INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES

1159

PUBLICATION INTERNE
N°

PROPOSING A NEW QUEUE SERVICE SCHEME FOR ABR MULTICAST FLOW

MUDASSIR TUFAIL AND BERNARD COUSIN

ISSN 1166-8687

IRISA
CAMPUS UNIVERSITAIRE DE BEAULIEU - 35042 RENNES CEDEX - FRANCE

INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES
Campus de Beaulieu – 35042 Rennes Cedex – France
Tél. : (33) 02 99 84 71 00 – Fax : (33) 02 99 84 71 71
http://www.irisa.fr

Proposing a new queue service scheme for ABR multicast flow

Mudassir TUFAIL and Bernard COUSIN

Thème 1 — Réseaux et systèmes
Projet Adp

Publication interne n ° 1159 — Août 1997 — 6 pages

Abstract: This paper proposes a queue service scheme destined to multiplexed unicast and multicast connections, pertaining to the ABR (Available Bit Rate) traffic class, and is implemented at the output ports of an ATM switch. Our proposed service scheme has the following important features: (1) each multicast queue, among replicated ones, is attributed a certain priority (high/low) in order to absorb the diversity in their respective service rates at different output ports, (2) this priority attribution to multicast queues does not, in any way, restrict unicast (even of the same traffic class) queues to get their fair share of bandwidth, (3) for all other queues belonging to other traffic classes the role of this service scheme is totally transparent, (4) the work required by the service scheme is of the order of n where n is the number of cell queues at the output port. We present simulation results of this service scheme when implemented to multiplexed unicast and multicast connections of ABR class.

Key-words: queue service scheme, fair queuing, multicast, ABR traffic class, homogeneous service.

(Résumé : tsvp)

email: {mtufail, bcousin}@irisa.fr

Résumé : Ce papier propose une discipline de service de files d'attente destinée aux connexions, appartenant à la classe de trafic ABR, multiplexées unicasts et multicasts. Cette discipline est mise en œuvre aux portes de sortie d'un commutateur ATM. Notre discipline a les importantes propriétés suivantes: (1) a chaque file d'attente multicast (pour celles qui sont dupliquées) est attribuée une certaine priorité (haute/basse) afin que les variations, dans le taux de service correspondant aux différentes portes de sortie, puissent être absorbées, (2) cette attribution de priorité aux connexions multicasts, dans aucun cas, empêche les connexions unicasts de profiter de leur part de bande passante, (3) pour les files d'attente, appartenant à d'autres classes de services, le rôle de cette discipline est totalement transparent, (4) la discipline exige un travail de l'ordre de n où n est le nombre de files d'attente présentes à une porte de sortie. Nous présentons les résultats de simulation de cette discipline lorsqu'elle est mise en œuvre pour les connexions multiplexées unicasts et multicasts appartenant à la classe ABR.

Mots clés : Discipline de service de files d'attente, ordonnancement équitable, multicast, classe de trafic ABR, service homogène.

1 Introduction

The congestion control is one of the important issues of ABR traffic class of ATM. An ABR source adapts its rate to the changing network conditions. Resource Management (RM) cells are emitted, periodically, by an ABR multicast source and are returned back by all the destinations[1] carrying the network congestion information to the source [2, 3]. The source is expected to modify its emission rate on the reception of RM cells. In an ABR multicast connection, source will receive back as many RM cells as number of destinations. The question here is that source adapts its emission rate at which instants and after having received how many RM cells? The ATM Forum [4] proposes that a switch may send fewer RM cells upstream than it has received back from destinations. We proposed in [3] to amalgamate the RM cells at switches so that a source receives back one RM cell and reacts as if the connection is unicast. The next question is that how to amalgamate the values of returned RM cells to one cell? The safest option is to consider the minimum value of rate, among the heterogeneously served replicated multicast queues, as proposed in [6] but this results in under-utilization of available resources. Therefore it becomes indispensable to equip the ATM switches with an intelligent service discipline which can assure a homogeneous service to all the replicated queues of a multicast connection.
The rest of the paper is structured as follows: We present the arbitration algorithm in section 2 and its working principles in section 3. We simulate the algorithm and analyze its result in section 4. The section 5 presents the concluding remarks on the paper.

2 The Arbitration Algorithm

At a switch[2] the incoming cells of an input port are led to different queues at output ports. The cells, belonging to a multicast connection, are replicated and are forwarded to their corresponding output port queues. Separate queues are maintained for different connections at an output port. Thus we have separate queues for unicast and multicast connections. We define two types of priorities, Normal Priority (NP) and Multicast Priority (MP), which are elaborated in the following sections.

2.1 Normal Priority (NP)

NP is assigned to all the ABR queues, regardless of their nature, unicast or multicast. The NP is function of following parameters and is updated every cell slot time.
• The percentage of buffer occupied by the queue.
• The number of times that a cell of the queue has been refused to be served by the scheduler.
• The rate at which the queue should be served. In other words, the fair share of queue for the available bandwidth.

[1]We are considering the point-to-multipoint connections only.

[2]In this paper, a switch means the switch where path segregation of multicast connection occurs.

We define the fair share on *max-min* criteria [7] which provides all VCs (Virtual Channels), that have a “low” demand of the capacity of resource, their entire requirement. The VCs, which have a “higher” demand, are provided at least an equal share of the left over bandwidth.
• The parameter MCR declared at the connection establishment time.

2.2 Multicast Priority (MP)

The MP is assigned to multicast ABR queues only. MP has three states: *Active* state, *Sleep* state and *Neutral* state. The attribution of a certain MP state to a multicast queue is associated to the service rates of its replicated queues at other output ports of the switch which means that the lengths of all replicated multicast queues, belonging to the same multicast connection, are to be measured at each cell slot time thus increasing the scheduling complexity by the order of m, where m is the number of replicated multicast queues. In order to avoid this additional complexity, we take help of the following parameters: max_threshold, min_threshold, α and β. The parameters max_threshold[3] and min_threshold[3] are defined for switches disposing output buffers and is same for all the queues. Where as the parameters α and β are connection dependent and are calculated independently, for each replicated multicast queue, as described below. Following are the important relations which will be employed to determine the values of α and β for the replicated queues of an ABR multicast connection with non-zero MCR value.

$$\begin{aligned} size_{buffer_i} &= buffer_{ABR} * queue_i / \Sigma_{j=0}^{n} queue_j \\ A_i &= rate_{emission_i} / rate_{fair_i} \\ B_i &= max\{size_{burst}, RTT_i\} \end{aligned}$$

All the multicast queues are, initially, in *Neutral* MP state.

2.2.1 *Active* state

A multicast queue i whose length exceeds α_i gets its MP in *Active* state. The value of α_i depends upon the multicast connection type and is calculated as:

$$size_{buffer_i} > \alpha_i \geq max\_threshold * A_i * B_i * rate_{emission_i}$$

An *Active* MP state means that the concerned multicast queue is to be served at top priority as long as other queues (unicast/multicast) on this output port are not deprived of their fair share.

2.2.2 *Sleep* state

A multicast queue i gets MP in *Sleep* state if its length falls behind the β_i value. The value of β_i is bounded as :

$$0 < \beta_i \leq \text{min\_threshold} * A_i * B_i * rate_{emission_i}$$

[3]It is given in percentage.

$buffer_{ABR}$	buffer available for ABR class
$size_{buffer_i}$	buffer available for queue i, calculated proportionally to queue length, regardless to MCR values of different ABR connections present at the output port
$queue_i$	current length of queue i
n	number of ABR queues (unicast/multicast) present at an output port
$rate_{emission_i}$	Current Cell Rate (CCR), as defined in [4], of queue i
$rate_{fair_i}$	fair share of queue i
MCR_i	Minimum Cell Rate which is guaranteed to an ABR (multicast/unicast) queue i
$size_{burst}$	Average burst size (in units of time) for the multicast ABR connection, taken zero if unknown
RTT_i	Round Trip Time (from source to destination) for queue i

Table 1: Notations used in this paper

A *Sleep* MP state means that the concerned multicast queue is to be served at lowest priority i.e. it is not be served as long as there are cells in other queues (unicast/multicast) on this output port.

2.2.3 $Neutral$ state

If none of above two cases are valid then the replicated multicast queues, on an output port, continue to have *Neutral* MP state. In this mode, the scheduler on the output port sees no difference between unicast and multicast queues.

2.3 Important Features

We would like to highlight the following important features of the arbitration algorithm.
• In order to ensure the fair distribution of bandwidth (fair share) among the queues, whether unicast or multicast, the algorithm will always serve the cell from a queue which tends to be full and its NP value is the largest among all those present at this output port. It is done regardless of MP values of multicast queues.
• The priority attribution is restricted to be effective within the traffic class, the queues belong to. It will, in no case, affect the service schedulers of other traffic classes for example those of Constant Bit Rate (CBR) traffic class and Variable Bit Rate (VBR) traffic class.
• As long as the additional bandwidth used by an ABR queue (unicast/multicast) is not at the expense of other ABR connections (i.e. if the other ABR connections are idle for the moment), it will not be penalized in the next scheduler's cycles by reducing its bandwidth allocation.

3 Working Principle

Each unicast queue is associated with NP and each multicast queue with both NP and MP values. Once the values of NP and MP are updated, at each cell slot time, the scheduler at each output port behaves as follows. For the sake of simplicity, we have assumed similar values of α_i and β_i for all replicated multicast queues (refer to the figure 1).
• A fair share of bandwidth is calculated for each queue contending for the same output port. At each serving of queue, its share value is updated (NP update) and scheduler then picks the eligible queue determined by its NP value.
• At the same time the scheduler may be directed (by MP value) to deviate, temporarily, from fair share principle.

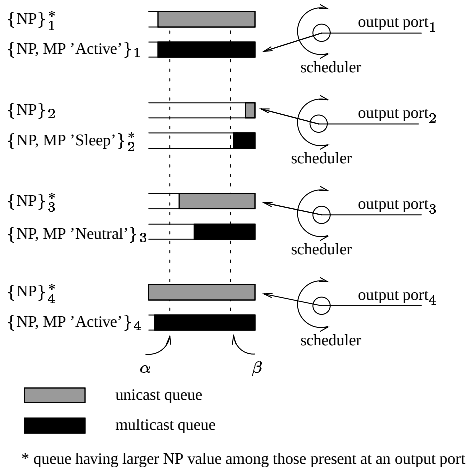

Figure 1: The scheduler's response to NP/MP values

- An *Active* MP state prioritizes the multicast queue over unicast one. The scheduler serves, now, the multicast queue despite of the fact that there is a unicast queue, at the same output port, with larger NP value. In order to have the service balance, the deprived unicast queue gets a credit addition in its NP value which will help it to recover its lost share of bandwidth in the next cycles.
  At the output port 1 (fig. 1) the schedulers serves the multicast queue as directed by its *Active* state where as the NP value of unicast queue is larger than that of multicast queue at the output port.

- A *Sleep* MP state lets the unicast queues to be served first (if they have cells) and may delay the multicast queue serving. The later results in incrementing the NP value of delayed multicast queue which is used, in next scheduler cycles, to recover the lost fair share of bandwidth.
  At the output port 2 (fig. 1), the scheduler does not serve the multicast queue which is in *Sleep* state despite of the fact that it has larger NP value.

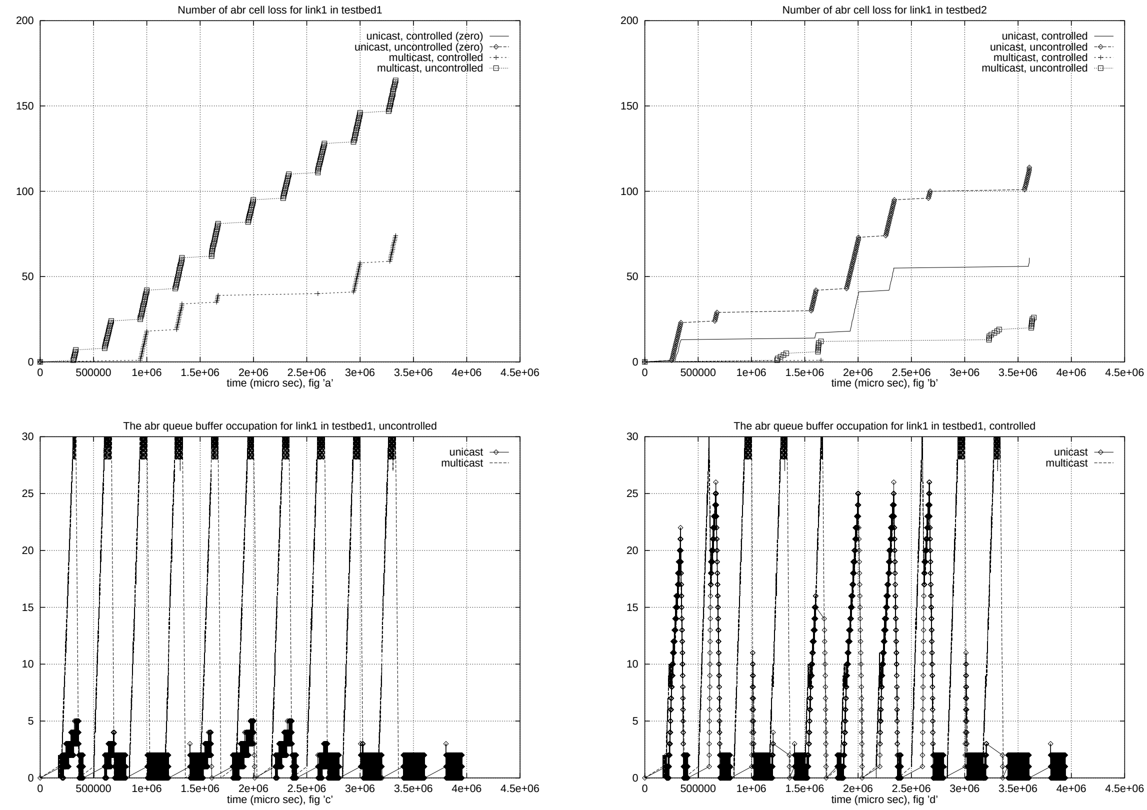

Figure 3: The simulation results

• With *Neutral* MP state, the scheduler serves the queue declared eligible by its NP value. If the queues happen to have the same NP values, then a cell from the largest queue (among those present at the output port) is selected.
The scheduler at output port 3 (fig. 1) decides to serve the unicast queue as it has larger NP value. The MP state of multicast queue is *Neutral*.
• If the unicast queue, having larger NP value, tends to be

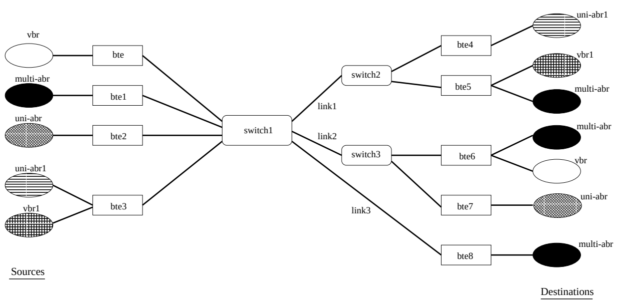

Figure 2: The testbed architecture

full then the scheduler will serve the unicast queue even if a multicast queue, at the same output port, has an *Active* MP state. It ensures that multicast queue's MP values do not hamper the unicast queues to get their fair share.
At the output port 4 (fig. 1), even the *Active* state of multicast queue could not make it eligible for the service. The unicast queue, having larger NP value, is served.

4 Simulation

The arbitration algorithm is being tested for ABR traffic class on several testbeds and we analyze, here, two of them having the same network architecture (figure 2). The NIST ATM simulator [5], after modifying the routing and queuing functions at a switch, was used for the simulation. For this paper, we have chosen to present only those simulation results taken with source applications having ON/OFF data emission flows because such flows are bursty and may analyze better the performance of the arbitration algorithm. We present the results observed at link1 whose bandwidth capacity is 100Mbps. The parameters of all the applications are selected such that an ABR application (unicast/multicast), when ON, has 50Mbps available for it on link1 as its fair share. Each source emits 1 Mbits of data. The two testbeds differ in the following aspect (refer to figure 2):
• In testbed1, the multicast ABR application (multi-abr) emits at 60Mbps and both unicast ABR applications (uni-abr and uni-abr1) emit at 50Mbps.
• In testbed2, the cell emission rate of both unicast ABR applications (uni-abr and uni-abr1) is 60Mbps and that of multicast application (multi-abr) is 50Mbps.

4.1 Results

Following are the simulation results:
• The simulations results show that the proposed method does not introduce any cell loss in unicast ABR while prioritizing, at certain instants, the multicast queues. In test-

bed1, there is no cell loss introduced by the method for unicast ABR connection (figure 3 'a'). Unicast cell loss, with and with out proposed method, is zero in both the cases.
• The method respects the fair share policy. This is proved by the testbed2 results where it has ensured that multicast ABR application does get its fair share and does not have cell loss because of "higher" demand of unicast ABR applications (figure 3 'b').
• Our proposed method reduces cell loss regardless of the nature (unicast/multicast) of ABR application(s). It can be witnessed by observing the decrease in cell loss in both testbeds (figure 3 'a' and 'b').
• The method does not interfere the cell scheduling of other traffic classes. As in both the tests, VBR (Variable Bit Rate) cell flow pattern did not get altered.
• From the figure 3 'c', it is evident that the unicast abr queue in testbed1 never fills up to its maximum capacity where as the multicast queue, at every burst arrival, overflows. Arbitration algorithm optimized the use of available buffer and thus reduced the multicast cell drop. This can be observed in figure 3 'd' which shows that the unicast cells were forced to stay in their queues (observe the rising unicast queues) if multicast queues were about to overflow. Similar behavior is also observed for testbed2.

5 Conclusion

For ABR traffic, cell loss determines the QoS. NP and MP values minimize the cell loss rate by the maximal usage of buffer occupancy. Our queue service scheme ensures a fair and homogeneous service to replicated queues and helps, the ATM switches, send upstream an amalgamated RM cell with an agreed/average rate value instead of minimum one [3, 4].
The proposed arbitration algorithm requires O(n) work for selecting a queue among n queues present at the output port. Note that amount of work is of the order of number of queues which is usually far less than the number of contending cells at the given instant. Implementation of a service discipline scheduling large number of queues is no more impracticable. Philips has recently developed an ATM switch [9] which can perform weighted round robin service among 2000 VCs (Virtual Channels), each served at 155Mbps. A complete congestion control scheme for ABR multicast connections is under development which will be enriched by our proposed queue service scheme.

References

[1] Hui Zhang and Srinivasan Keshav. Comparison of Rate-Based Service Disciplines. *Proceedings of ACM SIGCOMM'91*, Zurich, September 1991, pp. 157-173.

[2] M.Tufail and B.Cousin. ABR Congestion Control Schemes and the Multicast Problem. *Proc. of* 1^{st} *NDSU Workshop on ATM Networking*, Fargo, USA, August 1996, pp. 27-37.

[3] M.Tufail and B.Cousin. Timer Imposed and Priority Supported (TIPS) Congestion Control Scheme for Point-to-Multipoint Connections in ATM. *Proc. of* 11^{th} *Congrès De Nouvelles Architectures pour les Communications (DNAC) Large Bande: Internet ou RNIS*, France, 1996, pp. 65-79.

[4] ATM Forum Management Specifications version 4.0. *ATM Forum/95-0013R6* (June 7, 1995).

[5] Nada Golmie, Alfred Koenig, David Su. The NIST ATM Network Simulator, ver 1.0. *Computer Science Laboratory, Gaithersburg, MD 20899*, Aug 1995.

[6] Kai-Yeung Sui, Hong-Yi Tzeng. Congestion Control for Multicast Service in ATM network. *Dept. of Electrical & Computer Engineering, Univ. of California, Irvine. Tech. Rep. No.: 94-03-01*.

[7] K. K. Ramakrishnan and Peter Newman. Integration of Rate and Credit Schemes for ATM Flow Control. *IEEE Network*, March/April 1995.

[8] Flavio Bonomi, Kerry W. Fendick. The Rate-Based Flow Control Framework for The Available Bit Rate Service. *IEEE Network*, March/April 1995.

[9] Christian Cantou - Philips. FACE: an ATM switch component guaranteeing QoS. *Proc. of ATM97 Developments*, March 1997, Rennes, France, pp. 159.