Resource Allocation and Quality of Service for Distributed Multimedia Applications

Carlos C. Goulart[*]
goulart@dpi.ufv.br
Department of Informatics
Federal University of Viçosa
36570-000 Viçosa–MG

José Marcos S. Nogueira
jmarcos@dcc.ufmg.br
Department of Computer Science
Federal University of Minas Gerais
31270-010 Belo Horizonte, MG

Abstract

This paper presents a model of dynamic resource allocation to provide Quality of Service (QoS) to distributed multimedia applications. A brief review of some concepts related to QoS for distributed multimedia applications is presented from two different points of view: ATM (Asynchronous Transfer Mode) and Internet. The model of dynamic resource allocation is intended to be a general model that can be mapped according to both points of view presented. A description of the implementations that were made to ATM and to RSVP is presented and the major results are revisited.

1 Introduction

Multimedia applications combine different types of information such as audio, video, text, etc., that can be transmitted in separated streams or into a single stream. Synchronization among these different information is always needed. In order to implement distributed applications it is necessary to achieve some basic requirements. These requirements are mainly related to the capacity of the network to transport large amounts of data and a set of protocols that does this transport in an efficient way.

The first wide area networks had low speed lines and the processing capacity of the routers was enough to support the demand generated by the transmission lines. The bottleneck of the networks was located in the transmission lines. With the evolution of transmission technology the bottleneck moved to routers. With the development of optic fiber technology, transmission rates increased substantially, while processing capacity of routers did not increase by the same rate. Thus, protocols have to be modified in order to achieve higher packet processing rates. Moreover, the protocols used in the Internet and OSI stacks do not provide the necessary functions for transporting multimedia data. Amongst their main deficiencies we can cite the following ones: absence of synchronization mechanisms for the different types of information; absence of multicast support; and absence of support for quality of service definition. Some new proposals of protocols try to solve this type of problem. In particular, two of the main initiatives that are being developed by the Telecommunication community, inside the scope of

[*]This work was supported by CAPES-PICD and CNPq-RHAE programs

the ATM Forum, and by the Internet community, in the scope of the IETF(Internet Engineering Task Force).

The Asynchronous Transfer Mode (ATM) has been used to deploy most wide area high speed networks in combination with Synchronous Digital Hierarchy (SDH) in the physical layer. The key point for obtaining a fast switching in ATM is the simplification of the packet (called cell) format. In ATM, the cell has a fixed-size format of 53 octets, and the cell header is composed by only 5 octets. The set of fields defined in the header was reduced, when compared to the fields of an IP packet header, and this helps in reducing the processing complexity of cell switching.

The work developed by the IETF was motivated by the definition of version 6 of IP protocol (IPv6). The major goal was to increase the range of IP addresses. In addition to the increase of address fields from 32 to 128 bits, the header format was simplified when compared to the format of version 4. This simplification also had the purpose of simplifying the processing complexity. The solution proposed by the IETF considers the use of IPv6 protocol and the development of fast IP routers.

Currently a scenario is one where ATM/SDH is being used to deploy wide area high speed networks. On the other hand the biggest world-wide network, the Internet, uses IP protocol. Therefore, most information that circulates around the world has the format of IP packets. Thus, the use of IP over ATM became a necessity for applications on wide area high speed networks. Again the old debate comes back to the spotlight, since ATM typically offers connection oriented service, while IP protocol offers a connectionless service. Our work was developed in such a way that it can be mapped into ATM technology or into protocols defined by the IETF.

The rest of this paper is organized as follows: section 2 presents a brief discussion about quality of service and resource allocation, where the views of the ATM model and the Internet model are presented; in section 3, the model of dynamic resource allocation is described in its major aspects; the implementation experiences of the model of dynamic resource allocation are described in section 4. The implementations made for ATM and RSVP protocol are presented; section 5 presents the conclusion relative to this work.

2 Quality of Service and Resource Allocation

Quality of Service (QoS) is defined by the ISO (International Organization for Standardization) as "the collective effect of the performance of a service, which determines the degree of satisfaction of a user of the service" [13]. As it can be observed, this definition of QoS is sufficiently generic and must be better defined for a specific problem. In the case of telephonic systems, for example, an average time lesser than 2 seconds to hear the dial tone can be a concrete parameter of QoS. In the case of multimedia applications, some parameters of QoS can have a subjective component, since audio and video quality are related to the user's perception.

Resource allocation is the way to guarantee the QoS of multimedia applications. Resource allocation politics can determine three types of guarantee: best effort, deterministic and statistic. For the best effort guarantee type no resource is allocated, and in this case, there is no guarantee at all. In the deterministic guarantee the maximum value of each resource is allocated for the connection lifetime and that connection will always have available resources. The statistic guarantee is a trade-off solution between the two others. In this case, the amount of allocated resource will correspond to some value between the average and the maximum value. This strategy uses the statistical multiplexing concept, where the amount of resource allocated to a specific connection can be shared among other connections. So, it is possible to admit a bigger number of connections using statistic guarantee than using the deterministic guarantee. For many multimedia applications this commitment solution is very interesting since the deterministic guarantee is not necessary.

Another classification of resource allocation strategies considers the moment of the allocation. The static allocation is made at the service initialization and the amount of allocated resources remains the same for the connection lifetime. In the dynamic allocation the amount of resources can change during the connection lifetime. A survey about QoS for distributed multimedia applications can be found in Vogel et alli [20].

2.1 ATM and Quality of Service

ATM technology is proposed to serve as support for a wide variety of services and applications. Traffic control in an ATM network is related to the ability of the network to supply Quality of Service (QoS) to its applications [2]. Thus, the QoS is directly related to the traffic management. The primary function of the traffic management is to prevent congestion in the network, which by itself would make impracticable any concept of QoS. An additional function is to promote efficient use of network resources.

The definition of Quality of Service (QoS) in the ATM layer is made through the definition QoS parameters. One or more of these parameters can be defined to each connection to attend to the goals of performance requested by the user and those offered by the network.

The current ATM service architecture, defined by the ATM Forum, specifies five categories of service::

1. Constant Bit Rate (CBR): it is used by connections that request a fixed amount of bandwidth, which is available during the connection lifetime. This type of service is characterized by the Peak Cell Rate (PCRand it is intended to support real-time applications with strong restrictions in delay variation. The type of guarantee offered by CBR service is the deterministic guarantee;
2. Real-Time Variable Bit Rate (rt-VBR): it is intended for real-time connections with serious restrictions of delay and delay variation. These connections are characterized in terms of the PCR, Sustainable Cell Rate (SCR) and Maximum Burst Size (MBS). A burst corresponds to a period of time where the connection transmits using rates greater than its average rate. The guarantee for the parameters related to delay must be deterministic. The guarantee for the other parameters can be the statistic guarantee;
3. Non-Real-Time Variable Bit Rate (nrt-VRB): it is intended for non-real-time applications with burst traffic and it is characterized in terms of the PCR, SCR and MBS. Delay restrictions are not imposed. The type of guarantee offered by the service nrt-VBR is the statistic guarantee;
4. Unspecified Bit Rate (UBR): this service is intended for non-real-time connections. The network may or may not apply the PCR to CAC and UPC functions. Service UBR is indicated for connections with the best effort guarantee type;
5. Available Bit Rate (ABR): it is a category of service for which the network load status can influence in future connection establishments. ABR Service is not intended for real-time applications. In the connection establishment phase the user must inform the values

of PCR and the Minimum Cell Rate (MCR). The guarantee offered by ABR service is either deterministic for rates lesser or equal to the MCR or best effort for rates greater than MCR.

A document of the ITU-T [14] also defines 5 classes of service for ATM-based broadband Networks. The major difference is the definition of ABT (ATM Block Transfer) service. ABT Service introduces the concept of block of cells and it transports a block of cells with low cell loss ratio and with small delay variation. ABT Service uses on demand bandwidth allocation for the transmission of each block. In the connection establishment phase no bandwidth is allocated. Except for ABT service, all ATM service classes use static resource allocation at the connection establishment. This fact is explained by the philosophy of the development of ATM that is typically a connection oriented model.

2.2 Quality of Service on the Internet

Internet protocols traditionally offer best effort guarantee, following the datagrama service philosophy. Even IPv6, which have support for the definition of QoS, is basically a routing protocol and it does not implement QoS mechanisms. In this case, the QoS must be offered by protocols above IP. Some proposals are being analyzed by IETF work groups. Amongst these proposals, Reservation Protocol (RSVP) [11] is one possible protocol to be used to offer the QoS on the Internet. RSVP is used by a host, representing a connection, to request from the network a specific QoS level for a particular data stream. RSVP is also used by routers to send QoS control commands to all nodes in the connection's path in order to establish and keep the state that allows offering the requested service.

RSVP Protocol operates over IP protocol (version 4 or version 6), occupying the place of the transport protocol in the protocols stack. However, the RSVP does not transport application data, it is a control protocol of the Internet as IGMP, ICMP, or routing protocols. RSVP is not a routing protocol, but it was designed to be used with current and future routing protocols (point-to-point and point-to-multipoint). The RSVP is only responsible for the QoS of those packets that are forwarded in accordance to the routing.

In order to accomodate greater groups, with dynamic inclusions, exclusions and receivers with different requirements, RSVP leaves the responsibility for the QoS requests for the receivers. The QoS requests are directed in the reverse path of the data stream. At each node, a QoS request is passed to two local decision modules: admission control and politics control. The admission control module verifies if there is enough resource to support the QoS request and the politics control module determines if the user has permission to make the request. If one of the two controls fails the request will not be accepted.

2.3 Related Work

One of the first proposals for supporting QoS in services based on the OSI reference model is presented in Sluman [19]. As result of this and other initial works, ISO (International Organization for Standardization) issued a document with guidelines for the definition of QoS requirements for the OSI reference model [13]. Vukotic and Niemegurs [21] presented a study to adapt the upper layers of OSI reference model to support multimedia communications.

Service guarantee on integrated service networks has been studied since the beginning of this decade [17, 18] when major differences among the three types of guarantee were presented:

real-time services (deterministic guarantee); services that carry continuous media (statistic guarantee); and services without guarantee (best effort).

Resource reservation is fundamental to guarantee the QoS and this point has been presented in diverse works. In RFC1190 [10], ST-II protocol (Internet Stream Protocol) is presented. It was developed for audio and video applications using IP packets. SRP Protocol (Session Reservation Protocol) [1] was also considered for use on the Internet and allows reserving resources at terminal and intermediate nodes. RSVP Protocol [11] can be used in group communication with reservation being initiated by data receivers. RSVP was adopted by IETF for supporting QoS in integrated services approach. Another proposal in consideration by IETF for real-time applications on the Internet is presented in Jacobson et alli [9].

A lot of work has been developed in order to provide an environment that supports multimedia applications with quality of service guarantees. The solution proposed in Tenet protocols suite [22] consists of a protocol set that allows to establish and operate real-time channels on high speed networks of arbitrary topology.

An integrated architecture of protocols with support for quality of service is proposed in the QoS-A architecture[4]. QoS-A considers an integrated

vision of QoS including end-systems and the communication network. QoS-A uses the concept of stream to characterize the production, transmission and eventual consumption of a sequence of data of a single media as an integrated activity governed by a particular specification of QoS. Another approach is taken in Omega architecture which supports the definition of thresholds for transit delay and deals with bandwidth demands from clients [8].

3 A Model for Dynamic Resource Allocation

The proposed model for dynamic resource allocation is intended for applications where the connection lifetime can last for long periods. Video on demand is one of such applications, which can have short video clips and also long movies. During the connection lifetime, the connection requirements related to specified QoS level can vary substantially, according to the scene changes and quantity of movement. So the amount of resources must also be dynamically allocated.

The model of dynamic resource allocation [7] uses the definition of a minimum QoS level as its major point. The level of QoS is specified as function of some QoS parameter values. The initial negotiation must take into account the specified minimum value, as well as the peak value and the initial value for each QoS parameter. The key point in our approach is to use dynamic negotiation of resources. Although the model is generic, at this time, only bandwidth allocation is being considered for each connection.

The other parameters that must be considered are CPU cycles and end-to-end transmission delay. CPU cycles reservation is directly related to the ability of operating system at each local node to make reservation of this type of resource. In the case of transmission delay, this value is considered as an lower limit to define the time space between two consecutive QoS renegotiation requests. Thus, a connection request must specify two values for the bandwidth:

1. The minimum value of bandwidth to be allocated in a permanent form for the connection, referenced in the expressions to follow as $minimum\_BW_i$, where i indicates a particular connection;

2. The initial value for the bandwidth that is used together with the minimum value by each node's admission control algorithm. This parameter is referenced in the expressions to follow as $initial\_BW_i$;

The connection admission control algorithm and the of dynamic resource allocation algorithm are computed in a distributed way in the system. Thus, the computation must be made at each node j where a connection i passes. The criterion to accept a new connection is based on the following comparison:

```
if (total_BW(j) - allocated_BW(j)) >= initial_BW(i)
     accept_connection(i);
else reject_connection(i);
```

where:

$$allocated\_BW_{(j)} = \sum_{i=1}^{n} \begin{cases} minimum\_BW_{(i)}, & if\ current\_BW_{(i)} < minimum\_BW_{(i)} \\ current\_BW_{(i)}, & otherwise. \end{cases}$$

$total\_BW_{(j)}$ is the total channel capacity of a node (j) where the connection (i) is being admitted and $current\_BW_{(i)}$ is the amount of bandwidth that a connection (i) is using at the current period of time.

The value of $allocated\_BW$ is calculated for all the n active connections of node j. When a connection is using less resource than its $minimum\_BW$, the computation considers it as equal to $minimum\_BW$. To accept or to reject a renegotiation request, the comparison is made using the values of bandwidth that are in use $(current\_BW)$. The criterion to accept a renegotiation request considers the available bandwidth, as in the following comparison:

```
if (total_BW (j) - current_BW (j)) >= requested_BW(i)
     accept_renegotiation;
else reject_renegotiation;
```

where:

$$current\_BW_{(j)} = \sum_{i=1}^{n} current\_BW_i.$$

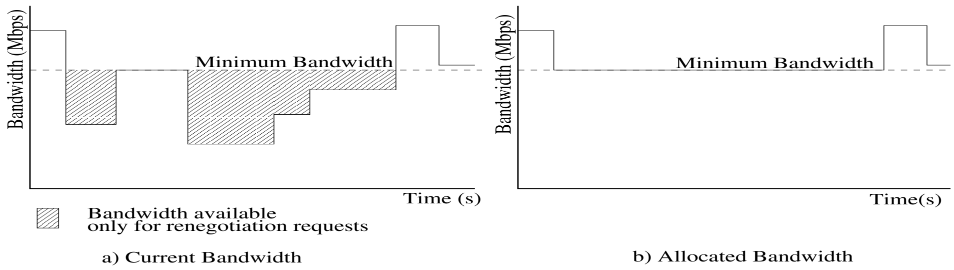

Figure 1: Minimum QoS reservation

In this case, the value of available bandwidth is always greater than or equal to the case of requesting a new connection. This occurs because the value of $current\_BW$ is always lesser than or equal to the $allocated\_BW$. The concept of available resource for each case is different. If one or more connections are using less bandwidth (their $current\_BW$) than their minimum value ($minimum\_BW$), then there is a certain amount of resource that can be used for the resource allocation algorithm and that can not be used for accepting a new connection. This situation is illustrated in figures 1.a and 1.b, for a given connection i.

Therefore, the use of the minimum level of the QoS serves to limit the number of accepted connections and at the same time it increase the probability of success for renegotiations, since a parcel of the resources is reserved exclusively for use of renegotiations. Thus, the use of minimum QoS level defines an implicit priority for renegotiations over new connection requests.

The proposed model was simulated for a service of video with connections with variable duration and with Variable Bit Rate (VBR) behavior. The proposed model was compared to a method of static allocation using as value of allocated bandwidth a value between the peak rate and the value of the average rate, in order to also offer a statistic guarantee. For the considered scheme of dynamic bandwidth allocation a maximum value for the reserved bandwidth (Res_max) was fixed. In this case, when the maximum value is equal the 4.0 Mbps, it means that the connections can reserve from zero to 4.0 Mbps. The figure 2 shows a comparison between the method of dynamic allocation (with different values for maximum reservation) and the method of allocation using the average bandwidth. The percentage of renegotiation rejects, which can represent packet loss due to lack of resource, was determined for different values of network load represented by the number of requested connections.

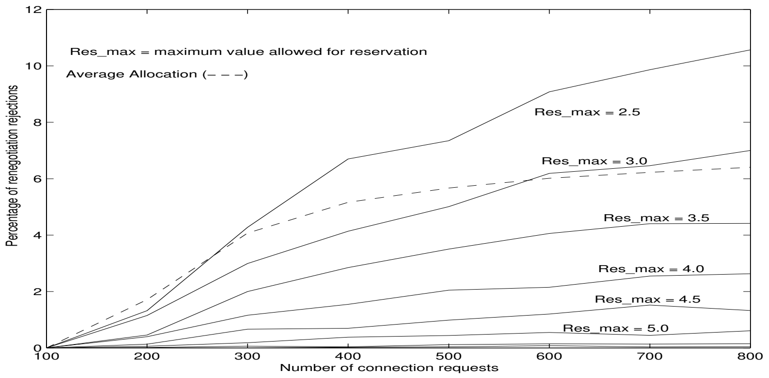

Figure 2: Comparison between dynamic allocation and average allocation

The simulation results had shown that for situations where the traffic load is low or medium the two models have a similar behavior, or either, both admit the same number of connections and the packet loss ratio due to lack of resource is very small. However, when the traffic load is high the model of dynamic allocation offers a higher level of guarantee for the accepted connections. The method of dynamic allocation restricts the number of accepted connections faster than the method of static allocation, but it guarantees a loss ratio below of 1%. The preliminary results of the simulation can be found in [6] while the complete results can be seen in [7].

4 Implementation of Dynamic Resource Allocation Model

4.1 Implementation using ATM

An alternative for the implementation of dynamic bandwidth allocation scheme is to use the services available in UNI(User Network Interface). It is possible to define service primitives and PDUs (Protocol Data Units)associated to the primitives. The implementation of a service for dynamic bandwidth negotiation in an ATM network is described in [3].

The set of service primitives was implemented inside a QoS Layer on top of the ATM Layer. This QoS layer allows a multimedia application to use the dynamic bandwidth allocation in an ATM network. For definition and implementation of QoS dynamic negotiation layer, the concepts of service and protocol had been used in accordance to OSI information model.

The service primitives offered to the user are: Setup_Connection, More_Resource, Release_Resource, and Finish_Connection. The UNI used in this implementation was UNI version 3.0 that does not allow renegotiation of QoS parameters during connection lifetime. Thus, the service of dynamic bandwidth allocation was implemented through CBR service. For each time interval between renegotiations, the value of the peak rate is updated taking into account the maximum rate for the next period. So, a new CBR connection is established with this new value as the Peak Rate, and the values allocated for the current connection is released. This control was made by an algorithm of connection admission control (CAC) implemented in a workstation instead of using the CAC algorithm of ATM switch. This was necessary because the ATM switch used in this experiment has a proprietary implementation, that does not allow modification in its CAC algorithm. The work was developed at the High-speed Network Laboratory of the Federal University of Minas Gerais. All the equipment used in this implementation follow the ATM Forum recommendation UNI version 3.0.

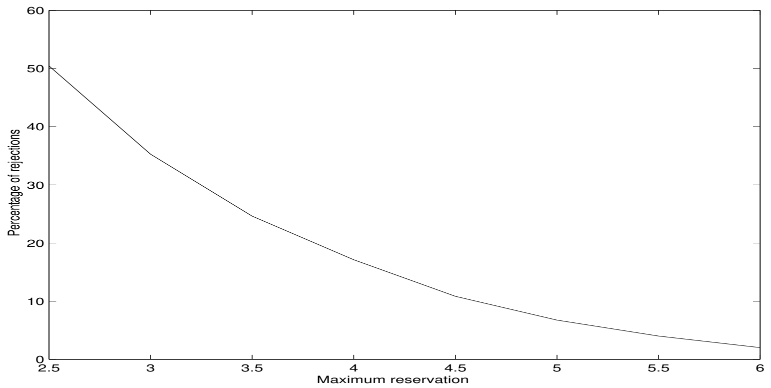

Figure 3: Renegotiation rejections

The results obtained with this implementation had been compatible with the simulation results, considering the differences between the two experiments and the restrictions imposed by the equipment. The figure 3 shows the behavior of the percentage of rejections as function of the value of maximum bandwidth reservation allowed to each connection. As in our simulations, the increase in maximum reservation reduces drastically the renegotiation rejection rate. The complete results relative to this implementation can be found in [3]. The service of dynamic bandwidth allocation was also used in the implementation of a service of video transmission using MPEG coding [16].

4.2 Implementation using RSVP

The philosophy of the RSVP is very similar to the philosophy of dynamic resource allocation model. Both use the idea of updating the load status that will be demanded by the network in the next period of time. The dynamic resource allocation model strengthens the idea of reservation based on the minimum quality allowed (minimum reservation). In the case of RSVP protocol, all requests either for more resource or only to confirm the current amount of resource, brings the value of bandwidth that will be necessary in the next period of time. Thus, it is necessary to make a small change in the reservation requests in their origin. A comparison between the requested value and the minimum value must be made. Every time the requested value is lesser than the minimum value, the stream description must be modified and the minimum value must be requested. This solution is facilitated by the merge mechanism of RSVP.

Another important detail is to guarantee that the requests of resource are generated in time intervals that can guarantee the maintenance of the soft state defined for the RSVP. In case where the time interval between two requests for resource is bigger than the limit interval defined by RSVP, extra requests must be generated to satisfy the specified limit. These extra requests must have identical values of last request, only to guarantee that the reservation done will not be cancelled.

RSVP protocol uses two main classes of messages: (1) *Resv* that is used by the receivers to make the QoS requests; and *Path* that is sent by the sender of data flow in order to keep and update routing information. As the RSVP was defined to be used with different lower level protocols, the format of its messages is defined in generic way. Instead of specifying formats based on fields, the specification is made in terms of objects.

Figure 4 shows the general format of RSVP messages. The Common Header has a fixed format for all messages. The *Vers* field identifies the protocol version number. The field *Flag* is not defined yet. The *Msg Type* field identifies the type of message (1 = *Resv*). RSVP Checksum is used for error detection. The *Send_TTL* field identifies the message lifetime in the origin. *RSVP Length* field specifies the total size of the message, including the header.

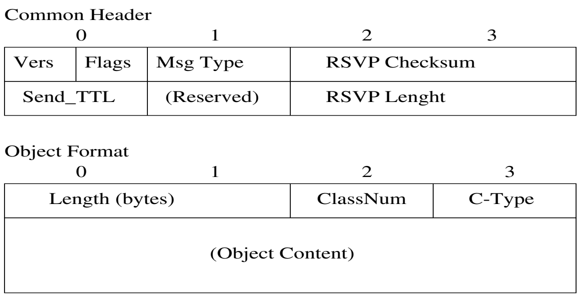

Figure 4: RSVP message format

The remainder of the message (body) has variable size and is composed by objects of any defined type. The format of each object is composed by a header (4 octets) and the content of the object which has a variable size. The header fields specify the total size of the object in octets (*Length*), the object class (*Class_Num*) and the type of the object that must be unique inside a class. The format of the content depends on the class and type of considered object.

RFC 2210 from the IETF [12] specifies the object format to be used inside of the vision of Internet integrated services. The model adopted for the QoS in the Internet allows the specification of two types of service: (1) the guaranteed QoS service (*Guaranteed QoS*) where the offered guarantee is deterministic; e (2) the service that offers a statistic guarantee for the QoS (*Controlled Load*). The format defined for objects can be used with any one of the two types of service.

Three types of objects are defined to represent the three different QoS views that exist in the system. These three objects can be used for the implementation of dynamic bandwidth allocation service, as described below:

1. *SENDER_TSPEC* it contains the specification of the traffic that the sender is going to generate. This message can be used so that the traffic source informs the traffic characteristics to be generated. In the case of a group communication application, the information can be coded in accordance to the definition presented in RFC2210 [12], where the traffic is characterized by information such as, peak rate, average rate, burst size, and so on, for a control based on the leaky bucket algorithm;
2. *ADSPEC* can be generated by the sender of data stream or by an intermediate node to indicate the lack of resource. This message is sent to the receivers and it is used to inform about a failure in allocating resources. So, all nodes between the receivers and the node where the failure happened, that had already reserved resources can release them. This type of message is used, also, to update routing information.
3. *FLOWSPEC* has information relative to resource reservation that can be made by the receivers. The values contained in this object also can be updated by intermediate nodes before arriving at the sender of data stream. The messages of this type will be generated by the receivers, based on the information received in *SENDER_TSPEC*. These messages will be forwarded to the sender so that resource reservation can be made along the path. The intermediate nodes must process the information contained in *FLOWSPEC* to either allocate or release resources. In case of lack of resource, the intermediate node can modify the information of *FLOWSPEC* before directing it to the next node in the path. In case of lack of resource, the intermediate node must also generate an *ADSPEC* message and send it in the reverse path in order to release resources previously allocated.

The dynamic resource allocation scheme was implemented using RSVP source code available at Information Sciences Institute (http://www.isi.edu/div7/rsvp) There are three public versions of RSVP code for different operating systems (Solaris, FreeBSD and IRIX). The version used in this implementation was for Solaris 2.5.1. The use of RSVP for implementing the dynamic bandwidth allocation scheme showed that it can increase the resource utilization on the network, and at the same time to offer a high level statistic guarantee (near 97%). The major limitation found in this first implementation was the incompatibility of the operating system with the resource reservation functions of RSVP. The operating system version used does not allow making resource reservation inside its kernel. Thus, the resource reservation requested was simulated. Detailed information about this implementation and the major results can be seen in [15].

5 Conclusions

Distributed multimedia applications demand a series of requirements from the network to be implemented. The concept of Quality of Service associated with the distributed multimedia applications requires that the tasks of reserving and allocating resources on the network be made in such a way to guarantee the existence of resource during the connection lifetime. Different solutions have been presented to solve the problem of guaranteeing QoS requirements for distributed multimedia applications. Particularly, the proposals developed in the scope of the ATM Forum and the IETF.

This work presented a model for dynamic resource allocation to guarantee the QoS for distributed multimedia applications and the results obtained in its implementation. The model was implemented for dynamic bandwidth negotiation, in order to better support connections with variable bit rate behavior. The main points of the model are: the definition of a minimum level for the bandwidth and the renegotiation of the allocated value during the connection lifetime. The model was defined in generic way in order to be mapped into different technologies. The model was simulated to verify its effectiveness in guaranteeing the QoS level of accepted connections.

The mapping for ATM was made through the implementation of a service of dynamic bandwidth negotiation using an API supplied by the manufacturer of the ATM switch. A set of primitives was defined to allow the negotiation of the bandwidth parameters of a connection. The results of this implementation were compatible with the results obtained in the simulation. Later this service of dynamic bandwidth allocation was used for implementing a service of MPEG video transmission.

The mapping for IETF protocols was made using RSVP protocol, that has a certain similarity with the proposed model, since both deal with dynamic resource allocation. The use of RSVP for the implementation of the dynamic bandwidth allocation scheme showed that it is possible to offer to a high level statistic guarantee when the network load increase.

References

[1] D. P. Anderson, R. G. Herrtwich, C. Schaefer, *SRP: A Resource Reservation Protocol for Guaranteed Performance Communication in the Internet*, Internal Report, International Computer Science Institute at Berkeley, 1991.

[2] The ATM Forum, *Traffic Management Specification - Version 4.0*, The ATM Forum, April, 1996.

[3] A. L. B. P. B. Diniz, C. C. Goulart, J. M. S. Nogueira, *A Dynamic Resource Allocation Service for ATM-based Networks Supporting Multimedia Applications*, In Proceedings of IEEE International Conference on ATM, France, June, 1998.

[4] A. Campbell, G. Coulson, D. Hutchison, *A Quality of Service Architecture*, ACM Computer Communication Review, vol. 24, n. 2, April, 1994.

[5] D. D. Clark, S. Shenker, L. Zhang, *Supporting Real-time Applications in an Integrated Services Packet Network*, Proceedings of the ACM SIGCOMM'92, pp. 14-26, Baltimore, USA, August, 1992.

[6] C. C. Goulart, J. M. S. Nogueira, G. W. Neufeld, *Dynamic QoS Negotiation for Multimedia Applications at Intermediate Nodes*, In Proceedings of IEEE Singapore International Conference on Networks, pp. 323-338, Singapore, april, 1997.

[7] C. C. Goulart, *A scheme for dynamic resource allocation to guarantee QoS for distributed multimedia applications*, PhD thesis (in Portuguese), Department of Computer Science, Federal University of Minas Gerais, April, 1998

[8] A. Hafid, G. v.Bochmann, B. Kerhervé, R. Dssouli, J. Gecsei, *On Quality of Service Negotiation for Distributed Multimedia Applications*, Publication Départamental n. 977, Départment IRO, Université de Montréal, May, 1995.

[9] V. Jacobson, S. Casner, R. Frederick, H. Schulzrinne, *RTP: A Transport Protocol for Real-Time Applications*, Internet Draft ietf-avt-rtp-new-00.txt, October, 1999.

[10] IETF – Internet Engineering Task Force, *RFC 1190 – Experimental Internet Stream Protocol*, version 2, October, 1990.

[11] IETF – Internet Engineering Task Force, *RFC 2205 – Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification*, September, 1997.

[12] IETF – Internet Engineering Task Force, *RFC 2210 – The use of RSVP with IETF Integrated Services*, September, 1997.

[13] JTC1/SC21/WG1. *ISO/IEC DIS 13236 – Quality of Service Framework - Outline*, February, 1992.

[14] ITU-T. *Recommendation I.371 - Traffic control and congestion control in B-ISDN*, August, 1996.

[15] A. A. F. Loureiro, V. L. Santos, C. C. Goulart, J. M. S. Nogueira. *A Study of Dynamic QoS Negotiation for Multimedia Applications in RSVP*, In Proceedings of VI IEEE/IFIP International Workshop on Quality of Service, Napa, CA, USA, May, 1998.

[16] F. F. N. Pereira, J. M. S. Nogueira. *Transmission of MPEG Video over ATM-based Networks utilizing Dynamic Bandwith Negotiations*, In Proceedings of Multimedia Management Network Systems Conference, Versailles, France, 1998.

[17] S. Keshav, *Report on Workshop on Quality of Service Issues in High Speed Networks*, ACM Computer Communication Review, vol. 22, n. 1, January, 1992.

[18] J. F. Kurose, *Open Issues and Challenges in Providing Quality of Service Guarantees in High Speed Networks*, ACM Computer Communication Review, vol. 23, n. 1, January, 1993.

[19] C. Sluman, *Quality of Service in Distributed Systems*, British Standards Institution, UK, October, 1991.

[20] L. A. Vogel, B. Kerhervé, G. von Bochmann, J. Gecse, *Distributed Multimedia and QOS: A Survey*, IEEE Multimedia, pp. 10-19, Summer, 1995.

[21] M. Z. Vukotic, I. G. Niemegurs, *Multimedia Communication Systems: Upper Layers in the OSI Reference Model*, IEEE Journal on Selected Areas in Communications, vol. 10, n. 9, February, 1992.

[22] L. Zhang, D. Ferrari, *Design and Implementation of the Real-time Internet Protocol*, IEEE Workshop on the Architecture and Implementation of High Performance Communication Subsystems, USA, February, 1992.