

# Feedback-Based Scheduling for Load-Balanced Crosspoint Buffered Crossbar Switches

Xiangcheng Wu<sup>a</sup>, Wuxing Mao<sup>b,\*</sup>

College of Physics and Information Engineering Jiang Han University Wuhan, China

<sup>a</sup>363940934@qq.com, <sup>b,\*</sup> 530919036@qq.com

**Abstract.** With the rapid development of the computer and Internet, users need high-bandwidth and high-stable network environment. Routing and switching, as an important network device, supporting high-quality services. It's always been a focus of researchers. In view of the combined input Crosspoint existing queue (CICQ) controlled by flow control delay, and the pure cross-point buffer (CQ) has the problem of "insufficient throughput performance in unbalanced traffic mode". There was an algorithm proposes a load-balanced cross-buffer scheduling algorithm, which can improve the performance of throughput, quantity and delay. But the load balanced cell has the problem of out-sequence of the cells in the data stream of the same output port. This paper proposes a feedback-based load-balanced cross-buffer scheduling algorithm to ensure throughput and delay performance. Through feedback scheduling, the problem of out-sequence is well solved.

**Keywords:** Crosspoint-queued, Load-balanced, Feedback, Switches.

## 1. Introduction

With the rapid development of the Internet, Internet users have exploded. to build a high-performance broadband network, routing switching must support high-performance bandwidth and high quality of service. Therefore, route switching has always been a focus for researchers.

The output queuing (OQ) structure has high throughput and delay performance due to its work-servicing characteristics, but its scalability is insufficient due to its need for an equal acceleration ratio to the port number  $N$ . The acceleration ratio of the input queuing (IQ) structure is 1, which has good scalability [1]. However, there is a heads-of-Line (HoL) blockage problem [2]. The Virtual Output Queuing (VOQ) structure [3] overcomes the HoL blockage problem, but VOQ is more complex in the scheduling control process.

As a new switching architecture, load balancing switching technology is generally composed of two switching matrices matching input and output ports periodically and buffer between two crosses over matrices. The load balancing level transforms the non-uniformly distributed traffic model into a uniformly distributed traffic model. So that the load balanced switching structure can achieve 100% throughput under certain conditions as the output queue switching structure. However, load balancing switching requires VOQ queues before the first stage switching, and it is easy to cause output sequence disorder.

Literature [6] proposes load balancing cross buffer scheduling, first stage load balancing, and second stage buffer scheduling. The problem that the first stage of the simple load balancing algorithm requires the VOQ queue is solved. By adding a load balancing matrix, the algorithm converts the burst or unbalanced traffic of the input port into the approximate balanced traffic on the intermediate input port. The problem of simple cross-point buffer unbalanced flow performance degradation is solved. It has better throughput and latency performance than almost OQ. However, with load balancing technology, cells in the same output port (referring to the cell sequence arriving from the same input port and going to the same output port) have a problem of out of order.

## 2. The System Architecture

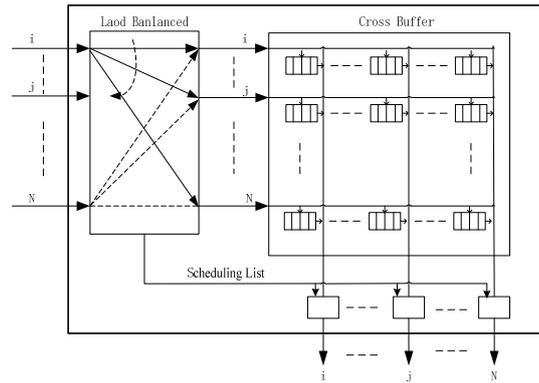


Fig 1. The structure of scheduling for load-balanced cross point buffered crossbar switches

We propose a novel architecture named LB-CQ-FS. The feedback based load balancing cross buffer scheduling switches include load balancing cross point buffer structure, and the load balancing intersection.

There are two stages of matrix in the fork buffer structure. The first stage is the  $N * N$  exchange Matrix, which uses the simplest timing polling method to match. Input and output port, at any slot  $T$  to the input port  $i$  of the first exchange matrix exchange through the predetermined mode. The current time gap exchange node configuration of the type, immediately switching to the middle port  $k$ .  $i, k$  of the second stage  $i$  is set to  $k = (i + t) \bmod N$ . The second stage exchange matrix is a cross-point buffer matrix that enters the terminal  $k$  from the middle to the second-level exchange of the cell meeting.

Directly to the corresponding cross-buffer point  $B_{kj}$ , the dispatcher on the output port  $j$  selects the corresponding cell output at each time slot. Further, each cell reaches the second stage port after the first stage load balancing, each stream. The divided letter element is added with a time stamp and a destination port number to form a linked list and sent to the corresponding scheduler. The scheduler is based on the chain.

The information of the table is scheduled in turn, which can ensure that the cells element is not out of order with way of FIFO, but it's going to go through a different queue. Delay may lead to the late arrival of the cell element is first dispatched out, causing the cell element to out of order.

## 3. Feedback-Based Scheduling

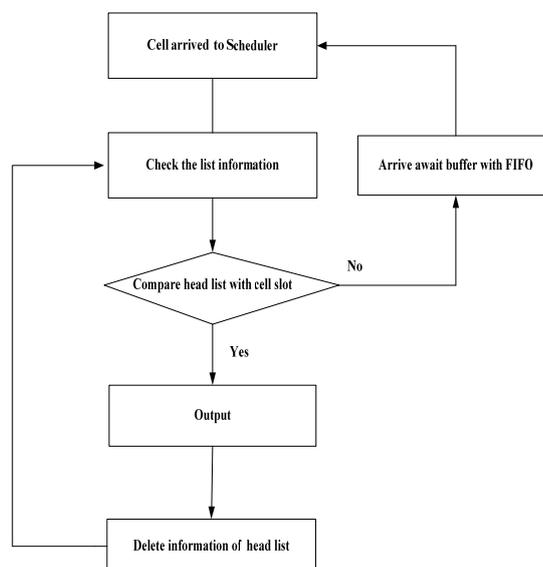


Fig 2. The flow chart of feedback scheduling algorithm.

The flow chart of scheduler scheduling algorithm is shown in Fig2. The specific scheduling algorithm is described as following:

Setp1. The cell arrives at the dispatcher, checks the existing dispatcher list, and compares the arrived cell slot with the slot information at the head of the list.

Setp2. If the information consistently indicates the correct order, the cell outputs directly and deletes the header information of the list at the same time, and the next cell slot is set at the top.

Setp3. If there is inconsistency, it means that the cells in the latter slot reach the scheduler first. At this time, we put the arrived cells into the waiting queue buffer s with FIFO order.

Setp4. The next cell arrives at the dispatcher and checks the link list comparison slot. If the consistency goes to Step2, the inconsistency goes to Step3.

#### 4. Performance Analysis

Similar to the LB-CQ exchange structure, the LB-CQ-FS switches structure is also stable under the best possible output scheduling algorithm.

The first stage load balancing scheduling without buffers, the input to the output port directly without any conflict and queuing delay, throughput is considered mainly in second stage of cross buffer. So, the analysis of the structure is as to the throughput performance of cross buffer.

$X_{kj}(n)$  indicates the length of the buffer queue in the first n time slot.

The average cell rate of i at any input port and for the purpose of any output port j obey the Bernoulli distribution with  $\lambda_{ij}$ .

If the input and output port pair meet the traffic

$$\sum_i \lambda_{ij} \leq 1$$

$$\sum_j \lambda_{ij} \leq 1$$

We say it admissible rate.

Definition 1: If an exchange scheduling algorithm is not available in the buffer of its schedule, it can always be transferred from the buffer cell to the output port, it is called Work Conserving.

Definition 2: The exchange structure that works under a particular scheduling algorithm has an acceptable rate

$$\lim_{n \rightarrow \infty} \frac{D_{ij}}{n} = \lambda_{ij}, \forall i, j = 1, \dots, N \quad (1)$$

Establish.

The switching structure is said to be stable, which means that the switching fabric can achieve 100% throughput under the control of the scheduling algorithm.

According to the fluid model presented in the literature, the fluid model of the load balanced cross buffered queuing system can be described by the following equation:

$$X_{kj}(t) = X_{kj}(0) + \lambda_{kj}(t) - D_{kj}(t) \geq 0 \quad (2)$$

$$D'_{kj}(t) = 1 \text{ If } X_{kj}(t) \geq 0 \text{ and } \pi_{kj} \in \prod(t) \quad (3)$$

Where t is a continuous time,  $\Pi(t)$  is t matching of time scheduling algorithm,  $D'(t)$  is the derivatives of D(t). The condition of the equation is that the rate of the intermediate level port to the output port satisfies the strong number theorem and is acceptable.

Proof:

Let

$$X_j(t) = \sum_k X_{kj}(t)$$

The fluid equations of the cross buffer can be rewritten as

$$X_j(t) = X_j(0) + \sum_k \lambda_{kj}t - D_j(t) \geq 0 \quad (4)$$

Obviously

$$D_j'(t) = 1,$$

And if

$$X_j(t) \geq 0$$

Scheduling algorithm will do it best.

If

$$X_j(t) \geq 0$$

Thus,

$$X_j'(t) = \sum_k \lambda_{kj} - 1 \leq 0 \quad (5)$$

According to the lemma,

If

$$X_j(t) \geq 0$$

And

$$X_j'(t) = \sum_k \lambda_{kj} - 1 \leq 0 \quad (6)$$

So that the fluid model is weakly stable, and the corresponding switch system is stable and it can provide 100% throughput.

**Theorem 1:** If the intersection point buffer exchange schedule uses the best possible scheduling algorithm. The LB-CQ-FS exchange structure is rate stable during the Bernoulli arrival process where the input packet is acceptable and the States traverse.

**Proof:** Since LB-CQ-FS does not change any structure of the load balancing level and the matching algorithm, the flow rate to any intermediate input port  $k$  to any output port  $j$  satisfies the strong number law and is acceptable.

**Theorem 2:** LB-CQ-FS can implement the secure output of the same data stream.

**Proof:** From the previous description algorithm for the LB-CQ-FS switching structure, it can be seen that all the signaling elements are marked with timestamps before scheduling, forming a feedback linked list sent to the output scheduler. Only the serial number and the time stamp of the letter element are output, so that the sequence output can be achieved, and their team head letters will not be output out of order.

## 5. Simulation Results

Through the above analysis, it is proved that the LB-CQ-FS exchange structure is stable, it has a good QoS performance, and realizes the exchange structure of the same data flow signal sequence output. The following simulation experiments will continue to verify performance of the LB-CQ-FS switching structure. We using the Longest Queue First (LQF) scheduling algorithm) realizes simulation, uses the interrupt Bernoulli arrival process (IBP) to simulate the probability distribution of traffic generated by each port, and the interrupt Bernoulli process is an ON-OFF process. It can be used to simulate burst traffic of variable-length packets. So, the resulting traffic is unbalanced.

In the balanced burst ON-OFF flow mode, the time-delay performance of LB-CQ-FS, LB-CQ, and OQ can be clearly seen in Fig3. There is almost no significant difference in the delay performance of the three under the balanced burst flow. It shows that the average delay of the LB-CQ-FS algorithm under burst traffic is close to OQ.

In the non-Balanced burst ON-OFF traffic mode, the time-delay performance of LB-CQ-FQ, LB-CQ-FS and LB-CQ is as shown in Fig 4, and the longest queue priority scheduling algorithm is also used. The delay performance of LB-CQ-FS is close to that of LB-CQ-FQ under non-equilibrium flow.

As shown in Fig 3. we can see from the throughput that when the cross-buffer capacity is small. Because the organization of virtual queues requires more buffers, the throughput performance of LB-CQ-FQ is poor, and the throughput of the feedback algorithm is close to LB-CQ.

Therefore, it is time for us to conclude that the introduction of the feedback link inherits the advantages of LB-CQ, which has good delay performance and high throughput under burst traffic.

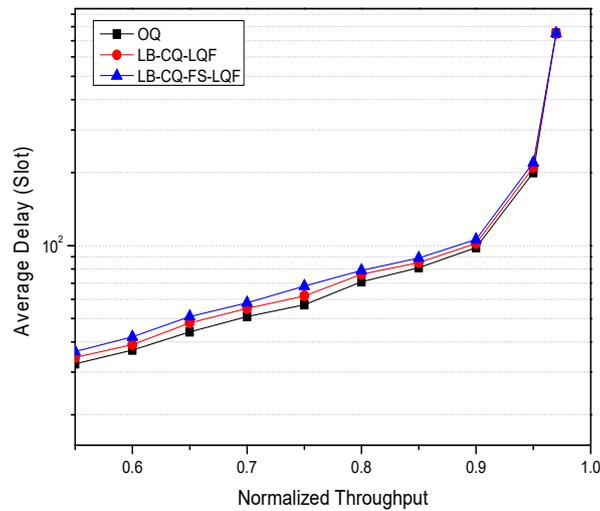


Fig 3. The average delay under burst traffic.

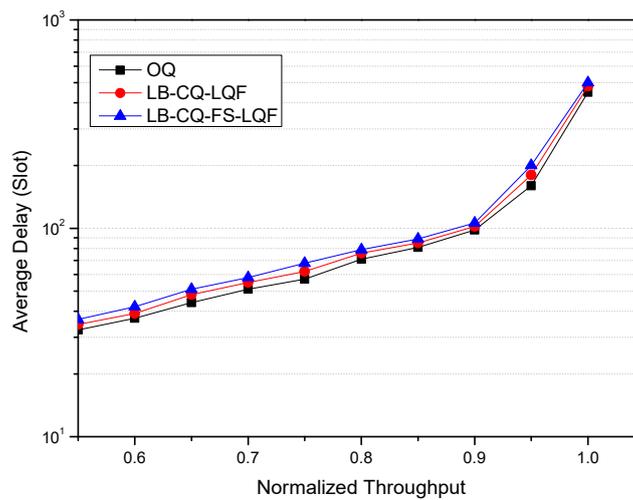


Fig 4. Average delay under unbalanced flow.

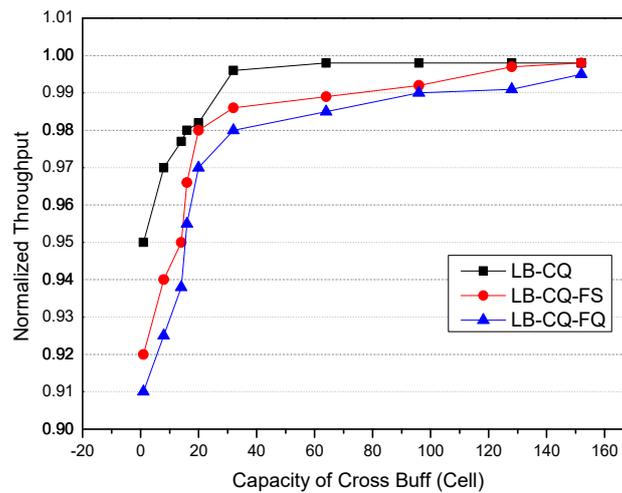


Fig 5. Maximum throughput under unbalanced traffic.

## 6. Conclusion

In this paper, we proposed the concept of feedback load balancing cross-buffer scheduling algorithm, which refers to the system structure of load balancing cross-buffer, then we put forward the idea of feedback scheduling, which solves the problem that load balancing easily causes cell disorder. Finally, the simulation show that the switches has high throughput and good delay performance.

## References

- [1]. Javidi T, Magill R, Hrabik T. A high-throughput scheduling algorithm for a buffered crossbar switch fabric [ A]. IEEE ICC' 01[ C]. St Petersburg., Russia, IEEE,2001. pp.1581-1587.
- [2]. Lotfi Mhamdi, Mounir Hamdi. CBF: A high-performance scheduling algorithm for buffered crossbar switches [ A] .4th High Performance Switching and Routing (HPSR ' 03) [ C]. Torino, Italy: IEEE, 2003. pp.67 -72.
- [3]. McKeown N, Mekkittikul A, Anantharam V. Achieving 100% throughput in an input-queued switch [ J]. IEEE Transactions on Communications, 1999, 47(8): pp.1260-1267.
- [4]. Zhang X and Bhuyan L N. An efficient algorithm for combinedinput-crosspointqueued (CICQ)switches [ A]. IEEE Globecom' 04 [C]. Dallas, USA: IEEE, 2004. pp.1168-1173.
- [5]. Hu B, Yeung K L. Feedback-Based Scheduling for Load-Balanced Two-Stage Switches[J]. IEEE/ACM Transactions on Networking, 2010, 18(4): pp.1077-1090.
- [6]. Xu Ning, Yu Shaohua. A Load-balanced Crosspoint-queued Switch Fabric [J]. China Communications. 2013, 10 (2): 134-142.
- [7]. Shen Y, Panwar S S, Chao H J. Providing 100 % throughput in a buffered crossbar switch [A] .8th High Performance Switching and Routing (HPSR' 07) [ C]. New York, USA: IEEE, pp.1 -8,2007.
- [8]. Shen Y, Panwar S S, Chao H J. Design and performance analysis of a practical load-balanced switch[J]. IEEE Transactions on Communications, 57(8): pp.2420-2429, 2009.