

# Buffer Dumping Management for High Speed Routers

Caroline Fayet,  
INT-GET, 9 rue Charles Fourier,  
91011 Evry Cedex, France  
[caroline.fayet@int-evry.fr](mailto:caroline.fayet@int-evry.fr),

André-Luc Beylot  
ENSEEIH, IRIT/TeSA, 2, rue Camichel,  
31071 Toulouse Cedex 7, France  
[andre-luc.beylot@enseeiht.fr](mailto:andre-luc.beylot@enseeiht.fr),

## Abstract

We propose an efficient solution to optimize buffer dumping allowing to reduce significantly the losses caused by buffer overflow. To highlight the proposed architecture we have analyzed its performance compared with different configurations of polling scheme.

## 1. Introduction

A router making use of QoS functions needs to propose the following processes: classification, policing and marking, queuing and scheduling. The step named “queuing” is composed of congestion control and of an efficient dumping of the buffers of the routers. Concerning the short time only appropriated mechanisms related to the queues management of the used equipment will be efficient. We propose a new hardware design allowing to optimize the buffer dumping of the routers. Traditionally the buffers are emptied by means of a polling process, where all the queues are scanned and served sequentially even though for empty queues. The main drawback of processing empty queues result in a useless waste of time. The process polling used for many ATM switches correspond to this description. We find in the literature an improved solution for routers considering a reduced service time for empty queues. We propose a mechanism which offers two advantages, firstly a very efficient and simple solution by allowing the polling process to ignore empty buffers, secondly lowering the waiting time in non-empty queues. We present, in section 2, a performance analysis of this proposed solution, compared with the ATM-like polling and the classical polling scheme.

The speed requirements put on broadband switches or routers impose a wired solution. A wired sequencer using digital circuits in a very high-speed technology available on the market today provides this hardware solution.

We consider the problem of multiplexing  $N$  input links to a single output. The situation occurs inside switching fabrics, high-speed routers or when matrices are interconnected in multistage switches. Most often (this is

the case when multiplexing for concentration) the sum of the speeds on each input link exceeds the capacity of the output link, so that buffering is compulsory. The proposed switching mechanism (Figure 1) is composed of a bus and a register RI. This Data Bus will receive sequentially an incoming information from the input registers. An incoming data located at the head of the elected queue is transferred to the bus, which carries it to the register RI where it is stored, waiting to be processed - i.e., redirected to the allocated output port. The filtering of the incoming data is processed by means of very high-speed 3-state outputs.

Each input owns its proper wired state machine whose role is firstly to control the presence or not of the information in the associate register, then eventually its priority level and also the state of the other inputs. A decision is then taken enabling one input register among  $N$  to be loaded on the bus. This shows a constant interaction between the different computer blocks.

As opposed to classical polling schemes, here empty buffers are ignored. Classically, an empty buffer sends an empty information or cell which is overwritten later on. Therefore, non-empty buffers are visited more often, increasing significantly the performance of such a system.

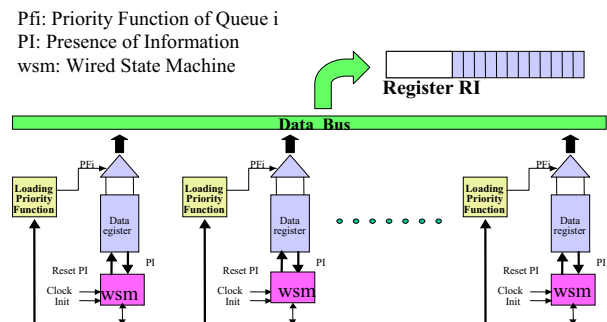


Figure 1. Block diagram of buffer dumping

## 2. Performance Analysis and Results

In the present paper, we will compare different configurations. A model with  $N$  infinite queues is considered. Packets arrive according to Poisson processes. The polling server has no initialization time. The service times of packets at queue  $i$  are constant. Unbalanced traffic patterns are considered: one input port is supposed to have a higher input rate,  $a$  times higher than the other input ports. The three following solutions are considered:

- **ATM-like polling**

In this solution, the input queues are polled and “served” even if there are empty. This solution has been extensively studied in the ATM context. Each queue can be analyzed through an M/G/1 model. For non empty queues, the service time is equal to the duration of a cycle. If the packet arrives in an empty queue, its waiting time is uniformly distributed between 0 and a cycle time [1].

- **Classical polling**

In this solution, all the input queues are polled even if there are empty; a constant switch-over time is considered. This solution is analyzed using the methods proposed in [2] (details on the implementation are presented in [1]).

- **Optimized polling**

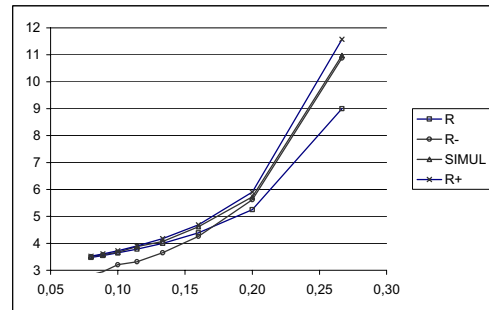
This solution corresponds to our optimized dumping management. Only non-empty queues are polled. The global mean waiting time is the one of an M/D/1 queue, the difference comes from the fact that in the proposed solution, the polling system may lead to a non-FIFO scheduling algorithm even if the scheduling is FIFO for each input queue. It leads to a lower bound for the heavily loaded input queue and an upper bound for the other queues (the mean length of the first queue is the largest).

Another lower bound can be derived by considering that the first queue is never empty and by analyzing the other queues like a symmetric 1-limited polling system with a constant switch-over time and using the method presented for the classical polling [2].

Finally, an upper bound of the response time of the first queue can be derived as follows. Let us consider a packet  $p$  which arrives in the first queue. As this queue is heavily loaded, the number of packets in the other queues may be lower than the number of packets in the first queue. Consequently, most of the packets that are in the system when packet  $p$  arrives will leave the system before  $p$ . In the worst case, all the packets that enter all the other queues when  $p$  is waiting will also be served before  $p$ .

The validation of the models has been done by comparisons with discrete event simulations. The number of input ports  $N$  is set to 8 and  $a$  is set to 5. The mean

response time are expressed in number of clock time periods. For the optimized polling mechanism we obtained the following results. R curves correspond to the mean response time in the symmetric traffic case, R- (resp. R+) to the lower bound (resp. upper bound). SIMUL results were obtained using discrete event simulations. The results are depicted as a function of the input load.



**Figure 2. Mean Response Time first input port**

It is shown that the model leads to a good estimation of the mean response time. The difference between simulation results and the lower bound is about 1% to 2%. The upper bound is also really good, the difference is about 5%.

When parameter  $a$  is equal to 5, the ATM-like solution is not stable for the first input queue. For instance, with an input load equal to 0.1, the mean response time of the first queue is equal to 5 clock time periods with the optimized polling, to 20 for the classical polling; For the other ports, we obtained a mean response time equal to 4 with the optimized polling, 9 with the classical polling and 17 for the ATM-like polling. Other results are presented in [1].

### 3. Conclusion

A new buffer dumping management mechanism for high speed routers has been proposed. It consists on allowing the polling process to ignore empty buffers. This mechanism improves the global performance of the system. Prospective works deal with the comparison with other polling techniques and the evaluation of the packet loss probability.

### 4. References

- [1] C. Fayet, A.-L. Beylot, Buffer Dumping Management for High Speed Routers, IRIT Technical Report, July 2003, <http://www.enseciht.fr/~beylot/IRITBeylot4.pdf>
- [2] S.W. Furman, Y.T. Wang, Mean Waiting Time Approximations of Cyclic Service Systems with Limited Service, Performance'87, pp. 253-365, Elsevier Science.