Dual Queue Approach to Improving Network Performance During Transient Congestion Episodes

David A. Hayes[*], Lachlan L. H. Andrew[†], and Michael P. Rumsewicz[*]

[*] SERC, RMIT, Australia ({david, mpr}@serc.rmit.edu.au)
[†] University of Melbourne, Australia (lha@ee.mu.oz.au)

I. INTRODUCTION

Despite extensive work on the network congestion[Jain, 1990], network congestion remains a real problem. The solution is not one answer, but many. Call admission controls try to limit the number of sessions according to the network's traffic capacity. Implicit and explicit feedback congestion control techniques are often too slow in reacting to transient conditions. Window, automatic code gapping, and percentage throttling based techniques all exhibit these shortcomings. Explicit feed back techniques may even make the congestion worse.

All these techniques play an important part in reducing congestion. However, transitory periods of congestion still occur. Reasons for this include the time taken to react to congestion, the nature of congestion control, and the character of packet traffic itself [Leland et. al., 1994]. For real time services, such as voice, this causes severe degradation in quality or complete loss of service. For interactive services, such as WWW browsing, delays become intolerable and packet retransmission further aggravates the problem. Unfortunately, such problems are likely to occur at relatively low average utilisation levels due to the self-similar nature of most data network traffic [Erramilli et al., 1996].

Given that transient congestion episodes are still a problem, it is necessary to make the network as tolerant as possible to these. This paper proposes an alternative philosophical approach to the problem and demonstrates that it can be implemented using a system of two queues.

II. PHILOSOPHY

Much recent work has focused on the concept of fair queuing [since Demers et al., 1989]. By allocating the available bandwidth as fairly as possible, it ensures that the performances of all connections are not degraded by the bandwidth requirements of a few. However, fairness does not always guarantee that most sources will receive good service. Under certain overload conditions a fair allocation of the available bandwidth means that most, or all, connections experience degraded performance during the congestion episode. These overload conditions occur when there are large numbers of connections, or the majority of connections having bursts at similar times.

We propose an alternative philosophical approach to the management of transient congestion in data networks. Our approach is to maximise the number of connections that receive good service regardless of the average or instantaneous utilisation levels. This is done at the expense of providing very poor service to selected connections during the congestion episode. Our philosophy is based on the following premises: from a network provider's point of view, it is much better to maximise the number of happy customers than have all equally unhappy; from a customer's point of view it is better to get good service as often as possible rather than varying degrees of degraded service.

III. DUAL QUEUE APPROACH

The Dual Queue approach uses two queues to implement this philosophy. The α queue is a short and efficient First In First Out (FIFO) queue (see slide 4). Its length corresponds to the longest delay that can be considered "good service". The β queue is significantly longer than the α queue. Its length is determined by the packet loss requirements of the system. All packets leave the system via the α queue. When a packet enters the system a decision is made as to whether it should be placed in the α or the β queue. The packet is usually placed in the α queue. However, if the queue length of the α queue exceeds some threshold and it is decided that the traffic from this source will significantly contribute to the α

queue overflowing, this packet, and subsequent packets from the source, are time stamped and placed in the β queue. As subsequent thresholds are crossed, further connections may be re-directed to the β queue.

When the length of the α queue reduces sufficiently, a packet from the β queue is moved into the α queue (see slide 5). The selected packet is from the last connection re-directed to the β queue. That is, Last In First Out (LIFO) is employed with respect to connections. LIFO connection order is used so as to send packets that are most likely to be useful first. The longer packets are in the β queue the less likely they will be useful once they reach their destination. Within a connection, the packets are moved in FIFO order to preserve sequencing.

Packets from a connection cease being redirected to the β queue when there are no packets from that particular connection in the β queue. Packets that have been in the β queue for too long are discarded. This ensures that packets which are no longer useful do not further waste network resources. This is especially important for real-time services, such as audio or video which are not delay tolerant.

Packets that have gone through the β are marked to allow for identification by other nodes in the network. At subsequent nodes these will be the first choice for redirection to the β queue during a congestion episode.

IV. SIMULATIONS

Simulation experiments with varying degrees of long range dependence have been conducted using on/off sources with Pareto distributed on times, and also with real MPEG traces (We use MPEG traces generated by Rose [Rose, 1995]). Both the generalised traffic simulations and the MPEG simulations yield very similar overall results. For brevity we present only MPEG results here.

Twenty different MPEG traces with randomly offset start times are fed into both a normal FIFO queue and the Dual Queue system. MPEG frames are segmented, if necessary, and sent in ethernet size packets. Both systems are given the same total buffer space. Good service, in this experiment has been defined as less than or equal to a 50ms delay. Results presented are overall results, or selected typical results.

MPEG traffic has of burstiness on two scales. The average data rate offered by the sources varies with the complexity of the scene and movement within the scene. The top figure on slide 6 shows the 1s average bit rates offered by one of the sources (log scale). Burstiness at the frame level is deterministic (lower figure slide 6). It occurs at the frame rate, 40ms in this case.

Frame delay is a good measure of the quality of service each source is receiving. Slide 7 shows the delay characteristics of typical sources over a 500s period. It is immediately observable that for a normal FIFO (referred to as Single Queue or SQ) all sources experience similar delays in a given 1s averaging interval. The same traffic passed through the Dual Queue system has quite a different performance characteristic. The majority of sources do not suffer degraded performance because of a minority. Even when the traffic offered to the system far exceeds the available bandwidth (not shown) as many sources as possible will still receive good service. This of course is at the expense of the other sources. However, in such case a single FIFO queue would give extremely poor service to all sources. A Fair Queuing system would also not give good service to any of the sources.

The Dual Queue approach dramatically improves the goodput. Slide 8 shows the percentage of bad frames in a 1s averaging interval. A bad frame is a frame which is incomplete (due to packet loss), or is made up of packets that have been delayed for more than what is considered “good service”. The figure is for all sources in 1s intervals. During transient overload conditions, the single queue gives “bad service” to all of the frames that pass through it. This is due to the delays caused by the queue being full, or by packet loss. The performance of the Dual Queue system is substantially better. There are far fewer bad frames. The Dual Queue system’s operation will generally make bad frames even worse, thus allowing good service for the rest of the frames.

The Dual Queue system is inherently unfair on small time scales. It sacrifices the performance of a limited number of sources for the benefit of the others. It attempts to choose the sources that are contributing the most to the congestion episode. It can, however, be considered fair over longer time periods, due to the probabilistic nature of redirection to the β queue. Slide 9 indicates how the bad frames are distributed amongst the 20 sources. It gives some idea of fairness over the longer term[1]. A "|" is drawn if there is a bad frame in the 1s measuring period. The figure shows the distribution of bad frames amongst the sources. During any transitory congestion period only a minority of sources will be redirected into the β queue. By definition of the Dual Queue's operation, frames from the other sources will be good.

A histogram of frame delays provides a good indication as to the robustness of a network of Dual Queue nodes. As a packet goes through a sequence of queues the delay distribution is a convolution of the delay distribution at each individual node. Slide 10 shows the delay histogram of frames from one of the sources, but is typical for all of them. The single queue histogram shows quite a heavy tailed distribution. This is to be expected due to the long range dependence of the MPEG traffic and the because the queue has some overflow. As the queue has some overflow it is more probable to find the queue filled to capacity, causing a hump near the 0.6s delay mark. The Dual Queue system shows a distribution concentrated near the origin with a very light tail. This gives an indication of the robustness of the queuing algorithm when applied to a network where each node would similarly convolve the delay distribution. Therefore the end to end variation in delay will be much less with the Dual Queue approach than with the single queue FIFO.

V. CONCLUSIONS

The Dual Queue approach has the following characteristics. All connections whose packets are placed in the first queue receive good service. Connections whose bursts of packets are likely to degrade the service of other connections are placed in the second queue. The FIFO-LIFO discipline of the second queue maintains packet order while serving packets that have the highest probability of being useful first. This also serves to maximise the number of connections receiving good service. This queuing discipline is not fair in dividing the available bandwidth among the connections over short time scales. However, short term fairness does not give the best network performance or customer satisfaction during congestion.

The results show the Dual Queue's superior performance over a standard FIFO queue. They also indicate the feasibility of real-time traffic over a "best effort" packet switched network if this algorithm was adopted.

VI. CURRENT AND FUTURE RESEARCH DIRECTIONS

Research is being conducted into techniques for best choosing and redirecting connections to the β queue. This work involves rate, level, and overflow based α queue management. The performance of the system is being explored under traffic loads from other real time sources, sources with higher level protocols such as TCP and the effects of flow control, and mixed real time and non-real time sources. A priority based implementation is being investigated with a corresponding pricing scheme based on it. Integration of the Dual Queue philosophy with weighted fair queuing mechanisms is also being investigated. Research is continuing into the performance of Dual Queue networks.

REFERENCES

A. Demers, S. Keshav, and S. Shenker. "Analysis and simulation of a fair queueing algorithm", Proceedings of the SIGOCMM '89 Symposium on Communications Architectures and Protocols, vol 19, no 4, pp1-12, 1989.

A. Erramilli, O. Narayan, and W. Willinger, "Experimental queueing analysis with long-range dependent packet traffic", in IEEE/ACM Transaction on Networking, vol 4, no 2, pp209-223, 1996.

R. Jain, "Congestion control in computer networks: issues and trends", in IEEE Network Magazine, pp24-30, May 1990.

W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, "On the self-similar nature of ethernet traffic (extended version)", in IEEE/ACM Transactions on Networking, vol 2, no 1, pp1-12, 1994.

O. Rose, "Statistical properties of MPEG video traffic and their impact on traffic modeling in ATM systems" in Proceedings of the 20th annual conference on local and computer networks, pp397-406, 1995.

---

[1] It should be noted that these are heterogenous sources so cannot be compared directly. The sources that have no bad frames in the 500s period have relatively lower average bit rates and burstiness (as measured by the Hurst parameter).