

一种基于 EPN 的拥塞价格机制^{*}

陶俊¹ 帅典勋^{1,2}

(华东理工大学计算机科学与工程系 上海 200237)¹

(清华大学智能技术与系统国家重点实验室 北京 100084)²

摘要 网络发生拥塞的根本原因在于用户需求大于资源供给。利用价格机制中价格随供求关系而波动,又反过来影响用户需求的原理,可以均衡负载,抑制拥塞。针对网络中价格波动及价格传递的问题,本文提出了一种基于 EPN (Explicit Price Notification) 的拥塞价格机制,该机制通过路由器的平均队列长度与队列阈值的关系来反映用户需求的变化,并将价格调整的信息快速传递给用户,激励用户做出正确的资源决策,使网络维持在高吞吐量,低拥塞的状态。同时通过仿真试验分析了队列阈值参数对网络性能的影响。

关键词 价格机制,供求关,RED/ECN,拥塞

EPN-based Congestion Pricing Mechanism

TAO Jun¹ SHUAI Dian-Xun^{1,2}

(Department of Computer Science and Technology, East China University of Science and Technology, Shanghai 200237)¹

(State Key Laboratory of Intelligence Technology and System, Tsinghua University, Beijing 100084)²

Abstract The basic reason of network congestion is the demand of user exceeds the supply of resource. In price mechanism, price is waved with the relation between supply and demand and on the other side effect the supply of user, so it can balance load and control congestion in network. In this paper, a kind of congestion price mechanism based on EPN (Explicit Price Notification) is proposed, which reflects the change of user demand through the relation between the average queue length of route and queue threshold, and transmit the information of price adjustment to user very quickly, in order to promote user to make the sound decision on resource and then to keep the network in high throughput and low congestion state. And by means of simulation, the influence of queue threshold to network efficiency is analyzed.

Keywords Pricing mechanism, Supply-demand relations, RED/ECN, Congestion

1 引言

计算机网络在过去十几年内经历了爆炸式的发展,随之而来的是越来越严重的拥塞问题。文[1]认为网络产生拥塞的根本原因在于用户(或端系统)提供给网络的负载大于网络资源容量和处理能力。从经济学的角度来看,即是用户对资源的需求大于供给,造成了供不应求的局面。因而近年来,一些学者提出将价格机制引入到带宽资源的分配中,利用价格的波动来反映带宽的稀缺程度,并激励用户做出正确的资源决策,以减轻网络拥塞,维持网络的效率。由于价格动态地反映了网络的拥塞程度,故又称为拥塞价格机制(congestion pricing mechanism)或自适应价格机制(adaptive pricing mechanism)^[2]。这类机制有灵巧市场(Smart Market)^[3]方法、动态能力协约(Dynamic Capacity Contracting)^[4]、公平定价策略(Proportional Fairness Pricing)^[5]等。这些方法存在着复杂程度过高、拥塞检测不精确等方面的问题,失去了实际应用的意义。本文基于随机早期检测(Random Early Detection, RED)^[6]以及显式拥塞指示(Explicit Congestion Notification, ECN)^[7]的思想,提出了一种基于显式价格指示(Explicit Price Notification)的带宽价格调节方法,并分析了 RED 中相关参数对网络吞吐量、延迟、丢包率等性能的影响。

2 RED/ECN

目前在 Internet 上使用的拥塞控制基本上建立在 TCP 的窗口控制基础上,是以端系统为核心、以丢包为拥塞指示的一种机制。在网络内部(如路由器或交换机的 IP 层)只采用最简单的先进先出(FIFO)及“去尾”(drop-tail)策略进行调度与队列管理。当包到达时缓存已满,则进行简单的丢包处理。而端系统的 TCP 层基于重传计数器(RTO)及 ACK 确认帧来判定网络的拥塞状态,并对相应的控制参数进行调整来消除拥塞。TCP 拥塞控制机制符合 Internet 的基本设计思想:将控制机制的复杂性推向网络边缘,在网关上只进行少量的操作。但其同时存在死锁(lock-out)、满队列(full queues)、TCP 全局同步(TCP global synchronization)等重要缺陷^[8]。

为了克服以上缺陷,IETF (The Internet Engineering Task Force) 推荐在路由器上实现主动队列管理(Active Queue Management, AQM),在网络设备的缓冲溢出之前就丢弃或标记报文^[9]。在众多的 AQM 算法中,RED 受到广泛的关注和研究,并被 IETF 推荐在下一代 Internet 上使用。

2.1 RED 机制

RED 的基本思想是通过监控路由器输出端口队列的平均长度来检测最初的拥塞。一旦发现拥塞逼近,就随机地选

^{*} 国家重点基础研究发展规划 973 资助项目(G1999032307);国家自然科学基金重点资助项目(60135010);国家自然科学基金资助项目(60073008);清华大学智能技术和系统国家重点实验室到开放课题基金资助项目。陶俊 博士生,研究方向:计算机网络仿真与网络行为;帅典勋 博士,教授,博士生导师,研究方向:人工智能、分布并行计算、人工生命、网络系统及网络信息利用。

择连接来通知拥塞,使它们在队列溢出导致丢包之前减小拥塞窗口,降低发送数据速度,从而缓解网络拥塞。由于 RED 是基于 FIFO 队列调度策略的,并且只是丢弃正进入路由器的数据包,因此其实施起来比较简单。

RED 算法主要分为两个部分:第一部分计算平均队列长度,以决定网关允许的流量突发的程度;第二部分计算不同拥塞程度下的丢包概率。

2.1.1 平均队长计算

RED 使用一个指数加权滑动平均(EWMA)低通滤波器来计算平均队列长度,因此突发流量带来的短时间队列长度增加不会使平均队列长度大幅增加。RED 平均队列长度使用下面公式来计算:

$$\text{avg}Q = (1 - w_q)\text{avg}Q + w_q q \quad (1)$$

q 是当前队列长度, $\text{avg}Q$ 是平均队列长度; w_q 为权值, 决定了低通滤波器的时间常数。

2.1.2 丢包概率计算

RED 网关设置两个和队列长度相关的阈值,分别是下限 min_q 和上限 max_q 。当一个数据包到达时,按式(1)计算平均队列长度 $\text{avg}Q$,并与两个阈值比较来判断是否存在早期拥塞。如果 $\text{avg}Q < \text{min}_q$,则将此数据包排队;如果 $\text{avg}Q > \text{max}_q$,所有的包均被丢弃;若 $\text{min}_q \leq \text{avg}Q \leq \text{max}_q$,则按概率 p_a 丢弃进入路由器的包。丢弃概率 p_a 由下式给出:

$$p_b = \frac{\text{avg}Q - \text{min}_q}{\text{max}_q - \text{min}_q} - \text{max}_p$$

$$p_a = \frac{p_b}{1 - \text{count} \cdot p_b} \quad (2)$$

max_p 为最大丢包概率, count 为从上次丢包开始到现在进入队列的包的数量。

2.2 ECN

在 RED 机制中,当平均队长超过一定的阈值时便开始丢包了。虽然使得 RED 有效地管理了平均队长,但也浪费了网络资源,并且对时延有一定要求的应用也不理想。因此,1994 年 Floyd 又以 RED 算法为基础,提出了显式拥塞指示(Explicit Congestion Notification, ECN)算法。ECN 在检测到网络的早期拥塞时并不丢弃数据包,而是按一定的概率在数据包中设置指示拥塞的标志位,继续转发该数据包至目的端。源端在接收到由目的端返回的携带的拥塞指示的 ACK 确认后,执行拥塞避免算法来降低发送速率。

为实现以上功能,ECN 需要在 IP 包头设置一个两位(bit)的 ECN 域:一个是 ECT(ECN-Capable Transport)位,由源端设置以显示源端节点的传输协议是支持 ECN 的;另一个是 CE(Congestion Experienced)位,由路由器设置,以显示是否发生了拥塞。

2.3 RED/ECN 的优缺点

RED/ECN 算法有以下优点:

- ①无状态算法,系统开销小,便于实现;
- ②能有效控制平均队长,吸收更多的突发数据包,以减少包的丢失率及排队延迟。
- ③避免了全局同步及对突发流偏见的现象。

但如同 TCP 的拥塞控制机制一样,RED/ECN 仍然存在着以下两个缺陷:

- ①没有提供一种激励促使用户执行拥塞控制算法。例如一些用户通过修改端系统上的 TCP,使得窗口的初始值很大并保持不变,来达到占有更多带宽的目的,却没有相应惩罚措

施来制止这种行为;另外一些应用采用 UDP 协议不能对拥塞作出反应,导致了带宽的不公平竞争。

②拥塞指示随数据包到达接受方以后,再由接收方在 ACK 包头中反馈给发送方,如 TCP 对拥塞的判断一样,存在反应滞后的问题。发送方在滞后时间内会继续以原速度发送包,而这些包的到达很可能导致拥塞的发生。

3 基于 EPN 的价格机制

为了解决上述两个缺陷,我们引入了价格机制,使资源的价格与网络的拥塞状态相关,促使用户做出正确的资源决策,并且将资源价格的调整建立在 EPN 上,使用户快速响应检测到的早期拥塞指示,使网络运行在高利用率、低延迟、低丢包率的状态。

3.1 价格机制

从经济学的角度,我们把网络用户和网络服务商(Internet Service Provider, ISP)分别看成带宽的消费者与生产者。设单位带宽的价格为 p ,用户 i 消费的带宽为 x_i ,用效用函数 $u_i(x_i)$ 来刻画带宽给消费者带来的满意程度,它有如下两条性质:(1) $u_i(x_i)$ 是单调增加的。(2) $u_i(x_i)$ 是凹的^[10]。用户作为理性的经济人,其目的是最大化自己的消费盈余,因而其对带宽的最优需求量为:

$$x_i(p) = \arg \max_x (u_i(x_i) - px_i) \quad (3)$$

由效用函数的性质可知, $x_i(p)$ 是一减函数,因而用户的总需求 $\sum_i x_i(p)$ 是价格 p 的减函数。这也反映了价格对需求的反作用。

网络的目的是最大化带宽的利用率。设网络中瓶颈带宽为 C ,则

$$\sum_i x_i(p) = C \quad (4)$$

为了求得最优价格 \bar{p} ,系统可以联立求解式(3)与式(4)。由于网络是一个复杂的分布系统,不可能知道每个用户的效用函数,而且用户的需求是动态变化的,最优价格也是动态变化,这势必给网络带来很大的额外负担。

因而,可以通过价格的自适应调整,将价格收敛于均衡价格 \bar{p} 。定义超额需求为 $z(p) = \sum_i x_i(p) - C$,价格 p 随时间按下式调节 $p \& \# = g(z(p))$,其中 $g(g)$ 是超额需求的平滑符号函数,则在一定条件下,价格 p 将收敛于均衡价格 \bar{p} ,这一过程称之为价格探索(Tatonnement)^[11]。

3.2 EPN

为了实现价格探索,提出了 EPN(Explicit Pricing Notification)算法。它以 RED/ECN 为基础,采用源端和路由器协作的方式来实现价格对用户需求的控制。该算法通过平均队长来反映用户需求的变化,当 $\text{avg}Q < \text{min}_q$ 时,认为供大于求,路由器将 CE 位置 0;当 $\text{avg}Q > \text{max}_q$ 时,认为供小于求,路由器将 CE 位置 1;当 $\text{min}_q \leq \text{avg}Q \leq \text{max}_q$ 时,认为供求基本平衡,依概率 p_a 路由器将 CE 位置 1, p_a 的计算同式(2)。路由器对 CE 位置定位后,立即向源端发送价格调整指示消息。用户在接收到价格调整指示后,计算出当前带宽价格。当 CE 位为 1 时,当前价格下调 Δp_2 个百分点;当 CE 位为 0 时,价格上调 Δp_1 个百分点。在计算出当前价格后,用户以最大化自己的消费盈余的原则来决定带宽的使用,即按式(3)来选取带宽。

3.3 仿真试验及结果

为了研究 EPN 对系统性能的影响,建立了如图 1 所示的

网络拓扑结构。其中用户数为 20,其应用为 FTP。路由器 A 与用户 u_1, u_2, \dots, u_{20} 相连接,各链路带宽为 5Mb,延时为 1ms;路由器 B 与 n_1, n_2, \dots, n_{20} 相连接,各链路带宽为 5Mb,延时为 1ms。主干链路带宽为 10Mb,延时为 2ms。设 TCP 包的长度为 1000 字节,仿真时间为 100s,路由器缓存最大包数 buf 为 1000 个。

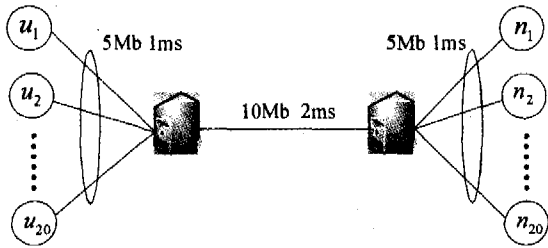


图 1 仿真网络拓扑

设用户的效用函数为 $u_i(x_i) = k_i \log x_i$, 其中 $k_i > 0$, 则 $u_i(x_i)' = k_i/x_i > 0, u_i(x_i)'' = -k_i/x_i^2 < 0$ 满足效用函数的两条性质。由式(3)知,用户 i 的最优带宽为 $x_i(p) = k_i/p$, 因而 k_i 的物理意义是用户的支付意愿。 k_i 越大,用户支出越大,其获得的带宽也越多。在仿真试验中,设 $k_i = 0.1 \times i, i = 1, 2, \dots, 20$ 。为了反映用户的动态变化,在仿真开始时,用户 1~10 加入网络,并持续到试验结束;用户 11~20 在 0~50s 内每隔 5s 依次进行连接,在 50~100s 内,又依次退出。

按照文[10],取权值 $w = 0.002, \max_p = 0.5, \min_{th}$ 与 \max_{th} 按下式取值进行相应的仿真试验:

$$\begin{cases} \min_{th} = 50 \times i, 0 \leq i < 20 \\ \max_{th} = 50 \times j, i < j \leq 20 \end{cases}$$

试验结束后,取系统的流量、延迟、丢包率等均值做图 2、3、4。

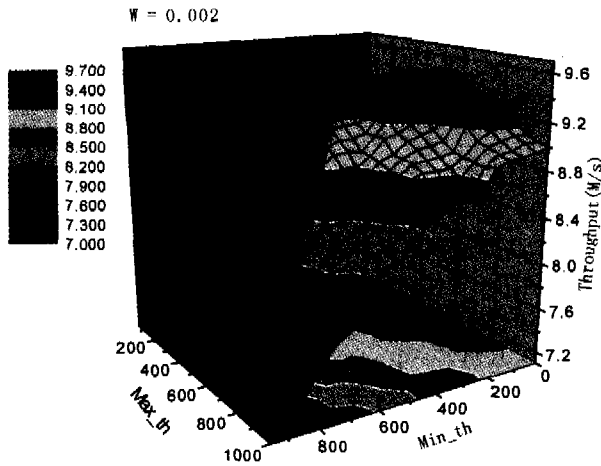


图 2 队列阈值对系统吞吐量的影响

从图 2 可以看出,当 $\min_{th}/buf = 0.1, \max_{th}/bu = 0.4$ 时,系统的平均流量达到最大值 9.45M/s,这是由系统链路利用率与系统丢包这一对矛盾相互作用的结果。如果 \min_{th} 过小,则链路利用率过低。但当 \max_{th} 过大时,它又没有足够的队列来容纳突发的数据包,导致大量丢包。

图 2、图 3 表明系统的平均延迟与平均丢包率都是随 $\min_{th}、\max_{th}$ 的增大而增大的。当 $\min_{th}/buf = 0.1, \max_{th}/bu = 0.4$ 时,系统的平均延迟为 17.3ms,平均丢包率为 0.98%,维持在一个比较理想的水平。

结论 针对目前网络带宽资源的分配所存在的问题,本

文从微观经济学的角度出发,提出了一种基于 EPN 的拥塞价格机制。其主要方法是利用 RED 平均队列的变化来反映网络供求关系的变化,路由器再根据当前的供求关系对价格进行调整,并快速将价格指示消息反馈到用户端。用户根据价格指示消息计算出当前带宽价格,并依据自己的效用函数做出自己的资源决策。利用这种自适应的机制,使系统快速收敛于供求平衡的均衡状态,使网络与用户的资产最大化。通过网络仿真试验,发现对 EPN 影响最大的是队列阈值的设定。当对 \min_{th} 与 \max_{th} 合理设置后,系统的吞吐量、延迟、丢包率将维持在一个理想的水平。

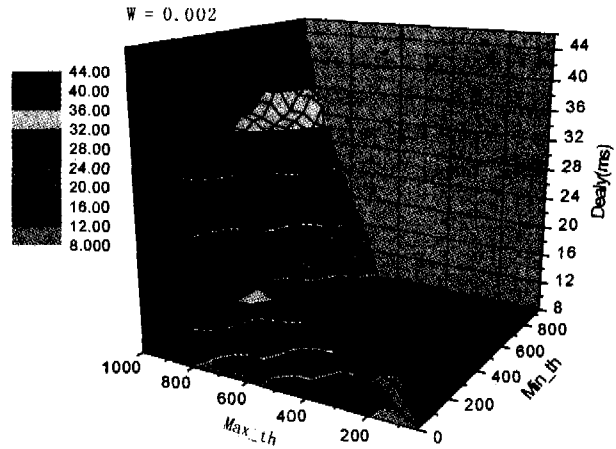


图 3 队列阈值对系统延迟的影响

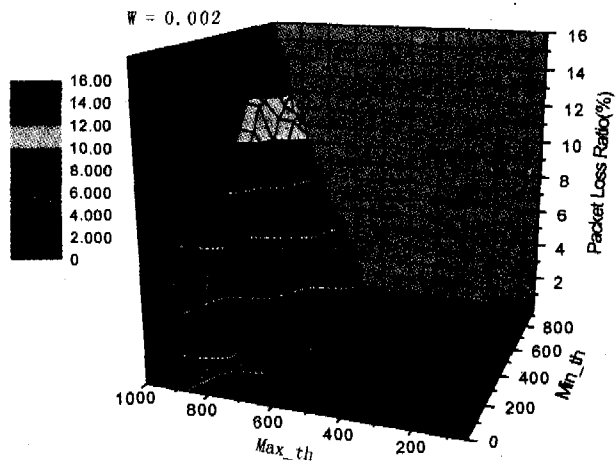


图 4 队列阈值对系统丢包率的影响

参考文献

- 1 章森,吴建平,林闯.互联网端到端拥塞控制研究综述.软件学报,2001,13(3):354~363
- 2 McKnight Lee W, Boroumand J. Pricing Internet Services: Approaches and Challenges. Internet Watch, 2000, 33(2):128~129
- 3 Varina H R, Mackie-Mason J K. Pricing the Internet. Public Access to the Internet, 1994
- 4 Clark D D. A model for cost allocation and pricing in the Internet. [Technical report]. MIT, August 1995
- 5 Gibbens R J, Kelly F P. Resource pricing and the evolution of congestion control. Automatica, 1999, 35:1969~1985
- 6 Floyd S, Jacobson V. Random early detection gateways for congestion avoidance. ACM/IEEE Transactions on Networking, August 1993
- 7 Floyd S, Ramakrishnan K. A proposal to add explicit congestion notification (ECN) to IP. RFC 2481, 1999
- 8 任丰原,林闯,刘卫东. IP 网络中的拥塞控制. 计算机学报, 2003, 26(9), 1025~1034
- 9 Braden B, Clark D, Crowcroft J. Recommendations on queue management and congestion avoidance in the internet. RFC 2309, 1994
- 10 周惠中. 微观经济学. 上海:上海人民出版社,1997
- 11 陈及. 西方经济学. 北京:中国财政经济出版社,1996