

提供端到端时延和时延抖动保障的 QoS 控制

邱 茵¹, 李玉峰^{1,2}, 邬江兴¹

(1. 国家数字交换系统工程技术研究中心, 河南郑州 450002; 2. 防空兵指挥学院信息控制系, 河南郑州 450052)

摘要: 提出了一类具有最大速率控制的速率保障 (Maximum Rate Control-Guaranteed Rate, MRC-GR) 算法, 可对流同时提供速率保障和最大速率控制. 当网络各节点执行 MRC-GR 算法时, 提供了确定网络端到端时延上限和下限的方法, 针对服从令牌桶模型和同步单元模型的业务源给出了网络时延上限和下限. 针对 MRC-GR 算法实例——具有最大速率控制的最差情形公平加权公平排队 (worst-case fair weighted fair queueing with maximum rate control) 调度算法进行仿真实验, 仿真结果验证了理论分析.

关键词: QoS 控制; 时延; MRC-GR 算法; 保障速率; 最大速率控制

中图分类号: TP393 **文献标识码:** A **文章编号:** 0372-2112 (2009) 03-0567-07

QoS Control for Providing End-to-End Delay and Delay Jitter Guarantee

QIU Han¹, LI Yu-feng^{1,2}, WU Jiang-xing¹

(1. National Digital Switching System Engineering & Technology R&D Center, Zhengzhou, Henan 450002, China;

2. Department of Information and Control, Air Defense Command College, Zhengzhou, Henan 450052, China)

Abstract: A class of maximum rate control-guaranteed rate (MRC-GR) algorithm is advanced, which can provide maximum rate control and guaranteed rate for flows simultaneously. For each network node employing scheduling algorithms belonging to MRC-GR, a method is presented to decide the upper bound and the lower bound on network end-to-end delay, and the delay bounds for source models conforming to token bucket and synchronization unit model are given. Simulations for worst-case fair weighted fair queueing with maximum rate control scheduling algorithm, which is an instance of MRC-GR, are carried on to validate the theoretic analysis.

Key words: QoS control; delay; maximum rate control-guaranteed rate (MRC-GR) algorithm; guaranteed rate; maximum rate control

1 引言

随着各种多媒体应用的开发和多媒体技术的发展, Internet 上的多媒体应用越来越普及, 尤其是各种交互式多媒体应用, 例如视频会议、IP 可视电话等. 交互式多媒体应用对时延和时延抖动都很敏感, 作为可感知的 QoS (Quality of Service) 参数, 网络时延和时延抖动直接影响着用户层 QoS^[1,2], 网络同时提供媒体流时延和时延抖动保障是实现交互式媒体流 QoS 保障的重要内容. 目前的网络 QoS 保障技术主要包括 IntServ, DiffServ, MPLS 和流量工程等, 它们提供的是数据流级别的粗粒度 QoS 控制, 不能提供精确的端到端延迟和延迟抖动. 网络提供精确时延和时延抖动 QoS 控制主要是在链路层实现的, 通过在网络节点对数据包转发的调度和队列管理进行.

自上世纪 90 年代以来, 研究人员开发出了大量尽力保障每连接或每队列性能的调度算法^[3-6], Pawan

Goyal 等^[7]将其归纳为保障速率 (Guaranteed Rate, GR) 调度算法, 证明了当业务源服从令牌桶模型 (文献 [4] 和 [7] 中称之为 leaky bucket, 而对模型的描述实属 token bucket, 故本文按照国际惯例称之为令牌桶模型) 时, 这类保障速率调度算法能够提供网络端到端时延上限, 但是通常难以提供端到端时延抖动上限, 可归纳成基于整形的速率控制服务规则. 已提出的非尽职尽责型调度算法, 包括 Jitter-EDD (Jitter Earliest-Due-Date)^[8], stop-and-go^[9], HRR (Hierarchical Round Robin)^[10] 和 RCSP (Rate-Controlled Static Priority)^[11] 等, 文献 [12] 将其归纳为一类基于速率控制服务规则, 该类算法能够基于时间帧策略或者基于分类优先级策略机制提供端到端时延上限、时延抖动上限, 但是需要配合速率控制器对业务进行整形, 需要维护两个队列, 且速率控制器实现难度大. 近年来提出的具有最大速率控制的 WF²Q (Worst-case Fair-Weighted fair Queueing with Maximum rate control, WF²Q-M) 调度算法^[13] 能够同时提供速率保障和最大速率控制,

收稿日期: 2008-04-25; 修回日期: 2008-11-11

基金项目: 国家重点基础研究发展规划 (973 计划) 项目 (No. 2007CB307102); 国家高技术研究发展计划 (863 计划) 课题 (No. 2007AA01Z2a1)

但关于时延、时延抖动性能并未进行深入研究。

本文提出一类 MRC-GR (Maximum Rate Control-Guaranteed Rate) 算法, 证明 GPS-M 模型和 WF²Q-M 调度算法属于 MRC-GR。在各节点均执行 MRC-GR 算法的情况下, 提供确定网络端到端时延上限和下限的方法, 针对业务源服从令牌桶模型和同步单元模型, 给出其网络端到端时延上限和下限。以 MRC-GR 算法中的 WF²Q-M 算法为例进行仿真实验, 验证理论分析的结果。对 MRC-GR 算法进行实用性分析, 说明 MRC-GR 算法可在实际网络环境中对多媒体业务的网络端到端时延上限和下限进行控制。

2 MRC-GR 算法

2.1 MRC-GR 算法

媒体流的媒体单元在网络上采取分组形式进行传输, 一个源发送的一系列分组可称为流^[5]。本文所定义的 MRC-GR 算法是针对流提供时延保障的一类基于速率控制的调度算法, 其时延保障基于每个分组的保障速率时钟值 (GRC, Guaranteed Rate Clock) 值和最大速率时钟值 (MRC, Maximum Rate Clock) 实现, 其中, GRC 定义沿用 GR 调度算法中的 GRC^[7] 定义:

定义 1 (GRC) 任取流 f , 假设其保障速率为 r_f (bit/s), p_f^j 表示流 f 的第 j 个分组, l_f^j 为分组长度, $A^i(p_f^j)$ 表示分组 p_f^j 到达服务器 i 的时间, $GRC^i(p_f^j)$ 表示分组 p_f^j 在服务器 i 处的 GRC 值, 那么流 f 分组 p_f^j 的 GRC 值定义如下:

$$GRC^i(p_f^j) = \begin{cases} 0, & j=0 \\ \max\{A^i(p_f^j), GRC^i(p_f^{j-1})\} + \frac{l_f^j}{r_f}, & j=1 \end{cases} \quad (1)$$

最大速率时钟值 MRC 定义如下:

定义 2 (MRC) 任取流 f , 假设其最大控制速率为 R_f (bit/s), p_f^j 表示流 f 的第 j 个分组, l_f^j 是分组的长度, $A^i(p_f^j)$ 表示分组 p_f^j 到达服务器 i 的时间, $MRC^i(p_f^j)$ 表示分组 p_f^j 在服务器 i 处的 MRC 值, 那么流 f 分组 p_f^j 的 MRC 值定义如下:

$$MRC^i(p_f^j) = \begin{cases} 0, & j=0 \\ \max\{A^i(p_f^j), MRC^i(p_f^{j-1})\} + \frac{l_f^j}{R_f}, & j=1 \end{cases} \quad (2)$$

利用 GRC 值和 MRC 值, 本文定义 MRC-GR 算法如下:

定义 3 (MRC-GR 算法) 对于流 f , 如果服务器 i 的调度算法保证流分组在时刻 $MRC^i(p_f^j)$ 之后, 时刻 $GRC^i(p_f^j) + \tau$ 之前传输, 则该调度算法属于 MRC-GR 算法, 其中 τ 是由调度算法和服务器决定的调度常数。

比较 GR 定义^[7]不难发现, MRC-GR 算法是一类特殊的 GR 调度算法, 不同点在于 MRC-GR 对流的最大速

率进行控制。

2.2 MRC-GR 算法实例分析

目前已公开的 WF²Q-M 调度算法是 GPS-M^[13] 模型的一个分组化实例, 分析及实验表明^[13], WF²Q-M 能够同时提供保障速率和最大速率控制。下面分别在确定保障速率下论证 GPS-M 模型和 WF²Q-M 调度算法属于本文定义的 MRC-GR 算法。

2.2.1 GPS-M 模型

GPS-M 模型是对 GPS 模型^[2]进行最大速率控制功能扩展的模型。在 GPS-M 中, 服务器 i 为每个流 f 分配正常数 ϕ_f^i , 表示每个流在分配处理资源时所占的权重, $B^i(t)$ 表示 t 时刻服务器 i 所有待处理流的集合, C^i 为服务器 i 的处理能力。对于最大控制速率为 R_f 的流 f 而言, 如果按照 GPS 为流 f 分配的服务速率大于最大控制速率 R_f , 则称该流饱和, 定义 $B_s^i(t)$ 表示服务器 i 在 t 时刻的饱和流集合, $\phi_{B_s^i(t)}^i = \phi_g^i$, $\overline{B_s^i(t)}$ 表示集合 $B^i(t)$ 除 $B_s^i(t)$ 外的部分。

GPS-M 模型下, 服务器 i 按照流的 $\overline{B_s^i(t)}$ 最大服务速率对 $B_s^i(t)$ 集合的流进行服务, $B_s^i(t)$ 集合中的流按照各自权重分享剩余带宽, 分配给流 f 的带宽 $r_f^i(t)$ 如下:

$$r_f^i(t) = \begin{cases} R_f, & \text{if } f \in B_s^i(t) \\ \phi_f^i \cdot (C^i - \sum_{g \in B_s^i(t)} R_g) / (\phi_{B_s^i(t)}^i - \phi_{B_s^i(t)}^i), & \text{otherwise} \end{cases} \quad (3)$$

其中, $\phi_{B_s^i(t)}^i = \sum_{g \in B_s^i(t)} \phi_g^i$ 。当 $B_s^i(t)$ 集合为空时, 所有流均为饱和流, $\phi_{B_s^i(t)}^i - \phi_{B_s^i(t)}^i$ 为 0, 流 f 的带宽按照最大控制速率分配, 不会出现 $\phi_{B_s^i(t)}^i - \phi_{B_s^i(t)}^i$ 作为分母的情况。此时, $\sum_{g \in B_s^i(t)} R_g < C^i$, GPS-M 模型处于非尽职工作状态。当 $B_s^i(t)$ 集合为空时, GPS-M 模型即为 GPS。

GPS-M 模型和 GPS 模型相同点在于服务器同时对所有流进行服务, 服务无限细分, 不同点在于 GPS 服务器以速率 $\left(\frac{\phi_f^i}{\phi_{B^i(t)}^i}\right) \cdot C^i$ 对流 f 进行服务, GPS-M 模型按照 $r_f^i(t)$ 分配给流 f 服务速率。

对于保障速率为 r_f , 最大控制速率为 R_f 的流 f , GPS-M 模型分配流 f 的权重 ϕ_f^i 必须满足: $\left(\frac{\phi_f^i}{\phi_{B^i(t)}^i}\right) \cdot C^i \geq r_f$, 则依据表达式 (3), 带宽 $r_f^i(t)$ 满足以下关系:

$$R_f \cdot r_f^i(t) \geq \phi_f^i \cdot \left\{ C^i - \sum_{g \in B_s^i(t)} R_g \right\} / \left\{ \phi_{B^i(t)}^i - \phi_{B_s^i(t)}^i \right\} > \phi_f^i \cdot \left\{ C^i - \frac{\phi_{B_s^i(t)}^i}{\phi_{B^i(t)}^i} C^i \right\} / \left\{ \phi_{B^i(t)}^i - \phi_{B_s^i(t)}^i \right\} = r_f \quad (4)$$

假设 $A^i(p_f^j)$ 为分组 p_f^j 到达服务器 i 的时刻, $L_{GPS-M}^i(p_f^j)$ 表示服务器 i 执行 GPS-M 下分组 p_f^j 离开时刻, 则存在:

$$\max\{A^i(p_f^j), L_{GPS-M}^i(p_f^{j-1})\} + \frac{l_f^j}{r_f} L_{GPS-M}^i(p_f^j)$$

$$\max\{A^i(p_f^j), L_{GPS-M}^i(p_f^{j-1})\} + \frac{l_f^j}{R_f}, j = 1 \quad (5)$$

令 $L_{GPS-M}^i(p_f^0) = GRC^i(p_f^0) = MRC^i(p_f^0) = 0$, 则由不等式(5), 可以推出:

$$MRC^i(p_f^j) \leq L_{GPS-M}^i(p_f^j) \leq GRC^i(p_f^j), j = 1 \quad (6)$$

对照定义 3, 如果服务器 i 分配给流 f 的权重 ϕ_j^i 满足 $\left(\frac{\phi_j^i}{\phi_{B^i(i)}^i}\right) \cdot C^i \leq r_f$, 则 GPS-M 模型对于流 f 而言属于 MRC-GR 算法, 且 $i = 0$.

2.2.2 WF²Q-M 调度算法

GPS-M 模型是一种理想的服务器模型, WF²Q-M 模型针对 GPS-M 模型进行了分组实例化, 引入了分组合格时间, 提出虚拟时间调节机制, 将 GPS-M 时间映射为 WF²Q-M 虚拟系统时间, 根据虚拟开始时间和虚拟结束时间的先后顺序对分组进行服务, 以此来实现饱和流外的其余待处理流的带宽分配, 详细实施步骤见文献 [13].

文献 [13] 中定理 5 证明了 WF²Q-M 性能受限于 GPS-M 模型, WF²Q-M 调度算法下分组 p_f^j 的离开时间 $L_{WF^2Q-M}^i(p_f^j)$ 满足如下关系:

$$L_{WF^2Q-M}^i(p_f^j) \leq L_{GPS-M}^i(p_f^j) + \frac{l_{max}^i}{C^i}, j = 1 \quad (7)$$

其中, l_{max}^i 为服务器 i 服务的最大分组长度. 由于 WF²Q-M 的分组合格时间定义与定义 2 完全相同, 则 $L_{WF^2Q-M}^i = MRC^i(p_f^j)$, 结合不等式(7) 即有: $MRC^i(p_f^j) \leq L_{WF^2Q-M}^i(p_f^j)$

$$L_{GPS-M}^i(p_f^j) + \frac{l_{max}^i}{C^i}, j = 1.$$

依照不等式(6) 可得:

$$MRC^i(p_f^j) \leq L_{WF^2Q-M}^i(p_f^j) \leq GRC^i(p_f^j) + \frac{l_{max}^i}{C^i}, j = 1 \quad (8)$$

对照定义 3, 如果服务器 i 分配给流 f 的权重 ϕ_j^i 满足 $\left(\frac{\phi_j^i}{\phi_{B^i(i)}^i}\right) \cdot C^i \leq r_f$, 则 WF²Q-M 调度算法对于流 f 而言属于 MRC-GR 算法, 且 $i = \frac{l_{max}^i}{C^i}$.

GR 调度算法除了包括 PGPS 外还有 VC (Virtual Clock) [5] 和 SCFQ (Self-Clocked Fair Queueing) [3] 等. 依照上面关于 GPS-M 和 WF²Q-M 的分析, 不难推断具有最大速率控制的 VC-M (VC with Maximum Rate Control) 和具有最大速率控制的 SCFQ-M (SCFQ with Maximum Rate Control) 在确定保障速率时也属于 MRC-GR 算法.

3 MRC-GR 算法网络时延界限确定

3.1 网络时延界限确定

假设流 f 在网络中需流经的节点总数为 K , 路径上的第 m 个节点用 m 表示, 0 和 $K+1$ 分别表示源和目的, 每个节点都执行 MRC-GR 算法, 对流 f 的保障速率均为 r_f , 最大控制速率均为 R_f , 节点 m 和 $m+1$ 间的传播时延为 $t_{m,m+1}$. 假设分组 p_f^j 到达节点 1 的时刻为 $A^1(p_f^j)$, 到达节点 $K+1$ 的时间为 $A^{K+1}(p_f^j)$, 则分组 p_f^j 经历的网络时延 d_f^j 可表示为:

$$d_f^j = A^{K+1}(p_f^j) - A^1(p_f^j) + t_{0,1} \quad (9)$$

依据定义 3, 路径上的任一节点 m 在 $MRC^i(p_f^j)$ 时刻之后, $GRC^m(p_f^j) + t_{m,m+1}$ 时刻之前必须传输分组 p_f^j , 则分组 p_f^j 到达的节点 $m+1$ 的时间与分组在节点 m 处的 MRC 值、GRC 值满足以下关系:

$$MRC^m(p_f^j) + t_{m,m+1} \leq A^{m+1}(p_f^j) \leq GRC^m(p_f^j) + t_{m,m+1} \quad (10)$$

MRC-GR 是具有最大速率控制的 GR 调度算法, 其 GRC 值与 GR 调度算法下的 GRC 值相同, 故文献 [7] 中引理 1 给出了 $GRC^m(p_f^j)$ 与 $GRC^{m+1}(p_f^j)$ 的迭代关系亦成立, 重写如下:

引理 1 如果流 f 流经的节点 m 执行 MRC-GR 算法, 则对于流 f 的分组 p_f^j 有:

$$GRC^{m+1}(p_f^j) \leq GRC^m(p_f^j) + \max_{i \in \{1, \dots, j\}} \frac{l_f^i}{r_f} + t_{m,m+1}, m = 1 \quad (11)$$

证明过程详见文献 [7] 引理 1 证明.

对于 MRC 值, 存在引理 2 所述关系:

引理 2 如果节点 m 处对流 f 执行 MRC-GR 算法, 则对于流 f 的分组 p_f^j 有:

$$MRC^{m+1}(p_f^j) \leq MRC^m(p_f^j) + \frac{l_f^j}{R_f} + t_{m,m+1}, m = 1 \quad (12)$$

证明: 由定义 3, 可得:

$$MRC^{m+1}(p_f^j) \leq A^{m+1}(p_f^j) + \frac{l_f^j}{R_f} \quad (13)$$

节点 m 对流 f 执行 MRC-GR 算法, 则依据定义 2 和定义 3:

$$A^{m+1}(p_f^j) \leq MRC^m(p_f^j) + t_{m,m+1} \quad (14)$$

综合不等式(13) 和(14), 可以推出不等式(12), 即引理 2 得证.

结合引理 1 和引理 2, 可得到引理 3:

引理 3 如果流 f 路径上的节点均对其执行 MRC-GR 算法, 则流 f 分组 p_f^j 所经历的网络时延 d_f^j 满足如下关系:

$$K \frac{l_f^j}{R_f} + t_{m=0}^{m=K} \leq d_f^j \leq GRC^1(p_f^j) - A^1(p_f^j) + (K-1) \max_{i \in \{1, \dots, j\}} \frac{l_f^i}{r_f} + t_{m=1}^{m=K} + t_{m=0}^{m=K} \quad (15)$$

证明:由于节点均对流 f 执行 MRC-GR 算法,则综合表达式(9)和不等式(10)可得:

$$\text{MRC}^K(p_f^j) - A^1(p_f^j) + \sum_{m=0}^{K, K+1} \frac{l_f^m}{R_f} \text{GRC}^K(p_f^j) - A^1(p_f^j) \quad (16)$$

运用引理 1 和引理 2 分别对不等式(16)的右侧表达式和左侧表达式进行 K 次迭代,即可得:

$$\text{MRC}^1(p_f^j) - A^1(p_f^j) + (K-1) \frac{l_f^m}{R_f} + \sum_{m=0}^K \frac{l_f^m}{R_f} \text{GRC}^1(p_f^j) - A^1(p_f^j) + (K-1) \max_{i \in \{1, \dots, j\}} \frac{l_f^i}{r_f} + \sum_{m=0}^K \frac{l_f^m}{R_f} \quad (17)$$

又由定义 2, $\text{MRC}^1(p_f^j) = A^1(p_f^j) + \frac{l_f^m}{R_f}$, 则不等式(15)成立. 引理 3 得证.

引理 3 给出了分组 p_f^j 经历的网络时延上限和下限,其中,上限仅与分组 p_f^j 在节点 1 处的 GRC 值及到达节点 1 的时间有关,下限仅与节点个数有关. 分析流到达过程,可得定理 1:

定理 1 设流 f 的第 j 个分组为 p_f^j , l_f^j 表示分组长度, $A^m(p_f^j)$ 为分组 p_f^j 到达节点 m 的时间, $AP_f^m(t_1, t_2)$ 表示流 f 在间隔 $[t_1, t_2]$ 内到达节点 m 的比特数,若流 f 在网络上流经的 K 个节点对其都执行 MRC-GR 算法,且保障速率均为 r_f , 最大控制速率均为 R_f , $r_f < R_f$, 则分组 p_f^j 的网络时延 d_f^j 满足以下关系:

$$K \frac{l_f^m}{R_f} + \sum_{m=0}^{m, m+1} \frac{l_f^m}{R_f} \text{GRC}^1(p_f^j) - A^1(p_f^j) + \frac{AP_f^1(A^1(p_f^k), A^1(p_f^j))}{r_f} + (K-1) \max_{i \in \{1, \dots, j\}} \frac{l_f^i}{r_f} + \sum_{m=1}^K \frac{l_f^m}{R_f} + \sum_{m=0}^K \frac{l_f^m}{R_f} \quad (17)$$

其中, $\frac{l_f^m}{R_f}$ 为节点 m 和 $m+1$ 间的传播时延.

证明:函数 $AP_f^m(t_1, t_2)$ 表示流 f 在间隔 $[t_1, t_2]$ 内到达节点 m 的比特数,只有当分组的最后一个比特到达节点时才认为该分组到达. 因此,到达函数 $AP_f^m(t_1, t_2)$ 由每个分组到达时刻脉冲构成,是右连续的,且 $AP_f^m(t, t)$ 表示时刻 t 到达的分组长度.

定义流 f 在节点 m 处的分组集合 S_f^m 为前一个分组 GRC 值小于该分组到达时间的分组序号集合,即: $S_f^m = \{h | h > 0, \text{GRC}^i(p_f^{h-1}) < A^i(p_f^h)\}$. 由定义 1 及 S_f^m 定义可知,对于任意分组 p_f^j 存在下式:

$$\text{GRC}^m(p_f^j) = A^m(p_f^k) + \sum_{n=0}^{j-k} \frac{l_f^{k+n}}{r_f}, k = \max\{h | h \in S_f^m\} \quad (18)$$

由于 $\sum_{n=0}^{j-k} \frac{l_f^{k+n}}{r_f} = AP_f^m(A^m(p_f^k), A^m(p_f^j))$, 则表达式(18)

变为:

$$\text{GRC}^m(p_f^j) = A^m(p_f^k) + \frac{AP_f^m(A^m(p_f^k), A^m(p_f^j))}{r_f} \quad (19)$$

对于节点 1 而言,流 f 直接由源经链路传送, $AP_f^1(t, t)$ 反映了流的业务源特征. 综合表达式(19)和引理 3 的结论(15),即可推出不等式(17),故定理 1 得证.

从不等式(17)中可以看出:定理 1 给出了确定网络时延上限和下限的方法,当 $A^1(p_f^k) - A^1(p_f^j) + \frac{AP_f^1(A^1(p_f^k), A^1(p_f^j))}{r_f}$ 有确定上限时,网络时延存在上限,而网络时延由分组长度、节点个数和最大控制速率决定,与具体业务源模型无关;时延上限和下限在 $r_f = R_f$ 时相差最小,即节点提供近似恒定速率服务时,网络时延范围变化最小.

定理 1 说明了 MRC-GR 算法对于任何业务源模型能够提供网络时延下限,下面针对令牌桶模型^[5]、同步单元模型^[14]进行具体的网络时延上、下限分析.

3.2 令牌桶模型网络时延界限确定

基于整形的速率控制服务规则采用令牌桶或漏桶进行业务整形,以协助调度算法实现 QoS 控制,令牌桶模型在限制数据平均传输速率的同时还允许某种程度的突发传输,本文以令牌桶为业务源模型确定其网络时延界限. 若流 f 服从参数为 (r_f, σ_f) 的令牌桶模型,即突发长度为 σ_f , 平均速率为 r_f , 则有:

$$AP_f^m(t_1, t_2) = \sigma_f + r_f(t_2 - t_1), 0 \leq t_1 < t_2 \quad (20)$$

依照定理 1, 分组 p_f^j 的网络时延 d_f^j 满足如下关系:

$$K \frac{l_f^m}{R_f} + \sum_{m=0}^{m, m+1} \frac{l_f^m}{R_f} d_f^j - \frac{l_f^m}{r_f} + \left(\frac{\sigma_f}{r_f} - 1 \right) (A^1(p_f^j) - A^1(p_f^k)) + (K-1) \max_{i \in \{1, \dots, j\}} \frac{l_f^i}{r_f} + \sum_{m=1}^K \frac{l_f^m}{R_f} + \sum_{m=0}^K \frac{l_f^m}{R_f} \quad (21)$$

当 $\sigma_f > r_f$, 网络时延 d_f^j 具有与分组 p_f^j 、 p_f^k 到达时刻无关的上限和下限,即:

$$K \frac{l_f^m}{R_f} + \sum_{m=0}^{m, m+1} \frac{l_f^m}{R_f} d_f^j - \frac{l_f^m}{r_f} + (K-1) \max_{i \in \{1, \dots, j\}} \frac{l_f^i}{r_f} + \sum_{m=1}^K \frac{l_f^m}{R_f} + \sum_{m=0}^K \frac{l_f^m}{R_f} \quad (21)$$

若 $\sigma_f < r_f$ 时,网络时延 d_f^j 上限与分组 p_f^j 、 p_f^k 到达间隔有关,随着时间的流逝分组 p_f^j 、 p_f^k 到达间隔可能越来越大,无法获得确定的上限.

由上述分析可知,针对服从令牌桶模型 (r_f, σ_f) 的流 f , MRC-GR 算法对其保障速率 r_f 满足 $r_f > \sigma_f$ 时,能够提供确定的网络时延上限和下限. 不等式(21)给出的时延上限与文献[7]中定理 3 结论相同,这表明最大速率控制对时延上限无改善作用.

3.3 同步单元模型网络时延界限确定

流媒体业务源模型很难描述,通常采用寻找业务

曲线一般形状的方法. 对于音/视频流而言, 流由多个同步单元组成, 一个同步单元是具有相同确定时延偏移的分组组合^[15], 可以用同步单元模型 (\min, \max, M) ^[14]来描述媒体流的一般性特征, 其中 \min 表示整个同步单元的平均速率, 而 \max 代表同步单元的峰值速率, M 表示同步单元中分组个数. 最常见的同步单元是语音中的共振峰和视频中的帧. 依照定理 1, 同步单元模型下网络时延 d_f^j 满足:

$$K \frac{l_f^j}{R_f} + \sum_{m=0}^{K-1} \max_{i \in \{1, \dots, j\}} \left(-\frac{1}{r_{f, \max}} + \frac{1}{r_{f, \min}} \right) (A P_f^j(A^1(p_f^k), A^1(p_f^j))) + (K-1) \max_{i \in \{1, \dots, j\}} \frac{l_f^j}{r_{f, \min}} + \sum_{m=1}^K \sum_{m=0}^{m-1} \dots$$

当 $r_{f, \max} = r_{f, \min}$ 时, 可得:

$$K \frac{l_f^j}{R_f} + \sum_{m=0}^{K-1} \max_{i \in \{1, \dots, j\}} \left(\frac{1}{r_{f, \min}} - \frac{1}{r_{f, \max}} \right) \frac{l_f^j}{r_{f, \min}} + (K-1) \max_{i \in \{1, \dots, j\}} \frac{l_f^j}{r_{f, \min}} + \sum_{m=1}^K \sum_{m=0}^{m-1} \dots \quad (22)$$

当 $r_{f, \max} = r_{f, \min}$ 时, 表明 MRC-GR 算法以峰值速率对流进行保障速率服务, 可以提供更加严格的时延上限:

$$(K-1) \max_{i \in \{1, \dots, j\}} \frac{l_f^j}{r_{f, \min}} + \sum_{m=1}^K \sum_{m=0}^{m-1} \dots$$

由于 \max 可能远大于 \min , 从提高网络资源利用率的角度看, r_f, R_f 一般满足如下关系: $\min r_f < R_f < \max$.

当流 f 中的分组等长且 $r_f = R_f, K=1$ 时, 不等式 (22) 给出网络时延的上限和下限之差比文献 [14] 中定理 5 的结论紧 $\frac{l_f^j}{R_f} - m$. 这是因为文献 [14] 中定理 5 的结论以尽职工作型调度算法为研究基础, 而不等式 (22) 针对 MRC-GR 算法, 当 $R_f < C_i$ 时 MRC-GR 算法工作于非尽职工作状态, 而 $R_f = C_i$ 时 MRC-GR 算法工作于尽职工作状态. 若将流 f 的最大控制速率设置为链路带宽, 即不提供最大速率控制, 此时 MRC-GR 算法下流 f 的网络时延上限与下限之差与文献 [14] 中定理 5 结论完全一致. 这表明了不等式 (22) 的正确性, 也揭示了网络时延上限和下限之差主要取决于业务的突发情况.

4 仿真实验

为了验证 MRC-GR 算法性能, 本文利用 ns-2 网络

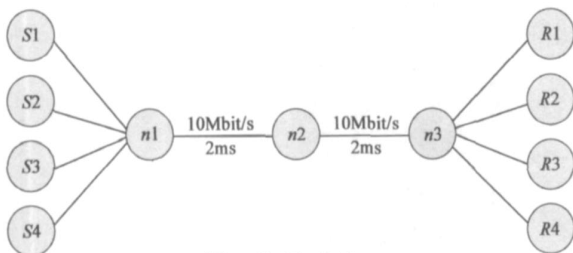


图1 仿真拓扑图

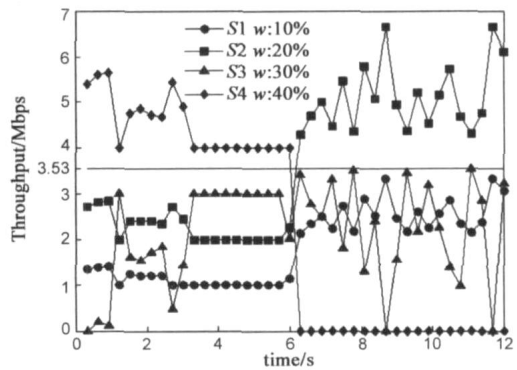
模拟器^[16]分别实现了 WF²Q 调度算法和 MRC-GR 算法的一个实例——WF²Q-M 调度算法, 并进行了速率控制和时延性能的对比仿真实验, 仿真拓扑如图 1 所示.

表 1 会话传输时间及权重

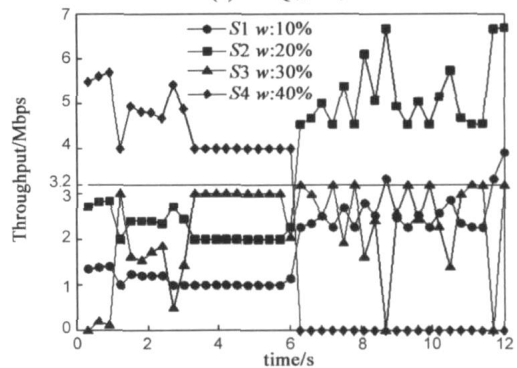
会话	时间 (s)	权重
S1-R1	0-12	10%
S2-R2	0-12	20%
S3-R3	0-12	30%
S4-R4	0-6	40%

图 1 中, 发送端 S1、S2、S4 分别产生 6Mbit/s 的 UDP CBR 业务, 其对应的接收端分别为 R1、R2、R4. 发送端 S3 采用 Pareto ON-OFF 过程产生 UDP 业务, 其对应的接收端为 R3, ON 和 OFF 状态的突发长度均为 300ms, UDP 业务在 ON 状态时以 6Mbit/s 恒定速率产生分组, 然后采用参数为 (22500 字节, 3Mbit/s) 的令牌桶进行业务整形. 每对发送端和接收端之间建立一个会话, 会话的传输时间及节点 n1、n2 分配给会话的权重如表 1 所示. 执行 WF²Q-M 调度算法时分配给会话 S3-R3 的保障速率为 3Mbit/s, 最大控制速率为 3.2Mbit/s. 分组长度为 1500 字节. 仿真实验对节点 n1、n2 和 n3 之间的吞吐率和时延进行测量.

图 2 (a) 给出了节点 n1、n2 执行 WF²Q 调度算法时节点 n2 和 n3 之间各会话的吞吐率, 图 2 (b) 给出了节点 n1、n2 执行 WF²Q-M 调度算法时节点 n2 和 n3 之间的吞吐率情况. 仿真时间 0-3 秒时, 由于 ON-OFF 源的强突发性, 令牌桶输出的速率在某些时间点低于保障速率, WF²Q 和 WF²Q-M 调度算法分配给会话 S3-R3 的速



(a) WF²Q 吞吐率



(b) WF²Q-M 吞吐率

图2 WF²Q和WF²Q-M调度算法下吞吐率

率即为令牌桶输出速率;仿真时间 3~6 秒时,令牌桶输出速率达到 3Mbit/s 以上,WF²Q 调度算法和 WF²Q-M 调度算法均以保障速率服务;仿真时间 6 秒以后,由于 WF²Q 调度算法分配给会话 S3-R3 的速率大于令牌桶输出的最大速率(3.5333Mbit/s),WF²Q 调度算法以令牌桶输出速率为会话 S3-R3 服务,而 WF²Q-M 调度算法为会话 S3-R3 提供保障速率服务同时将最大速率控制在 3.2Mbit/s 以内.对比图 2(a)和图 2(b)可知,WF²Q-M 作为 MRC-GR 算法一个实例能够提供严格的最大速率控制.

图 3(a)给出了节点 n1、n2 执行 WF²Q 调度算法时会话 S3-R3 在节点 n1 和 n3 之间经历的最大时延和最小时延以及理论上的时延上限^[7](70.4ms),图 3(b)给出了执行 WF²Q-M 调度算法时会话 S3-R3 的时延情况以及定理 1 给出的时延上限(70.4ms)和下限(11.5ms).从图 3(a)可以看出,经过令牌桶整形,在 WF²Q 调度算法下,会话 S3-R3 各仿真时刻的最大时延值均小于理论的时延上限,仿真时间 0~6 秒时,最大时延值逼近理论时延上限,而仿真时间 6~12 秒时,最大时延值明显小于时延上限,这是因为仿真时间 6~12 秒 WF²Q 算法分配给会话 S3-R3 的速率为 5Mbit/s,远大于令牌桶输出的最大速率,无需缓存突发分组.从图 3(b)可以看出,在 WF²Q-M 调度算法下,会话 S3-R3 各时间点的最大时延值均小于定理 3 给出的理论时延上限,仿真时间 0~6 秒时,最大时延值逼近时延上限;仿真时间 6~12

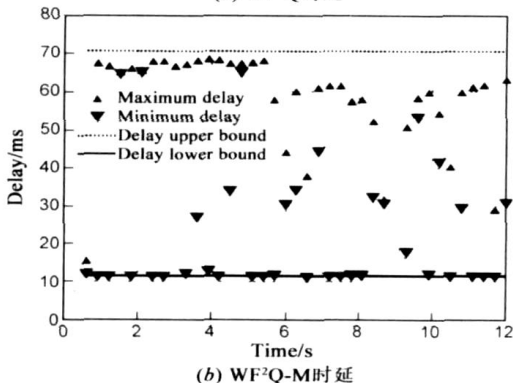
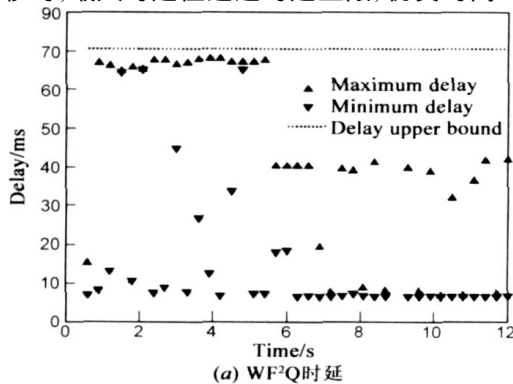


图3 WF²Q和WF²Q-M调度算法下时延

秒时,最大时延值略小于时延上限,而大于 WF²Q 调度算法的最大时延值.这一现象源于 WF²Q-M 算法对会话 S3-R3 执行最大速率控制,当令牌桶输出速率大于 3.2Mbit/s 时,以 3.2Mbit/s 进行服务,此时相当于保障速率为 3.2Mbit/s,对应的理论时延上限为 66.4ms,略小于 70.4ms.对比图 3(a)和图 3(b)可知,会话 S3-R3 在 WF²Q-M 调度算法下各仿真时间点的最小时延值大于 WF²Q 调度算法下的最小时延值,且均大于定理 1 给出的时延下限值,这表明 WF²Q-M 调度算法能够提供严格的时延上限和下限控制.

5 实用性分析

定理 1 提供了流 f 分组 p_f^i 的网络时延 d_f^i 上限和下限的确定方法,当 K 确定时,定理 1 即给出了流 f 的网络时延 d_f 上限和下限的确定方法.对于媒体流 f 而言,可以假设 K 是确定的,该假设基于以下两个事实:

实际网络中,同一媒体流的分组具有相同的源 IP 地址、目的 IP 地址等路由协议运行参数,因此通常具有相同的网络路径;

分组保序功能已经成为了网络节点设备的基本功能.

由此可知,MRC-GR 算法可实现媒体流的网络时延上限和下限控制.

时延和时延抖动是用户可以感知的 QoS 参数,不同应用的媒体流对系统端到端时延和时延抖动的要求可以通过测量获得,例如语音会话能接受的最大单向时延为 150ms,最大时延抖动为 10ms^[16],而时延抖动即为时延上限和时延下限之差,因此,MRC-GR 算法可以通过调节保障速率和最大控制速率实现不同媒体流要求的时延和时延抖动保障.

另外,MRC-GR 算法的具体实现可以采用 WF²Q-M 调度算法.WF²Q-M 调度算法的算法复杂度为 $O(N * M)$,其中 N 代表所有流数目, M 代表“最大速率控制”类流的数目.文献[13]分析表明,当 M 较小时,WF²Q-M 进行虚拟时间计算的开销将很高.实际算法实现中,只有当所有流的权值之和发生变化时,WF²Q-M 才重新计算虚拟时间,故实际计算开销大大降低.与 WF²Q 调度算法相比,WF²Q-M 需要额外的虚拟时间计算开销,当这些额外开销与 M 成正比,且 M 只占所有流的小部分比例时,这些额外开销是可以忽略的.对于路由器而言,WF²Q-M 调度算法带来的计算开销的增加相对于网络端到端时延上限和下限控制的实现是可以容忍的,工程上可实现.

6 结论

本文首先提出了一类基于速率控制的 MRC-GR 算法,能够针对流同时提供速率保障和最大速率控制,论

证表明 GPS-M 模型及其分组化实例 WF^2Q-M 调度算法属于 MRC-GR 算法。随后,在网络执行 MRC-GR 算法下提出了确定网络端到端时延上限和下限的方法,并且针对业务流服从令牌桶模型和同步单元模型分别给出了其网络时延上限和下限。接着,针对 MRC-GR 算法的一个实例—— WF^2Q-M 调度算法进行仿真实验,仿真结果验证了理论分析结果。最后,对 MRC-GR 算法的实用性进行了分析,说明在媒体流内分组的路径确定,并且路径上各个节点设备都能对分组按序转发的前提条件下,MRC-GR 算法能够实现对流媒体业务的时延和时延抖动保障。

MRC-GR 算法在保障流的服务速率同时对流传输的最大速率进行限制,将业务突发降低到一定范围,起到业务整形的作用,且无需维护两个队列,降低了网络节点的管理复杂度;执行 MRC-GR 算法的网络只需调节保障速率和最大控制速率即可实现网络端到端时延上限和下限控制,适于实际网络环境应用。

目前已公开的 MRC-GR 算法有 WF^2Q-M 调度算法, VC-M 及 SCFQ-M 调度算法有待下一步研究。

参考文献:

- [1] Ito Y, Tasaka S, Fukuta Y. Psychometric analysis of the effect of end-to-end delay on user-level QoS in live audio-video transmission [A]. Proceedings of IEEE ICC2004 [C]. Paris: IEEE Communication Society Press, 2004. 2214 - 2220.
- [2] Ribadeneura A F. An analysis of the MOS under conditions of delay, jitter and packet loss and an analysis of the impact of introduction piggybacking and reed solomon for FEC FOR VOIP [D]. Georgia State: College of Arts and Sciences Georgia State University, 2007.
- [3] Golestani S J. A self-clocked fair queueing scheme for high speed applications [A]. Proceedings of IEEE INFOCOM '94 [C]. Toronto, Canada: IEEE Computer Society Press, 1994. 636 - 646.
- [4] Parekh A K. A generalized processor-sharing approach to flow control in integrated services networks [D]. Cambridge, MA: The MIT Press, 1992.
- [5] Zhang L. Virtual clock: a new traffic control algorithm for packet-switching networks [A]. Proceedings of ACM SIGCOMM '90 [C]. New York: ACM Press, 1990. 19 - 29.
- [6] Bennett J C R, Zhang H. WF^2Q : worst-case fair weighted fair queueing [A]. Proceedings of IEEE INFOCOM '96 [C]. San Francisco: IEEE Press, 1996. 120 - 128.
- [7] Goyal P, Lam S S, Vin H M. Determining end-to-end delay bounds in heterogeneous networks [J]. Multimedia Systems, 1997, 5(3): 157 - 163.
- [8] Verma D, Zhang H, Ferrari D. Guaranteeing delay jitter bounds in packet switching networks [A]. Proceedings of Tricom 91 [C]. Chapel Hill, North Carolina: IEEE Press, 1991. 35 - 46.
- [9] Golestani S. A stop-and-go queueing framework for congestion management [A]. Proceedings of ACM SIGCOMM '90 [C]. Philadelphia: ACM Press, 1990. 8 - 18.
- [10] Kalmanek C, Kanakia H, Keshav S. Rate controlled servers for very high-speed networks [A]. Proceedings of IEEE GLOBECOM '90 [C]. San Diego: IEEE Press, 1990. 12 - 20.
- [11] Zhang H, Ferrari D. Rate-controlled static priority queueing [A]. Proceedings of IEEE INFOCOM '93 [C]. San Francisco, California: IEEE Press, 1993. 227 - 236.
- [12] Zhang H, Ferrari D. Rate-controlled service disciplines [J]. Journal of High Speed Networks, 1994, 3(4): 389 - 412.
- [13] Lee J F, Chen M C, Sun Y L. WF^2Q-M : Worst-case fair weighted fair queueing with maximum rate control [J]. Computer Networks, 2007, 51(6): 1403 - 1420.
- [14] Wang Z, Crowcroft J. Analysis of burstiness and jitter in real-time communications [A]. Proceedings of ACM SIGCOMM '93 [C]. San Francisco, California: ACM Press, 1993. 13 - 19.
- [15] Heffes H, Lucantoni D. A markov modulated characterization of packetized voice and data traffic and related statistical multiplexor performance [J]. IEEE Journal on Selected Area in Communications, 1986, 4(6): 68 - 76.
- [16] Tia, TSB-116-A, Telecommunications IP telephony equipment voice quality recommendations for IP telephony [M]. Virginia: Telecommunications Industry Association Press, 2006.

作者简介:



邱 菡 女, 1981 年 9 月出生于湖北随州。现为解放军信息工程大学通信与信息系统专业博士研究生。主要研究方向为宽带信息网络, 流媒体技术。

Email: qiuhan.ndsc@yahoo.com.cn



李玉峰 男, 1976 年 9 月出生于山东烟台。博士, 大学讲师。主要研究方向为路由与交换技术。

Email: liyufeng.ndsc@yahoo.com.cn



邬江兴 男, 1953 年 9 月出生于安徽金寨。博士生导师, 中国工程院院士。现为解放军信息工程大学校长。主要的研究方向为信息网络与交换技术。

Email: wjx@mail.ndsc.com.cn