A Throughput Analysis of Reliable Multicast Protocols in an Active Networking Environment

M. Maimour and C. D. Pham
Technical Report, TR01-2000
RESAM, Université Lyon 1, Claude Bernard
Bât. ISTIL, 43 bd du 11 novembre 1918
69622 Villeurbanne Cedex - France
email:{mmaimour,cpham}@resam.univ-lyon1.fr

Abstract

Reliable multicast protocols have recently been put in front of the scene with active services contribution where routers implement additional functionalities. Contributing mainly on feedback implosion problems, retransmission scoping and cache of data, these active protocols open new perspectives for achieving high throughput and low latency on wide-area networks. In this paper, we propose generic models of multicast protocols in an active networking environment and a throughput analysis. Comparisons between several mechanisms proposed in existing protocols, i.e. SRM, ARM, AER, are performed. Some insight on dimensioning active networks are also provided.

1 Introduction

Multicast is the process of sending every single packet to multiple destinations. Motivations behind multicast facilities is to handle one-to-many communications in a wide-area network with the lowest network and end-system overheads. In contrast to best-effort multicast, that typically tolerates some data losses and is more suited for real-time audio or video for instance, reliable multicast requires that all packets are safely delivered to the destinations.

Desirable features of reliable multicast include low end-to-end delays, high throughput and scalability. Meeting these objectives is not an easy task and reliable multicast is known to be a very difficult problem. In the past, there have been a number of propositions for reliable multicast protocols that rely on complex exchanges of feedback messages (ACK or NACK) [1, 2, 3, 4]. Most of these schemes were end-to-end oriented and suffered from bottlenecks both at the sender and intermediate routers stage because of the large resulting traffic of feedback

messages. Worth to mention, the approach taken in [5] where the authors propose a primary or additionally a set of secondary logging servers to perform retransmissions instead of the multicast source. Recently, the use of active network concepts [6] where routers themselves could contribute to enhance the network services by customized functionalities have been proposed in the multicast research community [7, 8]. Contributing mainly on feedback implosion problems, retransmission scoping and cache of data, these active reliable multicast protocols open new perspectives for achieving high throughput and low latency on wide-area networks.

There are multiple ways on how to implement the active services into routers. Whatever methods will be used, there will certainly be a cost incurred by the additional processing functions at the routers. Taking some costs into account, previous studies have successfully shown that active services can increase the performances of reliable multicast protocols [7, 8]. However, these studies are limited to their own proposed protocol compared to the non active solution. With the various mechanisms available for increasing the performance of reliable multicast protocols (feedback aggregation, caches, subcast...) and the number of parameters which have a direct impact on the performances of active routers, it is difficult to know what combination offers the best results. Therefore, there is definitely a need for a generic framework that would allow a deeper comparison of active reliable multicast protocols. Dimensioning active networks is also an essential aspect of this study that we would like to tackle.

We propose generic models for multicast protocols in an active networking environment and, as a first step, a throughput analysis of these protocols. The motivations behind the design of these models are to make possible the comparison between various mechanisms in active reliable multicast protocols and to better understand the contribution of active networking services. One additional desirable feature of the models is to be able to model existing protocols by instantiating some of the parameters involved in the models, thus justifying the "generic" property.

Our main contributions are in comparing various active strategies and identifying those that exhibit the best performances. In an attempt to help for active networks dimensioning, we also address the constraints put on active routers processing power for achieving better performances than traditional non active approaches. The rest of the article is organized as follows. Section II presents a brief overview of existing reliable multicast protocols with a focus on those proposed recently with active service contributions. The generic reliable multicast protocols are then presented in section III. Section IV presents the network models and the general assumptions associated to the study. The analytical generic models for the throughput

analysis are presented in section V and section VI presents the results. Section VII presents the previous related works on multicast protocol analysis and section VIII concludes.

2 Background: reliable multicast and active networks

2.1 Reliable multicast protocols

Earlier reliable multicast protocols took the end-to-end solution to perform loss recovery. Most of them fall into one of following classes: sender-initiated, receiver-initiated and receiver-initiated with local recovery protocols. In sender-initiated protocols, the sender is responsible for both the loss detection and the recovery [1]. These protocols do not scale well to a large number of receivers due to the ACK implosion problem. Receiver-initiated protocols move the loss detection responsibility to the receivers. They use NACKs instead of ACKs. However they still suffer from the NACK implosion problem when a large number of receivers have subscribed to the multicast session. In receiver-initiated protocols with local recovery, the retransmission of a lost packet can be performed by any receiver [2] in the neighborhood or by a designated receiver in a hierarchical structure [3]. All of the above schemes do not provide exact solutions to all the loss recovery problems. This is mainly due to the lack of topology information at the end hosts [9]. Therefore new schemes benefit from routers contribution in the loss recovery process.

2.2 Active network contribution

In active networking, routers themselves play an active role by executing application-dependent functions on incoming packets. For reliable multicast protocols these functions are mainly:

- the cache of data packets to permit local recovery of loss packets and reduce the recovery latency.
- the aggregation or the local suppression of NACKs to solve the NACK implosion problem.
- the subcast of repair packets to a set of receivers to limit both the retransmission scope and the bandwidth usage.

ARM (Active Reliable Multicast) [7] and AER (Active Error Recovery) [8] are two protocols that were recently proposed in the research community and that use active services. In both protocols, the main contribution of active services is a *best-effort* cache of data packets to permit

local recoveries. However, they differ in the strategy adopted for solving the NACK implosion problem. In ARM, a receiver experiencing a packet loss sends immediately a NACK to the source. Active services in routers then consist in the aggregation of the multiple NACKs. In contrast, AER uses a strategy similar to the one used by SRM and based on local timers at the receivers: prior to send a NACK packet, a receiver initiates a timer and waits for a random amount of time. If it receives a NACK for the same packet, it would cancel the timer, supress its NACK and behave exactly as if it has sent the NACK itself. In this case, the active routers have in charge to indicate to the receivers that one of them has issued a NACK for a given packet. Implementation of this service can be done by multicasting to the other receivers the NACK packet.

In addition, an active router in ARM would send the repair packet only to the set of receivers that have sent a NACK packet (subcast). In AER, the active router simply multicasts the repair packet to all its associated receivers.

3 Generic reliable multicast protocols

In this section, we describe 3 generic reliable multicast protocols from which the generic models are derived. These protocols are called $S1$, $S2$ and $S3$. All three protocols benefit from the cache of packets at the active routers. However, they differ from each other in the strategy for the NACK suppression and whether subcast is used or not.

A NACK packet is assumed to follow the reverse path of the data packet. This assumption is essential for taking advantage of active services. This behavior can be implemented by using a dedicated routing service for NACK packets or to make a receiver know the identity of the active router to which it is attached [8].

Also, in order to limit the overheads for processing duplicate NACKs, the source and the active routers do not consider any subsequent similar NACKs for a given amount of time. For the analysis, we will assume that the "duplicate discard" period is well chosen. Therefore, no duplicate NACKs will trigger a new retransmission.

Finally, we will call *local group* the set of receivers linked to the same active router. If there are several active routers on the path, the nearest one, from the receiver perspective, is taken.

3.1 Description of protocol S1

S1 uses the global suppression of NACK packets. A receiver experiencing a packet loss sends immediately a NACK to the source. The active routers have in charge the aggregation of NACK packets in order to forward only one NACK to the source. S1 has the following properties:

- the source multicasts data packets at the multicast address subscribed to by all the receivers.
- upon reception of a data packet, an active router stores the packet in its cache, if possible, and forward it downstream in the multicast tree.
- upon detection of a packet loss, a receiver sends immediately a NACK towards the source and sets a timer.
- upon reception of a NACK packet, an active router sends the corresponding repair packet, if available, to all the receivers composing its local group. Otherwise, it sends the NACK to the source.
- upon reception of a NACK packet, the source multicasts the repair packet to all receivers at the multicast address.

3.2 Description of protocol S2

S2 uses the NACK local suppression strategy. The receivers wait for a random amount of time prior to send a NACK to the source. The goal is to generate only one NACK per group of receivers in the multicast tree. S2 is similar to L1 described in [10] (which is itself a generic version of AER). Both uses NACK local suppression but S2 performs this suppression only within the end receivers linked to a same active router. In contrast, L1 would also do local suppression within intermediate nodes located at the same level in the multicast tree. S2 has the following properties:

- the source multicasts data packets at the multicast address subscribed to by all the receivers.
- upon reception of a data packet, an active router stores the packet in its cache, if possible, and forward it downstream in the multicast tree.

- upon detection of a packet loss, a receiver waits for a random amount of time. Only then it would send a NACK packet to the source and sets a timer.
- when a receiver is waiting to send a NACK, the reception of a similar NACK, from its active router, would make the receiver to cancel its NACK, set a timer and behave as if it has sent the NACK itself. If during the waiting time it receives the repair packet, then the receiver quits the recovery process.
- upon reception of a NACK from downstream, an active router multicasts the repair packet, if available, to its local group. Otherwise, it multicasts the NACK to both the source and its local group, excepting on the NACK incoming link. The desire behavior is to perform the NACK local suppression.
- upon reception of a NACK packet, an active router sends the corresponding repair packet, if available, to all the receivers composing its local group.
- upon reception of a NACK packet, the source multicasts the repair packet to all receivers at the multicast address.

3.3 Description of protocol S3

S3 is identical to S1 but performs subcast, meaning that only the set of receivers that really experienced a packet loss receives the repair packet.

4 Network models and hypothesis

The network model used in the study is similar to the one used in [10]: one source multicasts data packets to R receivers through a packet network composed of a fast core network and several slower edge access networks. We will call *source link* the set of point-to-point links and traditional routers that connects the source to the core network. Similarly, a *tail link* is composed of point-to-point links and routers connecting a receiver to the core network (see Fig. 1). We only consider active routers at the edge of the core network. This for two reasons:

- The core network is reliable. It has been shown in [11] that the losses occurred mainly at the edges of the core network: at the source link and at the tail links.
- The core network is a very high-speed network. Adding complex processing functions inside the core network will slow down the packet forwarding functions.

4.1 Base configuration

In the base configuration, active routers are associated to the tail links. However, it is possible that not all routers implement active services. We will consider that F routers among the N possible are active routers, $0 \leq F \leq N$ (Fig. 1). Each active router A_i is responsible of B receivers $R_{i1}, \cdots, R_{iB}$ forming a local group. A receiver associated with an active router is said *linked*. The other receivers are said *free*. Particular cases where $F = 0$ and $F = N$ will be considered in the study.

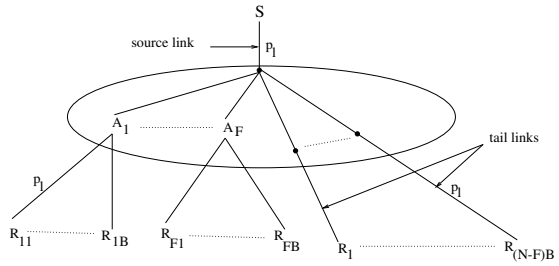


Figure 1: Base configuration.

4.2 Enhanced configuration

In addition to the base configuration, we propose an enhanced configuration where the source has an associated active router, noted A_S. The previously mentioned protocols will be noted $S1^+$, $S2^+$ and $S3^+$ when applied on the enhanced configuration.

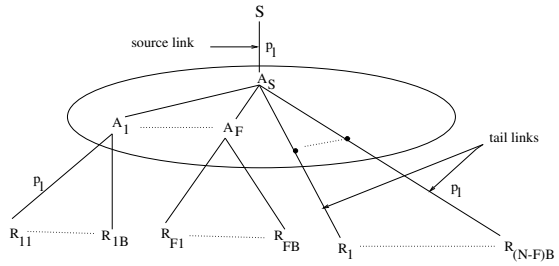


Figure 2: Enhanced configuration.

4.3 Hypothesis

The total processing time per packet at the source includes the initial transmission time, all the retransmissions, and NACK overheads necessary for achieving a correct reception of the packet by all the receivers. For the receiver, the total processing time per packet includes all the processing overheads until a randomly chosen packet is correctly received. This includes

the overheads for receiving the initial packet, all the necessary repair packets, the overheads for sending and receiving the NACK packets and the overheads for timeouts.

Regarding an active router, the total processing time per packet includes the overheads for sending/receiving and processing the packet and the corresponding NACK packets in case of losses. Kasera et al. also took in consideration the overheads of timeouts within active routers. Our generic protocols do not use local suppression between nodes located at the same level in the multicast tree. The reason is not to overload the active routers.

For the loss model, we will consider that the core network is reliable, as mentioned previously. For the other links (the source link or the tail links), the loss probability is noted p_l (this probability includes the losses due to congestion). Therefore, the end-to-end probability of a packet loss perceived by a receiver is $p = 1 - (1 - p_l)^2$. The losses at the source link are assumed to be temporally independent and those at the tail links are assumed to be mutually independent.

For this analysis, we will also assume that NACK packets are never lost. In addition, we also assume that sufficient memory resources are available for packet cache. Therefore, when a NACK arrives at an active router, this router will always have the corresponding packet in its cache.

5 Generic analytical models for the throughput analysis

Like the approach taken by Pingali et al. in [12], our analysis uses the processing requirements of the protocols rather than the bandwidth requirements. The maximum bandwidth achievable by a protocol depends on the processing time per packet at the various nodes in the multicast tree.

5.1 Notation

For the analysis, a similar notation to the one introduced in [12] is used. A number of additional parameters used in the study is listed in table 1. We will also use the notation $(x)^+$ to refer to $max(0, x)$. Table 2 summarizes the probability distribution and the mean of each random variable used. For space limitation, we will not be able to include the demonstrations in this paper. They can be found in [**?**].

$X^{S_j}, X^{S_j^+}$	respectively, the total processing time per packet at the source for S_j and S_j^+ protocols, $j = 1..3$.
$Y^{S_j}, Y^{S_j^+}$	respectively, the total processing time per packet at a linked receiver for S_j and S_j^+ protocols, $j = 1..3$.
$Y_*^{S_j}, Y_*^{S_j^+}$	respectively, the total processing time per packet at a free receiver for S_j and S_j^+ protocols, $j = 1..3$.
$A_i^{S_j}, A_i^{S_j^+}$	respectively, the total processing time per packet at an active router A_i ($i = 1..F$) associated to a local group of linked receivers for S_j and S_j^+ protocols, $j = 1..3$.
$A_s^{S_j^+}$	the total processing time per packet at the A_S active router for S_j^+ protocols, $j = 1..3$.
Λ_x^ω	mean throughput achieved by node x with the ω protocol.
Λ^ω	mean throughput achieved by the ω protocol.
X_p, X_n	respectively, the processing time for sending a data packet and to receive a NACK packet.
Y_p, Y_n	respectively, the processing time for receiving a data packet and to send a NACK packet.
Y_t	time to process a timeout
$(X_p^a), (X_n)^a$	respectively, the processing time for sending a data packet and to receive a NACK packet for an active router.
$(Y_p^a), (Y_n)^a$	respectively, the processing time for receiving a data packet and to send a NACK packet for an active router.
$M_S,\ M_{A_S}$	number of retransmission of a packet at the source, respectively at an active router, until all active routers associated to linked receivers and all free receivers have correctly received the packet.
$M_R,\ M_A$	same as M_S when $F = 0$, respectively $F = N$.
$M_{R'},\ M_{A'}$	same as M_{A_S} when $F = 0$, respectively $F = N$.
$M_b,\ M_B$	number of retransmission of a packet at an active router until one, respectively all, associated linked receivers have correctly received the packet.
M_a	number of retransmission of a packet at the source until an active router has correctly received the packet.
M_r	number of retransmission of a packet at the source until a receiver (linked or free) has correctly received the packet.

Table 1: Notation

Random var. X	Prob. distribution $P[X \leq m]$	Mean $E[X]$
$M_b = M_a$	$1 - p_l^m$	$1/(1-p_l)$ giving $E[(M_a-2)^+] = E[(M_b-2)^+] = p_l^2/(1-p_l)$
M_r	$1 - p^m$	$1/(1-p)$ giving $E[(M_r-2)^+] = p^2/(1-p)$
M_B	$(1-p_l^m)^B$	$1 + \sum_{m=1}^{\infty} 1-(1-p_l^m)^B$
M_S	$(1-p_l^m)^F\ (1-p^m)^{(N-F)B}$	$1 + \sum_{m=1}^{\infty} 1 - P[M_S \leq m]$
M_R	$(1-p^m)^R$	$1 + \sum_{m=1}^{\infty} 1-(1-p^m)^R$
M_A	$(1-p_l^m)^N$	$1 + \sum_{m=1}^{\infty} 1-(1-p_l^m)^N$
M_{A_S}	$(1-p_l)^{(N-F)B}$	$1 + \sum_{m=1}^{\infty} 1-(1-p_l)^{(N-F)B}$
$M_{R'}$	$(1-p_l^m)^R$	$1 + \sum_{m=1}^{\infty} 1-(1-p_l^m)^R$
$M_{A'}$	1	1

Table 2: Useful probability distributions.

5.2 Protocol S1

5.2.1 Base configuration

We begin by analyzing the processing requirements at the source: a data packet is sent M_S time until all active routers and free receivers have correctly received it. The source receives for this packet $F\ (M_a - 1)$ NACKs from active routers and $(N-F)B\ (M_r - 1)$ NACKs from free receivers. The processing time per packet at the source can therefore be written:

$$E[X^{S_1}] = E[M_S]\ E[X_p] + (F(E[M_a]-1) + (N-F)B\ (E[M_r]-1))\ E[X_n] \quad (1)$$

Replacing $E[M_a]$ and $E[M_r]$ by their expressions of table 2 gives:

$$E[X^{S_1}] = E[M_S]\ E[X_p] + \left(F\ \frac{p_l}{1-p_l} + (N-F)B\ \frac{p}{1-p}\right) E[X_n] \quad (2)$$

Note that the S1 protocol in the base configuration with no active routers ($F = 0$) is identical to the N_1 protocol described in [12]. By setting $F = 0$ in eq. (2) we have a similar result than eq. (11) stated in [12]:

$$E[X^{S_1}] = E[M_R]\ E[X_p] + R\ \frac{p}{1-p}\ E[X_n] \quad (3)$$

When all routers have active services, we find:

$$E[X^{S_1}] = E[M_A]\ E[X_p] + N\ \frac{p_l}{1-p_l}\ E[X_n] \quad (4)$$

The processing time at a receiver depends on whether it is linked or free. A linked receiver receives a data packet M_B times with probability $(1 - p_l)$. It sends back $(M_r - 1)$ NACKs towards the source, and so processes $(M_r - 2)^+$ timeouts. This gives the following expression:

$$E[Y^{S_1}] = (1 - p_l)\ E[M_B]\ E[Y_p]\ +\ \frac{p}{1-p}\ E[Y_n] +\ \frac{p^2}{1-p}\ E[Y_t] \tag{5}$$

In contrast, a free receiver receives the data packet M_S times with probability $(1 - p)$. It sends $(M_r - 1)$ NACKs and processes $(M_r - 2)^+$ timeouts, giving the following equation:

$$E[Y_*^{S_1}] = (1 - p)\ E[M_S]\ E[Y_p]\ +\ \frac{p}{1-p}\ E[Y_n] +\ \frac{p^2}{1-p}\ E[Y_t] \tag{6}$$

If $F = 0$, meaning that all receivers are free, we have an expression similar to that of protocol N_1 found in [12]:

$$E[Y_*^{S_1}] = (1 - p)\ E[M_R]\ E[Y_p]\ +\ \frac{p}{1-p}\ E[Y_n] +\ \frac{p^2}{1-p}\ E[Y_t] \tag{7}$$

For an active router A_i, it receives packets like a free receiver: a data packet is received M_S times with probability $(1 - p_l)$. It forwards the packet M_B times to its B linked receivers until they correctly receive the packet. It receives $B\ (E[M_r] - 1)$ NACKs from its B linked receivers and sends back $(M_b - 1)$ NACKs to the source. These NACKs have been issued by the linked receivers due to a packet loss at the source link so we have:

$$\begin{aligned} E[A_i^{S_1}] = \quad & (1 - p_l)\ E[M_S]\ E[Y_p^a]\ + E[M_B]\ E[X_p^a] \\ & + B\ \tfrac{p}{1-p}\ \ E[X_n^a] +\ \tfrac{p_l}{1-p_l}\ E[Y_n^a] \end{aligned} \tag{8}$$

If all routers are active, we find:

$$\begin{aligned} E[A_i^{S_1}] = \quad & (1 - p_l)\ E[M_A]\ E[Y_p^a]\ + E[M_B]\ E[X_p^a] \\ & + B\ \tfrac{p}{1-p}\ E[X_n^a] +\ \tfrac{p_l}{1-p_l}\ E[Y_n^a] \end{aligned} \tag{9}$$

5.2.2 Enhanced configuration

In the enhanced configuration, the source is "protected" by its active router A_S: the source sends a data packet until the active router receives it correctly, that makes M_a times. A_S then sends back $(M_a - 1)$ NACKs for the same packet, resulting in:

$$E[X^{S_1^+}] = \frac{1}{1-p_l}\ E[X_p] +\ \frac{p_l}{1-p_l}\ E[X_n] \tag{10}$$

The enhanced configuration brings no changes for a linked receiver. In contrast, a free receiver receives in this case a data packet M_{A_S} times with probability $(1-p_l)$. It sends back (M_r-1) NACKs and processes $(M_r-2)^+$ timeouts such that:

$$E[Y_*^{S_1^+}] = (1-p_l)\ E[M_{A_S}]\ E[Y_p]\ +\ \frac{p}{1-p}\ E[Y_n] +\ \frac{p^2}{1-p}\ E[Y_t] \tag{11}$$

If all receivers are free, case where $F=0$, we find:

$$E[Y_*^{S_1^+}] = (1-p_l)\ E[M'_R]\ E[Y_p]\ +\ \frac{p}{1-p}\ E[Y_n] +\ \frac{p^2}{1-p}\ E[Y_t] \tag{12}$$

The active router at the source receives only once a data packet from the source. It sends the packet M_{A_S} times to the other active routers and to the free receivers. In addition, A_S receives $(F+(N-F)B.E[M_b])$ NACKs from its sons. As A_S performs NACKs aggregation before sending them to the source it sends only M_b NACKs, so we have:

$$\begin{aligned} E[A_S^{S_1^+}] = \quad & E[Y_p^a]\ + E[M_{A_S}]\ E[X_p^a] + \\ & (F+(N-F)B)\ \tfrac{1}{1-p_l}\ E[X_n^a] +\ \tfrac{1}{1-p_l}\ E[Y_n^a] \end{aligned} \tag{13}$$

If all routers are active, $F=N$, we find:

$$E[A_S^{S_1^+}] = E[Y_p^a]\ +\ E[X_p^a] +\ \frac{N}{1-p_l}\ E[X_n^a] +\ \frac{1}{1-p_l}\ E[Y_n^a] \tag{14}$$

But, if only A_S is active, we have:

$$E[A_S^{S_1^+}] = E[Y_p^a]\ + E[M'_R]\ E[X_p^a] +\ \frac{R}{1-p_l}\ E[X_n^a] +\ \frac{1}{1-p_l}\ E[Y_n^a] \tag{15}$$

An active router A_i receives a data packet M_{A_S} times from A_S. It sends the packet M_B times to the receivers in its own local group. It receives B $(E[M_r-1])$ NACKs per packet but only sends $(E[M_b]-1)$ NACKs to the source. These NACKs come from losses that have occurred at the source link. So we have:

$$\begin{aligned} E[A_i^{S_1^+}] = \quad & E[M_{A_S}]\ E[Y_p^a]\ + E[M_B]\ E[X_p^a] + \\ & B\ \tfrac{p}{1-p}\ E[X_n^a] +\ \tfrac{p_l}{1-p_l}\ E[Y_n^a] \end{aligned} \tag{16}$$

If all routers are active, we have $E[M_{A_S}] = E[M_{A'}] = 1$ which gives:

$$E[A_i^{S_1^+}] = E[Y_p^a] + E[M_B]\ E[X_p^a] + B\frac{p}{1-p}\ E[X_n^a] + \frac{p_l}{1-p_l}\ E[Y_n^a] \quad (17)$$

5.3 Protocol S2

For the analysis of the S2 protocol, we will assume that the random waiting time of receivers are well chosen in order to generate only one NACK per local group.

5.3.1 Base configuration

In this configuration, both the source and the free receivers behave like in the S1 protocol. Consequently, we have the same equation for them: eq. (1) and eq. (6) respectively. For a linked receiver, we have:

$$E[Y^{S_2}] = (1-p_l)\ E[M_B]\ E[Y_p] + \frac{E[M_B]-1}{B}\ (E[Y_n] + E[X_n]) + \frac{p^2}{1-p}\ E[Y_t] \quad (18)$$

The first term indicates how many times a linked receiver receives a data packet. The second and the third terms correspond respectively to NACKs that were sent and received by a linked receiver in the process of NACK local suppression. The last term represents the processing overhead for the $(M_r - 2)^+$ timeouts.

An active router receives a data packet from the source M_S times with probability $(1 - p_l)$. It sends the packets to its linked receivers M_B times. Thanks to the NACKs local suppression task performed by the receivers, the active router receives only one NACK per packet loss. In total, it receives $(M_B - 1)$ NACKs. The active router would forward the NACKs to the source and to the other receivers if it does not have the packet: that makes $p_l\ (M_B - 1)$ NACKs where p_l is the loss probability at the source link. Finally we find:

$$E[A_i^{S_2}] = (1-p_l)\ E[M_S]\ E[Y_p^a] + E[M_B]\ E[X_p^a] + (E[M_B]-1)\ E[X_n^a] + p_l\ (E[M_B]-1)\ E[Y_n^a] \quad (19)$$

If all routers are active, we have:

$$E[A_i^{S_2}] = \begin{array}{l} (1-p_l)\ E[M_A]\ E[Y_p^a]\ + E[M_B]\ E[X_p^a] + \\ (E[M_B]-1)\ E[X_n^a] + \ p_l\ (E[M_B]-1)\ E[Y_n^a] \end{array} \quad (20)$$

5.3.2 Enhanced configuration

In the enhanced configuration, results for both the source and the linked receivers remain unchanged. For a free receiver, we have:

$$E[Y_*^{S_2}] = \begin{array}{l} (1-p_l)\ E[M_{A_S}]\ E[Y_p]\ + \left(\frac{E[M_{A_S}]-1}{F+(N-F)B}\right)\ E[Y_n] + \\ \left(\frac{E[M_{A_S}]-1}{F+(N-F)B}\right)\ E[Y_n] + \ \frac{p^2}{1-p}\ E[Y_t] \end{array} \quad (21)$$

The first term represents how many times a free receiver receives a data packet. The second and the third terms correspond respectively to NACKs that were sent and received by a free receiver in the process of NACK local suppression. The last term represents the processing overhead for the timeouts.

If $F = 0$ where no routers except A_S are active, we have:

$$E[Y_*^{S_2^+}] = \begin{array}{l} (1-p_l)\ E[M_{R'}]\ E[Y_p]\ + \left(\frac{E[M_{R'}]-1}{R}\right)\ E[Y_n] + \\ \left(\frac{E[M_{R'}]-1}{R}\right)\ E[Y_n] + \ \frac{p^2}{1-p}\ E[Y_t] \end{array} \quad (22)$$

If we look at the processing requirements at the active routers we have A_S that receives only once a data packet and sends it M_{A_S} times to the other active routers and free receivers. A_S receives one NACK per packet loss so $(M_{A_S} - 1)$ NACKs in total. It sends these NACKs to the source and to its sons with probability p_l (probability of a packet loss at the source link). We can then write:

$$E[A_S^{S_2^+}] = \begin{array}{l} E[Y_p^a]\ + E[M_{A_S}]\ E[X_p^a] + \\ (E[M_{A_S}-1)\ E[X_n^a] + \ p_l\ (E[M_{A_S}-1)\ E[Y_n^a] \end{array} \quad (23)$$

When $F = N$, A_S only performs data packet receptions and the forwarding of these packets to the other active routers. In this case, we find:

$$E[A_S^{S_2^+}] = E[Y_p^a] \ + E[X_p^a]. \tag{24}$$

If A_S is the only active router, it is more loaded as shown by the following expression:

$$\begin{aligned} E[A_S^{S_2^+}] = \quad & E[Y_p^a] \ + E[M_{R'}] \ E[X_p^a] + \\ & (E[M_{R'}] - 1) \ E[X_n^a] + \ p_l \ (E[M_{R'} - 1) \ E[Y_n^a] \end{aligned} \tag{25}$$

An active router associated to linked receivers receives a data packet M_{A_S} times. It sends the packet M_B times to receivers in its local group. These receivers will send back $(M_B - 1)$ NACKs. An active router would then send these NACKs to the source and to its sons with probability p_l (probability of a packet loss at the source link). Regarding NACKs local suppression, an active router receives $(E[M_{A_S}] - 1)$ / $(F + (N - F)B)$ NACKs from A_S. Finally we can write:

$$\begin{aligned} E[A_i^{S_2^+}] = \quad & E[M_{A_S}] \ E[Y_p^a] \ + E[M_B] \ E[X_p^a] \\ & + \ \left(\tfrac{1}{F+(N-F)B} \ (E[M_{A_S}] - 1) + \ (E[M_B] - 1) \right) \ E[X_n^a] \\ & + p_l (E[M_B] - 1) \ E[Y_n^a] \end{aligned} \tag{26}$$

When all routers are active routers, we find:

$$\begin{aligned} E[A_i^{S_2^+}] = \quad & E[Y_p^a] \ + E[M_B] \ E[X_p^a] + \\ & (E[M_B] - 1) \ E[X_n^a] + \ p_l (E[M_B] - 1) \ E[Y_n^a] \end{aligned} \tag{27}$$

5.4 Protocol S3

S3 is an enhanced version of S1 where an active router sends a repair packet only to the set of receivers that have experienced the packet loss (subcast).

5.4.1 Base configuration

In this configuration, the linked receivers benefit from the subcast feature since they receive a data packet only once. Therefore we have:

$$E[Y^{S_3}] = E[Y_p] \ + \ \frac{p}{1-p} \ E[Y_n] + \ \frac{p^2}{1-p} \ E[Y_t] \tag{28}$$

For the other nodes, we have the same results than for the S1 protocol.

5.4.2 Enhanced configuration

Having an active router at the source can greatly reduce the load of both the free receivers and the other active routers since they do not receive multiple copies of a data packet. The analysis is then identical to the base configuration, except for a free receiver and an active router. For them, we have respectively:

$$E[Y_*^{S_3^+}] = \frac{1}{1-p_l}\ E[Y_p]\ +\ \frac{p}{1-p}\ E[Y_n] +\ \frac{p^2}{1-p}\ E[Y_t] \tag{29}$$

$$E[A_i^{S_3^+}] = B\ \frac{p}{1-p}\ E[X_n^a]\ + E[M_B]\ E[X_p^a] +\ \frac{p_l}{1-p_l}\ E[X_n^a]. \tag{30}$$

6 Numerical results

In this section, we analyze the three generic protocols S1, S2 and S3. For each protocol, we first focus on the throughput achieved separately by each node and then we examine the overall throughput achieved by a protocol. In addition, we will first consider that active routers have the same processing time than receivers, then we will study the impact of varying the processing time of active routers.

6.1 Active routers with same processing time than receivers

The throughput Λ_x^ω achieved by node x in the protocol $\omega \in \{S1, S2, S3, S1^+, S2^+, S3^+\}$ is given by $\Lambda_x^\omega = 1/E[x^\omega]$ where $x \in \{X, Y, Y_*, A_1, \cdots, A_F, A_S\}$ and $E[x^\omega]$ is the mean processing time at node x in the protocol ω. Therefore, the overall throughput Λ^ω achieved by the protocol ω can be computed using $\Lambda^\omega = Min(\Lambda_x^\omega)$.

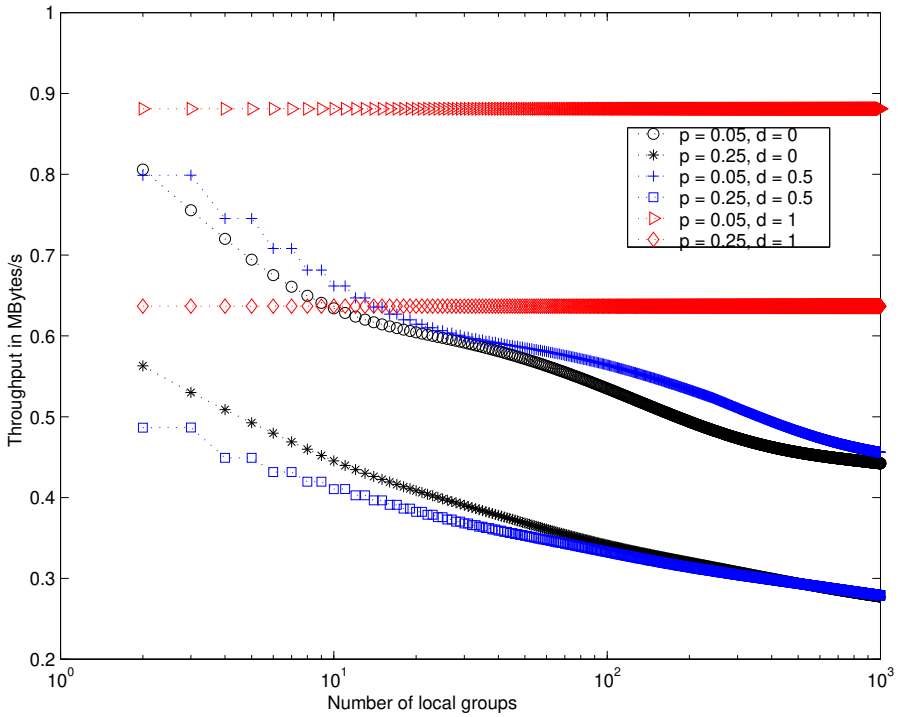
For the numerical evaluation, we take $E[X_p^a] = E[Y_p^a] = E[X_p] = E[Y_p] = 500\ \mu s$, $E[X_n^a] = E[Y_n^a] = E[X_n] = E[Y_n] = 85\ \mu s$ and $E[Y_t] = 32\ \mu secs$. For non active routers, these values are similar to those experimentally measured in [13]. For an active router, the data packet processing time is about 400 μs in [7]. Here we take $500\mu s$ for simplicity. For the throughput computation the packet size is 1024 bytes.

Fig. 3 shows the throughput achieved by each node in protocol S1 and S3 (base configuration) as a function of the number of local groups. We have 50 % of active routers and a loss rate of 5 %. We can observe that the minimal throughput in the system is first introduced by

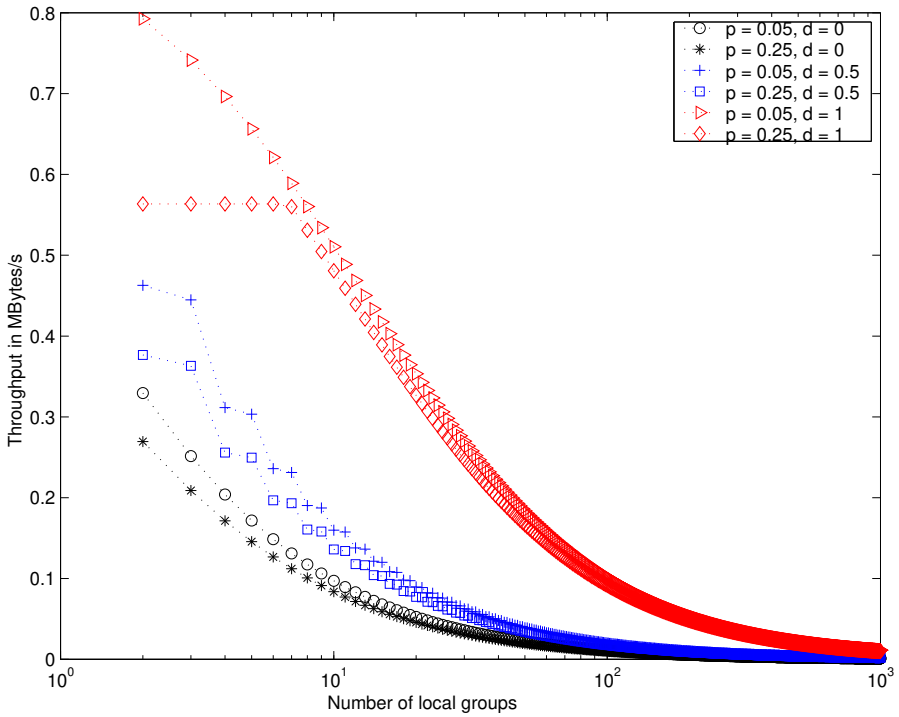
active routers. Then the source turns out to become the bottleneck. In S3, we can see that the linked receivers achieve a higher throughput compared to S1 which does not benefit from the subcast facility. Although not shown, S2 has similar curves than S1 but gives a slightly lower throughput.

Fig. 4 shows the throughput achieved by each node for the three protocols in the enhanced configuration. Compared to the base configuration, all nodes have a higher throughput. The source achieves the same throughput for any number of receivers because of its associated active router. Thanks to subcast facilities in $S3^{+}$, free receivers and active routers have a constant throughput.

Fig. 5 shows for S1 the benefit of active routers ($d \neq 0$) which permit a higher overall throughput, even with a high loss rate, and especially in the case of a large number of receivers. In the base configuration, local or global NACKs suppression strategies have very close performances in term of the overall throughput. Although not shown, S3 is identical to S1 in term of the overall throughput, S2 shows a slightly higher throughput.



(a)



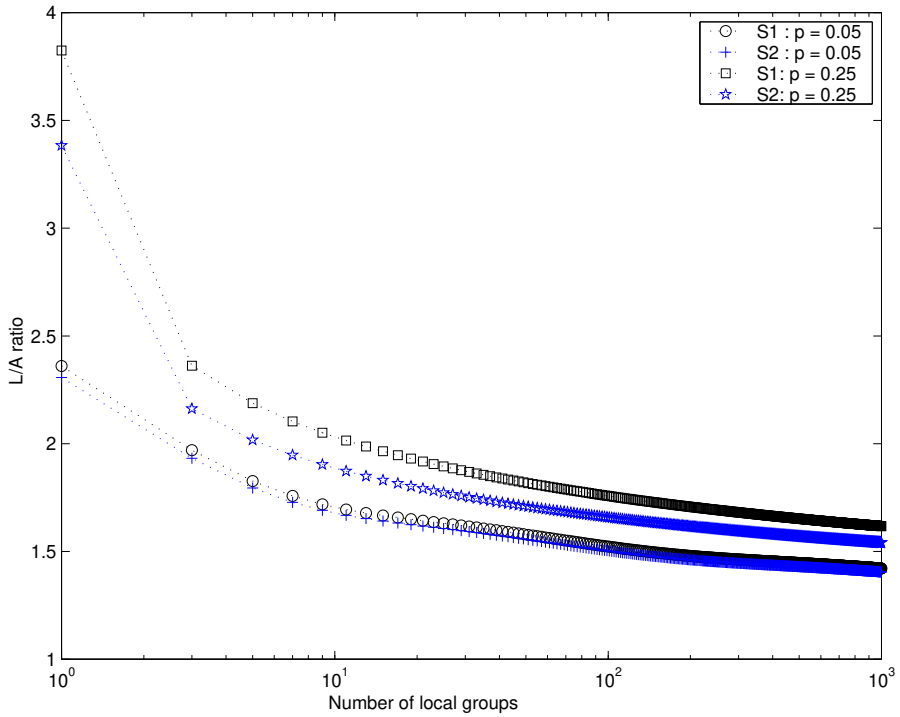
(b)

Figure 6: Throughput achieved by (a)$S2^+$. (b) $S3^+$. $B = 10$. ($S1^+$ is very similar to $S3^+$).

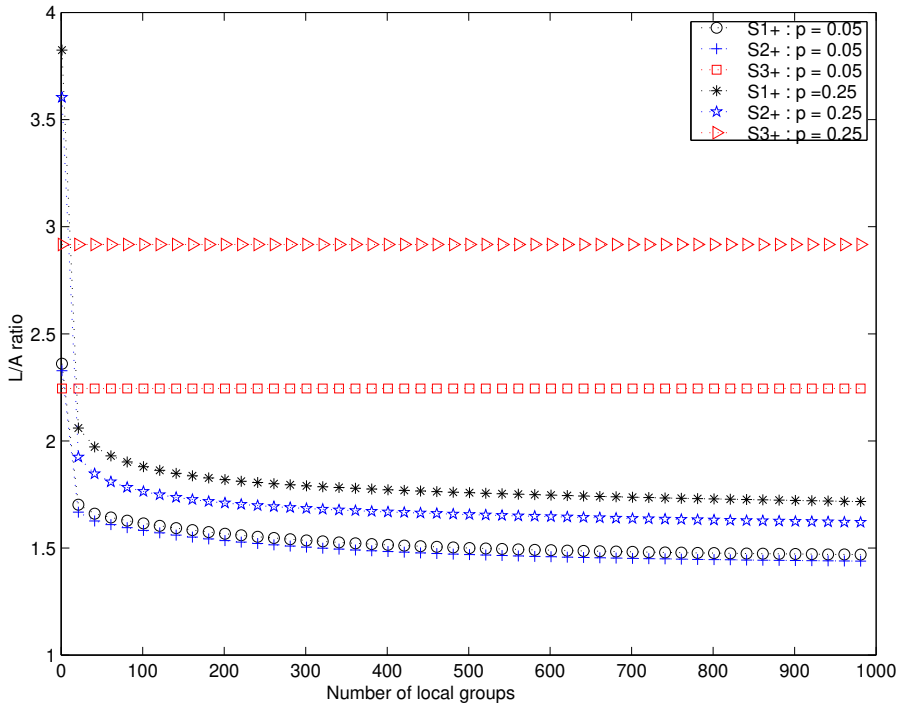
In the enhanced configuration (Fig. 6), $S2^+$ achieves a higher overall throughput than $S1^+$ and $S3^+$. The presence of the A_s active router in $S1^+$ and $S3^+$ is not profitable because, in contrast to $S2^+$, it is this active router that limits the throughput. When all the routers are active, $S2^+$ exhibits a constant throughput because the NACK local suppression overhead is independent of the size of the local group.

6.2 Best and worse cases analysis

The previous throughput analysis has shown that the overall throughput is limited by the active routers. The question now is: "How much the active routers processing power must be increased so that the network is never the bottleneck?". To answer this question, we plot the ratio between the throughput achieved by a free receiver (L) and an active router associated to the receivers (A) (free receivers are the end nodes that introduce the lowest throughput). This ratio gives the multiplicating factor put on an active router's processing power in order to achieve the same level of performance than a free receiver.



(a)



(b)

Figure 7: Active router/free receiver throughput ratio. (a) base configuration. (b) enhanced configuration. $B = 10$ and 50% of active routers.

Fig. 7(a) shows the L/A ratio for S1 and S2 (S3 behaves like S1). For instance, for one local group (i.e. 10 receivers) active routers in S1 need to be almost 4 times faster to achieve the same level of performance than free receivers. Moreover, we can observe that S2 needs a lower processing power at the active routers than S1. Fig. 7(b) shows the L/A ratio for the three protocols in the enhanced configuration. $S2^+$ is always the protocol that needs a lower processing power at the active routers. This is explained by the fact that an active router processes only one NACK per packet lost and per local group. $S3^+$ needs a higher processing power than the free receivers because they exhibit a higher throughput thanks to the subcast service.

Fig. 8 shows the impact of the active router density on the protocol's performances in term of the overall throughput in the base configuration. The figure plots the overall throughput gain as the number of active routers is increased compared to no active routers case. Several multiplicating factors to the active routers' processing power are applied. We can see that with the same processing time at the active routers and the receivers, the overall throughput can be an order of magnitude higher if all the receivers are linked. Most interestingly, if the active router's processing power is divided by 10, we can still double the overall throughput provided that 55 % of routers are active.

In the enhanced configuration (Fig. 9), if the number of active routers increases, $S1^+$ and $S3^+$ achieve a higher overall throughput compared to the no active routers case, even when the active router's processing power is considerably reduced. Unlike $S1^+$ and $S3^+$, $S2^+$ shows no improvements when the processing power in multiplied by 2 or more: active routers are not the bottleneck in $S2^+$ and increasing the number of active routers does not affect considerably the performances in term of overall throughput.

7 Previous analytical analysis of reliable multicast protocols

Cache of data implemented as an active service permits the local recovery of loss packets. Kasera et al. [10] have analytically shown that protocols using local recovery exhibit better performances. They compared 3 different strategies, noted L1, L2 and L3, for local recovery. L1 is called *server-based* and benefits from the cache of packets from *repair-servers* co-located with traditional routers at the edge of the core network. L2 is a generic version of SRM and L3 uses a logical tree structure for the recovery. Kasera et al. assume that the repair servers have

sufficient processing power and are never the bottleneck. The network model used is the one introduced by Pingali et al. [12]. Using a different model than the one introduced by Pingali et al., the authors in [14] have quantified the performances of recovery strategies with active services. Their analysis is based on the recovery latency and has shown that active approaches always perform better provided that a minimum storage capacity is available in the active routers.

None of these works have analytically investigated the two other active services: NACK suppression and subcast. One of our contributions in this field consists in the analysis of NACK suppression strategies and the potential of subcast in an active networking environment.

8 Conclusions

Reliable multicast is a difficult problem. Contributions of active routers within the multicast tree can be used for performing additional functionalities such as cache of data, feedback aggregation or subcast. In order to evaluate the potential of these mechanisms, we proposed three generic analytical models derived from a corresponding protocol's description: S1 that uses the global NACK suppression strategy, S2 that uses local suppression instead, and S3 that differs from S1 by the subcast facility. All three benefit from the cache of data.

Our analysis uses the processing requirements of the protocols to derive the achievable throughput at the various nodes in the multicast tree. The results showed that the active routers are the bottleneck in the system. Also, it appeared that local suppression, under the assumption that the random waiting time at a receiver is well-chosen, requires less processing power than global suppression. Subcast usually allows for a constant throughput at the receivers with more scalability.

We also studied the impact of the active router density on performances. Without an active router associated to the source, all three protocols have the same behavior: the achievable throughput increases as the number of active routers increases. Most interestingly, if we have slow, even ten times slower active routers, increasing their number allows for more performances. With an active router associated to the source, it appeared that the global suppression strategy performs better than the local one: increasing the number of active routers in $S2^+$ (local suppression) does not improve performance. In contrast, performances of $S1^+$ and $S3^+$ (both perform global suppression) improve greatly as the active router density increases.

Appendix

M_a and M_b

M_a is the number of retransmission until an active router has correctly received the packet. From $P[M_a \leq m] = 1 - p_l^m$, we have:

$$E[M_a] = 1 + \sum_{m=1}^{\infty} 1 - p_l^m = \frac{1}{1 - p_l}$$

M_b est le nombre de transmissions nécessaire à partir d'un routeur actif A_i $(i = 1..F)$ jusqu'à la réception correcte d'un paquet par un de ses récepteurs liés. From $P[M_b \leq m] = 1 - p_l^m$, we have:

$$E[M_b] = E[M_a] = \frac{1}{1 - p_l}$$

From [12], we can find:

$$E[M_a - 1] = E[M_b - 1] = \sum_{m=1}^{\infty} 1 - p_l^m = \frac{p_l}{1 - p_l}$$

$$E[M_a - 2] = E[M_b - 2] = \frac{p_l^2}{1 - p_l}$$

Variable aléatoire M_r

M_r est le nombre de transmissions nécessaire, à partir de la source jusqu'à la réception correcte d'un paquet par un récepteur (lié ou libre). From $P[M_r \leq m] = 1 - p^m$, we have:

$$E[M_r] = 1 + \sum_{m=1}^{\infty} 1 - p^m = \frac{1}{1 - p}$$

$$E[M_r - 1] = \sum_{m=1}^{\infty} 1 - p^m = \frac{p}{1 - p}$$

$$E[M_r - 2] = E[M_b - 2] = \frac{p^2}{1 - p}$$

Variable aléatoire M_B

M_B est le nombre de transmissions nécessaire, à partir d'un routeur actif A_i $(i = 1 \,..\, F)$ jusqu'à la réception correcte d'un paquet par tous ses B récepteus liés.

$$P[M_B \leq m] = \prod_{i=1}^{B} P[M_b \leq m] = (1 - p_l^m)^B.$$

$$E[M_B] = 1 + \sum_{m=1}^{\infty} 1 - (1 - p_l^m)^B$$

Variables aléatoires M_S, M_R et M_A

M_S est la variable aléatoire qui représente le nombre de transmissions nécessaire, à partir de la source jusqu'à la réception correcte d'un paquet par tous les F routeurs actifs et les $(N-F)B$ récepteus libres.

$$P[M_S \leq m] = \prod_{i=1}^{F} P[M_a \leq m] \prod_{j=1}^{(N-F)B} P[M_r \leq m]$$

or $P[M_a \leq m] = 1 - p_l^m$ et $P[M_r \leq m] = 1 - p^m$, on trouve :

$$P[M_S \leq m] = (1 - p_l^m)^F \ (1 - p^m)^{(N-F)B}$$

Pour calculer sa moyenne, on utilise la formule:

$$E[M_S] = 1 + \sum_{m=1}^{\infty} 1 - P[M_S \leq m]$$

Si $F = 0$ alors $M_S = M_R$. On a alors:

$$P[M_R \leq m] = (1 - p^m)^R$$

et par conséquent:

$$E[M_R] = 1 + \sum_{m=1}^{\infty} 1 - (1 - p^m)^R$$

Si $F = N$ alors $M_S = M_A$. On a alors:

$$P[M_A \leq m] = (1 - p_l^m)^R$$

et la moyenne est comme suit::

$$E[M_A] = 1 + \sum_{m=1}^{\infty} 1 - (1 - p_l^m)^N$$

Variables aléatoires M_{A_S} et M'_R

M_{A_S} est le nombre de transmissions nécessaire, à partir du routeur actif A_S jusqu'à la réception correcte d'un paquet par tous les F routeurs actifs et les $(N-F)B$ récepteus libres.

Or la probabilité de perte d'un paquet dans le backbone est nulle, alors M_{A_S} revient au nombre de transmissions nécessaire jusqu'à la réception correcte du paquet par les $(N-F)B$ récepteurs libres, soit :

$$P[M_{A_S} \leq m] = (1 - p_l^m)^{(N-F)B}$$

La moyenne est donnée par la formule:

$$E[M_{A_S}] = 1 + \sum_{m=1}^{\infty} 1 - (1 - p_l^m)^{(N-F)B}$$

Si $F = 0$ alors $M_{A_S} = M'_R$. On trouve:

$$P[M'_R \leq m] = 1 + \sum_{m=1}^{\infty} 1 - (1 - p_l^m)^R$$

Si $F = N$ alors $M_{A_S} = 1$. il suffit une seule transmission pour que le paquet soit correctement reçu par les routeurs actifs A_i $(i = 1..F)$ à partir du routeur A_S

References

[1] XTP Revision 4.0. *Xpress Transport Protocol Specification.* XTP Forum, Santa Barbara, March 1995.

[2] S. Floyd, V. Jacobson, C.G. Liu. *A Reliable Multicast Framework for Light Weight Session and Application Level Framing.* in Proc ACM SIGCOMM'95, pp 342-356, August 1995.

[3] S. Paul, K. K. Sabnani. *Reliable Multicast Transport Protocol (RMTP).* IEEE Journal of Selected Areas in Communications, Special Issue on Network Support for Multipoint Communication, Vol. 15, No. 3, pp 407-421, April 1997.

[4] R. Yavatkar, J. Griffioen, and M. Sudan. *A Reliable Dissemination Protocol for Interactive Collaborative Applications.* Proceedings' of the ACM Multimedia'95, November 1995.

[5] H. W. Holbrook, S. K. Singhal, D. R. Cheriton. *Log-Based receiver-Reliable Multicast for Distributed Interactive Simulation.* Proceeding of ACM SigComm'95, Vol.25, No.4, pp 328-341, October 1995.

[6] D. L. Tennehouse, J. M. Smith, W. D. Sincoskie, D. J. Wetherall, and G. J. Winden. *A Survey of Active Network Research.* in IEEE Communications Magazine, pp 80-86, January 1997.

[7] Li-Wei H. Lehman, S. J. Garland, and D. L. Tennenhouse. *Active Reliable Multicast.* Proceedings of IEEE INFOCOM'98, San Francisco, CA, March 1998.

[8] S. K. Kasera, S. Bhatacharyya, M. Keaton, D. Kiwior, J. Kurose, D. Towsley and S. Zabele. *Scalable Fair Reliable Multicast Using Active Services*, IEEE Network Magazine's Special Issue on Multicast 2000.

[9] C. Papadopoulos, G. Parulkar and G. Varghese. *An Error Control Scheme for Large-Scale Multicast Application.* In Proceedings of the IEEE INFOCOM, pp 1188-1996, March 1998.

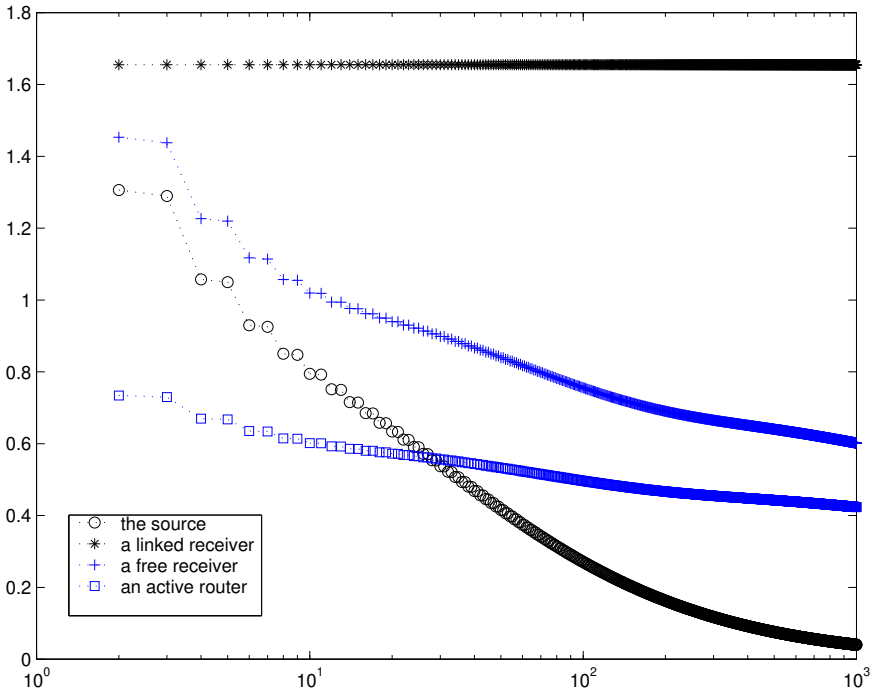
[10] S. K. Kasera, J. Kurose and D. Towsley. *A Comparaison of Server-Based and Receiver-Based Local Recovery Approaches for Scalable Reliable Multicast.* In Proceedings of IEEE INFOCOM, San Francisco, CA, USA, March 1998.

[11] M. Yajnik, J. Kurose, and D. Towsley. *Packet Loss Correlation in the MBone Multicast Network.* in Proceedings of Global Internet Conference, November 1996.

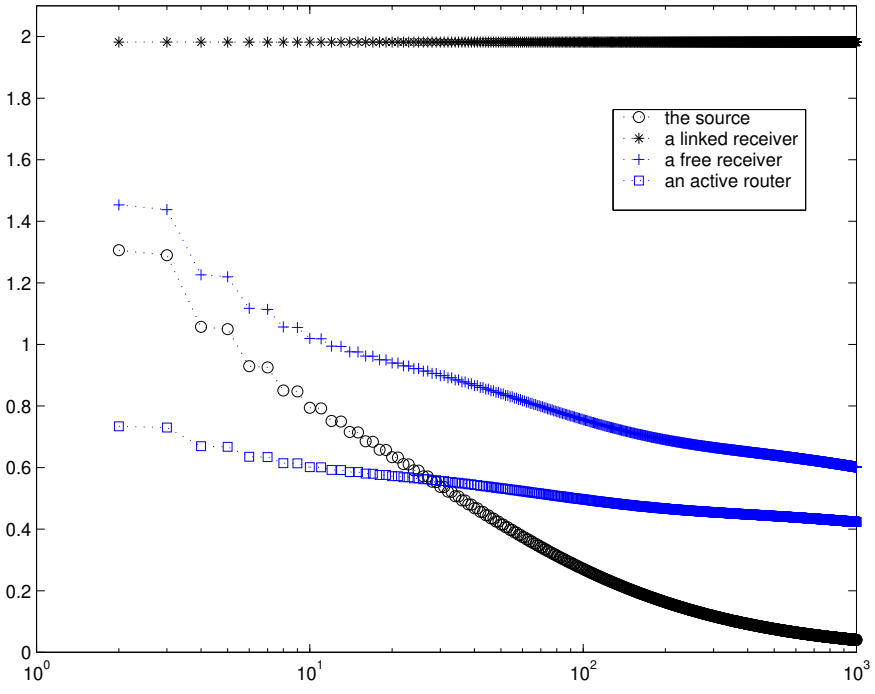
[12] S. Pingali, D. Towsley, and J.F. Kurose. *A Comparaison of Sender-initiated and Receiver-initiated Reliable Multicast protocols.* in Proceedings of ACM SIGMETRICS'94, Vol 14, pp 221-230, 1994.

[13] S. K. Kasera, J. Kurose and D. Towsley. *A Comparaison of Server-Based and Receiver-Based Local Recovery Approaches for Scalable Reliable Multicast.* CMPSCI Technical Report TR 97-69, December 1997.

[14] S. Bhattacharjee, K. L. Calvert, E. W. Zegura. *Active Networking and the End-to-End Argument.* Proc ICNP'97, Atlanta, GA, October 1997.

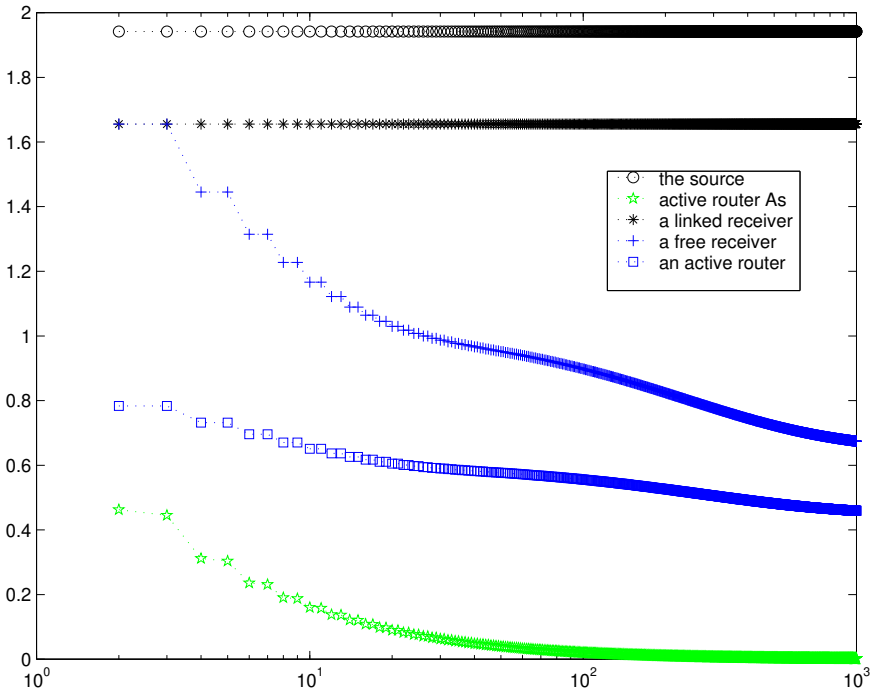


(a)

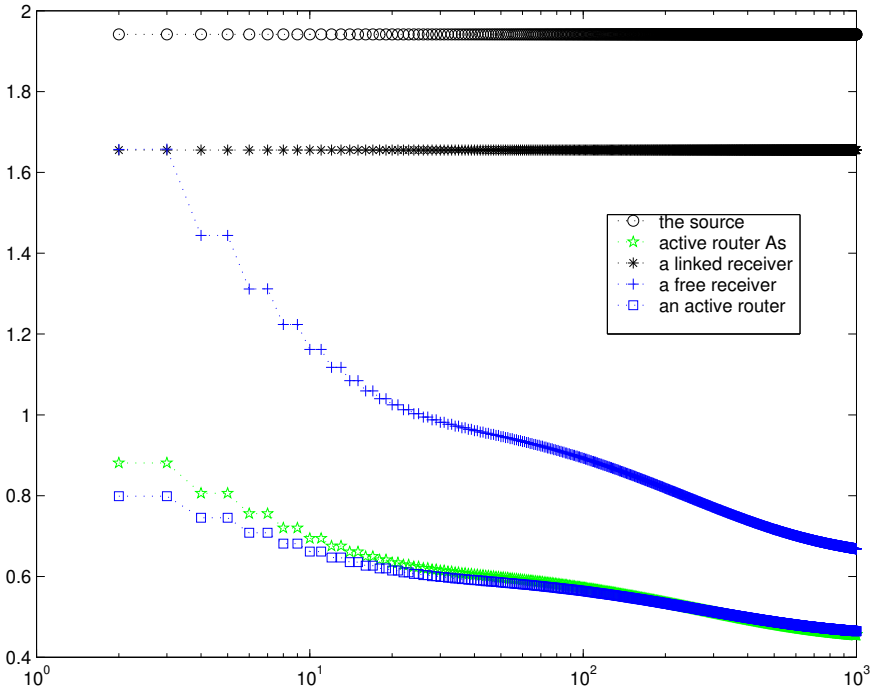


(b)

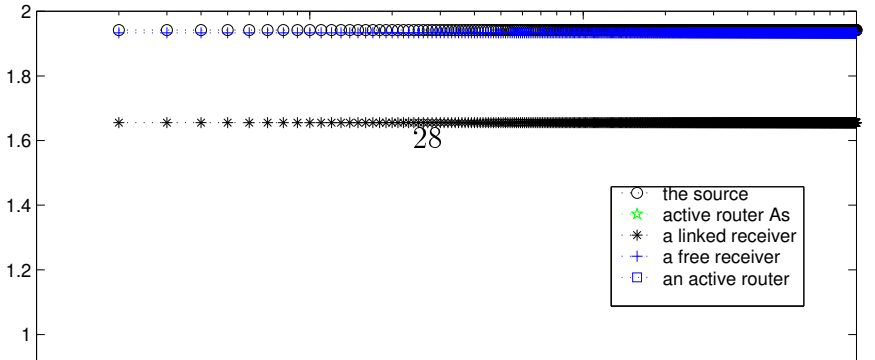
Figure 3: Throughput of the various nodes with (a) S1 and (b) S3. p = 0.05, B = 10, 50% of active routers.


the source
active router As
a linked receiver
a free receiver
an active router

(a)


the source
active router As
a linked receiver
a free receiver
an active router

(b)


the source
active router As
a linked receiver
a free receiver
an active router

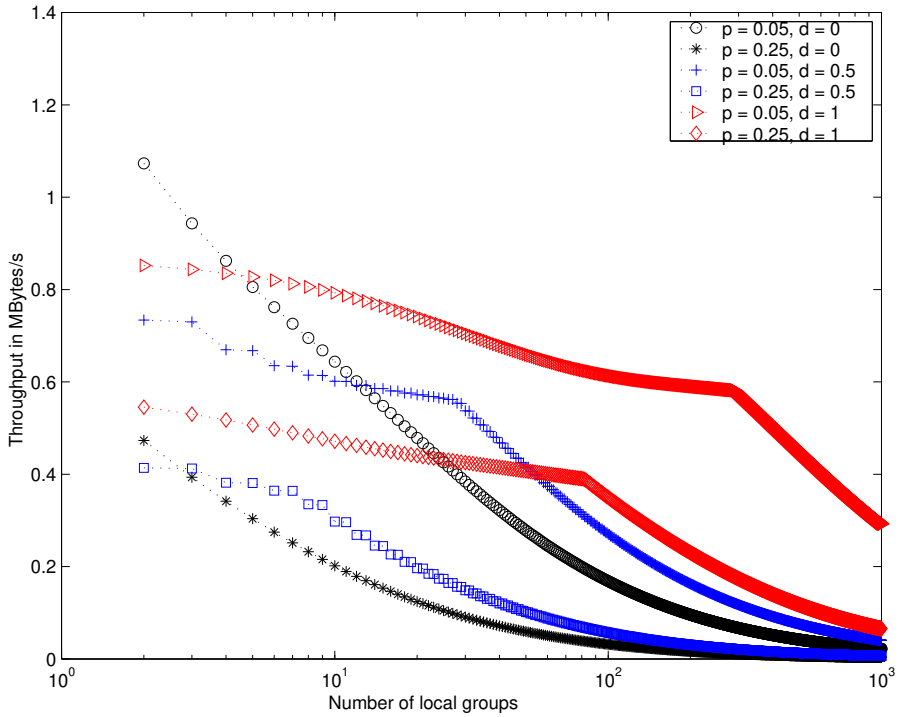


Figure 5: Throughput achieved by S1.

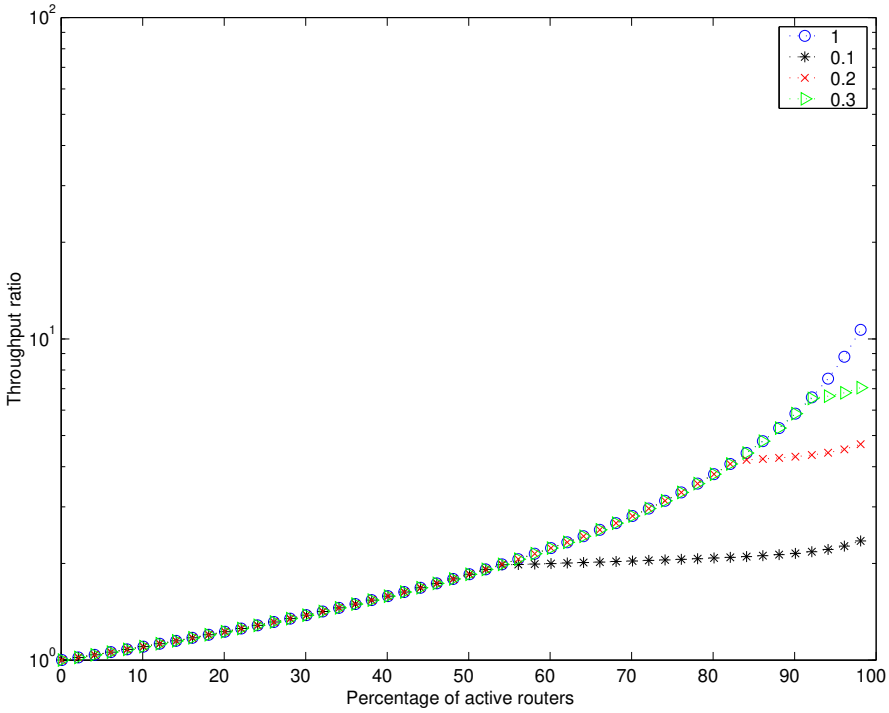
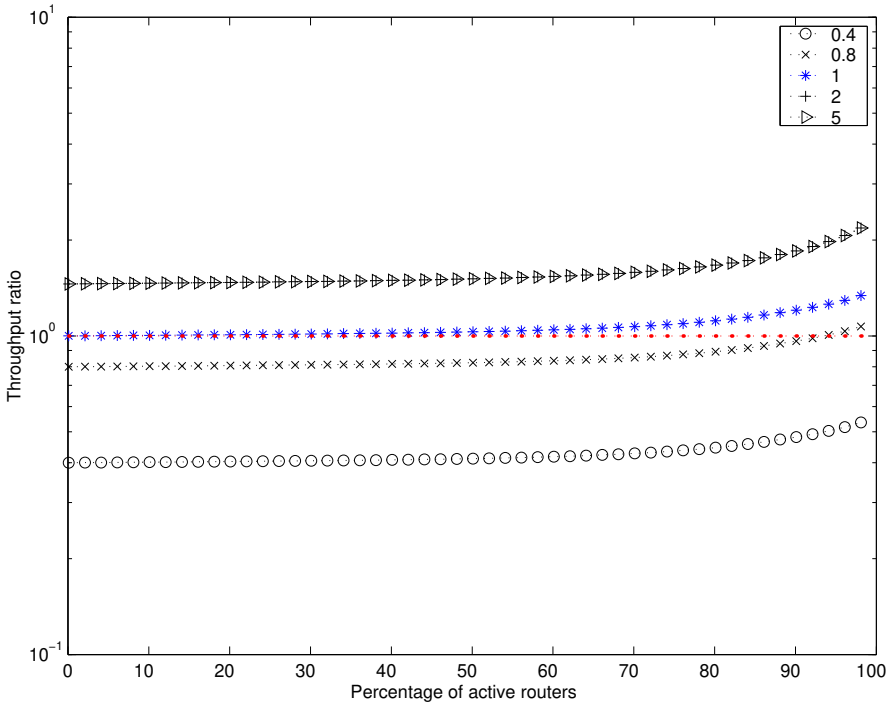
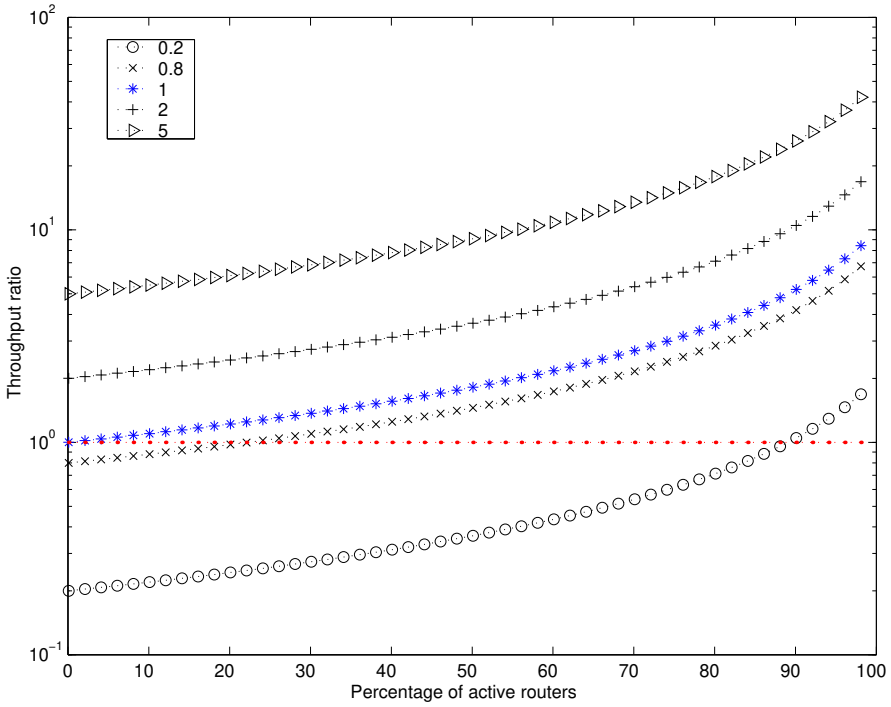


Figure 8: Overall throughput achieved by the 3 protocols in the base configuration with F active routers vs no active routers. $p = 0.05$, $B = 10$.



(a)



(b)

Figure 9: Overall throughput achieved by (a) $S2^+$, (b) $S3^+$, with F active routers vs no active routers. $p = 0.05$, $B = 10$. ($S1^+$ is similar to $S3^+$)