INSTITUT NATIONAL DE RECHERCHJE EN INFORMATIQUE ET EN AUTOMATIQUE

Quality and Usage enhancement in TELESIA

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

N° 000

July 1995

Actions de Développement

Projet MÉDIA

Rapport de recherche

ISSN 0249-6399

Quality and Usage enhancement in TÉLÉSIA

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

Development

Projet MÉDIA

Rapport de recherche n˚000 - July 1995

33 pages

Abstract: During the last years, the TÉLÉSIA application, remote-seminar remote-meeting, was the base of a large number of experiments. These experimentations, in France and over the world, was the field tests for advanced multimedia services and evaluation technologies research and developments. The usages are rather different due to the user's needs for a remote-seminar or a remote-meeting. Know-how, expertise exchange is enrich with the tele-presence of the speaker: the listeners are free from the information decoding. The technology is crucial for the tele-presence, which is the main principle of the tele-seminars. Therefore, the audio, the video processing and networking needs to prevent the weakness of the low quality networks: off sequence packets, lost packets, variable transit delays.

This report presents an analysis of the major defects encountered during the experiments, the solutions designed and realized, to embrace the objective and subjective quality of audio and video processing. Finally we present the analyze of the results trough "in-the-field tests" and campaigns of measures done during December 1994 and May 1995.

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

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: 27 - September - 1995

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

L’amélioration qualitative et d’usage dans TÉLÉSIA

Résumé : L’application de télé-séminaire et télé-réunion TÉLÉSIA a servi ces dernières années de base de tests en France en Europe et dans le monde. Ces expériences ont été un champs d’investigation primordial pour les services multimédia avancés et d’évaluation de technologies de recherches et de développements. Le Télé-séminaire, la télé-réunion correspondent à des usages, des besoins différents. Le savoir faire, l’expertise échangé est enrichi grâce au principe de télé-présence qui libère l’auditeur d’un travail fantasmagorique sur son interlocuteur. Les aspects technologiques apparaissent ainsi cruciaux pour l’amélioration qualitative de la télé-présence. Le traitement du son, de la vidéo, du réseau ressortent comme les éléments permettant de limiter les défauts des réseaux actuels : dé-séquencement et pertes de paquets, délais de transit variables.

Ce rapport, du projet MÉDIA, analyse à la lumière des expériences les défauts majeurs rencontrés, présente les solutions conçues et testées pour l’amélioration qualitative de la vidéo et du son. L’analyse a été réalisée pendant des campagnes de mesures, grandeur nature, de Décembre 1994 et Mai 1995.

Mots-clé : rapport de recherche, TÉLÉSIA, télé-conférence, travail collaboratif, architectures parallèles, systèmes distribués, réseaux.

1 The context

For many years, computer networks have allowed data communication between distant sites. A large use of such technics is now effective in banking management systems, train or plane seats booking, electronic mail and files transfer. But, however, these applications do not allow audio and video communication. It has been necessary to wait for the early 90s to see new application software integrating such features. This new kind of software provides today a good user interface for CSCW (Computer Supported Cooperative Work).

Allowing people to share information and ideas through computer and data networks requires a new approach to data processing and man-machine interface [17]. Among the new required functionalities of such system, one is to give to the user the feeling of a real human communication. That is defined as tele-presence paradigm [3, 15, 16,].

Several domains have to be explored to develop and make usable some applications introducing a good approach to this kind of problems. Technical solutions already exist for data transport, co-operative and distributed processing, multimedia data processing, application management. Organizational matter concerns Internet and Multicast Bone activities, European and whole-world projects studying the CSCW solutions and the advanced services on Informations Highways [Race, ACTS, Telematics European programs].

Some of the works are more focused on basic technics in network protocols, user interface, data processing. Some others intend to gain experience in the development of advanced services by experimenting real distributed activities. Before today, Multimedia tools were developed to compensate for the lack of industrial solutions. Since 92, there are some well known prototypes carrying out local or world-wide experiments on Internet [Vat, Nevot, Nv, IVS, JVTOS, Télésia, CU-SeeMe]. The manufacturers also provide proprietary solutions. We do not forget the more classical approach of dedicated solutions as the vide-conferences systems based on ISDN and H261 codec. Their commercial development is limited by economical and technological matter (cost and specificity of material support).

In any case, it appears that the more difficult thing is to find a real field of interest for end-user avoiding him to be a passive partner for technical evaluation of the new information technologies. In other words, how to interest sufficiently the end-user to have a good feedback information on his own needs and how to make the technology transparent for a natural use of advanced distributed services.

TÉLÉSIA Project, started in 91 with the cooperation of Aristote association, has focused on distributed tele-seminar services development and evaluation. To have a real interest for the end-user, interactive seminars were regularly performed on the RENATER network.

The activity was shared between “in-the-field” experimentations and enhancements of TÉLÉSIA software to improve the functionalities and the service’s quality in technical and usage point of view.

In this document, we present the work done, since summer 94, to enhance audio and video quality for a better use of tele-seminar tool in terms of image readability, voice understanding and listening, end-user comfort. We give also a summarized report of the trials made for usage enhancement in terms of seminar management, control, and global status of distributed activity reporting.

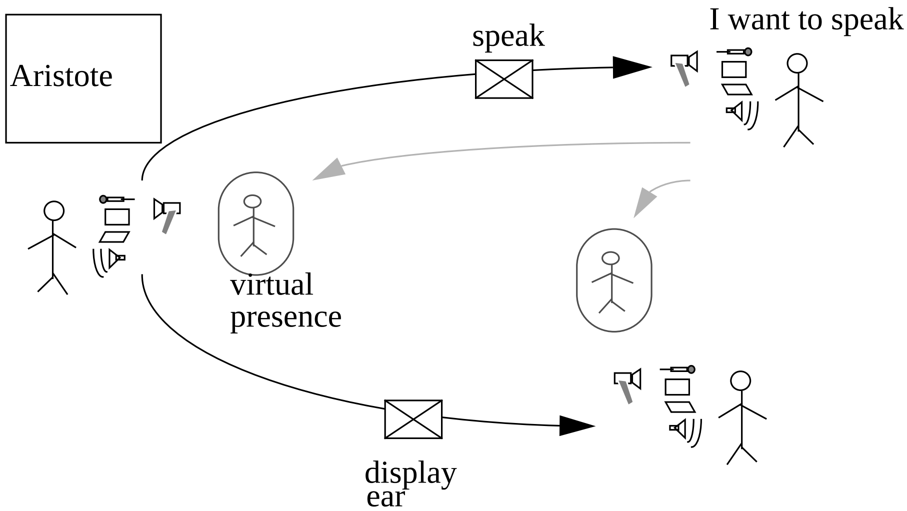
2 Overview of TÉLÉSIA activity

Experiments are organized to cover a large set of usages: seminar, meeting and free talk. Several public manifestations have been broadcasted: 25 years of INRIA, 5 years of Paris - Dauphine university, French Regional backbones promotion, and «La Science en fête» allowing us to give an access to our tools to a large public improving the natural aspect of the interface and the efficiency of the offered services. During the Interop'94 (oct. 94), the "CANAL News+I" tv channel was broadcasted on MBone from the CNIT La Défense (Paris).

But, for us, a good prefiguration of professional activity is the broadcast of Aristote's* on RENATER and Internet network.

2.1 Aims

The juxtaposition of audio and video technologies is not sufficient to realize neither a meeting nor a seminar. The progress of this type of events cannot be reduced to a simple broadcast of live images enriched by sound or vice versa.



Picture 1: Seminar protocol: control and enslavement

A seminar is a work-group activity with particular rules well known by the participants. That avoid us to develop a particular algorithm to simulate some work-group activity. The dissymmetric

* Aristote association is a French users group coming from universities, schools and research centers : CEA, CNES, CNET, CNRS, Ecole Polytechnique, EDF,...INRIA..., they work in the field of advanced computing developments. The main objectives are: to share resources for prospective studies, to promote new services and products, to organize knowledge dissemination through seminars. The studies are organized in working groups focusing on the following topics : ARES for the security, X500 directory, ARNIS connectivity between RENATER (Internet) and the national ISDN network, GARI network management, graphic and human interface, high speed networks, advanced mail, multimedia activities, seminars organization

dialog (one speaker to n listeners at every time with particular rules changing the roles) allows to adapt the size of the experiment to the end users availability and/or the computing resources (network and computing facilities). This kind of activity give the privilege to the evaluation and enhancement of the usage. This pass in quality enhancement of the technology, user environment and activity management.

2.2 Organization

TÉLÉSIA seminar is a networked activity in many senses:

- it imply several sites and workstations in an interactive exchange,
- it is performed by several people as end users, developers, engineers and event producers.

The Polytechnique School provides the seminar room, the local network infrastructure, the audiovisual service for analog image and sound processing.

The participants are located in France and connected on French MBone. Some seminars are broadcasted for Europe and the world.

The technical tele-seminar organization consists in establishing the Polytechnique multicast branch, seminar announcement on MBone, seminar room preparation (workstations, audio and video analog signal processing, network connection, partners' assistance to join conference, TÉLÉSIA software delivery).

The organization of the tele-seminar is included in the seminar organization (Aristote and Polytechnique school).

These simple functionalities need to be integrated into global service and became complex to manage. Despite non standard service on computer networks today and non commercial activity, the quality of the service need particular attention to be perfect.

2.3 Technology

Beside the computerized technology developed for this activity which is something quite familiar for us, we had to gain experience in tele-seminar production in terms of audiovisual peripherals connection and usage, synchronization between lecturer and remote participants, network infrastructure administration at the national level and social event control.

2.4 Networking

All experiments are based on MBone and Internet. The observations and the work done in this context are described further.

2.4.1. TÉLÉSIA software

Today there are no fundamental obstacles for audio and video transmission between few people connected to the network. The basic technologies for that new interactivity sketch already exist:

- the terminal: a standard workstation including audio capture and play features, specific equip-

ments for the video data capture,
- standard communication protocols: Internet Protocols including the new emerging one: RTP, multicast, and in the near future IPV6,
- voice or/and video oriented data processing methods (for computing or/and communication management): H261, MPEG[27], JPEG, postscript, PCM, ADPCM, ulaw.

However, because this experiment is performed on public networks, we have to adapt the data flow parameters to cope with the available network bandwidth. Several parameters influence the troughput of the tele-seminar service. This depends on the data compression policy chosen for audio and video. It depends also on video frame rate and number of simultaneous audio and video streams for a given session.

The description of the TÉLÉSIA software is not the main topic of this report. It is just necessary to know that we use the H261 CCITT [5] norm to manage video compression and encoding. The application consist on an integration of several “methods” into a global application. It is used as a test-bed for technical developments in the field of networked multimedia adapted to the tele-activities.

The modular architecture allows us to interchange functional modules to evaluate some new compression protocol, transport protocol or scheduling scheme. We also adapt the supervision (control of the tele-seminar, management of the tele-presence...) to the new features highlighted by the live experiments.

2.4.2. Seminary production

Beside the computerized part of the tele-seminar there is production activity which describes the “show” controlled. Indeed, because the remote people do not participate physically to the seminar sharing the same room, it’s quite unavoidable to perform a real animation with sound and images.

The tele-seminar must be produced like a TV show, with a real management of camera effects, and good audio system. We decided to use a slide projector fitted out with its own camera to improve the quality of the slides’ images. The sound is mixed by engineers. The main difficulties concern the quality of the receiving equipments. They are private choice of end users and some time we remarks some non compliant workstations with too low processor performances, display quality and network load. This affects the subjective quality and the visibility of such experiments. But, for the moment, we are not well armed against the TV fashion.

2.5 TÉLÉSIA working tracks

2.5.1. Networking and data transport

Internet is a critical resource. We try to cope with the limitations inherent in the current infrastructure. But they are conflicting with the end user needs in terms of high video frame rate, high fidelity sound and wide conferences (number of full participants).

That leads us to search the best compromise between network load/limits and service performances/quality. The fields to investigate were and remain data networking, user interface adaptation to control appropriate network parameters, global network architecture management for a given event.

Further you will found the experiments for this domain.

2.5.2. Multimedia data processing and seminar service developments

There are two tracks. The first one is the basic technic development or evaluation as compression/decompression algorithms and software, capture of images and rendering, audio processing, environment supervision.

The other one is the seminar software developments as an integrated view of several functional modules. It can be divided in three parts:

- **Seminar emission**, software part for audio and video broadcast as well as supervision services,
- **Seminar reception**, software part to follow the seminar and correspond with the seminar moderator,
- **Seminar organization**, software part for supervision and development of cooperative administration. It contains also development of evaluation tools: network statistics and services functions reporting. The seminar services development is not directly pointed by this report.

Information on seminar software developments are given to have a better understanding of TÉLÉSIA work and to situate the context of technical proposals made further for data packetization and network traffic analysis.

2.6 Some major impacts

The use of TÉLÉSIA has revealed a lot of new requirements impacting seminar organization and software development. Many things may be done for a better presence rendering. Obviously, we have made choices and taken priorities. Indeed, without a reliable communication system and a good audio and video rendering, it could be not possible to improve the seminar quality of service. This is the reason why we have take pain to master the network and also the basic audio and video processing.

2.6.1. Tele-seminar organization

To evaluate the real impact of this advanced service, it was necessary to have the best service as possible. To do that, the communication infrastructure were controlled, and the seminar environment adapted as for the tv broadcast.

Concerning the usage, there are few problems due to the technical conditions of the experiment. The end user has variable results in terms of quality of service. This is due to the data transport influence. The audio and video quality are affected by the packets' losses. The people who are familiar with the Internet know that it is more and more often saturated. Therefore the real-time applications are made sensible by the necessity to process continuously the audio and video streams. Consequently, they can render different quality on the receiving part.

That affects directly the presence rendering, and limit the end user's capacity to follow the seminar. This is the reason to identify the major causes and solve them as priority. The conclusions of these in-the-field tests have led us to identify, at least, two tracks for analysis:

- Quality of service based on the technical aspects,
- Quality of service based on the usage facility.

3 TÉLÉSIA services' enhancement.

3.1 Objective and subjective analysis

TÉLÉSIA is an automated medium between source and destination. Without any intermediary, information received is theoretically the same as the one produced: no decoding, no distortion, no interpretation. The use of TÉLÉSIA implies the following question: are the informations received the same as produced ? In other words: "Is the tool a good medium ? " Is it acceptable for the end-user ?

The used technologies are not exempt of limits that can be harmful to the service's quality: compression and non-deterministic networks, imply damages. We have studied these damages, to know if the tool could really be adopted by an end user for a daily use. We also want to determine what kind of technology could be adapted to minimize distortion and induce a good feeling to users.

The difficulty encountered to lead this study is inherent to the type of transmitted data: audio and video. So we had to investigate two fields:

- **objective or non-empirical**: concerns what could be effectively measured with evaluation tools and interpreted in a deterministic way, i.e.: off sequence packets of sound, lost packets, throughput. There are well-known technics to manage such technical hitches, i.e. connected mode oriented transport. But these solutions does not work in a multi-points environment,
- **subjective or empirical**: concerns the satisfaction induced by the service to both the lecturer and the listener. It's a human and psychological analysis, that is the threshold which determines whether the information is understandable, and how to measure or/and identify it. Even if Rao [3] or Torkzadeh [4] explore and define some tools in order to identify the satisfaction criteria and to qualify them, there is a lack of determinism between criteria retained for analysis and technological elements to change and/or to develop.

In order to do that, we need to perform regular, interactive experiments in real context and data. We need also to receive pertinent reports. The technical people who participate the experiments may give subjective and objective comments.

To summarize, the interest of the Aristote seminars broadcasts lie in the following points:

- the seminars regularly organized are a meeting point between the users and a new software release,
- the seminars are interesting. The remote listener does not want to play with a new trend of application but desires to participate a conference,
- the organization of broadcasts implies a minimal respect of listeners, lecturers, and Aristote association. That implies a professional approach of that activity,
- at the moment, several listeners are technicians, and they know different aspects of distributed multimedia applications and network management. So, they are an interested in analyzing with us the problems encountered.

3.1.1. Communication analysis.

During the experiment, several people have declared that the reception was not equal along the day. Some telecommunication experts explained the variable quality of the reception with the delay of the network transit. Facing this report of poor audio quality and distortion affecting video, we organize different campaigns to analyze the network's behaviors and performances in:

- transit delay,
- packet losses,
- off sequence packets.

3.1.1.1. Basic network quality enhancement.

These studies are organized to determine if either the IP level, or MBone level or application level can be enhanced and in which way.

The transit delay is the time for the information to arrive. This delay is noticeable if it can not be measured in seconds.

Problem due to off sequence packets can be solved using a reordering policy at reception. The use of the connectionless mode induces critical packets losses in this application.

At the IP level, the measurement were good, meaning that the RENATER management was good, excepted for a regular and brief network lock affecting the transit delay, the packet losses and the packet sequencing. The routing policy was suspected to introduce this typical effect. We send a report and explanations to the Internet service provider who fixed that problem. It result a more regular throughput and some hiccough disappeared in data streams.

However, it is clear that the network has to be increased to support the drastically growing traffic on it.

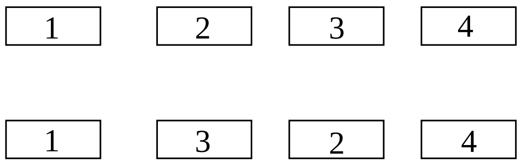
3.1.1.2. Packet sequencing influence

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 seminar day. But we decided to not 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 says the speaker. 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 actual packetization technic [1]).

The impact on audio is evident. For the video, the off sequence packets are, theoretically, not a problem because the image is split-up by packetization. The image rebuild could be done continuously. Nevertheless, the H261 being a continuous stream, the off sequence packets disrupt the data flow if there is no reordering policy:



The packets 3, 2, 4 are de- synchronized,
the stream is break for three packets.

Picture 2: Off sequence packets in a video stream

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

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

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

- the number of off sequence packets,
- average off sequence of the packet’s distance: 1 2 3 4...
- the number of lost packets,
- the number of lost packets sorted by size of the packet’s group: 1 2 3 4...
- the number of rejected packets: they are the one that arrive 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 a way that this profile should determine the conditions of use, depending on the nature of the data:

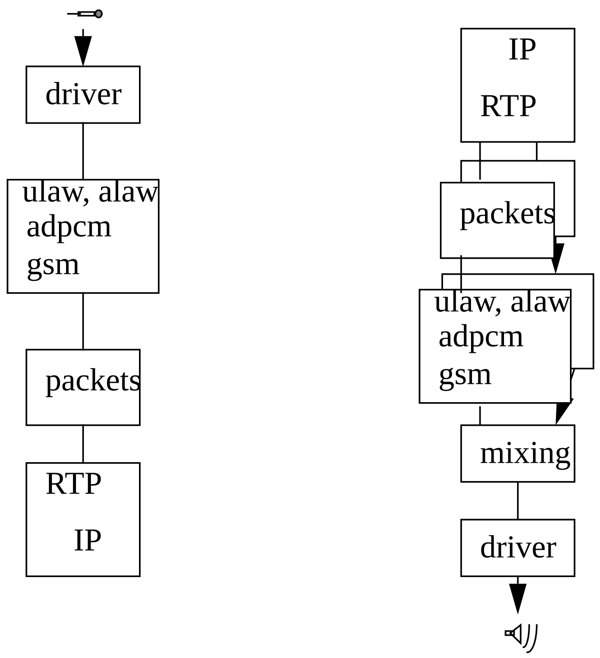
- the control channel of the application, due to the low bandwidth available and the contextual independence of the packet, is set to 0: no reordering,
- for the video and audio, the size of the window is initialized to 2,
- due to the use of either the manual control or the automatic detection of audio,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

The impact of the new feature on the code is very light. The current reading interface did not change. The rest of the interface was only enriched in order to accept the setting initialization and the tuning of the size of the window. During the execution of the code, the main impact is the delay implied by the management of the window. Of course, that delay depends on the size of the window.

Here, at INRIA, we noticed a off sequence packets ratio of 1. But for some sites, the poor quality of the connection does not allow a better reception than a delay of 3 or 4 packets.

3.1.2. The audio

Firstly, we will describe the architecture of the sound management. Two processes are designed for this aspect: one for the encoding of the sound stream output, and one for the decoding of the multiple input streams:



Picture 3: Sound codec, architecture

Where:

- **driver**, is either a driver for the devices, or an interface for the DEC AF server,
- **ulaw,...**, is the coding - compressing stage, ulaw, alaw, ADPCM, and GSM,
- **packets**, defines the way the samples are put into IP packets. Depend on the choice of the coding. The RTP header contains the coding identifier, the sample follows the RTP header, the original sample is read from the device in ulaw or alaw format. The size of the sample is 1024 bytes. The result in ADPCM is 512 bytes, in GSM is 160 bytes,
- **RTP - IP**, is the Real Time Protocol and IP network interface. It manages the RTP header formatting and the IP interface through a UNICAST or a MULTICAST mode,
- **mix**, is the mixing of the multiple sound streams received by the decoding process.

During the conferences, audio was affected. The main symptoms could be resumed by:

- there are holes inside the stream, implying some very short amounts of time of complete silence,
- there are disturbing “*clicks*” sounds inside the stream at the beginning and at the end of the pre-

ceding described holes of sound,
- there are off sequence samples,
- there is an echo between the microphone and the loud speakers.

3.1.2.1. Some answers

One of these problems, the off sequence samples, has been already described and solved in the preceding chapter: protocol, by a general algorithm in the RTP layer.

The disturbing “clicks” sounds where generated for two different reasons. Inside the stream, they appear between two contiguous samples, meaning that time is wasted between these samples. The solution is to manage a “time to play the sample”. This management is the first step of a complete synchronization, with a “time to play the conference”. The second origin of these sounds was that the holes, because of the lost packets, began and ended by a too violent front wave. We solved it, by filling the holes with a pause.

For the echo, we know that the question is not simple and that there is, at this time, no real solution, no efficient algorithm. So we designed two simple mechanisms that solved partially the question, or avoid the problem:

- firstly, we added a “*push to talk*” button. That button has two functions: it stops the audio play, and allows the new speaker to talk. The management of the inversion forbids anybody to cut immediately the current audio stream,
- secondly, we added a level of detection of the audio. When this level, defined by the user, is reached, the samples are sent. That technic is also a solution to limit the saturation of the network bandwidth, by stopping the sending out of samples when someone stops to talk.

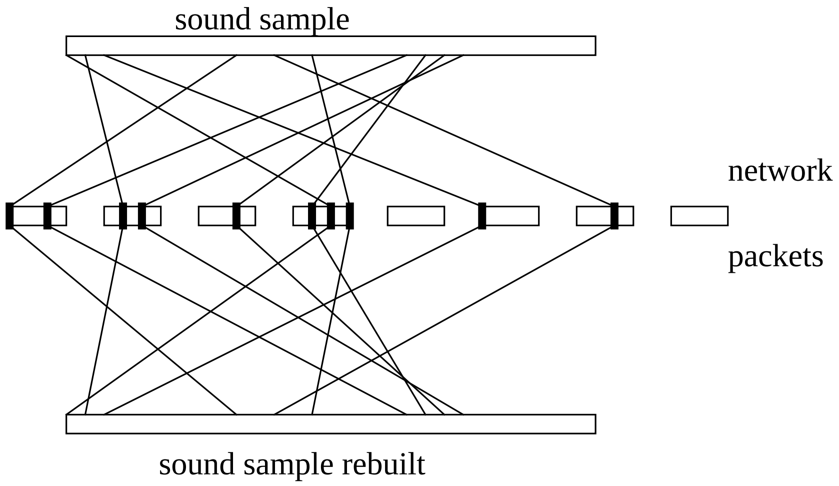
Both systems help to limit the bad effects of the echo, and we decided not to spend more time on it. More interesting questions needed to be solved to increase the quality of the conferences system.

3.1.2.2. A new packetization

One of the more profitable track concerning the quality of the audio, is to solve the losses of sound sample 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 to the delay implied by this protocol.

So, we need to minimize the effect of the lost samples. As seen previously, firstly we fill the holes by pauses to avoid the disturbing “clicks” at the beginning and at the end of the holes. But the holes still exist and the sound stream remains broken and hard to follow.

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, it’s description is the following:

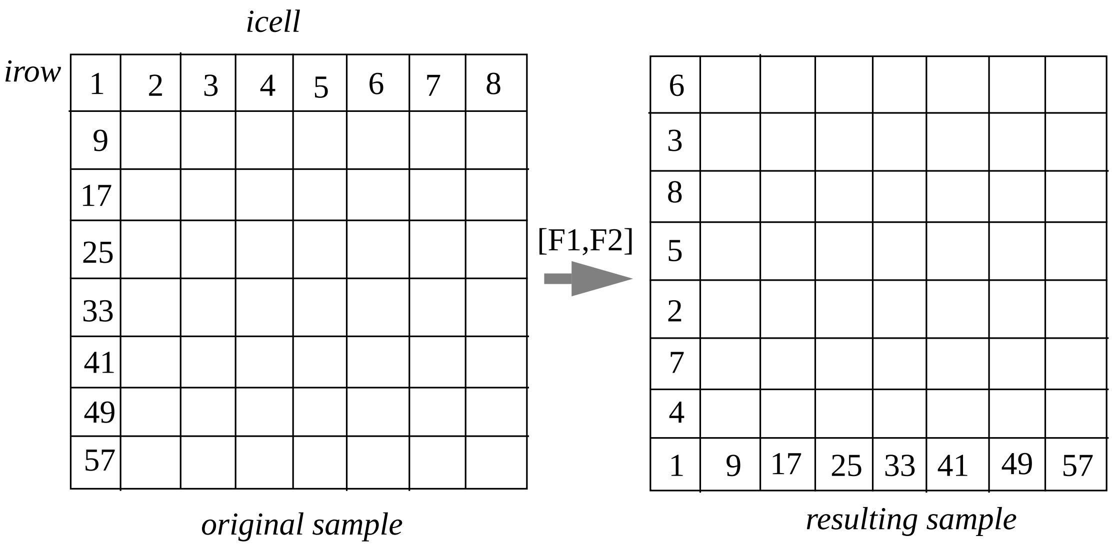


Picture 4: Sample distribution into the network packets [TÉLÉSIA TEAP mode]

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.

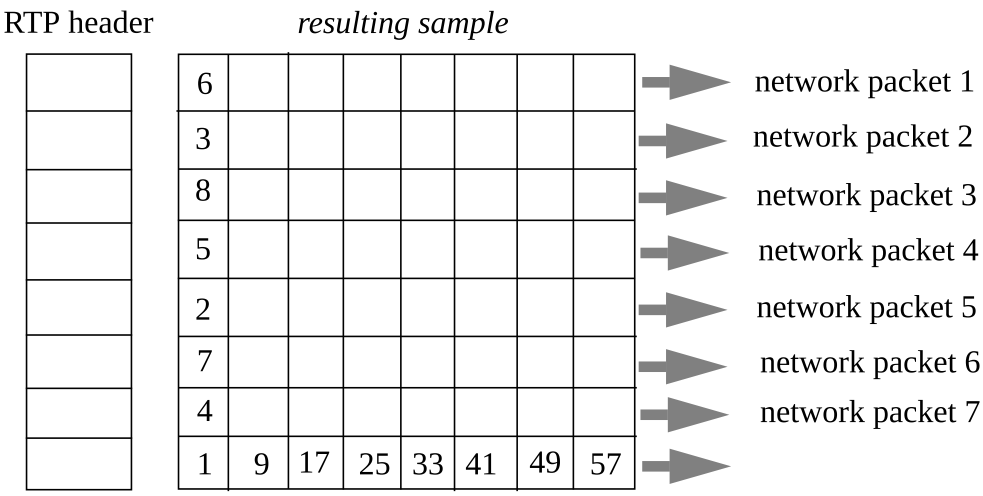
The acquired sample is called the *original sample*, the compute sample is called the *resulting sample*. The *original sample* is projected into a two dimensions table of 8 by 8 cells of sound. The rows are indexed by *irow* and the cells are indexed in a row by *icell*. The *resulting sample* is computed so:

- *resulting_sample* [F1(*icell*), F2(*irow*)] = *original sample* [*irow, icell*



Picture 5: Distribution formula

In this formula F1 and F2 are two bijective functions that are designed to spare the sound samples. It means copying the sound samples into sufficiently spaced network packets, and avoiding a creation of holes of sound at the receiving part (see Picture 7) in case of losses:

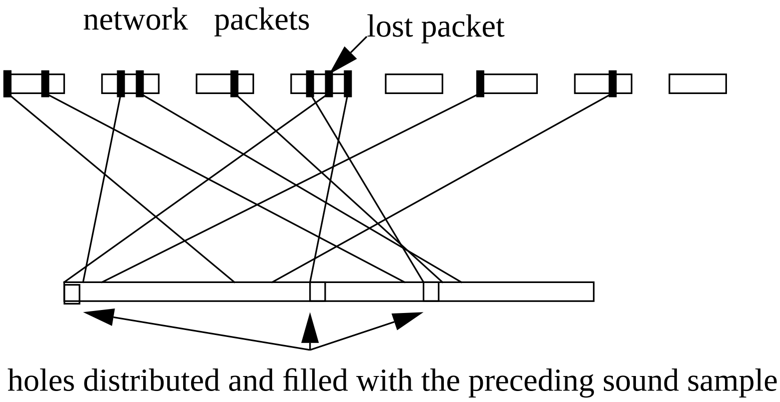


Picture 6: Packetization scheme

Now, you see what could be the impact of a lost packet. A network packet contains non-contiguous sound samples. 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, each hole, is filled with the preceding sample in the rebuilt sound sample. The scale used in our 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 7: Impact of a lost packet

3.1.2.3. RTP usage

The packetization is formatted with the RTP layer. On the emitting side, nothing particular is done. We choose a factor aligning the original sound sampling with a power of 2 sequence number. On the receiving side, the sequence number is used as a parameter of the inverted distribution function. The impact on RTP is very low. The following sketch is the packet structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                          RTP header                           .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 8 ulaw/alaw samples = 128 bytes               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Picture 8: Audio packet structure

3.1.2.4. Results

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 analyzing the impact of the packet size on the lost packets' statistics. The experiment context defined for the test is the following: TÉLÉ-

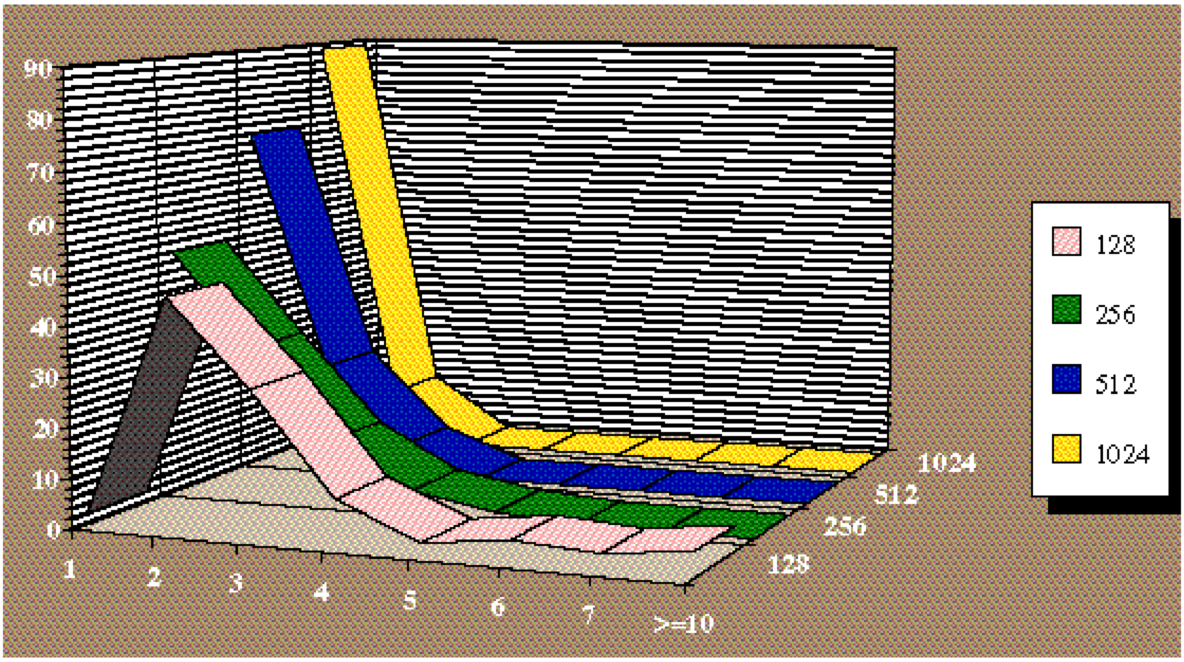
SIA sends the audio stream through the INRIA's Ethernet LAN which is connected to the RENATER network, the output bandwidth is 2 Mbits, then the stream goes all over a closed circuit of work-stations in France before coming back.

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. The following array shows the result of measurements and the subjective comments:

Type of coding	Packets size	Lost percent	Comments
ulaw	1024	8%	non regular, non-comprehensible
ulaw	512	13%	regular, comprehensible but hard to follow
ulaw	256	24%	regular, non-comprehensible, to hard to follow, stream largely broken
ulaw	128	31%	regular, non-comprehensible, stream largely broken
new TEAP	128	30%	comprehensible, voice with tunnel effect when packets are lost

Table 1 : Audio packetization tests

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 9: Distribution of the holes size. Ratio per packet

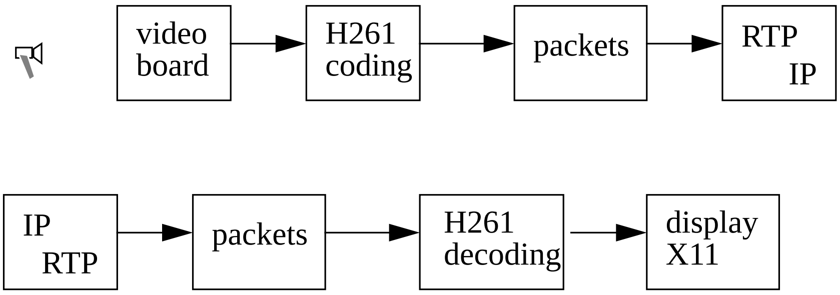
After these different experiments, we can conclude that the throughput is clearly a part of the network aspect. That's the most often criteria used for the quality of service evaluation. However, the size of the packets and, by implication, the number of packets commuted a second is important too. The routing system seems to be more sensible to that last factor. Because 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) and the size of the packets sent to the network.

3.1.2.5. Audio conclusion

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 (morpho analysis) and/or more robust in case of lost packet (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.

3.1.3. H261 coding and decoding

The video codec architecture is equivalent to the audio: one process dedicated to the video coding, another to the video decoding. The difference reside in the decoding process that handles only one video stream:



Picture 10: Video codec, architecture

Where:

- **video - board**: is the direct access to a frame through:
  - Capture board and os driver,
  - Graphic library.
- **H261 - coding**: generate the H261 stream with:
  - movement detection by comparison of two consecutive frames,
  - coding of Macro Blocks, in INTER or INTRA modes,
  - extended information for the packetization layer.
- **packets**: formats the RTP - IP packets with the H261 stream and the extended information. The packets vary from 20 to 1446 bytes,
- **RTP - IP**: the RTP layer shared with the audio processing. There is no specific profile for the H261 stream,
- **display - X11**: the decoded frames are displayed in RGB and depending on the screen depth in full color, 8/4/2 bits mode. A dithering is done for less than 24 bits depth. Each frames can be zoomed by 4 or 1/4. Brightness and contrast control are available.

IVS - Sophia first released the H261 packetization described in *draft-ietf-avt-video-packet-02.txt* [2].

During the seminars, we noticed two major disruptions:

- the frame is frizzed for a while,
- a part of the frame is shifted on the left doing patchwork effect on the display window.

Subjectively, we could say that the more bothering point is the second one, even if the end user gives the frame rate as a quality criteria. 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. Actually, after analyzing the code, we conclude that off sequence packet where already treated in the H261 decoder. But we noticed real problems for the decoder to re-synchronize itself after a loss of 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:

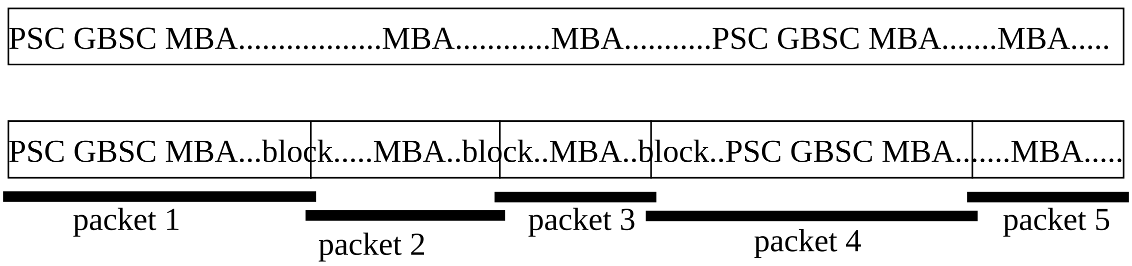
- the S bit: the packet includes the beginning of the encoding of a GOB,
- the E bit: the packet contains the end of a GOB.

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.

H261 stream



Picture 11: Current packetization

In this sketch, the Blocks are very often dealt between two packets out.

3.1.3.1. 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.

3.1.3.2. Some proposals to improve the H261 packetization

A part of the proposition is to reduced the number of sent packets. That can be achieved by reducing the throughput of the H261 stream generated by the coder. That has been achieved by adapting the quantization level or changing detection threshold. But in the same time the frame rate increases thank’s 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 packet alone.

The current H261 packetization is context free. One solution to solve the points 1 and 3 is to complete the H261 header with:

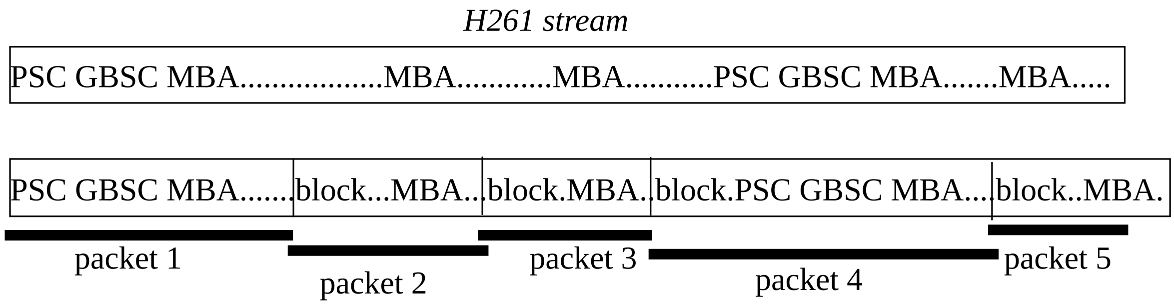
- the **GOB number**, number of the Group of Block which begins the packet,
- the **MB number**, number of the Macro Block which begins the packet,
- the **last MB number**, number of the last Macro Block decoded,
- the **type** of the last object of the packet: Picture, Group Of Block, Macro Block.

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:

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

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 12: 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 analyzes 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 analyzes 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.

3.1.3.3. RTP usage

To implement the new H261 packetization and use the RTP layer, it has been necessary to re-organize and rewrite completely the code. Now, the decoder is an automate with an internal state, external tokens, and semantics actions for the H261 stream. The original sequencing is suppressed and already done by the RTP layer. The implications with the RTP layer reside in the use of the header, which is no longer used to record H261 specific information, excepted the type of the conference channel.

Structure of a new H261 packet:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                           RTP header                          .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          H.261 header                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         H.261 stream...                       .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Picture 13: H261 packet

New H261 header:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    GOB no     |    MB no      | last MB       |     SBIT      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     EBIT      | D |  F  | I |V|C| |    format     |     MBZ       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Picture 14: H261 header

GOB no (8bits)	GOB number in the Frame
MB no (8bits)	MB number in the GOB
last MB (8bits)	Last encoded MB number
SBIT (8bits)	Start bit position number of the GOB

Tableau 2 : Fields' explanations

EBIT (8 bits)	End bit position number of the last GOB
D (2 bits)	type of the beginning of the packet: Frame, Gob, Macro Block
F (2bits)	type of the end of the packet: Frame, Gob, Macro Block
I (1 bit)	Full Intra Image flag. Set if it is the first packet of a full intra image.
V (1 bit)	Movement Vector flag. Set if movement vectors are encoded. All V bits of the same frame must be identical.
C (1bit)	Color Frame?
format (8bits)	Frame format: Qcif, Cif, Scif
MBZ (8 bits)	Must Be Zero.

Tableau 2 : Fields' explanations

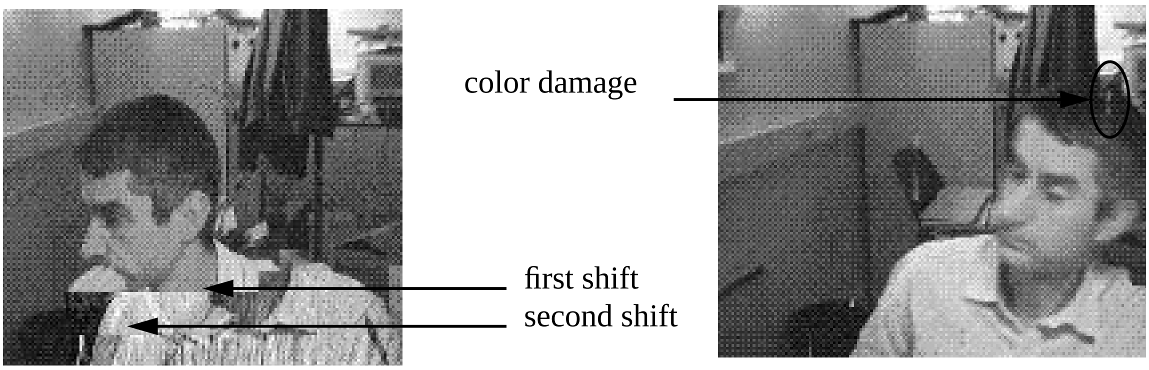
3.1.3.4. Results

Quality enhancement:

The enhancements analysis is done through both points of view, defined in the 3.1 chapter:

- subjective: the quality of the images,
- objective: qualification through the network traffic, and the frame rate.

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 15: Corrected Video damages

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

Still remain the color 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 moved objects of 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 enough frequent to imply a specific treatment by the application.

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.

The measure campaigns:

The measures campaigns were organized both during specific sessions, for the objectives aspects, and during in-field-tests with the Aristote seminars, mainly for the subjective one. The measures campaigns' date are December 1994 and May 1995. The tool used to extract these measures is integrated in the H261 packetization layer. Two kind of video record were used to evaluate the old and new packetization.

The workstation used for the campaigns is a Sun workstation:

- processing unit: Sun SS10 sx - bipro 512,
- memory: 64 Mo,
- digitalization: Sun video,
- system: Solaris.

Example 1:

The first material is a 3 minutes video film report on grape harvest in Burgundy. The video is very animated, the frames change very often.

The following table shows the increasing throughput in terms of packets and frames per second. We notice that the rise of packets throughput is more important due to the lower filled packets ratio:

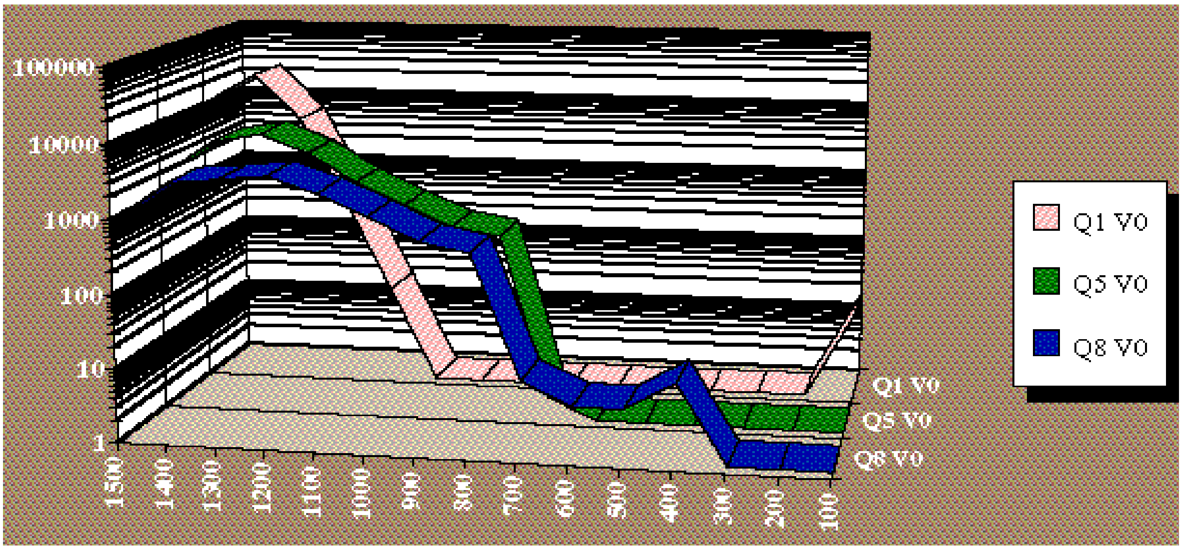
	Frame * sec.			Throughput in Kbit * sec.			Packets * sec.		
Quant	V0	new	%	V0	new	%	V0	new	%
1	6.06	7.01	+15.5	1583	1728	+9.15	146.0	182.1	+24.7
5	7.9	9.6	+20.8	708.2	839.5	+18.5	71.9	102.4	+42.4
8	9.6	11.7	+21.9	503.1	595.2	+18.3	53.7	75.5	+40.5

Table 3 : Speed enhancement

Formal notations: **Quant** or **Qx** design the quantization number, **V0** designs the first release of the packetization, the release of the TÉLÉSIA software is December 1994, and **new** designs the last release May 1995.

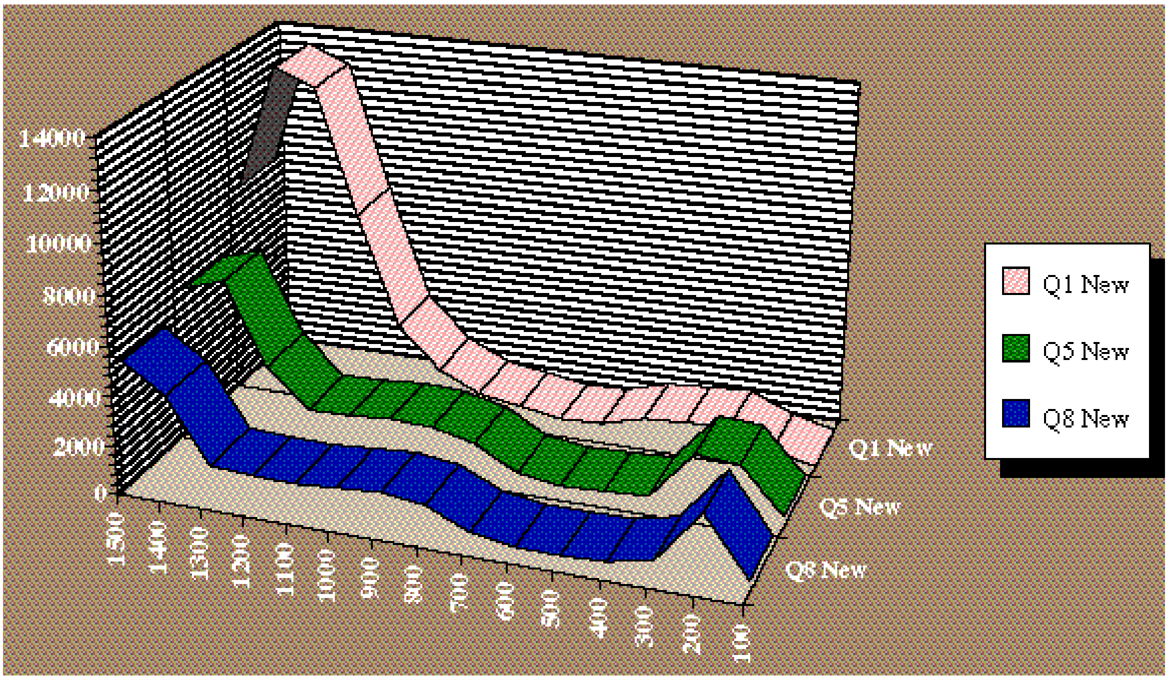
The two following graphics give the of the packets' size distribution. We noticed rapidly the effect of the choice of the Block granularity as a cutting border:

- larger distribution over the scale of the packets size,
- smaller packets



Picture 16: Old packetization: packet's size distribution.

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



Picture 17: New packetization: packets' size distribution. Linear scale

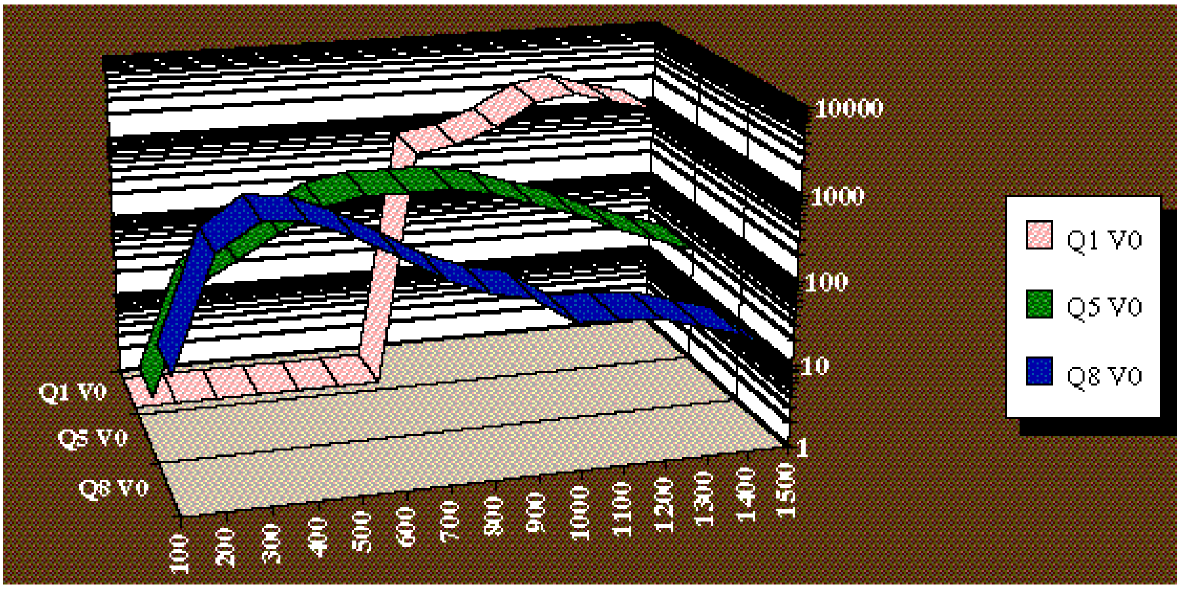
Example 2:

Now the same measures have been done with the video tape of a conference[20], same time: 3 minutes. The main difference with the preceding test is in the fixed video image background: only the speaker moves. The video contains in the same proportion speaker talking and graphics slides presentation. The following table compare the video report and video conference ratio (frames, throughput, packets, per second):

	Frame * sec.			Throughput in Kbit * sec.			Packets * sec.		
Quant	report	conf.	%	report	conf.	%	report	conf.	%
1	7.01	5.7	-18.7	1728	458.4	-73.5	182.1	59.3	-67.4
5	9.6	15.3	+59.4	839.5	117.1	-86.0	102.4	18.1	-82.3
8	11.7	16.4	+40.2	595.2	62.8	-89.4	75.5	17.4	-76.9

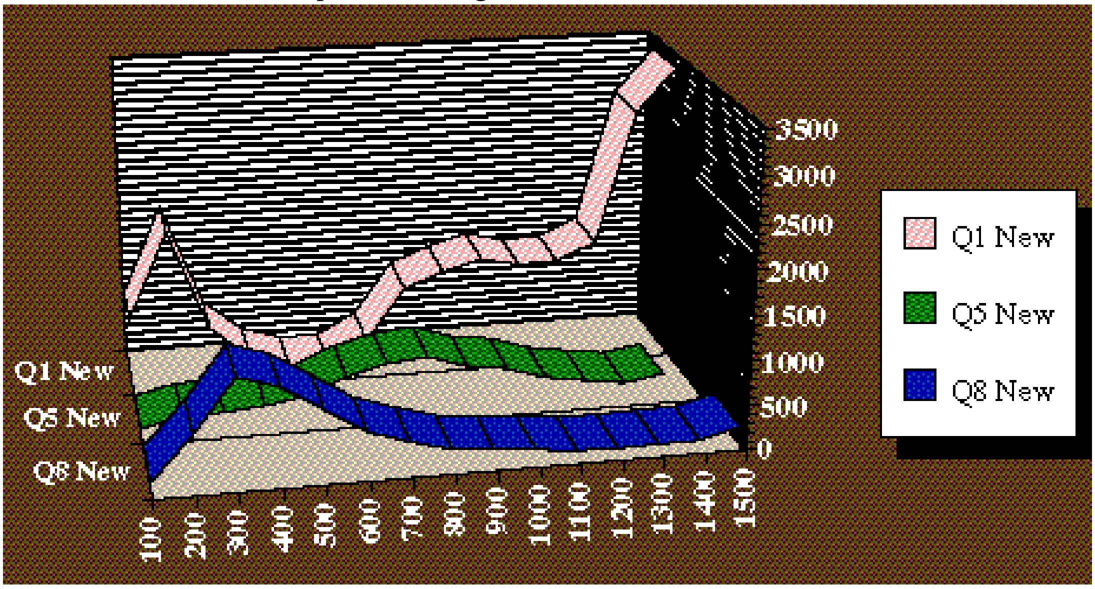
Table 4 : Difference between video report and conference

Then, follow two graphics on the packets' size distribution. 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 seminar broadcasting.



Picture 18: 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 19: New packetization: packet's size distribution. Linear scale

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

4 Usages enhancement

4.1 Network organization

Network organization implies a global connectivity and also a local background.

Before july 93, the MBone connectivity between sites had to be established or checked before each seminar. The problems encountered to have a good quality network led us to organize the French Multicast Bone which is structured now.

The global FMBone topology for France is a tree build between academic and R&D sites over RENATER. The national main router, the root router, is installed at EDF and connected to RENATER with a 34 Mbits link. The EDF router has tunnels with other countries. The main management activity is to equilibrate the load of tunnels.

Concerning the local background, new participating sites are integrated before each seminar. It is necessary to cope with security procedures installed by every site. There are some firewalls to build or routing filter to establish the incoming multicast traffic. For the moment we do not have any standard procedure to establish the virtual network for a given multicast group.

At Polytechnique School, a particular network is created in the seminar room and a tunnel is established directly with the national main router.

4.2 Workstation configuration

The TÉLÉSIA application runs on two workstations:

One is used to capture and broadcast information about lecturer.

- It generates audio and video streams:
  - audio is encoded in ulaw (64 Kb/s or ADPCM 32 Kb/s),
  - video is encoded using H261 standard in QCIF size (minimize the bandwidth used on the network),
- it manages the seminar and the right to speak.

The other generates the slides video stream (until they can be generated electronically and broadcasted using such solution like w3 based tools):

- slides are encoded in H261 in CIF (they can be zoomed on the receiving workstations).

4.3 Software functionalities

One workstation is chosen at random as the moderator. A signalization mechanism allows the moderator to control all the participants. The seminar supervision is build on this signalization. The main functions are:

- audio and video coding remote driving (selection of streams to decode, lock on emission functions),
- dialog between moderator and participants to negotiate the right to speak,
- announcement of changes in each participant context.
- maintenance of communication activity between participants. Due to the multicast limits, in terms of protocol reliability, there is no warranty for the good reception of the data. This feature allows to repeat signalization messages. There are different policies to manage the duplicated messages, depending on the type of information.

Using the supervision functions of TÉLÉSIA, the moderator can drive the remote sites to select audio stream that are to be processed. This orders are permanently actualized to allow the user a total transparency of seminar reception.

4.4 User Interface

When you participate the seminar, you are only to listen and ask, maybe questions. For remote listeners, the tele-presence must be improved to allow a good feeling with the event. Therefore, the signalization reports a lot of informations to the user interface to know:

- where is the moderator,
- who is speaking,
- if your request is registered,
- if you have the right to speak,
- if you have a feedback from the seminar environment.

The end user has at his disposal:

- several audio controls as sound parameters, the push to talk button, free hand microphone mode, visual feedback with color status (for pending request, for registered request, for allowed right to speak...),
- several video control as the image size, color, brightness control and also frame rate, compression ratio and color/black processing mode,
- global status report as list of participants, status of them, video frame rate/s, used network bandwidth, moderator name, status of audio and video peripherals.

That data and reports are frequently changed for a better information and understanding of the end-users accordingly with experience gained during seminars or public experiments.

4.5 Seminar administration

To evaluate a real advanced service, it is necessary to conduct experiment with a maximum of realistic characteristics. This includes seminar announcement, participant registration, live management, ending. To give a real transparency to end user, some usual administration technics have to be integrated. The difficulty resides in the networked activity which does not facilitate a global view of participants.

4.5.1. Seminar Installation

This identify the seminar install, addresses definition, advertising,...

For the moment, the most used tool is Sd which allows the MBone management. It maintains the events database with technical informations. This features require more accurate data. A cycle of seminars is announced for the beginning of the school year. But, the announcement should include not only informations about technical matter as multicast address or software to use, but also registration forms, seminary advertising, administrative registration conditions and interfaces.

We investigated the X500 [25] directory approach to store all relevant informations about a seminar before, during or after it. We wanted to take advantage of the world wide availability of such directory and the potential of available services. This allows to manage in a uniform context all relevant data for announcement, registration, live control and ending.

This directory can be used as a “names” directory (logical declaration of technical data: mapping between logical names and physical multicast addresses, participants’ names, sites’ names...) or a “service” directory (localization, name resolution...). This last functionality is viewed as the one defined in OMG [23] model in the broker environment [24] and Open Distributed Processing [22]. In this case, several information concerning the TÉLÉSIA distributed activity can be accessed in real time to process dynamically the links between services.

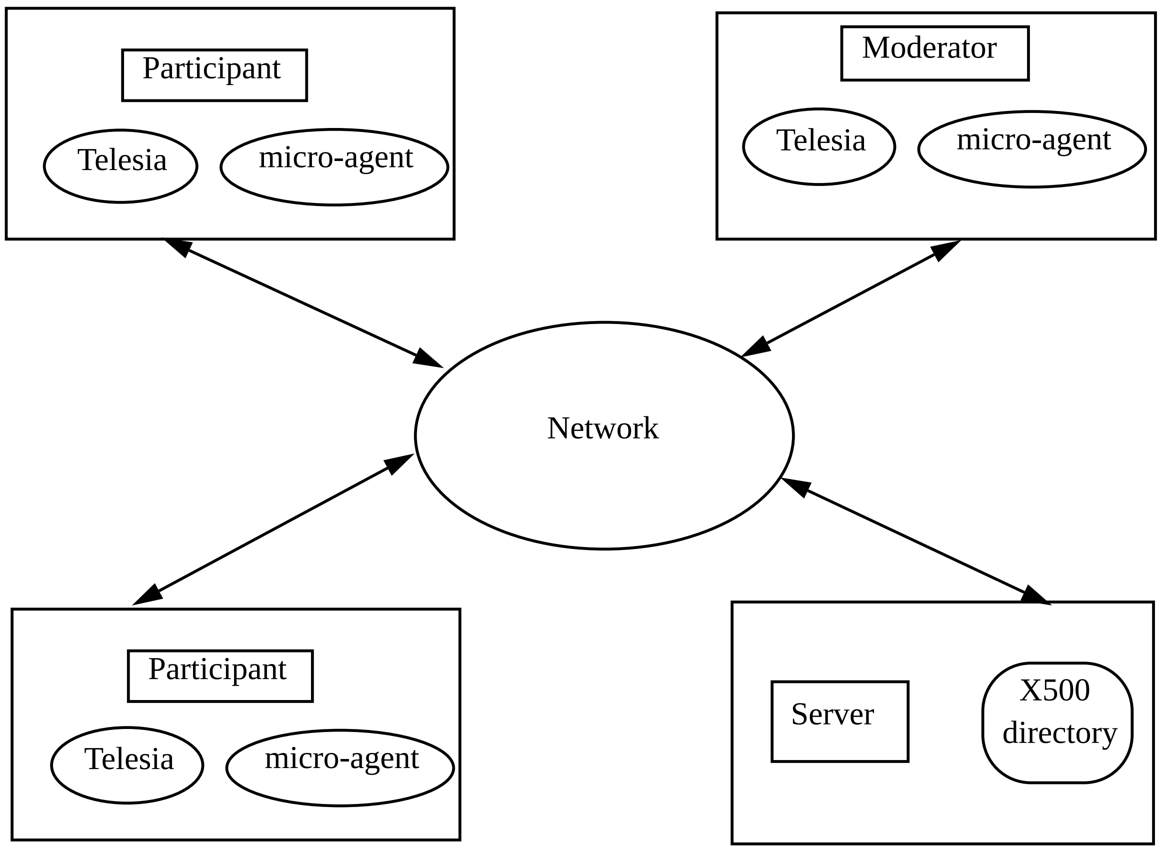
4.5.2. Participant registration

For the moment, the only way to participate a seminar is to reach the conference during its process. There is not satisfying in terms of standard management or more economical activity. In real world, the registration phase gives the organizer the possibility to prepare, install and control the required resources (network links, computing resources, security).

The X500 approach, as previously said, is one way to collect data and exploit them. It represents a good interface for information exchanges between the organizer and participants. The seminar issue can be accessed by cooperative processing between administration or multimedia services programs or web for user information and interaction with.

This features are not stabilized for the moment and need more investigations. Therefore, a minimum of characteristics have been tested for seminar naming and participants list management.

4.5.3. Real time seminar management



Picture 20: Cooperative administration: schematic diagram

4.5.3.1. Seminar Opening

This phase is performed by the moderator who initializes the seminar. Audio, video and supervision packets are sent to the dedicated multicast address(es). There is an authentication phase for every actor, including the moderator. The seminar database is used. If every thing is right, the workstation is admitted to the conference, the seminar and the administration context are launched and the cooperative management activity can begin.

4.5.3.2. Distributed activity management

In 1994, we designed and tested an administration's prototype policy for TÉLÉSIA sites involved in a conference. It is inspired from the SNMP protocols for network management. The aim is the development of a net of cooperative agents reporting to each other. Beside every running TÉLÉSIA application, there is an "agent" reporting to both directions (to local application and to other).

A part of administration activity is automated and transparent to the user. It concerns the reporting of user interface status, multimedia data processing, transport statistics.

But, the data stored by the agents can be also consulted (trough user interface) by every participant application to give a better understanding of the remote environment:

- reception quality in terms of packet loss,
- the really encoded/decoded streams,
- the workstation status (coding/decoding performances),
- the local multicast parameters.

The availability of such information leads us to enhance the seminar technical and organizational supervision. Indeed, the emission or reception parameters can be remotely tuned at level requested by the networking or cooperative work conditions. By the way, some non compliant activities can be traced or stopped.

This cooperative administration could be used to manage documents distribution or access and memorize history of the group events.

4.5.3.3. Seminar ending

No particular action to end the seminar. Some economical matter have to be processed like addresses liberation, access closing and reporting (statistics...).

5 Conclusion, Open doors and future works

The main purpose of that report was not an exhaustive presentation of CSCW applications. With TÉLÉSIA, video and audio were used as convenient information for collaborative working, which has not to be proved now [18]. However, we feel that tele-presence and human protocol reproduction need to be improved by giving the user a more natural feeling of his distributed activity. That improvement could be seen at different levels: technology, and usage.

A part of the improvement could be found through better performances and new encoding functions for audio and video. For example and non exhaustive:

- physio-acoustic compression is more efficient but need a larger computing power. That compression method will allow the use of a better quality of sound 16Khz x 16 bits,
- MPEG for video is not real time for the moment, it also needs a larger computing power.

Pragmatic work could be done with the concept of the telepresence. Audio and video treated as a logical unique stream. A simple and pragmatic concept that has implication, in both the interface of the application and the networking. In the network point of view, we will study the piggy-backing of the RTP audio packets and the RTP video packets. Piggy backing is easy to implement with the new RTPV2 release, a great expectation for the limitation of the packets number throughput, a cause of the overloading of the Internet routing system.

The network congestion control and its automatic adaptations are already implemented in some applications. Limiting the video emission according with the processing capacity of the destinations introduce an oscillation of the computed troughput that can converge to epsilon. Without any bandwidth reservation, reducing the throughput does not grant that the stream will be well networked. Because some other streams will immediately used the freed throughput.

We have designed a concept based on a re-structuring of the data stream into multiple independent but complementary network streams. Each receiver chooses a set of network streams depending on either its capacity to process the data and the capacity of the network to route correctly the streams.

The second aspect of the improvement for a tele-seminar and in general for the cooperative work applications are in the usages, following are some tracks in progress:

- *the administration*: a seminar is not only the emitting of video and audio: subscription, MULTICAST address choice, authorization number, control of the telepresence, control of the following, control of the satisfaction level, control of the network traffic,
- *tele-meeting*: the main aspect of a meeting is the free access to the right to speak and the possibility for every body to saturate the meeting without any control. Coupling other experiments in the decision field is an interesting track. For example, the right to speak could be managed by a decision help software which will distribute it equally or with different criteria,
- *complete view of the participants*: the visualization of the hall either for the speakers and for the lecturers is not satisfactory. The view of the participants is either a hard listing of names or is limited by the computing power of the workstation and the place on the screen if you display the images. A compromise needs to be found between an absence of telepresence and a partial one. A new presentation of the participants based on a slower and smaller image moving for each one, a designation concept to get some complementary informations. Could the virtual reality give some answer to improve the global vue and the selectivity ?
- *organization*: as seen previously, the seminar must be considered as a TV program, that means that a lot of efforts has to be done to combine the audio-visual know-how and multimedia automation. At this time due to the lack of resources, this activity is not sufficiently oriented to a professional production.

6 Acknowledgments

We want to thank our partners: the Aristote association and all its members. We wave (by Télésia) to the colleagues who have joint the Télésia experiment and helped us in producing and broadcasting these events:

Thierry Turletti and the RODEO project, for the part of the way made together,

the INRIA sites, and in alphabetical order:

CEA, one of the first experimenter,

CNAM, for their patience and encouragement as end-users

CNET, for the collaboration on MBone and ATM,

CRIHAN, for their assistance and participation since the early days,

EDF, for the collaboration on MBone, Télésia development and experiment,

POLYTECHNIQUE SCHOOL, for the big help in networking, seminar production and animation,

UNIVERSITIES, for the participation in a lot of experiments and also networked seminar promotion,

RENATER and its technical team for the continuous partnership and friendliness.

We apologize to other friends and travellers on the Information Highways, that we have not mentioned here.

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] Queueing 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.

Annexe: A. Realization

From the beginning of 1994 the TÉLÉSIA application has been rewritten in C++ for the modularity and multiple contexts management facilities. With the C++ rewriting, the networking has been divided in three layers:

- the IP access: manage all the IP address, socket characteristics, and some statistics computing. Provides the read and write access for the upper layers. The IP management is simple enough to be in one C++ class. The main interface is: *create*, *send*, *recb*, *getfromaddr*,
- the RTP layer: manage the RTP protocol V1, headers, data description, and include a reordering layer based on the sequential stamping of the RTP packets. The RTP layers provides also the interface to tune the reordering features in coordination with the profile of the application. Call the IP layer for the network access. Three C++ class are designed:
  - an advance management of RTP typed buffers,
  - a reordering class, that manages the window and its dynamic tuning,
  - a set of dynamic ratio displays concerning the received packets,
  - the RTP protocol itself. The main interface is: *create, send*, *recb*, *getoption*, *push*, *getdata*, *getsequence*,
- the RTCP layer: call the RTP layer for the network access. The main interface is: *create*, *send_desc*, *send_bye*, *send_tele_cde*, *recb*. RTCP provide the distributed control of the application:
  - the standard protocol aspects,
  - the TÉLÉSIA specific aspects: seminar mode and the free meeting mode,
  - activate the user interface and drive the subprocesses: coder, decoder.

Both video and audio packetization have been suppressed from the coder/decoder:

- the H261ip manages the H261 packetization over the RTP packets. In output it uses the H261 stream from the coder and a specifically designed table which describes the different H261 objects in the stream. In input it decodes the H261 header and provides the decoder the token and address concerning the first H261 object in the stream. A set of ratio are computed and displayed on the output packetization. H261ip calls the RTP layer for the network access. The main interface is: *create*, *send*, *ReadBuffer*, *inith261image*,
- the audio_mixage manages the mixing of the multiple audio channels, and calls the new packetization. It calls the RTP layer for the network access. The main interface is: *rentre*, *sortir*, *pret*. Two C++ classes are designed:
  - the mixing one that sorts the audio channel and calls the packetization,
  - the reordonne one that realizes the new packetization.

The sources of this original software are available at the following address:

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

Unité de recherche INRIA Lorraine, technopôle de Nancy-Brabois, 615 rue du jardin botanique, BP 101, 54600 VILLERS-LÈS-NANCY
Unité de recherche INRIA Rennes, IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, domaine de Voluceau, Rocquencourt, BP 105, LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
Inria, Domaine de Voluceau, Rocquencourt, BP 105 LE CHESNAY Cedex (France)

ISSN 0249-6399