International Journal of Engineering

Journal Homepage: www.ije.ir

RESEARCH NOTE

Starvation Free Scheduler for Buffered Crossbar Switches

N. Narayanan Prasanth[a*], K. Balasubramanian[b], R. Chithra Devi[c]

[a]*Department of Information Technology, National College of Engineering, Maruthakulam, India*
[b]*Department of Computer Science & Engineering, Mepco Schlenk Engineering College, Sivakasi, India*
[c]*Department of Information Technology, Dr. Sivanthi Aditanar College of Engineering, Tiruchendur, India*

PAPER INFO

Paper history:
Received 17 November 2013
Received in revised form 28 November 2014
Accepted 29 January 2015

Keywords:
Buffered Crossbar Switch
Starvation
Router
Throughput
Delay

ABSTRACT

Need for high speed internet connectivity has led to a substantial research in switching systems. Buffered crossbar switches have received a lot of attention from both research and industrial communities due to their flexibility and scalability. Designing a scheduling algorithm for buffered crossbar switches without starvation is a major challenge as of now. In this paper, we propose a Delay based prioritized queue with round-robin scheduler (DPQRS) which uses no speedup. Simulation result shows that DPQRS reduces the starvation considerably with maximum throughput and minimum delay comparable to PQRS and LQF-RR.

***doi**: 10.5829/idosi.ije.2015.28.04a.05*

1. INTRODUCTION

The annual growth of internet users is 8% compared to the population growth of 1.14% [1]. This shows the impact of internet usage in the day-to-day life. Several techniques [1] are used by the researchers to measure the internet traffic beyond its numerous limitations [2]. It is the responsibility of the researchers to provide the necessary technology to attain high speed internet with enough bandwidth. High speed connectivity can be achieved with the adequate networking devices such as switches and routers. Among the various switching architectures, buffered crossbar switches (BCS) are considered due to the two advantages [3]. First, the adoption of internal buffers makes the scheduling totally distributed, dramatically reducing the arbitration complexity. Second, and most importantly, these internal buffers reduce (or avoid) output contention as they allow the inputs to make cell transfers concurrently to a single output.

*Corresponding Author's Email: narayana.prasanth@gmail.com (N.Narayanan Prasanth)

1. http://www.internetlivestats.com/internet-users

Many scheduling algorithms have been recently proposed for the BCS architecture. Round Robin Scheduling (RRS) [4] is the simplest algorithm to implement on both the arrival and departure schedule. Another simplest algorithm, oldest cell first (OCF) [5] on the arrival schedule and RRS in the departure schedule is also employed. longest queue first (LQF) algorithm along with RRS is also proposed for BCS which offers good performance [6] in uniform and non-uniform traffics compared to RRS and OCF-RR. Both LQF-RR and OCF-RR suffers from starvation, implementation complexity and are also time consuming during input schedule [7]. With a speedup of 2, algorithm proposed in other works [8] can achieve 100% throughput under any admissible traffic. The speed requirement can be further reduced to 2-1/N in the algorithm proposed elsewhere [9]. In the literature [10], the proposal required an infinite-size buffer in the crosspoint to achieve maximum throughput. SQUID [11] is a centralized algorithm which can achieve 100% throughput under any admissible traffic. Due to its communication complexity, starvation and latency *i.e.* delay, it is impossible to implement this algorithm in large scale high-speed switching systems. Author in another work [12] proposed a distributed algorithm *DISQUO* which results in high degree of starvation.

Please cite this article as: N.Narayanan Prasanth, Dr. Kannan Balasubramanian, R.Chithra Devi, Starvation Free Scheduler for Buffered Crossbar Switches, International Journal of Engineering (IJE), TRANSACTIONS A: Basics Vol. 28, No. 4, (April 2015) 523-528

MCBF [13] considerably reduces the hardware complexity with optimal stability performance but suffers from starvation. From these studies, it is understood that no algorithm can achieve 100% throughput for any admissible traffic without starvation and speedup.

In this paper, we propose a delay based prioritized queue Round-robin scheduler (D-PQRS) for Buffered Crossbar Switches. During arrival schedule, selection of queue from Virtual Output Queue (VOQ) is based on Updated Queue Priority Value (UQPV). Every queue in a VOQ has a queue priority and at each timeslot a bonus value is added with queue priority which results in UQPV. Bonus value is computed based on the waiting time of a queue. We simulated a NxN BCS with VOQ where N=4, 16, 64, 256 and 512 and results show the starvation is considerably reduced in all VOQ's of a switch and interestingly, the throughput and delay performance are also improved comparatively to PQRS and LQF-RR algorithms.

2. BUFFERED CROSSBAR SWITCH

As an alternative to Bufferless Crossbar Switch, Buffered Crossbar Switch is introduced to improve the switching throughput. Buffering is exclusively introduced at each crosspoint to avoid the cell loss. Initially, BCS was developed [14] with buffer as a separate unit but the implementation failed to match the expectations in terms of cell loss. Later, buffer was embedded on chip along with the switching circuits. Here, the numbers of ports were limited by the memory size that can be implemented in a module chip. Therefore the size of the memory determines the number of cells switched at a single timeslot. In order to overcome the excessive on-chip memory requirement, buffered crossbar switches with input queues were proposed [15, 16]. Virtual output queue scheme is introduced to ensure that there is no packet drop during the arrival process. Therefore Buffered crossbar switches with input VOQs (VOQ/BCS) have recently received high attention. VOQ/BCS also supports variable sized packets without the need of segmentation and reassembly (SAR) circuitry and speedup is not required [17].

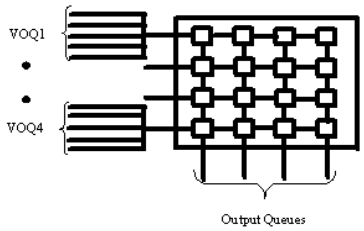

Figure 1. Buffered crossbar switch

Figure 1 shows the N x N VOQ/BCS with N input cards maintaining N VOQ's, one for each output. The buffer available in the crosspoint is assumed to be of one cell size and if there is no packet drop during switching process then the switch (BCS) is said to be stable.

3. BCS SCHEDULING

The objective of the scheduling algorithms is to maintain synchronization between the cell switching processes. Flow of cells, in and out of the switch depends on the employed scheduling algorithm. BCS uses two independent scheduling algorithms named *input or arrival scheduler* and *output or departure scheduler*. Input scheduler selects a cell from the queue and passes to the vacant buffer of the switch whereas output scheduler transfers the cell from the buffer to output queue. A switching process is said to be complete if a cell passes the input as well as output schedule in order [16, 17].

4. DELAY BASED PRIORITIZED QUEUE With ROUND-Robin SCHEDULER (D-PQRS)

Proposed algorithm uses Delay based Priority Queue algorithm (DPQ) as arrival schedule and Round-Robin (RR) algorithm as Departure schedule.

4. 1. Arrival Schedule

1. For any traffic, switching of cells from VOQ to crosspoint is based on Updated Queue Priority Value (UQPV).
2. Every queue in a VOQ has a priority value called Queue Priority (QP). QP is an integer value which is based on the number of cells occupied by the queue during the initial schedule.
3. During every switch, a Bonus value (B_i) is distributed to all the queues in the VOQ except the one which is used in the current time slot.
4. Bonus value is an integer value which is the number of timeslots the queue is waiting.
5. Queue priority is added with the bonus value to get the Updated Queue Priority Value (UQPV).

$$UQPV = QP + B_i,$$

where B_i is the Bonus value in the i^{th} timeslot

6. UQPV for the currently used queue will be reduced by one. *i.e.* UQPV = UQPV - 1
7. Queues with same UQPV have to follow the under said.
   - A queue will not be selected for schedule for consecutive number of times unless all other queues are empty

- Queue which is not scheduled at least once, will be given the next opportunity
- Otherwise follow FIFO schedule

8. If high prioritized queue is empty then opt for the next highest value.

4. 2. Departure Schedule For each departure schedule S^D at time t, if the buffer $B^D \neq 0$ then Round Robin (RR) schedule is used. If all the crosspoint buffers are empty then $S^D = 0$. Here both the scheduling algorithms are independent to each other. Since RRS is a proven scheduler in output queued switches, it is used in the departure schedule.

4. 3. An Example Consider a 4x4 BCS which uses non-uniform iid traffic and it is assumed that all the input cells are arrived to the queue at the same time. At each timeslot, we evaluated the UQPV of every queue in a VOQ. Schedule for the next timeslot is based on the UQPV of a queue. Table 1 shows the computation of UPQV for timeslots T_0-T_3. At timeslot T_0, queue priority of the first VOQ is assumed as {7, 6, 9, 8}, respectively. Third queue is selected for switching process as it holds the highest priority. Meanwhile a bonus value is distributed to all the queues except for the third one where the bonus value is reduced by 1. At T_1, T_2 and T_3 the queue priority is computed as {8, 7, 8, 9}, {10, 9, 9, 8} and {9, 12, 11, 9}, respectively. The process continues till the queue becomes empty.

5. PERFORMANCE EVALUATION

5. 1. Simulation Environment Proposed scheduler is simulated using BCSSIM [18] with switch size N=4, 16, 64, 256 and 512. For any simulation, buffer size used is one cell and the switch is operated with speedup under non-uniform traffic patterns.

5. 2. Simulation Results

5. 2. 1. Average Waiting Time (AWT) Analysis Average Waiting Time is measured for the example stated in 4.3. It is assumed that all the cells arrived to the queue at the same time. Difference between the maximum and minimum AWT (D-AWT) of a queue is evaluated to identify the occurrence of starvation. Figure 2a shows the D-AWT of a queue under Bernoulli non-uniform iid traffic which is 0.5ms for D-PQRS comparable to 1ms and 3ms for PQRS and LQF-RR, respectively. Figure 2b shows the D-AWT of a queue under Bernoulli non-uniform bursty traffic which is 0.6ms for D-PQRS comparable to 1.3ms and 3.2ms for PQRS and LQF-RR, respectively. Therefore it is proved that D-PQRS considerably reduces the starvation comparable to PQRS and LQF-RR.

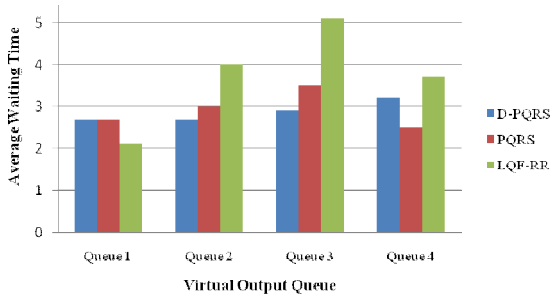

Figure 2a. Average waiting time of a VOQ under Bernoulli non-uniform iid traffic

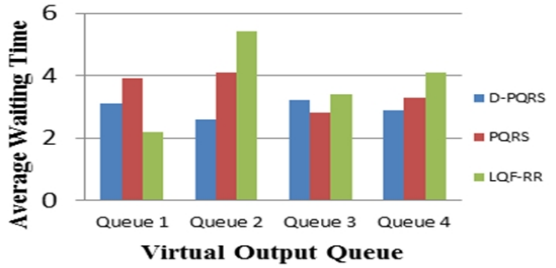

Figure 2b. Average waiting time of a VOQ under Bernoulli non-uniform bursty traffic

TABLE 1. Computation process of UPQV

Queue priority	Queue	Timeslot (s)			
		T_0	T_1	T_2	T_3
	1	7	8 (+1)	**10** (+2)	9 (-1)
	2	6	7 (+1)	9 (+2)	**12** (+3)
	3	**9**	8 (-1)	9 (+1)	11 (+2)
	4	8	**9** (+1)	8 (-1)	9 (+1)

*Values inside the ( ) shows the bonus value of the queue at each timeslot

5. 2. 2. Throughput Analysis We examine the throughput of the scheduling algorithm by varying the switch sizes. Figure 3a depicts the relationship between the throughput and load under Bernoulli non-uniform iid traffic. There is no much difference in throughput for different switch sizes. Almost 95% throughput is achieved by all the switches when minimum load is offered. Throughput is almost maintained until 60% of load is supplied, then starts dropping to 80% at maximum load. Figure 3b depicts the relationship between the throughput and load under Bernoulli non-uniform bursty traffic. Similar to iid traffic, 95% of throughput is achieved by the algorithm for all the switch sizes, and drops to 75% when maximum load is offered.

5. 2. 3. Average Cell Latency Analysis Average cell latency is analyzed for the algorithm with different switch sizes. Figure 4a and 4b depict, when minimum load is offered, less than 5% of latency is occurred

under non-uniform iid and bursty traffic. Over the time, latency constantly gets increased with respect to the load. A difference of 5% latency occurs between the switches 4x4 and 512x512 which confirms there is a little impact on switch size over latency.

5. 2. 4. Comparative Analysis We examined the throughput and delay performance of DPQRS, PQRS and LQF-RR algorithms for switch size=256 under non-uniform traffic patterns. Figure 5a shows the throughput performance of all the algorithms under Bernoulli non-uniform iid traffic. Throughput of DPQRS is 90% at minimum load and 80% at maximum load which betters PQRS and LQF-RR by 5-10%. Figure 5b shows the throughput performance of all the algorithms under Bernoulli non-uniform bursty traffic. Throughput of DPQRS is 90% at minimum load and 78% at maximum load which betters PQRS and LQF-RR by 5-15%.

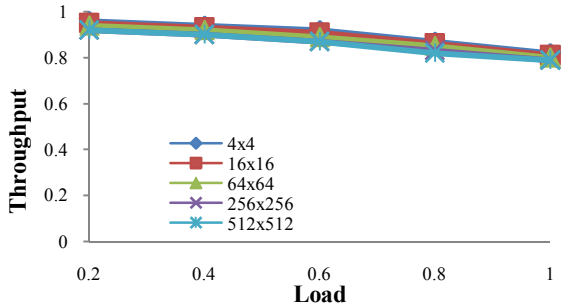

Figure 3a. Throughput performance of DPRQRS for different switch sizes under non-uniform iid traffic

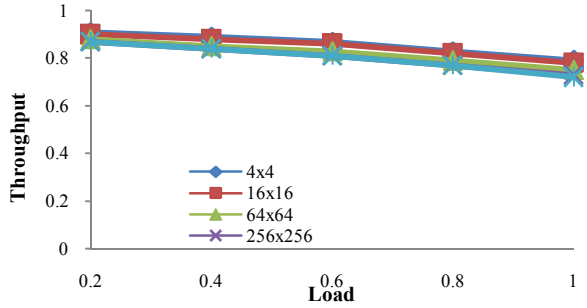

Figure 3b. Throughput performance of DPRQRS for different switch sizes under non-uniform bursty traffic

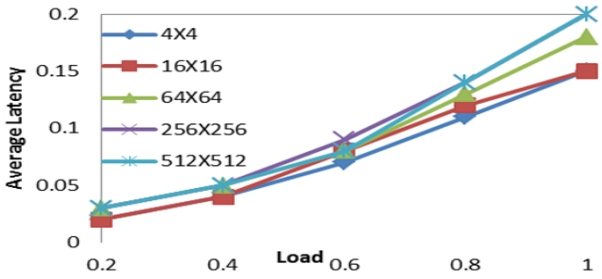

Figure 4a. Delay performance of DPRQRS for different switch sizes under non-uniform bursty traffic

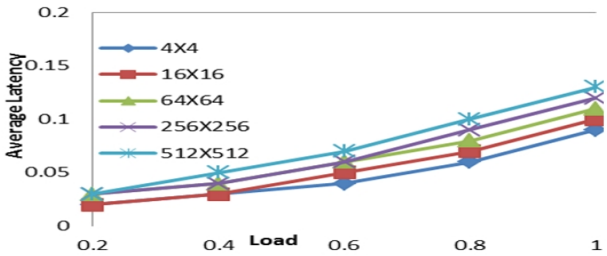

Figure 4b. Delay performance of DPRQRS for different switch sizes under non-uniform iid traffic

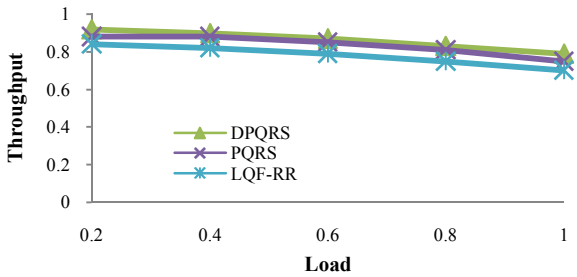

Figure 5a. Throughput performance of various algorithms for switch size=256 under non-uniform iid traffic

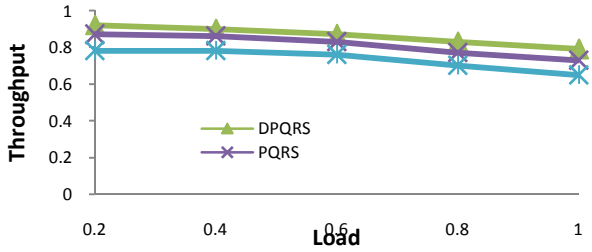

Figure 5b. Throughput performance of various algorithms for switch size=256 under non-uniform bursty traffic

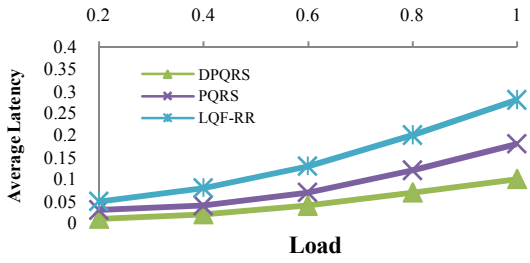

Figure 5c. Delay performance of various algorithms for switch size=256 under non-uniform iid traffic

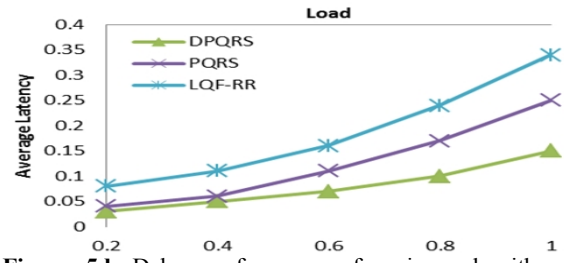

Figure 5d. Delay performance of various algorithms for switch size=256 under non-uniform bursty traffic

Figure 5c shows the delay performance of all the algorithms under Bernoulli non-uniform iid traffic. For DPQRS, average cell latency is nearly avoided at minimum load and 10% at maximum load which betters PQRS and LQF-RR by 7 and 17%, respectively. Figure 5d shows the average cell latency performance of all the algorithms under Bernoulli non-uniform bursty traffic. Average cell latency of DPQRS is 5% at minimum load and 15% at maximum load which betters PQRS and LQF-RR by 7 and 15%, respectively.

6. CONCLUSION

Designing an algorithm to achieve 100% throughput without starvation is a major challenge. D-PQRS reduces the starvation considerably and also achieves 80-95% throughput at different load structures, which is increased by 10% comparable to existing algorithms. The work can be extended to examine the algorithm with different buffer sizes and more research is also required to nullify the effect of starvation.

7. REFERENCES

1. Yaghmae, M. and Safavi, S., "A fuzzy based three color meter/marker for diffserv networks (research note)", ***International Journal of Engineering-Transactions A: Basics***, Vol. 17, No. 3, (2004), 243-250.
2. Valinataj, M., "Reliability and performance evaluation of fault-aware routing methods for network-on-chip architectures", ***International Journal of Engineering***, Vol. 27, No. 4, (2014), 509-516.
3. Elhanany, I. and Hamdi, M., "High-performance packet switching architectures, Springer, (2007).
4. Rojas-Cessa, R., Oki, E., Jing, Z. and Chao, H.J., "Cixb-1: Combined input-one-cell-crosspoint buffered switch", in High Performance Switching and Routing, Workshop on, IEEE, (2001), 324-329.
5. Nabeshima, M., "Performance evaluation of a combined input- and crosspoint-queued switch", ***IEICE Transactions on Communications***, Vol. 83, No. 3, (2000), 737-741.
6. Javidi, T., Magill, R. and Hrabik, T., "A high-throughput scheduling algorithm for a buffered crossbar switch fabric", in Communications, International Conference on, IEEE. Vol. 5, (2001), 1586-1591.
7. Mekkittikul, A., "Scheduling non-uniform traffic in high speed packet switches and routers", (1998).
8. Chuang, S.-T., Iyer, S. and McKeown, N., "Practical algorithms for performance guarantees in buffered crossbars", in INFOCOM. 24th Annual Joint Conference of the Computer and Communications Societies. Proceedings, IEEE. Vol. 2, (2005), 981-991.
9. Berger, M.S., "Delivering 100% throughput in a buffered crossbar with round robin scheduling", in High Performance Switching and Routing, Workshop on, IEEE, (2006).
10. Giaccone, P., Leonardi, E. and Shah, D., "On the maximal throughput of networks with finite buffers and its application to buffered crossbars", in INFOCOM. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings,. Vol. 2, (2005), 971-980.
11. Shen Y., Panwar S.S. and Chao H.J., "Squid: Providing 100% throughput in a buffered crossbar switch", ***Proceedings of IEEE HPSR***, (2007).
12. Ye, S., Shen, Y. and Panwar, S., "Disquo: A distributed 100% throughput algorithm for a buffered crossbar switch", in High Performance Switching and Routing (HPSR), 2010 International Conference on, IEEE, (2010), 75-81.
13. Mhamdi L. and Hamdi M., "MCBF-A high performance scheduling algorithm for buffered crossbar switches", ***IEEE Communication Letters***, Vol. 7, No. 9, (2003), 451-453.
14. Nojima, S., Tsutsui, E., Fukuda, H. and Hashimoto, M., "Integrated services packet network using bus matrix switch", ***Selected Areas in Communications, IEEE Journal on***, Vol. 5, No. 8, (1987), 1284-1292.
15. Katevenis M., "Fast switching and fair control of congested flow in broad-band networks", ***IEEE Journal on Selected Areas in Communications***, Vol. 5, No. 8, (2006), 1315- 1326.
16. Gupta, A., Barbosa, L.O. and Georganas, N., "16× 16 limited intermediate buffer switch module for atm networks", in Global Telecommunications Conference,. GLOBECOM'91.' Countdown to the New Millennium. Featuring a Mini-Theme on: Personal Communications Services, IEEE, (1991), 939-943.
17. Katevenis, M. and Passas, G., "Variable-size multipacket segments in buffered crossbar (cicq) architectures", in Communications, IEEE International Conference on, IEEE. Vol. 2, (2005), 999-1004.
18. Prasanth, N.N., Balasubramanian, K. and Devi, R.C., "Prioritized queue with round robin scheduler for buffered crossbar switches", ***ICTACT Journal on Communication Technology***, Vol. 5, No. 1, (2014).

Starvation Free Scheduler for Buffered Crossbar Switches

RESEARCH NOTE

N. Narayanan Prasanth[a], K. Balasubramanian[b], R. Chithra Devi[c]

[a]*Department of Information Technology, National College of Engineering, Maruthakulam, India*
[b]*Department of Computer Science & Engineering, Mepco Schlenk Engineering College, Sivakasi, India*
[c]*Department of Information Technology, Dr. Sivanthi Aditanar College of Engineering, Tiruchendur, India*

P A P E R I N F O

Paper history:
Received 17 November 2013
Received in revised form 28 November 2014
Accepted 29 January 2015

Keywords:
Buffered Crossbar Switch
Starvation
Router
Throughput
Delay

چکیده

نیاز به اتصال به اینترنت با سرعت بالا منجر به تحقیقات قابل توجهی در سیستم های سوئیچینگ گردیده است. سوئیچهای کراس بار بافر به دلیل انعطاف پذیری و مقیاس پذیری برای تحقیق و در جوامع صنعتی مورد توجه زیادی قرار گرفته اند. طراحی الگوریتم زمان بندی برای سوئیچ کراس بار بافر بدون قحطی یک چالش عمده از هم اکنون است. در این مقاله، ما یک صف بندی اولویت بندی شده بر اساس تاخیر با زمانبند (D-PQRS) ارائه می کنیم که از هیچ افزایش سرعتی استفاده نمی کند. نتیجه شبیه سازی نشان می دهد که DPQRS قحطی را بطور قابل توجهی با حداکثر توان و حداقل تاخیر کاهش می دهد که قابل مقایسه با PQRS و LQF-RR است.

***doi**: 10.5829/idosi.ije.2015.28.04a.05*