DOI: 10.21917/ijct.2014.0126

# PRIORITIZED QUEUE WITH ROUND ROBIN SCHEDULER FOR BUFFERED CROSSBAR SWITCHES

**N. Narayanan Prasanth[1], Kannan Balasubramanian[2] and R. Chithra Devi[3]**
[1]*Department of Information Technology, National College of Engineering, India*
E-mail: narayana.prasanth@gmail.com
[2]*Department of Computer Science and Engineering, Mepco Schlenk Engineering College, India*
E-mail: kannanbala@mepcoeng.ac.in
[3]*Department of Information Technology, Dr. Sivanthi Aditanar College of Engineering, India*
E-mail: chitra_rajan2001@yahoo.co.in

***Abstract***
***Research in high speed switching systems is in greater demand as the internet traffic gets rapid increase. Designing an efficient scheduling algorithm with high throughput and low delay is an open challenge. Most of the algorithms achieve 100% throughput in uniform traffics but failed to attain the same performance under non-uniform traffics. Moreover these algorithms are also suffers from starvation leads to extended waiting time of VOQ. In this paper, Prioritized Queue with Round Robin Scheduler (PQRS) is proposed for Buffered Crossbar Switches. We proved that our proposed scheduler can achieve 85% throughput under any non-uniform traffic without starvation.***

***Keywords:***
***Buffered Crossbar Switch, Delay Performance, Scheduling Algorithms, Starvation, Throughput***

## 1. INTRODUCTION

Need for high speed internet is the major requirement for the internet society, as the usage has been widened in the last few years. 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 accepted for high switching efficiency through fixed length switching technologies. BCS can accept packets of variable length which are then segmented into fixed sized cells and are transmitted. On the other side, the cells are reassembled to form an original packet [1]. As input queued switch suffers from throughput limitation and output queued switch suffers from output port contention, BCS avoids these shortcomings. Also input queued switch suffers from Head of Line blocking [2] which can be overcome through the introduction of Virtual Output Queue (VOQ)[3], BCS uses VOQ to input the packets. A buffer is introduced at every crosspoint of BCS to hold the incoming cell. Size of the buffer can be assigned based on architecture requirement but normally it is one cell size. To improve the switching efficiency, buffer size can be enlarged but it introduces huge implementation cost. Switch performance is based on the effective utilization of the switch i.e. all the baselines should be used at every timeslot. For effective utilization, proper scheduling algorithm must be employed.

The primary objective of the scheduling algorithms is to achieve 100% throughput with no delay in all sorts of traffic. Usually the performance of the algorithms is not the same for both uniform and non-uniform traffics. Round Robin scheduler (RRS) [4]-[6] is used at input and output schedule to achieve 100% throughput under uniform traffic whereas for non-uniform traffic it lacks its performance. Moreover the average waiting time of the VOQ is more than 5ms for a 4 × 4 switch. Then the RRS at input schedule is replaced by Longest Queue First algorithm [7]-[8] which offers 100% and 75 % throughput in Bernoulli uniform and non-uniform traffic respectively. Moreover it offers worse latency and fairness. To further improve the throughput at both the traffic formats, some algorithms uses a speedup of 2 in certain architectures [8]-[9] but speedup will provide only the half of the aggregate line throughput and also introduces a need for output queue leads to increase in implementation complexity. The authors in [10]-[11] achieved 100% throughput under uniform traffic through their proposed distributed algorithm but their performance gets dropped a maximum of 30% under non-uniform traffic. In [8]-[12], author(s) proposed an algorithm Oldest Cell First (OCF) which is simple to implement but offers poor performance in Bernoulli bursty traffic. Most Critical Buffer First (MCBF) [13] offers good stability and high performance but requires internal buffer state information for scheduling, thereby complexity gets increased. SQUISH and SQUID [14] achieves 100% throughput without speedup for any Bernoulli admissible traffic but its extended waiting time at the VOQ leads to starvation. Starvation will halt the movement of cells from a particular queue in the VOQ.

From our study, it is understood that most of the algorithm achieves 100% throughput under uniform traffic with or without speedup but their performance reduce upto 30% under non-uniform traffic. At a maximum, 70% throughput has been achieved by Longest Queue First with Round Robin scheduler (LQF-RR) under non-uniform traffic [4]-[6]. In this paper, we proposed a **P**rioritized **Q**ueue with **R**ound-robin **S**cheduler (PQRS) for Buffered Crossbar Switches (BCS) with no speedup. Through simulation, average waiting time, throughput and average cell latency is measured for different load structure under Bernoulli non-uniform iid Traffic and Bernoulli non-uniform Bursty Traffic. The outcome is compared with the LQF-RR and is considerably very good. It is understood that designing a starvation free scheduling algorithm to achieve 100% throughput under non-uniform traffic is an open challenge and we made an attempt. The structure of the paper is as follows. In section 2, we defined the Buffered Crossbar Switch along with its properties. In section 3, Priority based BCS algorithm is proposed. Section 4 comprises of simulation results for Bernoulli non-uniform traffics and its comparative analysis. Finally section 5 concludes the paper.

## 2. BCS SCHEDULING

The performance of the Buffered Crossbar Switch is based on the scheduling algorithm which is employed. Scheduling algorithm decides when and where to switch the packets, therefore designing an appropriate scheduling is must for any BCS [15]. Here, every switch requires two schedules: an Arrival Schedule (AS) and a Departure Schedule (DS). At each timeslot, Arrival Schedule selects the cell which is transferred from VOQ to BCS buffer and Departure Schedule selects the cell transferred from BCS buffer to output Queue. An arrival schedule is possible only if any of the buffers in the crosspoint is empty. If the buffer size is unlimited, then there is no need for arrival schedule that is input can be directly sent to the buffer. Because of the implementation cost, unlimited buffer size is practically not possible and in this paper the buffer used is 1 cell size. Fig.1 shows the structure of Buffered Crossbar Switch.

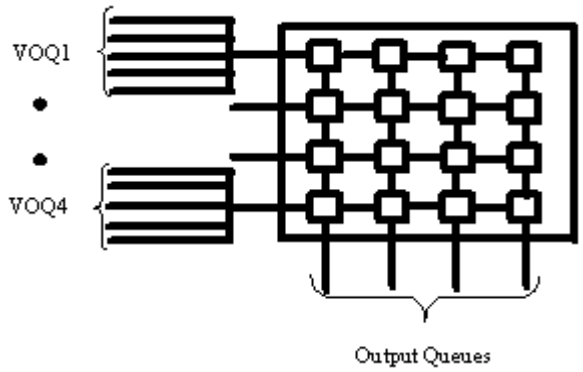


Fig.1. Buffered Crossbar Switch

For a BCS, let '*a*' denote arrival,'*d*' denote departure, '*B*' denote buffer, '*t*' denote timeslot and 'Q' denote queue then $L_{ad}(n)$ is the Queue length for any VOQ of size n. The crosspoint buffer $B(n)$ where $n = 0$ or 1 for all the iterations. Let $B_a$ is the Buffer occupancy through Arrival Schedule and $B_d$ is the Buffer occupancy through Departure Schedule then the Buffer occupancy $B_{ad}(n)$for a particular timeslot is given in Eq.(1)

$$B_a \le B \,\&\, B_d \ge B \tag{1}$$

Total number of cells available in VOQ during arrival schedule is denoted as $C_{ad}(n)$ and for every schedule at each time slot $t_{ad} \le 0$. Cell arrival from input port through VOQ is a stochastic process $A_a(t)$ and the arrival rate is denoted by $\lambda_{ad}$. Therefore $A_{ad}(t)$ denotes the arrival process from input port to output port. For an N × N switch, the arrival schedule at time is represented as $S^A(n)$. If the buffer is empty, then a switch of atleast one cell from input queue to buffer is possible as given in Eq.(2)

$$\sum S^A{}_{ad}(n) \le 1.S^A{}_{ad}(n) = 0 \tag{2}$$

$S^D(n)$ is the Departure Schedule, where a switch of atleast one cell is possible from buffer to output queue, If the buffer is not empty then it is given in Eq.(3)

$$\sum S^D{}_{ad}(n) \le 1.S^D{}_{ad}(n) = 0 \tag{3}$$

Finally, for every timeslot *t*, a switch is possible which includes both arrival and departure schedule as shown in Eq.(4)

$$\forall S^D{}_{ad}(n) = \left[S^A : S^D\right] \tag{4}$$

Every BCS has a set of properties which exhibits its character and is given as.

i. For each timeslot, a switch can use independent scheduling algorithms for arrival as well as departure schedule
ii. Every schedule should transfer atleast one cell from input port to buffer or/and from buffer to output port
iii. During every timeslot, input schedule is followed by output schedule
iv. For every timeslot, input schedule is possible only if any one of the buffer is empty and in parallel output schedule is possible only if the buffer is not empty

For any BCS, the switch is stable [14] if the algorithm used is Maximum Weight Matching (MWM) and the available queue size is bounded.

## 3. PRIORITIZED QUEUE WITH ROUND-ROBIN SCHEDULER

In this section, we propose the **P**rioritized **Q**ueue with **R**ound-robin **S**cheduler (PQRS) for Buffered Crossbar Switches. It uses independent algorithms for arrival and departure schedule and PQRS works based on the principle of Maximum Weight Based (MWB) algorithms. The algorithm is as follows.

### 3.1 ARRIVAL SCHEDULE – PRIORITY QUEUE SCHEDULER (PQS)

i. For any non-uniform traffic, selection of queue for cell transfer from VOQ to a crosspoint is based on the Queue priority
ii. Queue priority is the number of cells occupied in the queue. For every switch, a bonus priority value of 1 will be distributed to all the queues in the VOQ and is summed with the actual priority. Bonus is not applicable to the queue which is used in the current timeslot.
iii. Queues with same priority 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 atleast once, will be given the next opportunity,
   - Otherwise follow FIFO schedule
iv. If high prioritized queue is empty then opt for the next highest

### 3.2 DEPARTURE SCHEDULE

For each departure schedule $S^D$ at time *t*, if the buffer $B^D \ne 0$ 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 RR is a proven scheduler in output queued switches it is used in the departure schedule. Furthermore Extended RR is also a better option for departure schedule.

# 4. PERFORMANCE EVALUATION

## 4.1 SIMULATION ENVIRONMENT

We implemented a simulator in java *BCSSIM* that models the buffered crossbar switch of size N × N. In general for all the experiments, we used a 4 × 4 VOQ/BCS switch with a buffer size of 1 and no speedup is introduced at any stage. Input for *BCSSIM* is supplied through Bernoulli non-uniform iid traffic and Bernoulli non-uniform bursty traffic.

## 4.2 SIMULATION RESULTS

Under Bernoulli non-uniform iid traffic, the Average Waiting Time (AWT) of the entire queue in a VOQ is computed and difference between maximum and minimum AWT is less than 1ms which is shown in Fig.2. For Bernoulli non-uniform bursty traffic, the difference is slightly greater than 1ms. Therefore it is understood that all the queues are equally served for both the traffic patterns and hence it avoids starvation.

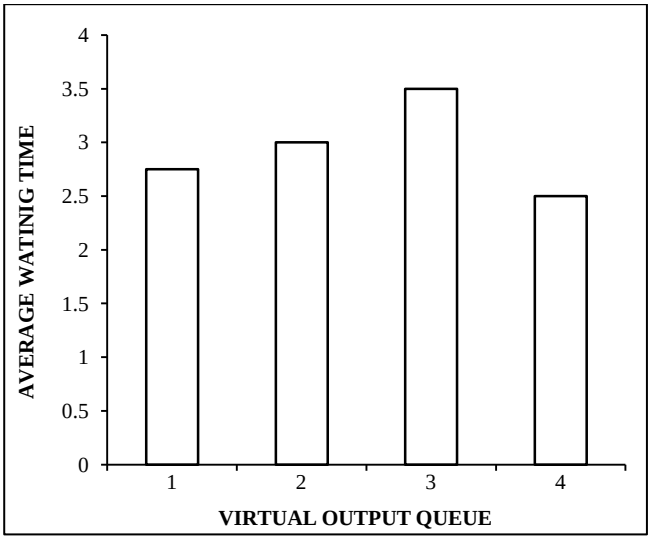


Fig.2. Average Waiting Time of a VOQ

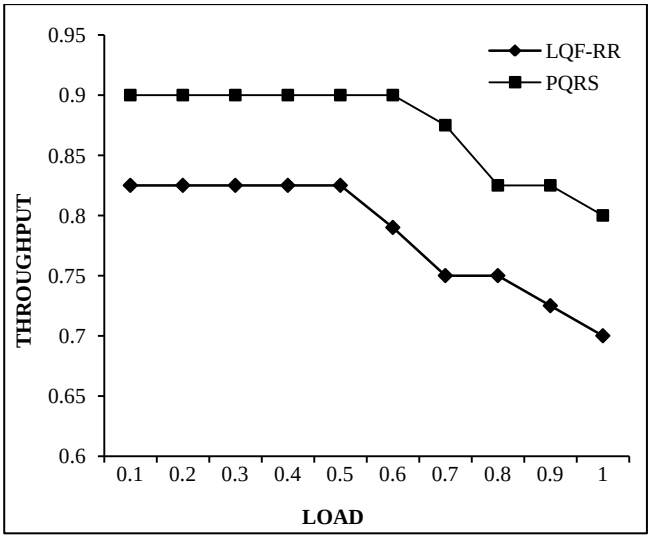


Fig.3. Throughput as a function with respect to load for Bernoulli non-uniform iid traffic

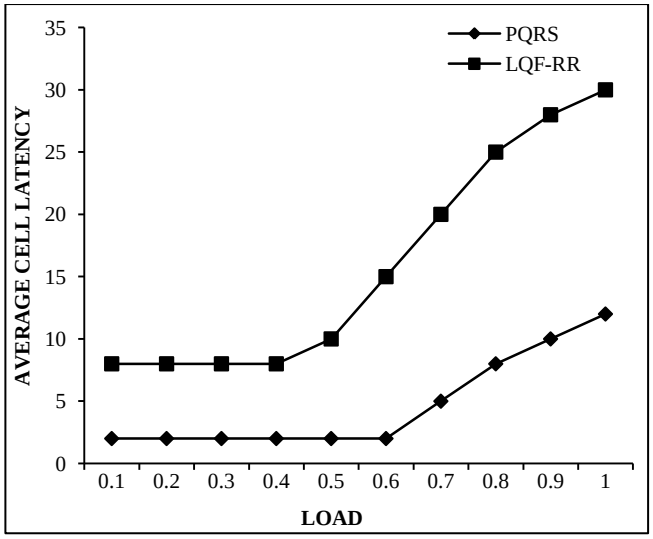


Fig.4. Average Cell Latency as a function with respect to load for Bernoulli non-uniform iid traffic

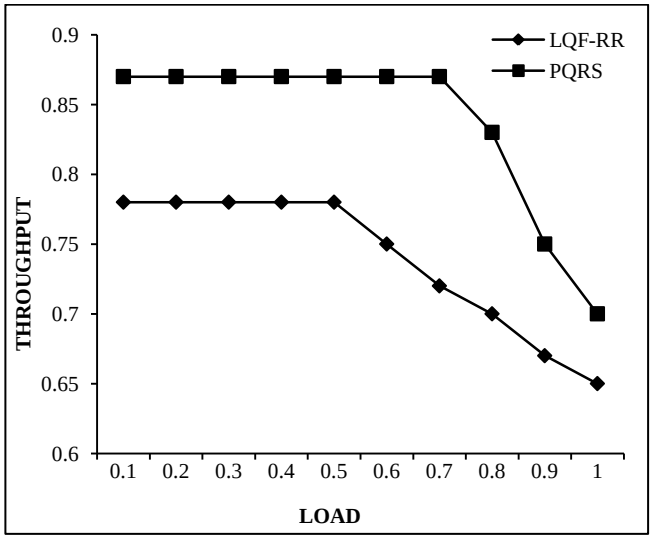


Fig.5. Throughput as a function with respect to load for Bernoulli non-uniform bursty traffic

We implemented the PQRS to compute its throughput and delay performances for various non-uniform traffic patterns and compared the outcome with LQF-RR. During the simulation, the arrival rate considerably varies between $\lambda_a$ = 0.3 to 0.7 to introduce Bernoulli non-uniform iid traffic to the switch. For such traffic, the switch behaves optimistically as shown in Fig.3 and Fig.4 until the arrival rate is < 0.6. That is, 90% throughput is achieved for arrival rate ≤ 0.6 and it decreases to 84% for the load beyond that. Average Cell Latency (ACL) is less than 5% until 50% load offered and decreases upto 10% for maximum load. Comparing to LQF-RR, PQRS extend very good delay and throughput performance by 10%.

The Fig.5 shows the throughput analysis of PQRS and LQF-RR under Bernoulli non-uniform bursty traffic with respect to load. Above 85% throughput has been achieved by PQRS until 60% of load is offered and decreases to 70% when maximum load is offered. In all the cases PQRS outperforms LQF-RR by more than 10%. Fig.6 shows that the ACL of PQRS is 5% until

60% of the load is offered and drops upto 18% when maximum load is offered. Comparing to LQF-RR, PQRS offers minimum delay performance by more than 15%.

From the results it is understood that the throughput and delay performance gets decreased when the load exceeds 70%. However the waiting time of the VOQ is stabilized for any load.

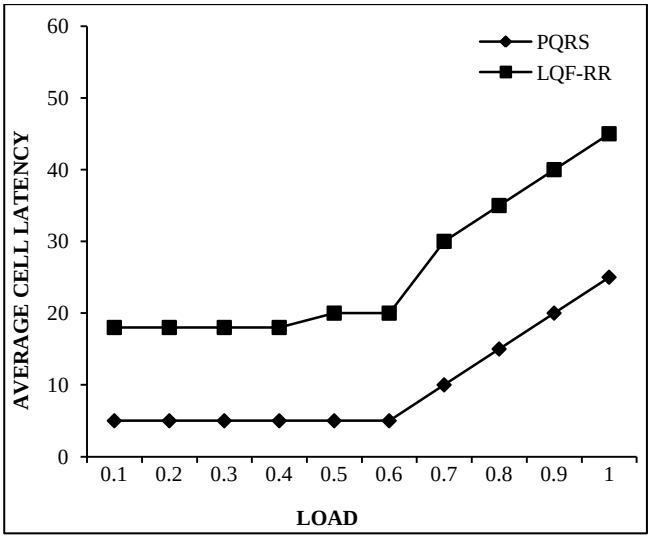


Fig.6. Average Cell Latency as a function with respect to load for Bernoulli non-uniform bursty traffic

## 5. CONCLUSION

This paper presents the Prioritized Queue with Round Robin scheduler for buffered crossbar switches. It uses a Prioritized Queue Scheduler on the arrival schedule and Round Robin algorithm on the departure schedule. These combined scheduling schemes got the essence of reducing the total waiting time involved in the VOQ. From the simulation results, it is proved that PQRS to be the better option for BCS scheduling in non-uniform traffic environments.

## REFERENCES

[1] M. Katevenis, et al., "Variable Packet Size Buffered Crossbar (CICQ) Switches", *Proceedings of the IEEE International Conference on Communications*, Vol. 2, pp. 1090-1096, 2004.

[2] M.J. Karol, M.G. Hluchyj and S.P. Morgan, "Input versus output queuing on a space-division packet switch", *IEEE Transactions on Communications*, Vol. 35, No. 12, pp. 1347-1356, 1987.

[3] R. Cessa, et al., "CIXB-1: Combined Input-One-Cell-Crosspoint Buffered Switch", *Proceedings of the IEEE Workshop on High Performance Switching and Routing*, pp. 324-329. 2001.

[4] M.S. Berger, "Delivering 100% throughput in a Buffered Crossbar with Round Robin Scheduling", *Proceedings of the Workshop on High Performance Switching and Routing*, 2006.

[5] Y. Li, S. Panwar and H.J. Chao, "On the Performance of a Dual Round-Robin Switch", *Proceedings of Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies*, Vol. 3, pp. 1688-1697, 2001.

[6] A. Mhamdi, et al., "High-performance switching based on buffered crossbar fabrics", *Elsevier – Computer Networks*, Vol. 50, pp. 2271-2285, 2005.

[7] T. Javidi, R. Magill and T. Hrabik, "A high throughput scheduling algorithm for a buffered crossbar switch fabric", *Proceeding of the IEEE International Conference on Communications*, Vol. 5, pp. 1586-1591, 2001.

[8] N. Prasanth and K. Balasubramanian, "A Study on Buffered Crossbar Switch Scheduling Algorithms", *International Journal of Computer Networking, Wireless and Mobile Communications*, Vol. 3, No. 1, pp. 13-26, 2013.

[9] D. Pan and Y. Yang, "Localized Independent Packet Scheduling for Buffered Crossbar Switches", *IEEE Transactions on Computers*, Vol. 58, No. 2, pp. 260-274, 2009.

[10] K. Balasubramanian and C. Sindhu, "ISA-Independent Scheduling Algorithm for Buffered Crossbar Switch", *International Journal of Advanced Networking and Applications*, Vol. 3, No. 3, pp. 1194-1198, 2011.

[11] P. Momcilovis, "A Distributed Switch Scheduling Algorithm", *Elsevier – Performance Evaluation, 26th International Symposium on Computer Performance, Modeling, Measurements and Evaluation*, Vol. 64, No. 9-12, pp. 1053-1061, 2007.

[12] X. Li and I. Elhanany, "Stability of a Frame Based Oldest Cell First Maximum Weight Matching Algorithm", *IEEE Transaction on Communications*, Vol. 56, No. 1, pp. 21-26, 2008.

[13] L. Mhamdi and M. Hamdi, "MCBF-A High Performance Scheduling Algorithm for Buffered Crossbar Switches", *IEEE Communication Letters*, Vol. 7, No. 9, pp. 451-453, 2003.

[14] Y. Sen, S. Panwar and H. Chao, "SQUID: A Practical 100% Throughput Scheduler for Crosspoint Buffered Switches", *IEEE/ACM Transaction on Networking*, Vol. 18, No. 4, pp. 1119-1131, 2010.

[15] M. Radonjic and I. Radusinovic, "Impact of Scheduling Algorithms on Performances of Crosspoint Queued Switch", *Annals of Telecommunications*, Vol. 66, No. 5-6, pp. 363-376, 2011.