Proceedings of ITC-CSCC 2000, Pusan, Korea

A Simple Scheduling Algorithm Supporting Various Traffics in ATM Networks

Jae-Jeong Shim, Jae-Young Pyun, and Sung-Jea Ko
Multimedia Communications Lab., Dept. of Electronics Engineering, Korea University
1, 5-Ga, Anam-Dong, Sungbuk-Gu, Seoul 136-701, Korea
Tel:+82-2-929-2805, Fax:+82-2-928-0179
E-mail: sjj@dali.korea.ac.kr

Abstract: A new scheduling algorithm called the Adaptive Weighted Round Robin with Delay Tolerance (AWRR/DT) is presented. The proposed scheme can reduce the average delay of non-real-time (NRT) class while maintaining the QoS of real-time (RT) classes. Our scheme can also reflect the traffic fluctuation of networks with a small processing burden.

I. Introduction

The ATM is a key technology to support various types of services over a high-speed backbone network. The ATM requires a cell-switching methodology that enables ATM networks to satisfy various service requirements and diverse traffic characteristics. Weighted Round Robin (WRR) scheduling [1], which is an extension of round robin scheduling, is commonly used in current ATM switches due to its simplicity and bandwidth guarantee. However, the WRR cell scheduling technique may experience somewhat large delays caused by the burstiness generated in the network.

Some variations of the WRR scheduler have been proposed to reduce the effects of burstiness [2], [3]. Since, these variations are still based on the fixed weight assignment to each queue, it is not easy to determine the proper weight of each queue. Moreover, these schemes cannot adapt to time-varying traffic in the networks efficiently.

There have been many previous researches on scheduling for RT and NRT traffic [4], [5], [6], [7]. Although these methodologies show good performance, the extra processing overhead required to implement these schemes is crucial. For example, the time stamp technique in MLT (Minimum Laxity Threshold) [5] and the management of a list of cell arrival times in HOL-PJ (Head-Of-the-Line with Priority Jumps) [6] can result in increasing processing delay for a cell. In the high-speed network such as ATM network, this type of excess overhead would cause a critical problem especially in high loaded (or congested) period. Moreover, the MLT model has only 2 types of class (RT class and NRT class) such as QLT [5] or DQLT [7], and does not consider heterogeneous types of traffic classes.

In this paper, we propose a new computationally efficient scheduling algorithm, called the AWRR/DT which do not require any extra information for each cell. Traffic classes of the proposed method are classified into a single NRT traffic class and RT traffic classes according to their delay characteristics. Each RT class has a unique delay tolerance which is the maximum acceptable delay. The weight of each RT class is determined such that any cell going into the queue of RT class can be served within this delay tolerance. Taking the delay characteristics and input traffic into account, the proposed scheme determines the weight of each class at the beginning of every cycle.

II. The proposed scheduling algorithm

The basic concept of our proposed algorithm is that all the cells going into the queue of a class should be served within the maximum acceptable delay of that class. For example, if the maximum acceptable delay of a class is 3 times the period of a cycle, any cell going into the queue of the class should get service by the end of next 3 cycles to maintain QoS. Our weight allocation mechanism is modeled to satisfy this concept.

The conceptual model of the proposed scheme is shown in Fig. 1. There are N-1 RT classes (Class 1, Class 2, $\cdots$, Class N-1) and a single NRT class (Class N). Each class gets service according to the weight allocated to that class at the beginning of each cycle. In our mechanism, the RT classes get service by the weighted round-robin in advance of the NRT class in a cycle. At the end of the service for RT classes, the NRT class is served.

We use a time-slotted model where the duration of a slot is the time required to process a single cell in the scheduler. The period of a cycle, P_c, is a number of slots. Each RT class has a delay tolerance (denoted by T_i, $T_i \geq P_c$) and $k_i = \lfloor T_i/P_c \rfloor$ counters, where $\lfloor x \rfloor$ is the greatest integer less than or equal to x. Note that k_i represents the maximum number of cycles by which all the cells going into the Q_i should be served to satisfy the delay requirement of Class i. A counter, s_i^j, represents the number of cells in the Q_i that should be served within the next jth cycle(s). And c_i^j denotes the cluster containing these s_i^j cells. Before the weight allocation of each cycle, counters are updated. That

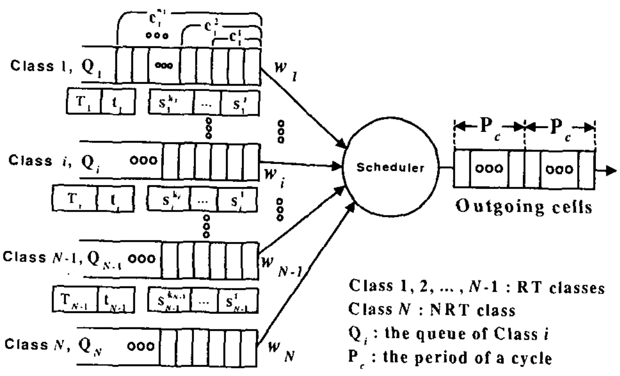

Fig. 1. Conceptual model of AWRR/DT scheduling

is, $s_i^{k_i}$ is set to the total number of cells stored in the Q_i, and $s_i^{j-1} = s_i^j - [\textit{number of cells served or discarded during the previous cycle}]_i$, where $j = 2, 3, \cdots, k_i$.

In the proposed scheduler with N-1 RT classes and a single NRT class, the weight of RT class i is obtained by

$$w_i = max\left\{\left\lceil s_i^{k_i} \cdot u_i^{k_i}\right\rceil, \cdots, \left\lceil s_i^j \cdot u_i^j\right\rceil, \cdots, \left\lceil s_i^1 \cdot u_i^1\right\rceil\right\}, \tag{1}$$

where $\lceil x \rceil$ is the smallest integer that is greater than or equal to x, and u_i^j is the urgency factor of cluster c_i^j. Each element $\left\lceil s_i^j \cdot u_i^j\right\rceil$ in (1) is the weight satisfying the QoS requirement of the associated cluster. To obtain a weight which satisfies the QoS requirement of every clusters in Q_i, the maximum value w_i is selected among the weight for each cluster. The urgency factor, u_i^j, is defined by

$$u_i^j = \frac{P_c}{T_i - t_i - (k_i - j) \cdot P_c}, \tag{2}$$

where t_i $(0 < t_i < P_c)$ represents the number of slots elapsed from the arrival of the latest queued cell in the Q_i during the previous cycle. The t_i is reset whenever a cell arrives at Q_i. The reciprocal of u_i^j, $1/u_i^j$ (a real number, $0 < (u_i^j)^{-1} < k_i$), represents the residual number of cycles by which all the cells in c_i^j should be served to satisfy the delay requirement of Class i. Note that the weight for c_i^j is s_i^j (the number of cells in c_i^j) divided by this residual number of cycles. To obtain the smallest integer greater than or equal to each element in (1), the ceiling function $\lceil\cdot\rceil$ is used.

After the allocation of weight to each RT class, residual slots are assigned to NRT class. Thus, the weight of NRT class is obtained by

$$w_N = P_c - \sum_{i=1}^{N-1} w_i. \tag{3}$$

The weight of each class in the proposed scheduling algorithm is determined by (1), (2) and (3) at the beginning of every cycle, periodically. If the sum of RT weights is greater than P_c, the proportional weight of each class, w_i', is redetermined using (4) so that its sum is equal to P_c. In this case, NRT class is not served ($w_N' = 0$).

$$w_i' = min\left\{\left\lceil\left(w_i / \sum_{m=1}^{N-1} w_m\right) \cdot P_c\right\rceil, r_{i-1}\right\}, \tag{4}$$

where $r_{i-1} = P_c - \sum_{m=1}^{i-1} w_m'$ and $i = 1, 2, \cdots, N$.

In general, RT cells arrived at the destination beyond its maximum acceptable delay might be discarded. Hence the proposed algorithm incorporates a cell discarding scheme to reduce the QoS degradation incurred by the congestion of networks. With this scheme, some RT cells which may not be transmitted within the delay tolerance of the corresponding class are discarded in advance to satisfy overall QoS requirements. If $w_i' < min\left\{\left\lceil s_i^{k_i} \cdot u_i^{k_i}\right\rceil, \cdots, \left\lceil s_i^j \cdot u_i^j\right\rceil, \cdots, \left\lceil s_i^1 \cdot u_i^1\right\rceil\right\}$, none of the clusters in Q_i could be served properly. In this case, the scheduler discards cells from Q_i and the number of cells to be discarded, d_i, is equal to

$$d_i = min\left\{\left\lceil s_i^{k_i} \cdot u_i^{k_i}\right\rceil, \cdots, \left\lceil s_i^j \cdot u_i^j\right\rceil, \cdots, \left\lceil s_i^1 \cdot u_i^1\right\rceil\right\} - w_i'. \tag{5}$$

The determination of weights can be summarized as follows:

/*The period of cycle, P_c, and the delay tolerance of Class i, T_i are given. There are N-1 RT classes (Class 1, 2, $\cdots$, N-1) and a single NRT class (Class N)*/

Step 1 : Initialization

Calculate k_i for $i = 1, 2, \cdots, N-1$.

Initialize t_i and s_i^j for $i = 1, 2, \cdots, N-1$ and $j = 1, 2, \cdots, k_i$.

Do the following steps (Steps 2, 3, and 4) at the beginning of every cycle.

Step 2 : Update s_i^j and determine w_i.

1) Update s_i^j for $i = 1, 2, \cdots, N-1$ and $j = 2, 3, \cdots, k_i$.

- $s_i^{j-1} = s_i^j - [\textit{number of cells served or discarded during the previous cycle}]_i$.
- $s_i^{k_i} = \textit{the total number of cells stored in the } Q_i$.

2) Determine w_i using (1), (2), and (3).

Step 3 : If the sum of RT weights is greater than P_c, the weight of each class, w_i', is redetermined using (4).

Step 4 : If the redetermined weight of an RT class i, w_i', is less than the minimum weight, determine the number of cells to be discarded from Q_i using (5).

III. Experimental results

Our model for discrete-event simulation has two RT classes (Classes 1 and 2) and a single NRT class (Class

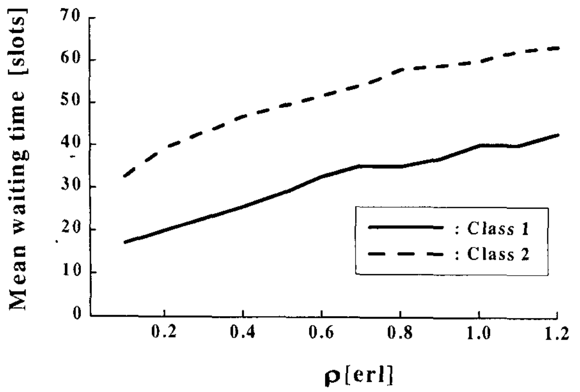

Fig. 2. Mean waiting time versus traffic load for Classes 1 and 2 of the AWRR/DT

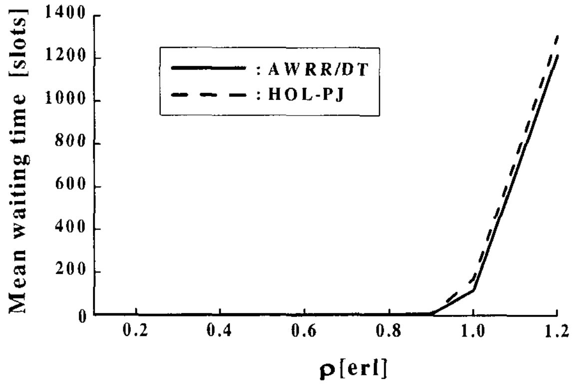

Fig. 3. Mean waiting time versus traffic load for Class 3

3). Assume that the input to Q_1 is the superposition of 10 on-off sources and the inputs to Q_2 and Q_3 are independent and exponentially distributed sequences. We assume no buffer size limitation for the queues.

In the simulation, the delay tolerance values of Classes 1 and 2 equal to 60 and 80, respectively, and the period of the cycle is 20. The time-out markers of Classes 2 and 3 for HOL-PJ are set to 20 and 50, respectively. The traffic ratio among Classes 1, 2, and 3 is 1:1:2 (RT:NRT=1:1).

Figs. 2 and 3 show the simulation results. The mean waiting times for RT classes (Classes 1 and 2) of the AWRR/DT are within their corresponding delay tolerance, 60 for Class 1 and 80 for Class 2 (see Fig. 2). Those of HOL-PJ are below 5 both for Class 1 and 2 because the HOL-PJ always schedules RT traffic first. The cell loss ratio (delay tolerance violation ratio including the number of discarded cells) for RT classes of the AWRR/DT is minimal (below 3×10^{-4} when $\rho = 0.9$). In the case of the NRT class (Class 3), the proposed scheme shows smaller mean waiting time than HOL-PJ in relatively high-loaded condition (see Fig. 3). In the region of lower load, the difference is indistinguishable. The proposed policy shows satisfactory performance with significantly reduced complexity compared with HOL-PJ.

The performance results associated with the cell discarding scheme are shown in Table I. It is seen that the proposed scheme with the discarding method produces the lower cell loss ratio (including the number of discarded cells) and the smaller mean waiting time than that without the discarding scheme.

TABLE I

COMPARISON OF MEAN WAITING TIME AND CELL LOSS RATIO (WITH AND WITHOUT THE CELL DISCARDING METHOD)

AWRR /DT	Without the discarding method				With the discarding method			
Class	Class I		Class II		Class I		Class II	
Input traffic load (ρ)	Mean waiting time [slots]	Cell loss ratio (%)	Mean waiting time [slots]	Cell loss ratio (%)	Mean waiting time [slots]	Cell loss ratio (%)	Mean waiting time [slots]	Cell loss ratio (%)
1.0	44.23	0.49	63.53	0.49	44.12	0.42	63.30	0.52
1.1	47.88	5.31	66.08	2.47	47.05	2.79	63.40	0.60
1.2	48.72	6.08	67.97	4.28	48.00	3.02	67.11	1.73

IV. Conclusion

We showed that our proposed algorithm can reduce the average delay of NRT class while maintaining the QoS of RT classes, without the use of extra information for each cell that is required in MLT or HOL-PJ. Therefore, the proposed scheme can be efficiently employed in high-speed switching networks.

Acknowledgments

The support from KOSEF (Korea Science and Engineering Foundation, 97-0100-1501-5) is greatly appreciated.

References

[1] M. Katevenis, S. Sidiropoulos, and C. Courcoubetis, "Weighted round-robin cell multiplexing in a general-purpose atm switch chip," *IEEE Journal of Selected Areas in Communications*, vol. 9, pp. 1265–1279, Oct. 1991.

[2] Y. T. Wang, T. P. Lin, and K. C. Gan, "An improved scheduling algorithm for weighted round-robin cell multiplexing in an atm switch," *in Proc. of IEEE ICC'94*, pp. 1032–1037, 1994.

[3] H. Shimonishi and H. Suzuki, "Performance analysis of weighted round robin cell scheduling and its improvement in atm networks," *IEICE Trans. on Communications*, 1998.

[4] D. Hong and T. Suda, "Congestion control and prevention in atm networks," *IEEE Network Magazine*, vol. 9, pp. 10–16, July 1991.

[5] R. Chipalkatti, J. Kurose, and D. Towsley, "Scheduling policies for real-time and non-real-time traffic in a statistical multiplexer," *Proceedings of IEEE Infocom'89*, pp. 774–783, Apr. 1989.

[6] Y. Lim and J. E. Kobza, "Analysis of a delay-dependent priority discipline in an integrated muticlass traffic fast packet switch," *IEEE Trans. on Communications*, vol. 38, pp. 659–665, May 1990.

[7] J.-S. Jang, S.-H. Shim, and B.-C. Shin, "Analysis of dqlt scheduling policy for an atm multiplexer," *IEEE Communications Letters*, vol. 1, pp. 175–177, Nov. 1997.