Increasing the Quality of Service of CICQ-VCQ switches by Combining Push-Out and Wait-to-Drain Algorithms

Seyyed Aboalfazl Sayyadi[†],

Islamic Azad University Ardestan Branch, ARDESTAN, IRAN

Summary
Existence of limited space in the core of CICQ switches makes these switches unable to support large buffers in switch Crosspoints, proportional to increasing of RTT delay. This factor leads to reduction of output throughput of today multi-cabinet CICQ switches with lengthy RTT delays. In order to support the increasing of RTT delay, despite limitation in size of Crosspoint buffers, a new structure namely CICQ-VCQ is at the center of attention. Although its core is considerably smaller than that of CICQ switches, offers much better output throughput. However, this switch encounters two important problems to support multiple priority levels. These are HOL Blocking and Buffer Hogging which extends the delay of sending high priority packets, and therefore reduces the quality of presenting services by these switches. In this paper, to solve these problems, the input scheduler of CICQ switch is implemented by combining two algorithms, namely *Push-Out* and *Wait-to-Drain* and the resultant switch is called PW-CICQ-VCQ. The delay of sending packets from this switch was compared with CICQ-VCQ switch, by means of simulation. It was seen that the delay of sending high priority packets in the new switch structure has reduced about 10% comparing with the old one.

Keywords:
CICQ-VCQ, Push-Out, Wait-to-Drain, PW-CICQ-VCQ

1. Introduction

Today, Combined Input Crosspoint Queued, CICQ, Switches have been considered extensively due to the property of scalability [1]. In these switches to solve the problem of contention in the inputs and outputs of "Crossbar" switches, for each outputs a few small buffers, "Crosspoint" Buffers, CB, are inserted at switch crosspoints. Also in the inputs, some Virtual Output Queues, VOQs, are established. Since CICQ switches directly work on the packets with variable length, they don't need to Segmentation and Reassembly, SAR, circuits and also Speedup [2].

Figure (1) shows the structure of an N×N CICQ switch with P Priority levels. VOQi,j indicates the buffer in the input i for output j. VOQ-Si is the scheduler of input i and CBi,j indicates the buffer in the crosspoint of switch that buffers input packets from VOQi,j. also CB-Sj is scheduler of output j. input and output schedulers operate independently and simultaneously with the policy of Round Robin. In this switch, a Credit-based flow control mechanism provides lossless transmission between input ports and CB Buffers [3].

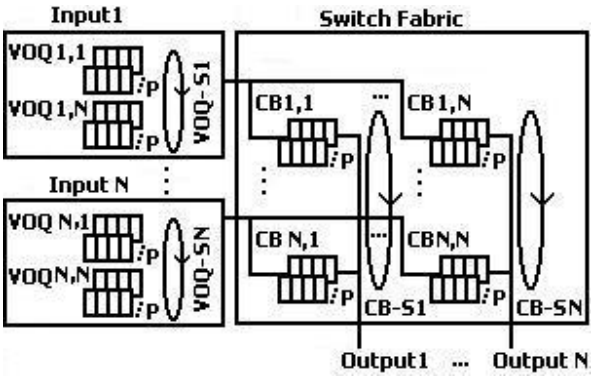

Fig. 1 An N×N CICQ switch with P Priority levels.

In order for a CICQ switch to be able to ensure that 100% of its output throughput for a completely unbalanced input traffic, when an input is sending the whole of its traffic to only a particular output, the switch should have a CBi,j buffer size according to Eq. 1 [4].

$$CBi,j \geq maximum\ size\ of\ packets + RTT \times Line\ rate \quad (1)$$

In the Eq. 1, line rate indicates the rate of switch input lines. As it can be seen, the output throughput of CICQ switches is a function of Round Trip Time, RTT, delay and size of CB buffers. Multi-cabinet CICQ switches due to the large distance between input lines and switch core, and therefore long RTT, require larger CB buffers for their output throughput not to reduce. But, due to existing limited memory in the core of these switches, it is not practical to increase the size of CB buffers. As a result, the output throughput of CICQ switches decreases, as RTT increases.

Lots of studies have been performed on CICQ switches to make them support large RTT against CB buffers with limited size. A load-balanced CICQ switch was proposed in [5]. In this switch an extra switch stage which plays a balancing role for input traffic, is inserted

Manuscript received October 5, 2025
Manuscript revised October 20, 2025
https://doi.org/10.22937/IJCSNS.2025.25.10.23

between input ports and CB buffers. In this structure, the size of CB buffers is reduced by a factor of N independent of the RTT value. However an additional cost should be borne for designation of load-balancer. Nowadays a new structure for crossbar switch is proposed which has better performance than CICQ switches, regarding utilized space in switch core, output throughput and supporting larger RTT.

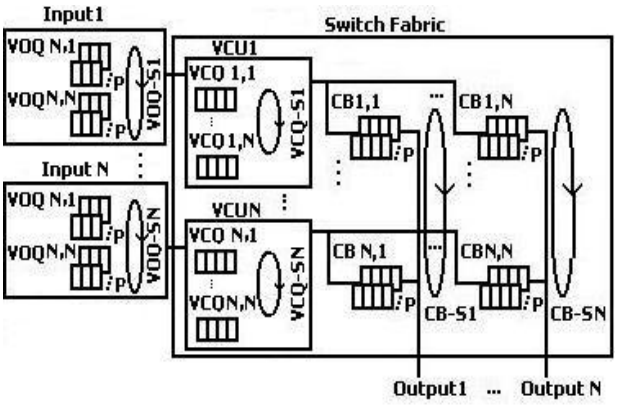

Fig. 2 An N×N CICQ-VCQ switch with P Priority levels.

This structure is called CICQ-Virtual Crosspoint Queued, CICQ-VCQ switch [6]. As it can be observed from figure (2), in the new structure there are N units of Virtual Crosspoint Units, VCUs, inside the core of switch as well. These units of memory should work with the same rate of input lines to increase the output throughput of switch, and meanwhile decrease the required memory of switch core. Each of these VCUs has a specified scheduler for its own, named VCQ-Si. Memory unit for input i, VCUi, is divided into N Virtual Crosspoint Queues, VCQs each one relating to a CB and appropriate VOQ. In this switch too, the Credit-based flow control mechanism is used to eliminate the probability of losing packets during transmission between input ports and CB buffers.

A CICQ-VCQ switch can dynamically allocate all available VCQs in a VCU to buffer completely unbalanced traffic. Therefore, the size of each VCUi require to ensure that 100% of output throughput is allocated to a completely unbalanced traffic must be equal to the RTT. The required space to perform the core of a CICQ-VCQ switch with *P* priority levels and *N* input and N Output in comparison with a respective CICQ switch with ensuring that 100% of output throughput is indicated in table (1). As it can be seen, the required memory for switch core of CICQ-VCQ is much smaller than that of CICQ switches.

As an example presented in figure (6) of [6], with a priority level, P=1, Bernoulli unbalanced input traffic, a 32×32 switch with RTT=64 Cell time, VCQ=128 Cells, CB=32 Cells, simulating results show that improvement of output throughput for CICQ-VCQ switch versus CICQ switch in a completely unbalanced traffic is 75% Meanwhile it reduces the size of switch core about 68%.

Table 1: Comparison between two switch cores memory

Switch	*Core Size*
CICQ-VCQ	N×RTT + N×N×P×Max Size of Packets
CICQ	N×N×P× (Max Size of Packets + RTT)

2. HOL Blocking and Buffer Hogging

When a VCQ is shared between multiple priority levels, it is possible that a high priority packet comes behind a low priority one. In this case the former should suffer large delay to exit from VCQ, because VCQ scheduler puts off servicing to low priority packet, due to its priority level, and does not know that a high priority packet comes afterward. This phenomenon is called Head of Line Blocking, "HOL Blocking". Sometimes it is possible that low priority packets occupy the whole space of VCQ. In this case, high priority packets should wait in the input line, because no buffer space is allocate for entering VCQ. This phenomenon is called "Buffer Hogging" [7]. Existence of these two factors makes a CICQ-VCQ to support multiple priority levels of packets postpone servicing to important and high priority levels of packets, and consequently reduces the quality of service offering of the switch. Therefore, it seems necessary to find a way that can remove or minimize the two problems in CICQ-VCQ switches.

3. Push-Out and Wait-to-Drain Algorithms

The highest priority level between packets of a VCQ is called "Effective Priority" of that VCQ. Suppose that a VCQ is shared among multiple priority levels. In this case, a high priority packet enters the VCQ but stands behind a low priority packet. If we push the low priority packet out of VCQ, the high priority packet can reduce its delay behind the low priority one. This algorithm is called "Push-Out". The high priority packet can do it by stating to VCQ scheduler that there is a high priority packet in the queue. Disadvantage of this algorithm is that some low priority packets get out of VCQ before high priority ones and are serviced earlier.

The same effect may happen in the form of Buffer Hogging. This occurs when VCQ is occupied by low priority packets. In this case, if a high priority packet in the input line is going to enter the VCQ, it is not possible. In such a case, if the high priority packet increases priority of low priority packets in the VCQ, instead of waiting in the input line, in order to have they serviced earlier and get out of VCQ as soon as possible, and then the high priority packet can decrease its delay for entering VCQ. This can be performed by input scheduler telling VCU scheduler that there is a high priority packet in the input line. Consider opposite conditions. if the next packet in the input line have less priority with comparison to effective priority of respective VCQ, input scheduler waits until all higher priority packets than present packet in the input line to exit the VCQ and then sends the lower priority packet ,provided that another high priority packet doesn't enter to input line, it gives the chance to high priority packets recently reaching the input line to pass from lower priority packets awaiting in the input and enter VCQ before them and be serviced. This algorithm which is called "Wait-to-Drain" can reduce the delay of sending high priority packets to the output. Meanwhile, the algorithm may increase the delay of low priority packets which are less important comparing to high priority ones.

4. Implementing PW-CICQ-VCQ switch

Combination of Push-out and wait-to-Drain algorithms in designation of CICQ-VCQ switch can reduce disadvantage of these switches, (i.e. long delay in sending high priority packets out of switch). For this purpose, some changes should be made in the structure of CICQ-VCQ switch input scheduler. In this paper changes according to pseudo code indicated in figure (3) in the input scheduler of CICQ-VCQ switch were made, and resultant switch was called PW-CICQ-VCQ.

- Input_id: the number of switch input port
- VCU (Input_id): the number of VCU with respect to Input_id
- P: the quantity of priority levels supported by the switch
- VOQ (Input_id,Output_id,k): the buffer in the Input_id for Output_id with priority of k
- VCQ (Input_id,Output_id): the buffer of present VCQ in VCU(Input_id) for Output_id
- Effpr[VCQ(Input_id,Output)]: the effective priority in the buffer of VCQ available in VCU(Input_id) for Output_id
- HOL: the packet available at the beginning of the queue

```
For an VOQ (Input_id, Output_id) 1≤ id ≤ N
Loop for each time slot
Loop {1≤ K ≤P} // top most priority is 1
If VCQ (Input_id, Output_id) in VCU (Input_id) _
                                        NOT FULL
  If VOQ (Input_id, Output_id, k) is NOT EMPTY
   If K > Effp [VCQ (Input_id, Output_id)]
      Send HOL (VOQ (Input_id, Output_id, k)) to _
                          VCQ (Input_id, Output_id)
      Push-Out
   Else if K=Effp [VCQ (Input_id, Output_id)]
      Send HOL (VOQ (Input_id, Output_id, k)) to _
                          VCQ (Input_id, Output_id)
   Else
      Wait-to-Drain
   End if
  End if
Else
  If VOQ (Input_id, Output_id,k) is NOT EMPTY
    If K >Effp [VCQ (Input_id, Output_id)]
      Push-Out
    Else
      Wait-to-Drain
    End if
  End if
End if
K= (k+1) mod P
End loop
End loop
```

Fig. 3 Pseudo code of PW-CICQ-VCQ input scheduler

5. Simulation

In this paper software named "Simscript" [8] was used for simulating and the packet transmission. By means of this software, 32×32 CICQ-VCQ and PW-CICQ-VCQ switches with port speed 10Gbps were implemented. Also RTT=400ns, VCU=1000 byte, CB= 1500 byte and 4 priority levels P0>P1>P2>P3 was taken into account. Furthermore, packets header and switch speedup were ignored. Packets length was considered variable and destination of packets was distributed equally in the switch. The probability of entering all priorities was considered the same. In addition, the "Burst60" of "Poisson process" was utilized for implementing the input traffic to the switch. In this model, each of traffics includes 60 back-to-back packets and the length of non-

traffic periods is distributed exponentially. The packets of traffic have the same destination and priority level. The size of each packet is chosen independently by "Pareto distribution" [7]. In this case, the average length for a packet is 400 bytes and the length of smallest and longest packets are considered 40 and 1500 bytes respectively. Thereby the mean size of traffic equals 23KBytes. This traffic model simulates the strictest case of a real traffic and detects the system problems against Buffer Hogging and HOL Blocking as much as possible.

In this paper, comparison about the delay of two priority levels (namely P0 and P3) of packets in the two switches (i.e. CICQ-VCQ and PW-CICQ-VCQ) was made. The delay is defined as follows: the time that a packet is going out of the switch core minus the time that this packet going into the input port. Figure (4) shows the delay of priority level P0 and figure (5) shows the delay of priority level P3 of two switches versus increase of input traffic load from 50% to 100%.

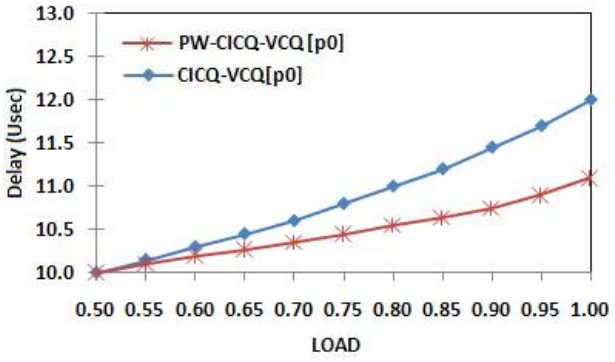

Fig. 4 Comparison between delay of priority level P0

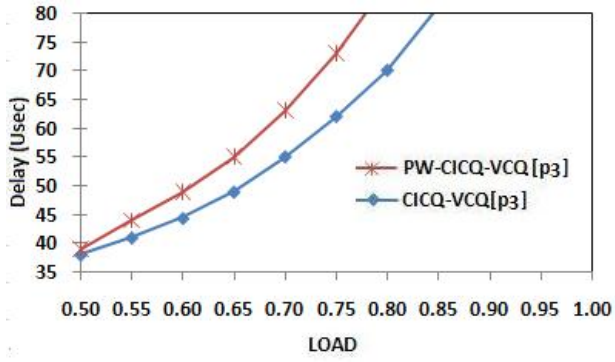

Fig. 5 Comparison between delay of priority level P3

6. Conclusions

As it is indicated from figure (4), the delay of sending high priority packets in PW-CICQ-VCQ switch is less than that of CICQ-VCQ switch. This reduction in the input traffic of 100% is amount to 10%. This improvement is due to the new algorithms of Push-Out and Wait-to-Drain performed by the input scheduler of PW-CICQ-VCQ switch. However, the result is different for low priority packets. Because the two algorithms give the priority to high priority packets and prevent low priority packets from entering the switch until there are packets with higher priority in the respective VCQ. With regards to figure (5), the delay of sending low priority packets in PW-CICQ-VCQ switch is longer than CICQ-VCQ one. Nevertheless, since low priority packets are less important than high priority ones, in the quality of service giving of switch, this increasing of delay doesn't have much effect in the efficiency of PW-CICQ-VCQ switch. Therefore, the quality of service offering in PW-CICQ-VCQ switch is much better than that of CICQ-VCQ one.

Acknowledgments

This paper therefore research plan titled "Create Dynamic Scheduling and Routing Algorithm for VCQ-Crossbar Switch to Increase its Output Throughput" has been performed.

References

[1] M. Nabeshima, "Performance Evaluation of a Combined Input and Crosspoint Queued Switch," IEICE Trans. on Comm., vol. E83-B, no. 3, pp. 737-741, March 2000.

[2] M. Katevenis, G. Passas, D. Simos, I. Papefstathiou, and N. Chrysos, "Variable Packet Size Buffered Crossbar (CICQ) Swithes," IEEE/ICC, vol. 2, pp. 1090-1096, June 2004.

[3] H. T .Kung, and R. Morris, "Credit-Based Flow Control for ATM Networks," IEEE Network Magazine, vol. 9, 1995, pp. 40-48.

[4] R. Rojas, E. Oki, Z. Jing, and H. J. Chao, "CIXB-1: Combined Input-One-Cell Crosspoint Buffered Switch," Proc. IEEE/HPSR, pp. 324-329, May 2001.

[5] R. Rojas, and Z. Dong, "Load-balanced Combined Input Crosspoint Buffered Packet Switch and Long Round-Trip Times," IEEE Comm., vol. 4, pp. 661-663, July 2005.

[6] K. Yoshigoe, "The CICQ Switch with Virtual Crosspoint Queues for Large RTT," Proc. of IEEE/ICC, pp. 299-303, June 2006.

[7] N. Chrysos, and M. Katevenis, "Multiple Priorities in a Two-Lane Buffered Crossbar," Proc. of IEEE/GLOBECOM, vol. 2, pp. 1180-1186, November 2004.

[8] Simscript II, http://www.simprocess.com/products/simscript.cfm

Seyyed Aboalfazl Sayyadi received the M.S. degree in Poly Technique University of Tehran in 2006. Ever since 2006, he is a Faculty of Islamic Azad University Ardestan Branch in Esfahan-Iran. His Research is done about switch architectures and packet switching networks.