A GENERIC ARCHITECTURE FOR MULTISERVICE ROUTING

Xiaglin Li and Frank Ball
Distributed Research Group
School of Computing and Mathematical Sciences
Oxford Brookes University
XL.Li,fball,@brookes.ac.uk

Abstract

This paper presents our proposal for the development of a prototype multiservice router that will be used to form the basis of an experimental multiservice network. It discusses the findings of our previous work into network level support for multiservice, in particular the results for a comparative evaluation of WFQ and CBQ, and relates them to the development of this model

1. Introduction

In order for a packet switched network to provide multiservice it must employ some form of bandwidth partitioning mechanism. Furthermore, to support the guaranteed classes of service, some form of resource allocation process must also be provided. Currently there are two main approaches to resource control in packet switched networks: Sorted Priority; and Class Based Queueing (CBQ).

In previous work we have carried out an evaluation of the two approaches where each was presented with identical traffic profiles [Callinan, 2000]. An important consideration of this work was to assess the effectiveness of each system in supporting guaranteed services to continuous media sources, and not just different classes of computer data traffic. In addition to the comparative evaluation of CBQ and WFQ, we have been carrying out investigations into other aspects of network level support for multiservice. This work is divided into three main areas: measurement based flow control for use with CBQ in packet switched networks [Ball.]; reactive congestion control for adaptive continuous media [Tater, 2000]; and mapping IPv6 flows onto ATM cell streams [Basu, 2000]. The majority of this work has been theoretical and mainly involved analytical modelling and simulation. However, we now intend to use the results from these previous studies in the development and implementation of a prototype router that will facilitate the deployment of appropriate multiservice support mechanisms.

As with our previous work an important consideration in the development of this router will be the provision of guaranteed services to continuous media, thereby allowing good quality audio and video applications to be supported. A mayor motivation for carrying out this implementation work is to further validate the control mechanisms we have developed during the theoretical phase of our research. However, once completed this multiservice router will be used as the building block of an experimental multiservice network that will be used to support experimental multimedia applications in addition to further network level research. Therefore an important influence in the development of this router will be the need to support controlled experiments.

This paper presents our proposal for the general architecture of a prototype multiservice-router, which has been developed using the finding of our previous work. It describes work in progress toward its implementation and pays particular attention to the link-sharing mechanism, which is the main focus of our current work. The remainder of the paper is organized as follows: section 2 outlines the main findings from our previous work; section 3 discusses a general model for multiservice that we have developed from these findings; section 4 presents general architecture of the proposed router; and finally, in section 5 we conclude.

2. MAIN FINDINGS FROM PREVIOUS STUDIES

As a research group we have been investigating a number of issues relating to the problems of providing multiservice in packet switched networks. Our work has been focused mainly on the development and evaluation of bandwidth partitioning and resource control mechanisms. The development work specialising in support for continuous media traffic. Our approach has been to consider these problems in a general context with a focus on the medium to long-term future. Therefore our investigations have not been limited by legacy factors or the constraints of current commercial considerations.

2.1 Comparative Evaluation of CBQ and WFQ

Sorted priority algorithms allow reordering of the output queue so that flow delays can be controlled. They can provide a mathematically provable, if somewhat conservative, delay bound to each source, provided that the source's traffic is appropriately shaped. The general goal of a sorted priority algorithm is to provide a guaranteed weighted share of the bandwidth when the network is heavily loaded, and a weighted fair share of any surplus bandwidth when the network is lightly loaded.

CBQ and link-sharing mechanisms [Floyd, 95] divide the link bandwidth between a numbers of different classes of traffic, and if required, can provide isolation between these classes. However, if individual sources within a class are to be given a service guarantee, an additional acceptance procedure must be provided. Generally, this acceptance procedure will employ some form of analytical queueing model that will provide a probabilistic, and much less conservative, delay bound than that provided by sorted priority algorithms. CBQ can provide a class with either a fixed share of the link bandwidth or allow classes to take an additional share of any unused bandwidth.

This evaluation set out to compare the performance of the two systems in three main areas: the ability to provide bandwidth partitioning and isolation between different flows of traffic, particularly under conditions of heavy network loading; the response time delay characteristics, again under heavy load conditions; and the ability to be fair when distributing surplus bandwidth between competing flows, at times of lighter network loading. Full details of the comparative evaluation and its results are given in [Callinan, 2000]. In this section we simple present a summary of the findings that are relevant to the development of our proposed experimental system.

Although it was shown that both WFQ and CBQ could offer adequate isolation to continuous media sources, CBQ was shown to offer a much lower response time delay at times of heavy loads, provided that the sources were grouped into appropriate classes. WFQ was shown to give a lower response time delay once the network load drops below a certain level. This would be expected due to being a work-conserving algorithm. However, this was not considered to be of any real advantage to continuous media since the response time delay under heavy loads would be the important factor for jitter smoothing at the receiver.

Another important observation was that the mathematically proven bound offered by WFQ (the Parekh Bound) [Parekh, 93] was seen to be far too conservative to be of any real use. The mechanism we describe in section 2.2 can provide a probabilistic delay bound for a CBQ system. This method can predict the higher percentiles of the response time delay for any class using a non-work-conserving rate controller. Such percentiles, for example the 99%-tile, provide a far more pragmatic delay bound.

Both systems were shown to be capable of guaranteeing a minimum throughput and providing a fair share of surplus bandwidth to competing flows of computer data. However, the work-conserving rate controller used with CBQ for the computer data classes [Floyd, 93] was seen to be sensitive to traffic characteristics. Certain parameters of this mechanism need to be adjusted to cope with either random or self-similar traffic sources. On the other hand, WFQ was shown to be completely insensitive to this problem, and worked equally well with both random and self-similar traffic without any need for adjustment.

We conclude from these findings that generally, continuous media will be best served by a simple CBQ system using a non-work-conserving rate controller. However, until a fair work-conserving rate controller that is not sensitive to traffic characteristics can be developed, computer data would be best controlled by WFQ. Fortunately, it is possible to implement a hybrid solution, and thereby meet the needs of both types of traffic in a single system. This is the approach we will take in the development of the first prototype multiservice router.

2.2 Measurement Based Flow Acceptance Control (M-FAC)

Although CBQ can provide bulk resource allocation to a particular class of traffic by providing isolation from other classes, it cannot directly provide a guarantee to individual flows within that class. In order to achieve this it is also necessary to control admission to the individual class. Admission test are generally based on certain parameters of the traffic's characteristics and usually involves some for of analytical model. A major problem associated with such tests is the difficultly in obtaining accurate a priori estimates of the traffics characteristic. One solution to this problem is to employed measurement based techniques.

We have developed a measurement based flow admission control mechanism that operates in conjunction with CBQ [Ball, 98, 99a]. This mechanism uses a combination of prediction and measurement techniques in order to maintain the Quality of Service (QoS) requirements of continuous media traffic. It can be used to provide admission tests to individual streams, or alternatively, it may be used to dimension the resources needed to support aggregated streams. It is similar in operation to a measurement based ATM CAC scheme proposed by Crosbie et. al. [Crosbie,] except that it is designed to work with variable length packets.

The rate assigned to a class is initially predicted from a traffic specification of the aggregated arrival stream and the desired delay/loss characteristics. When a request is made for a new stream, its traffic specification is combined with the traffic specification of the existing aggregated stream to form an updated traffic specification. This updated specification is then used to predict a new value for the assigned rate. If this new value is below the permitted rate threshold then the new stream is accepted, otherwise it is rejected.

Although, in the event of a new stream the assigned rate is determined by prediction, it may be adjusted later if the actual delay characteristics differ from the desired delay characteristics. Between the acceptances of new streams, a monitoring process will control adjustments in the rate. This monitoring process will not only adjust rate, but will also update the aggregated traffic specification. We have completed development of the prediction mechanism, and currently we are experimenting with appropriate monitoring techniques.

2.3 Reactive Congestion Control for Adaptive Continuous Media.

Although a guaranteed service is the most appropriate for supporting continuous because it will provide consistency, certain continuous media applications may be able to tolerate a degree of fluctuation in quality. Therefore it would be useful for the network to also offer a class of service that could meet the requirements of adaptive continuous media sources.

The requirements of adaptive continuous media are very different from those of computer data. Whilst maximising throughput, and providing fairness between different sources may be the main requirements for computer data, avoiding loss, and graceful adaptation are more important for continuous media steams. Therefore it would seem appropriate that these two types of traffic should be controlled by separate mechanisms.

We have been carrying out an investigation into the requirements for network support of adaptive continuous media applications. This investigation has found that whilst certain aspects of feedback control developed for computer data may offer some benefits to adaptive continuous media, generally they are inappropriate to its needs. Techniques such as RED (Random Early Detection) and its variants have been shown to provide a more optimal control with TCP sources [Floyd, 93]. However, our work has shown that they will not offer the same benefits to continuous media sources. Generally, these methods focus on controlling the average delay at each node, whereas controlling a higher percentile of delay is more appropriate for continuous media. Furthermore, the current version of RED employs packet loss as an implicit signal of congestion.

We propose that adaptive continuous media traffic should be provided with explicit congestion feedback from the network nodes, and that the feedback mechanism should be optimised to the needs and characteristics of adaptive video. Currently, we are experimenting with the use of Forward Congestion Indication (FCI) and we have developed a new congestion monitoring technique that monitors and controls the higher percentiles of delay. This mechanism could also work with Backward Congestion Indication (BCN).

2.4 Mapping IPv6 Flows onto ATM Cell Steams

The next generation of Internet Protocol, IPv6 provides some additional features to aid support to guaranteed services. One of such feature is the flow label identifier (ID) field that identifies individual flows. The joint combination of IP address field and flow label id field can be used to identify each individual flow on the Internet and hence provides the framework to provide flow based QoS.

Although future IP networks will be able to support a similar kind of service guarantee as ATM networks, there will still be many cases where IP traffic will have to be transferred over an ATM network. Where this is the case, previous techniques designed for this purpose will no longer be sufficient, as performance will now also be an important factor. An important requirement will be to map the QoS requirements of the IPv6 flows onto the traffic parameters of the corresponding ATM VCs.

We are investigating the problem of mapping IPv6 packets flows onto ATM virtual circuits (VC), with an overall emphasis on providing a seamless end-to-end QoS guarantee [Basu, 2000]. We have developed a mechanism that operates at the AAL level the objective of which is to choose the minimum possible peak cell rate that can satisfy the delay constraints at the IP-ATM interface. This mechanism is based on previous work into traffic smoothing at the AAL level [Ball, 96a, 96b].

3. A MODEL FOR MULTISERVICE

This section presents a general model for multiservice, which is based our on the findings presented above. A general model for a multiservice router is shown below in figure 1.

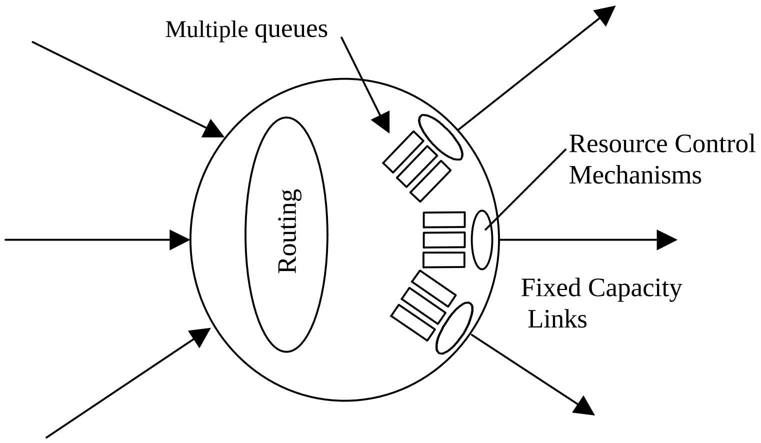

Fig. 1 A General Model for a Multiservice Router

Our current work is not primarily concerned with the routing stage of the model. Ideally the routing engine should be engineered to be sufficiently powerful to allow the routing process to be internally "non-blocking", thereby ensuring that resource contention takes place only at the output links. The main focus of our current work is on the resource control model associated with the output stages of the router. Figure 2 shows the basic structure of our model which is a hybrid CBQ-WFQ system that is very similar to a model we tested using simulation. In this particular example there five basic traffic classes: guaranteed video (G_VIDEO); guaranteed audio (G_AUDIO); adaptive video (A_VIDEO); adaptive audio (A_AUDIO); and computer data (DATA). Although the diagram only shows one instance of each continuous media class, the model allows for multiple instances of any particular class.

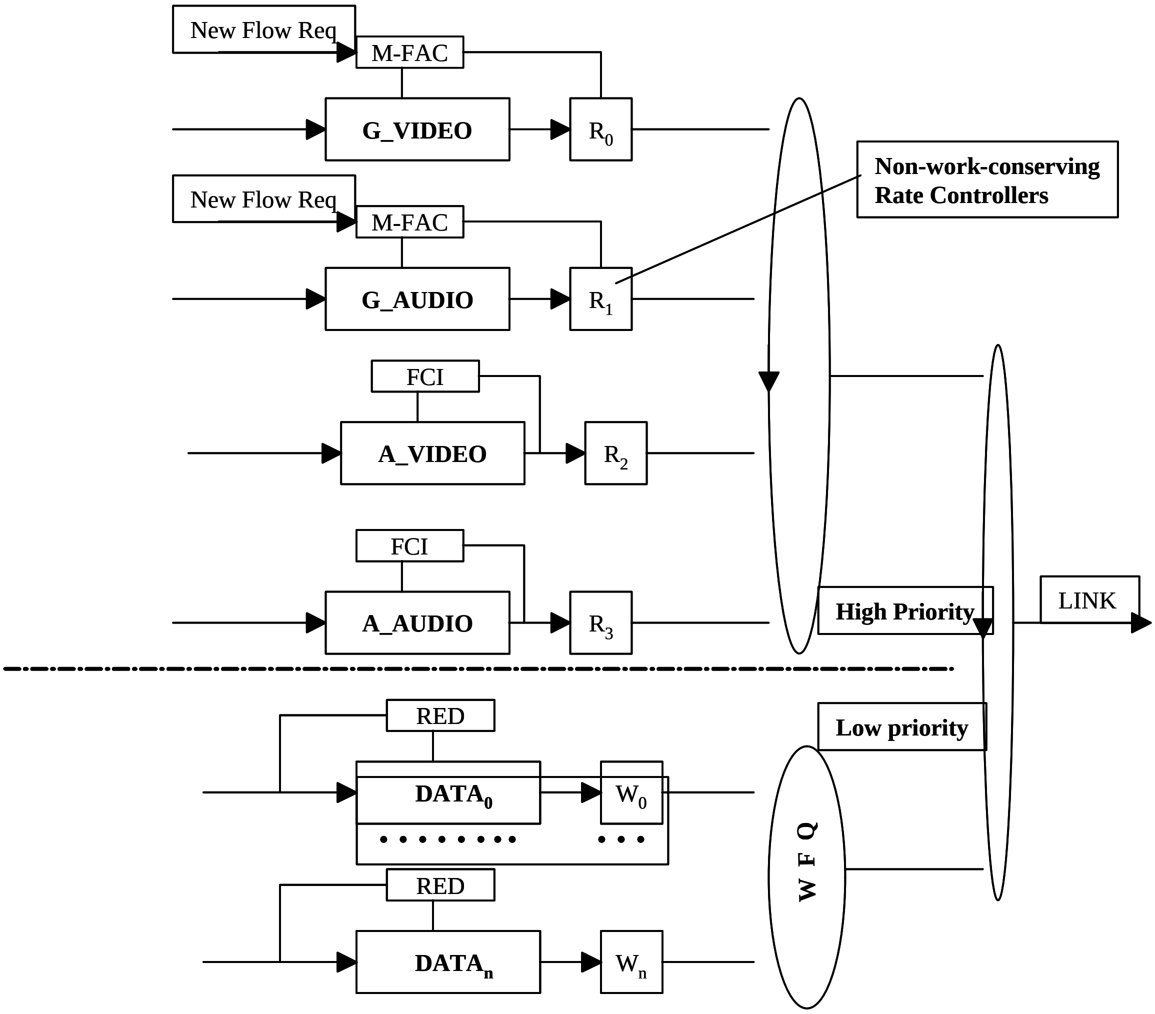

Fig. 2 Bandwidth Sharing and Resource Control Model.

All the continuous media classes are regulated by a simple non-work-conserving rate control mechanism. Each of these rate control mechanism is an assigned a rate, which may be changed by certain events. However, this rate will remain fixed between any two successive events that may effect such a change. In the case of the guaranteed continuous media classes M-FAC is used to control the admission of new flows. When a request is made for a new flow M-FAC will predict what rate is required to accommodate the flow in addition the existing flow. If this new rate is below the agreed maximum for the particular class, M-FAC will update the rate allocation for that class and accept the new flow. In the case of the adaptive continuous media classes there will be no admission test for new flows. However, flows belonging to these classes will be required to adapt their output according to network conditions. Signals requiring the sources to reduce their output will be sent via. the FCI mechanism. FCI messages will be triggered by the new monitoring techniques that we are currently developing.

Flows of computer data will be regulated by a WFQ mechanism, this in turn will be given a share of the link bandwidth by the priority based general scheduler of the CBQ system. Each data class will be assigned a weight, which represents the minimum throughput it can expect to receive. In general the data sources will be able to make use of whatever bandwidth is unused by the continuous media classes. However, the continuous media classes will be limited to a certain percentage of the link bandwidth to ensure that data sources are always given their minimum throughput.

A resource reservation protocol such as RSVP could be used for establishing new flows in the guaranteed continuous media class, and to assign and reassign weights to the data classes.

4. ROUTER ARCHITECTURE

The first task in the implementation stage of our work has been to define a general architecture that meets the requirements of the multiservice model. This is an abstract architecture and does not imply any particular form of implementation. However, it is intended not only define the functional components of the router and their interaction, but also the performance requirements associated with each of these components. It is important that both the computational and performance models are accurately represented in the final implementation to ensure the fidelity of the traffic control models. This work is in its early stages and so far only an outline definition of the architecture has been developed, which identifies the basic functional components as shown below in figure 3.

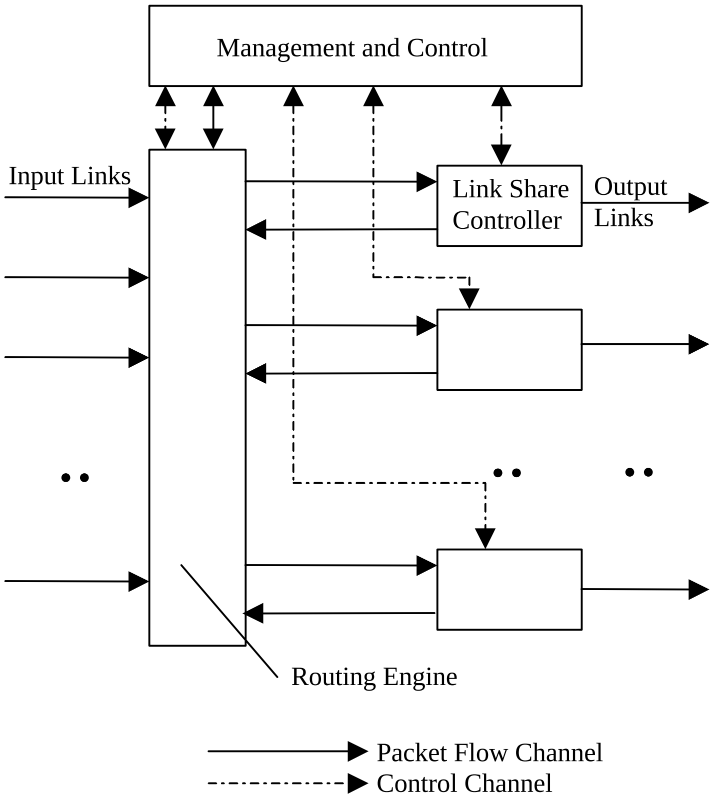

Fig. 3 General Architecture of the Router

Our current work is mainly focused on the Link Share Controller, details of which are shown below in figure 4.

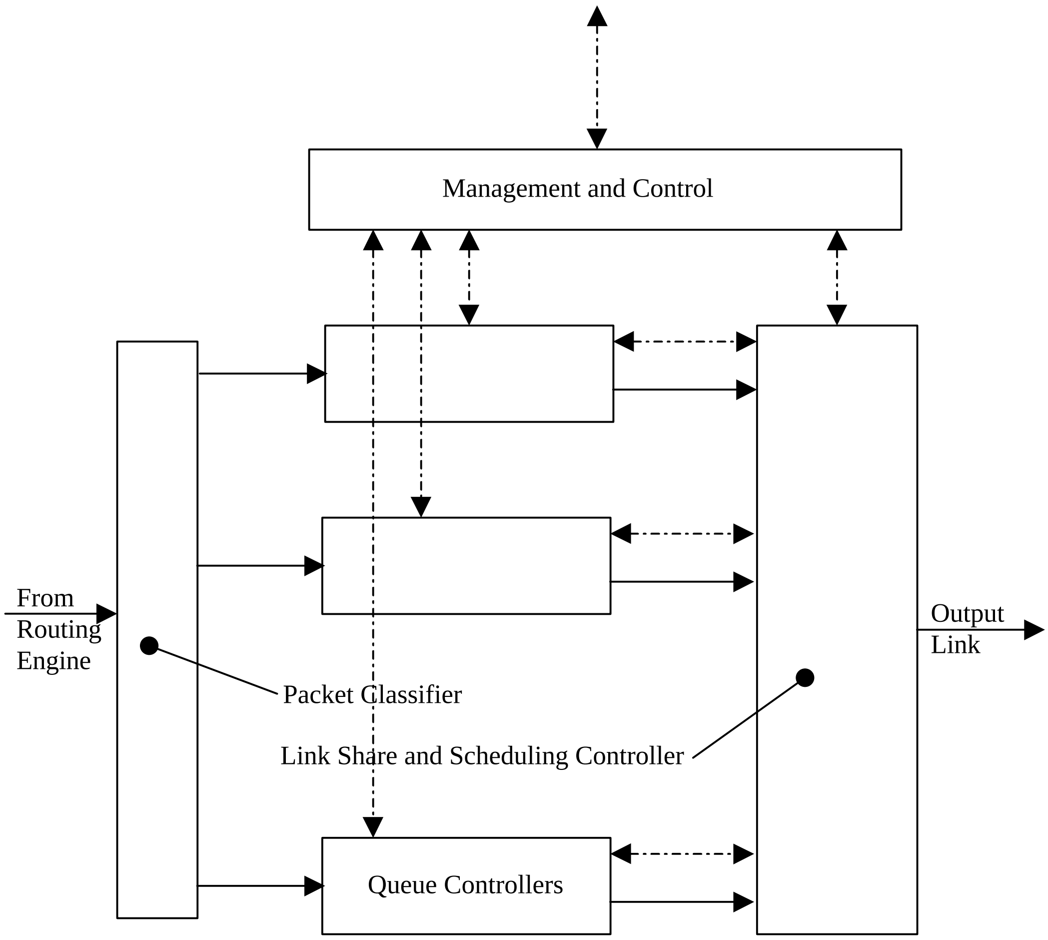

Fig. 4 Link Share Controller Architecture

From our previous work we have identified the sub-components of this mechanisms and we are now considering the range of functionality that each will be required to provide. In order to provided maximum flexibility the Queue Controllers will need to be able execute any of resource control mechanism that we presented above in addition to those developed by other researchers. The Link Share and Scheduling Controller will also need to provide flexibility in the hybrid CBQ/WFQ that we have proposed. Once the functional specification of this component has been completed we will develop a model of the performance requirements. We will then design a particular implementation that incorporates both the function and performance requirements. This link share controller is a basic building block not only for the output stage of the router but will also be required for the multimedia end-systems if QoS is to be provided on an end-to-end basis.

5 CONCLUSIONS

This paper has discussed our proposal for the development of a prototype multiservice router that will be used as the basis for an experimental multimedia communications infrastructure. We have presented an architecture for this router that we have developed from the findings of our previous work into network level support for multiservice.

References

[Ball, 96a] F.Ball, D. Hutchison and D.D. Kouvatsos, " VBR Video Traffic shaping for ATM Networks", **ATM, Networks and LANs NOC'96**, vol II, D.W Faulkner and A.L.Hamer (Eds), IOS press, 68-74,.ISBN 90-5199-272-2

[Ball, 96b] F.Ball, D. Hutchison and D.D. Kouvatsos, " VBR Video Traffic Smoothing at the AAL SAR Level", **IFIP 4th Workshop on Performance Modelling and Evaluation of ATM Networks**, (8th-10th July 1996) Ilkley UK.

[Ball, 98] F. Ball, P. Callinan, D.D. Kouvatsos and C Skianis, "Dynamic Dimensioning for Guaranteed Services in Packet Switched Networks." **IFIP 6th Workshop on Performance Modeling and Evaluation of ATM Networks**, (July 1998) Ilkley, UK.

[Ball, 99a] F.Ball, P.Callinan, D.D.Kouvatsos and C Skianis, "A Measurement Based Admission Control Mechanisms for use with CBQ in Packet Switched Networks", **15th Annual UK Performance UKPEW'99**, (22nd-23rd July 1999) Bristol, UK. ISBN 0-9524027-8-5.

[Ball, 99b] F. Ball and S. Tater, "Supporting Adaptive Video Applications in Future IP Networks", **European Workshop on DISTRIBUTED IMAGING**,(18th -19th Nov 1999) London, UK

[Callinan, 2000a P. Callinan, M.Witwit, F. Ball, "A Comparitive Evaluation of Sorted Priority Algorithms and Class Based Queueing using Simulation", **Advanced Simulation Technologies Conference ASTC 2000**, (16th-20th April 2000) Washington DC.

[Basu, 2000] K. Basu, F. Ball and D.D. Kouvatsos," IPV6 over ATM: Flow mapping for guaranteed services", To be presented at the **16th UK Teletraffic Symposium UKTS** (22nd -24th May 2000) Harlow, UK

[Floyd, 93] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", **IEEE/ACM Transactions on Networking,** August 1993.

[Floyd, 95] S. Floyd and V. Jacobson, " Link-Sharing and Resource Management Models for Packet Networks", **IEEE/ACM Transactions on Networking**, 3, 4 August 1995.

[Parekh, 93] Abhay K. Parekh, "A Generalised Processor Sharing Approach to Flow Control in Integrated Services Networks", LIDS-TH-2089, MIT Laboratory, a doctoral dissertation.

[Tater, 2000] S Tater and F.Ball " Supporting Adaptive Video Applications in Future IP Networks", **18th IASTED International Conference Applied Informatics AI'2000** (14th-17th Feb 2000) Innsbruck, Austria.