TCP/IP 拥塞控制算法研究

江 艳,李宝林
(西华师范大学 计算机学院,四川 南充 637002)

摘 要:TCP/IP 拥塞控制的研究是计算机网络最活跃的领域之一,其目的是解决 TCP/IP 拥塞控制所面临的自相似性问题、效率问题、公平性问题。文中从控制理论角度对开环与闭环控制拥塞模式进行了阐述,从造成 TCP/IP 拥塞原因入手,并对其相应的策略做出改进,此外还对闭环拥塞控制的基本策略:许可控制、依赖路由选择、隐式拥塞信令、显式拥塞信令及抑制包策略进行了分析并提出了相应的抑制包策略改进方法,并取得了很好的成效。

关键词:拥塞控制;开环控制;闭环控制;TCP/IP
中图分类号:TP301.6 **文献标识码**:A **文章编号**:1673-629X(2015)01-0119-04
doi:10.3969/j.issn.1673-629X.2015.01.027

Research on TCP/IP Congestion Control Algorithm

JIANG Yan,LI Bao-lin
(College of Computer,China West Normal University,Nanchong 637002,China)

Abstract:Research on TCP/IP congestion control is one of the most active fields in the computer network,its purpose is to solve the problem of self similarity,efficiency,fairness in the TCP/IP congestion control. From the perspective of control theory,the open-loop and closed-loop congestion control model are described,starting from the TCP/IP congestion reasons,and the corresponding strategies to improve. In addition,for the basic strategy of closed loop congestion control,such as admission control,routing relying selection,explicit and implicit congestion signaling,and inhibition of package strategy are analyzed and put forward the corresponding control strategy improved methods,which has achieved good results.

Key words:congestion control;open-loop control;closed-loop control;TCP/IP

0 引 言

有关 TCP/IP 拥塞研究的伊始,源于在拥有大量网络传输数据的因特网中怎样准确高效的传输。从 20 世纪 80 年代中期开始,Internet 用户呈指数级增长趋势,其主要网络协议 TCP/IP 的应用也越来越普及,但在网络应用和新的网络技术不断发展的背景下,网络拥塞问题也开始凸显[1]。另外,随着互联网的不断扩大和网络用户的持续增加,新旧技术的并存、不同版本之间的兼容以及网络的异质性差别导致了许多不匹配的问题。如链路速度不匹配、网络各个中间节点处包的到达速度与处理速度不匹配等引发了排队等待和拥塞现象。如今是网络信息经济时代,用户对网络服务质量要求更高,同时影响网络性能的拥塞控制问题也备受关注。据统计,互联网上 95% 的数据流和 90% 的报文数使用的是 TCP/IP 协议,此外,由于当前网络拥塞控制的大部分工作是由 TCP 协议完成的,因此,基于互连的 TCP/IP 网络协议中的拥塞控制机制研究也就对控制网络拥塞有着特别重要的意义,并且它一直以来都是互联网拥塞控制问题的一个研究热点。

1 TCP/IP 的拥塞控制

1.1 拥塞的定义及产生原因

拥塞指的是在包交换网络中由于传送的包数目太多,而存储转发节点的资源有限造成网络传输性能下降的情况。

网络产生拥塞的根本原因在于对网络中某一资源的需求超过了该资源所能提供的可用部分。在网络发生拥塞时,会导致吞吐量下降,严重时会发生“死锁”现象。具体表现为数据包时延增加、丢弃概率增加、上层应用系统性能下降等[2-3]。具体可以用图 1 来进行

收稿日期:2014-02-06 **修回日期**:2014-05-11 **网络出版时间**:2014-10-23
基金项目:四川省科技支撑计划项目(2012GZ0020)
作者简介:江 艳(1989-),女,硕士,研究方向为计算机网络、物联网应用技术;李宝林,博士,副教授,研究方向为物联网应用技术、电子商务。
网络出版地址:http://www.cnki.net/kcms/detail/61.1450.TP.20141023.1102.028.html

描述。

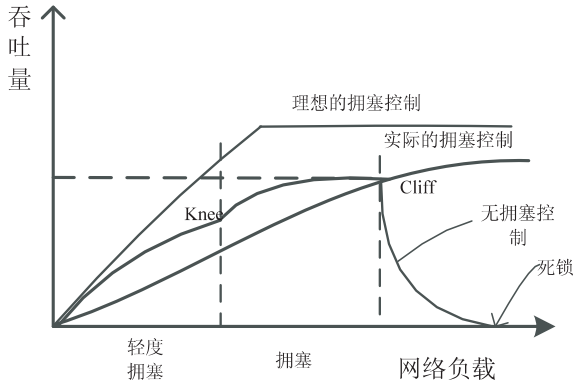

图 1 网络负载和吞吐量模型图

当网络具有理想的拥塞控制机制且吞吐量未达到饱和之前,网络吞吐量等于网络提供的负载,吞吐量曲线为 45°斜线。然而当负载超过一定限度即吞吐量达到饱和时,吞吐量不再呈线性增长而保持为水平线;但是在实际的网络中,拥塞控制并不会像想象的那么完美。由图 1 可知,当网络负载增大时,网络吞吐量出现缓慢增长并在达到某一吞吐量后保持水平。这就说明在一开始到吞吐量达到饱和之前就有很多数据分组被丢弃;然而在无拥塞控制的网络环境中,前期吞吐量随着负载的增长而增长;当到达临界点 Knee 时,吞吐量增长速度减缓且响应时延增加;而当吞吐量达到 Cliff 点时,若网络负载继续增加会导致吞吐量下降。然而当网络处于极度拥塞时,通信链路上传输的数据基本上都是重传数据,从而导致链路的有效吞吐量下降。由此引起的恶性循环一定会导致局部死锁,甚至可能导致整个网络系统瘫痪,最终迫使网络的有效吞吐量趋近为零。

拥塞产生的直接原因有以下三点[4-11]:

(1)存储空间不足。当某个节点缓冲区的容量大小不足以容纳到达该节点的数据时,会导致到达该节点的数据因无空间缓存而被丢弃。然后由于无限制地扩大该节点的缓存容量能让到达该节点的分组在缓存队列中进行排队,从表面上看似乎缓解了分组被丢弃的现状。但事实上这一矛盾一直存在并被激化。因为输出链路的容量和处理机的速度并没有得到同步提高,因此,在该队列排队等待的大部分数据分组会在此等待很长的时间,使得数据发送端误认为该分组已被丢弃,从而迫使其对该分组进行重传。

(2)带宽容量受限。据不完全统计,大多数网络拥塞的发生都是因为将高速的数据流输入到低速链路上。又因为在网络通信过程中,网络的信道带宽与它的数据传输能力存在基本的稳定关系,因此,当进入网络的数据流量大于信道带宽时,会迫使流经该节点的数据不断被压入缓存进行等待,从而使网络发生拥塞。

(3)处理机性能受限。路由器中的 CPU 主要执行缓存区排队、更新路由表、进行路由选择等功能,如果其工作效率不能满足高速链路的需求,就会造成网络拥塞。

1.2 拥塞控制的目的及策略

1.2.1 拥塞控制的目的

基于上述三个原因,如果单单从提高其中某一部分的部分性能,例如单一地增加存储空间或者是加快处理器速度,则根本无法避免拥塞,其结果是将拥塞的瓶颈问题转移到网络的其他部分中去。然而理想的拥塞控制策略是采取有效措施控制输入到网络链路上的数据流量,使其尽量小于网络负载临界点 Cliff,并保持高速高效率的运行状态。事实上大多数时候网络的运行状态都不会处于理想状态,然而为了在不影响用户正常通信的情况下适当增加通信量,可以在网络吐吞量下降时采取适当的拥塞避免策略,使整个网络的负载在如图 1 所示的 Knee 处震荡。但是当突发的通信量使网络负载接近 Cliff 临界点时,必须采用拥塞控制策略,确保网络资源的使用率达到最佳。

1.2.2 拥塞控制的基本策略

由于计算机网络是一个很复杂的系统,故考虑从控制理论的角度来看待拥塞控制的问题。这样不管从哪个方面看,开环控制和闭环控制都是拥塞控制的两个主流,其中基于开环控制的网络一经搭建运行就不再进行修改,这就要求提供的网络环境良好且不发生拥塞。但是由于网络环境瞬息万变,根本无法确保最初设置好的理想网络环境不会受外界因素的影响,因此选择开环控制策略不一定能达到拥塞控制的目的。相反建立在反馈链路上的闭环拥塞控制或许是不错的选择,因为基于它的显式反馈和隐式反馈都有一套相对较完善的拥塞控制方法,其中隐式反馈能通过局部信息推断链路是否发生拥塞而显式反馈是通过检测到的拥塞节点向数据发送端发送警告包的形式来控制拥塞。接下来着重介绍闭环系统的三个组成部分。

(1)监视系统:通过检测丢弃包的比例、平均队列长度、超时和重发包的数量、平均包延迟来检测何时何地发生拥塞;

(2)反馈拥塞信息:通过检测到拥塞的路由器向源端发送警告包、设置包中某位或某字段通知它的邻接节点、主机或路由器定期发送探测包显示询问拥塞状况;

(3)调整系统操作:通过增加系统资源和降低负载来减少拥塞。

TCP 的拥塞控制采用的是基于窗口的端到端的闭环控制方式[12],接下来将从端系统的角度来分析各种算法。

(1)许可控制策略。基本思想是当检测到网络虚电路中一旦出现拥塞节点,便不再对该拥塞节点创建任何虚电路,直到拥塞解除。具体如图 2 所示。其中黑圈代表拥塞节点,当节点 i、j 处出现拥塞时,网络在创建虚拟电路时会直接取消该节点中的所有虚拟电路[13-14]。

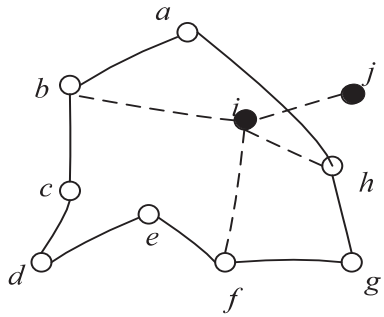

图 2　许可策略虚拟电路图

(2)依赖路由选择策略。基本思想是建立新的虚电路绕开拥塞区域;利用链路延迟信息影响新包产生速率。

(3)隐式拥塞信令策略。基本思想是源端通过检测到传输时延的增加(以反压效应或者目的站对源站的回应时间明显延长)以及连续的分组的丢弃[15],以此判定网络出现拥塞并据此减缓网络流量消除网络拥塞。而基于隐式信令的拥塞控制不需要任何网络中间节点的参与,可直接通过端系统来完成。因此,对于在以数据报交换的网络和基于 IP 的互联网中,采用隐式拥塞信令策略控制网络拥塞是一种不错的选择。此外,面向连接的网络也常常使用此策略。

(4)显式拥塞信令策略。此策略的目的是使得网络中的可用容量得到充分利用,但同时以公平的方式对拥塞作出及时反应;此策略适用于面向连接的网络;其拥塞信令的传递方式分为正向和反向。具体流程如图 3 所示。

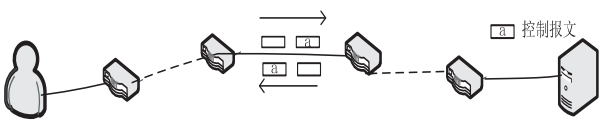

图 3　拥塞信息传递流程图

(5)抑制包策略。基本思想是拥塞的节点产生抑制包,并且传给源端以限制通信流量。而其实现技术是周期性地对线路的瞬间利用率进行取样 f,可得到 U 的近似值。

假设线路的利用率用实型变量 U 表示,U 的取值范围在 0.0 ~ 1.0 之间;a 为常数,代表路由器忘记历史的最快速度。

当 a 等于 0 时说明线路的利用率得到了一个新的取样;当 a 的值为 1 时,线路的利用率没有发生改变,具体如公式(1)所示:

$$a=\begin{cases}0, u_{\text{new}}=f\\ 1, u_{\text{new}}=u_{\text{old}}\end{cases} \tag{1}$$

抑制包在两个产生节点处的工作流程:

(1)在路由器处:路由器通过监视输出线路和其他资源的利用率→当利用率达到某个阈值时相应的线路进入“警告”状态→检查每个新到的包,如果其输出线路处于“警告”状态则向包的源端发送一个抑制包→将新包打上标记以免为此产生更多的抑制包;

(2)在源端:收到抑制包后以 X 比率减少发往该目的地的报文→等待一段时间后再检查是否有新抑制包的到来→如果有来自同一目的地的抑制包,应再次降低发送速率否则应加大通信量。

如果收到抑制包的源端不采取措施来减少发往拥塞点的包,系统将会采取公平队列的方式来避免拥塞的发生。其优势在于即使源端发送更多的包也得不到更多的机会。设计流程是路由器为每条输出线路设置多个队列对应于每个源端,当输出线路空闲时,路由器循环扫描各个队列,一次从各个队列中取出下一个包发送。具体如图 4 所示。

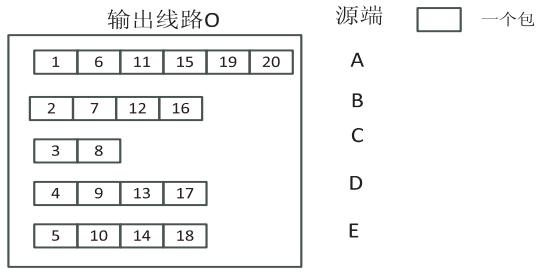

图 4　路由器发送包示意图

即当线路空闲,可以发送数据包时,路由器会依次扫描源端 ABCDE,并从各个源端队列中分别取出下一个数据包依次发送,此时数据包的发送顺序是 1、2、3、4、5。

然而使用上述方式发送数据包,会存在有些很小的数据包等待时间过长的问题。基于上述问题,提出了公平队列的改进方式,即以包的长度作为发送的次序,这样可以解决因发送时延而导致信息无法迅速传达的问题。具体可参照图 4 进行分析,由该图可知数据包的发送次序依次为:CBDEA。但这种方式的公平队列也存在长的数据包占用较大带宽的问题,为此可以通过加权队列的方式做进一步改进,即给不同的主机以不同的优先级,优先级高的可每节拍扫描两个或更多个字节。

一般的抑制包的工作流程是源端以 X mbps 的速度向接收端发送数据包,在某一个节点 d 产生拥塞,在该节点处立即向源端发送抑制包,但此抑制包并不会瞬间就到达源端,而是逐个节点的传送,传送的过程中会经过一段时间,但在这段时间内源端并不知道接收

点处已经发生拥塞,仍以相同的速度向接收端发送数据包,这样在抑制包为到达之前网络拥塞不会减缓反而会加剧[16],具体流程如图 5 所示。

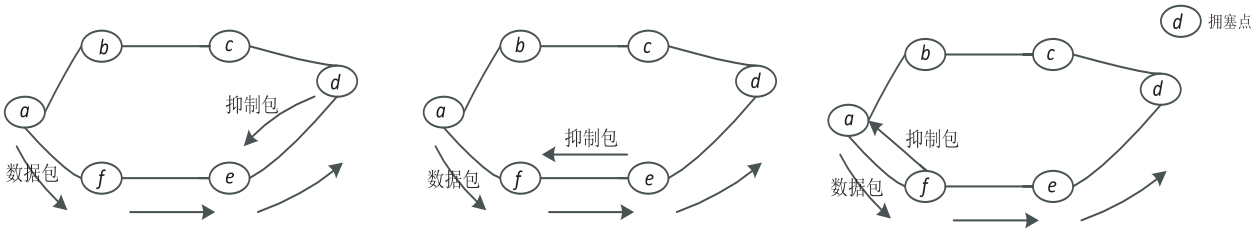

图 5 一般抑制包工作流程图

为此,进一步提出了跳-跳(hop-to-hop)的抑制包改进方法避免了上述拥塞加剧问题。具体工作模式是在发送数据包的过程中一旦发现有拥塞节点,拥塞节点会立即往源端发送抑制包,此抑制包会逐一节点起到抑制作用,也就是说从 d 点发出的抑制包首先会到达 e 点,此时 e 收到抑制包立即减缓从 e 点发往 d 点的数据包的速率。此后每个节点处逐一减缓,这样就从根本上缓解了网络拥塞问题。

2 结束语

随着互联网的广泛应用,其健壮性越来越依赖于 TCP/IP 的拥塞控制,而当前对 TCP/IP 拥塞控制的研究大多是在特定网络环境假设的前提下进行,并基于此采取仿真实验的策略对算法进行改进,缺乏理论支撑;此外,网络公平性问题仍然没有得到完全解决。主要表现在对网络资源使用的不公平以及由于使用窗口尺寸和 RTT 大小或者数据包大小的不同导致占用带宽大小不同而产生的不公平;最后,TCP 与 IP 拥塞控制协议的相互作用问题在一定程度上也会影响网络拥塞的产生,因为它们工作在不同的网络协议层,当对拥塞进行控制时,会因为二者的控制机制不同而影响各自的控制效率,从而导致整个拥塞控制机制的效率下降。因此找到一种性能较好的拥塞控制策略来提高网络性能已迫在眉睫。

参考文献:

[1] 冯彦君,孙利民,钱华林,等. MANET 中 TCP 改进研究综述[J]. 软件学报,2005,16(3):434-444.

[2] 罗万明,林 闯,阎保平. TCP/IP 拥塞控制研究[J]. 计算机学报,2001,24(1):1-18.

[3] 王 东,陈 明,张大方. 一种基于多路复用的多媒体流 TCP 友好拥塞控制机制[J]. 电子学报,2006,34(3):567-572.

[4] 陈 晶,郑明春,孟 强. 基于路由器的网络拥塞控制策略研究[J]. 计算机工程与科学,2002,24(4):24-27.

[5] 潘永辉,张海朝,潘世超. 延迟预警在 TFRC 拥塞控制中的应用[J]. 计算机工程,2009,35(14):114-116.

[6] 刘宝宝,徐国雄,卜应敏. 移动 IPv4 与 IPv6 工作机制的研究[J]. 计算机技术与发展,2011,21(4):126-128.

[7] Paxson V, Floyd S. Wide area traffic: the failure of Poisson modeling[J]. IEEE/ACM Transactions on Networking, 1995, 3(3):226-244.

[8] Hasegawa G, Matsuo T K, Murata M, et al. Comparisons of packet scheduling algorithms for fair service among connections on the Internet[C]//Proc of IEEE INFOCOM 2000. Tel Aviv: IEEE Computer Society, 2000:1253-1262.

[9] Floyd S, Fall K. Promoting the use of end-to-end congestion control in the Internet[J]. IEEE/ACM Transaction on Networking, 1999, 7(4):458-472.

[10] 李 强,张新荣. 基于延迟抖动分析的 TCP 友好拥塞控制算法[J]. 计算机工程与科学,2007,29(6):18-20.

[11] 梁 柱. 组播拥塞控制策略设计与仿真研究[J]. 重庆邮电大学报:自然科学版,2009,21(5):642-646.

[12] Morris R. Scalable TCP congestion control[C]//Proc of IEEE INFOCOM 2000. Tel Aviv: IEEE Computer Society, 2000: 1176-1183.

[13] Padhye J, Towsley D, Firoiu V, et al. Modeling TCP throughput: a simple model and its empirical validation[C]//Proc of ACM SIGCOMM. [s. l.]: ACM, 1998.

[14] Hamann T, Walrand J. A new fair window algorithm for ECN capable TCP (New-ECN)[C]//Proc of IEEE INFOCOM 2000. Tel Aviv: IEEE Computer Society, 2000:1528-1536.

[15] 赵 炯,张树京. TCP 拥塞控制分析模型[J]. 同济大学学报:自然科学版,2004,32(6):786-791.

[16] 汪小帆,孙金生,王执铨. 控制理论在 Internet 拥塞控制中的应用[J]. 控制与决策,2002,17(2):129-134.

TCP/IP拥塞控制算法研究

作者： 江艳， 李宝林， JIANG Yan， LI Bao-lin
作者单位： 西华师范大学 计算机学院,四川 南充,637002
刊名： 计算机技术与发展 ISTIC
英文刊名： Computer Technology and Development
年，卷(期)： 2015(1)