Computer Communications 32 (2009) 357–367

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

LGRR: A new packet scheduling algorithm for differentiated services packet-switched networks

A. Ghaffar Pour Rahbar [a,*], Oliver Yang [b]

[a] Computer Networks Research Lab, Department of Electrical Engineering, Sahand University of Technology, Sahand New Town, Tabriz, Iran
[b] CCNR Lab, School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada K1N 6N5

ARTICLE INFO

Article history:
Received 21 August 2007
Received in revised form 2 November 2008
Accepted 7 November 2008
Available online 18 November 2008

Keywords:
Class-based packet scheduling
Buffer management
Round Robin
Bandwidth loan

ABSTRACT

Since Quality of Service (QoS) support is a mandatory requirement in the next-generation networking, each router in a packet-switched network must provide a better service to higher-priority packets under any situation such as congestion. We propose in this paper the loan-grant based Round Robin (LGRR) packet scheduler for use in each output port of a router in a DiffServ network. LGRR is a frame-based scheduler to pass traffic streams according to their class types and to their immediate upstream source routers. It uses a loan-grant scheme so that a higher priority traffic stream can be processed quickly by requesting a bandwidth loan from the scheduler. To control the amount of transmitted bits from each stream and to prevent malicious abuse, the bandwidth loan must be paid back from the quantum values acquired in future. LGRR gives a fair opportunity to different traffic streams to access to the network bandwidth. It performs better than MDRR+, MDRR++, and OCGRR in handling traffic under both normal and bursty traffic, but it also gives a better loss and delay performance to the higher-priority traffic when traffic load is very high.

1. Introduction

Under Differentiated Services (DiffServ) [1], edge routers are in charge of classifying, marking, dropping, and shaping IP packets, while core routers perform high speed routing of these packets that have been classified as Expedited Forwarding (EF) [2], Assured Forwarding (AF) [3] and Best Effort (BE). In general, EF traffic needs low loss, low latency, low jitter, and assured bandwidth. AF traffic requires a guaranteed forwarding, and BE traffic has no service guarantee.

The common approach to support the DiffServ traffic in an output port of a router is to use a shared First-Come-First Served (FCFS) buffer for each class, e.g., [8–13]. Here, the same-class packets from different sources are all saved in a shared buffer. Although using a shared buffer for all traffic streams buffer management would be easier [6], it is difficult to control the service order of packets from different sources because a malicious source in a class may cause a higher delay and even loss for well-behaved traffic streams within that class.

Scheduling techniques such as Deficit Round Robin (DRR+) [7], Priority Queuing (PQ) [23], Token Bank Fair Queuing (TBFQ) [4], Weighted Round Robin (WRR) [5], PQWRR [11], Dynamic WRR, e.g. [22], Output-Controlled Grant based Round Robin (OCGRR) [19], and DRR++ [14,15] can be used in a class-based domain. They all support variable-length packets, except those originally designed for fixed-size packets (e.g., TBFQ and WRR). DRR++ suffers from head-of-line blocking when scheduling more than one latency-critical traffic stream. This problem can be resolved by increasing the service time complexity to $O(n)$. Other schemes have a fairness problem. PQ is unfair to the lower-priority traffic. PQWRR is unfair to AF and BE by using PQ for the EF traffic, while WRR is unfair to the AF and BE traffic. DRR+ has a tendency to generate bursty output when serving a traffic stream, thus leading to a higher jitter and startup latency. All algorithms that use DRR as their basic method, e.g. [16,17], must know not only the maximum packet length in order to achieve the per-packet work complexity [7] of $O(1)$, but also the packet size in order to schedule each packet at the head of a queue.

Common approach to deal with the bursty traffic for bursty traffic streams is to borrow bandwidth from the available bandwidth, e.g., [4,18]. However, this cannot guarantee the desired QoS for a higher priority traffic stream when traffic arrival to an output port of a router is higher than the available bandwidth. As mentioned, the techniques in [4,18] are appropriate for fixed-size packets and not for variable-sized packets. Note that traffic burstiness results from the common pattern observed in the Internet traffic [21].

OCGRR has a nice feature of reducing the inter-transmission time from the same traffic stream and achieving a smaller jitter and startup latency. However, it may not be able to serve higher-priority traffic well under router congestion arising from bursty traffic and/or high network traffic load. This is because traffic

* Corresponding author. Tel.: +98 411 5263374.
E-mail addresses: ghaffarpour@sut.ac.ir (A. Ghaffar Pour Rahbar), yang@site.uottawa.ca (O. Yang).

0140-3664/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2008.11.007

streams in OCGRR gain their bandwidth quantum at a rate proportional to the average traffic arrival rate, but not proportional to traffic fluctuations. So there is a need to extend OCGRR to serve higher-priority traffic well even under router congestion.

The objective of this paper is to design the LGRR (loan-grant based Round Robin) algorithm that can serve high-priority traffic well. LGRR has extended and improved OCGRR so that higher-priority classes can be served quickly. This is achieved by giving more opportunity to higher priority traffic to obtain the necessary bandwidth when traffic is bursty and/or when the traffic load is high. Even if there is no available bandwidth, LGRR would allow higher-priority traffic streams to take out "bandwidth loan" to pass their traffic as soon as possible. This is at the expense of momentarily degrading lower-priority traffic streams. This is usually not a problem because lower-priority traffic can experience higher delay and even loss. Similar to OCGRR, the LGRR algorithm schedules variable-size packets in a multiple-class domain using a packet-by-packet scheme from the traffic streams inside each class. This smooth scheduling can improve the performance of downstream switches or routers, in terms of lower traffic burstiness/congestion and less buffer requirement. Like OCGRR, we may isolate traffic streams from each other in each class to combat the behavior of a malicious stream. Through performance evaluation, we shall demonstrate that traffic streams enjoy shorter queuing delay and loss inside a LGRR-enabled router.

The contribution of this paper is to extend OCGRR by using a bandwidth loan policy to manipulate traffic burstiness, and to allow a stream to loan some bandwidth from the scheduler to quickly pass its bursts. The extension allows LGRR to efficiently serve higher-priority classes under congested router, while keeping the OCGRR capabilities such as smooth output, lower jitter, and lower latency.

This paper is organized as follows. Section 2 provides the network operations and modeling. Section 3 describes the loan-grant concept for traffic scheduling. The LGRR scheduler is discussed in Section 4 followed by performance evaluation in Section 5. Section 6 gives a brief conclusion. For the remainder of this paper, the following general symbols and notations pertain.

$B_{J\ i}$ the size of buffer allocated to stream i in class J, $J = 1(\text{EF}),\ldots,M\ (\text{BE})$.
B_w output bandwidth
C_p coherence parameter
C_J index of class J, $J = 1(\text{EF}),\ldots, M(\text{BE})$.
$G_{J,i}$ available grant for stream i in class J, $J = 1(\text{EF}),\ldots, M\ (\text{BE})$.
K grant intensity allocation parameter used in LGRR scheduling
L traffic load
L_{max} largest packet size in system
M number of traffic classes generated from each one of the source routers
N_{PJ} partition size for a buffer in class J
$Q_{J,j}$ quantum for stream i in class J, $J = 1(\text{EF}),\ldots, M(\text{BE})$.
R number of input source routers connected immediately to the core router
T_f expected frame time period
$U_{J,i}$ used-grant for stream i in class J, $J = 1(\text{EF}),\ldots, M\ (\text{BE})$.
$q_{J,i}$ the size of traffic available in the buffer allocated to stream i in class J, $J = 1(\text{EF}),\ldots, M\ (\text{BE})$.
t virtual time that is equal for all streams and classes
Γ expected Frame Length
α_r traffic loan installment coefficient
β deduced bandwidth loan installment
γ maximum bandwidth loan coefficient
λ_J average arrival rate of class J traffic streams, $J = 1(\text{EF}),\ldots, M\ (\text{BE})$.
$\lambda_{J,i}$ average arrival rate of stream i in class J, $J = 1(\text{EF}),\ldots, M\ (\text{BE})$.
λ_{min} the smallest AAR among all streams of all classes
ρ average packet length
$£_{J,i}$ outstanding bandwidth loan for stream i in Class J, $J = 1(\text{EF}),\ldots, M\ (\text{BE})$.
$£_{J,i,max}$ maximum payable bandwidth loan to stream i in class J, $J = 1(\text{EF}),\ldots, M\ (\text{BE})$.

2. Network operations and modeling

We consider a DiffServ core router (hereafter referred to as core) in a packet-switched network, which is physically connected to R immediate upstream source routers (R is usually fixed for a given network topology). A source router is referred to as a *source*. The bandwidth of each output port of the core router is B_w. Packets in the network are classified into M classes. So there are M classes of traffic from each source and they can be found in each output port of the core router. For $M = 2$ used in this paper, the class-1 and class-2 traffic are called EF and BE, respectively. We define a *stream* to be the traffic of the same-class packets from a source routed to an output port of the core router. So, there are only $R \times M$ streams arriving at the core router destined to leave the router from the same output port.

An LGRR scheduler resides in each output port (Fig. 1) of the core router that is responsible for bandwidth sharing and transmitting order of packets from the streams. The scheduler system consists of some sorting buffers, and a traffic monitoring unit to provide the information on the arrival rate of streams.

For each output port, class-based sorting buffers are used to hold different classes of traffic. To fairly process streams, a dedicated buffer is allocated to each stream in each class. Let $B_{J,i}$ be the size of the buffer allocated to stream i in class J. It can also be argued that a shared buffer may also be used for all traffic streams within each class so that the packet drop rate due to the burstiness of the traffic can be minimized if desired. Here, overflow packets from buffer i dedicated to stream i within class J will be redirected to the shared buffer of class J. Network manager may allocate 1, R, or $R + 1$ buffers to traffic classes. At the output port

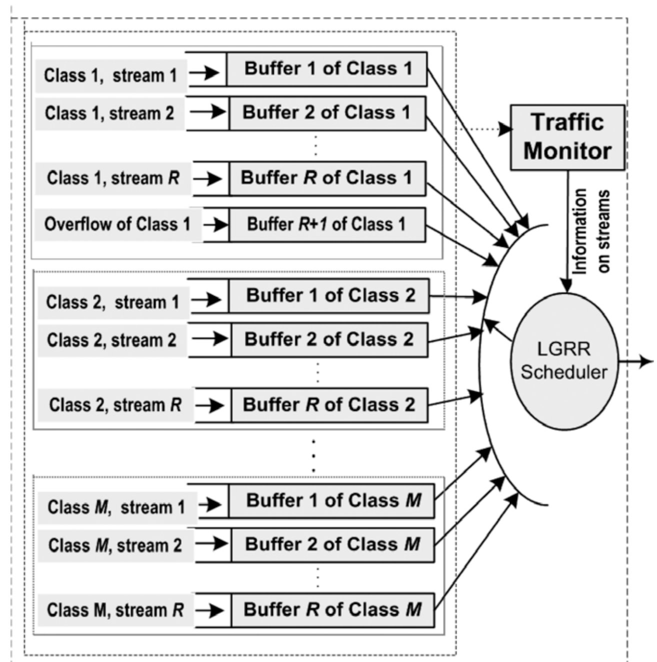

Fig. 1. The model of one output port in a core router.