Macrothink Institute™

Network Protocols and Algorithms
ISSN 1943-3581
2011, Vol. 3, No. 4

Performance Evaluation on Traffic Control in Combined Input and Cross-point Queuing Switches

Qiang Duan

Information Science & Technology Department

The Pennsylvania State University Abington College

Abington, PA, USA 19001

Tel: 1-215-881-7872 E-mail: qduan@psu.edu

Received: December 2, 2011 Accepted: December 20, 2011 Published: December 31, 2011

DOI: 10.5296/npa.v3i4.1246 URL: http://dx.doi.org/10.5296/npa.v3i4.1246

Abstract

A wide range of emerging networking applications demand computer networks to provide Quality of Service (QoS). Packet switches play a crucial role in computer communications and traffic control in packet switches is a key to QoS provisioning. The Combined Input and Cross-point Queuing (CICQ) switch employs the crossbar-based switching fabric structure with distributed output buffer at switch fabric cross-points. The CICQ switch is being widely deployed in computer networks; therefore control traffic in CICQ switches to provide QoS guarantee becomes an important issue. The research reported in this paper investigates QoS performance of CICQ switches with a variety of traffic control algorithms. Since it is well known that the Output Queuing (OQ) switch is an ideal case that provides the optimal QoS performance, the achievable performance of CICQ and OQ switches with the same traffic control algorithm is compared. The performance analysis in this paper is based on a simulation system developed using OPNET Modeler.

Keywords: Combined Input and Cross-point Queuing (CICQ) switch, traffic control, packet scheduling, QoS performance, simulation.

1. Introduction

A wide variety of networking applications demand computer networks to guarantee Quality of Service (QoS). Packet switches in computer networks play a crucial role in QoS provisioning. The traffic control system in a packet switch forwards each packet from its arrival input to its destined output, and coordinates the packet forwarding for different traffic classes. The switching fabric structure and the queuing scheme are two factors that have the most significant influence on traffic control performance in packet switches. Currently the crossbar-based structure becomes the dominating switching fabric structure for high-speed switches. The typical queuing schemes for packet switches include Output Queuing (OQ), Input Queuing (IQ), Virtual Output Queuing (VOQ), and Combined Input and Output Queuing (CIOQ) [1]. The Combined Input and Cross-point Queuing (CICQ) switch is a switch architecture that employs the crossbar-based switching fabric and a variant of the CIOQ scheme with distributed output buffer at switching fabric cross-points [2]. CICQ switches are being widely deployed in computer networks due to its potential to achieve high performance with relatively simple implementation. Therefore control traffic in CICQ switches to provide QoS guarantees becomes an important issue.

The research reported in this paper investigates QoS performance of CICQ switches with various traffic control algorithms. Since it is well known that the OQ switch is an ideal case that provides the optimal QoS performance, the performance of the CICQ switch is compared with that of OQ switch with the same traffic control algorithm. In order to achieve this objective, a simulation system is developed by using the OPNET Modeler to simulate both CICQ and OQ switches with various widely deployed packet scheduling algorithms, including round-robin (RR), weighted round-robin (WRR), and weighted fair queuing (WFQ) [3]. Performance analysis in this paper is focused on average packet delay because it is one of the most important network QoS parameters.

The rest of this paper is organized as follows. Section 2 gives a description of the CICQ packet switching system and its traffic control mechanism. Section 3 introduces the developed simulation system. Simulation results and analysis are given in Section 4. Section 5 draws conclusion.

2. CICQ Packet Switching System

A block diagram of an *N*x*N* CICQ switch is shown in Fig. 1. This switch has a set of ingress port modules, X_i, i=1, 2, …, *N*, a set of egress port modules, Y_j, i=1, 2,…, *N*, and a switching fabric. The switching fabric is based on a crossbar-based structure. There is a set of horizontal lines, each of which is connected to one ingress port module, and a set of vertical lines, each of which is connected to one egress port module. There is a buffer at each cross-point of this crossbar, which is referred to as a *cross-point buffer*. The cross-point buffer between ingress port module X_i and egress port module Y_j is denoted as $M_{i,j}$.

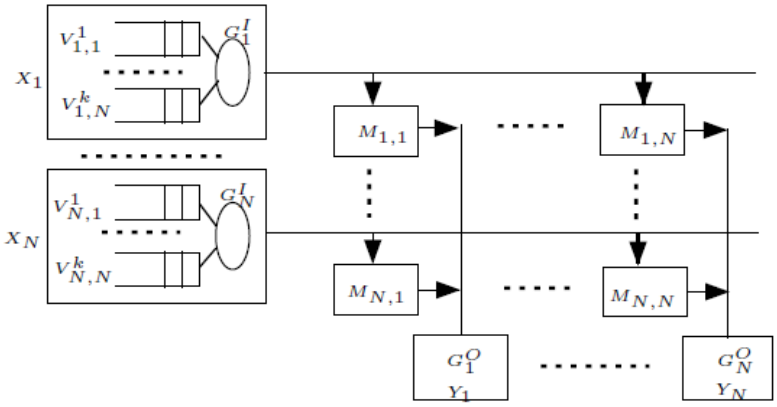

Figure 1. Combined Input and Cross-point Queuing (CICQ) switching system

Buffers at each ingress port module are organized as one virtual output queue (VOQ) for each egress port. If there are multiple traffic classes between a pair of ingress and egress ports, input buffers are organized as one VOQ for each class destined to each egress port. All cross-point buffers on the same crossbar vertical line can be accessed directly from the egress port that is connected with this vertical line, thus constituting a distributed output queue for the egress port.

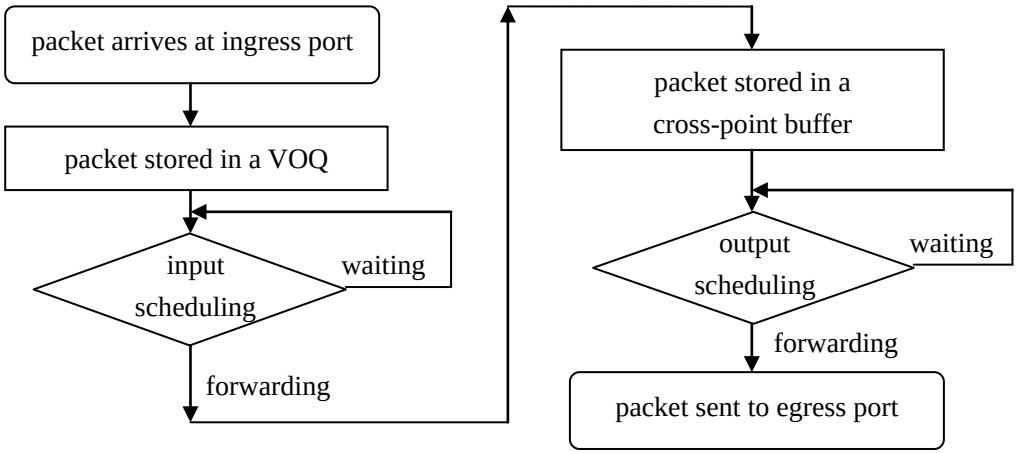

Figure 2. Traffic control procedure for packet forwarding in CICQ switch

To control traffic in the CICQ switch, each ingress port has an input scheduler, which in each time slot selects one packet from the input buffer to be forwarded into a cross-point buffer. Each egress port has an output scheduler, which in each time slot chooses one packet in the distributed output queue for sending out from the egress port. Various scheduling algorithms can be employed at input and output schedulers to control packet forwarding in the CICQ switch. Typical algorithms for packet scheduling include the Round-Robin (RR), Weighted Round-Robin (WRR), and Weighted Fair Queuing (WFQ) algorithms. The flowchart given in Fig. 2 shows the traffic control procedure for packet forwarding in CICQ switch.

To avoid losing packets due to cross-point buffer overflow, a credit-based flow control mechanism is applied between ingress port modules and cross-point buffers. A finite pool of

credits is maintained to support each flow; that is, there is one credit pool for each VOQ at each ingress port module. Initially, all credits are in the credit pools at the ingress port module. Each time a packet is transferred from the ingress port module to a cross-point buffer, one credit is moved to the credit pool in the switching fabric. Each time a packet departs a cross-point buffer, one credit is returned to the arrival ingress port module of that packet. At each input schedule, only packets in the VOQs that have available credits at the ingress port module can selected by the scheduling algorithm to be forwarded into a cross-point buffer. By setting the total number of credits for to a VOQ to be equal to the cross-point buffer space allocated to the corresponding traffic flow traversing that VOQ, the credit circulation mechanism between input and cross-point buffer controls the amount of traffic that can be forwarded by the input scheduler into cross-point buffer, thus avoiding cross-point buffer overflow.

3. Related Work

CICQ switch has attracted extensive research interest. Nabeshima [4] studied CICQ switch with VOQs at each ingress module and proposed the oldest packet first algorithm for output scheduling. Simulation in this paper shows that the average packet delay for a switch using the proposed scheduling algorithm is smaller than that for switches using ring arbitration-based algorithms. Javild and coauthors of [5] examined high-throughput scheduling algorithms for CICQ switch containing one packet buffer per cross-point and proposed the longest queue first algorithm for input scheduling and round-robin algorithm for output scheduling. Yoshigoe [6] studied a parallel-polled CICQ switch that supports variable length packets. Such switch architecture is shown to have lower average packet delay under heavy offered load than an IQ switch using the iSLIP [7] algorithm for both fixed or variable length packet. Rojas-Cessa and his colleagues proposed a CICQ switch architecture called CIXOB-*k* in [8], which employs the round-robin algorithm at both input and output schedulers and credit-based flow control between VOQ and cross-point buffers. It is shown by simulation that this switch provides 100% throughput under uniform and non-balanced traffic load and has smaller average delay than crossbar switches using iSLIP and PIM [9] algorithms.

The aforementioned research progresses mainly focus on the throughput and overall average delay performance for the entire switch. Average packet delay of individual traffic flow is an important QoS parameter that has a direct impact on networking application performance. In order to support QoS guarantee in packet switches, a scheduling algorithm to emulate the OQ switch is developed in [10], in which it is proved that a speedup factor of 2-1/*N* is necessary and sufficient for *N*x*N* CIOQ switch to emulate an OQ switch. Magill *et al.* [11] studied the issue of matching output queuing in CICQ switches, and proved that a CICQ switch with a speedup factor of two can emulate the behavior of an OQ switch with FIFO output scheduling algorithm. In [12], the authors developed an analytical modeling method for CICQ switches using some network calculus [13] techniques and gave delay bound for individual traffic flows forwarded through the switches. The method developed in [12] was

employed in [15] to show that CICQ switches may achieve equivalent performance in terms of the maximum packet delay for each traffic flow as the OQ switch without speeding up the switching fabric.

Although the output queuing match-based approach can provide QoS guarantee in a CICQ switch as in an OQ switch, such a desirable result has to be achieved at the cost of higher switch traffic control complexity with special algorithms that are not practical. The author's previous work in this area reported in [12], [15], and [16] focused on analytical modeling and theoretical performance analysis. It is also important to obtain a thorough understanding about the traffic control performance in CICQ switch with various practical scheduling algorithms. It is useful to compare the achievable performance of the CICQ switch with the ideal performance of OQ switch to see how much difference exists with various impact factors such as scheduling algorithms, traffic load, etc. The rest of this paper addresses this research problem through simulations. This new additions to the author's previous study on CICQ switches reported in this paper are mainly about simulation-based investigation on the performance of some practical traffic control algorithms employed in CICQ switches under more realistic traffic loads.

4. Simulation System for Traffic Control in CICQ Switch

OPNET Modeler is employed as the simulation tool for studying CICQ switch performance in this paper. OPNET Modeler is one of the industry's leading tools for network modeling and simulations. OPNET Modeler allows users to design and study networks, devices, protocols and applications with great flexibility and scalability in a hierarchical mode. The hierarchical network modeling approach reflects the structures of real networks; it can be divided into three domains: network domain, node domain, and process domain, each of which is provided with an editor. The network editor and the node editor are object-oriented and the process editor is based on the finite state machine model. The network editor graphically represents the topology of a network consisting of nodes and links. The node editor specifies a networking device by interconnecting a set of modules. Each module can send packets to and receive packets from other modules within the node. Each module typically consists of multiple processes, which are specified by the process editor in form of a finite state machine. The functions performed at each state can be described in C/C++ codes with support of user-defined or Modeler provided libraries.

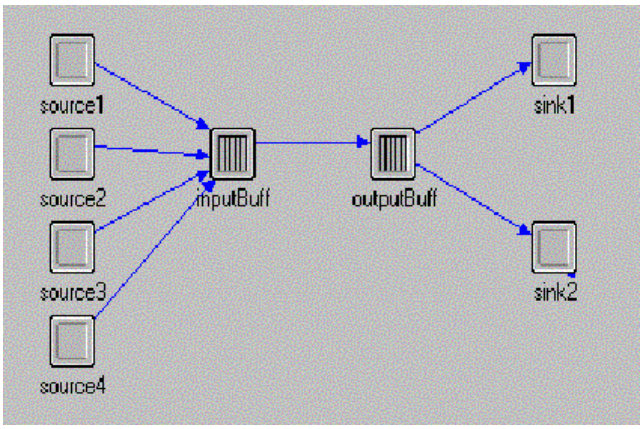

Figure 3. The OPNET Modeler simulation system for traffic control in CICQ switch

Fig. 3 shows the simulation system that is designed in this paper for studying traffic control in CICQ switches. The CICQ switch has buffers and a scheduler at each input interface, and a distributed queue and a scheduler for each output interface. Based on this structure, the simulation system consists of two interfaces: the *input buffer* module and the *output buffer* module that simulate traffic control at the input interface and the output interface respectively. In this simulation system, input traffic generated from four sources arriving at an input interface will be forwarded to an output interface, from where the traffic goes to two destinations. The input buffer module simulates VOQs and input schedulers and the output buffer simulates cross-point buffers and output schedulers. The scheduling algorithms simulated by this system include Round-Robin (RR), Weighted Round-Robin (WRR), and Weighted Fair Queuing (WFQ) algorithms. They are employed at both input and output schedulers and it is assumed that the same algorithm will be used for both input and output scheduling at the same time.

Four traffic flows are implemented in the simulation system, each of which is an IP datagram stream. The flow f_1 is from the source1 to the sink1. The other three flows are generated from source2, source3, and source4 respectively, and are destined to the sink2. For measuring delay performance, each packet gets a stamp to record the arrival time when it enters the input buffer. The sink checks the arrival time of each packet and subtracts the arrival time from it to obtain the delay time. The average packet delay time for a flow is calculated as the mean of the delay time of all packets in this flow.

Although the output queuing scheme is not practical for high-speed switches, it is an ideal case that can achieve optimal performance in terms of packet delay and throughput. Therefore, the OQ switch is also simulated and is used as a reference for comparing simulation results of the CICQ switch. The simulation system for the OQ switch is similar to that for the CICQ switch except that it only has the output buffer module, which simulates the output buffers and output schedulers. In order to compare the performance achieved by CICQ and OQ switches, the same set of scheduling algorithms is used in the OQ switch as in the CICQ switch.

5. Simulation Results and Analysis

The simulation system described in the previous section has been used to study traffic control performance of both CICQ and OQ switches with a variety of packet scheduling algorithms, including the Round-Robin (RR), Weighted Round-Robin (WRR), and Weighted Fair Queuing (WFQ) algorithms. In the simulation the switch I/O port rate, which is equal to the total capacity of output scheduler, is assumed to be R = 10M packets/second. In order to study the performance that CICQ switches can achieve without speeding up the switching fabric, the total capacity of input scheduling, which is equal to the maximum rate for packet forwarding into cross-point buffer from each ingress module, is assumed to be equal to the switch I/O port rate. The arrival rate from each traffic source can be changed to generate various traffic loads to the simulated switch. The traffic load ratio is defined as α = r /R, where r is the traffic rate generated from a source and R is the scheduler capacity. The number of packets arriving per second has a Poison distribution and the packet length is a constant 1024 bits. A relatively short and fixed packet length (128 bytes) is chosen in the simulation in order to reflect the packet forwarding in typical high-speed packet switches, where IP-datagrams are divided into packet segments with a fixed and relatively short length to facilitate hardware process. In this paper, we focus our study on the average packet delay since it is one of the most important performance parameters for most networking applications.

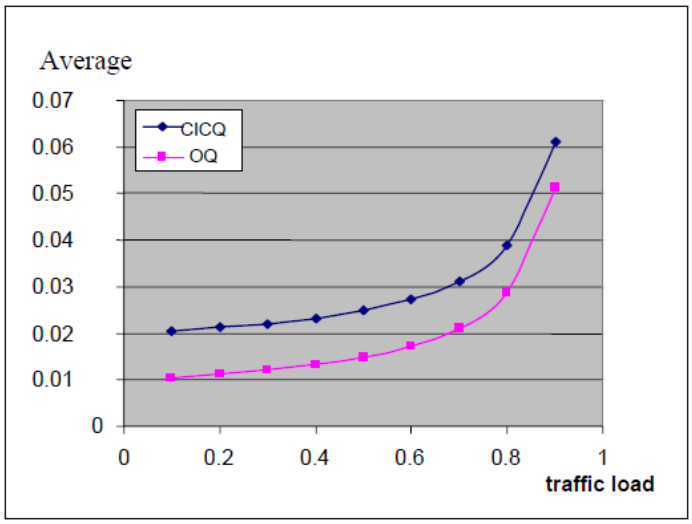

Figure 4. The average packet delay in CICQ and OQ switches.

The average packet delay for a flow *f* under different traffic load ratios is measured from the simulation results for both CICQ and OQ switches with RR scheduling algorithm. The results are shown in Fig. 4. We can see from this figure that in both CICQ and OQ switches, the average packet delay increases with the traffic load. The delay increasing rate is slower in light traffic load (α < 0.6) than the increasing rate under heavy load (α >0.6).

By comparing the average delay for the same flow in these two types of switches we can see that the CICQ switch causes longer delay than the OQ switch, as we expected. The delay difference decreases when the traffic load increases. This implies that the CICQ switch delay

performance approaches to the optimal OQ switch delay performance as traffic load increases. Fig. 5 gives CICQ-OQ switch delay increment percentage and the arrival traffic load. We can see that the delay difference is almost 90% under light traffic load (α = 0.2) while just about 30% under heavy load (α = 0.8).

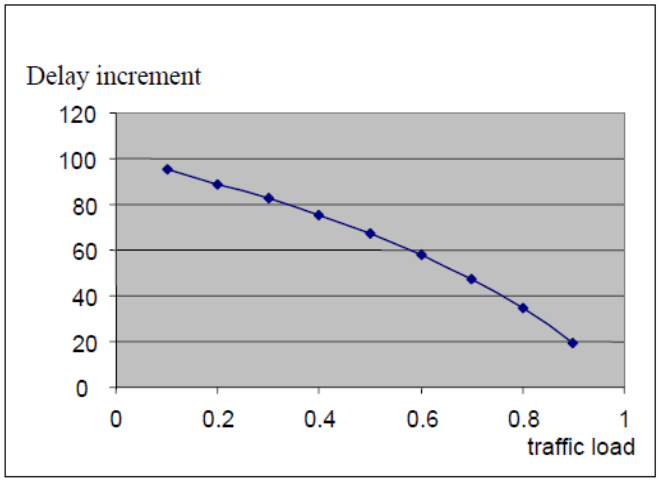

Figure 5. The average delay increment in CICQ and OQ switches.

The maximal queue length for the flow f at is measured at input buffer of the CICQ switch and at the output buffer of OQ switch. The obtained results are plotted in Fig. 6. This figure shows that the input queue length of CICQ switch and the output queue length in OQ switch are almost identical. This implies that the same amount of input buffer space will be occupied in the CICQ switch as the output buffer space in the OQ switch under the same traffic load. Both queue lengths increase with traffic load, which implies that larger buffer space is required to support a heavier traffic load at the switch. The queue lengths increase in an approximately linear rate when the load α <0.7 and in an approximately exponential rate when $\alpha \geq 0.7$.

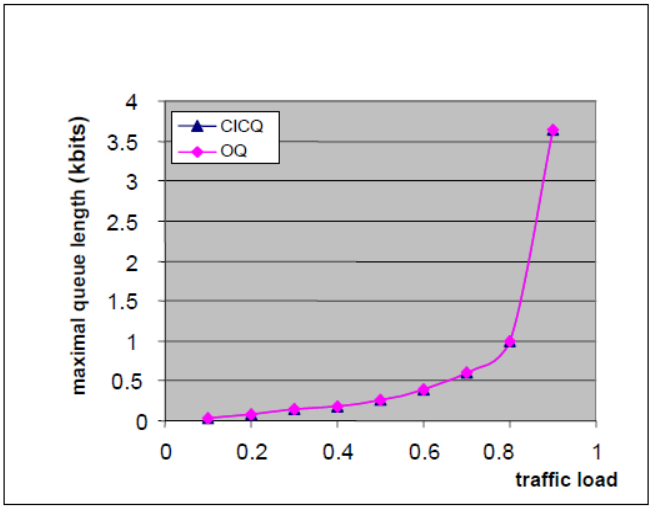

Figure 6. The maximum queue lengths in CICQ and OQ switches.

The performance of both CICQ and OQ switches with WRR scheduling algorithm is also evaluated through simulations. The average packet delay for the flow f_1 is first measured with different assigned weights and 30%, 50%, and 70% traffic load respectively in the CICQ switch. The results are shown in Fig. 7. From this figure we can see that under all traffic loads, the average delay decreases with the increment of assigned weight. This is because a larger weight means more bandwidth allocated to the flow, thus decreasing the average waiting time for packets of this flow. By comparing the average delay for a specific assigned weight under different traffic loads we can see that average delay increases with the arrival traffic load to the switch. This is because a fix assigned weight means a fixed amount bandwidth allocated to the flow, therefore the heavier the traffic load is, the longer is the average waiting time for each packet. We also noticed that the average delay increment with traffic load becomes smaller for larger assigned weight, and is close to zero for weights greater than 0.8. This is because for such weights, most bandwidth of the WRR scheduler is allocated to this measured flow. In such a situation, the waiting time in buffers for each packet is very short, and the transmission time instead of waiting time becomes the dominating part of the total packet delay, which will not be influenced by traffic load.

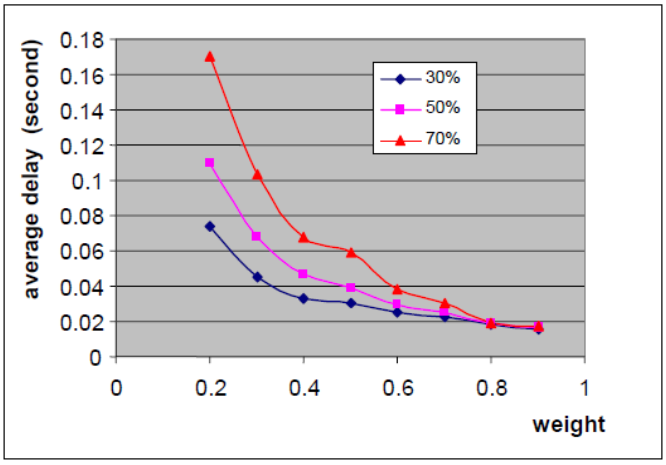

Figure 7. The average packet delay in CICQ switch with WRR scheduling.

Simulation results for average packet delay of the same traffic flow in an OQ switch are shown in Fig. 8. We can see very similar shape for each curve except the delay values are smaller in the OQ switch.

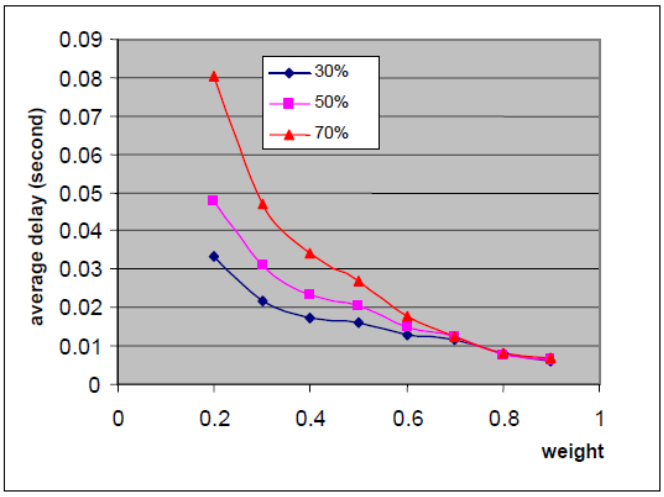

Figure 8. The average packet delay in OQ switch with WRR scheduling.

The obtained simulation results about performance of CICQ switch with the WRR algorithm also show that the average packet delay for a flow will be influenced by the weight assignment of other flows that share the same WRR scheduler. Two cases of weight assignment were tested. Suppose four flows (f_1, f_2, f_3, f_4) sharing the scheduler and the flow f_1 is assigned a weight φ_1. For case 1, $\varphi_2 = (1- \varphi_1)/2$ is assigned for f_2, and both f_3 and f_4 get a weight $(1- \varphi_1)/4$. For case 2, each of the three flows is assigned a weight $(1- \varphi_1)/3$. The average delay for f_1 was tested under 70% traffic load for both cases and the results are shown in Fig. 9. We can see that the average delay for case 1 is lower than that for case 2 when $\varphi_1 < 0.6$, and the delay for both cases are very close when $\varphi_1 \geq 0.6$. This observation implies that when a flow is assigned a relatively small weight, e.g. $\varphi_1 \leq 0.5$, the delay performance for this flow at the WRR scheduler is determined not only by the weight assigned to the flow but also the weight assignment for other flows at the scheduler. Simulation results indicate that more uniform weight assignment among other flows yields a lower average delay in this situation.

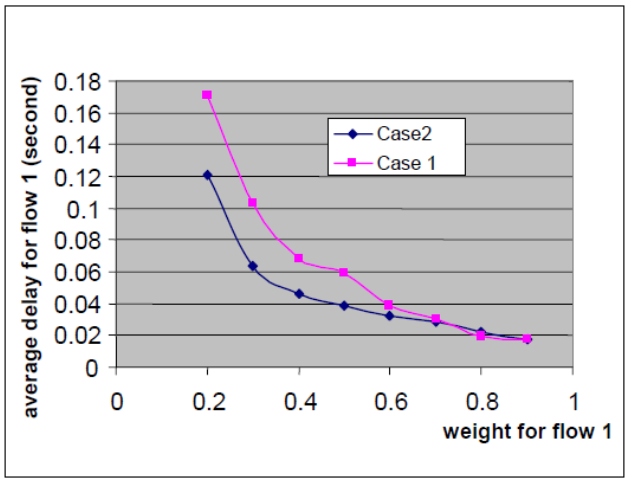

Figure 9. Average delay for the flow f with different weight assignments.

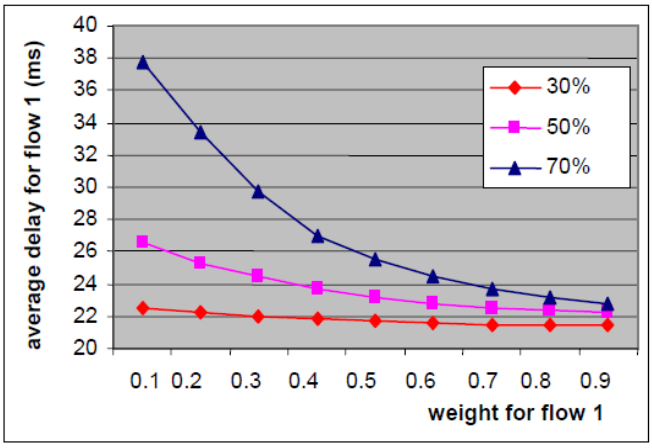

Figure 10. Average delay for flow f_1 with different weight assignments in the CICQ switch.

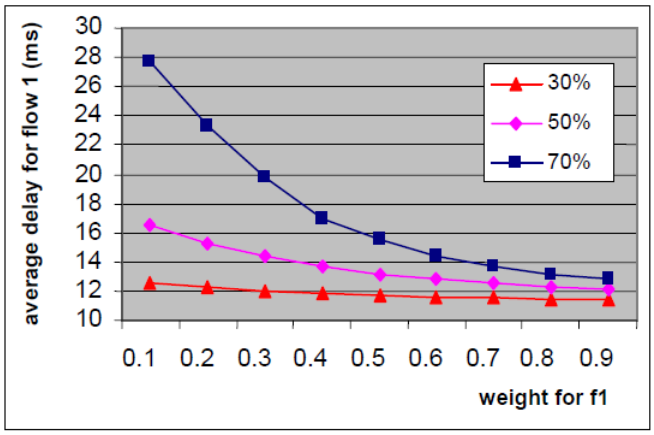

Figure 11. Average delay for flow $f_1$with different weight assignments in the OQ switch.

The WFQ is another major scheduling algorithm widely employed in packet switches. Traffic control in CICQ and OQ switches, both using WFQ algorithm, have also been simulated. The average delay for a traffic flow f_1 is measured with different assigned weights under 30%, 50%, and 70% traffic load. The obtained results for CICQ and OQ switches are plotted in Fig. 10 and 11 respectively. From these two figures we can see that under all traffic loads, the average delay decreases with the increment of assigned weight. The simulation results indicate that although having different average delay values, CICQ and OQ switches with WFQ algorithm achieve a similar delay pattern in that the average packet delay of both switches is a increasing function of traffic load and a decreasing function of the available switching capacity allocated to the traffic flow.

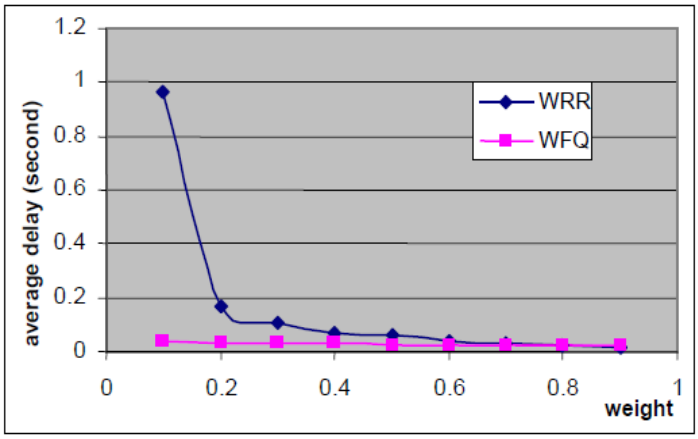

Figure 12. Delay performance comparison between WRR and WFQ schedulers.

Comparison is conducted between the average delay of the same flow achieved from WFQ and WRR schedulers with various assigned weights under 70% traffic load, and the results are given in Fig. 12. This figure shows that when the assigned weight is light (φ_1 <0.4), WFQ achieves better delay performance than WRR. Nonetheless, with increasing assigned weight, the delay difference decreases and becomes almost indistinguishable when assigned weight is large (φ_1 >0.6). This is because under a large assigned weight, most of the bandwidth is assigned towards the flow, the waiting time is no longer a factor and is replaced by the transmission time as the dominate factor for the delay. This implies that WFQ algorithm achieves almost equivalent performance as WRR algorithm does when one of the flows has a weight value that is much larger than the weight values assigned to other flows.

6. Conclusion

In this paper traffic control performance in Combined Input and Cross-point Queuing (CICQ) switch evaluated through simulations. Traffic control operations in CICQ switch with various scheduling algorithms including Round-Robin (RR), Weighted Round-Robin (WRR), and Weighted Fair Queuing (WFQ) are simulated. All these scheduling algorithms have been widely deployed in packet switches. We measured the average packet delay for traffic flows in CICQ switch and analyzed some impact factors on the achieved delay performance, including traffic load ratio and the assigned weights to different flows. Because the Output Queuing (OQ) switch has been shown as an ideal case with optimal performance, traffic control performance of CICQ switch is compared with that of the OQ switch. The obtained simulation results show that the average packet delay and the maximum queue length in a CICQ switch increase with traffic load. Performance comparison between CICQ and OQ switches shows that they have similar performance features in terms of average delay and queue length. The delay performance of CICQ switch approaches that of OQ switch when traffic load increases.

References

[1] A. K. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: the single-node case," *IEEE/ACM Transaction on Networking*, Vol. 1, Iusse 3, pp. 344-357, June 1993

[2] K. Yoshigoe and K. J. Christensen, "An evolution to crossbar switches with virtual output queuing and buffered cross points," *IEEE Network Magazine*, vol. 17, no. 5, pp. 48-56, Sept. 2003. http://dx.doi.org/10.1109/MNET.2003.1233917

[3] H. Zhang, "Service disciplines for guaranteed performance service in packet-switching networks," *Proceedings of the IEEE*, pp. 1374-1398, Oct. 1995. http://dx.doi.org/10.1109/5.469298

[4] M. Nabeshima, "Performance evaluation of a combined input and cross-point queued switch," *IEICE Transaction on Communications*, E83-B:3, pp. 737-741, March 2000.

[5] T. Javidi, R. Magill, T. Hrabik, "A high-throughput scheduling algorithm for a buffered crossbar switch fabric," *Proceedings of the 2001 International Conference of Communication*, pp. 1586-1591, Amsterdam, June 2001

[6] K. Yoshigoe and K. J. Christensen, "A parallel-polled virtual output queued switch with a buffered crossbar," *Proceedings of the 2001 Workshop on High Performance Switching and Routing*, pp. 271-275, Dallas, May 2001

[7] N. McKeown and T. E. Anderson, "A quantitative comparison of scheduling algorithms for input queued switches," *Computer Networks and ISDN Systems*, Vol. 30, Issue 24, pp. 2309-2326, Dec. 1998. http://dx.doi.org/10.1016/S0169-7552(98)00157-3

[8] R. Rojas, E. Oki and H. J. Chao, "CIXOB-k combined input and cross-point output buffered packet switch," *Proceedings of the 2001 Global Communications Conference*, pp. 2654-2660, San Antonio, Nov. 2001

[9] T. E. Anderson, "High-speed switch scheduling for local-area networks," *ACM Transactions on Computer Systems*, Vol. 11, Issue 4, pp. 319-352, Nov. 1993. http://dx.doi.org/10.1145/161541.161736

[10] S. T. Chuang, A. Goel, N. McKeown, and B. Prabhakar, "Matching output queuing with a combined input/output-queued switch," *IEEE Journal of Selected Areas in Communications*, Vol. 17, Issue 6, pp. 1030-1038, June 1999. http://dx.doi.org/10.1109/49.772430

[11] R. B. Magill, C. E. Robhrs, and R. L. Stevenson, "Output-queued switch emulation by fabrics with limited memory," *IEEE Journal of Selected Areas in Communications*, Vol. 21, Issue 4, pp. 606-615, May 2003. http://dx.doi.org/10.1109/JSAC.2003.810532

[12] Q. Duan and J. N. Daigle, "Resource allocation for quality of service provision in multistage buffered crossbar switches," *Elsevier Journal of Computer Networks*, vol. 46, no. 10, pp. 147-168, Oct. 2004.

[13] J. L. Boudec and P. Thiran, *Network Calculus: a Theory of Deterministic Queuing Systems for the Internet*, Springer Verlag, 2003

[14] Q. Duan, "Quality of Service Provision in High-Speed Packet Switches," Chapter 8 in book *Computer Networking and Networks*, pp. 189-237, Nova Science Publishers Inc., September 2006.

[15] Q. Duan, "Quality of Service Provision in Combined Input and Cross-point Queued Switches without Output Queuing Match," *Elsevier Journal of Computer Communications*, Vol. 30, Issue 4, pp. 830-840, Feb. 2007.

[16] Q. Duan, "Resource Allocation for Quality of Service Provision in Buffered Crossbar Switches with Traffic Aggregation," *International Journal of Computers and Applications*, Vol. 29, Issue 3, pp. 283-290, Aug. 2007.

Copyright Disclaimer