 成正比关系。②针对慢启动后期,窗口增长过快问题,考虑在窗口接近慢启动阈值时,逐渐减慢窗口增加速度,减少多包丢失的发生。

新算法的基本思想是:在连接建立初期,cwnd 的增长方式与标准慢启动相同,即呈指数增长。当 cwnd 增加至 ssthresh/2 时,引入与 RTT 相关的窗口增长控制因子 $(ssthresh - cwnd) (RTT / (RTT_{min} + RTT_{max}))$,使得 RTT 较大的 TCP 连接窗口增长较快,RTT 较小的 TCP 连接窗口增长较慢,即每个 RTT 内拥塞窗口增加 $(ssthresh - cwnd) (RTT / (RTT_{min} + RTT_{max}))$,逐步逼近门限值,直到 $(ssthresh - cwnd)$ 小于设定的调节因子 δ $(ssthresh/2 \geq \delta \geq 2)$,同时将当前 cwnd 设为门限值,从而进入拥塞避免阶段。拥塞窗口 cwnd 的大小表示为:

$$cwnd(t+T) = \begin{cases} 2cwnd(t) & \text{若 } cwnd < ssthresh/2; \\ cwnd(t) + \frac{RTT}{RTT_{min} + RTT_{max}} (ssthresh - cwnd(t)) & \text{若 } ssthresh/2 \leq cwnd(t) < ssthresh \\ & \text{且 } ssthresh - cwnd \geq \delta; \\ ssthresh & \text{其他} \end{cases} \quad (1)$$

式中: t 为当前时刻; T 表示1个RTT的时间; RTT 为当前TCP连接的RTT值; RTT_{min} 和 RTT_{max} 分别为共享同一瓶颈链路的不同TCP连接的最小值和最大值,即最短时延连接和最长时延连接的RTT值。

新算法借鉴了P-Start算法的分阶段调整拥塞窗口思想,在慢启动后期同样设定了调节因子 δ 。但是,新算法与P-Start算法有很大不同,P-Start算法在慢启动后期每个RTT内窗口增加 $(ssthresh - cwnd)/2$,没有考虑不同连接RTT的不同;而新算法在慢启动后期引入与RTT相关的窗口增长控制因子,可根据RTT不同,对不同TCP连接采用不同的拥塞窗口增长速度,即每个RTT内窗口增加 $(ssthresh - cwnd) (RTT / (RTT_{min} + RTT_{max}))$ 。这样,RTT较小的TCP连接窗口增长较慢,RTT较大的TCP连接窗口增长较快,在一定程度上解决了不同RTT的TCP连接共享瓶颈链路的不公平性问题。

2 SS+ 算法性能分析

假设 $ssthresh = 2^{N+1}$,SS+算法与标准慢启动算法相比,具有以下特点:

1)cwnd 增长到 ssthresh/2 之后,cwnd 增长的速度随 RTT 的不同而不同,每个 RTT 内增加 $(ssthresh - cwnd)$ 的 $(RTT / (RTT_{min} + RTT_{max}))$ 倍。可看出,随着 RTT 的减小,cwnd 增长的速度逐渐减慢,这样,RTT 较大的 TCP 流获得较多的增加带宽,RTT 较小的 TCP 流获得较小的增加带宽,缓解了不同 RTT 的 TCP 流之间竞争带宽的不公平性问题。

2)拥塞窗口增长到 ssthresh/2 之后,随着 cwnd 的增大,cwnd 逐渐减缓增长速度。当 $RTT = (RTT_{min} + RTT_{max})/2$, $\delta = gssthresh/8$ 时,SS+算法的慢启动窗口大小变化见图1。cwnd 在第 $N+1$ 个RTT时,从 $ssthresh/2$ 增加到 $3ssthresh/4$;在第 $(N+2)$ 个RTT时,cwnd 从 $3ssthresh/4$ 增加到 $7ssthresh/8$;此时, $(ssthresh - cwnd) \leq \delta = ssthresh/8$,因而在第 $(N+3)$ 个RTT时,cwnd 从 $7ssthresh/8$ 增加到 ssthresh。由图1可看到,在整个慢启动期间,SS+算法所消耗 RTT 次数只比标准慢启动多2次,但新算法在慢启动后期 cwnd 增长趋于缓慢,能较平滑的进入拥塞避免阶段,降低拥塞发生的可能性;同时,在一定程度上减少了慢启动阶段的同一个窗口多个分组丢失的现象,避免了慢启动阶段的重传超时。

此外,随着 RTT 的增大,到达慢启动阈值 ssthresh 所消耗的 RTT 次数减小,由慢启动所需要的时间等于 RTT 与所需要消耗的 RTT 的次数的乘积可知,采用 SS+ 算法在一定程度上缓解了不同 RTT 的 TCP 流竞争带宽的不公平性。

3)SS+算法实现简单、开销小,能够适用于任意 TCP 源端算法。即只比原算法增加了一个比较运算和一个算术运算,对于当前的高速计算机来说,性能上几乎没有影响。

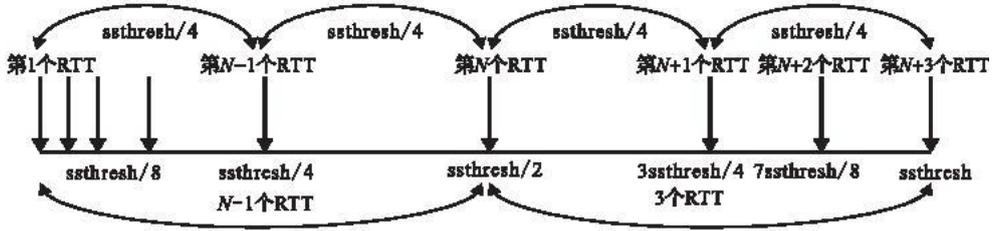


图1 SS + 算法慢启动窗口大小变化 ($RTT = RTT_{max}$, $\delta = ssthresh/8$)

Fig. 1 Congestion window change of SS + algorithm when $RTT = RTT_{max}$ and $\delta = ssthresh/8$

3 仿真结果与分析

网络仿真拓扑结构及相关参数取值见图2。2个路由器之间的链路为瓶颈链路,发送端 $S_1 - S_3$ 以 Reno 方式分别向接收端 $D_1 - D_3$ 传输 FTP 流。所有链路均为双向链路,为了更好的比较,假设各发送接收端对的 RTT 成比例关系,取 $RTT_1 = 400$ ms, $RTT_2 = 200$ ms, $RTT_3 = 100$ ms。瓶颈链路采用 Drop-tail 算法,网关采用 FIFO 的分组调度算法,网关路由缓存容量设为 50 个分组, $\delta = 2$ 。分组大小设为 1 kB,则 ssthresh 为 64 个分组。仿真运行 100 s。

图3给出了 SS + 算法中拥塞窗口增长与所需 RTT 次数的关系。图4和图5分别给出了标准慢启动和 SS + 算法中不同 RTT 流的 cwnd 随时间变化情况。

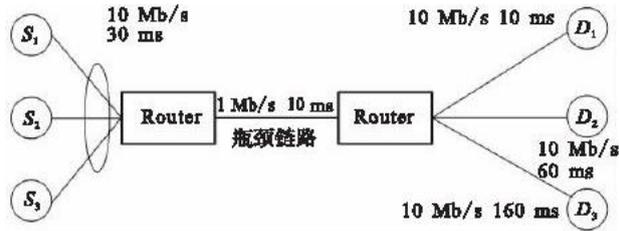


图2 网络仿真拓扑结构图

Fig. 2 Simulation topology

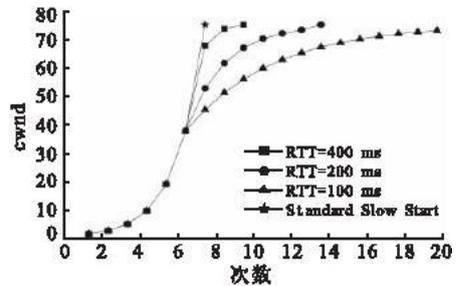


图3 SS + 算法拥塞窗口大小与 RTT 次数关系图

Fig. 3 Relation of congestion window size and the number of RTT

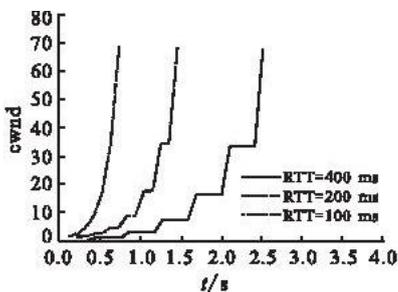


图4 标准慢启动中拥塞窗口随时间的变化

Fig. 4 Congestion window changing with time of standard slow start

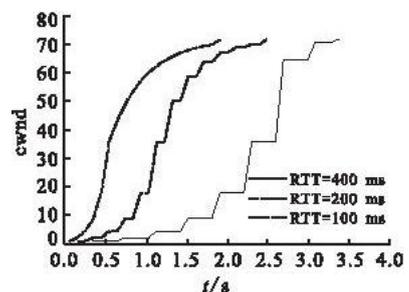


图5 SS + 算法中拥塞窗口随时间的变化

Fig. 5 congestion window changing with time of SS + algorithm

由图3可见,慢启动初期,新算法与标准慢启动算法均用了5个RTT使cwnd达到ssthresh/2;但当cwnd > ssthresh/2时,标准慢启动只用了1个RTT就达到了ssthresh;而新算法中,当RTT = 400 ms时,经过3次RTT达到ssthresh;当RTT = 200 ms时,经过7次RTT达到ssthresh;而当RTT = 100 ms时,需要经过13次RTT才达到ssthresh。由此可看出,新算法在慢启动后期减缓了拥塞窗口增长速度的同时,也削弱了RTT较小的TCP连接在竞争带宽时的侵略性,因为cwnd从ssthresh/2增长到ssthresh所用的时间增加了,所占用的带宽增长速度就减缓了。

由图4、图5可见cwnd从ssthresh/2增长到ssthresh,当RTT = 400 ms时,标准慢启动耗时0.4 s,SS + 耗

时 1.2 s;当 $RTT = 200$ ms 时,标准慢启动耗费 0.2 s,SS + 耗时 1.4 s;当 $RTT = 100$ ms 时,标准慢启动耗时 0.1 s,SS + 耗时 1.4 s。即当 RTT 分别为 400 ms,200 ms,100 ms,cwnd 从 $ssthresh/2$ 增长到 $ssthresh$ 时,标准慢启动和 SS + 耗时比值分别为 4:2:1 与 0.86:1:1;整个慢启动期间耗时比值分别为 4:2:1 与 1.8:1.3:1。可见,SS + 在一定程度上保证了不同 RTT 的 TCP 连接共享带宽的公平性。

4 结束语

针对具有不同 RTT 的 TCP 连接竞争相同瓶颈带宽时存在的不公平性问题;同时考虑慢启动后期拥塞窗口增长速度过快问题,对标准 TCP 慢启动算法进行改进。新算法 SS + 在连接建立初期,拥塞窗口仍保持指数增长;当 cwnd 增加至 $ssthresh/2$ 时,cwnd 的增加与 RTT 成正比,即每个 RTT 内 cwnd 增加 $(ssthresh - cwnd) \cdot (RTT / (RTT_{min} + RTT_{max}))$ 。性能分析和仿真结果表明,SS + 算法在慢启动后期不仅减缓了拥塞窗口的增长速度,与慢启动算法的设计初衷相吻合,而且削弱了 RTT 较小的 TCP 流在竞争带宽时的侵略性,在一定程度上保证了具有不同 RTT 的 TCP 连接共享带宽的公平性。

参考文献:

- [1] 杨 征, 吴玲达. 一种适合于宽带卫星网络的均滑式慢启动策略[J]. 小型微型计算机系统,2007,28(6): 987-990.
YANG Zheng, WU Lingda. A Smooth Slow-start for Broadband Satellite Networks [J]. Mini-Micro System,2007,28(6): 987-990. (in Chinese)
- [2] Wang H, Xin H, Reeves D S, et al. A Simple Refinement of Slow-Start of TCP Congestion Control [C]//In Proc of ISCC 2000. Los Alamitos:[s. n.],2000:98-105.
- [3] Floyd S. Limited Slow Start for TCP with Large Congestion Window [R]. RFC 3742,2004.
- [4] CHEN Zhigang, DENG Xiaoheng, ZHANG Lianming, et al. A New Parameter-config Based on Slow-Start Mechanism [J]. Journal of Communication and Computer,2005, 12(5): 56-62.
- [5] 邓晓衡, 陈志刚, 张连明, 等. TCP Yuelu: 一种基于有线/无线混合网络端到端的拥塞控制机制[J]. 计算机学报, 2005,28(8): 1342-1350.
DENG Xiaoheng, CHEN Zhigang, ZHANG Lianming, et al. TCP Yuelu: An End-to-End Congestion Control Mechanism for Wired/Wireless Hybrid Networks [J]. Chinese Journal of Computers, 2005, 28(8): 1342-1350. (in Chinese)
- [6] 刘文远, 冯 波, 龙承念, 等. 一种新的 TCP 拥塞控制慢启动策略[J]. 小型微型计算机系统,2005, 26(1): 23-25.
LIU Wenyuan, FENG Bo, LONG Chengnian, et al. New Strategy of Slow-Start for TCP Congestion Control [J]. Mini-Micro Systems, 2005, 26(1): 23-25. (in Chinese)
- [7] 茹新宇, 刘 渊. 一种新的 TCP 慢启动策略[J]. 计算机工程, 2008, 34(5): 116-119.
RU Xinyu, LIU Yuan. COS-Slow-Start: A New Strategy of TCP Slow-Start [J]. Computer Engineering, 2008, 34(5): 116-119. (in Chinese)
- [8] Wang R, Pau G, Yamada K, et al. TCP Startup Performance in Large Bandwidth Delay Networks[C]//Proc of the IEEE Infocom 2004. Piscataway:IEEE, 2004:796-805.
- [9] 王建峰. GEO 卫星链路 TCP 拥塞控制算法研究[D]. 西安:空军工程大学,2008.
WANG Jianfeng. TCP Congestion Control Algorithms Based on GEO Satellite Links [D]. Xi'an:Air Force Engineering University, 2008. (in Chinese)
- [10] 黄 敏, 张鹏丽, 段 焰. 基于 Petri 网的 Internet 拥塞控制慢启动改进算法[J]. 长沙交通学院学报, 2007, 23(2): 69-72.
HUANG Min, ZHANG Pengli, DUAN Yan. An Improved Slow-start Algorithm of Internet Congestion Control Based on Petri Net [J]. Journal of Changsha Communications University, 2007, 23(2): 69-72. (in Chinese)

(编辑:徐楠楠)

A New TCP Slow Start Algorithm to Improve RTT Fairness

WANG Si-si¹, ZHU Jin-peng¹, WANG Jian-feng²

(1. Air Force Command College, Beijing 100097, China; 2. Telecommunication Engineering Institute, Air Force

Engineering University, Xi'an 710077, China)

Abstract: To solve the problems such as the unfairness existing when standard TCP connects with different RTT competing bandwidth, the window over - fast growth caused in the late phase of slow start and etc. , a new TCP slow start algorithm for improving, termed SS + , is proposed. In the early slow start phase, using SS + can keep the congestion window exponential grow; in the late phase, the window control growth factor in direct proportion to RTT is introduced to increase the congestion window of TCP connecting with larger RTT faster, otherwise the increase will be slower. Performance analysis and simulation results show that by using SS + algorithm the congestion window growth rate is slowed down and the TCP connection with smaller RTT in the aggressive bandwidth competition is weakened in the late slow start phase. To a certain extent, by using SS + algorithm a fair sharing of available bandwidth between TCP connections with different RTT is obtained. And SS + algorithm is easy to realize and only requires the TCP source revised. This SS + algorithm has some theoretical and practical application value.

Key words: congestion control; TCP; slow start algorithm; RTT - based fairness

(上接第 49 页)

ISAR Imaging Analysis of Ship Target

ZHU Zi - jian¹, XU You², MA Zhi - qiang¹, LIU Ling - xia¹

(1. Telecommunication Engineering Institute, Air Force Engineering University, Xi'an 710077, China 2. Department of Training, Air Force Engineering University, Xi'an 710051, China)

Abstract: ISAR imaging has been the focus in recent years. Based on Doppler, the one - dimensional imaging is obtained. Through range alignment and auto - focus, the target can be equivalent to a rotation model. Based on the rotation model, the movement of the ship target will be divided into a horizontal component and a moving component. The first one is useless in the ISAR imaging, so it must be compensated. This is the important step. Based on the rotation model, for the moving smoothly ship, the two - dimensional ISAR imaging can be got through range alignment and auto - focus. But the apriori knowledge of ship target motion is generally unknown, because of their uncooperative moving, the echoes have an obvious time - variant character. Based on the periodicity of Doppler frequency, the whole dwell time is divided into several intervals, then, based on the above, the instantaneous ISAR imaging of every time interval can be got and simulations are done to verify the method. The verification result shows that the method is effective for the non - stationary ship imaging.

Key words: inverse synthetic aperture radar; rotation - model; movement compensation; Instantaneous ISAR Imaging