Computación y Sistemas Vol. 3 No.1 pp. 50-57
 ISSN 1405-5546 Impreso en México

Resource Reservation Management of Video over Internet

V. M. González y Gónzalez
Unidad de Sistemas de Información y Cómputo
Universidad de Guanajuato, México.
Fax: +52 473 21804 Tel: +52 473 20006
E-mail gonzvic@quijote.ugto.mx

M. Ghanbari
Multimedia Communications Research Laboratory
Department of Electronic Systems Engineering
University of Essex, United Kingdom
Fax: +44 1206 872900 Tel: +44 1206 872434
E-mail ghan@essex.ac.uk

Article received on January 22, 1999; accepted on May 14, 1999

Abstract

This paper describes a method of Resource Reservation Management (RRM) mechanism that optimises bandwidth reservation in IP routers. The technique uses the original specification of the Resource Reservation Protocol (RSVP) with minimum modifications. In the proposal it is assumed that the video services are coded at multiresolution bit rates, each providing a different quality of service. The paper analyses the behaviour of the proposed RRM in an Internet network with different levels of congestion. The results show that the proposed mechanism can deliver an acceptable quality of service by dynamically adjusting the demanded bandwidth.

Keywords:

Packet video, video over Internet, ATM, Internet.

1 Introduction

It is believed that changes on the role of the Internet from being a pure data-network into a multimedia network are very immanent. The original structure of the Internet network evolved with the aim of providing a reliable data-transport service by using retransmission and automatic reconfiguration of routers. However delivery of audio and video services (real-time applications) has always experienced problems within Internet.

The Transport Control Protocol (TCP) is a well-known retransmission protocol for non-real time applications, which can guarantee correct reception of the data (Hunt, 1994; Thomas, 1997). The same technique when applied to real-time applications may not work, since retransmission only adds additional latency to the received data. TCP puts much emphasis on ensuring that information is delivered safely, rather than on time. For real-time applications if a packet does not arrive on time or it is erroneous, it is better to be discarded, since with interpolations some of the lost information can be retrieved (Ghanbari and Seferidis, 1993).

To resolve this problem, Resource Reservation Protocol (RSVP) has been set-up to cater the required resource reservation to support real-time services over the Internet. The protocol is used by routers to deliver requested Quality of Service (QoS) to all the nodes along the path(s) of the flows and to establish and maintain a state to provide the requested service. RSVP will generally, although not necessarily, result in resources being reserved in each node along the data path. (Braden *et all.*, 1996; Zhang *et all.*, 1993). However this still is not enough to provide smooth delivery of video over Internet. In the following we first review the RSVP protocol, and show how addition of a simple extra flow control to RSVP facilitates transmission of multi-layer bit streams over Internet.

Since RSVP is a receiver-oriented protocol, to establish a reservation for a particular flow between a source and its destination two steps are needed. First, the receiver needs to know what traffic the sender is likely to send so that it can make an appropriate reservation. With a best-effort service model we can only indicate to the network where the packets should be sent. A real-time service requires the user to provide a specification of the flow (*flowspecs*) to the network. Second, after the sender's specifications, it should know the passage where message will follow from sender to receiver, so that it can make the reservations in each hop of the network. It is important to indicate that the path from sender to receiver may be different to the path from receiver to sender.

The path and specification (Tspec - transmission specification) could be sent to the receiver in the same message, called PATH. Each router looks at the PATH message, as it passes by, and works out the reverse path that will be used to send reservations from the receiver back to the sender. With the PATH message, the receiver has enough information to request a reservation. This is done by using a RESV message, which contains the sender's Tspec and an Rspec (Receiver specification) describing the requirements of the receiver.

Each router in the path looks at the RESV message and tries to allocate the necessary resources. If the reservation can be made, the RESV message is passed to the next hop in the path. If not, the router sends an error message to the receiver with the reason for denying the request. There are many reasons why a request may be denied: the requested bandwidth is unavailable, the delay time-limit cannot be met, bad flow specification, etc.

In the case of rejection, the application may consider two options: Indicate to the user that the connection is not possible or try a new request for fewer resources (i.e. bandwidth, delay, etc.). Whether the latter is possible or not depends on the sender who should have different versions of the same stream, provided that the source is multilayer coded (Ghanbari, 1989). Other alternative is to transcode the non-layered bit stream into lower rates (Assuncao and Ghanbari, 1997).

2 Resource Reservation Management

Multimedia applications over the Internet could work best if they are adapted to the varying bandwidth. Adaptive applications reduce the amount of data transmitted during periods of congestion; this results in either a lower-quality transmission (known as throughput adaptation) or a reduced frame frequency (known as delay adaptation) (Emmerson and Greetham, 1996).

Resource Reservation Management (RRM) is a framework used for real-time applications to manage the allocation of resources in the network. It specifies the way to request resources using layers of reservations. The proposed framework could be implemented as a separate piece of software between the application and the drivers of the network layer. Also, it could be implemented as part of the functions within an application that uses real-time data (e.g. a browser, video tools or audio tools). RRM does not actually specify a protocol and format of messages. Instead it sketches a framework for a protocol, defining the basic roles and operations. Table 1 shows the position of the RRM in the layered model.

Application
Resource Reservation Management
Network Layer Implementation
Data Link Layer Implementation
Physical Link Layer Implementation

Table 1 : RRM in the layered model

RRM in fact makes it possible the multi-layer bit stream supported by RSVP to be transported over IP networks. At the beginning the RRM will operate at a minimum level and gradually, (and if the level of congestion in the network allows it), will increase the quality by requesting a higher layer, until reaches its maximum level

2.1 RRM functions

The functions performed by the RRM are:

Specification of requirements

The application will transfer three parameters to the RRM: The minimum reservation, the amount of increment, and the top reservation. Below the minimum reservation it is not possible to operate.

Setting-up of the connection

On behalf of the application, the RRM will send a RESV message to all the routers between sender and receiver. This message includes the minimum level of resources required. If the minimum reservation cannot be granted by the network, the application is notified by an error. This error could be called "fatal" because the communication in such situation comes to a halt.

Re-negotiation of requirements

As with the specification of RSVP, every 30 seconds the receiver should refresh the reservation in order to maintain it. The refresh message is a RESV message that repeats the allocation of resources throughout the routers. The RRM will keep a record of how many resources it has allocated and if they are less than the maximum level, it will send a RESV message requesting an increment. The network will send an

RSVP message indicating whether or not the reservation can be granted. In case the request cannot be allocated the router keeps the last reservation granted.

Congestion messages

When one of the routers along the path becomes congested (i.e. the memory buffers are full or have reached a predetermined congestion level) a corrective action should be carried out to avoid an excessive drop of packets. The action consists of checking the table of allocated resources and select the host with maximum resources. If this host has a level of resources above the specified minimum level, the router will send a message asking resources to be decreased by one level.

When all the hosts are at their minimum reservation level, the application served on a best effort basis (not supported by RRM-RSVP) could lose packets. They will be dropped by the router if the congestion level continues and reaches the saturation level of the buffers.

Increment and Reduction messages

Because RSVP is a receiver-oriented-reservation mechanism it is necessary to note that an increment or reduction in the reservation should assume a previous arrangement between source and destination. When an increment is requested (to receive a stream with higher quality) the new reservation of resources must be established before sending a notification to the source. A reduction of reservation implies that the host must request a stream with lower quality from the source. Therefore, the router should not decrease the reservation until a new RESV message with fewer resources arrives. This is done to prevent loss of packets from the source with reservation, in the interval that begins when the router sends a Reduction and ends when the host receives it.

2.2 Modifications to RSVP to Allow Multiple Layers of Reservations

The RSVP routers have tables to keep information about allocated resources to specific hosts. Packets whom posses reservations are identified by a flow label. RSVP periodically sends refresh messages to keep the reservation. Routers update their tables by these messages and cancel any reservation that has not been refreshed. The time-out of each reservation could be 60 seconds or any other fairly short periods (Peterson and Bruce, 1996).

The actual specification of RSVP allows the sources to change the reservation requested every time that a RESV message is sent [4]. The first time that a router receives a RESV message it applies the Admission Control algorithms to the request. The next time that a RESV message arrives with the same specifications in the request, the router marks its table and resets its clean-up timer. If the specifications are different from the original ones, the reservation is considered new and the admission process must be done again, so it is possible that in this case the request could be rejected.

To support RRM, RSVP should consider a modified way to increment resources allocated for a flow. A RESV message could be used but, in this case, the function of the message would be to request more resources and, at the same time, refresh the reservation. If the increment succeeds, the application receives a message of acknowledgement. If it fails a negative acknowledgement is received but the router keeps the last reservation granted.

Also, the RSVP must include a message to indicate a reduction in the reservation of a certain host. This message could consist of a very small code that indicates a request for reduction, but not the amount of resources that the source should reduce. The information related to the amount of reduction stays in the host with the reservation. The RESV message should be modified to include the minimum reservation, the amount of increment and the top reservation. These three numbers could refer to bandwidth, buffers reserved, delay restrictions, or any other values used to specify the QoS.

3 Simulation Analysis of Layered Reservation

The aim of the simulation is to analyse the behaviour of the proposed mechanism of Resource Reservations Management (RRM), and to show how the RRM adjusts the reservation according to the condition of congestion in a router. We analyse the performance of RRM as the percentage of time that a reservation stays at a certain level and how the mechanism reacts to congestion in the router.

In the test, a source with multiple levels of reservation is combined with a background traffic without reservation. The level of background traffic, served on a Best Effort basis, and the packet loss in the router due to the saturation of its buffer is calculated. It is important to note that these results have a direct relation to the scheduling algorithm that the router is using. However, the analysis gives us an idea of how a variable reservation may disturb the traffic that does not have reservation.

The router has an output link with a maximum bandwidth of 2 Mbit/s. The maximum level of congestion that the router can reach in its buffer, before carrying out a corrective action, is set to 85%. During the simulation the size of used buffer was 30 packets. The background source can generate traffic at different rates, varying from 500 to 2000 kbit/s with a mean message length of 1000 bytes. The mean message length for RRM source was 100 bytes. The small length of the packets is necessary to allow the transport of delay sensitive traffic. All the simulations consisted of thirty-minute sessions, close to the duration of a typical videoconferencing scenarios.

When packets arrive at a router they are served immediately, unless the link is busy. If the packet is sent to the queue, there is a verification of the state of congestion. The state of congestion is determined by the percentage of buffer space filled.

The simulation model was consisted of a router and two sources: A RSVP source with RRM and a background source with a best-effort service. The RSVP source has a minimum rate of 64 kbit/s with increments of 64 kbit/s up to a maximum rate of 320 kbit/s. The router was configured with a maximum congestion level of 85%, maximum queue length of 30 packets with an output link rate of 2 Mbit/s.

Figures 1-6 show the performance of RRM for each level of the background traffic varying from a mean bit rate of 500 to 1750 kbit/s. They show how RRM can take advantage of the level of congestion in the network and increases theamount of resources requested. If the assumed lowest level of 64 kbit/s is considered to be the minimum acceptable image quality, as the figures show, for all the background traffic, a video with the least bandwidth can always be delivered.

With a low level of background traffic (500 kbit/s) the source gradually reaches the top quality (320 kbit/s) and stays there for the whole session (30 minutes). When the mean background traffic reaches 1000 kbit/s, the RSVP-RRM source adjusts its reservation more than ten times. After a reduction, the source then increments its reservation every 30 seconds. As the mean background traffic level increases, there are less chances of using the higher rates. At the maximum background traffic of 1750 kbit/s, the RRM-RSVP source mainly uses 64 kbit/s. It should be noted that in none of the cases, RRM-RSVP source looses packets. This is achieved at the expense of loosing packets from the best-effort background source, as shown in figure 7.

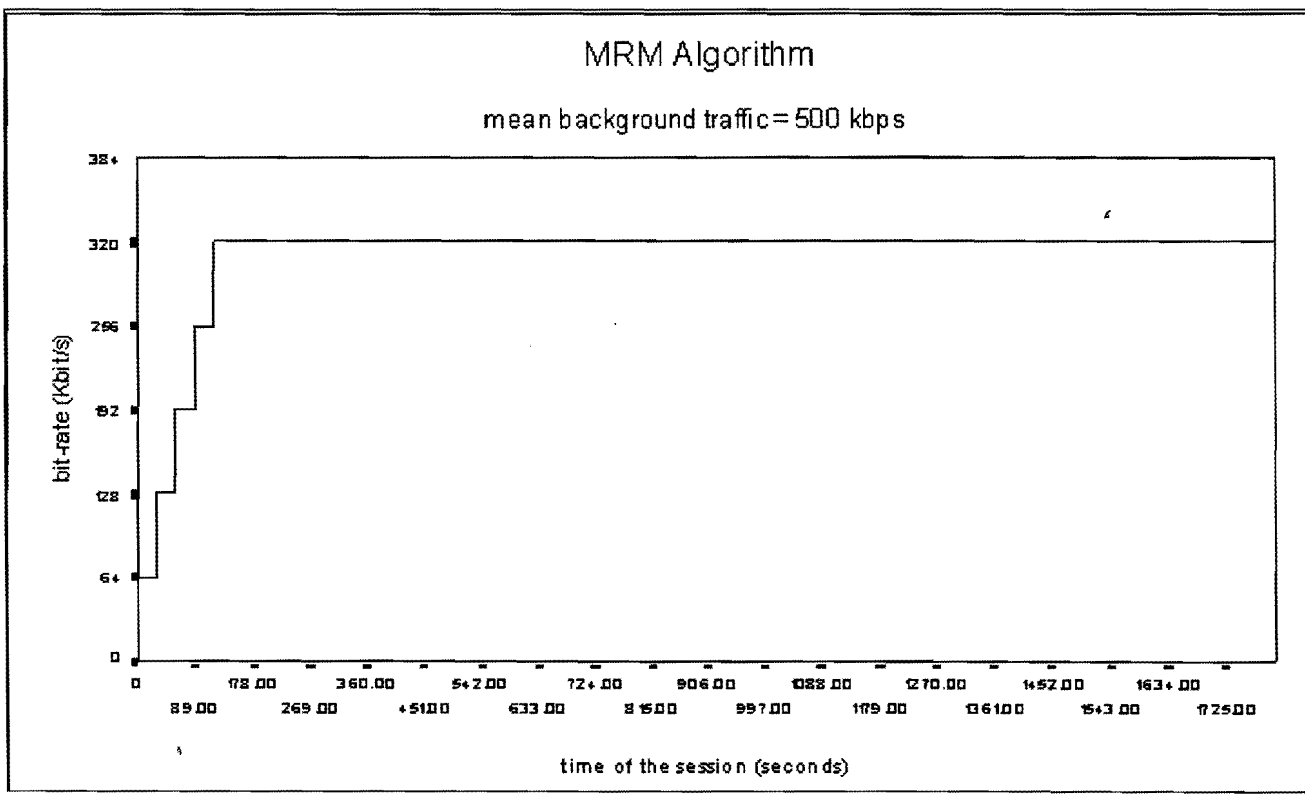


Figure 1: RRM Algorithm - mean background traffic 500 kbit/s

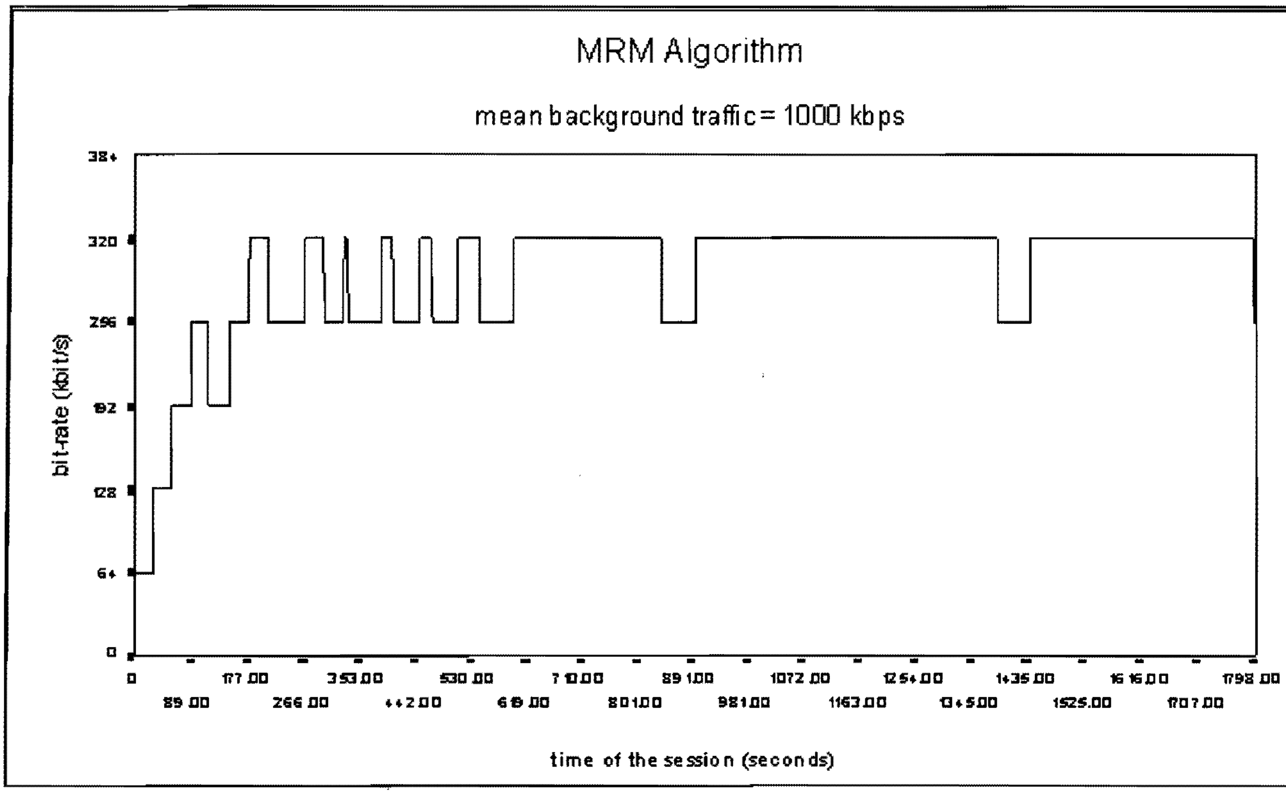


Figure 2: RRM Algorithm - mean background traffic 1000 kbit/s

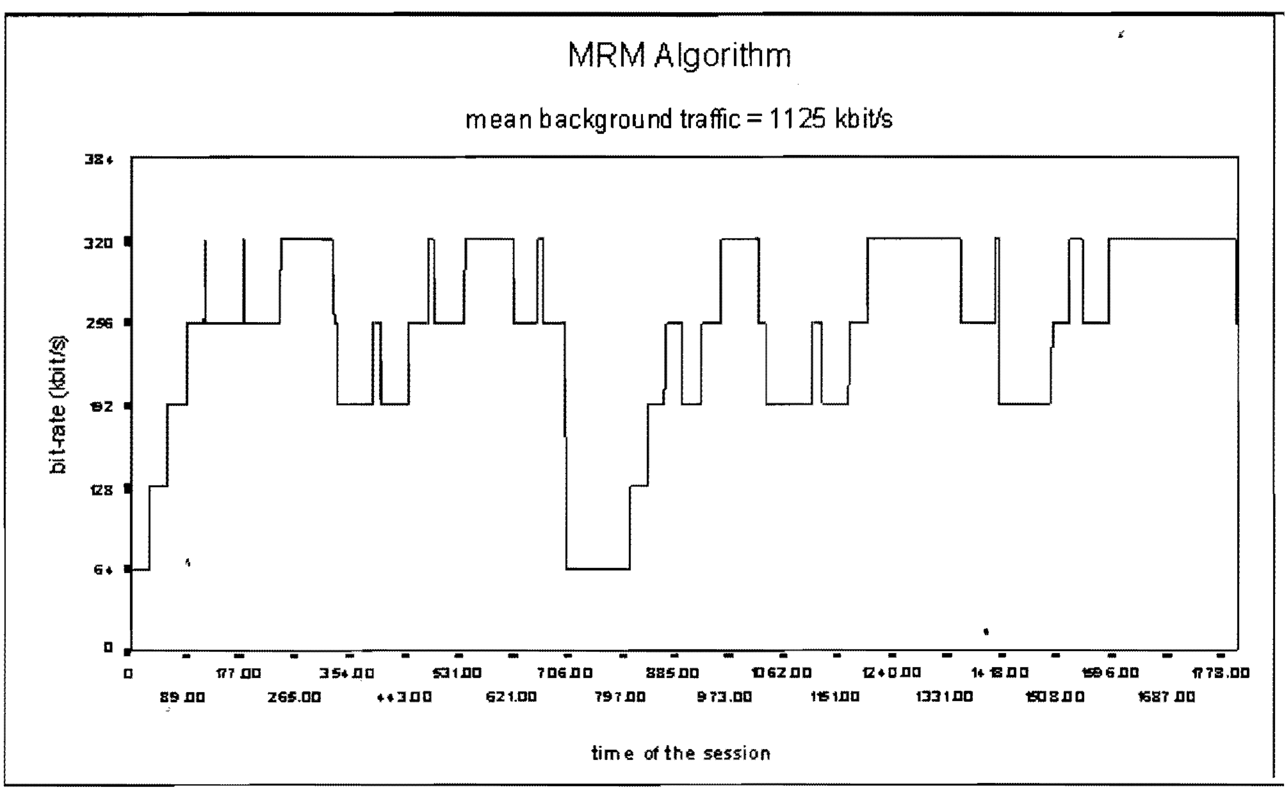


Figure 3: RRM Algorithm - mean background traffic 1125 kbit/s

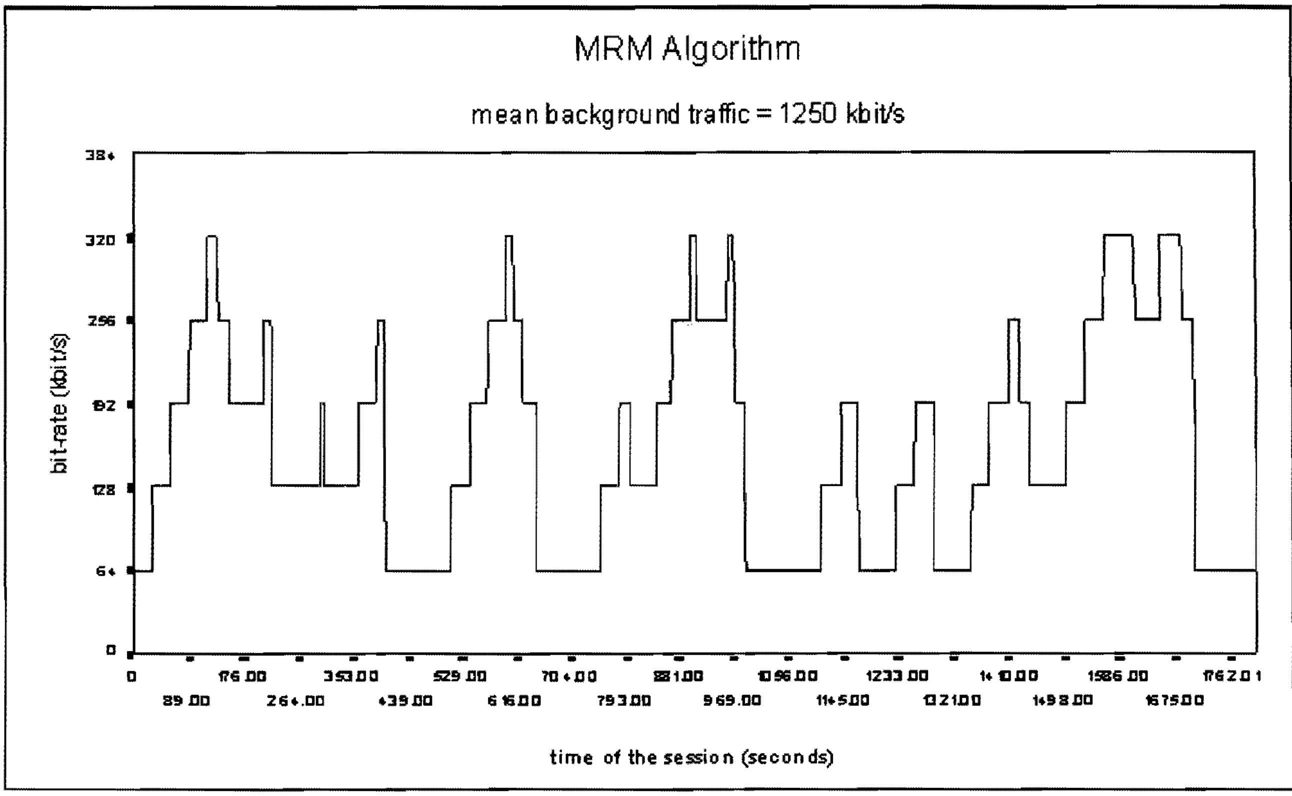


Figure 4: RRM Algorithm - mean background traffic 1250 kbit/s

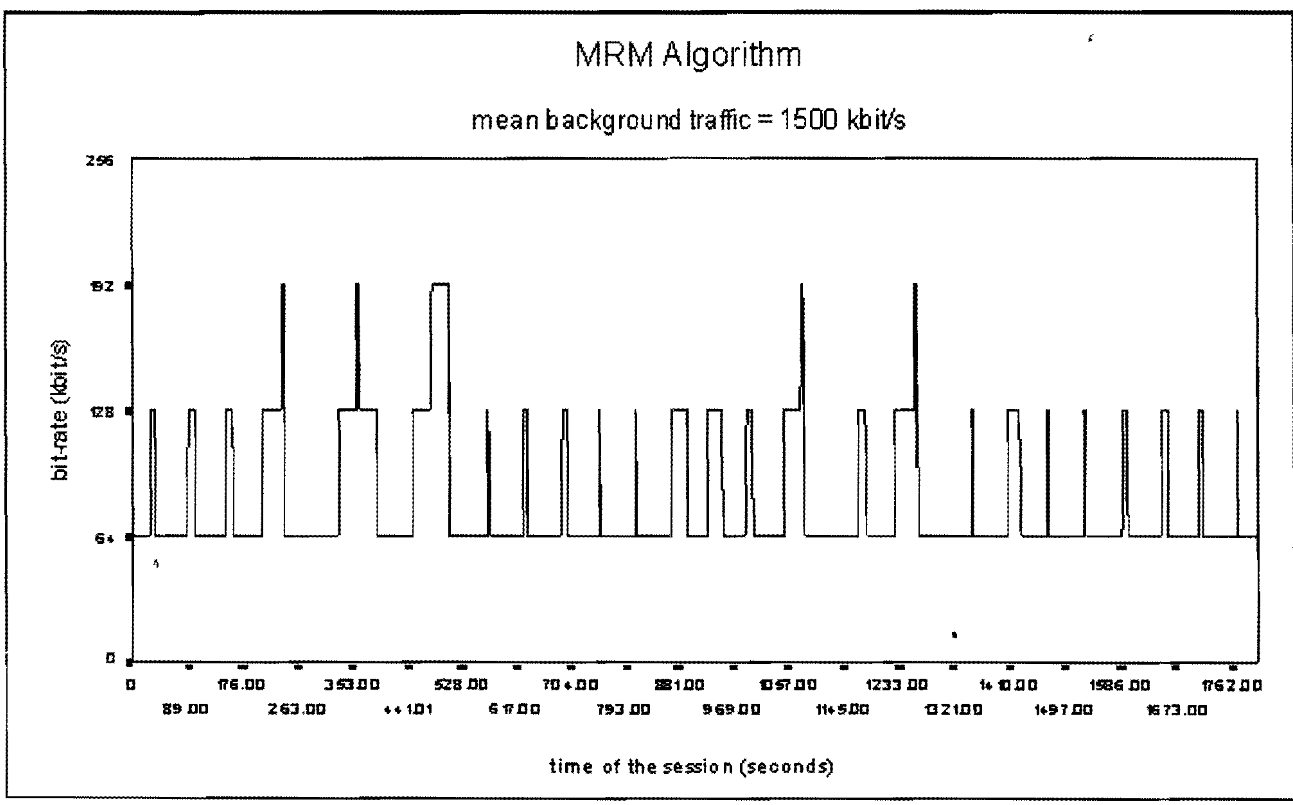


Figure 5: RRM Algorithm - mean background traffic 1500 kbit/s

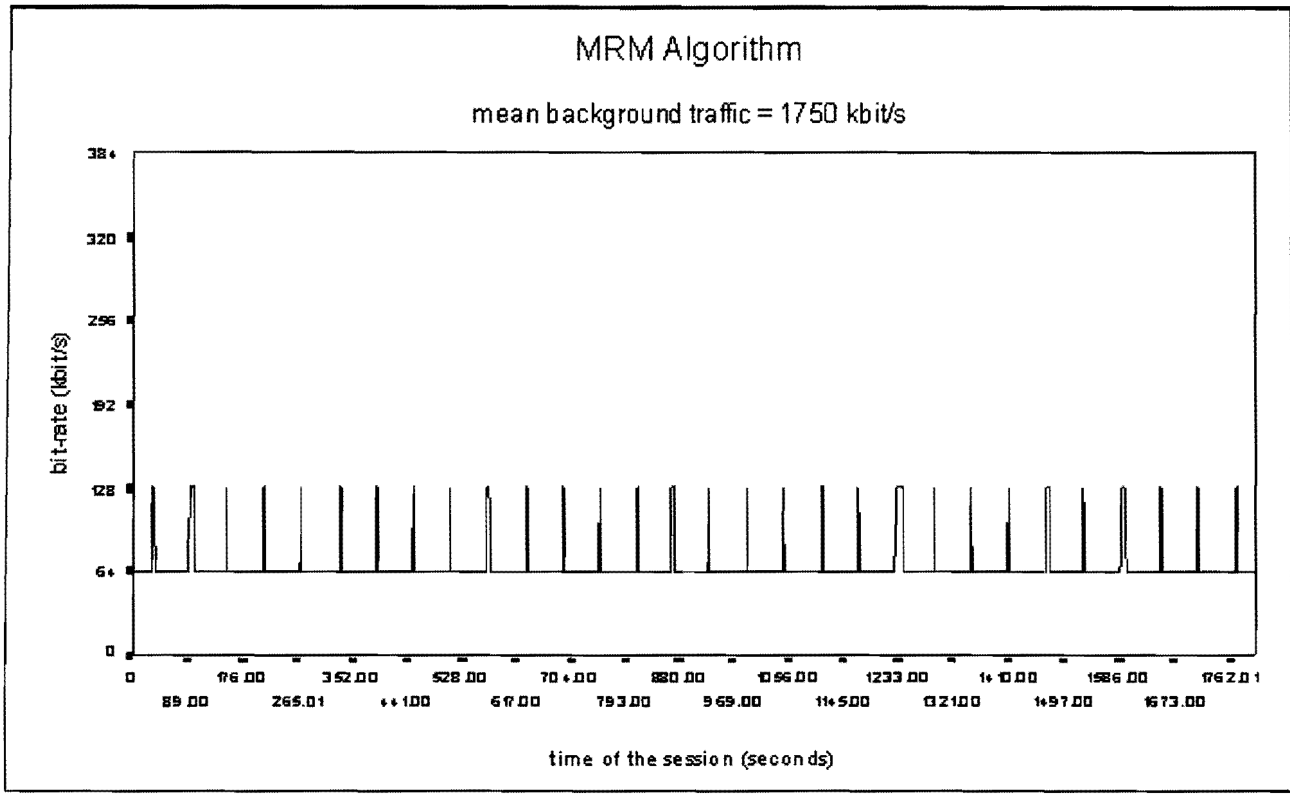


Figure 6: RRM Algorithm - mean background traffic 1750 kbit/s

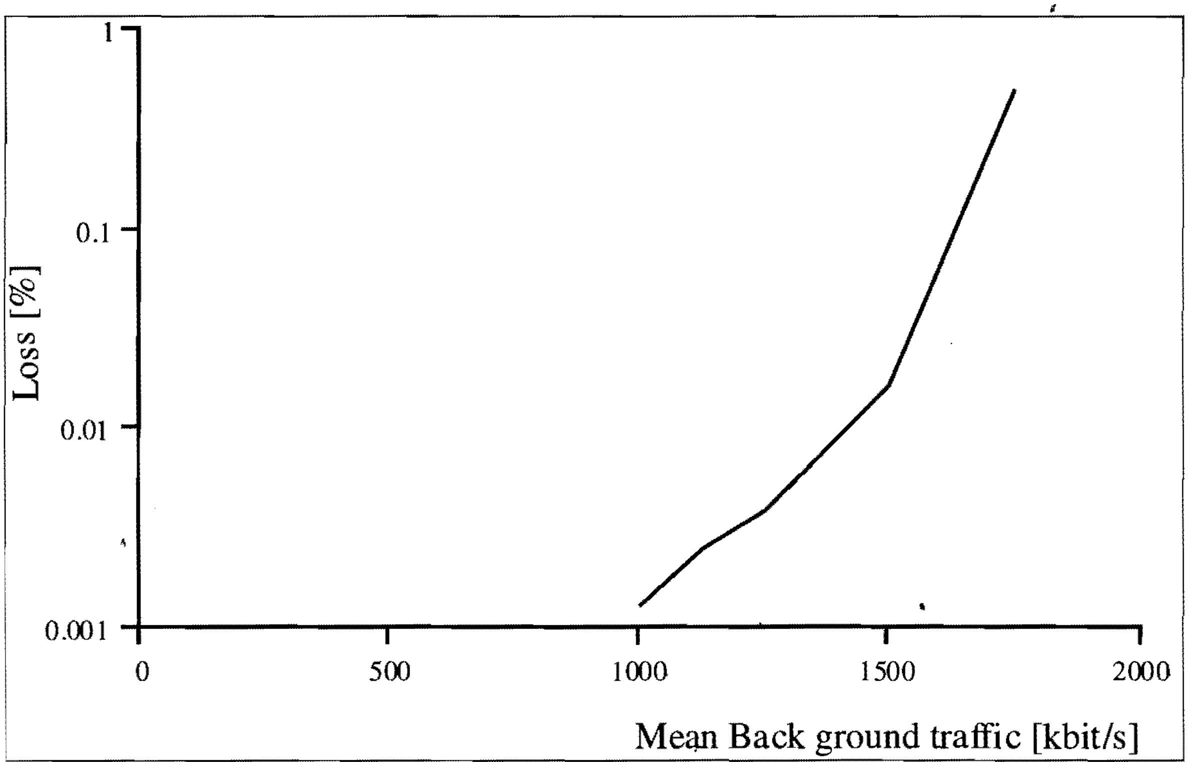


Figure 7: Percentage of packet loss from the best effort back ground source

4 Conclusions

The use of Internet to deliver real-time data streams is now possible using modifications to the basic assumptions of a packet-data network. Resource Reservations Management is a viable solution to the problem of making efficient use of the resources in the routers. It tries to keep a best effort model for the reservations by considering a minimum level of reservation. RRM tries to allocate as much as possible bandwidth but adjusts the reservation granted according to the congestion condition. With more than one source with reservation, the mechanism would request a reduction to the source with the largest resources granted.

RRM is not a protocol, it is a framework that defines the main functions and roles that should be included to support a layered system of reservations. RRM could be used as an enhancement to the actual RSVP.

The simulation results show that, provided video bitstream is generated by a multi-layer CODEC, it is always possible to deliver it, better than the lowest acceptable quality. The quality can vary according to the activity of the background traffic, but it never falls below the minimum requested one. Moreover, the imposition due to the priority mechanism, no packets from the transmitted video within the reserved bandwidth is lost.

However, these fluctuations in the perceived video quality should not be mis-interpreted by the system instability. What the proposed reservation mechanism does, is to try to maximise network utilisation for a better delivery of video. If fluctuation in quality appears to be annoying, then with the user choice, within the interval of frequent change in bandwidth, only the lowest one is used and the extra bandwidth is made available to other users (e.g. the best effort uses).

In accordance with the actual development of the Internet and the increasing interest in delivering audio and video over it, we consider that a Layered Reservations framework used by adaptive applications is a good alternative in the near future.

References

Assuncao P.A.A. and **M. Ghanbari**, *"Transcoding of single layer MPEG video into lower rates"*, *IEE Proceedings, Vision, Image and Signal Processing* , vol 144, no. 6, pp. 377-383, December 1997.

Braden R., L.Zhang, S. Berson, S.Herzog and **S.Hamin**, *"Resource Reservation Protocol (RSVP) — Version 1 Functional Specification"*. Internet Draft, February 1996.

Emmerson B. and **D. Greetham** , *"Adaptive multimedia applications will aid collaborative work.* Byte Magazine. July 1996.

Ghanbari M., *"Two-layer coding of video signals for VBR networks" IEEE Journal on Selected Areas in Communications*, Issue on Packet Speech and Video, 7:5 (June 1989) pp. 771–781.

Ghanbari M. and **V. Seferidis**, *"Cell loss concealment in ATM video codecs"*, *IEEE Transactions on Circuits and Systems for Video Technology* , Special Issue on Packet Video, 3:3 (June 1993) pp. 238-247.

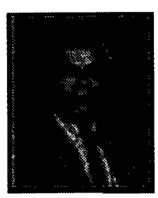
Hunt C., *"TCP/IP Network Administration"* 1994 O'Reilly & Associates, Inc. Sebastopol, CA. USA.

Thomas S. A., *"IPng and the TCP/IP Protocols Implementing the next generation Internet"*, Wiley, New York U.S.A., 1997

Peterson L. and **D. Bruce** *"Computer Networks: A Systems Approach.* 1996 Morgab Kaufmann. San Francisco, CA. USA.

Zhang L., S. Deering, D. Estrin, and **D. Zappala**, *"RSVP: A New Resource Reservation Protocol, IEEE Network Magazine*, March 1993.

Víctor Manuel González y González *works in the Universidad de Guanajuato since 1995 where leads the Unit of Information and Computer Systems. He is graduated from the Tecnológico de Monterrey Campus Querétaro(1994), where he studied an BSc in Electronics and Communications Engineering. He holds a Master in Science Degree from the University of Essex -Telecommunications and Information Systems (1997). Areas of interest: Multimedia, CSWC, Web-based education and Knowledge Management Technologies.*



Mohammed Ghanbari *is a professor in the Department of Electronic Systems Engineering with the main research interest in the areas of Video Networking. He was born in Saveh Iran in 1948 and graduated from Aryamehr University of Technology, Tehran, Iran with a BSc degree in Electrical Engineering in 1970 and MSc in Telecommunications and PhD in Electronics from the University of Essex, England in 1976 and 1979 respectively. After 10 years work in Radio and Television broadcasting he started his academic career in 1986 as a Research Fellow working on Video coding for Packet Networks. He was then appointed as a Lecturer at the Department of Electronic Systems Engineering, University of Essex, in 1988 and was promoted to Senior Lecturer then Reader in 1993 and 1995 respectively. He was appointed a personal chair in 1996.*