INRIA ROCQUENCOURT

The Tele-seminar networking with TELESIA

Pierre Léonard, Alain Caristan, Pierre de La Motte, Andrzej Wozniak

Programme Development

Projet MÉDIA

24 pages

Abstract: Since 1992, the TÉLÉSIA project takes the challenge to broadcast the bimonthly seminar organized by the Aristote association. This broadcasting is realized through the French high performance National Network for Technology, Education and Research: RENATER. TÉLÉSIA project is focused on distributed multimedia applications, designed and experimented for collaborative work. Tele-seminar is a pilot activity involving several sites and their users. The number of experiments on high speed networks increase started to a few months ago. A lot of them remain in the local network area and are basic network technology oriented. TÉLÉSIA performs "in-the-field" tests as an effective way to evaluate some technical aspects of computerized distributed activities, such as network's architecture and performances, software functions, user interface, service specification and evaluation.

TÉLÉSIA experiment showed the limits of the networks in terms of loss of packets, protocol policies (multicast through Internet), bandwidth and throughput. To evaluate the quality of the tele-seminar service, it is necessary to integrate several parameters depending on the features required by the end user. Without giving an exhaustive list, video frame rate, sound quality, interactivity ratio (transit delay, number of participants) have to be taken into account.

Networks protocol type and features are chosen to cope with the Quality of service. Therefore, the multimedia streams have to be processed accordingly with the transport requirement. For example, in ATM the size of the cells does not fit with the current data coding and packetization which has to be modified. Therefore, the high speed networking introduces enhancements in transit delay and response time, which modifies drastically the end user perception and allows to define new adaptable criteria for QOS negotiation between high and low level in application architecture.

TÉLÉSIA tele-seminar allowed various experiments through ATM networks: National HOST conference and Proposers day (Madrid Oct.' 1994), INRIA sites interconnection and video-conferencing experiment and use (Sept.' - Dec.' 94). This paper present the knowledge gained during these experiments, the impact on application management and also data networking (audio and video packetization). The benefits and drawbacks for ATM network will be explored.

Key-words: research reports, TELESIA, cooperative work, visio-conference, parallel architecture, distributed systems, network

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : (33) 39 63 55 11 - Télécopie : (33) 1 39 63 51 14

Pierre.Leonard@inria.fr *http://magoo.inria.fr/pl/cv.html*
Alain.Caristan@inria.fr
delamot@wagner.inria.fr IFATEC
Andrzej.Wozniak@inria.fr

(Résumé : tsvp)

Release date: 1 - December - 1995

1 Introduction

On all times, man was preoccupied by the access to knowledge and its transfer. The evolution of the technologies for the treatment, conveying and stocking of the information leads new ways for computerized applications. Means of communication, outputs and wavebands of which regularly increase, allow to imagine new kinds of professional, social or private activities leaning on the services offered by the information highways. The most obvious seem to be those linked to the transfer of knowledge. Tele-teaching finds there new spaces to explore and new forms to imagine and experiment.

The transfer of knowledge relies on the learning of a teacher or a lecturer who delivers his informations through natural ways: voice and image. The use of a network widens the area of exploration of an old activity, even if it is not really dominated, technology as well as practical experience. That explains why several experiments of tele-teaching are developed using different multimedia tools but the evaluation of the networks, the machines and the methods is not over.

The Telesia project is concentrated on multimedia through networks with, as a purpose, the evaluation -through experimentation- of the possibilities and the restraints of the technologies of the information applied on tele-activities. Telesia is a service of tele-seminar firstly created to broadcast the Aristote seminars. It is a basic tool for the development of technological bricks, their putting in situation and the validation of concepts linked to tele-activities through Telematics networks. Since 1992, Telesia contributed to the development of Teleconferencing through Internet, to the organization of the french network Multicast and the experimentation of interactive tele-seminars. The present paper gives the context of tele-teaching experimentation, the main statements and puts forward some technical suggestions to increase the objective and subjective quality of the tele-seminar.

1.1 Project’s context and purposes

1.1.1. A general description of both the application and the experimentation

An Aristote seminar is divided into two separate actions:

1. In one hand, it is a standard seminar. The lecturer addresses the participants meeting in premises. The rules are simple and usual: the interventions are controlled by the lecturer, the president or the session regulator if ever there are several lecturers.

2. In the other hand, the seminar is transmitted through Internet from an emitting workstation to Internet-connected Workstation.

For this are used the Telesia software including a video codec H261, an audio codec and a seminar supervision unit.

The supervision offers several services:

- the creation of a regulator (1 per conference),
- the creation of participants (unlimited),
- the control of both the audio and video flows in emission as well as in reception,
- the management of the enslavement of the remote stations receiving the order to commutate the sources to be decoded,
- the management of the right to speak from the regulator’s station,
- the possibility to speak over the stations taking part.

The messages are conveyed through Multicast network. A tunnel links, on that occasion, the Ecole Polytechnique and french Multicast’s root conveyer (set in E. D. F.). The sites taking part in the seminar are notified a few days before by the session directory (S. D.). Several sites, accepting Multicast and connected on RENATER play a part in a seminar or a lesson relayed by Telesia. The users give their impressions relative to the quality of the sound, of the image and of the exchanges between the regulator and them. In this case, it is about having a better understanding of the control of the activity, of the follow through of the lesson and, of course, knowing the level of reliability of the application. Insofar as the lessons drawn can induce modifications, either on the Telesia coding or on the conditions of running (environment, procedures of working), the experimental conditions evolve with time. Different users population, on various sites, are requested to value in a better way the needs of the users and the variations of the performance on the network

This paper attempts particularly to evaluate the conditions of running of the means of communication and their repercussions on the Telesia application and on the tele-teaching activity. It states the differences between the experimentation on local networks (L. A. N.) and on wide networks (W. A. N.)

2 Results from current experiments

2.1 Interactive multimedia application and networks

Tele-presence quality plays a major part in knowledge's spreading. It influences directly the participate's attention capacity for a cooperative activity. For tele-teaching, it's something necessary to have an accurate synchronization between the audio and the video flows or between the sites, to reproduce an event. The human reaction time, 300ms, doesn't leave much room for manoeuvre.

Interactive multimedia application create high flows of data needing to be restored at their best. The problems caused by the conveying of these flows lead us to control the production of the multimedia information and to evaluate the consequences of the network problems of the application's performances. This evaluation leans on the objective measurement of the quality (rate of packet loss, transit delay, desequencing). The activity is more or less easy, depending of the quality of restitution. That is why it is also necessary to make a subjective evaluation of the quality (audio and video comfort, mutual comprehension between the users, tele-presence) and to evaluate the influence of the characteristics of running of the network.

Telesia tele-teaching experiment occurs in various contexts and through different kinds of networks, that is why it allows these quality observations.

Exploration results are different whether it is a local or a wide network. Evaluation criteria are the same but evaluation modes and networks management technic do not have the same effects on the performance of the conveying service. The high flow notion is relative. It seems more accurate to take about gradient. It is to analyse, when possible, the variation of the quality depending for the throughput available on the network.

2.2 Utilized networks

Most is multicast experiment.

2.2.1. The local networks

They are mostly Ethernet broadcasting networks at 10 Mbits or FDDI at 100Mbits. The setting of the native ATM is still experimental, that is why it is impossible for us to communicate observations or results.

A local network is intrinsically a structure easy to control, administratively talking as well as technically talking. Its availability and the throughput can be optimised by playing with the topology, the filtering or the packet routing and finally the control of the traffics.

The principle of locality induces short transit delays whose average value do not fluctuate a lot for the distances to be covered by the signal remain short compared to the delays of spreading of the signal.

2.2.2. The wide networks

Several kinds of networks were used.

2.2.2.1. RENATER

- RENATER means National Network for Technology, Education and Research. It federates regional plates offering access from 64 Kbits to 34 Mbits to interested sites. The most usual is a 2 Mbits connection. The EDF/DER, that owns a 34 Mbits access, shelters the root conveyer of the french multicast network (FMbone).
- 2 Mbits access: The Ecole Polytechnique, where the seminars broadcast from, owns a 2 Mbits access. A tunnel is established with the national conveyer. The 2 Mbits flow is sufficient for Telesia if the traffic to or from the Ecole Polytechnique is not excessive. This experimentation is the most regular through high flow french Internet.
- 64 Kbits access: in May 1995, a video course was realized between the ENPC -Ecole Nationale des Ponts et Chaussées- and the INRIA. The interest was the limitation of the throughput due to the 64 Kbits access of the ENPC. It was necessary to push the video codec within its limits (compression of the image flow to less than 32 Kbits, black and white, with a maximum of images per second), to use a much compressed sound (13 Kbits GSM) and to choose an audio packetization limiting the loss effect or packet arrival delays. This experiment permitted firstly to evaluate the Telesia possibility low flows and secondly to show the pupils the importance of the networks throughput for the realization of distributed multimedia services.

2.2.2.2. Transrel ATM

- The Transrel ATM experiment was the connection of three INRIA sites (Rocquencourt, Grenoble and Sophia Antipolis) through a Transrel link bearing ATM. The links were set to 13.5 Mbits. The services were CBDS on ATM.
- Most multimedia applications were tele-teaching or tele-formation tools. TELESIA/IVS was also used, through this connection, for the meetings of the group of evaluation of the Transrel-ATM experiment. An electronic white board, Liveboard by Xerox, was tested for split conferences.
- Using TELESIA/IVS, two improvements were noticed: in one hand, there is a better synchronization between the audio and video flows. In the other hand, a bigger synchronization between the sites (pseudo isochronism). This subjective betterment of the transit delay and to the increasing of the throughput available.
- It was not possible to measure the packet losses and the error rates according to the load of the ATM Network. The average transit delays are 20 ms for 64Ko packets, if the network is not very loaded. If ever the global load increases, the delay may be multiplied by five and the standard deviation increases.

2.2.2.3. International ATM

- During the National Host conference, an ATM link between Madrid and Lannion was set to show that an International high flow network was possible. TELESIA had a 10 Mbits circuit. An other experiment (Picturetel) was simultaneously set on the same link but with other circuits. It was possible to see the differences of working of the applications using two different kinds of IP conveying: the UDP conveying used by TELESIA and the TCP conveying used by Picturetel.
- The conditions of running are equal to the Transrel-ATM connection. Even if it can't be confirmed, the experiment was too brief, the UDP using application seems to resist better the transmissions problems. It can be explored, considering the TCP protocol is more affected by

transmission errors, mainly in chains, by the breaking of a connection management, i.e., to the application level.

2.2.2.4. RNIS

- In the field of wide networks, RNIS network is used for national or international show as a specialized link between local networks. These links are set using RNIS/IP interface boxes. The number of RNIS channels, depending of the kind of box (allowing from 128 to 668 Kbits), varies from 2 to 6. Such connections were set for national or international Aristote seminars (interactive seminars between Montreal (Canada) or Bonn (Germany) and the Ecole Polytechnique in Palaiseau.
- The quality of the link is generally good, the difficulties coming from the differences of throughput between the RNIS link and the connected networks. The variation of the flow lead us to the low flow Renater access case, mainly during the use of the multicast to Canada and before the multicast addresses were systematically filtered in the network conveyers supporting the connection boxes.

2.3 Comments

Despite the variety of networks met, it is obvious that there are constants in the way of working. Our main interest was the consequences of the problems encountered on the subjective quality. That why, after having presented the main observations, we will discuss the main elements for a subjective and objective analysis, for instance the packet losses, desequencing, transit delays. It will lead us to propose modifications to the basic technology, to avoid this sort of problems.

3 Main assessment of the experimentation

3.1 Results

3.1.1. RNIS

It is the weaker element of our experiments. That is why it plays a major part in the evaluation of the subjective qualification in touch with the throughput. In the case of a wide network, the RNIS is a betterment leading to local congestions in the supplying conveyers (see multicast).

To obtain a good compromise between the cost of the experiment and the throughput available, the usual standard used is 256 Kbits. In the case of such flows, it is a good definition. We have to find the good compromise between the quality image, the animation and the throughput available.

In term of performances only, there are two kinds of problems: the business of the link due to the Internet traffic and the business due to the tele-teaching application. In the first case, protocols appear because of the multicast traffic, that is still difficult to filter or because there is another traffic between the two networks connected. The multicast traffic can be partly controlled by prunning technics but these must be no other conference on the RNIS section or not too many floods open for the same conference.

In the second case, the exceeding of the authorized band can be explained by the impossibility to set an accurate value for the video flow, contrary to the audio flow, the video flow depending on the kind of images sent. There is few coded information in the case of a scene showing a rather static person (H261 coding, differential mode). For scenes very animate, the changes are more important in every shot, which induces an increasing of the fluidity, it is obliged to work at the limit of the threshold of business of the link, that becomes very sensitive to the variation of the data flow.

The transit delay is constant for one RNIS section. The variation come from the coupling equipment (local network <=> RNIS). The constant delay depends on the distance of communication. Tele-presence can be deteriorated on high interaction phases if the delay is too important: RNIS is often used for tele-teaching. It allows reliable broadcasting offering a satisfactory comfort, but comprehension is limited when the image is used as a mean of information. In the current day state of the treatments of the multimedia information (audio and video Codec limited in compression and in image rate), under 256 Kbits, its use is delicate for real-time multi-point tele-teaching experiments. Nevertheless, its adaptability makes of it a good way of entering the area of IP networks.

3.1.2. Local networks

As previously said, the main feature of the video flow is the variation of amplitude at constant sample rate. It might be the source of problems on a local network whose sections are different in terms of throughput. Either their throughput are theoretically different or they are saturated with the traffic they support. The consequences are the same as for the RNIS.

A current management technic is to separate the sections conveying the multimedia traffic. That way, the results are satisfactory and, depending on the real load, it is possible to increase the out flow of the application to increase the quality of the image and its fluidity.

It is not noticeable that the limitations of the multimedia flow are more linked to the hardware than to the local network: you need specific hardware to create flows able to saturate the network. If, of course, the number of channels simultaneously open is strictly controlled.

Anyway, the consequences of such problems or choices is that the loss of packets is more or less important according to the rate of congestion. We are attached then to work in reasonable conditions and the technics of compensation of the errors are more securities than absolute needs. The transit delays observed, since they are weak and the rate of dispersal (compared to the human time of reaction) is weak too, remain in limit allowing an acceptable synchronization of the users on the cooperative activity. Even in the case of lateness due to local phenomena of congestion, the case of an unbearable delay must firstly throw back into question the network before an intervention on the application.

Desequencing are seldom observed on local networks, because of their mode of managing. In the opposite case, we are lead to the wide network case.

3.1.3. Wide networks

A light loaded wide network is comparable to a local network, that is to say there is no problem of congestion.

The values of the significant parameters are, on the contrary, totally different and the cause of problems also. Renater is an heterogeneous network. It is very loaded on certain points. That is why the reception of a single piece of information is not the same on different point.

Transit delays, because of the links, commutation problems in the nodes, control of the duration of the life of the packets make the losses common and then we have to protect the flows from them.

Local congestions and the adaptive conveying include breaking off in the multimedia flows.

3.1.3.1. Transrel-ATM link, NH Madrid - Lannion link

This kind of link can be considered as the high flow counterpart of the RNIS.

This service of virtual way acts like a booking adjustable flow specialized link. This kind of link bring us back to the RNIS ease of connection, to the local network for the topology and to the wide network for the management. Indeed local end network are coupled on a node that is currently a sub-network with its own conveying and its own management.

Losses exist and are du to conveying flows in the Transrel-ATM network itself. Several causes are known:

- loss of an ATM cell inducing the loss of a UDP packet, part of it being included in the ATM cell. The problem of the protocols on ATM is that TCP is not effective, compared to the throughput allocated. The question remain about the adaptation of the conveying protocols reliable on ATM that do not slow down the flow of information. This problem is discussed,
- congestions in commutation switch inducing losses of packets in commutation switch's queue.

For TELESIA tele-teaching, this configuration does not bring new problems and the statements are the same as for the wide networks.

The interest of ATM as a long distance packet commutation technology leans, with the obvious gain of throughput, on the stability of the transit delay and on its weak dispersal in normal conditions of use. That is why the video conferences experiments worked well, once the connections mastered.

3.2 Objective and subjective criteria analysis

3.2.1. Loss of packets

None of the network used are 100% reliable. The rate of losses is more or less important, according to the congestion, the conveying problems or the protocol mistakes. The throughput used is always at the limit of the acceptable load on the network. This for the low flow networks such as RNIS, for parameters need to be adapted to fit the possible limits, as well as high flow networks since they are dedicated to a single communication at once. The consequences of the losses on the different channels (audio, video and supervision) are of different natures.

3.2.2. Image sequences

The description of an image is distributed among several packets. The loss of packets leads to the alteration of this description as well as a loss of information on the flow if the cut is bad. It is more or less harmful to the quality of restitution. the first H261 packetizations did not take care of this phenomenon. Fourth chapter's proposals describe solutions to this problem.

When the problem of the correlation of the packets of a video sequence is settled, the restitution of an image can tolerate the loss of some of its blocks. Its depends on the rate of renewal of the image directly to the image rate.

High flows, by furthering the increase of the image rate, by the way the renewal of the information, make the loss of packets less sizeable. We enter the area of subjective quality for the visual impression depends, from a given threshold, on every individual. The purpose of the technical developments is to obtain an acceptable level of subjective quality and then to ameliorate it at each individual need. The first works had an influence on basic methods for the treatment and conveying of the multimedia information. Thus the standards. The second open the way to proprietary considerations.

Finally, as the usual acceptable standard is 25 to 30 images per second, which set a boundary to the quality of information to transmit, the evolution to high flows must allow to gain space and time to make the broadcasting protocols reliable (redundancy, re-emitting on request) aimed to overcome the loss of packets.

3.2.3. The audio

The audio coding used does not create any determinism. The loss of a packet of sound has an immediate effect on the quality of restitution. A hole is created, which is sometimes unpleasant.

From a subjective point of view, this hole has consequences more or less harmful to the understanding according to the nature of the sound or word it touches. The human brain can, sometimes overcome the lack of information with the natural redundancy existing in the subject. Though, talking of tone or troubles incompatible with a good attention.

In other way, compression methods too drastic cause an alteration of the audition comfort. The interest would reside in the decreasing of the flow, what could allow the redundancy of a packet of sounds. High flow allow the picture sound treatments giving a better audition quality as well as more reliable protocols, from the beginning to the end, in an interval of time compatible with real-time acquiring and restitution. Through well dimensioned local networks and in normal load conditions, the quality of sound is satisfactory. Through wide networks, the problem is more delicate for the mastering of the sections is not equal. But on a well mastered wide network, it is possible to realize good transmissions. On open Internet part of the solution is the evolution to high flows.

3.3 Transit delays, desenquencing

Desequencing are the consequence of the conveying politics. Through local networks, they are non-existent excepted particular cases bringing back to wide networks context. This phenomenon has few effect on the audio and video return, as long as it does not come with a sensitive increase of the transit delay of a desequenced packet. It leads to a more important management of the multimedia flows, in terms of application protocol, if the resequencing is not realized at the network's level. Time is a critical restraint in such real-time applications. A desequencing too important can equal a loss of packet, since the queuing packets whose lifetime is over are abandoned.

The increase of the flow inducing shorter time delays, if it is not sort out the desequencing problems, allows to manage period at the user's. To play again the information received, an acceptable delay from the applications and the user's point of view can be set. The decrease of the conveying delay increase the waiting delay at the user's while not modifying this comfort.

4 Interface, interaction, protocol

4.1 Reordering the packets

The packet losses due to the failure of the routing activity were significant but rapidly fixed. The rest of the lost traffic differs from hours to hours during the day of the seminar. But we decided not to waste time with this problem at the IP protocol level. The real-time nature of activity implies a different approach.

There is not a lot of off-sequence packets. It depends mainly on the packet's route through the network. The average value for off-sequence packets is 2. In case of no reordering policy, the impact is clear on audio and video rendering:

- it's more difficult to listen to the audio and understand what the speaker says. To simplify: a word **tatoti** could be heard: **tatito**,
- the video H261 decoder lost the synchronization with the stream because some information depends on the packet sequence (according to the current packetization technic [1]).

To avoid this, the first action was to add a reordering feature to the RTP layer. We use the sequence number of the RTP header to reorder the packets received. The main features are the following:

- packets are stored in a window of variable size,
- the size of the window is either fixed to 0, meaning no reordering, or to any other value,
- the size of the window is automatically computed by analysing the arriving sequence numbers.

By the way, the RTP layers also provides statistics on the packet's stream:

- the number of off-sequence packets,
- the number of off-sequence sorted by lateness: 1 2 3 4...
- the number of lost packets,
- the number of lost packets sorted by the number of contiguous lost packets: 1 2 3 4...
- the number of rejected packets: they are the packets arrived too late because of the current size of the window,
- number of doubloons of packets.

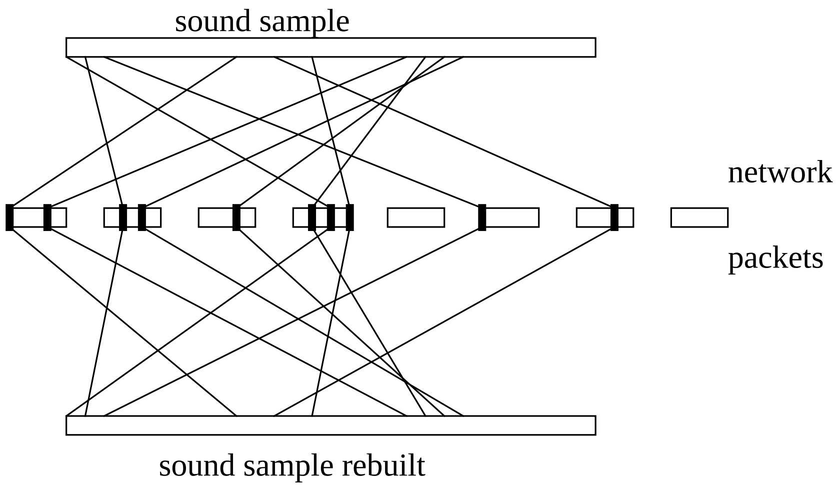
The use of the reordering policy implies that the application should be profiled in such way that this profile should determine the conditions of use, depending on the nature of the data:

- the control channel of the application, due to the low bandwidth available and the contextual independence of the packet, is set to 0: no reordering,
- for audio and video, the size of the window is initialized to 2,
- due to the use of either the manual control or the automatic detection of audio, we have designed a context re-initialization of the number of the packets sequence. The re-initialization, detected at the arrival, allows the window to be flushed and the sound to be completely played,
- to confirm the re-initialization of the number of the sequence, the packet is sent several times.

4.2 Re-designing the audio packetization

One of the more profitable track concerning the quality of the audio, is to solve the losses of sound sampled packets. The use of the conferences system on wide area network, does not allow to ask for another transmission of lost packets. The real time aspect of the application will not survive the delay implied by this protocol.

Then, we design a new way of building the packets that divides and distributes the sound samples into multiple network packets. This new packetization is called TEAP, standing for TÉlésia Audio Packetization mode, its description following:

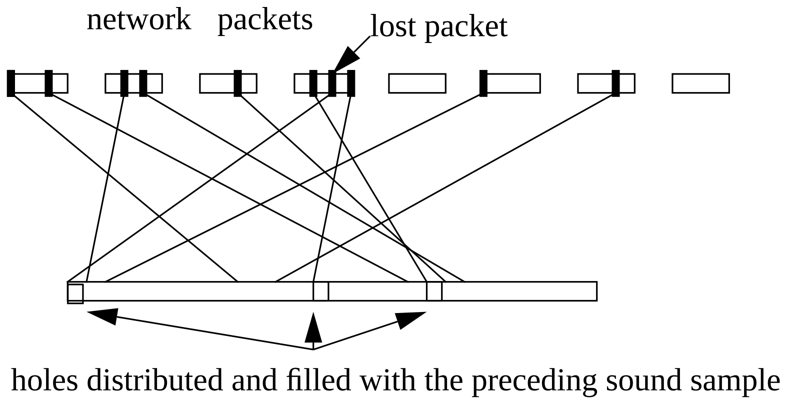


Picture 1: Sample distribution into network packets [TÉLÉSIA TEAP mode]

This building of the packets is not a simple division of a sample and a multiplication of the packets. The distribution also use a function based on sequences to compute the relations between the samples and the network packets. The main aim of the algorithm is to transmit in various packets the sound samples that are contiguous at the acquisition level.

The bad impact of the holes is then redistributed into the rebuilt sound sample. The holes are now very small, and to avoid a multi-breaking of the stream effect, the preceding sample is copied in the hole. The scale used for the experiment is the following:

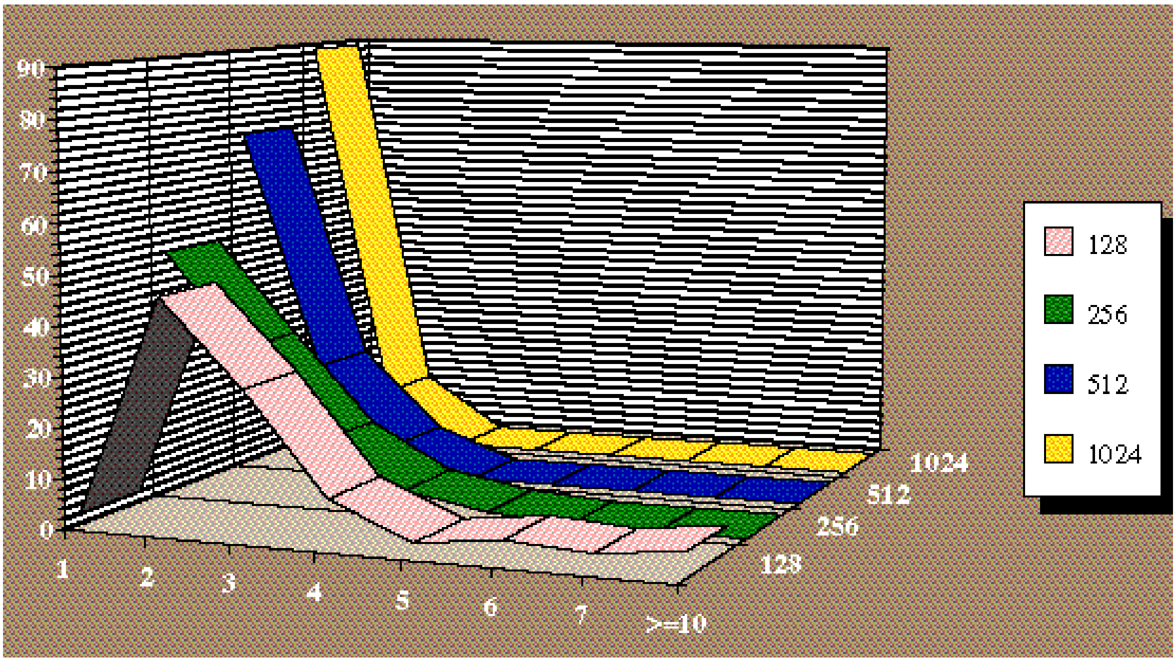
- size of the *original sample*: 1024 bytes, 1/8 second in ulaw mode,
- number of network packets per original sample: 8,
- size of the network packet, without RTP header: 128 bytes,
- number of *cell* per network packets: 8,
- size of the *cell*: 16 bytes, duration: 1.95 ms



Picture 2: Impact of a lost packet

4.2.1. Results analysis

The analysis of this new packetization design, compared with the old one, was done either objectively by computing the lost packets' statistics and subjectively by hearing the resulting sound. It became clear, that the conclusions lead us to enlarge the study by analysing the impact of the packet size on the lost packets' statistics. The measures campaigns were organized both during specific sessions and during in-field-tests with the Aristote seminars. The measure campaigns took place in January 1995 and April 1995. The tool used to extract these measures is integrated to the RTP layer and particularly the reordering part of our RTP implementation. An other influencing point is the size of the holes generated by the lost packets. Some parameters could be added to the functions (F1, F2) used to compute the sample distribution into the network packets. These parameters will affect the distribution by limiting dynamically the impact of the holes:



Picture 3: Distribution of the size of the holes. Ratio per packet

After these different experiments, it can conclude that the throughput is clearly a part of the network aspect. That's the most often criteria used for the evaluation quality of service. However, the size of the packets and, by implication, the number of packets commuted a second is important too. The routing system seems to be more sensible to that last factor. Since the size of the packet influences directly the size of the holes in data stream, in case of losses of sound samples, a compromise has to be found between the acceptable holes (several packets lost in a sequence), the size of the packets sent to the network, and the delay induced by the reordering.

We have proposed a way of coupling the subjective character of the sound rendering quality with an objective and pragmatic solution. Despite the relative simplicity of the distribution algorithm, we think that a window is opened to another approach of the sound packetization. For example, this approach can be used with compression algorithms more adapted with the sound structure (psycho-acoustic analysis MPEG/AUDIO) and/or more robust in case of loss of packets (some predictive algorithms, for instance). It could be interesting to adapt dynamically the packetization (size of packets and samples dispersion) to the network quality of service. Due to the specificity of the native ATM network, this sampling and packetization method needs to be tested on the ATM cells framework.

4.3 Re-designing the H261 packetization

Subjectively, between a frame frizzed for a while and a frame laid out like a draughtboard, we could say that the most bothering point is the second one, even if the end user gives the frame rate as a quality criteria. In one hand, with the current H261 coding, the frame rate is unsettled, depending on the changes between two frames. The user integrates partially that fact. Thus, if the frame is completely frozen, the trouble is not really disturbing for a few moment. On the other hand, if the frame is distorted, that can divert the user's attention.

There are two explanations for that disturbance: off-sequence packets and lost packets. For a better and clearer understanding we describe the H261 structure of the current packetization. The H261 document from CCITT[5] describes four main objects:

- **picture**: starting pattern Picture Start Code, *PSC*,
- **Group Of Block**: starting pattern Group of Block Start Code, *GBSC*,
- **Macro Block**: starting pattern Macro Block Address, *MBA*,
- **Block**: ending pattern End Of Block, *EOB*.

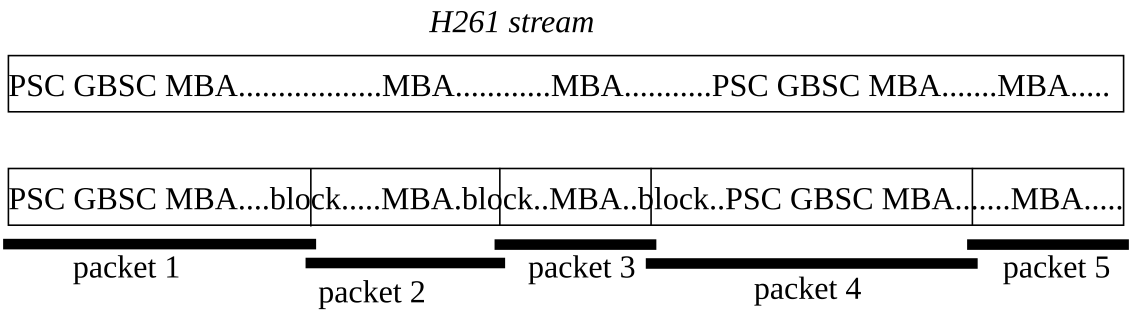
The H261 packetization describes the packet format in a specific header. The relations between the H261 stream and the IP packets are given in the header option fields:

- the S bit: the packet includes the beginning of the encoding of a GOB,
- the E bit: the packet contains the end of a GOB.

There are also some choices weakening the protocol resistance to the transport failure:

- a packet can contain several GOB,

The stream is cut arbitrarily at the end of the IP packet.



Picture 4: Current packetization

In this sketch, the Blocks are very often split in two packets.

4.3.1. Discussion.

Firstly, in the H261 stream structure, the EOB determines either an End Of Block or an End Of Macro Block or an End Of Group of Block, or an End Of Picture. The decoder needs to read ahead to decide.

Secondly, to cut the H261 stream in the middle of an H261 object causes a de-synchronization. After a loss of a packet, the decoder has no information to link the new data with the remaining part of the stream.

Thirdly, the H261 stream Block address is relative to the MBA. Thus, a packet beginning with a MB could be interpreted at the current address of the frame, just after the last decoded Block, without preserving the original part of the picture.

4.3.2. Improving the H261 packetization

A part of the proposition is to reduce the number of packets sent. That can be achieved by reducing the throughput of the H261 stream generated by the coder: with a better quality of the H261 processing (quantization level or changing detection threshold). But in the same time the frame rate increases, thanks to a more efficient processing of the image. It does not solve the packet loss side effect.

Therefore we concentrated our work on the packetization to limit the sensitivity of the decoder to the lost packets alone. The current H261 packetization is context free. One solution to solve the points 1 and 3 is to complete the H261 header with:

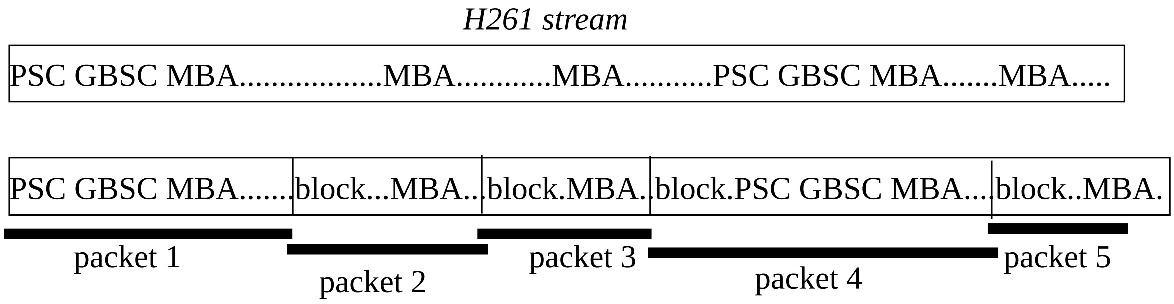
- the **GOB number**, number of the Group of Block beginning the packet,
- the **MB number**, number of the Macro Block beginning the packet,
- the **last MB number**, number of the last Macro Block decoded,
- the **type** of the last object in packet: Picture, Group Of Block, Macro Block.

Now, the decoder interprets the H261 part of the stream with its context. In the case of a packet loss, the decoder knows how to process a Picture, a GOB, or a MB changing. The decisions, can be to display or not the current frame. Another one is to salvage the decoder context to restart with a coherent internal state (consistent sequence, entire blocks...).

With the added context information, we change the cutting policy of the stream to be able to:

- always start an IP packet with an H261 object,
- have a maximum of flexibility to tune the size of the packet.

The finest granularity copes with the smaller H261 object: the Block. Accordingly, an H261 Block is never cut between two packets and is always wholly decoded in one pass:



Picture 5: New packetization

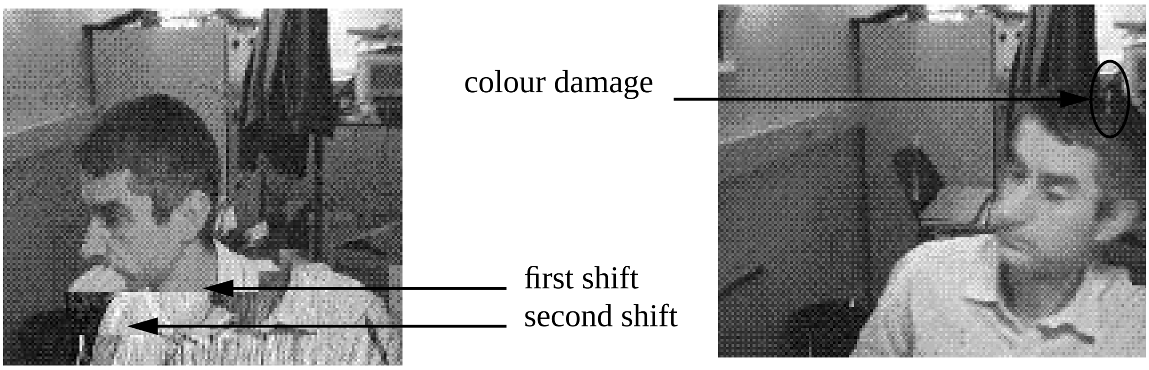
The new H261 coder generates a H261 Object table containing for each Object the address in the stream and its type. To process this information, a new H261 packetization layer has been added. It analyses the Object table and decides how to cut the stream in order to generate the best filled IP packet.

On the decoder side, the packetization layer analyses the H261 header and gives to the decoder the pattern corresponding to the beginning or the end of a Picture, a Group Of Block, a Macro Block, a Block and the related number.

4.3.3. Results analysis

4.3.3.1. Subjectivity

The subjective quality of the images is determined by a part of the damages that have disappeared and another part that still remains. The main disturbing aspect of the image damages is the shift to the left of a part of the image due to the lost packets. With the new header and the absolute addresses inside it, the stream could always be interpreted correctly:



Picture 6: Corrected Video damages

That damage disappeared completely and the video acquire now a better stability.

Still remain the colour changes implied by the losses of the blocks coded in INTER mode. As the video information are coded in differential, each lost packets is a lost display information for the blocks. That damage is represented by a halo around the objects moved in the image. The only answer is to send the frame in INTRA mode as often as possible (i.e. vick). In a well controlled network these damages are not frequent enough to imply a specific treatment by the application.

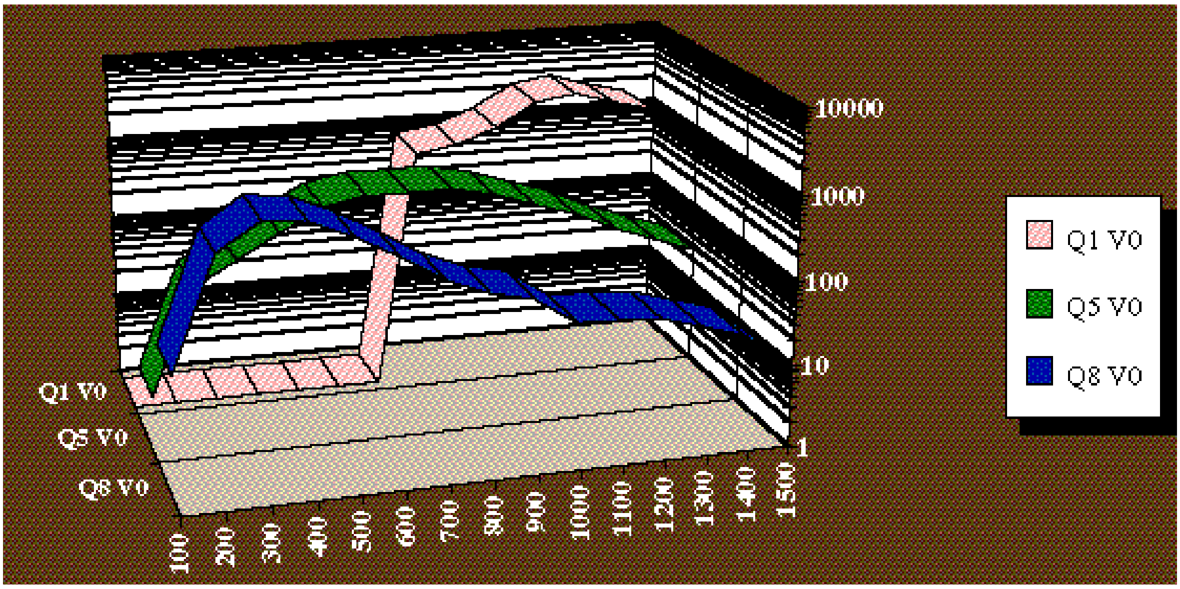
4.3.3.2. Objectively

The objective quality: The revision of the coder architecture and some changes in the DCT improve the efficiency by a rate of 15 to 20 percent. Moreover, the number of packets also increases, in a more important ratio. We previously saw the effect of the packet number throughput on the routing system. The reason for the increase of the packet number is the choice of the block granularity for the cutting out of the packets. In fact, the block frontier does never fit exactly with the maximum size of the RTP packet, which implies a lower ratio of filled packets.

Example:

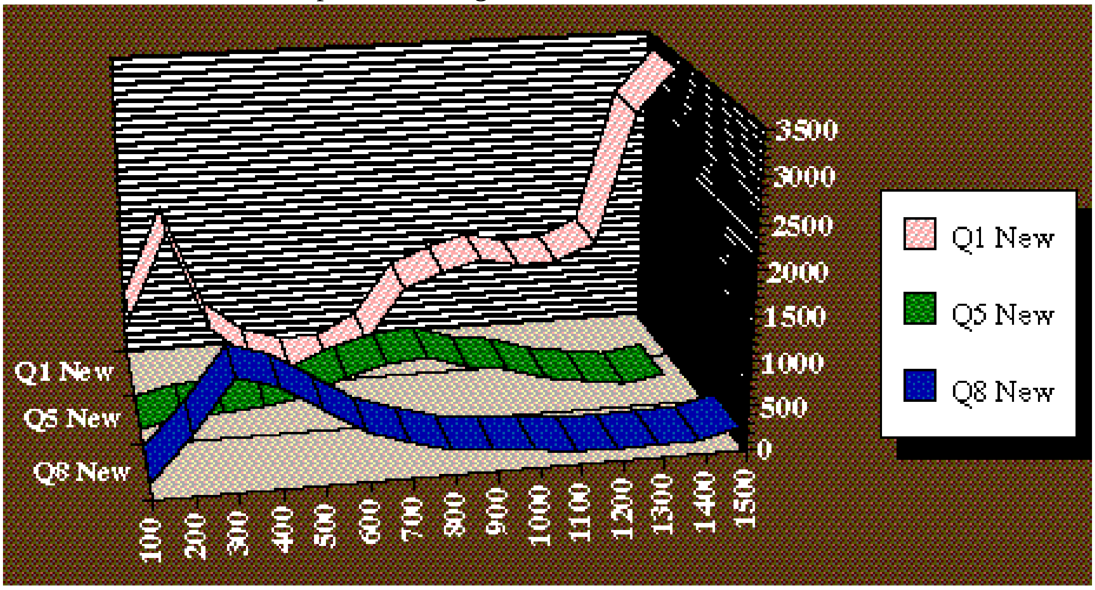
One of the tests was done with a conference[20] on tape. The duration is 3 minutes. The background is fixed, only the speaker moves. The video contains in the same proportion speaker talking and graphics slides presentation.

Two graphics on the packets' size distribution are following. We notice that the differences between the two packetizations are more sensible when the H261 throughput is important, either with a small quantization number or an animated video. On the contrary, the bad effect of the spreading of the packet's size is no longer critical when you use a 5 or 8 quantization factor, which are the values currently used for the broadcasting of seminars.



Picture 7: Old packetization: packet's size distribution.

The horizontal scale designs the size of the packets in bytes and grows from left to right. The vertical scale is the number of packets during the session.



Picture 8: New packetization: packet's size distribution. Linear scale

The bad effect of the smaller packets is largely balanced by the robustness capacity acquired at the decoding part.

More technical details could be found in the following paper:

http://magoo.inria.fr/pl/telesia-pap2.html

5 Evolution

To enhance the global teleteaching service, there are two axes. The first is to improve the content and the way to deliver it. Not only by given the right information but also by performing very well session. The quality of the teaching performance is at least as important as the technical environment quality. Indeed, we observed that, in case of very attractive presentation, the people is so captivated that it doesn't care about the technical conditions. Obviously if they are sufficiently acceptable, And here we have to introduce the second axe, the technical matter.

In this paper we focused on correction of standard networking incidents. There is a lot of developments to control at the same time networking, workstation configuration, software functionalities integration, user interface and distributed activity administration.

Before to introduce advanced technologies in this domain, it has been necessary to well manage the basic functions of such activity.

5.1 Network mastering

As said above, in a controlled network, the congestion can be avoided to reduce the packets loss rate.The mastering of French Multicast Bone is organised since july 93. One of the major aims is to balance the multicast traffic between routers depending on their switching capabilities.

The difficulties encountered to easily obtain a good subjective quality of video-conference, comes partially from a lack of good evaluation technics and criteria. Contrary to the computer performances evaluation, for whom a lot of technics based on benchmarks are available, the performance of networked multimedia depends on image quality, image frame rate, synchronization. For example, he frame rate varies with the CODEC performances, image types an size, sequence stability, network bandwidth. For the moment this parameters are modified independently of each other.

This process doesn't converge to the best compromise, because there is no well adapted tools. A good example comes form the joystick driving method. Using the two control wheels (horizontal movement and vertical movement) separately to drive the cursor, you will not get a good control of the trajectory. Using the stick, allows accurate movement. We think that the same thing could be done for video-conferencing parameters control. This problem is not mundane, due to the number of criteria. However, one can imagine to develop a method which allows the end user to easily move a pointer into a multidimensional space. Several times some parameters are fixed by the context like network maximum bandwidth, maximum frame rate and transit delay. That gives the limit conditions for tuning functions and reduces the number of degrees of liberty. The best way is to get these values dynamically. Thus, one can imagine to give end user or conference manager a better understanding on what he does.

5.2 Mastering Audio and video processing

Hardware peripherals are becoming more performing. The audio chips offer a variety of sampling methods. The video capture board can now store 25/30 frames/s on disk using DMA. The remaining problem concern the digitized data format. For the moment, the standard compression technics are not implemented in such manner that they can be used for advanced CODEC. Indeed, the same technics (FFT, quantization, DCT,...) are used in all CODEC on nxn bits array. It would be useful to have API allowing easy inter-operability between hardware functions and software.

For a good teleteaching, it is recognized that the automated camera-man improve the comfort. Another solution is the remote control of audio and video peripherals. But the price of well adapted equipment constitute an obstacle to the implementation of software advanced control functions.

5.3 Mastering the seminar

The seminar environment must be organised to take into account the necessary announce, preparation and production of the tele-conference. For the moment, there is no specific action to announce the seminar, register the participants, except SD registration.

The seminar administration concerns three phases:

The seminar installation in terms of advertising, addresses allocation, technical environment preparation. We investigated the X500 directory approach to store all relevant informations about the seminar before, during and after the event.

The real-time seminar management. The opening is under moderator control who initializes the contexts in terms of image to decode, audio channel to play. This is a real distributed activity we are usually calling distributed collaborative administration. Indeed, the information must be available for each partner and they have to exchange information control and statistical reports. This activity requires a good response time end therefore the maximum reliability in communication systems. Particularly, when the seminar is broadcasted, because it is not easy to check all participants individually. In 1994, a model of administration based on SNMP protocol approach has been designed and prototyped. A dedicated agent runs near each instance of the application and is able to report on in and outgoing events. Objective quality in terms of loss packets, frame rate, local end user parameters can be fetch by every member of the distributed group.

The seminar closing doesn’t introduce specific action, but must be well managed to close the service and keep available all informations that could be useful for further activity.

The seminar mastering brings some new requirements for networking services. In this case, the challenge is less to transport a great amount of data than allow good synchronisation capabilities.

6 Conclusions.

This paper intended to introduce the problems encountered in teleteaching development in the context computers networks. The new challenge is the emergency of broadband networks. However, a very high speed network doesn’t constitute an objective by itself. That depends on what it is expected to offer the teacher and also the student. Because of computerised information processing and networking, it’s necessary to develop specific technics to configure well adapted network allowing at the same time, the best subjective quality.

The failures detection and avoidance of them by the technical proposals presented here are being tested in terms of compliance with the real audio and video control process needs.

We think, that taxonomy technics can be helpful to classify the signal processing technics and build reference tables:

- for parameters tuning during the teleteaching activity,
- for network loading estimation in anticipation of a given activity,
- to prevent some unexpected event during teleteaching activity.,
- to determine the optimal parameters in terms of number of participants, network bandwidth reservation, transit delay adjustment.

The used bandwidth doesn’t depend not only on the audio and video. More than this is necessary to manage the teleteaching, more information has to be generate. For example, using the web client server protocol can improve drastically the used bandwidth. The difference resides in the fact that the exchanges are less sensible to the synchronisation. They also don’t represent continuous streams and can be treated asynchronously.

We have explained above that it is necessary to reach a minimum level of quality to pretend realize a remote activity and to draw the remote participant's attention to the teacher. That explains the effort for a better rendering of audio and video even if the network doesn't work very well. But clearly the challenge is not only on a technological domain.

We have observed, in the case of very attractive or amazing presentation, that the people could be very captivated even if the subjective quality is bad.

Thus, the content seems to be more important than the container above certain limits. In this case, the technology, such as high speed network, can improve the comfort of the end user. That introduce also some new problems for the teacher who have to cope with sophisticated or not easily manageable environment. For the moment, as long as the networking is not driven in a very user-friendly manner, it remains an hindrance to the progression of the tele-teaching comfort and efficiency.

7 Bibliography

- [1] Packetization of H.261, Thierry Turletti, Christian Huitéma, September 9 1994, draft-ietf-avt-video-packet-02,
- [2] Thierry Turletti, H.261 software codec for video-conferencing over the Internet INRIA Research Report no 1834, January 1993.
- [3] Dimensions of Satisfaction in Teleconferencing: An exploratory Analysis, V. Srinivasan Rao, HICSS94,
- [4] The measurement of End User Computing Satisfaction, Doll, W.J. and G. Torkzadeh, MIS Quarterly, Vol 12, No 2, June 1988, p 259-274,
- [5] Video codec for audiovisual services AT p x 64 kbit/s, CCITT Recommendation H261
- [6] RTP : A Transport Protocol for Real-Time Application, draft-ietf-avt-rtp-06.txt, H. Shulzrinne, S. Casner, AT&T/ISI, November 1994
- [7] La communication entre individus au sein d'une visioconférence, Corinne Potiron, INRIA, Septembre 1993,
- [8] schéma Directeur des Moyens Informatiques de l'INRIA 1991 - 1994, INRIA, Novembre 1991,
- [9] Queuing Study of a 3-Priority Policy with Distinct Service Strategies, Randall Landry, Ioannis Stavrakakis, Member, IEEE, IEE/ACM Transactions October 1993,
- [10] PAGEIN, Pilot Applications in a Gigabit European Integrated Network, WP report, R2031/RUS/AWS/DS/L/007/b2, Ulrich Lang, Alain Caristan, Lei Wang, Andrea Rozek,...,
- [11] CHORUS, a new technology for Building UNIX Systems, François Armand and all, EUUG Autumn 88 Conference Proceedings, Cascais, Portugal.
- [12] The Cyclades Computer Network - Towards Layered Network Architecture, Louis Pouzin, Monograph Series of the ICCC, 2, Elsevier Publishing Company, Inc., New-York, NY, (1982), ISBN 0-444-86482-2.
- [13] Etude des mutations observées dans les entreprises, influence des technologies d'information, innovations techniques et modèles de production typographique, A.M. Laulan, A. Marshall, Journées COMICS 92, Juin 1992.
- [14] Bilan d'évaluation préliminaire du réseau expérimental de visio-communication médicale Béthune-Salengro, Christophe Deveaux, Centre Hospitalier Régional Universitaire de Lille, unité LOGINAT, Avril 1994.
- [15] Design for privacy in ubiquitous computing environments, Victoria Belloci and Abigail Sellen, Proceedings of the Third European Conference on Computer-Supported Cooperative Work, 13-17 September 1993 Milan Italy.
- [16] A spatial model of interaction in large virtual environments, Steven Bedford and Lennart Fahlén, Proceedings of the Third European Conference on Computer-Supported Cooperative Work, 13-17 September 1993 Milan Italy.
- [17] Telework: exploring the borderless office, Mary Elizabeth Watson Fritz, Kunihiko Higa, Sridhar Narasimhan, Proceedings of the twenty-Seventh Annual Hawaii International Conference on System Sciences, 1994.
- [18] PAGEIN, Pilot Applications in a Gigabit European Integrated Network, WP report, R2031/NLR/IR/DS/R//031-032/b1, Frédéric Buissez, Ulrich Lang, Alain Caristan, Lei Wang, Andreas Wierse, Sylvain Causse, Michel Grave, Harald Nebel, Bruno Troff.
- [19] Evaluation du service Transrel ATM à l'INRIA, Georges Aziza, Denis Joiret, Rapport interne INRIA Janvier 1995.
- [20] Bases de données multimédia réparties, LICEF, Université du Québec, Télé-Université.
- [21] AudioFile: A network - Transparent System for Distributed Audio Applications. Thomas M Levergood, Andrew C. Payne, James Gettys, G. Winfield, and Lawrence C. Stewart. Dec Cambridge Research Lab. June 11, 1993.
- [22] Basic Reference Model of Open Distributed Processing, ISO/IEC JTC1/SC21/Wg7, DEC

1991.

- [23] Object Management Architecture Guide, Revision 2.0, September 1, 1992, OMG TC Doc 92.11.1 (Contains OMG Object Model).
- [24] The Common Object Request Broker: Architecture and Specification, Revision 1.1, December 6, 1991, OMG TC Doc 91.12.1. Other OMG document: server@omg.org with “help” in Body.
- [25] X500 ISO/IEC 9594: 1993 (E)
- [26]MPEG: a Video Compression Standard for Multimedia. Didier Le Gall. Communications ACM (Special Issue on Multimedia), Vol 34, N4, April 1991, P 47-58.
- [27] MPEG DIS 11172.
- [28] A Tutorial on MPEG/Audio Co,pression. David Pan - Motorola - IEEE Multimedia Summer 1995.