Reliable Multicast Transport by Satellite: a Hybrid Satellite/Terrestrial Solution with Erasure Codes

F. de Belleville[1,3], L. Dairaine[1,3], J. Lacan[1,3], C. Fraboul[2,3]

[1]ENSICA, 1, Place Emile Blouin, 31056 Toulouse, Cedex 5, France
{Florestan.de-Belleville, Laurent.Dairaine,Jerome.Lacan}@ensica.fr
[2]ENSEEIHT /IRIT : 2, rue Charles Camichel, BP 7122 - F 31071 Toulouse
Christian.Fraboul@enseeiht.fr
[3]TéSA : Laboratoire coopératif en Télécommunication Spatiales et Aéronautiques 2, rue Charles Camichel BP 7122 - F-31071 Toulouse

Abstract. Geostationary satellites are an efficient way to provide a large scale multipoint communication service. In the context of reliable multicast communications, a new hybrid satellite/terrestrial approach is proposed. It aims at reducing the overall communication cost using satellite broadcasting only when enough receivers are present, and terrestrial transmissions otherwise. This approach has been statistically evaluated for a particular cost function and appears to be advantageous. Then since the hybrid approach relies on Forward Error Correction, several practical aspects of MDS and LDPC codes are investigated in order to determine impact of code selection.

1 Introduction

The support for multipoint communications is an interesting extension of the services proposed in nowadays Internet. Numerous applications like multimedia streaming and software updates would indeed take advantage of such a service. Today, solutions based on the ubiquitous Internet Protocol (IP) exist: IP Multicast [1]. However availability of IP Multicast at end-user for large group communication over the whole internet is still limited for the time being. This lack of deployment is mainly due to technical concerns and economical issues [2].

A practical solution to implement IP multicast service may consist in using a geostationary satellite. The broadcast nature and the large coverage zone of such systems make it possible for a source to reach a huge number of receivers with only one hop. Satellite broadcasting seems expensive at first sight, but per-receiver cost becomes less than using terrestrial network when the number of receivers increases. For this reason this study is focused on large scale reliable multipoint communications. Moreover we consider applications with no time constraints, because the long transmission delay of satellite links is not really compatible with such applications. Software updates or cache feeding in content delivery networks are examples of possible target applications.

During a satellite transmission the receivers which experienced losses have to wait for more information to recover those losses. In consequence more time is required to transmit data to them. Thus the number of receivers which continue to be interested by the satellite broadcast will decrease along the transmission. Our approach consists in transmitting data via satellite only when the number of receivers is sufficiently high (i.e. when per-receiver cost is less than with terrestrial network) and via terrestrial transmissions otherwise.

Reliable multicast communications have been extensively studied in the past decade. As it is not possible to design a *one-size-fits-all* reliable multicast transport protocol, numerous propositions exist [8]. Among these researches, the use of Forward Error Correction (FEC) for reliable multicast transport protocols has been shown to be especially interesting [3] because it increases their scalability. At transport layer FEC implies use of an erasure code like Maximum Distance Separable (MDS) [4] or Low Density Parity Check codes (or like) [5]. MDS codes are optimal codes in terms of the amount of data needed to decode information, but imply long processing times which limit the amount of data that can be encoded. On the opposite, LDPC codes are not optimal, but do not have the limits of MDS codes. To determine the appropriate code for the hybrid approach presented, practical aspects of these codes must be studied.

The remaining of the paper is organised as follows: in part 2 the context of the study is specified, and a new approach for reliable multicast transmission is proposed and statistically evaluated. Then the overhead and processing times of MDS and LDPC codes are evaluated in part 3. Finally conclusions and future work are presented in part 4.

2 A Reliable Transport Protocol Designed for Hybrid Satellite-Terrestrial Network

In this section, the context of the study is specified. Then the target transport service is described, as well as a proposition of a hybrid satellite/terrestrial approach to achieve a reliable multicast transport service. The last part presents a statistical study of this approach.

2.1 Communication System

The system considered is a hybrid satellite/terrestrial network, i.e. end-users are connected to both satellite and terrestrial networks. Assumptions are similar to the DIPCAST project [20]: end-users are connected to terrestrial network via a high speed access network (e.g. xDSL or LAN), and to satellite system either via a high speed access network, or directly with a Very Small Aperture Terminal (VSAT).

The satellite system uses a geostationary satellite, and proposes a best effort multipoint communication service based on IP Multicast. This supposes that a protocol which manages joining and leaving procedures is integrated to the satellite system, as well as tree establishment algorithm. This problem of integration is out of the scope of this paper and is not further investigated.

In order to be representative of today Internet, terrestrial network is not supposed to support multipoint transmissions. Thus any terrestrial communication in this hybrid system is a point-to-point transmission.

Finally the following assumptions are made on the application:

– The application transmits data from one source to a large group of receivers (one-to-many communications)
– This application does not have strong delay constraints.
– A service transmitting session characteristics (file properties, start time, associated group address, etc.) is available. Those transmissions can be done via out of band means (e.g. e-mail), or via session management tools.
– No receiver can join the session after the beginning (late joining is not supported).

2.2 Multicast Transport Service

Since satellites are really advantageous for large scale data transmissions, applications which transfer files to an important numbers of receivers (several hundreds or more) are considered in the present paper. Furthermore we consider applications which must be assured that the whole group has received transmitted information. An example of such applications is the transport of multimedia files (e.g., video, music, games, etc.) towards a large set of users. Another issue concerns the overall communication cost. The utilisation of satellite links is indeed quite expensive. Nevertheless when the number of receivers increases, the per-user cost decreases. Thus any application being charged according to its bandwidth utilisation may prefer protocols which carefully watch communication cost.

Considering file transport using the best effort protocol IP Multicast over satellite, a fully reliable transport protocol must be used. In the following paragraphs some features of this protocol required in the above-mentioned context are exposed.

In the first place the transport protocol must guarantee that all receivers in the group receive transmitted information. Several multicast transport protocols propose a statistically reliable service [8]. Although it allows designing transport protocols with no return channel, full reliability is not ensured because there is no adaptation to really occurred losses. This technique is then not convenient for the aimed purpose.

In the second place, since satellite bandwidth is expensive, the transport protocol must ensure that any useless satellite transmission is avoided. This confirms that statistical reliability is not recommended in the presented context because systematic coding of information implies a potential waste of bandwidth [6].

Eventually, as considered applications are designed for transmissions towards very large groups, underlying protocols (and specifically transport protocol) must scale very well. In particular for the transport layer, mechanisms of feedback suppression like [7] must be studied and configured for a satellite link.

2.3 A Hybrid Satellite/Terrestrial Approach

Numerous multicast transport protocols have been designed in the last few years to achieve efficient and scalable multicast transmissions [8]. From all researches on reliable multicast transport, a technique referenced as Hybrid ARQ type II [3] has emerged. It is an efficient way to diminish used bandwidth and improve scalability. It consists in using FEC combined with Automatic Repeat reQuest (ARQ) in the following way: after a transmission, the source asks for the maximum number M of missing packets. Then it generates and transmits M new encoded packets. As those M packets have not already been transmitted, they are useful for any receiver which experienced losses. Hybrid ARQ type II allows to greatly reduce the amount of retransmitted information in multipoint transmissions, and is then particularly interesting for any large scale full reliable transport protocol [3].

According to authors, most of the protocols presented in [8] are usable with satellite links because they support asymmetric transmissions. Nevertheless no previous work considers the overall communication cost. With the reliability mechanisms used, either packets are systematically retransmitted to the whole group, or each missing packet is retrieved with a point-to-point connection. In our context when point-to-point retransmissions are used, if a packet is requested by numerous receivers, satellite broadcasting may be profitable. On the contrary, if every missing packet is transmitted to the group via satellite, when only few receivers are concerned, few terrestrial point-to-point connections are cheaper. According to this simple statement, the trade-off between satellite and terrestrial bandwidth use may be studied, and it is our belief that optimisation of this trade-off is an interesting way to reduce the overall communication cost.

An interesting approach may be to define a threshold R_{min} representing the minimum number of receivers for satellite broadcasting to be advantageous (taking economical costs into account). A session would then behave as follows: at a predefined time, the satellite transmission starts. During this transmission, the source periodically estimates session size. A receiver is considered to belong to the session as long as it has not received the whole transmitted information. Several papers have addressed the question of estimating multicast session size [9]. Although proposed mechanisms would have to be adapted -or at least configured- for satellite links, we assume in this paper that an effective mechanism is available. Thus once the entire initial information has been transmitted, session size is likely to decrease (all receivers which experienced no losses quit the session). The source then goes on estimating session size while it transmits encoded redundancy packets to repair losses. When estimated session size goes below R_{min}, the satellite transmission stops. Receivers which do not have enough information to decode received data (i.e. receivers experiencing high loss rates) then contact other receivers to recover missing data. This terrestrial recovery can be done using peer-to-peer services for example. When all receivers have fully received the information, the session stops.

2.4 Statistical Study

The present paragraph gives an illustration of the gain generated by the hybrid satellite/terrestrial approach. For that purpose a statistical study is presented, with the following assumptions:

Errors and Fades. For terrestrial transmissions, packet losses are only supposed to be produced by network congestions. These congestions cause at the transport layer a Packet Loss Rate (PLR) of 5% for point-to-point communications [16] and 10% for multipoint communications [17].

For satellite transmissions, three categories of receivers are considered. The first one corresponds to all receivers under a clear sky. Assumption is made that these ones do not experience any losses (possible issues due to scintillation are not considered) and that this category encompass 90% of the receivers. The second category is supposed to include 9.9% of the receivers, and corresponds to end-users under a rainy sky. Fades due to light rain are supposed to cause a PLR of 20% at transport level. Finally, the last category encompass receivers under a stormy weather. This category is supposed to include 0.1% of the receivers which experience a PLR of 60%. According to [18] those values are realistic for a satellite communication using the Ka Band.

Note that all losses are supposed to be independent and uniformly distributed among packets. This assumption is not realistic, but since transport layer is supposed to implement ARQ type II technique, only the amount of lost packets is important for our computations. Assumption was also made that no packet are lost on the return channel.

Cost Function. In order to compare costs generated by hybrid satellite/terrestrial multipoint communications with pure terrestrial and satellite multipoint communications, it is necessary to first define a cost function. We choose to adopt a per-packet cost approach, and then define a cost function as follows:

$$F_X(R,K) = \alpha_X \times C_X(R,K), \quad X \in \{TU, TM, SM\}, \tag{1}$$

were K is the number of packets to transmit, R the number of receivers and α_x the per-packet transmission cost. $C_X(R,K)$ represents the average number of packets passing through the network in order to transmit K packets to R receivers (taking losses into account). The indices TU, TM and SM correspond to Terrestrial Unicast, Terrestrial Multicast and Satellite Multicast communications.

For terrestrial point-to-point communications R connections experiencing independent and uniformly distributed losses (represented by the PLR) are used to transmit K packets to R receivers. Then:

$$C_{TU}(R,K) = K \times R \times \sum_{i=1}^{\infty} i \times (1 - PLR) \times PLR^{i-1} = \frac{K \times R}{1 - PLR}. \tag{2}$$

For multipoint communications, the probability P(N,PLR,R) that exactly N packets are needed so that all end-users receive K packets can be expressed as: the probability that each end-user receive K packets among N, minus the probability that R receivers receive K packets among (N-1). Then P(N,PLR,R) can be calculated as:

$$P(N, PLR, R) = \tag{3}$$

$$\begin{cases} \left[\sum_{i=0}^{N-K}\binom{N}{i}PLR^i(1-PLR)^{N-i}\right]^R - \left[\sum_{i=0}^{(N-1)-K}\binom{N-1}{i}PLR^i(1-PLR)^{(N-1)-i}\right]^R, when\, N > K \\ \left[(1-PLR)^K\right]^R, when\, N = K \end{cases}$$

For terrestrial multicast communications, according to [18] transmitting data to R receivers is equivalent to $R^{0.8}$ point-to-point connections. The average number of packets passing through the network is then:

$$C_{TM}(R,K) = \left[\sum_{N=K}^{\infty} N.P(N, PLR, R)\right] \times R^{0.8}. \tag{4}$$

For satellite multipoint communications, since we defined three categories of receivers and since no multicast tree is established, we have:

$$C_{SM}(N,K) = \max\left\{\left[\sum_{N=K}^{\infty} N.P(N, PLR_i, \beta_i R)\right], i \in [0,2]\right\} \tag{5}$$

where $PLR_0=0$, $PLR_1=20\%$, $PLR_2=60\%$, $\beta_0=90\%$, $\beta_1=9.9\%$ and $\beta_2=0.1\%$.

For hybrid satellite/terrestrial transmissions, the cost is defined as the sum of the costs generated by terrestrial network utilization, and satellite system utilization.

Results. Using the cost function defined above, R_{min} has been computed for group sizes ranging from 100 to 600,000. Then the hybrid satellite/terrestrial approach has been compared to terrestrial point-to-point and multipoint communications as well as to pure multipoint satellite communications. Figure 1 shows the results for $\alpha_U = \alpha_M = 1$, $\alpha_S = 100$ and $K = 100$. The different levels perceptible on the curve representing the hybrid communication cost are due to the model definition: when group size increases, the number of receivers under a rainy sky or a stormy sky increases as well. Thus it becomes necessary to repair more and more losses using satellite link. In this example, the hybrid satellite/terrestrial approach induces a gain ranging from 10% to 50% compared with the most advantageous of the three classical approaches.

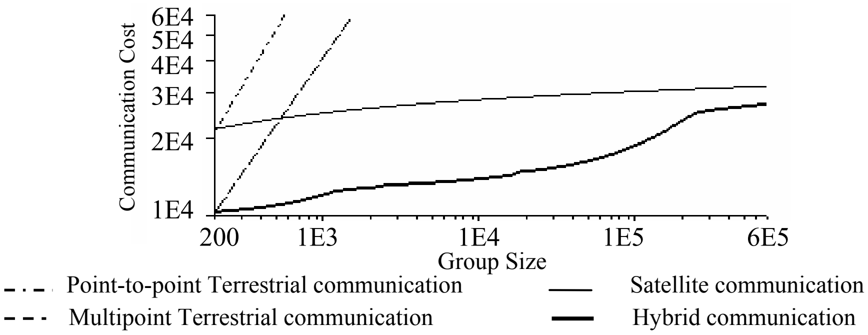

Fig. 1. Cost generated by terrestrial, satellite and hybrid communications. $\alpha_{TU} = \alpha_{TM} = 1$, $\alpha_{SM} = 100$ and K = 100

The hybrid satellite/terrestrial approach relies on hybrid ARQ type II. As mentioned previously this technique implies to encode transmitted information. The remaining of the paper is focused on the problem of code selection for the reliability mechanism.

3 Codes for an Erasure Channel in Hybrid Transport Protocol

At transport layer, because lower protocol layers detect and discard corrupted packets (using e.g. checksums) protocols only have to deal with missing packets in a stream (referenced as erasures). In consequence the hybrid ARQ type II technique supposes that a code is available which allows transport protocol to encode redundancy packets on-demand, and receivers to decode information assuming a sufficient set of packets is received. Two types of codes exist which have such properties: MDS codes [4] and LDPC (or like) codes [5].

One of the early propositions to introduce Maximum Distance Separable (MDS) codes at transport layer was presented in [13]. MDS erasure codes are an important class of erasure codes since they are optimal in terms of required data redundancy for error repair: for K packets encoded in N packets, reception of any K packets among the N allows receivers to decode information. Nevertheless the number N_{max} of packets potentially generated is limited because processing time become prohibitive when N_{max} increases (usually $N_{max} < 2^{16}$). Since $K < N_{max}$ the number of packets potentially encoded is also limited. Then when K is fixed, using hybrid ARQ type II technique induce that at most N_{max} -K losses occur.

Unlike MDS codes, LDPC codes [5,10] are not optimal: when N packets are generated from K original packets, $K(1 + \varepsilon)$ packets are required to decode information. Though ε tends to zero as K approaches infinity, for reasonable values of K, at least 5% of additional data is necessary to decode information [11]. Such codes still have great advantages: coding and decoding speeds are really faster than the MDS ones, and there is almost no limit to the number of redundancy packets generated.

Implementation issues for these two type of codes are investigated in the following paragraphs.

Processing Speeds. Standards implementations of MDS [14] codes have been used to evaluate processing speeds with recent computers. Tests were done on a 1GHz Pentium class, and the tested MDS codes are limited to $N_{max} = 2^{16}$ (so that the number of generated packet is not too much limited). For values of K ranging from 16 to 256, coding and decoding speeds were evaluated to be superior to 1MB/s. So the use of MDS codes up to K = 256 is acceptable as long as sending rate is inferior to 1MB/s. For LDPC codes, the order of magnitude of processing speeds is of several thousands MB/s [11].

Computer Memory Requirements. Memory occupation was also evaluated with standard implementations of MDS and LDPC codes. Several objects have to be stored so that it is possible to encode and decode information. The amount of memory occupied by those objects was evaluated to less than one megabyte for both types of codes, which seems an acceptable value.

Source packets storage is another cause of memory occupation. With satellite transmissions, losses occur in bursts. For MDS codes it is recommended to interleave several blocks of encoded packets in order to distribute losses among blocks (then the probability to exceed the maximum number of redundancy packets potentially generated decreases). For example, all fade listed in [19] last at most 6000 seconds. With a packet size of 1500 bytes, K = 256, $N_{max} = 2^{16}$ and a sending rate of 1Mb/s, 62 blocks have to be interleaved so that the maximum number of packets that can be generated is not exceeded. This implies that approximately 24MB of memory is occupied by packet storage. Note that Byers et al. pointed out that block interleaving may imply a reception overhead [10], but since these results do not consider losses in bursts, they are not representative of satellite transmissions.

For standard LDPC codes with large K values, the occupied memory may be too large depending on the chosen implementation. For this reason, we consider in this paper that K equals at most 50,000. Such K is sufficiently large so that the overhead factor is not prohibitive, and the implementation of such codes is not impossible (storage of all packets represents approximately 75 MB with a packet size of 1500 bytes).

Transport Protocol Header Overhead. All information that must be sent to receivers is defined in RFC 3452 [12]. Information that must be embedded in data packets header is referenced as *FEC payload ID*. The predefined FEC payload ID 128 and 129 specifies possible structures for information related to FEC. Both consist in 8 bytes added to protocol header, and for FEC payload ID 129 (adapted to MDS codes), only 6 bytes are used to identify encoded object.

Backchannel Traffic. The difference between the numbers of packets potentially encoded in a block has also an impact on the traffic generated on the return channel. ARQ type II technique implies that the source is informed about the missing packets for each block. If the same scalable mechanism is used to sent information to the source (like the one presented in [7]), the same average number of messages N_{Back} is generated. Then for each block it is necessary to send the number of missing packets and the block number. The total length of this information is 8 bytes for LDPC codes, and at least 6 bytes for MDS codes. Thus the minimum ratio of amount of traffic generated by MDS codes and LDPC codes is:

$$R = \frac{B_{MDS} \times 6 \times N_{Back}}{B_{LDPC} \times 8 \times N_{Back}}. \qquad \textbf{(6)}$$

Were B_X is the number of blocks used by the MDS or LDPC codes. B_X can be expressed as:

$$B_X = \left\lceil \frac{F}{K_X \times MTU} \right\rceil,\ X \in \{MDS, LDPC\}, \qquad \textbf{(7)}$$

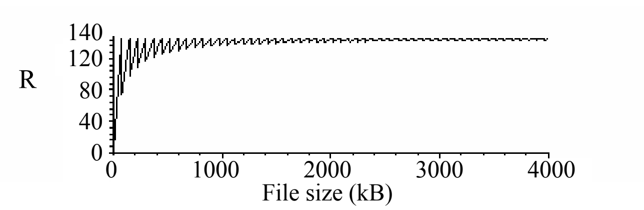

Fig. 2. Evolution of the ration of generated traffic on backchannel versus file size. $K_{MDS} = 256, K_{LDPC} = 50{,}000, MTU = 1496$

were F is the file size in bytes, K_X the number of encoded packet per block, and MTU the Maximum Transfer Unit in bytes. As depicted in Figure 2, this ratio is at least equal to 70 for transmitted files of 100 megabytes or more. The large variations are due to the operator *ceil( )*.

Discussion Results presented above lead to a simple conclusion: in the specified context, MDS codes are designed for bandwidth optimisation, whereas LDPC codes are designed to release constraints on end-users. Thus the code selection is up to the transport protocol designer. If network resources are cheap enough, a LDPC code make it possible for the transport protocol to increase end-user satisfaction: as coding and decoding processes are fast, information is available quickly.

On the other hand, as soon as losses occur, the use of LDPC codes implies an intrinsic waste of bandwidth due to the overhead factor. As the main goal of the hybrid satellite/terrestrial approach is to limit satellite bandwidth use, MDS codes have to be used to be coherent. In this case sending rates must not exceed encoding speed, (1MB/s with K=256 in the tested configuration). Moreover since the amount of feedback traffic is proportional to the number of blocks, information related to all interleaved blocks has to be aggregated in order to limit traffic generated on the return channel.

When available bandwidth is greater than encoding speed of MDS codes, a solution may consist in encoding packets before the session starts. This technique requires a large amount of memory, as all encoded packets have to be stored, but all those packets may be stored on hard drive. If it is not possible because of memory considerations, or because source system can not spend time induced by encoding process, the only way to use all the available bandwidth is to use LDPC codes.

4 Conclusion and Future Work

In the context of hybrid satellite/terrestrial network a proposition of large scale reliable multicast transport protocol is presented. Unlike previous works, our main objective is to reduce the overall communication cost. To achieve this goal, a new approach for a reliable multicast transport protocol is proposed. This approach consists in using satellite transmission only when the number of receivers is sufficiently large, and terrestrial transmissions otherwise.

Coupled with this principle, the use of hybrid ARQ type II is an efficient technique for achieving reliability and scalability. This technique implies selection of a MDS or of a LDPC code. The service provided by each type of codes and practical implementation issues have been investigated. As a conclusion, though MDS codes are optimum for bandwidth utilization, code selection depends on connection characteristics and on the importance given to the end-user satisfaction.

To achieve the proposed approach, an issue is still to be studied: how receivers retrieve missing information with the terrestrial network once satellite transmission stops. After satellite broadcasting, a set of encoded packet is distributed among receivers. Several papers underlined the advantages of using peer-to-peer techniques in such a situation [15]. In consequence it may be interesting to further study integration of a peer-to-peer mechanisms to retrieve losses.

References

1. S. Deering： Host Extensions for IP Multicasting. RFC 1112 (1989).
2. Christophe Diot and Brian Neil Levine and Bryan Lyles and Hassan Kassem and Doug Balensiefen: Deployment issues for the IP multicast service and architecture. IEEE Network, vol.14, N. 1, p.78-88 (2000)
3. J. Nonnenmacher, E. W. Biersack, D. Towsley: Parity-based loss recovery for reliable multicast transmission. IEEE ACM Transactions on Networking (1998), vol.6 , p.349-361.
4. R.E.Blahut: Theory and Practice of Error Control Codes. Addison Wesley, MA, 1984.
5. R.G. Gallager,: Low-Density Parity-Check Codes. IRE Transaction Information Theory (1962), IT-8, p. 21-28.
6. M. Jung, J. Nonnenmacher and E. Biersack: Uni-directional versus Bi-directional Communication. Kommunikation in Verteilten Systemen, p. 264-275 (1999)
7. Jorg Nonnenmacher and Ernst Biersack: Optimal Multicast Feedback. INFOCOM (1998), p. 964-971
8. M.W. Koyabe, G. Fairhurst: Reliable Multiast by Satellite: A Comparison Survey and Taxonomy. International Journal of Satellite Communications (2001), Vol:24(1), p. 21-26.
9. S. Alouf, E. Altman, P. Nain: Optimal on-line estimation of the size of a dynamic multicast group. INFOCOM (2002), New York, USA
10. Byers, J.W., Luby, M., Mitzenmacher, M. and A. Rege: A Digital Fountain Approach to Reliable Distribution of Bulk Data. Proceedings ACM SIGCOMM (1998)
11. J. S. Plank, M. G. Thomason: On the Practical Use of LDPC Erasure Codes for Distributed Storage Applications'. Technical Report UT-CS-03-510, University of Tennessee, 2003.
12. M. Luby, L. Vicisano, J. Gemmell, L. Rizzo, M. Handley, J. Crowcroft: Forward Error Correction (FEC) Building Block. Experimental RFC 3452 (2002)
13. A. J. McAuley: Reliable Broadband Communication Using a Burst Erasure Correcting Code. Proc. of SIGCOMM (1990), p. 297-306
14. J. Blomer, M. Kalfane, R. Karp, M. Karpinski, M. Luby and D. Zuckerman: An xor-based erasure-resilient coding scheme. Technical report (1995), International Computer Science Institute, Berkeley, California.
15. J. Lacan, L. Lancérica, L. Dairaine: When FEC Speed up Data Access in P2P Networks. Proceedings of Interactive Distributed Multimedia Systems Conference (2002), Coimbra, Portugal.
16. V. Paxson: End-to-end Internet packet dynamics. In Proc. ACM SIGCOMM, pp. 139--152, September 1997

17. M. Yajnik, J. Kurose, and D. Towsley: Packet loss correlation in MBone multicast network. In Proc. IEEE Global Internet Conference, Part of GLOBECOM'96, Nov. 1996, pp. 94--99.
18. John C.-I. Chuang, Marvin A. Sirbu: Pricing Multicast Communication: A Cost-Based Approach. Telecommunication Systems, Vol. 17, p. 281-297 (2001).
19. J. Pinder, L. J. Ippolito, S. Horan: Four years of Experimental Results from the New Mexico ACTS Propagation Terminal at 20.185 and 27.505 GHz. IEEE on Selected Areas in Communication, Vol.17, N° 2, p.153-162, 1999
20. DVB comme support d'IP multiCAST par satellite. RNRT project No. 67. http://www.dipcast-satellite.com/