

分级指导随机提早丢弃队列管理机制 CGRED

宋光农,赵永祥,邢新宇,陈常嘉

(北方交通大学电信学院通信系,北京 100044)

摘 要: 本文提出的分级指导随机提早丢弃 CGRED 是在因特网上为用户提供有区别服务的单队列管理机制,它同时满足了对服务质量的精细控制和实现简单性两方面的需求. CGRED 采用在线测量的方法跟踪业务类的实际带宽占用情况,根据给定的优先级关系确定调整业务类的指导策略,并以采用自适应算法计算得到的丢弃指导概率作为分组丢弃概率的边界,成功地解决了将不同业务类的实际带宽占用情况转化成如何丢弃不同业务类分组的实际操作. 仿真结果表明 CGRED 是一种能有效提供 DiffServ 的实用网络控制机制.

关键词: 有区别服务; 链路共享; 带宽维护; 队列管理

中图分类号: TN91 **文献标识码:** A **文章编号:** 0372-2112 (2002) 09-1409-04

CGRED: Class Guided Random Early Discarding

SONG Guang-nong, ZHAO Yong-xiang, XING Xin-yu, CHEN Chang-jia

(Dept. of Electronics & Communication, Northern JiaoTong Univ., Beijing 100044, China)

Abstract: Class Guided Random Early Discarding (CGRED), a novel scheme to provide services differentiation in today's Internet, is proposed. It can precisely maintain the assigned priority and pre-allocated bandwidth for each service class. As a single queue management mechanism, CGRED possesses all advantages of the simplicity, robustness and scalability of schemes like RIO and WRED. And it has the capability to maintain the pre-allocated bandwidth for each class of service precisely just as CBQ. The key enforcement of CGRED to standard RED is to discard packets of different service classes adaptively with different probabilities according to the bandwidth actually occupied by different classes and the pre-allocated bandwidth. To accomplish this, a couple of class-specific guidelines, the 'adjustment direction' and 'guiding probability', are introduced to conduct the adaptive determination of class-differential discarding probabilities for each class of service. The guidelines for each class of service are updated periodically based on online measured bandwidth utilization. When discarding packets, CGRED sets the class-specific guidelines as the boundary of the packet dropping probability of standard RED. In this way, the unbalance of bandwidth consumption among different service classes is successfully transferred to differential operations in the buffer management among different service classes. Extensive simulation work has been conducted to verify the performance of CGRED and robustness of parameters in its algorithm. The simulation results show that CGRED is an effective scheme to provide Differential Services in networks with pre-allocated bandwidth and pre-assigned priority hierarchy among service classes.

Key words: DiffServ; link sharing; bandwidth maintaining; queue management

1 引言

因特网采用普遍公平和尽力而为(best effort)的网络服务机制,对网络用户的传输带宽和时延不做任何承诺. 这种服务方式保证了网络业务的普及性,同时也阻碍了对网络服务质量要求较高的新技术和新业务的开展,因此有必要为网络用户提供一定程度的服务质量保证.

目前,提供网络服务质量保证的两个基本框架是综合服务(IntServ)和有区别服务(DiffServ). IntServ 需要维护大量的连接信息,可扩展性差;简单易行的 DiffServ 只对业务类提供服务质量的保证,网络只需按照分组的业务类标记进行优先级处理. 目前对业务类的优先级处理方案有:维护业务类预分配

带宽和不维护业务类预分配带宽两类.

在维护业务类预分配带宽方案中,路由器为每个业务类预分配一个带宽值,路由器管理的目标是保证每类业务的汇聚流能够得到为其预先分配的带宽. 业务质量的差别表现在不同业务类的汇聚流的特征上,而不是体现在不同业务类的单个流的特征上. 文献[1]在链路共享的概念下讨论了建立业务类间等级关系的一般性方法和协调准则, CBQ 调度策略^[3-5]将文献[1]的思想开发成一个实用的控制机制. 因此,可以说 CBQ 是提供有区别服务的优先级处理机制中维护业务类预分配带宽的典型代表^[6].

在不维护业务类预分配带宽的方案中,路由器试图在一

些基础性能方面保证业务类汇聚流间的区别;目前的工作主要集中在分组丢弃率方面,如 RIO^[7]与 WRED^[8]通过改造 RED^[9]单队列管理机制,使不同等级的业务类得到不同的分组丢失质量.其它一些基于调度的工作,如 MDP 和 EDF,由于调度方法本质上的复杂性,并没有成为研究的主流.

CBQ 虽然在理论上较为完美,但基于如下两个因素,我们认为 CBQ 不会有很好的应用前景:第一, CBQ 是基于调度的管理策略,其实现复杂性远大于队列管理机制的实现复杂性,很难适应网络规模的增长速率;第二,虽然 CBQ 在处理不实施端到端拥塞控制的媒体流方面有独到的优势,但从网络业务的发展趋势看,网络不会允许这类流发展成规模,并且大量如何惩罚这类流的研究也在积极进行之中,因此 CBQ 的这个优势未必能得到发挥.

RIO^[7]和 WRED^[8]机制实现简单,并且只适用于以端到端拥塞控制机制为主导的网络(如 Internet).但是,由于不具备对业务类进行带宽分配和维护的能力,RIO 和 WRED 不能对不同业务类的质量区分进行较为精细的控制,存在高等级用户过分抢占低等级用户带宽的问题^[10,11].

基于上述因素,我们认为:如何在 Internet 上提供有区别服务的问题并没有得到解决,并且 Internet 迫切需要一种在实现复杂性方面与 RIO 和 WRED 相当、在业务类的质量控制方面类似于 CBQ 的新机制.为此,本文提出了一种提供有区别服务的新机制—分级指导随机提早丢弃队列管理机制(CGRED).

基于单队列管理的 CGRED 由三个主要部分组成,其中二、三部分是本文的主要贡献.第一部分沿用文献[1]中链路共享的思想,根据规定的业务类等级关系和实际业务类的带宽占用情况推断出指导每个业务类进行调整的丢弃策略;第二部分根据丢弃策略、业务类的实际带宽占用情况,采用自适应算法动态地调整每个业务类的指导丢弃概率;第三部分根据每个业务类的丢弃策略、指导丢弃概率和系统队列的实际占用情况,实施分组丢弃.在实施分组丢弃时以标准 RED 的分组丢弃概率作为参考、指导丢弃概率作为分组丢弃概率的限制边界.仿真结果表明,CGRED 具有类似 CBQ 机制的带宽维护能力,并且实现简单.因此,CGRED 是提供有区别服务的一种有效的网络控制和管理机制.

本文内容安排如下:第二节给出 CGRED 实现方案的详细描述;第三节介绍我们的仿真工作;第四节是本文总结.

2 分级指导随机提早丢弃队列管理机制 CGRED

本文假定需要对 N 个业务类提供有区别的服务,且编号越小的业务类 $s \in (0, \dots, N-1)$ 的优先等级越低.与 CBQ 类似,CGRED 通过对每个业务类实际占用带宽的定期测量来了解各业务类 S 的资源占用情况,在比较带宽预分配方案和实际资源占用情况后做出对不同优先等级的业务类进行调整的决断.CGRED 与 CBQ 的区别在于实施调整时前者采用单队列管理方式而后者采用分组调度的方式.

2.1 CGRED 的工作原理

假定各业务类 s 的分配带宽为 $\{W_s, 0 \leq s < N\}$.CGRED 根据到达分组的业务类标识区分分组的业务类型,分组的业务

类标识不在本文讨论范围之内(可用 IP 分组头中的 TOS 字段承载该标识).在每个测量时间间隔内,每到达业务类 s 的一个分组 P_s ,CGRED 按 $\hat{w}_s = \hat{w}_s + l(P_s)$ 更新业务类 s 的瞬时带宽 \hat{w}_s ,其中 $l(P_s)$ 为分组的数据长度.为减小测量值的波动,CGRED 采用低通过滤的方法求业务类 s 的占用带宽估计值 \hat{W}_s ,即 $\hat{W}_s = \alpha \hat{W} + (1 - \alpha) \frac{\hat{w}_s}{\tau}$,其中 α 为低通系数.

CGRED 为每个业务类 s 设置一个丢弃策略指示 g_s 和一个指导丢弃概率 $q_g^{(s)}$. g_s 有三个可能的取值:抢占调整、退让调整或无须调整,而 $q_g^{(s)}$ 是 0 到 1 之间的一个实数.根据各业务类的占用带宽估计值 $\{\hat{W}_s, 0 \leq s \leq N\}$ 和分配带宽值 $\{W_s, 0 \leq s \leq N\}$,CGRED 按下列原则计算各业务类的 $\{g_s, 0 \leq s < N\}$ 和 $\{q_g^{(s)}, 0 \leq s \leq N\}$:

- 对每个业务类 $s: 0 \leq s \leq N$,
- 如果 $(\hat{W}_s > W_s)$ 且 $(\sum_{j>s} \hat{W}_j < \sum_{j>s} W_j)$ 则 $g_s =$ 退让调整;
- $q_g^{(s)} = q_g^{(s)} - \max p_s^b (1 - \frac{\hat{W}_s}{W_s}) + \max p_s^r (1 - \frac{\sum_{j>s} \hat{W}_j}{\sum_{j>s} W_j})$;
- 如果 $(\hat{W}_s < W_s)$ 且 $(\sum_{j<s} \hat{W}_j < \sum_{j<s} W_j)$ 则 $g_s =$ 抢占调整;
- $q_g^{(s)} = q_g^{(s)} - \max p_s^b (1 - \frac{\hat{W}_s}{W_s}) + \max p_s^r (1 - \frac{\sum_{j<s} \hat{W}_j}{\sum_{j<s} W_j})$;
- 否则: $g_s =$ 无须调整;

其中更新系数 $\{\max p_s^b, 0 \leq s \leq N\}$ 和 $\{\max p_s^r, 0 \leq s \leq N\}$ 的选择与当前各业务类占用带宽和分配带宽有关.

对每个到达分组 P_s ,CGRED 根据业务类 s 的当前 $g_s, q_g^{(s)}$ 和标准 RED 的丢弃概率 q_R ,采用如下算法确定分组 P_s 的实际丢弃概率 q_P ,并更新 $q_g^{(s)}$:

- 如果 $g_s =$ 抢占调整,则 $q_P = \min\{q_R, q_g^{(s)}\}$;
- 如果 $g_s =$ 退让调整,则 $q_P = \max\{q_R, q_g^{(s)}\}$;
- 如果 $g_s =$ 无须调整,则 $q_P = q_R$;
- 令 $q_g^{(s)} = q_P$.

CGRED 得到 P_s 的实际丢弃概率 q_P 后,产生一个在 0,1 间均匀分布的随机数 x .如果 $x < q_P$,则丢弃该分组,否则接纳该分组进入缓存单元.

2.2 CGRED 的工作模式

从上述 CGRED 的工作原理可以看出,系统有两个工作模式:当所有业务类的丢弃策略指示均为无须调整时,系统工作在 RED 模式;否则,系统工作在指导模式.

以 2 个业务等级的系统为例,如果高业务等级的占用带宽大于其分配带宽,且低业务等级的占用带宽小于其分配带宽,则系统处于 RED 模式.此时,(1)若高业务等级的单流速率高于低业务等级的单流速率,则在 RED 的控制下(当系统中的流均使用端到端拥塞控制并有大致相同的端到端 RTT 时,RED 的控制倾向是将所有流的速率拉向一致),系统向减小高业务等级的总占用带宽和增大低业务等级的总占用带宽的方向调整.一旦当高业务等级的占用带宽小于其分配带宽,或低业务等级的占用带宽大于其分配带宽,则系统进入指导模式,指导机制将使系统向增大高业务等级的占用带宽和减小低业

务等级的占用带宽的反方向重新调整.否则,直至所有流的速率被调整到一致,系统仍处于 RED 模式.(2)若高业务等级的单流速率低于低业务等级的单流速率,则在 RED 的控制下,系统向增大高业务等级的总占用带宽、减小低业务等级的总占用带宽的方向调整,若没有外部负荷的变化,直到所有流的速率趋于一致,系统的 RED 模式不会变化.CGRED 就是这样在保证各业务类实际使用带宽稳定在分配带宽值的同时,保证了高业务等级单个用户的使用带宽不会明显低于低业务等级单个用户的使用带宽(由于 RED 算法本身会带来吞吐量的波动除外).

3 对 CGRED 的仿真工作

3.1 仿真环境

本文中的仿真方案都用 ns^[2]实现.仿真拓扑如图 1 所示,20 个源端用户(S_0, S_1, \dots, S_{19})通过一条带宽为 2Mbps、时延为 50ms 的共享链路分别连接到 20 个终端用户(D_0, D_1, \dots, D_{19}).用户接入链路的带宽为 10Mbps,时延为 0.001ms.仿真中考虑了三个业务类:用 G, M 和 B 分别代表最高、中间和最低优先等级的业务类.所有的端到端流采用 TCP 协议,并以 FTP 作为其上层应用.

验证 CGRED 性能的仿真工作分为三部分:①基础实验是在一组典型算法参数设定下检验 CGRED 是否能够达到维护业务类分配带宽的目标;②算法参数稳定性实验则是检验参数变化对系统性能的影响;③检验瓶颈链路带宽以及带宽分配比例变化对 CGRED 性能的影响.

表 1 参数稳定性实验中的数据

	基础实验	参数稳定性实验	
		下限值	上限值
瓶颈链路带宽(Mbps)	2	2	
测量间隔时间 τ	1s	-	2s
自适应算法更新参数:			
$Maxp_i^b = Maxp_i^r$	0.01	0.02	0.02
低通过滤参数 α	0.9	0.5	0.99
缓存长度 $buflen$	100	50	200
RED 参数:			
L_{min}	0	-	$0.2 buflen$
L_{max}	$0.5 buflen$	-	$0.8 buflen$
$Maxp$ (斜率倒数)	20	10	30
总用户数	0.05	-	0.02

3.2 基础实验

基础实验中的参数设定见表 1.以每个业务类至少有一个流为约束条件,我们在仿真开始 50 秒后记录数据,仿真持续时间为 500 秒.记录的数据是每个业务类 $s, \{B, M, G\}$ 的流个数 N_s 和每个流的吞吐量 $R_s^{(i)}, i = 1, \dots, N_s$,并根据这些记录值计算业务类 s 的

$$\text{总占用带宽: } T_s = \sum_i R_s^{(i)};$$

$$\text{单个流的平均速率: } R_s = T_s / N_s;$$

$$\text{流速率的标准差: } \sigma_s = \frac{\sqrt{\sum_i (R_s^{(i)} - R_s)^2 / N_s}}{R_s}.$$

仿真结果如图 2 所示,3 个子图分别表示 CGRED 的带宽维护

能力、单流速率和流速率的标准差;每个子图中的 3 条曲线分别对应 G, M 和 B 业务类的特征随各业务类的流个数变化的规律.需要注意的是:观察不同业务类的特征曲线时,横轴代表的是相应业务类的流的数目,如考察 G 业务类的特征时, X 轴仅代表 G 业务类的流的个数.

图 2(a)表示 CGRED 的带宽维护特征.根据各业务类流个数的多少,大致可以将 CGRED 的带宽维护特征分为三种情况:极少流个数、中等流个数和大量流个数的情况.当流的个数为中等数目时,CGRED 表现出对分配带宽极强的维护能力,不会出现在 RIO 和 WRED 中存在的、低等级业务类的带宽被全部挤占的情形^[10,11].极少流个数和大量流个数是实际应用时的特例,在这些情况下 CGRED 有可能脱离对分配带宽的维护,表现出链路共享的特征.在仿真中,当 G 业务类的只有一个流时, G 业务类没有使用预先分配的全部带宽,因为 CGRED 使用单队列管理机制,不同业务类的分组共享单个缓存,缓存上的排队时延限制了 G 业务类的流的 TCP 协议吞吐量.即使在这种情况下, G 业务类没有占用的带宽资源也没有被浪费;CGRED 通过增大 M 和 B 业务类的单流速率,这部分带宽由 M 和 B 业务类的用户共享了.进一步讲,提供有区别服务时,高优先级业务类的分配带宽一般较大而且是以相当数量的用户共同使用为目标设计的,因此当该业务类的用户很少时,由这些少量用户分享这个大带宽并不合理,并且实际应用一般也用不完这个带宽,所以将剩余带宽分给低等级业务类共享在实际应用中更为合理.

当高等级业务类的流数目极多(大于 14)时,CGRED 也会脱离对分配带宽的维护,出现高等级业务类抢占低等级业务类带宽的现象,以保证高等级业务类的单个流能得到与低等级业务类的单个流在数值上相当的吞吐量.由此造成了图 2(a)中低等级业务类的流数目极小处对应的曲线低于分配带宽的现象.

图 2(b)表示 CGRED 的单流平均带宽.理想带宽维护能力所导致的平均单流带宽曲线应有 $1/x$ 的形状,图 2(b)中各业务类的单流平均带宽的曲线形状大致吻合这个结论.对于 G 业务类,曲线尾部趋向于一条值近似为 100Kbps 水平的直线,而 100Kbps 正是 20 个流共享 2Mbps 带宽时的带宽平均值.造成这个现象的原因正是前面提到的、CGRED 保证高等级业务类单流速率与低等级业务类单流速率相当的功能.

图 2(c)表示 CGRED 中流速率的不均匀特征,业务类 G 的流速率间的差别不超过 10%.业务类 M 和 B 的流速率间的差别分别小于 20% 和 40%.流速率的一致性变差是因为 CGRED 要不断调整才能保证对分配带宽的维护,但是 CGRED 基本能保证同业务类中各个流所得带宽基本相同.

3.3 参数稳定性的实验

基础实验通过一组参数验证了 CGRED 具有维护分配带宽、提供有区别服务的能力,但只有同时证明 CGRED 对算法所涉及的各个参数(如表 1 所示)的变化不敏感,我们才能宣称这个算法是实用的.仿真结果^[11]表明,单个参数的改变对 CGRED 控制性能的影响很小(限于篇幅略去仿真结果).

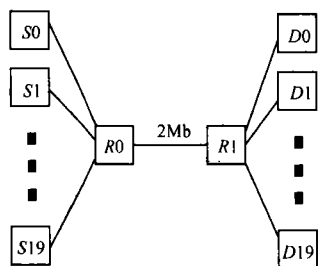


图 1 仿真网络拓扑

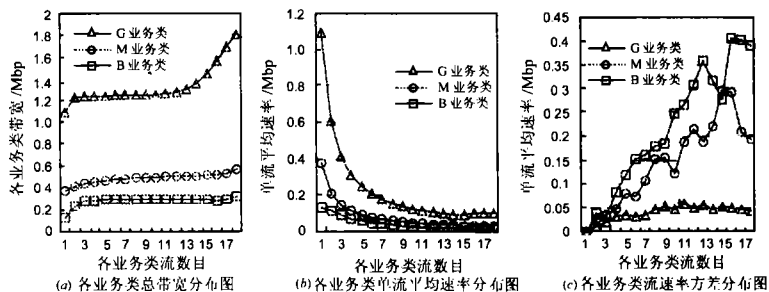


图 2 基础实验的仿真结果

3.4 瓶颈链路带宽以及带宽分配比例变化对 CGRED 性能的影响

为检验瓶颈链路带宽以及带宽分配比例变化对 CGRED 性能的影响,我们将链路带宽增加到 10 Mbps,同时改变总用户数(50 对)和带宽分配比例(6:3:1)。图 3 所示的仿真结果表明:CGRED 保持了与基础实验相类似的特点。当各业务类的

流的个数为中等数目时,各业务类的带宽维持在预先分配的带宽值附近,充分表现出 CGRED 对分配带宽极强的维护能力。此外,当某业务类的流个数偏少时,其未被使用的带宽会被其它流数目偏多的业务类占用。如前所述,这样的好处是提高瓶颈链路的带宽利用率。

4 结论

本文论述了 DiffServ 为什么需要一种既具有带宽控制能力又能保持实现简单性的优先级业务类链路共享机制,并提出了实现这一特征的单队列管理机制—CGRED 机制。通过大量的仿真工作我们证明 CGRED 机制具有如下特点:

(1)CGRED 保持了有区别服务简单易行、扩展性好的特点;

(2)CGRED 能较为精细地维护带宽分配,从而较好地进行业务类质量的区分;

(3)CGRED 的性能稳定,对参数变化不敏感;

(4)CGRED 能保证高优先级业务类的最低单流带宽基本不低于低优先级业务类的单流带宽;

因此 CGRED 是一个实用的提供有区别服务的网络控制机制。

参考文献:

[1] Sally Floyd, Van Jacobson. Link-sharing and resource management models for packet networks [J]. IEEE/ACM Transactions on Networking, 1995, 3(4): 365 - 386.
 [2] NS. Network simulator (Version 2. 1b6) [CP/OL]. URL: http://www.isi.edu/nsnam/ns.html.
 [3] Sally Floyd. Notes on CBQ and guaranteed service [Z]. 1995, 7. URL: http://www.icir.org/floyd/papers/guaranteed.ps.
 [4] Sally Floyd. Notes of class-based queueing: Setting parameters [Z]. 1996, 2, URL: http://www.icir.org/floyd/papers/params.ps.Z.
 [5] Sally Floyd, Michael Francis Speer. Experimental result for class-based queueing [Z]. Jan 1998, URL: http://www.nrg.ee.lbl.gov/floyd/cbq/notes.html.
 [6] 刘紫千. CBQ 的研究与仿真 [D]. 北京:北方交通大学, 2001.
 [7] Clark D, Fang W. Explicit allocation of best-effort packet delivery service [J]. IEEE/ACM Transactions on Networking, 1998, 6(4): 362 - 373.
 [8] Technical specification from Cisco, distributed weighted random early

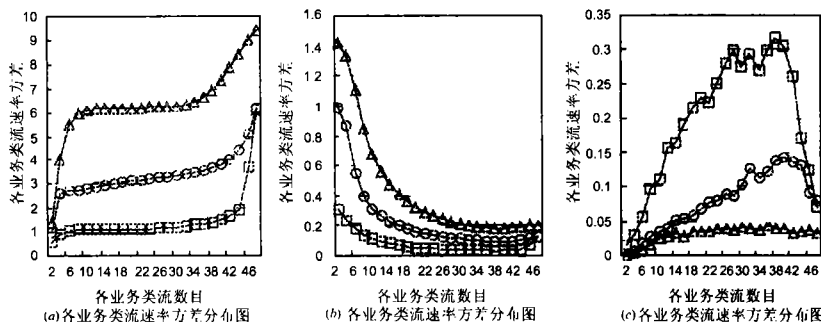


图 3 瓶颈链路实验的仿真结果

detection [Z]. URL: http://www.cisco.com/univercd/cc/td/doc/product/software/ios111/cc111/wred.pdf.

[9] Sally Floyd Van Jacobson. Random early detection gateways for congestion avoidance [J]. IEEE/ACM Transactions on Networking, 1993, 1(4): 397 - 413.
 [10] Ulf Bodin, Olov Schelén, Stephen Pink. Load-tolerant differentiation with active queue management [J]. ACM SIGCOMM Computer Communication Review, 2000, 30(4): 4 - 16.
 [11] 邢新宇. 一种保证 QoS 的新算法-CGRED 的研究与仿真 [D]. 北京:北方交通大学, 2001.

作者简介:



赵永祥 男, 1970 年 7 月出生于云南昆明, 北方交通大学讲师, 博士, 主要研究方向: 网络的调度和缓存管理, 拥塞控制。

邢新宇 女, 1978 年生于天津武清, 北方交通大学通信系在读硕士生, 研究方向: 网络服务质量。

陈常嘉 男, 1949 年 10 月生于澳门, 北方交通大学教授, 主要研究领域: 宽带通信网络技术。

宋光农 女, 1973 年 11 月 4 日生于四川冕宁, 北方交通大学通信系在读博士生, 主要研究方向: 网络服务质量, IP 网中多媒体流的拥塞控制。E-mail: zhangsonggn@sina.com.