

WFQ 不是 PGPS 吗?

))) 对/ 一种新的基于 GPS 的分组公平调度器0 一文的评注

齐望东¹, 彭来献², 董 民²

(11 解放军理工大学计算机工程系, 江苏南京 210007; 21 解放军理工大学电信工程系, 江苏南京 210007)

摘 要: 邬海涛等^[1]声称/ 证明了 WFQ、WF²Q 等算法并不是 P2GPS(基于分组的 GPS), 也就不能保证 P2GPS 的时延及服务特性0. 不幸的是, 该文对 WFQ 系统中一个关键概念/ 阻塞的连接0 的理解是错误的, 因此导致该文的结论无效.

关键词: 调度算法; GPS; WFQ; PGPS

中图分类号: TN915. 05 文献标识码: A 文章编号: 0372-2112 (2003) 12-1882-02

Is WFQ an Algorithm Different from PGPS?

))) Comments on/ A Packet Fair Scheduler Based on GPS0

QI Wangdong¹, PENG LaiXian², DONG Min²

(1. Department of Computer Engineering, Institute of Communications Engineering, Nanjing, Jiangsu 210007, China;

2. Department of Telecommunications Engineering, Institute of Communications Engineering, Nanjing, Jiangsu 210007, China)

Abstract: Wu Haitao et al^[1] claim that they/ have shown the two scheduling algorithms WFQ and WF²Q are not PGPS (packet by packet GPS) so that delay and service characteristics of the latter do not hold in the former two0. Unfortunately, there is a fatal misunderstanding in their paper about an important concept, the set of busy sessions, in the WFQ algorithm, leading to the voidness of their claims.

Key words: scheduling algorithms; GPS; WFQ; PGPS

1 引言

邬海涛等发表在5电子学报6第 30 卷第 4 期(2002 年 4 月)上的/ 一种新的基于 GPS 的分组公平调度器0 一文^[1] (以后简称/ 邬文0)的中心观点是:/ WFQ 和 WF²Q 等算法并不是 P2GPS(基于分组的 GPS), 也就不能保证 P2GPS 的时延及服务特性0. 这个观点可分成两部分来讨论. 一方面, 邬文并没有否认 PGPS 的时延及服务特性. 简单说, 这种特性就是在信源的通信量满足一定条件的情况下(比如漏桶速率控制), 在网络中采用 PGPS 就可以保证该信源的分组时延不超过一个特定的数值. 这个特性是综合服务(Integrated Service)体系结构中确保服务(Guaranteed Service)的理论基础之一. 另一方面, 邬文否认 WFQ 算法是 PGPS 也就是说将 WFQ 用于网络调度并不能提供上述的时延上限保障. 如果邬文的结论是正确的, 那么由于 WFQ 算法的关键地位, 该文得到的就是一个非常重要的发现. WFQ 算法的地位既体现在它被广泛认为是一种可以提供时延上限保障的算法, 也体现在它是很多后来提出的更实用的算法的基础. 可惜的是, 邬文对 WFQ 的理解是错误的. 我们先提供一点背景的情况, 然后再讨论其问题所在.

2 有关术语的精确含义

WFQ(Weighted Fair Queuing)算法的思想在文献[2]中第一次被提出, 但是术语 WFQ 本身并未在该文献中出现. 文献[2]

中提出并详细地分析和仿真了 FQ(Fair Queuing)算法. 在 FQ 算法中各个会话(有人用/ 流0 这个术语, 在面向连接的网络中也称为/ 连接0. 这些术语在本文中含义相同)的权重是相同的. 只是在最后的讨论部分文献[2]才指出这种情况可以推广到各个会话的权重不同的情况. 在文献[3]中, 作者提出了 GPS(Generalized Processor Sharing)和 PGPS(Packet by packet GPS)算法, 并且指出 PGPS 和文献[2]中的 WFQ 算法是同一个算法. 一般文献从此就把 PGPS 和 WFQ 两个术语开始混用(例如文献[4]). 其中 WFQ 这个术语使用得更广泛一些.

值得注意的是, WFQ 这个术语除了指一种特定的算法(这时它就是 PGPS)之外还有其他含义. 有时 WFQ 是指一般的加权公平调度策略, 那么它所涵盖的算法就包括 SCFQ、WF²Q 等等的很多算法. 本文以及邬文的 WFQ 都是指算法而非策略.

我们下来仔细讨论邬文中对术语 WFQ 和 PGPS 的使用方式. 邬文对 PGPS 的定义是文中的定义 2/ 假定在时间 S 以后系统无分组到达, 调度器选择在相应的 GPS 系统里最先完成服务的分组进行服务, 则该调度器为 P2GPS(Packet by Packet GPS). 这个定义和文献[3]中第 3 节对于 PGPS 的说明是一样的. 对于 WFQ, 邬文中最接近定义的陈述是/ 其分组选择策略是 SFI0(最小完成时间标签优先)并且/ 其系统虚时间都按照0 文中的式(3)进行计算. 该式为:

$$\begin{cases} P(0) = 0 \\ P(t_{j-1} + S) = P(t_{j-1}) + \frac{S \cdot @r}{E_{i \in B_j}} \end{cases}$$

$$S[t_j - t_{j-1}, j = 2, 3, \dots]$$

$P(t)$ 称为系统虚时间函数,这里假定在时间段 (t_{j-1}, t_j) 系统内阻塞的连接没有变化,且这些阻塞的连接集合用 B_j 表示。

该定义所描述的算法本质上就是文献[2]中的WFQ,因为当令 $c_i = r_i/r$ 而把这里所用的变量符号 $P(t)$ 换为 $V(t)$ 时上述的公式就是文献[3]中的式(10),而文献[3]中的式(10)和式(11)就是文献[2]提出的WFQ的定义的一种复述。

3 阻塞集与虚阻塞集)))问题的关键

文献[3]声称WFQ和PGPS在分组调度上是等价的,而邹文则认为WFQ不是PGPS。邹文的错误在于对WFQ定义中的式(3)的理解。邹文以为 B_j 是在实际的分组调度系统中阻塞的连接集合,而实际上式(3)中的 B_j 表示的是在GPS系统中阻塞的连接集合。我们注意到,在文献[2]/Fair Queueing0一节的Definition0小节中明确定义了/当 $R(t) \leq F_i^A(i = \text{MAX}(j) | t_j^A \leq t)$ 时,则会话A活动,其中 $R(t)$ 、 F_i^A 和 t_j^A 分别相当于邹文中的 $P(t)$ 、 C_A^A 和 a_A^A ,而术语“活动”相当于邹文中的“阻塞”。这就是说,阻塞会话的定义是依据BR(Bit-by-bit Round robin)得到的,而BR就是GPS在各个会话的权重相同时的版本UPS(Uniform Processor Sharing)。从文献[3]中第3节第A小节的讨论也可以明确得知/ B_j 是在GPS系统中阻塞的连接集合。按照邹文对 B_j 的理解,邹文得到了文中图1所示的虚时间序列和WFQ的分组调度序列。这两个序列都是错误的。邹文根据WFQ分组调度序列所得到的有关WFQ与GPS调度序列差异的反例也是无效的。正确的虚时间序列和WFQ的分组调度序列实际应该如该文的图2所示的情况。为了便于参考,我们把邹文的图1和图2都复制在下面。

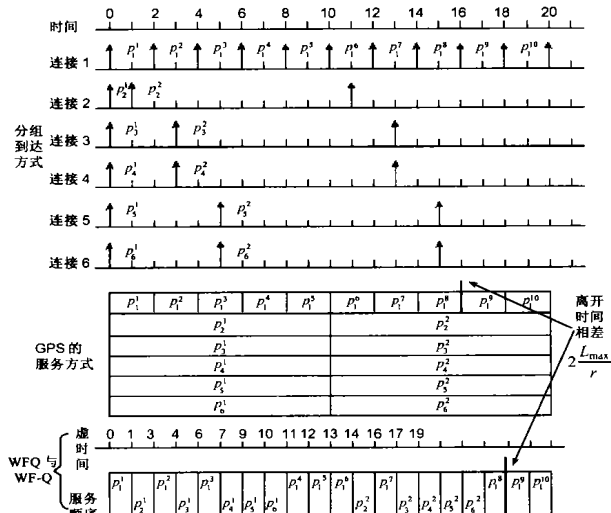


图1 错误的虚时序列WFQ分组调度序列

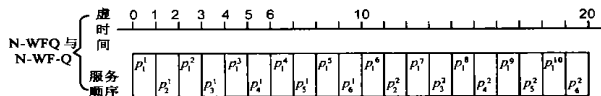


图2 正确的WFQ的调度序列是图中N2WF2Q的服务序列因此,WFQ就等价于PGPS,如同文献[3]所说。

邹文后面部分(第4节/GPS的正确模拟方法0和第5节/算法的实现0)的动机是由于/ WFQ不是PGPS0,因此提出了N2WFQ算法和N2WF2Q算法。这些算法的核心是/虚阻塞集0的概念。其实,由于邹文的虚阻塞集合与GPS的阻塞集合相同(依据邹文引理2证明中的陈述),因此也即文献[3]中的阻塞集合。因此,所谓的N2WFQ即文献[2]中的WFQ,而N2WF2Q算法即文献[5]中的WF2Q。邹文并没有提出新的分组调度器。此外,邹文中关于/虚阻塞集合0(定义4)和/分组系统的阻塞更新事件0(定义5)的思想已在文献[6]中出现过。

另外,邹文关于WF2Q不是PGPS的说法没有问题。PGPS等价于WFQ,当然就不可能等价于WF2Q。

致谢 感谢田畅博士对本文提出的有价值的建议。

参考文献:

- [1] 邹海涛,王重钢,隆克平,程时端.一种新的基于GPS的分组公平调度器[J].电子学报,2002,30(4):460-463.
- [2] A Demers, S Keshav, S Shenker. Analysis and simulation of a fair queueing algorithm[J]. J. Internetworking Res Experience, 1990, 1(10): 3-26.
- [3] A K Parekh, R G Gallager. A generalized processor sharing approach to flow control in integrated service networks: the single node case[J]. IEEE/ACM Trans Networking, 1993, 1(3): 344-357.
- [4] H Zhang. Service disciplines for guaranteed performance service in packet switching networks[J]. Proceedings of the IEEE, 1995, 83(10): 1374-1396.
- [5] J C R Bennett, H Zhang. WF2Q: Worst case fair queueing algorithm[A]. IEEE INFOCOM 96[C]. San Francisco, CA, 1996.
- [6] Ian Marsh. Implementation of Weighted Fair Queuing including RSVP in a BSD kernel[EB/OL]. http://www.sics.se/~ianm/WFQ/wfq_descrp.ps.gz, 1997.10.

作者简介:

齐望东 男,1968年2月生于陕西省西安市,副教授,主要研究方向为计算机网络。



彭来献 男,1978年3月生于安徽省阜阳市,博士研究生,主要研究方向为宽带网络技术。