Available online at www.sciencedirect.com
SCIENCE DIRECT
Computer Communications 28 (2005) 929–946
computer communications
www.elsevier.com/locate/comcom

Three new approaches for adjustment and improvement of the RR scheduler in a dynamic resource environment

John Tsiligaridis[a,*], Raj Acharya[b]

[a]Computer Science and Engineering, 236 Bell Hall, Buffalo, NY 14260 SUNY at Buffalo, USA
[b]Computer Science and Engineering, 220 Pond Laboratory, Penn. State University, University Park, PA 16802-6106, USA

Received 26 July 2003; revised 12 January 2005; accepted 12 January 2005
Available online 19 February 2005

Abstract

Many methods of packets' service have been developed, such as the Deficit Round Robin (DRR) and the Surplus Round Robin (SRR), that have to do with a borrowing of credits technique for use in the next rounds. In the present work, we propose a new variable flexible credit scheme, for *adjustment of the weights*, instead of a fixed credit scheme with a lower maximum weight. The service of various short duration flows with large size packets (impossible to be serviced in a round), is achieved by using gradual or direct weight increase mechanisms. A *feedback mechanism* is used for adjusting the weight size taking into account the parameter of the current packet size. This constitutes an alternative way of defining the *most suitable size of the quantum* versus the traditional one of the weight's computation (like SLA). The advantage of our best weight adjusting method lies on both the restriction of the otherwise required, repeated, unexpected and uncertain changes of the next round's credits (as it happens with the uninformed for the large packets arrival Deficit Round Robin) and the guarantee of the next packet's service with no extra overhead ('empty round'). With this ability, the DRR has again O(1) complexity, *without making any assumption* for the maximum size of the packets.

A family of algorithms, with O(1) complexity, that work *on the fly*, using different and *systematic* weight request methods are: the Direct Increasing Weight (DIW), the Compound Round Robin (CRR) and the Proportional Increasing Weight (PIW). In the case of large packets sequence's arrival, our algorithms outclasses the low static quantum DRR. Upgrading the DRR using the PIW, is a different approach that outperforms the DRR (fixed weight).

Their relative fairness along with the delay bounds are derived. Simulation experiments support the algorithms' significance and useful results are also provided in comparison with other scheduling algorithms. This kind of algorithms can be applied for next generation internet routers, achieving short time service for delay sensitive flows.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Scheduling; DRR; SRR; ERR

1. Introduction

As far as the rate servers are concerned, keeping the fairness in the allocation of resources among the users, is essential. In most computer networks, no firewalls exist and these networks are subject to the badly behaved sources. A systematic approach of the *weight sufficiency problem* is much more preferable than the continuous and endless change of the credits' amount in every round as it happens in the Deficit Round Robin (DRR). It is interesting to develop an *adaptively adjustment credit scheme* instead of a fixed weight scheme.

Unfortunately, many fair scheduling disciplines cannot be implemented in high-speed switches due to their time complexity. A typical example is the Weighted Fair Queuing (WFQ) algorithm with O(log(n)), where n is the number of flows.

Many algorithms have been developed *making assumptions on the packet's size* before beginning the transmission. More efficient schedulers such as the Deficit Round Robin

* Corresponding author. Address: Department of Mathematics and Computer Science, Heritage University, 3240 Fort Road, Toppenish, WA 98948, USA. Tel.: +1 509 865 8500, ext. 3802; fax: +1 509 865 8679.
E-mail addresses: tsiligaridis_j@heritage.edu (J. Tsiligaridis), jt32@acsu.buffalo.edu (J. Tsiligaridis).

0140-3664/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2005.01.011

Table 1
Performance of queuing algorithms

Type of scheduler	Relative fairness (RF)	Complexity
DRR	3*M*	O(1)
SRR	3*M*	O(1)
ERR	3*M*	O(1)
DIW	6*M*	O(1)
CRR	3*M*	O(1)
PIW	3*M*	O(1)

(DRR) and Surplus Round Robin (SRR) operate, provided that the maximum size of the packet is *known*, and the O(1) complexity is achieved. It has to be mentioned that the SRR uses the method of borrowing credits from the next round.

In the following, we provide some definitions and notes necessary for our analysis.

Definition 1. *A large packet* (*l-packet*), unlike the normal or small size, cannot fit to the size of the weight and be transmitted in a round.

To sustain a certain data rate, the scheduler has to complete a packet scheduling cycle within a certain time limit. The *l*-packet allows a higher time budget for a packet scheduler because the less size of the quantum *makes more cycles* for transmission.

Definition 2. An *empty round* appears, when the scheduler cannot service a packet of a flow, due to the fact that the amount of the previous remainder of bandwidth plus the weight is less than the size of the packet.

Because of the empty rounds, we can have unlimited borrowing of quantum for the following periods. This work has a bad influence on the end-to-end latency, on the latency of the other queues and on the overall performance of the scheduler (Table 1). The *importance of the empty round* is examined at the following. Under DRR policy, a flow with n *l*-packets needs $2n$ rounds to be serviced because for each *l*-packet, one round is needed for search and another one for the quantum increase. Thus, for a sequence with n *l*-packets, $2n$ rounds are needed. The loss is exactly $2n-n=n$ rounds. This amount is a significant waste of resources, especially when the flow is delay-sensitive.

Definition 3. *Quantum* is the *fixed* service defined by contract that the flow should receive during each round robin service opportunity.

Definition 4. *Weight* is the *variable, adjustable* amount of service that a flow receives in a round. Finally, after its appropriate increase, the weight's size is called quantum. It must be proportional to the packet's size.

Definition 5. *MSLP* stands for the maximum size of the *l*-packets, a parameter very important for the weight's size.

It is evident that normally *a set of various size packets* can be serviced by a quantum in a round. When this size exceeds a threshold, then an empty round may occur and an increase of the weight is necessary. *Quite often, a fixed-credit scheme using a lower maximum (for some flows), does not perform as well as the flexible variable credit scheme*. In [3], the DRR clearly states that the quantum for each flow must be greater than the largest possible packet size. In some cases (*l*-packets arrival), the fixed value of the DRR weight is not sufficient to provide O(1) complexity.

The probable reasons that make *the instantaneous increase of the weight* of a flow a necessity are:

- *The MTU factor*. Let us consider the TCP protocol. The sending and receiving TCP entities exchange data in the form of segments [13]. Each network has a maximum transfer unit (MTU) within which each segment must fit. If a segment passes through a sequence of networks without being fragmented, and hits one whose MTU is smaller than the segment, then the router at the boundary fragments the segment into two or more smaller segments. If a segment is too large for a network and has to be transferred, then it must be broken up into multiple segments by a router. Each new segment gets its own IP header. The fragmentation by the routers increases the total overhead, and delay, because each additional segment adds 20 bytes of extra header information to the form of an IP header [13]. An algorithm that can adjust and increase instantaneously the weight, is considered as necessary, so that the fragmentation, due to the probable short size of the MTU to be avoided.
- *The SLA factor*. In general situation, the contract for each flow such as the SLA (service level agreement) finally determines the necessary weight. The bad behavior of a flow is usually the result of the internal conditions of the network and not of the particular source's behavior. Thus, the core routers may benefit from an increase of the value of the weight for limited time.
- *The Latency factor*. The instant increase of the weights can be considered mandatory because the service delay of some flows can cause further delay to all the other flows as well.
- *New services (or applications) factor*. For any new service (like video on demand), an increase of the flow's weight is probably needed, so that the empty rounds are avoided. All the routers along a delivery chain must be able to support the desired type of QoS. Any router that cannot be a potential entrance point for the latency (resulting jitter that breaks the time-domain needs of applications), may cause the multimedia reproduction to suffer.

It is assumed that the maximum size of the packets is unknown for the DRR [9]. Additionally, in [14] it has clearly been stated that in order to get an O(1) complexity, the quantum has to be greater than the largest size of the packet, that may have arrived during the execution of the scheduler. *The problem arises when the l-packets of*

Download English Version:

https://daneshyari.com/en/article/10338528

Download Persian Version:

https://daneshyari.com/article/10338528

Daneshyari.com