Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computer Communications 28 (2005) 1711–1720
computer communications
www.elsevier.com/locate/comcom

Weighted deficit earliest departure first scheduling☆

Teck Meng Lim, Bu Sung Lee*, Chai Kiat Yeo

Centre of Multimedia and Network Technology, School of Computer Engineering, Nanyang Technological University, Nanyang, Singapore

Received 3 December 2004; accepted 3 December 2004
Available online 13 January 2005

Abstract

Deficit Round Robin (DRR) and Weighted Deficit Round Robin schedulers have shown their ability in providing fair and weighted sharing of network resources for network devices. However, they are unable to provide fairness in terms of waiting time of packets among flows. We propose a new packet scheduling discipline, Weighted Deficit Earliest Departure First scheduler. It provides fairness in terms of throughput and differentiation in average and worst-case queuing time among queues that require different rate and delay guarantees. Simulation shows that it achieves near-perfect throughput fairness that is comparable to DRR scheduler and delay differentiation that is comparable to Earliest Departure First scheduler.

Keywords: DRR; EDF; Quality of service; Scheduling

1. Introduction

Link scheduling in packet network is an important mechanism to achieve Quality-of-Service (QoS) as it directly controls packet delays. Existing QoS architectures like Integrated Services (IntServ) [1] and Differentiated Services (DiffServ) [2] rely on link scheduling to provide the differentiated bandwidth fairness and delay among queues on each router. Many schedulers have been proposed to address the issues.

The Weighted Round Robin (WRR) [3] scheduler is the pioneer that shows differentiated fairness among queues. Packets, from different flows, are queued in separate queues and the scheduler polls each queue in a cyclic manner in proportion to a weight pre-assigned to each queue. Each poll serves a packet from any non-empty queue encountered. Flows are protected as a misbehaving flow only overflows its own queue. Weighted Deficit Round Robin (WDRR) [4] performs well when all packets have the same size. The Deficit Round Robin (DRR) [5] scheduler is a modification of the WRR scheduler that handles variable packet sizes without knowing the mean packet size of each flow in advance. The DRR scheduler provides near-perfect throughput fairness and flows isolation at low implementation cost. Each queue is allocated a fixed quantum and is serviced in a round robin manner. The allocated quantum is reduced by the length of each packet sent. This deduction stops at the packet that will result in a deficit of quantum. The final balance of quantum is preserved and added to the quantum to be assigned in the next round if the queue is not emptied during the current round. This preservation method ensures fairness and allows packets with length greater than the quantum to be sent. The WDRR scheduler is a variation of the DRR scheduler to provide unfair bandwidth allocation to different queues in the scheduler. Different queues are allocated a different quantum value using a proportional weighted function. These schedulers do not guarantee any delay differentiation.

Weighted Fair Queuing (WFQ) [6,7] is a distinctive scheduler that provides a good emulation of the Generalised Processor Sharing (GPS) [6]. It guarantees both

☆ This work is supported by Singapore Advanced Research and Education Network (SingAREN) Broadband Grant. (Grant number: ICT/00/013/08).
* Corresponding author.
E-mail addresses: teckmeng@ieee.org (T.M. Lim), ebslee@ntu.edu.sg (B.S. Lee), asckyeo@ntu.edu.sg (C.K. Yeo).

0140-3664/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2004.12.015

differentiated fairness and delay among its queues. It is comparable to the WDRR scheduler in terms of weight sharing ability but it has higher algorithmic complexity than the WDRR scheduler. The Earliest Departure First (EDF) [8] scheduler is one of the best-known algorithms for Real-Time processing that has been applied to solve packet-scheduling problems. Although the EDF scheduler does not assure any differential fairness, it is an optimal dynamic algorithm that assigns highest priority to the active packet with the most imminent deadline. The assignment is dynamic because packet priority changes as time passes.

This paper introduces the Weighted Deficit Earliest Departure First (WDEDF) [9] scheduler which aims to maintain the weighted share of bandwidth among queues while reducing the queuing delay of packets, thus, reducing the number of deadline misses in the delivery of packets. The remainder of the paper is structured as follows. The Section 2 discusses the basic operation details and algorithm of the WDEDF scheduler. Section 3 discusses some analytical results of the WDEDF scheduler. Section 4 evaluates the performance of WDEDF scheduler in comparison to the DRR, WDRR, WFQ and EDF schedulers. Finally, Section 5 concludes the WDEDF scheduler.

2. Weighted deficit earliest departure first scheduler

In the WDEDF scheduler, each queue is allocated a weighted quantum using a proportional weighted function like the DRR scheduler except that the WDEDF scheduler does not restrict the setting of minimum quantum size to be the Maximum Transmission Unit (MTU) [10] of a router.

Each queue i reserves a bandwidth of rate R_i on the scheduler. Based on the reservation rates, each queue is allocated ω_i worth of bits in each service round. Let us define ω^{MIN} as the minimum quantum size of WDEDF and $R^{\mathrm{MIN}} = \min(R_j)$ where $\mathbb{R}$ is the set of reservation rates and $R_j \in \mathbb{R}$. The quantum allocated to queue i is simply

$$\omega_i = \frac{R_i}{R^{\mathrm{MIN}}} \times \omega^{\mathrm{MIN}}, \qquad (1)$$

with $\omega^{\mathrm{MIN}} \geq \mathrm{MTU}$. This quantum allocation method ensures that at least one packet from each queue can be transmitted in a service round.

A deficit counter (DC) that increases by ω_i at the start of each service round is kept for each queue. The WDEDF scheduler decreases the DC_i by the size of the packet forwarded for queue i. A queue is blocked for service in the current service round when it exceeds its allocation ω by forwarding its next packet, i.e. the WDEDF scheduler blocks a queue when its DC value is insufficient to allow the forwarding of its next packet.

The WDEDF scheduler works in service rounds. A service round ends when all queues are either emptied or blocked. Empty queues are set as inactive and their DCs set to zero. The final balance of the DCs of blocked queues are preserved and summed with the quantum value to be added in the next round. All blocked queues are set as active queues at the start of a new service round. An inactive queue becomes active again with its DC set to its allocated ω when it has a new packet.

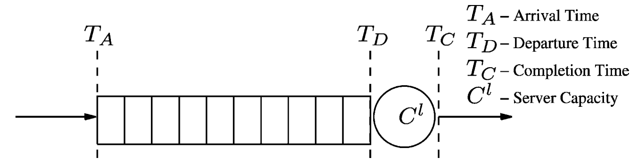

Fig. 1. General Queue timing diagram.

The DRR scheduler services packets in a queue until the DC of the queue is insufficient for forwarding the Head-Of-Line (HOL) packet. This criterion can implicitly minimise the delay of a packet from the queue in service at the expense of increasing the delay of other queues waiting to be served. An example of this phenomenon is illustrated in Example 1.

The WDEDF scheduler improves on the DRR scheduler to provide a lower queuing delay by forwarding the packet at the head of the queue with the earliest departure time. The departure time (T_{D}) (see Fig. 1) is the time that a packet needs to leave the scheduler.

$\widehat{T_{\mathrm{D}}}$ can be computed using Eq. (2), where T_{A} is the arrival time of a packet, L is the packet size, R is the reservation rate of the packet and C^l is the service rate of the server, can be used to compute the expected departure time, $\widehat{T_{\mathrm{D}}}$, of each arrived packet.

$$\widehat{T_{\mathrm{D}}} = T_{\mathrm{A}} + \frac{L}{R} - \frac{L}{C^l} \qquad (2)$$

Example 1. There are two non-empty queues (Queue A and B) set up on a router with a 10 Mbps link. A_i represents the ith packet of Queue A; likewise for Queue B. Queue A is reserved 2 Mbps and Queue B is reserved 3 Mbps on the router. Packet A_1 arrives at the router at time (t_{A}) µs. B_1 arrives at time ($t_{\mathrm{A}}+100$) µs. A_2 arrives at time ($t_{\mathrm{A}}+160$) µs. The departure times of the packets are tabulated in Table 1 and queues' quantum are set to 1000 bits. At time $t=(t_{\mathrm{A}}+162)$ µs. The router starts to

Table 1
Packet departure times used in Example 1

Packet	L (bits)	Queuing time (µs)	Service time (µs)	T_{D} (µs)
A_1	400	200	40	$t_{\mathrm{A}}+160$
B_1	500	166	50	$t_{\mathrm{A}}+216$ ($t_{\mathrm{A}}+100+116$)
A_2	400	200	40	$t_{\mathrm{A}}+320$ ($t_{\mathrm{A}}+160+160$)