SCHEDULING INPUT-QUEUED SWITCHES BY SHADOW DEPARTURE TIME ALGORITHM

Jinhui Li *Nirwan Ansari*

Abstract: A new scheduling algorithm for input-queued cell switches, referred to as shadow departure time algorithm (SDTA), is introduced. Simulations demonstrate that the SDTA's cell delay distribution is more desirable than those of the GS-OCF and GS-LQF algorithms in terms of cell overdue probability.

Introduction: The input-queued (IQ) switching architecture is attractive for high speed switch implementation owing to its scalability. The throughput of an IQ switch using the longest queue first (LQF) [1] algorithm and oldest cell first (OCF) [2] algorithm, can achieve 100% throughput under all admissible independent traffic. As the approximations of LQF and OCF, GS-LQF algorithm and GS-OCF algorithm were introduced [3] using stable matching [4], and thus have lower complexity than LQF and OCF which employ maximum weight matching. The disadvantage of LQF (as well as GS-LQF) is that it can lead to starvation under certain condition [2].

In this Letter, shadow departure time algorithm (SDTA), which is starvation-free, is proposed in order to achieve better performance than GS-LQF and GS-OCF with the same level of complexity.

The Shadow Departure Time Algorithm (SDTA): Consider an $N \times N$ IQ cell switch, consisting of N inputs, N outputs, and a non-blocking switch fabric, with virtual output queuing (VOQ) [2], in which multiple VOQs directed to different outputs are maintained at each input. Assume there exists a shadow $N \times N$ FIFO output-queued (OQ) switch, and the exactly same traffic going to the IQ switch is fed into the shadow switch concurrently. $SDT(\mathbf{c})$, the Shadow Departure Time of a cell $\mathbf{c}$, is defined as the time that the cell departs the shadow switch. Since FIFOs are used both in the VOQs of the IQ switch and the output queues of the shadow switch, in the same VOQ the cell that arrives later will have a larger SDT, and the head-of-line (HOL) cell will have the smallest SDT among all the cells belonging to the same VOQ.

Let $Q_{i,j}$ denote the VOQ directed to output j at input i. In order to apply SDTA, $w_{i,j}(n)$, the weight of $Q_{i,j}$ at time slot n, is defined as:

$$w_{i,j}(n) = \begin{cases} SDT(\mathbf{c^0_{i,j}(n)}) - n & \text{if } Q_{i,j} \text{ is not empty,} \\ \infty & \text{otherwise,} \end{cases}$$

in which $\mathbf{c^0_{i,j}}(n)$ is the HOL cell of $Q_{i,j}$ at time slot n. According to the above definition, the cell that has a smaller weight is more urgent to leave the switch.

SDTA searches for a stable matching [4] between inputs and outputs by setting the preference list for every input and output following the rule: input i prefers output j with smaller $w_{i,j}(n)$, and ties are broken randomly. Conversely, output j prefers input i with the smaller $w_{i,j}(n)$, and ties are also broken randomly. Stable matching [4] seeks to match N inputs with N outputs so that there is no pair consisting of an input and an output which prefer each other to the "partners" with which they are currently matched.

It can be shown that the VOQ which has the smallest weight will always be chosen to transmit the HOL cell. If a HOL cell of a VOQ is not served in a time slot, its weight will decrease by one, thus eventually becoming small enough to be served. Hence, SDTA is a starvation-free algorithm.

Performance: In order to simulate the bursty nature of real traffic, on-off traffic model was used in the simulations. The on-off traffic model assumes that the source has two states: OFF and ON state. In the OFF state, the source does not send any cells. In the ON state, the source sends data cells at the peak cell rate (P). At each time slot, the source in the OFF state changes to the ON state with a probability α. Similarly, the source in the ON state changes to the OFF state with a probability β. There is no correlation between the two probabilities.

The performance of SDTA, GS-OCF and GS-LQF algorithms was simulated in an 16×16 switch. 256 *i.i.d.* flows, each belonging to a different input-output pair, were created in the simulations. Bursty traffic was generated based on the on-off traffic model. β was chosen to be 0.1 and the peak cell rate was set to be the link capacity. Each simulation lasted for 1 million time slots.

Fig. 1 shows the cumulative distributions of cells versus cell delay using GS-OCF, GS-LQF and SDTA algorithms under *i.i.d.* on-off traffic with a traffic load of 80%. The curves approximate the cumulative density functions of cell delay. The figure indicates that the cumulative distribution of SDTA are always larger than that of GS-OCF, implying that for a given delay bound more percentage of cells can be transmitted within the delay bound using SDTA than using GS-OCF. In other words, cells using SDTA have a lower probability of overdue than those using GS-OCF. When cell delay is large, the cells cumulative distributions of SDTA and GS-OCF are almost identical, but that of GS-LQF is smaller, implying that more cells experience much longer latency than those using GS-OCF and SDTA.

Conclusions: A new algorithm, SDTA, which employs stable matching, thus having a lower complexity than algorithms which use maximum weight matching, has been proposed to improve upon existing stable matching algorithms in terms of QoS features. It is proved that SDTA is starvation-free. Simulations also show that SDTA has a larger cumulative distribution of cell delay than GS-OCF which implies that switches using SDTA have a lower probability of cell overdue than those using GS-OCF.

Acknowledgments: This work was partially supported by Lucent Technologies, and was carried out in part while N. Ansari was on leave at the Department of Information Engineering, Chinese University of Hong Kong. The authors would like to acknowledge the fruitful discussion with S. Li on input queuing.

Jinhui Li and Nirwan Ansari (New Jersey Center for Wireless Telecommunications, New Jersey Institute of Technology, Newark, NJ 07102, USA)

REFERENCES

[1] N. Mckeown, V. Anantharam and J. Walrand, "Achieving 100% throughput in an input-queued switch," *Proc. INFOCOM'96,* pp. 296-302, San Francisco, CA, March 1996.

[2] A. Mekkittikul and N. McKeown, "A starvation-free algorithm for achieving 100% throughput in an input-queued switch," *Proc. ICCCN'96*, pp. 226-231, Oct. 1996.

[3] N. McKeown, "Scheduling algorithms for input-queued cell switches," Ph.D. dissertation, Univ. of California, Berkeley, 1995.

[4] D. Gale and L.S. Shapley, "College admissions and the stability of marriage," *American Mathematical Monthly,* Vol. 69, pp9-15, 1962.

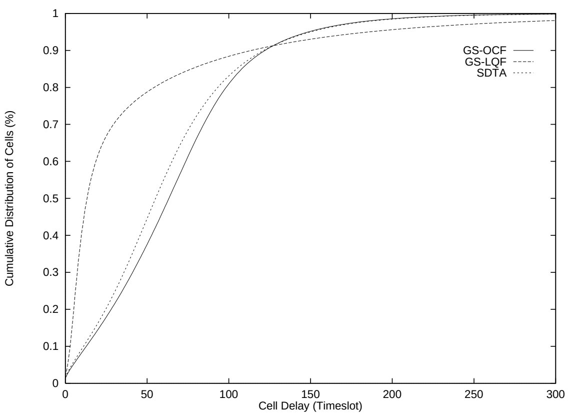

Figure 1. Cumulative distributions of cells under *i.i.d.* **on-off traffic with a load of 80%.**