

动态权重调整 RED

江 昊, 晏蒲柳, 吴 静, 周建国

(武汉大学电子信息学院, 湖北武汉 430079)

摘 要: IETF 推荐 RED(Random Early Detection) 作为下一代 Internet 路由器的拥塞控制算法. 但 RED 中平均队列长度对实际队列变化反应较慢, 导致出现拥塞的时间增加, 丢包增多, 本文在现有 RED 算法的基础上通过改进 RED 算法的平均队列长度计算, 使用平均队列长度的均值, 判断拥塞的发生和结束, 动态调整权重, 提高 RED 对拥塞发生和拥塞结束的反应速度. 仿真结果表明算法减少了拥塞时间, 能够维持较低的实际队列长度, 进而减少了数据包排队时间.

关键词: RED; 平均队列长度; 拥塞控制

中图分类号: TP393

文献标识码: A

文章编号: 0372-2112 (2005) 03-0574-04

A Dynamic Weight Tune RED Gateway

JIANG Hao, YAN Pu liu, WU Jing, ZHOU Jian guo

(School of Electronic Information, Wuhan University, Wuhan, Hubei 430079, China)

Abstract: The IETF has addressed the congestion control by advocating the deployment of active queue management mechanisms, such as RED, in the network. But RED has some problems, such as the average queue length changes slowly when congestion starts and ends. To alleviate this problem, we propose a dynamic weight tune RED which can quickly be conscious of the congestion and dynamically tune the weight. Dynamic Weight RED can alleviate the congestion and drop when the congestion is over. The time of congestion is shorter than the RED.

Key words: RED; dynamic weight; congestion control;

1 引言

拥塞控制算法对保证 Internet 的稳定以及实现 QoS 具有十分重要的作用. 最初的 TCP 协议中只有流控制(Flow Control)而没有拥塞控制. 1986 年 10 月, 由于拥塞崩溃(Congestion Collapse)的发生, 美国 IBL 到 UC Berkeley 的数据吞吐量从 32Kbps 跌落到 40bps^[1]. 在那之后, 拥塞控制领域开展了大量的研究工作.

根据算法的实现位置, 可以将拥塞控制算法分为两大类: 链路算法(link algorithm)和源算法(source algorithm). 源算法中使用最广泛的是 TCP 协议中的拥塞控制算法, 但是这些改进算法并没有使链路的拥塞有太大改观, 因此 IETF 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^[2]. 在众多 AQM 算法中 RED(Random Early Detection)^[3, 4] 受到了广泛的关注和研究, 并被 IETF 推荐在下一代 Internet 上使用.

目前国际上对 RED 的研究, 主要集中在 RED 算法的动态调整标记/丢弃的概率方面, 提出了 ARED、SRED、BLUE 等算法. 但在 RED 的基础部分平均队列的计算部分, 目前只有

LPF/ODA^[5] 算法. 最新的研究^[6, 7]表明, 由于 RED 中队列长度与丢失反馈紧密接合, 导致了负荷与队列水平相关, 从而产生了更高的排队延迟.

本文从改进平均队列的权重的角度, 提出了新的平均队列的权重动态调整算法, 很好的解决了 RED 的平均队列对实际队列响应慢的问题, 同时又保证了对突发流量的容忍和算法的简洁, 能有效的减少队列处于拥塞状态的时间. 该算法的主要思想是, 判断平均队列的变化趋势, 并根据这个趋势动态的调整权重 w , 使得平均队列在拥塞发生时更接近实际队列, 进而在初期就发现拥塞, 同时在长时间的拥塞结束后, 平均队列能跟随实际队列快速下降, 避免在非拥塞时期的无用丢包. 该方法在提高平均队列速度以及减少拥塞时间等方面有较好的效果.

2 动态权重调整 RED-DWRED

2.1 RED 的动态调整权重

在 RED 中平均队列长度的计算是整个 RED 的基础, 其中在平均队列的计算中低通滤波器的权重 w 的设置最为关键, 它决定了低通滤波器的时间常数^[4]. 需要强调的是 DWRED

是以包数为单位表示队列长度。

RED 使用一个低通滤波器来计算平均队列长度, 因此由突发流量产生的短时间队列长度的增加不会导致平均队列长度剧烈的增加。RED 平均队列长度使用(EWMA) 来计算:

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

低通滤波器过滤掉了当前实际队列长度中部分信息, 相当于高频部分, 而保留了历史的平均队列长度, 相当于低频部分。在这中间权重 w 用以控制区分高低频的门限, 当 w 较小时, 平均队列长度中包含当前实际队列长度的信息就少, 必然导致对当前实际队列长度变化的不敏感对拥塞反应慢, 如果 w 较大, 平均队列长度就能更多的反应当前实际队列长度的变化, 但又带来了抖动的问题。

RED 之所以使用这个低通滤波器计算平均队列, 就是防止实际队列长度的剧烈变化, 而产生的抖动。 w 由 RED 的最小阈值 \min_{th} 与路由器允许的突发流量 L 决定的, 其必须满足下面等式:

$$L + 1 + \frac{(1 - w)^{L+1} - 1}{w} < \min_{th} \quad (2)$$

w 的大小影响着对突发流量的处理, RED 的一大优点就是能容忍短时间的突发流量和拥塞, 容忍程度由 w 决定。 w 越小, 能容忍的突发流量就越大, 但是在 RED 的设计中存在问题, 就是如果 w 设置太小, 平均队列长度对实际队列长度变化的反映就很慢, 在这种情况下, RED 就不能在拥塞发生的初期就发现拥塞, 这也就达不到 RED 将总体队列长度维持在一个较低水平的目的, 同样也不能使平局队列长度在拥塞结束后快速下降进而减少无为的丢包。动态的调整权重是对该问题一个很好的解决方法, 但主要的问题是在何时调整 and 如何调整。

2.2 平均队列长度变化趋势的判断

因此 RED 计算的平均队列在一段时间上的变化是平稳的。 RED 是在每个包到达时计算平均队列, 相邻两个计算得到的平均队列长度在时间上是很接近的, 而且从平均队列的计算公式可发现实际队列长度在平均队列长度中所占比例是很小的。

因此直接使用平均队列长度判断队列变化情况是比较困难的, 我们应用均值法结合变化幅度判断 RED 中平均队列长度的变化趋势。方法如下:

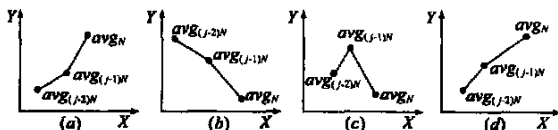
(1) 统计 N 个平均队列长度 avg , 计算 N 个 avg 的平均值

$$\text{avg}_N = \frac{\sum_{i=1}^N \text{avg}_i}{N}$$

(2) $D_j = \text{avg}_N - \text{avg}_{(j-1)N}$, $D_{j-1} = \text{avg}_{(j-1)N} - \text{avg}_{(j-2)N}$,

avg_N 为第 j 个 avg_N 。比较 D_j, D_{j-1} , 大致会出现四种情况。

- (I) $D_j > 0, D_{j-1} > 0$ 且 $D_j > D_{j-1}$
- (II) $D_j < 0, D_{j-1} < 0$ 且 $|D_j| > |D_{j-1}|$
- (III) $D_j < 0, D_{j-1} > 0$ 或 $D_j > 0, D_{j-1} < 0$
- (IV) $D_j/D_{j-1} > 0$, 但 $|D_j| < |D_{j-1}|$, 而 $|D_j| > m$, m 是



$|D_j|$ 变化的门限值

在以上 4 种情况中, (a) 和 (b) 平均队列的变化趋势分别是上升和下降, (c) 平均队列的变化趋势是有波折。因此判断在情况 (a) 和 (b) 发生时, 是拥塞过程初期和拥塞结束过程的初期情况, 所以需要调整权重以跟随实际队列的变化。 (d) 是在拥塞发生后或结束前平均队列长度在平稳大幅度上升期间可能出现的一种情况, 虽然 $|D_j| < |D_{j-1}|$, 但 $|D_j|$ 比较大, 超过一定门限值, 即 D_j 的绝对值 $|D_j|$ 大于 m , 而 m 的估计要考虑每次平均队列的统计数 N 。情况 (a) 和 (b) 针对的是在拥塞发生初期和拥塞结束初期平均队列长度急剧变化, 情况 (d) 针对的是在拥塞发生后或结束前平均队列长度平稳变化, 但变化幅度很大。

这里首先需要确定的参数是每次平均队列的统计数 N 。 N 要考虑平均队列变化的周期, 而平均队列长度的变化是由 TCP 发送速率的变化引起的。 TCP 改变发送速率 (无论增加还是减小) 都是在收到 ACK 后, 即一个 RTT 后。

在这里我们将进入 RED 网关的流量抽象为多个稳定的 TCP 流。因此 N 在时间轴上要在一个平均 RTT 内。我们得到公式

$$N \leq \text{RTT} \cdot S$$

S 是进入数据包 RED 网关的平均速率, 在这里近似等于平均吞吐量, 极限情况为链路的带宽。

m 的估计是基于 N 的, 通常拥塞发生时, 实际队列长度是大于平均队列长度的, 且实际队列在短时间内是急剧上升的。 $|D_j|$ 变化的极限情况是在到达 N 个包的区间内, 每到达一个包实际队列长度就加 1。

根据平均队列的计算公式, 以及拥塞发生初期平均队列长度与实际队列长度在 N 个包内变化的趋势。

$$\text{avg}_{N-1} = (1 - w)^{N-1} (\text{avg}_0 - wq_0) + \sum_{i=0}^{N-1} (1 - w)^{N-1-i} wq_i \quad (3)$$

因为 $w = 0.002$, $(1 - w)^i \approx 1.0$, 所以 $\text{avg}_{N-1} \approx \text{avg}_0 - wq_0 + \sum_{i=0}^{N-1} wq_i$, $\text{avg}_{N-1} - \text{avg}_0 \approx w \sum_{i=1}^{N-1} q_i$, 因为是拥塞发生时, q_i 在理论上是大于 q_0 , 且小于 $q_0 + N - 1$, 假设 q_i 在区间 $[q_0, q_N - 1]$ 上等概率分布, 我们可以假定 $\text{avg}_{N-1} - \text{avg}_0 \approx w(2q_0 + N)N/2$ 。为了适应一般情况, 我们将判断拥塞时的 m 定为 0.3, 即 $D_j/D_{j-1} > 0$, 但 $|D_j| < |D_{j-1}|$ 时, 如果 $|D_j| > 0.3$ 可认为拥塞发生。同理估算拥塞结束过程中, 由于可能在 N 个包到达的时间内, 实际队列减少的值会大于 N , $\text{avg}_{N-1} - \text{avg}_0 > 0.6$, 情况更为复杂, 为了统一计算 m 也为 0.3。

2.3 权重的调整

在判断了平均队列变化的趋势后, 就要调整权重 w 的大小。根据权重公式, 以及允许的确突发流量, 可以计算权重的最大值。参考文献[4] 在这里我们给出最小权重 $w_{\min} = 0.002$, 在仿真中也都以此为权重下限。

权重的调节要根据 D_j 和 D_{j-1} 的变化以及 $|D_j|$ 的幅度, 来自适应调节, 在这里我们借鉴了自适应滤波的思想, 情况 (i) 和情况 (ii) 时根据 D_j 和 D_{j-1} 比值以及最小权重 w_{\min} 和最大权重 w_{\max} 来调节权重, 在情况 (IV) 时根据 $|D_j|$ 的幅度来调节权重。 w_{\max} 由 RED 的最小阈值 \min_{th} 与路由器允许的突发流量 L

决定的.

定义 1 如果 $D_j/D_{j-1} > 1$, 且 $D_j/D_{j-1} < w_{\max}/w_{\min}$

$w_c = w_{\min} \times D_j/D_{j-1}$

如果 $D_j/D_{j-1} > 1$, 且 $D_j/D_{j-1} \geq w_{\max}/w_{\min}$

$w_c = w_{\max}$

如果 $0 < D_j/D_{j-1} < 1$, 且 $|D_j| \geq m$

$w_c = w_{\max}$

如果 $D_j/D_{j-1} < 1$, 且 $|D_j| < m$

$w_c = w_{\min}$

3 仿真实验

在这一部分我们将用仿真的方法表明, 动态调节权重的方法能很好的跟随实际队列的变化, 在拥塞发生初期就对其进行控制, 在长期拥塞结束后能减少不必要的丢包, 并能有效的减少拥塞时间.

图 1 是仿真实验拓扑:

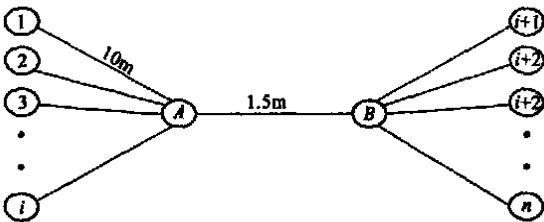


图 1 仿真实验拓扑

节点 A 和节点 B 是路由器, 节点 A 与节点 B 之间的链路 LAB 是瓶颈链路带宽 1.5m, LAB 两端都配置的是 DWRED, DWRED 的参数设置为: 平均队列最小值 \min_{th} 是 5, 最大值 \max_{th} 是 15, 队列最大容量是 50, 最小权重 w 是 0.002, 根据权重公式最大权重 w 是 0.0042. 为了更好地验证 DWRED 的优点, 本文使用了一种与网络上的流量比较接近的流量模型^[8]. 该模型具有如下特点: (1) 既有长生存期的 FTP 流量, 也有短生存期的 Web 流量, 以 Web 为主, 约占总吞吐量的 90%; (2) 考虑到了反向路径的流量; (3) 往返时间的值是一定值域内的随机数. 节点 A 和节点 B 的两端各配有 20 个 FTP 源, 300 个 HTTP 源. 这个实验模型在许多文献中, 例如文献^[7, 9], 得到广泛的使用, 是一个比较好的仿真模型.

由于 DWRED 能有效地在拥塞发生初期就检测到拥塞并作出反应, 而且在拥塞结束后能发现实际队列长度的下降, 并调整平均队列从而减少不必要的丢包重传, 所以在整个过程中拥塞时间要比采用传统的 RED 情况下拥塞的时间要短. 在这里我们定义实际队列长度超过 \max_{th} 的情况为严重拥塞, 介于 \max_{th} 和 \min_{th} 之间的情况为可控制拥塞, 小于 \min_{th} 的情况为无拥塞. 在 100s 的实验中, 比较 DWRED 与 RED 对拥塞的控制能力. 表 1 是结果.

表 1 DWRED 减少整体拥塞时间

	严重拥塞/总时间 (%)	可控制拥塞/总时间 (%)	无拥塞/总时间 (%)
DWRED	39.2	40.2	20.6
RED	44.5	35.8	19.7

由上表可以看出 DWRED 在控制拥塞方面比 RED 有一定优势, 它的严重拥塞时间要明显比 RED 少, 无拥塞时间要比 RED 多, 可控制拥塞时间也比 RED 多, 这说明 DWRED 控制拥塞的能力较 RED 强.

图 2(a)、图 2(b) 是 DWRED、RED 的实际队列长度.

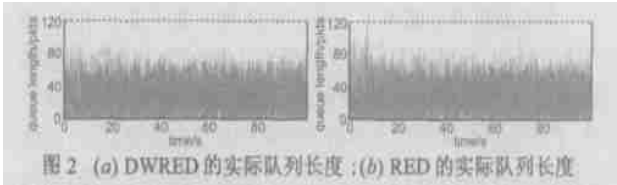


图 2 (a) DWRED 的实际队列长度; (b) RED 的实际队列长度

由图知, DWRED 表 2 DWRED 和 RED 实际队列长度比较的实际队列长度总体上小于 RED, 表 2 就是实际队列长度的统计数据.

	实际队列长度			
几率 (%)	< 15	< 30	> 50	> 60
DWRED	24.1	53.8	11.2	3.2
RED	22.2	52.8	11.2	3.6

由表 2 可以看出 DWRED 处于较小实际队列长度的几率要比 RED 大.

图 3(a)、图 3(b) 是 DWRED、RED 瓶颈链路的链路延迟:

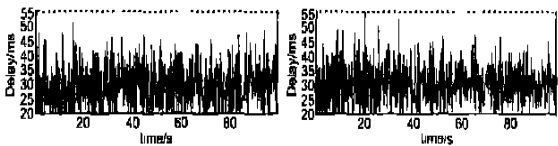


图 3 (a) DWRED 延迟; (b) RED 延迟

由表 3 可以看出 DWRED 的平均延迟和丢包率要小于 RED, 但代价是

表 3 DWRED 和 RED 性能比较

	延迟(ms)	吞吐量(字节/带宽)(%)
DWRED	30.569	97.64
RED	31.194	97.76

吞吐量的降低. 从以上的实验结果可知, DWRED 可以用在拥塞比较严重的网路上, 它可以有效地减少队列处于“严重拥塞”的时间.

4 总结

本文提出了一个新的 RED 拥塞控制改进算法, 其主要方法就是利用 RED 平均队列的变化趋势来动态调整 RED 的低通滤波器的权重. 使得在拥塞发生初期路由器就能检测到拥塞并作出反应, 同时在拥塞结束后能使平均队列长度更贴近实际队列, 从而减少不必要的丢包.

通过仿真实验证明, DWRED 能有效地减少队列拥塞时间, 而且 DWRED 比 RED 在拥塞发生时和拥塞结束后在延迟、丢包率等几个指标上有更好的性能. 同时也保留了 RED 的优点, 如算法简单、允许突发流量等.

下一阶段将重点从控制论的角度对 DWRED 做出理论分析, 并尝试克服 RED 的队列在流数目增加的情况下, 平均队列长度会逐渐增加的缺点.

参考文献:

[1] JACOBSON V. Congestion avoidance and control[J]. ACM Computer Communication Review, 1988, 18(4): 314- 329.
[2] BRADEN B, et al. IETF RFC 2309, Recommendations on Queue Management and Congestion Avoidance in the Internet[S]. April 1998.

- [3] FLOYD S, JACOBSON V. On traffic phase effects in packet switched gateways[J]. Internetworking: Research and Experience, 1992, 3 (3) : 115- 156.
- [4] FLOYD S, JACOBSON V. Random early detection gateways for congestion avoidance [J]. IEEE/ ACM Transaction on Networking, 1993, 1 (4) : 397- 413.
- [5] BING ZHENG, MOHAMMED ATQUZZAMAN. Low pass filter/ over drop avoidance(LPF/ ODA): An algorithm to improve the performance of RED gateways[R]. CS-TR 01-001, technical report, <http://www.cs.ou.edu/~atq/papers/>; February 2001. 3- 4.
- [6] HOLLOT C, et al. A control theoretic analysis of RED[A]. Proceedings of INFOCOM 2001[C]. USA: IEEE, 2001. 1510- 1519.
- [7] HOLLOT C, et al. On designing improved controllers for AQM routers supporting TCP flows[A]. Proceedings of INFOCOM 2001[C]. USA: IEEE, 2001. 1726- 1734.
- [8] FLOYD S, EDDIE K. Internet research needs better models[A]. Proceedings of HotNet II[C]. USA: Princeton University, October 2002. 300 - 312.

- [9] REN FENGYUAN, LIN CHUANG. A robust active queue management algorithm based on sliding mode variable structure control[A]. Proceedings of INFOCOM 2002[C]. USA: IEEE, 2002. 64- 79.

作者简介:

江 昊 男, 1976 年生, 2001 年在武汉大学获硕士学位, 目前博士在读, 研究兴趣为计算机网络拥塞控制性能分析优化. E-mail: jianghaow@ 263. net



晏蒲柳 女, 1962 年生, 武汉大学教授, 博士生导师, 研究方向为高速信息网络.

电子学报

2005 年第 3 期 Acta Electronica Sinica No. 3 2005

(总期 257 期) (Monthly) (Series No. 257)

主管单位 中国科学技术协会
主办单位 中国电子学会
协办单位 南京才华科技有限公司
编辑 《电子学报》编辑委员会
主编 王 守 觉
总编辑 刘 力
通信处 北 京 1 6 5 信 箱
(邮政编码 100036)
电 话 (010) 68279116, 68285082
传 真 (010) 68173796

China Association for Science and Technology
Published by the Chinese Institute of Electronics, Beijing
Nanjing Caihua Technology Co., Ltd.
Edited by Editorial Board of Acta Electronica Sinica
Chief Editor: WANG Shou jue
Director: LIU Li
Add: Editorial Office of Acta Electronica Sinica
(POBox 165, Beijing 100036, China)
Tel: 86-10-68279116, 68285082
Fax: 86-10-68173796

Home page: <http://www.elecjournal.org>; <http://dzxu.chinajournal.net.cn>

Email: cje@elecjournal.org; dzxu@chinajournal.net.cn

排版印刷 北京育兴达印刷厂
国内总发行 北京市报刊发行局

Printed by Yuxingda, Beijing, China
Distributed by

国外总发行 中国国际图书贸易总公司
国内订购处 全 国 各 邮 电 局

Domestic: Beijing Baokan Faxingju, China
Foreign: China International Book Trading Corporation
Subscription Office — All Local Post Offices in China

国内统一刊号: CN11- 2087/TN

邮发代号(国内/ 国外) : 2- 891/M436

国内定价 ¥ 32. 00