Chapter Five
Supporting High-Priority Traffic

In Sec. 3.5, we discuss uses for the queues of a multiqueue buffer other than simply increasing the saturation throughput of the network or reducing the average latency. There are situations where one may desire ‘‘especially good’’ service for packets of a particular class. Examples given in Sec. 3.5 include messages between kernel processes whose latency impacts a large number of user processes, messages controlling remote i/o devices which may have particular timing constraints (real-time communication), and flow control messages which must be able to bypass the network congestion which it is attempting to resolve. This chapter evaluates the ability of the DAMQ buffer to provide support for ‘‘high-priority packets’’ which require ‘‘especially good service’’.

There are two aspects to the ‘‘especially good’’ service needed by high-priority packets. First, the latency of packets in this class must be low — significantly lower than the latency of ‘‘normal’’ packets. Second, the latency of packets in this class must be predictable (especially when being used for real-time applications). To be predictable, packet in this class must be oblivious to the throughput of ‘‘normal’’ packets in the network *and* have a low variance in latency. In other words, given a percentage of packets being marked high-priority, the average latency of these packets should remain relatively constant as the throughput of normal packets is varied, and the worst-case latency of the high-priority packets should remain near the average latency as the throughput of normal packets is varied.

In Sec. 4.5, we presented figures showing average latency vs. throughput for multistage interconnection networks. This performance metric does not present the information necessary in order to evaluate the effectiveness of multi-queue support for high-priority packets. Presenting the average and maximum latencies of high-priority packets for a simulation, however, is inconclusive. The maximum latency experienced by a given simulation is an increasing random variable; the longer the simulation runs, the closer the maximum latency will approach its theoretical bound (if such a bound exists).

This chapter utilizes the same simulator that was used in Ch. 4, but the 99^{th} *percentile latency* is used as the metric of a network's support for high-priority traffic (in contrast to *average latency*). The 99^{th} percentile latency of packets for a given simulation is the lowest latency of the slowest one percent of the packets. While the maximum packet latency of a simulation continues to climb long after the simulation has reached steady state, the 99^{th} percentile latency of the simulation remains relatively constant.

The problem of high worst-case latency through a network is demonstrated by the results of simulations of a 64×64 Omega network composed of switches with FIFO input port buffers. As shown in Fig. 5.1, the maximum latency through the network may be many times larger than the average latency. Furthermore, if the buffer size is increased from four packets slots to eight packet slots, the ratio of maximum latency to average latency is not improved, nor is the ration of 99^{th} percentile latency to average latency. Hence, increasing the size of the FIFO buffers is not a solution to the problem of very large worst-case latency through the network.

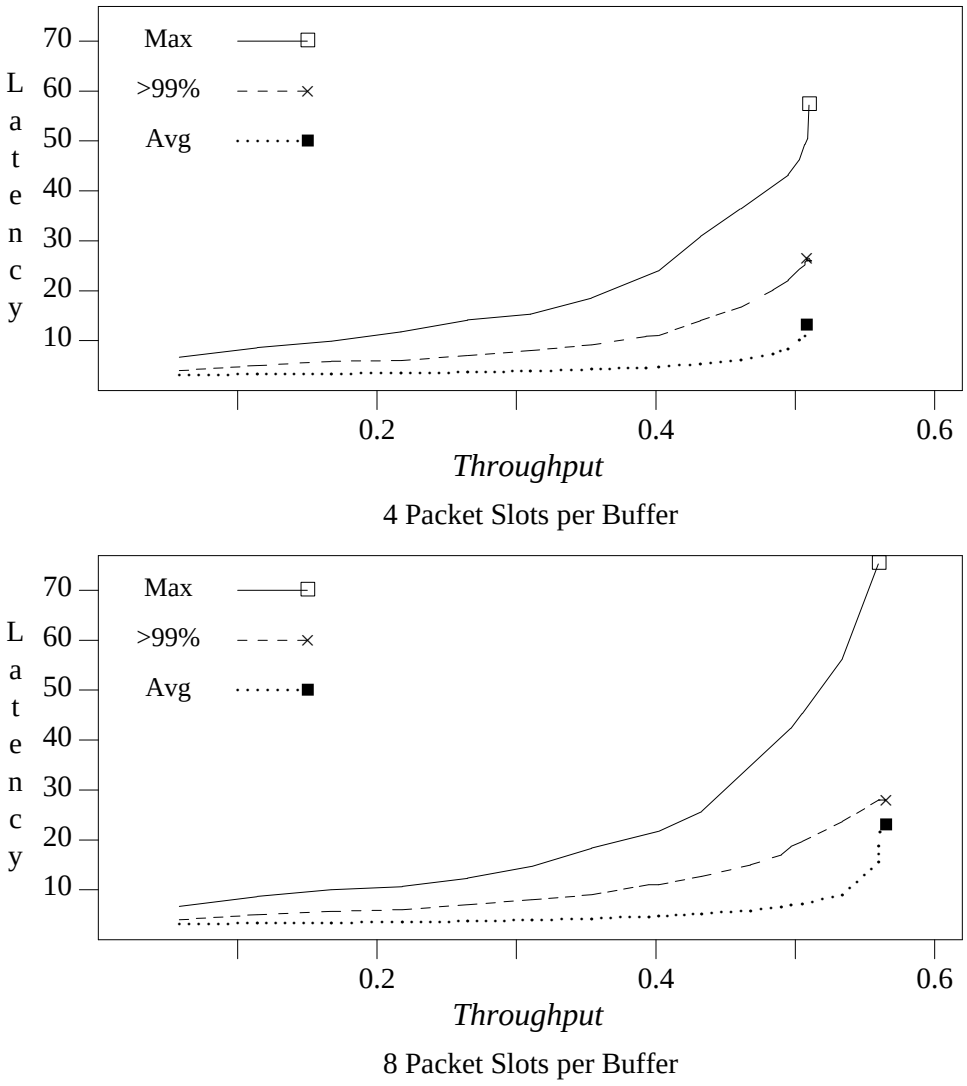


Figure 5.1: *Variance of Latency, FIFO Switches.* Maximum, 99th and Average Latency vs. Throughput. Four packet slots per buffer.

We have previously shown that networks composed of DAMQ, SAMQ, or SAFC switches perform significantly better than networks composed of FIFO switches (Figure 4.1). As shown in Table 5.1, these three practical switches as well as the CBDA switch perform better than FIFO switches in terms of worst-case

	Buffer Type	Throughput									
		0.10		0.20		0.30		0.40		0.50	
> 99% / avg	FIFO	4.75	8.10	5.95	11.11	7.78	15.03	10.97	23.77	23.48	45.08
		3.14		3.38		3.79		4.65		9.34	
	SAMQ	5.76	9.80	6.75	13.88	9.00	19.79	12.00	26.74	17.88	39.25
		3.24		3.58		4.09		4.90		6.57	
	SAFC	5.38	10.10	6.73	13.78	8.16	20.06	11.00	21.60	14.38	32.83
		3.22		3.50		3.88		4.42		5.28	
	DAMQ	4.76	7.76	5.67	10.24	7.00	13.74	8.88	16.87	11.11	22.27
		3.14		3.36		3.68		4.16		4.91	
	CBDA	4.39	6.43	5.00	8.11	6.00	9.60	7.00	11.75	8.00	14.40
		2.76		3.29		3.50		3.80		4.19	

Table 5.1: *Variance of Latency, All Switches.* Maximum, 99th Percentile and Average Latencies vs. Throughput. Four slots per input port (a total of sixteen slots in the CBDA switch). Uniformly distributed traffic, 100% of packets normal.

packet latency under heavy load (throughput). However, even for these ‘‘better’’ switches the maximum (and 99th percentile) latencies are several times larger than their average latency. Even for the CBDA switch, which attempts to achieve fairness by arbitrating between incoming packets based upon the amount of time spent in the previous switch, the 99th percentile latency is double the average latency at 50% throughput, and the maximum latency is double that. These results indicate that the problem of large variations in the latency of packets, some of which may be ‘‘high priority,’’ cannot be solved by using different buffer organizations without special support for the high-priority packets.

In the remainder of Ch. 5 we investigate switch designs that provide support for high-priority packets. Alternatives designs are presented in the order of

increasing complexity and hardware cost. For each switch organization we describe specific design considerations, simulations results, and a brief summary of the major conclusions from the simulation results. The next subsection describes *priority arbitration*, in which the arbiter gives priority to queues which have high-priority packets at their head. Section 5.2 introduces the use of dedicated queues for high-priority packets (the support mechanism discussed in Ch. 3). Multiple dedicated queues for high-priority packets, one queue per output port, are discussed in Section 5.3. Finally, Sec. 5.4 evaluates the use of dedicated buffers for the high-priority packets (i.e. two physically distinct buffers per input port, one dedicated to high-priority packets).

5.1. Arbiter Support for High-Priority Packets

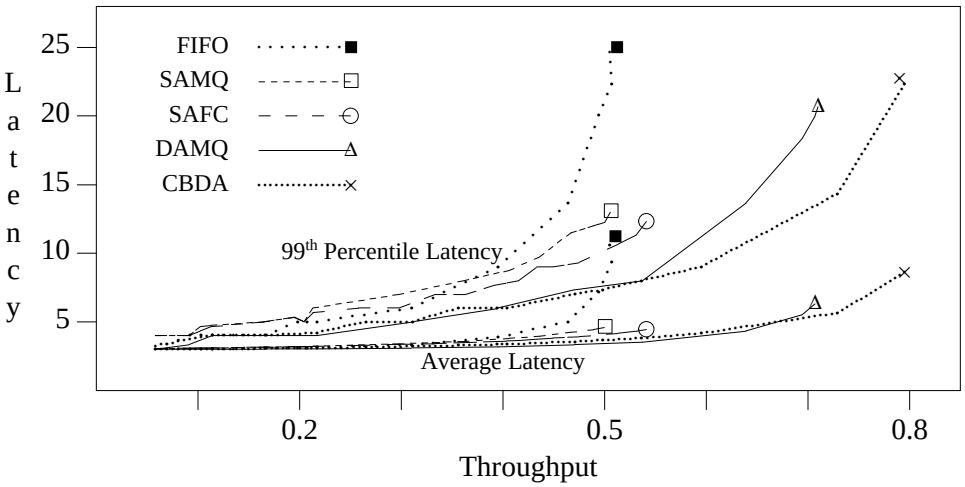
One way to support high-priority packets is to modify the crossbar arbiter — the controller that determines which input buffer is connected to which output port. The purpose of such a modification is to provide preferential arbitration to high-priority packets. Specifically, when determining the crossbar configuration, the modified arbiter will give priority to those queues which have a high-priority packet at their head. At each cycle, the arbiter first attempts to connect queues with high-priority packets at their heads to output ports, and then connects queues with normal packets at their heads to any remaining available output ports. For the CBDA switch, support for high-priority traffic in the arbiter means that if multiple packets destined for the same output port arrive simultaneously, the high-priority packets will be placed in the queue for that output port ahead of any normal packets. Furthermore, if the number of available slots in the buffer of the CBDA

switch is less than the number of senders to the switch, senders with high-priority packets at the head of their queue will be given priority.

In order to evaluate the effectiveness of the above scheme we ran simulations of the 64×64 omega network where a randomly selected 5% of all the packets generated were marked as high priority. Figure 5.2.a compares the 99^{th} percentile latency to the average latency of the high-priority packets only in networks using priority arbitration. Figure 5.2.b compares the 99^{th} percentile latency of high-priority packets for networks where the switches provide preferential arbitration for high-priority traffic to the 99^{th} percentile latency of packets in a network with no support for multiple packet classes. The results in Fig. 5.2.b indicate that, especially for networks with non-FIFO buffers, the modified switches improve high-priority performance over a system with no special support. However, as shown in Fig. 5.2.a, the 99^{th} percentile latencies are still much worse than the average latencies, especially under high network load. Supporting for high-priority packets in the arbiter alone is not sufficient; since the high-priority packets often wait in queues behind low-priority packets and receive preferential treatment for only the brief time they spend at the heads of queues. There must be a mechanism for the high-priority packets to bypass the normal packets.

5.2. Dedicated Queues for High-Priority Traffic

The SAMQ, SAFC, DAMQ, and CBDA switches achieve low average latencies by employing multiple queues of packets, one queue associated with each output port, to prevent packets which are destined for different output ports from blocking each other. A possible extension of this idea is to implement an



a) Priority arbitration, comparing 99^{th} percentile and average latencies.

	Buffer Type	Throughput				
		0.10	0.20	0.30	0.40	0.50
>99% Latency $\frac{prior.\ arbit.}{normal\ switch}$	FIFO	4.00	5.00	5.89	9.34	21.08
		4.75	5.95	7.78	10.97	23.46
	SAMQ	4.55	5.13	7.05	8.65	12.25
		5.76	6.75	9.00	12.00	17.88
	SAFC	4.27	5.15	6.06	7.78	10.21
		5.38	6.73	8.16	11.00	14.38
	DAMQ	3.59	4.00	4.89	6.09	7.63
		4.76	5.67	7.00	8.88	11.11
	CBDA	3.81	4.13	5.00	6.00	7.37
		4.37	5.00	6.00	7.00	8.00

b) Priority arbitration and no support for priority, comparing 99^{th} percentile latencies between the two.

Figure 5.2: *Priority Arbitration Performance.* Preferential arbitration for high priority packets. Four slots per input port. 5% high-priority packets. Throughputs shown are *total* network throughputs.

additional queue at each input port dedicated to high-priority packets. This should allow high-priority packets to be routed through the switch as soon as they arrive in preference to any normal packets waiting to be transmitted out of the same output port. This idea cannot be applied to a FIFO buffer since it has only a single queue; FIFO switches with the priority arbitration scheme discussed above are used in the following comparisons.

Implementing a dedicated queue for high-priority packets in the SAMQ and SAFC buffers involves adding additional storage (at least one packet slot) as well as a slight increase in the complexity of the arbitration circuitry. Since each queue in a SAMQ or SAFC buffer has at least one packet slot, the minimum size SAMQ or SAFC buffer in a 4×4 switch with support for high-priority traffic is five packet slots. On the other hand, adding the dedicated queue to the DAMQ buffer does not require any additional buffer storage — the available storage is dynamically shared between queues. A few (two or three) more control registers are needed to manage an additional linked list. In order to present a fair comparison of the performance of networks with the different buffer types, many of the results presented in this section were obtained using a total buffer size of five packet slots per input port. For the CBDA switch, special queues for high-priority traffic mean that the switch maintains, in its central buffer, a separate FIFO queue for each priority level at each output port. For each output port, packets are sent from the queue of normal packets only if the high-priority queue is empty. Furthermore, if the number of available slots in the buffer of the CBDA switch is less than the number of senders to the switch, senders with a non-empty high-priority queue are given priority.

We have simulated 64×64 omega networks constructed out of switches which

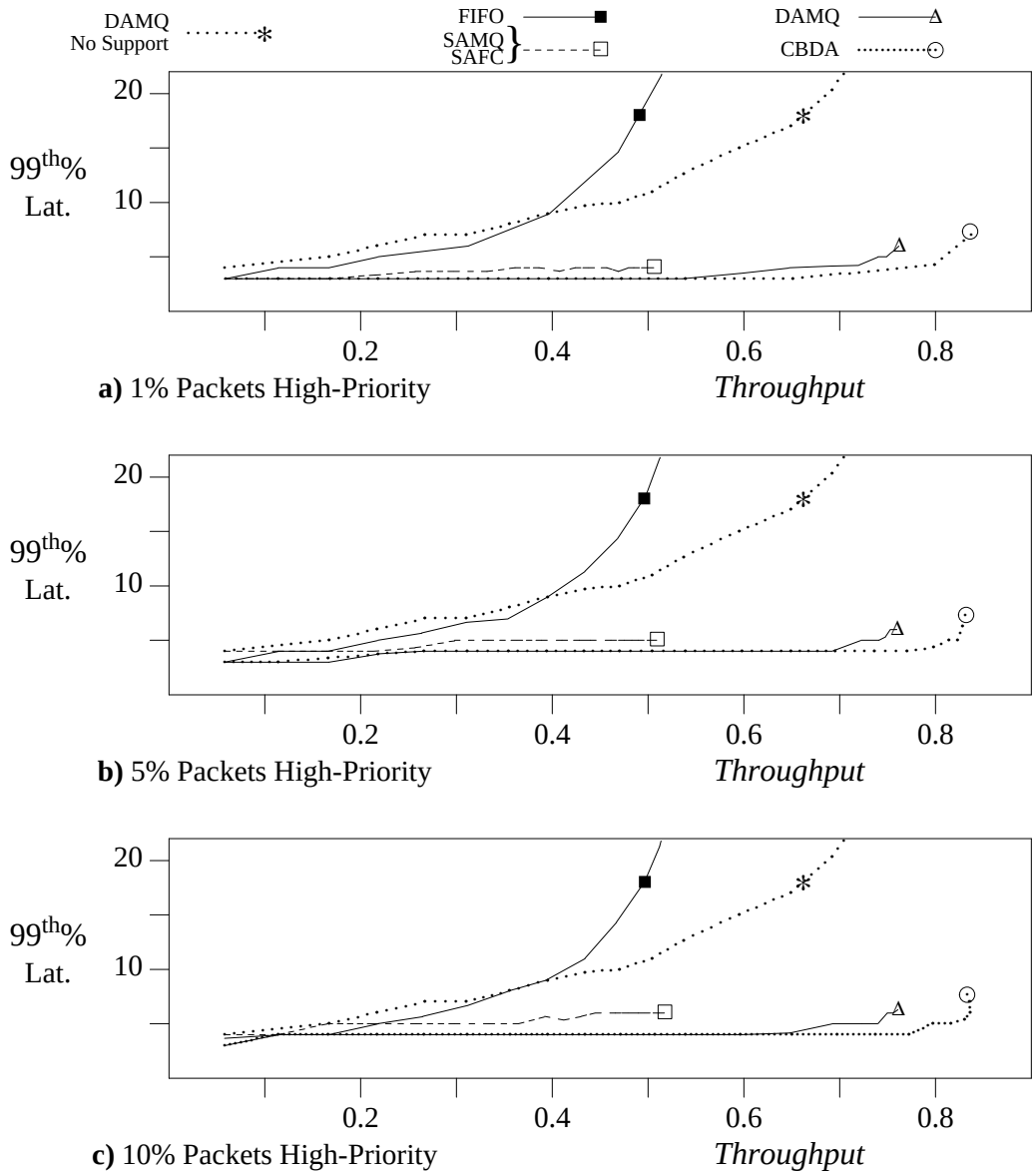


Figure 5.3: *Priority Queue, Various Percentages of High-Priority Packets.* The impact of a dedicated queue for high-priority packets at each input port. Shown are the 99th percentile latencies of high-priority packets vs. total throughput. Five slots per input port (a total of twenty packet slots in the CBDA switch).

include a dedicated queue for the high-priority packets at each input port. The 99th percentile latencies of the high-priority packets for networks with 1%, 5% and 10% high-priority traffic are shown in Fig. 5.3. For comparison, the figure also shows the 99th percentile latencies of a network with five-slot DAMQ buffers with no support for high-priority packets. The DAMQ buffer was chosen since, as shown in Tab. 5.1, this buffer results lower 99th percentile latencies than the other practical switches. These simulation results demonstrate the benefits of dedicated high-priority queues. Networks composed of the three practical switches which can support a separate high-priority queue achieved significantly lower 99th percentile latencies compared to those networks without separate high-priority queues. Furthermore, the 99th percentile latencies for high-priority packets are lower than the *average* latencies with switches that do not have support for high priority traffic (Figure 4.1).

Switches with DAMQ buffers outperform the SAMQ and SAFC switches by achieving lower 99th percentile latencies for high-priority traffic. The DAMQ switch is better at handling high-priority packets for the same reason it provides lower average latency for normal traffic — while only a fraction of the buffer space of the SAMQ and SAFC buffers is available to any single packet arriving at an input port (one fifth, in this case), the entire buffer is available to packets arriving at the DAMQ switch. Thus, the DAMQ switch better utilizes the available space, blocks incoming packets less often, is capable of handling higher throughputs, and delivers packets with lower latencies.

Figure 5.4 shows the 99th percentile latency of high-priority packets in a DAMQ network with high-priority queues, with priority arbitration, and with no

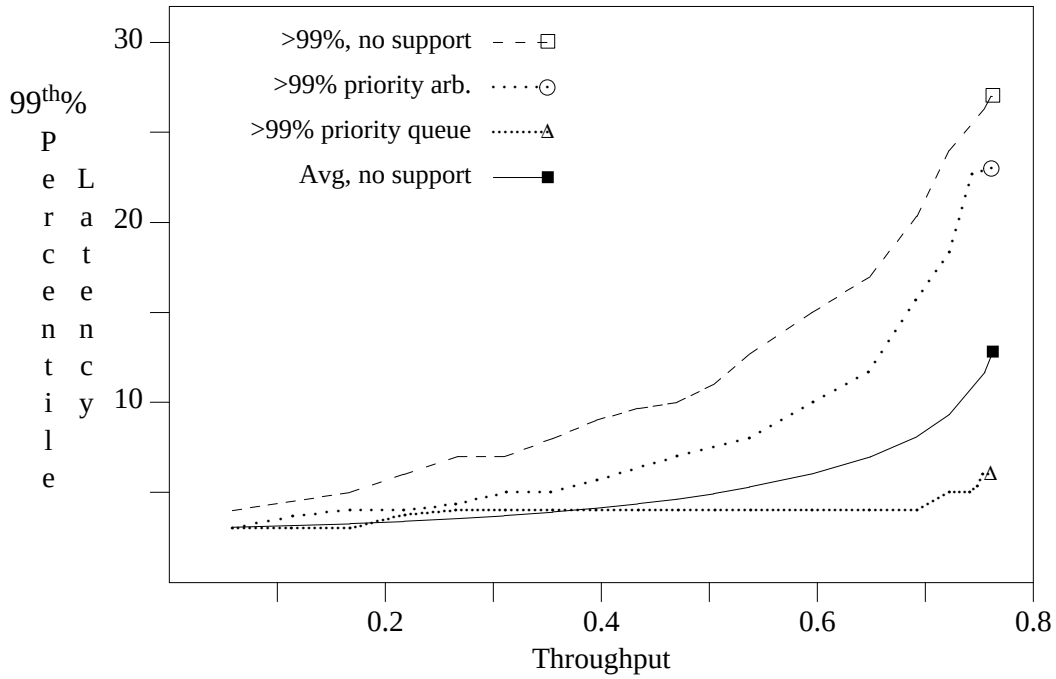


Figure 5.4: *Increasing Levels of Support for High-Priority Packets.* The average latency for normal traffic and the 99^{th} percentile latency for high-priority traffic without support, with priority arbitration and with priority queues vs. total network throughput. DAMQ buffers with five slots per input port. 5% high-priority traffic.

support for high-priority traffic. The average latency in a DAMQ network with no high-priority support is also shown. In all cases the buffers have five packet slots and 5% of the packets are high priority. With dedicated queues for high-priority traffic, the 99^{th} percentile latency for this traffic is stable at near the minimum network latency for a wide range of throughputs. With the dedicated queues, at medium and high throughputs, the 99^{th} percentile latency of the high-priority traffic is actually less than the average latency for normal traffic. An interconnection network with switches using DAMQ buffers with a high-priority queue can thus provide low worst-case latencies for high priority packets even

under network loads approaching saturation.

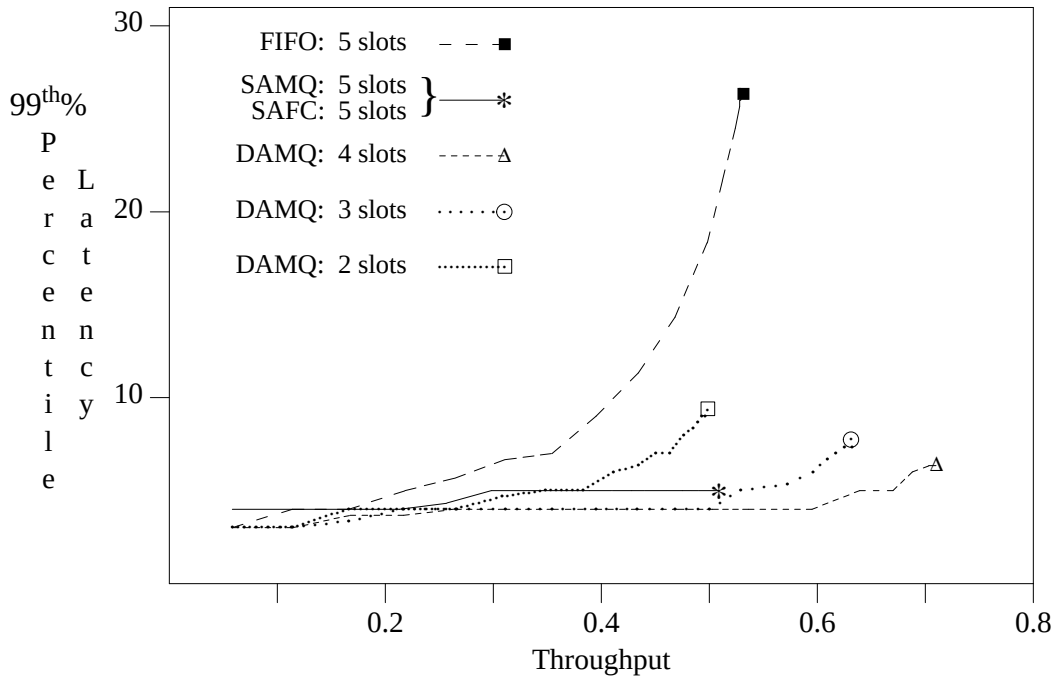


Figure 5.5: *Priority Queue, DAMQ Buffer Efficiency.* The 99th percentile latency of high-priority traffic vs. total network throughput. FIFO, SAMQ, and SAFC buffers with five packet slots per input port, DAMQ buffers with four, three and two packet slots per input port. 5% high-priority traffic. Dedicated queues for high-priority packets.

The advantages of the DAMQ buffer are further demonstrated by the results shown in Figure 5.5. The 99th percentile latencies of a network composed of DAMQ switches with two, three, and four slots are compared with the 99th percentile latencies for networks composed of FIFO, SAMQ, and SAFC switches with five slots. The DAMQ switch with only two slots significantly outperforms the FIFO switch with five slots. The DAMQ switch with three slots outperforms the SAMQ and SAFC switches with five buffer slots.

Figure 5.6 shows the impact of buffer size on the effectiveness of using dedicated queues for high-priority traffic with DAMQ buffers. For all buffer sizes, the dedicated queues result in significantly better performance for the high-priority packets than in a network where there is no special support for such packets. With very small buffers (two packet slots), even with dedicated queues, the 99^{th} percentile latency of the high-priority packets increase significantly as the total network throughput increases above 0.3. Furthermore, this 99^{th} percentile latency of the high-priority traffic is always higher than the average latency in a network without the dedicated queues. The reason for this rather poor performance with two slot buffers is that at most two queues can exist in the buffer at any stage cycle.

The maximum *total* throughput of the network increases significantly as the buffer size is increased from two to four slots and then six slots (Figure 5.6). With the larger buffers, more queues can exist in the buffer simultaneously and, at any point in time, the probability of a full buffer is lower than with two slot buffers. Hence, with dedicated queues for high-priority packets, it is much less likely that such a packet will be blocked due to a full buffer in the next stage. As a result, with four slot and six slot buffers with dedicated queues for high-priority traffic, the 99^{th} percentile latency of the high-priority packets remains very low (4 stage cycles) even as the network throughput increases up to 0.6. Furthermore, this 99^{th} percentile latency of the high-priority traffic is always *lower* than the average latency in a network without the dedicated queues.

This subsection has demonstrated the benefits of using dedicated queues for high-priority packets. DAMQ buffers were shown to be a particularly effective mechanism for exploiting such dedicated queues. We have now shown that

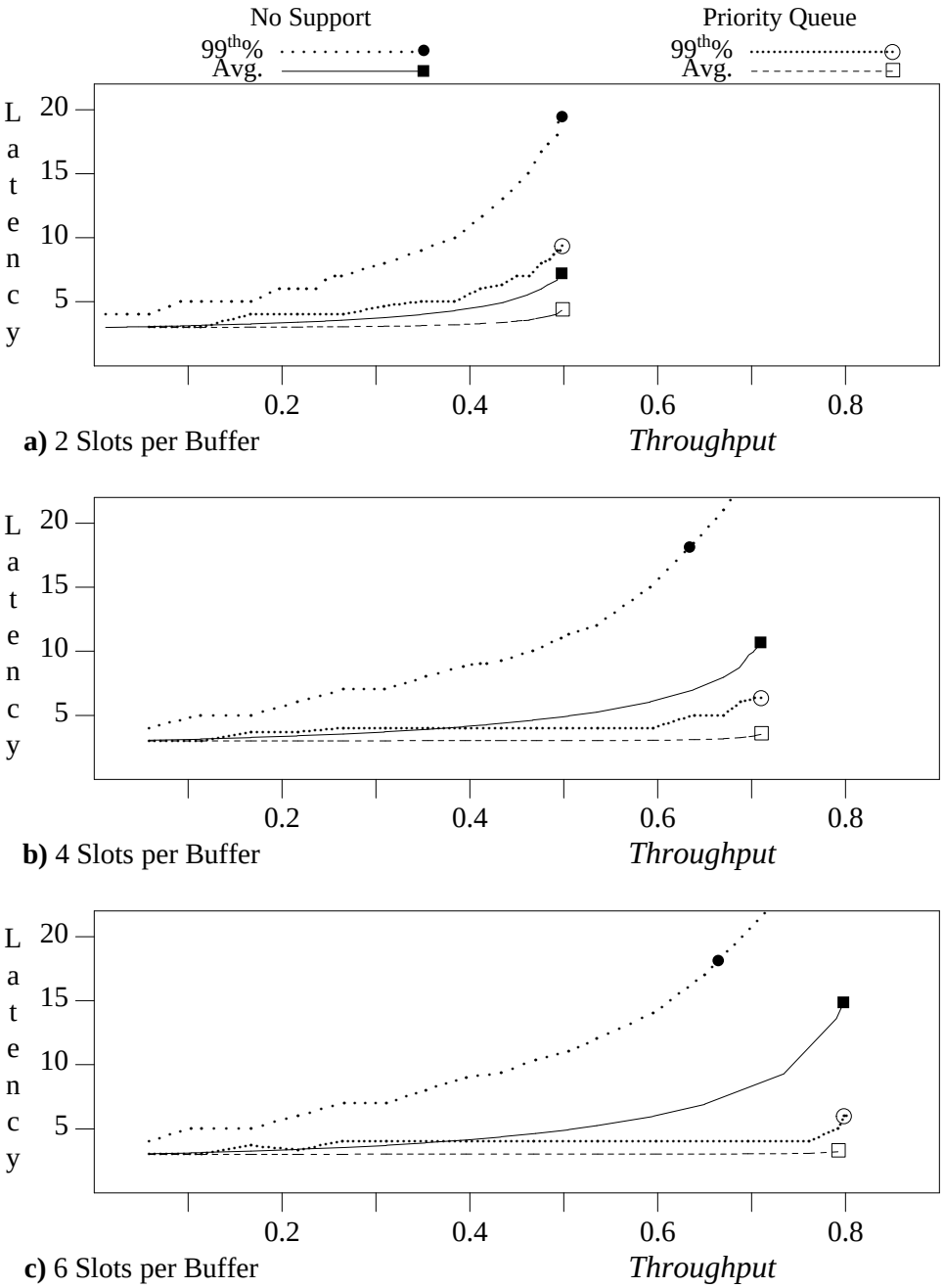


Figure 5.6: *Impact of Varying Buffer Size, DAMQ Buffers.* The average and 99^{th} percentile latencies for the 5% high priority traffic. Network with no support for high-priority packets and network with dedicated queues.

DAMQ buffers achieve higher performance than other practical buffers for high-priority traffic and for normal traffic (Sec. 4.5). Thus, for the rest of this chapter, DAMQ buffers with dedicated queues for high-priority traffic, will serve as a benchmark against which alternative designs will be compared.

5.3. Multiple Dedicated Queues for High-Priority Traffic

The previous subsection proposes support for high-priority traffic using, at each input port, a DAMQ buffer with a separate queue for high-priority packets. For normal traffic, DAMQ buffers achieve superior performance to FIFO buffers by, at each input port, providing a separate queue for each output port [Tami88b]. Hence, it should be possible to achieve better performance for high-priority traffic by providing, at each input port, a dedicated queue for high-priority traffic for each output port, in addition to the separate queues for normal traffic. With this scheme, for a 4×4 switch, each buffer will contain eight queues, instead of the five queues proposed in the previous subsection. The additional hardware required by this scheme may be worthwhile for applications where minimizing worst-case latency is particularly critical. SAMQ and SAFC switches with multiple dedicated queues for high-priority traffic require even more hardware since additional buffer storage is needed for each additional queue. Since, in addition, DAMQ switches with high-priority queues have already been shown to achieve superior performance compared to SAMQ and SAFC switches, only DAMQ switches are considered in this subsection.

Figure 5.7 compares the 99^{th} percentile latencies of the high-priority traffic with eight-queue and five-queue DAMQ switches, varying the total throughput and

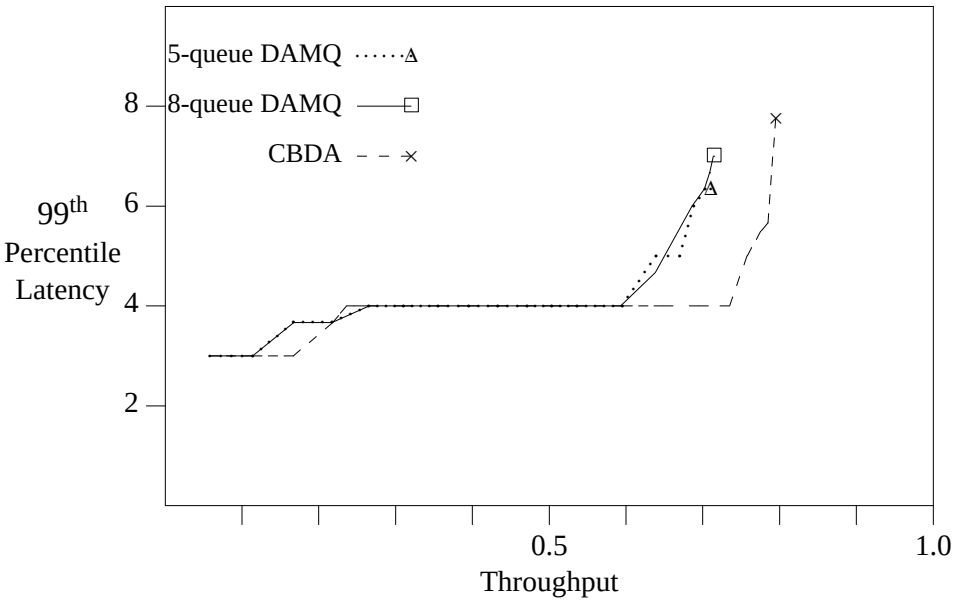


Figure 5.7: *Five-Queue and Eight-Queue DAMQ Buffers.* The 99th percentile latency of high-priority traffic vs. the total network throughput. Four packet slots per input port. 5% high-priority traffic.

keeping the percentage of high-priority packets constant at 5%. The performance of a network with CBDA switches is shown as a reference point for the level of performance that can be achieved. Under these conditions, the two DAMQ switches behave almost identically. The reason for this is that multiple queues are beneficial only when there are more than one packet in the buffer simultaneously. With only 5% high-priority packets it is rare that more than one high-priority packet is in a buffer at any instant.

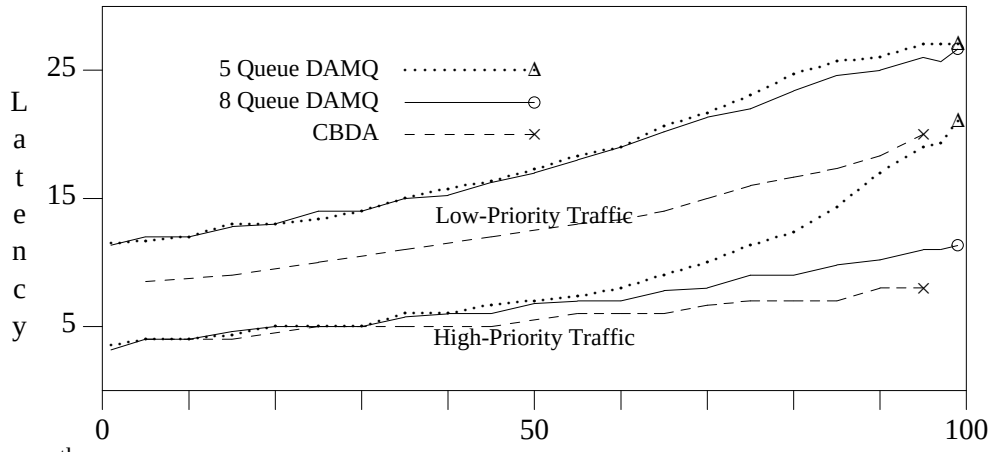
In a network constructed from switches that support high-priority traffic, it is useful to know up to what level (proportion) of high-priority traffic will the separate queues be able to maintain good worst-case performance, i.e., low 99th percentile latency, for the high priority packets. To answer this question we have

simulated the 64×64 omega network of four-slot DAMQ buffers under a constant total network throughput of 50% with varying proportions of high-priority traffic. It should be noted that without support for high-priority traffic, a network of DAMQ switches operating at this throughput achieves an average latency of 4.9 stage cycles and a 99th percentile latency of 11.1 stage cycles (Table 5.1).

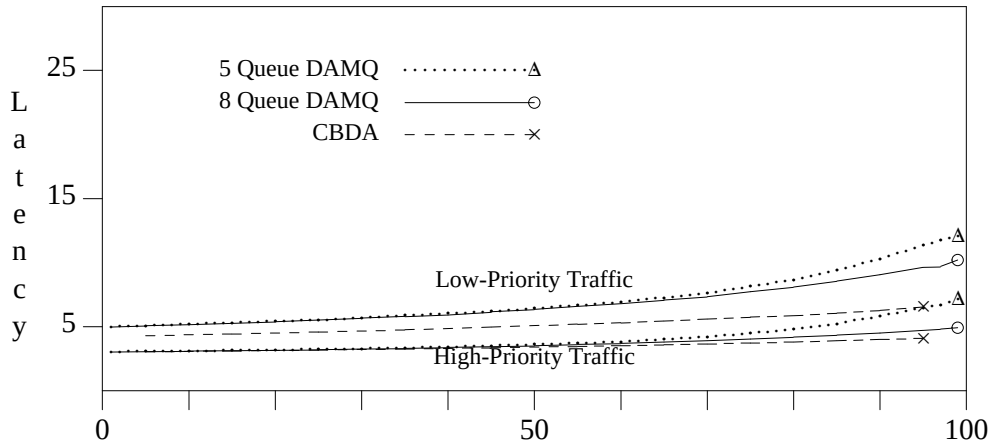
Figure 5.8 shows the 99th percentile and average latencies of both the high-priority and low-priority (normal) traffic for a varying percentage of high-priority traffic. For a low percentage of high-priority packets, the five-queue DAMQ buffers perform well, i.e., result in low 99th percentile latency for the high priority traffic. With up to 18% high priority packets, the 99th percentile latency for such packets is below the overall average latency for a standard network using switches with DAMQ buffers (Table 5.1).

The average latencies and 99th percentile latencies of the low priority (normal) packets (Figure 5.8) indicate that, for low percentages of high-priority packets, the support for high-priority traffic has little impact on network performance with respect to normal traffic. Specifically, as the percentage of high-priority traffic varies from 0% to 10%, the change in average latency of the low priority traffic is negligible.

As the proportion of high-priority packets increases beyond 75%, the 99th percentile latency of a network with the five-queue DAMQ buffers is *larger* than 11.1 stage cycles, i.e., larger than the 99th percentile latency for the network composed of switches without support for high-priority traffic (Table 5.1). The reason for this behavior is that the DAMQ switch has four queues at each input port for normal traffic, but only a single queue for high-priority traffic. Hence, with



a) 99th Percentile Latency



b) Average Latency

Figure 5.8: *Varying Percentage of High-Priority Packets.* Shown are 99th percentile latencies and average latencies of high-priority and low-priority traffic for a fixed total network throughput of 50%. Four packet slots per input port. Five-queue DAMQ switches, eight-queue DAMQ and CBDA switches.

respect to high-priority traffic, the DAMQ switch with a high-priority queue performs as a FIFO switch. As shown in Fig. 5.1.a, at a throughput of 0.37, the 99th percentile latency of a network with FIFO switches is 10 stage cycles. With the network of DAMQ switches with high priority queues, 0.50 network

throughput with 75% high priority traffic, the absolute throughput of high priority traffic is 0.375. Since part of each input buffer is occupied by normal packets, it is not surprising that the high priority traffic receives worse ‘‘service’’ (higher 99^{th} percentile latency) than the packets in the FIFO network with a throughput of 0.37.

Figure 5.8 shows that when more than 50% of the packets are high-priority, the eight-queue DAMQ switch performs significantly better than the five-queue DAMQ buffer, maintaining a relatively low 99^{th} percentile latency for the high-priority traffic. Furthermore, the improved performance with respect to high-priority packets does *not* come at the expense of the performance for low-priority packets. If the high-priority traffic is less than 50% of the total network load, there are no benefits to using the eight-queue DAMQ buffers. Eight-queue DAMQ buffers should be considered only if the *majority* of the packets are high priority. In this latter case, a better solution might be to use five-queue DAMQ buffers with a single queue dedicated to the low-priority traffic and four queues dedicated to high-priority traffic.

5.4. Dedicated Buffers for High-Priority Traffic

In the previous two subsections we have considered the use of one or more dedicated queues for high-priority packets and demonstrated the advantages of using DAMQ buffers that include such queues. With the DAMQ buffers, the normal packets and high-priority packets share the available buffer storage, even though they are organized in separate queues. Another design option for supporting high-priority traffic is the use of two independent buffers at each input port — one for normal packets and one for high-priority packets. Examples of

such a scheme include the use of an n slot DAMQ buffer for normal packets together with an m slot FIFO buffer for high-priority packets, or an n slot DAMQ buffer for normal packets with an m slot DAMQ buffer for high-priority packets. The motivation for such a scheme is to isolate the high-priority buffering from the normal packet buffering. This would prevent a situation in where a high-priority packet could not be forwarded to the next stage due to a full buffer containing normal packets. As with high-priority queues, for a synchronous network, the high-priority traffic does not compete for link bandwidth with the low-priority traffic since, whenever there is a conflict, the high-priority traffic is always given priority over the low-priority traffic. Thus, with dedicated buffers for high-priority traffic, network performance with respect to such traffic will be *independent* of low priority traffic.

We have evaluated schemes for using dedicated buffers for high-priority packets by comparing their performance, with respect to high-priority traffic, to the performance of a network of DAMQ switches where the buffer storage is *shared* between the high-priority traffic and the normal traffic (Sec. 5.2). For a varying percentage of high-priority traffic, the comparisons were done for a fixed total network throughput of 50% (at which, as shown in Fig. 4.1, a network of switches with four-slot FIFO buffers is already in saturation). We have previously shown that the five-queue DAMQ buffer with the dedicated high-priority queue provides good support at a low implementation cost for high-priority traffic. At 50% total network throughput, the network composed of switches using these buffers is quite heavily loaded, so contention for buffer storage between the high-priority and low-priority traffic might be expected. Hence, this network serves as a good

benchmark, relative to which dedicated buffer schemes can be compared.

We have considered several input port organizations based on a four-slot four-queue DAMQ buffer for normal packets and a dedicated buffer for high-priority packets. DAMQ buffers for normal packets are assumed since they have been shown to have significant performance benefits for normal traffic [Tami88b]. In choosing the buffer for high-priority traffic, we must consider the hardware cost for additional storage and control, as well as performance. Single-slot FIFO buffers are a low-cost choice for the dedicated high-priority buffers. Multi-slot high-priority buffers can improve the performance of dedicated high-priority buffer schemes, but require more hardware. In particular, we have considered the use of dedicated two-slot FIFO buffers and two-slot DAMQ buffers for high-priority traffic.

Figure 5.9 shows the performance of networks with the dedicated high-priority buffer schemes described above as well as the performance of networks with DAMQ buffers which include dedicated queues for high-priority traffic but where storage is shared by all traffic. The simulation results presented are for a fixed total network throughput of 50%. For varying percentage of high-priority traffic, Figure 5.9.a shows the 99^{th} percentile latency of high-priority packets, while Fig. 5.9.b shows the 99^{th} percentile latency of normal packets. Since none of the three shared DAMQ buffers is in its saturation region, they all result in identical performance. With a one slot dedicated buffer, handling of high-priority packets is poor, despite using more total buffer space than the shared four-slot DAMQ buffers. Performance with respect to the normal packets is identical with all the schemes for the proportion of high-priority packets of interest (under 30%).

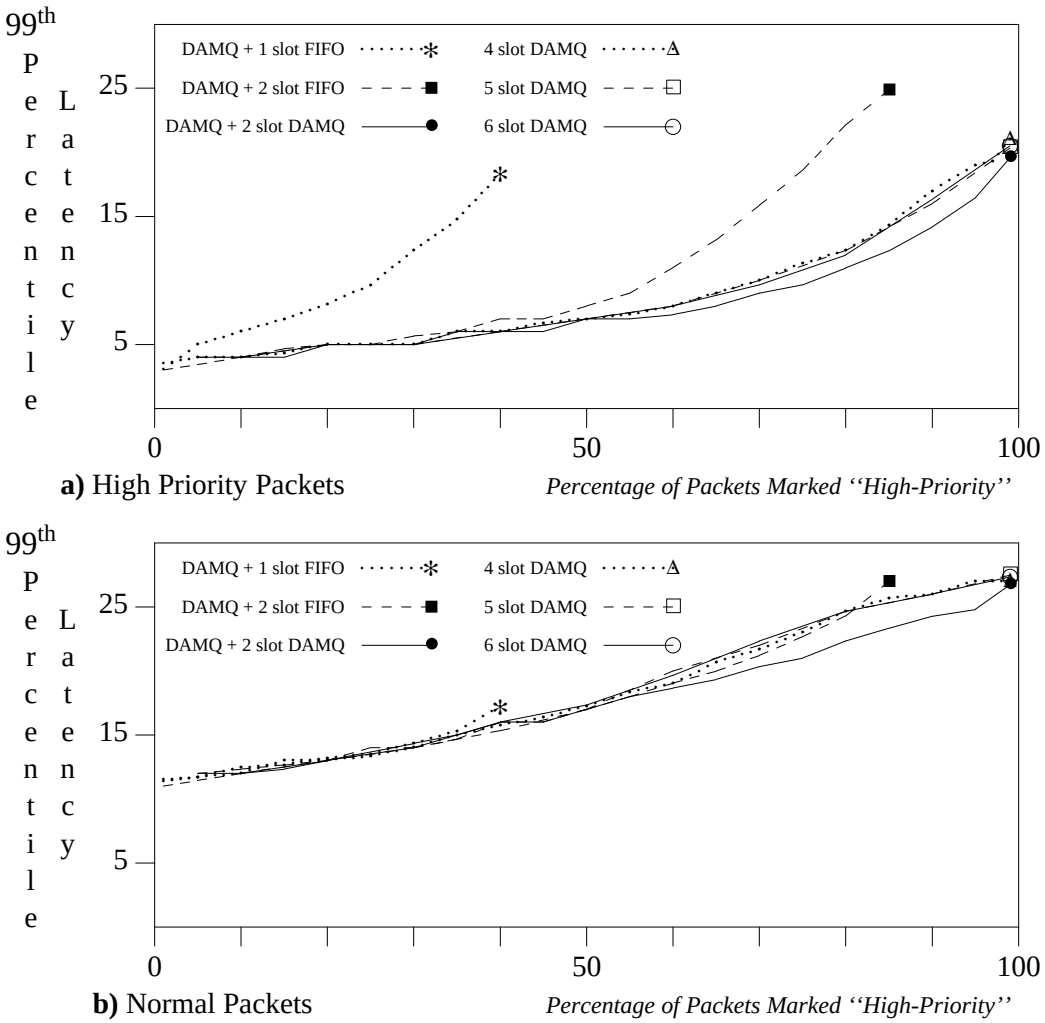


Figure 5.9: *Dedicated Buffers for High-Priority Packets.* The 99th percentile latency of high-priority and of normal packets vs. the percentage of high-priority traffic. The total network throughput is 0.5. Comparing dedicated buffer (4-slot, 4-queue DAMQ plus another) to dedicated queue (5-queue DAMQ) support.

At a low percentage of high-priority traffic (under 30%), the shared DAMQ buffers deliver equal performance to the performance of the two slot dedicated buffer schemes, despite the fact that the shared buffers are also supporting the total network throughput of 0.5. Since DAMQ buffers are already used to achieve high

performance for normal traffic, very little additional hardware is required to support dedicated queues. Dedicated two-slot buffers require significantly more hardware for both storage and control. In addition, the two buffers schemes are less likely to adapt well to a variety of conditions (such as very high throughput with *no* high-priority packets) due to the *static* partitioning of the available storage. Thus, dedicated buffers for high-priority traffic should be considered only if there is some reason why DAMQ buffers cannot be used. Under these conditions, it is useful to know that similar performance, with respect to high-priority traffic, can be achieved using this alternative approach.

5.5. Summary: Support for High-Priority Traffic

In order to use large multiprocessors and multicomputers for real-time applications, it must be possible to guarantee, with a high degree of confidence, that high-priority traffic can be transmitted through the network with low specified latency. With conventional interconnection networks used in multiprocessors and multicomputers, the worst-case latency of traffic through the network can increase dramatically as the load on the network increases. This may prevent the use of these interconnection networks for critical real-time applications or force their use with very low utilization (and thus high cost/performance ratio) in order to guarantee low maximum latency.

For interconnection networks composed of small $n \times n$ switches, we have shown that simply increasing the size of conventional buffers in the switches does not result in improved performance for high priority packets. There is thus a fundamental need for new buffer organizations which support high-priority

packets. We have developed a technique for efficiently supporting high-priority traffic, while maintaining good performance for normal traffic. Our scheme is based on a small modification of the *dynamically allocated multi-queue* (DAMQ) buffer. The modifications to the DAMQ buffer in order to support high-priority traffic involve a few additional control registers and somewhat more complex arbitration of the crossbar switch. Overall these modifications are expected to require only a small percentage increase in total buffer area.

We have evaluated alternative approaches to providing support for high-priority packets in $n \times n$ switches. This evaluation was based on implementation complexity and simulation studies of a multistage interconnection network transmitting packets with two levels of priority. Our simulations have demonstrated five key points. (1) In a conventional network with a single priority level for all packets, worst-case latency for packets can be several times higher than average latency. Hence, there is a need to identify and provide preferential treatment to those packets for which fast service is particularly important. (2) Using the priority of packets as the determining factor in arbitrating contention within each switch does not, on its own, provide sufficient support for high-priority traffic. Such arbitration does not reduce the 99^{th} percentile latency of the high-priority packets to the level of average normal packet latency, even under moderate network load. (3) As long as the proportion of high-priority traffic is low, dedicated queues for high priority traffic reduce the 99^{th} percentile latency of the high-priority packets to the level of the average latency for normal traffic under a wide range of network throughputs. (4) When the proportion of high-priority traffic is large and the network is heavily loaded, multiple dedicated queues for

high-priority traffic can reduce the 99th percentile latency relative to the single dedicated queue approach. However, this performance advantage is marginal and the associated implementation cost is relatively high. (5) Dedicated buffers for the high-priority traffic can provide the same performance advantage as dedicated queues in shared DAMQ buffers. However, since DAMQ buffers are needed to maximize performance for normal traffic, the implementation cost for dedicated buffers is much higher than the cost of adding dedicated queues to DAMQ buffers. Hence, there is no reason to use dedicated buffers.

Our results indicate that the DAMQ buffer with a single dedicated queue for high-priority packets provides support for high-priority traffic which is superior to the support provided by alternate switch designs based on a dedicated high-priority queue. This resulting network performance, with respect to high-priority traffic, is very close to the performance of a network based on ‘‘ideal’’ switches for which a practical implementation is not feasible. Hence, given the low hardware overhead of the scheme based on a DAMQ buffer with a single high-priority queue, it is clearly the preferable option.