More Precise Fairness Bounds of Deficit Round Robin Scheduler

Kos, Anton; Tomažič, Sašo

***Abstract*— Fairness is a much desired property of scheduling algorithms. In a fair scheduler each data flow gets its fair share of the available bandwidth, and this share is not affected by the presence of other flows and their possible misbehaviour. Deficit Round Robin (DRR) scheduler is easy to implement, it provides fairness for flows with variable packet lengths, it allows bandwidth reservation, and it has *O*(1) complexity. The fairness of schedulers has frequently been evaluated using one of the two fairness measures: the absolute fairness measure (AFM) and the relative fairness measure (RFM). In this paper we carry out a thorough fairness analysis of DRR and we derive more precise bounds of both RFM and AFM. We show that our new fairness bounds are mathematically more accurate and that they give tighter approximation of the worst-case fairness behaviour of DRR scheduler, than the bounds derived earlier.**

Index Terms—DRR, Deficit Round Robin, Fairness, Fairness Measure, Packet Networks, Scheduling

1. Introduction

In packet networks with statistical multiplexing (like internet) overload causes congestion that is solved either by delaying or by dropping excess packets. There are different solutions that try to solve a challenge of assuring high resource utilization and high application performance at the same time. Essentially they can be grouped into two categories: end-system based solutions and router based solutions. In this paper we analyse router based solutions.

Many multimedia applications rely on the ability of the network to provide some sort of quality of service guarantees. The term Quality of Service (QoS) can generally be defined as a set of network mechanisms that satisfy the varied quality of service levels required by applications, while at the same time maximizing bandwidth utilization. Applications rely on traffic scheduling algorithms in switches and routers to guarantee performance bounds and meet the agreed QoS. There are several measures that are to be considered when choosing a scheduling algorithm. The most important are: fairness, latency and complexity.

Manuscript received May 4, 2009.
Kos, Anton is with the Faculty of Electrical Engineering, University of Ljubljana, Slovenia (e-mail: anton.kos@fe.uni-lj.si).
Tomažič, Sašo is with the Faculty of Electrical Engineering, University of Ljubljana, Slovenia (e-mail: saso.tomazic@fe.uni-lj.si).

The paper is organised as follows. In section 2 we briefly describe why fairness is an important property of each scheduler. In section 3 we give the advantages of our fairness measures over the existing ones. We continue with the explanation of the basics of Deficit Round Robin scheduler in section 4 and the definition of fairness measures in section 5. Sections 6 and 7 are the core of this paper where we derive the new relative and absolute fairness measures for the Deficit Round Robin scheduler. We conclude with the comparison of fairness measures of some well known schedulers in section 8.

2. Importance of Fairness

Packets belonging to different flows often share links in their transmission paths. Fairness is a very desirable property in the allocation of bandwidth on such links. In multiuser/multiapplication environments the protection guaranteed by fair scheduling improves the isolation between flows. Isolation offers more predictable performance of the system to users applications. Fair allocation of bandwidth ensures that the performance of one flow is not affected when another, possibly misbehaving flow, tries to send packets faster than its reserved rate. In addition, strategies and algorithms for fair management of network traffic can serve as a critical component of QoS mechanisms to achieve certain guaranteed services such as delay bounds and minimum bandwidths.

Since a scheduling algorithm should always provide the best possible QoS it has to be fair, it has a bounded maximum delay limit, low computational cost (complexity), easy implementation, and high efficiency.

3. Our Fairness Measures

Fairness bounds of DRR have been given or derived in many papers that discuss Latency-Rate or Round Robin like schedulers. Let us mention just the most important papers in this topic written by Shreedhar and Varghese [3] and Stiliadis and

Varma [7]. Their fairness bounds seem very similar to our bounds. Unfortunately they are inaccurate as they have made a slight mistake in presumption about the maximal size of a deficit counter and they have not taken into account the discrete nature of packet size. The detailed explanation of these differences is given in [1] and [2].

In sections 6 and 7 we give the simplified derivation of our fairness bounds. The exhaustive derivation with proofs can be found in [1] and [2]. Further comments about the differences in fairness bounds are given within the section 6.

The most important result of our paper is the fact that our fairness bounds for DRR are better than the ones calculated in [3] and [7].

4. Deficit Round Robin

Scheduling algorithms can be broadly classified into two categories: sorted priority schedulers and frame-based schedulers. Shreedhar and Varghese proposed one of the most popular frame-based scheduling algorithms in 1996 [3]. The main characteristic of all Deficit Round Robin (DRR) like scheduling algorithms is their ability to provide guaranteed service rates for each flow (queue). DRR services flows in a strict round-robin order. It has low complexity and it is easy to implement. Its latency is comparable to other frame-based schedulers. Below is the list of variables used in our analysis:

R	transmission rate of an output link,
N	the total number of active flows,
r_i	the reserved rate of flow i,
w_i	weight assigned to each flow i,
Q_i	quantum assigned to flow i,
DC_i	Deficit Counter of flow i,
F	frame size,
L_{max}	the maximum possible packet size.

A better fairness property of DRR comparing to Round Robin (RR) is achieved by maintaining a variable called *Deficit Counter* for each of the flows. It is denoted with DC_i, for flow i, and it stores the value of the deficit that a flow has accumulated in its active period (a flow is active during any period of time when packets belonging to that flow are continuously queued in the system). DRR assigns a *quantum* of service Q_i to each flow i. When a flow is on its turn for service, its DC_i is incremented by the quantum of that flow. A flow is served only if the packet at the head of the flow i is smaller or equal than the current value of DC_i. Otherwise the scheduler begins servicing the next flow in the round robin sequence. Flow i is being serviced until packets at the head of the flow remain smaller that the current value of DC_i. When a packet is removed from the flow i, the DC_i is decremented for the length of the packet. A detailed operation of DRR algorithm can be found in [3].

Because all flows share the same output link, a necessary constraint is that the sum of all reserved rates must be less or equal to the transmission rate of the output link:

$$\sum_i r_i \leq R \tag{1}$$

Let r_{min} be the smallest of r_i: $r_{min} = \min_{\forall i} r_i$. Each flow i is assigned a weight that is given by:

$$w_i = \frac{r_i}{r_{min}}. \tag{2}$$

Note that $\forall i \in 1, 2, \cdots, N$ holds $w_i \geq 1$. Each flow i is assigned a quantum of Q_i bits, that is a whole positive value, i.e. $Q_i \in \mathcal{N}$. This quantum is actually the amount of service that the flow should receive during each round robin service opportunity. Let us define with Q_{min} the minimum of all the quanta. Then the quantum for each flow i is expressed as:

$$Q_i = w_i Q_{min}. \tag{3}$$

5. Fairness

When there is contention for resources, it is important for resources to be allocated fairly. Among scheduling algorithms significant discrepancies may exist in service provided to different flows over the short term. For example, two scheduling algorithms may have the same delay guarantees but can have very different fairness behaviours.

There is no commonly accepted method for estimating the fairness of a scheduling algorithm. In general, we would like the system to always serve flows proportional to their reserved rate and distribute the unused bandwidth left behind by idle flows proportionally among active ones. In addition, flows should not be penalized for excess bandwidth they received while other flows were idle.

Based on this intuitive definition of fairness the Generalized Processor Sharing (GPS) [5] is identified as an ideal resource sharing discipline. GPS is defined with respect to a fluid-model, where packets are considered to be infinitely divisible. The share of bandwidth reserved by flow i is represented by a real number r_i. Let us assume that there is N active flows in the system during the time interval (t_1, t_2) and let us with $SENT_i(t_1, t_2)$ denote the amount of service received by flow i during the same time interval. According to [7] the next inequality holds:

$$SENT_i(t_1, t_2) \geq \frac{r_i}{\sum_{j=1}^{N} r_j} R(t_2 - t_1) \tag{4}$$

where R is the link rate. The minimum service that a flow can receive in any interval of time is:

$$\frac{r_i}{\sum_{j=1}^{V} r_j} R(t_2 - t_1). \tag{5}$$

where V is the maximum number of active flows in the server at the same time. It is also true that $N \leq V$ and

$$\sum_{j=1}^{V} r_j \leq R \tag{6}$$

Thus, at each instant GPS serves each active flow with a maximum rate equal to its reserved rate; in addition, the excess bandwidth available from flows not using their reserved rate is distributed among all others backlogged ones in proportion to their reservations. This results in perfect isolation among all flows and the ideal fairness. However, GPS scheduler is not implementable since in real packet-switched networks data are forwarded in units of packets and not in infinitesimal quantities.

The fairness of scheduling algorithm used in communication networks has frequently been evaluated using two fairness measures. One of them is known as Absolute Fairness Measure (AFM), and it is based on the maximum difference between the service received by a flow under the discipline being measured and the service it would receive under the ideal GPS policy. The AFM of a scheduler is frequently hard to obtain. Therefore an alternative measure, known as the Relative Fairness Measure (RFM) is used. It is based on the maximum difference between the services received by two different flows under the scheduling algorithm being measured. In this paper we will use both of them to give a better perspective of the fairness of DRR.

To give more appropriate definitions of these measures we define some figures of merit. Since arriving packets of different flows are stored in different queues, let us denote with: $SENT_i(m)$ the number of bytes sent out for flow i in round m, and by $SENT_i(p, m)$ the number of bytes sent out for flow i starting from round p to round m. According to [9] we have definitions:

Definition 1: Absolute fairness bound of flow i over time interval (t_1, t_2), is denoted by:

$$AF_i(t_1, t_2) = \left| \frac{SENT_i(t_1, t_2)}{w_i} - \frac{SENT_{gps_i}(t_1, t_2)}{w_{gps_i}} \right| \tag{7}$$

where $SENT_{gps_i}$ denotes the total service received by flow i under GPS scheduling policy and w_{gps_i} is the weight assigned to flow i in GPS scheduling policy.

Definition 2: The absolute fairness, AF, of an algorithm over some interval of time (t_1, t_2) and absolute fairness measure, AFM, of an algorithm are given with:

$$AF(t_1, t_2) = \max_{\forall i} AF_i(t_1, t_2) \tag{8}$$

$$AFM = \max_{\forall (t_1, t_2)} AF(t_1, t_2) \tag{9}$$

Relative fairness measure can be meaningfully defined only with respect to flows that are all backlogged during the time interval under consideration. Following previous work we will assume that all flows under consideration are active at all instants of time during that interval.

Definition 3: The relative fairness with respect to a pair of flows (i, j) over time interval (t_1, t_2), denoted with $RF_{(i,j)}(t_1, t_2)$, is:

$$RF_{(i,j)}(t_1, t_2) = \left| \frac{SENT_i(t_1, t_2)}{w_i} - \frac{SENT_j(t_1, t_2)}{w_j} \right| \tag{10}$$

Definition 4: The relative fairness with respect to flow i over some interval of time (t_1, t_2), denoted with $RF_i(t_1, t_2)$, is:

$$RF_i(t_1, t_2) = \max_{\forall j} RF_{(i,j)}(t_1, t_2) \tag{11}$$

Definition 5: The relative fairness, RF, of an algorithm over some interval of time (t_1, t_2) and relative fairness measure, RFM, of an algorithm are:

$$RF(t_1, t_2) = \max_{\forall i} RF_i(t_1, t_2) \tag{12}$$

$$RFM = \max_{\forall (t_1, t_2)} RF(t_1, t_2) \tag{13}$$

Definition 6: A scheduling algorithm is **fair** *if AFM and RFM are both bounded by some small constant.*

6. Relative Fairness Measure

We can now proceed with the analysis of a DRR algorithm. For the purpose of the analysis we need lemmas that are proved in [1] and [2].

Lemma 1: For all integers i, the following statement holds:

$$0 \leq DC_i \leq L_{max} - 1 \tag{14}$$

where L_{max} denotes the size of the maximum packet that can arrive.

Lemma 2: Let us consider any execution of the DRR scheme and any interval (t_1, t_2) of that execution, such that flow i is constantly backlogged during given time interval, and let m be the number of round robin service opportunities received by flow i during the time interval in progress. Then the following inequality holds:

$$mQ_i - L_{max} + 1 \leq SENT_i(1, m) \leq mQ_i + L_{max} - 1$$

Lemma 3: A basic invariant of DRR algorithm is that during arbitrary interval in which two queues, i and j, are backlogged, between any two round robin service opportunities given to queue i, queue j must have had a round-robin opportunity.

Theorem 1: For an arbitrary interval in time and in any execution of DRR service discipline the following inequality holds:

$$\begin{aligned} RFM &= \left|\frac{SENT_i(m)}{w_i} - \frac{SENT_j(m')}{w_j}\right| \\ &\leq \frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} + Q_{min} - \frac{1}{w_i} - \frac{1}{w_j} \end{aligned} \tag{15}$$

where m is a number of round robin opportunities given to flow i in the interval (t_1, t_2) and m' is a number of round robin opportunities given to flow j in the same interval.

Proof 1: Let us consider an arbitrary time interval (t_1, t_2) in any execution of DRR algorithm and arbitrary two flows i and j that are backlogged during this interval. According to the Lemma 3 we have:

$$|m - m'| \leq 1 \tag{16}$$

Now according to Lemma 2 we have:

$$\frac{SENT_i(1, m)}{w_i} \leq \frac{mQ_i}{w_i} + \frac{L_{max}}{w_i} - \frac{1}{w_i} \tag{17}$$

$$\frac{SENT_j(1, m')}{w_j} \geq \frac{m'Q_j}{w_j} - \frac{L_{max}}{w_j} + \frac{1}{w_j} \tag{18}$$

We have used only the one side inequalities stated in Lemma 2, for queue i its right hand side and for queue j its left hand side. Since from (3) we have:

$$\frac{SENT_i(1, m)}{w_i} \leq mQ_{min} + \frac{L_{max}}{w_i} - \frac{1}{w_i} \tag{19}$$

$$\frac{SENT_j(1, m')}{w_j} \geq m'Q_{min} - \frac{L_{max}}{w_j} + \frac{1}{w_j} \tag{20}$$

Then it follows

$$\begin{aligned} & \left|\frac{SENT_i(1, m)}{w_i} - \frac{SENT_j(1, m')}{w_j}\right| \\ \leq & \left|mQ_{min} + \frac{L_{max}}{w_i} - \frac{1}{w_i} - m'Q_{min} + \frac{L_{max}}{w_j} - \frac{1}{w_j}\right| \\ \leq & |Q_{min}(m - m')| + \left|\frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} - \frac{1}{w_i} - \frac{1}{w_j}\right| \\ = & Q_{min}|m - m'| + \frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} - \frac{1}{w_i} - \frac{1}{w_j} \end{aligned} \tag{21}$$

In order to justify equality in the last line of expression (21), we have to prove that the expression $\frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} - \frac{1}{w_i} - \frac{1}{w_j}$ is always greater or equal to zero. Reorganizing the expression to:

$$\frac{L_{max} - 1}{w_i} + \frac{L_{max} - 1}{w_j}$$

and since $L_{max} \geq 1$ and w_i, $w_j \geq 1$ we can easily see that our statement holds. Now we get:

$$\begin{aligned} & \left|\frac{SENT_i(1, m)}{w_i} - \frac{SENT_j(1, m')}{w_j}\right| \\ \leq & Q_{min}|m - m'| + \frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} - \frac{1}{w_i} - \frac{1}{w_j} \end{aligned} \tag{22}$$

Using expression (16) and inequality (22) we can conclude:

$$\begin{aligned} & \left|\frac{SENT_i(1, m)}{w_i} - \frac{SENT_j(1, m')}{w_j}\right| \\ \leq & Q_{min} + \frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} - \frac{1}{w_i} - \frac{1}{w_j} \end{aligned} \tag{23}$$

■

Comment 1: Our objective is to find the maximum value of the expression on the right hand side of inequality (23) and along with that the upper bound of RFM. Since our interest is only to find the local maximum, we can proceed by analysing every variable separately. L_{max} is defined by system properties, so we can treat it as a constant.

$$Q_{min} + \frac{L_{max} - 1}{w_i} + \frac{L_{max} - 1}{w_j} \tag{24}$$

So the expression (24) has its maximum when both w_i and w_j take their minimal possible value that is 1.

We can conclude that for a given Q_{min}, the upper bound of RFM occurs when at least two flows have their quanta equal to Q_{min}. Legitimacy of this statement lies in the definition 5 because we are only interested in the worst possible case between any two flows.

Scheduling algorithm	Relative fairness bound
GPS	0
FIFO	∞
PGPS = WFQ	$\max_{i,j}\left(\max\left\{C_j + \frac{L_{max}}{w_i} + \frac{L_j}{w_j}, C_i + \frac{L_{max}}{w_j} + \frac{L_i}{w_i}\right\}\right),$ $C_i = \min\left\{(N-1)\frac{L_{max}}{w_i}, \max_{1\le n\le N}\frac{L_n}{w_n}\right\}$
Virtual Clock Fair Queuing	∞
SCFQ	$\frac{L_i}{w_i} + \frac{L_j}{w_j}$
Worst Case Weighted Fair Queuing	$\frac{L_i}{w_i} + \frac{L_j}{w_j}$
Frame Based Fair Queuing	$\max\left\{\frac{2F-Q_i}{R} + \frac{L_i}{w_i}, \frac{2F-Q_i}{R} + \frac{L_j}{w_j}, \frac{L_{max}}{w_i} + \frac{L_j}{w_j}, \frac{L_{max}}{w_j} + \frac{L_i}{w_i}\right\}$
Packet Based Round Robin	∞
DRR - old	$\frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} + Q_{min}$
Surplus Round Robin	$\frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} + Q_{min}$
Elastic Round Robin	$\frac{L_i}{w_i} + \frac{L_j}{w_j} + L_{max}$
DRR - new	$\frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} + Q_{min} - \frac{1}{w_i} - \frac{1}{w_j}$

TABLE 1

Relative fairness bounds for some well known scheduling algorithms

The RFM bound expression (23) may seem very similar to the RFM bound calculated in [3] and [7]. But it is not. In addition to the fact, that in [3] RFM has not been correctly derived, it should be emphasized that our RFM bound is smaller by some margin than the one calculated in [3] or in any other work to our knowledge. We can conclude that our RFM result is always better than the one calculated in [3] and [7] for the original DRR algorithm.

7. Absolute Fairness Measure

Absolute fairness measure is intuitively closer to us because it shows the difference between the fairness of scheduler being analysed and the fairness of ideal GPS scheduler. We use the relationship between the RFM and AFM derived in [9].

Lemma 4: The relationship between the relative and absolute fairness measures, under any work-conserving scheduling policy, is described with the inequality:

$$AFM \le \left(1 - \frac{Q_{min}}{F}\right) RFM \tag{25}$$

where F denotes frame size, i.e. $F = \sum_i Q_i$.

Using Lemma 4, that was proved in [9], we have that for an arbitrary interval in time and in any execution of DRR service discipline we have:

$$\begin{aligned} AFM \quad \le \quad & \left(1 - \frac{Q_{min}}{F}\right)\left[\frac{L_{max}}{w_i} + \frac{L_{max}}{w_j}\right. \\ + \quad & \left. Q_{min} - \frac{1}{w_i} - \frac{1}{w_j}\right] \end{aligned} \tag{26}$$

When Q_{min} is large the absolute fairness can be low even if the corresponding relative fairness measure is high.

We have already discussed relative fairness bound and its worst and best case, what is left to be analysed is the factor $(1 - \frac{Q_{min}}{F})$. Its minimum value is being reached when $\frac{Q_{min}}{F}$ reaches its maximum value, i.e. when Q_{min} reaches its maximum value, since frame size F is a constant. We have previously concluded that Q_{min} takes its maximum value when all the quanta are equal. In that case AFM is the smallest and the furthest from RFM.

8. Comparison of DRR with Some Other Schedulers

Let us compare fairness bounds of DRR and some of the schedulers from Table 1 that lists fairness bounds for some of the most known scheduling algorithms. Some of them belong to the class of sorted-priority schedulers and others to the class of frame-based schedulers. Thus, this table shows descriptive differences in these algorithms looking through their fairness measure properties.

We observe that the bounding function of PGPS takes higher values than the one just derived for DRR scheduler. The difference between PGPS and DRR is only in factors C_i and $Q_{min} - \frac{1}{w_i} - \frac{1}{w_j}$. If we look closer into the definition of C_i,

$$C_i = \min\left\{(N-1)\frac{L_{max}}{w_i}, \max_{1\leq n\leq N}\frac{L_n}{w_n}\right\}, \qquad (27)$$

we see that since both factors: $(N-1)\frac{L_{max}}{w_i}, \max_{1\leq n\leq N}\frac{L_n}{w_n}$ are bigger than Q_{min}, their minimum is bigger too. It follows that the discussed minimum is also bigger than $Q_{min} - \frac{1}{w_i} - \frac{1}{w_j}$. This leads us to conclusion that DRR algorithm has better fairness property than the PGPS scheduler.

When we compare expressions for SCQF and DRR given in Table 1, we see that they differ in factor $Q_{min} - \frac{1}{w_i} - \frac{1}{w_j}$. If this factor is bigger than zero, the fairness of DRR algorithm is worse, since in original DRR algorithm the minimum value of quanta assigned to flows should be bigger than $L_{max} - 1$ for DRR to have an $O(1)$ complexity. So SCFQ scheduler has better fairness properties than DRR scheduler.

In the class of frame-based schedulers DRR has slightly better fairness bound than the rest of them. The difference is in two subtracting factors, $\frac{1}{w_i}$ and $\frac{1}{w_j}$, which contribute to its better fairness bound.

9. Conclusion

In this paper we have derived a new and improved fairness bounds for DRR scheduling algorithm that contribute to more exact analysis of DRR. We have shown that our fairness bounds are tighter than other previously derived bounds for DRR scheduler. More detailed DRR fairness analysis can be found in [1] and [2].

References

[1] Anton Kos, *Zagotavljanje različnih stopenj kakovosti storitve v omrežjih s paketnim prenosom podatkov*, Doctoral Thesis, Faculty of Electrical Engineering, University of Ljubljana, February 2006

[2] Anton Kos, Jelena Miletić: *Sub-Critical Deficit Round Robin*, Technical Report, Faculty of Electrical Engineering, University of Ljubljana, July 2005

[3] M. Shreedhar, George Varghese: *Efficient Fair Queuing Using Deficit Round Robin*, IEEE/ACM Transactions on Networking, Volume 4, Issue 3, June 1996

[4] J. M. Arco, D. Meziat, B. Alarcos: *Deficit Round Robin Alternated - A New Scheduling Algorithm*, Conference on Protocols For Multimedia Systems - PROMS2000, Cracow, Poland, October 22-25, 2000

[5] Abhay K. Parekh, Robert G. Gallager: *A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-Node Case*, IEEE/ACM Transactions on Networking, Volume 1, Issue 3, June 1993

[6] Chuck Semeria: *Supporting Differentiated Service Classes - Queue Scheduling Disciplines*, White Paper, Juniper Networks, Inc., December 2001

[7] D. Stiliadis, A.Varma: *Latency-Rate Servers - A General Model for Analysis of Traffic Scheduling Algorithms*, IEEE/ACM Transactions on Networking, Volume 6, Issue 5, October 1998

[8] Nihal Pekergin: *Stochastic Bounds on Delays of Fair Queueing Algorithms*, Proceedings of the Conference on Computer Communications (IEEE Infocom), New York, March 1999

[9] Yunkai Zhou, Harish Sethu: *On the Relationship Between Absolute and Relative Fairness Bounds*, IEEE Communications Letters, Vol. 6, No. 1, January 2002

[10] Jian Zhang, Timo Hamalainen, Jyrki Joutsensalo: *Stochastic Analysis of Upper Delay Bound of GPS-Based Packetized Fair Queuing Algorithms*, Dept. of Mathematical Information Technology, University of Jyvaskyla, Finland

[11] S. Golestani: *A self-clocked fair queueing scheme for broadband applications*, Proc. IEEE INFOCOM, Toronto, ON, Canada, June 1994