Multimedia Integrated Conferencing for European Researchers (MICE): Piloting Activities and the Conference Management and Multiplexing Centre

Mark J. Handley
Peter T. Kirstein
M. Angela Sasse

Department of Computer Science
University College London
Gower Street
London WC1E 6BT, UK

Abstract

In 1992, the CEC agreed to finance a one-year project called "Piloting of Multimedia Integrated Services for European Researchers (MICE)". This paper outlines the background, purpose and aims of the project. The technical work required to interconnect existing facilities, and issues relating to influencing factors - such as standards and infrastructure - are discussed in detail. In conclusion, we provide an outlook on the future of multimedia conferencing and highlights some of the next steps to be taken.

1 Introduction

Multimedia Conferencing, in the context of the activities to be discussed in this paper, implies that *voice, video and computer data* are to be shared amongst geographically distributed groups. A great deal of research has been done in the individual areas, to the extent that communication Service Providers (PTOs) now offer video/voice facilities as commercial services, whilst communication equipment providers are offering video-conferencing studios complete with workstations.

In practice, most of the facilities which are currently available fall into one of the following categories:

1. *High quality studios*, often used by large commercial organisations, and connected by high quality codecs at fairly high data rates (1.5Mbps (T1) or 2Mbps (E1)), or by analogue TV channels. The communication channels are circuit-switched, and usually only two site point-point, though sometimes multi-point is possible.

2. *Simpler, lower quality studios connected* over lower rate circuit-switched channels. These usually operate at 384 - 1024 Kbps, though more recently some use has been made of Basic Rate ISDN (64 or 128 Kbps) to connect such studios.

3. *Lower-end workstations which* operate using one or two Basic Rate ISDN channels (64 or 128 Kbps); if one channel only is used for the A/V, it is often possible to put data on the second channel of a Basic Rate ISDN termination. Such workstations usually require a *codec board* to perform video compression.

4. Multimedia conferencing facilities have been set up on *higher-end workstations. Software compression of video* in such workstations is possible, typically at speeds of around 64 - 512 Kbps. Usually the frame rate or picture quality is lower than video from hardware codecs at comparable data rates, due to limitations on available processing power. Such systems [1], [2], have been deployed to a limited extent in the research communities on wide area networks This systems are particularly interesting as the operate over packet switched networks rather than circuits.

In conference room systems, the traffic is usually carried over circuit-switched channels. This requires a constant bit rate service for video, and that any voice or data facilities have a negligible impact on the additional communication requirements. Clark [3] has shown that this is an erroneous assumption, and that there can be considerable advantages in packet-switched operation.

Under *DARPA* auspices, there are several facilities for providing multimedia services over the US Internet; DARTNET [3] and the SIMNET (Defense Simulation Internet, [4]) are two examples. Both facilities use codecs, and different mechanisms for conference control and multiplexing have been incorporated. Typically, they operate at 128 - 384 Kbps speeds, with bandwidth allocations in the routers to ensure that adequate performance is achieved, so that the codecs do not lose synchronisation. Since the European research networks have much the same technological characteristics as the US Internet, it would be desirable to provide multimedia facilities for European researchers over the same communication infrastructure, and link them to the US. This is what the MICE project aims to provide on a small scale.

In the European context, there are some significant differences to the earlier US activities. Firstly, there is a recent international standard for codecs (H.261 for video coding, H.221 for multimedia streams over serial lines, and various other standards for audio coding). In accordance with European policies, MICE will use these standards [5] wherever possible. Secondly, ISDN has become widely available: 2-4 channels at a speed of 128 - 256 Kbps are compatible with the bandwidth that should be available over the research packet networks. MICE therefore aims to provide access and coordination for conference rooms or workstations using H.261/H.221 codecs operating both over research packet-switched networks and over ISDN.

Over the last few years, workstations have become increasingly powerful. Manufacturers have been determined to develop new applications, and multimedia has tremendous sales potential. Since codecs have been prohibitively expensive, most manufacturers' activities have concentrated on the Local Area Network (LAN), but they have recognised also the need for interoperation and standardisation. Two standards have emerged: the Joint Photographic Experts Group (JPEG, [6]) for still pictures, and the Motion Picture Experts Group (MPEG, [7]) for moving pictures. Both standards provide very good results at speeds above about 512 Kbps, and again specialised chip sets are becoming available for them.

Even without these special chipsets, current workstations are capable of providing codec functionality in software, albeit with reduced performance. Such software is often aimed purely at packet network users. There are users who want multimedia facilities with more modest video which can be provided software codecs.

In summary, the current situation requires facilities for users of both conference rooms and workstations, plus connectivity between the two approaches.

1.1 The MICE Aims and Participants

The MICE project, coordinated by UCL, addresses these requirements. It aims to provide multi-way integrated conferencing service between a number of European pilot sites, and link them to the appropriate communities in the US. MICE demonstrates a range of possible ways of interworking between the project partners who are using heterogeneous environments, rather than develop new systems or a special infrastructure. The intention is to interconnect up to half a dozen conference rooms, and up to a dozen workstations both by the ISDN and research networks, and to pilot multi-way conferencing services between them. The project has demonstrated multi-way conferencing between 11 sites in 7 countries within 8 months (though only four at a time), and will demonstrate multipoint with up to 4 simultaneous users from 12 sites in 7 countries within 12 months.

The initial partners in the MICE project are:

ULB (Belgium), INRIA (France), GMD (Germany), ONERA (France), Stuttgart U (Germany), NTR (Norway), Nottingham U (UK), Oslo U (Norway), SICS (Sweden), UCL (UK).

In addition, two National Research Network Operators (the German DFN and the British JNT) are associated with the project without funding. Both they and the CEC are planning to attach their own conferencing systems for interactions with ones provided in MICE. Others sites have asked to join the project, and the consortium would welcome further partners if the project is able to continue beyond the initial 12 month period.

The MICE project relies heavily on developments, expertise and facilities provided in other national and international projects; only the international aspects, the technology integration, and the operational coordination are financed by the CEC. Therefore, the project uses a range of existing conferencing equipment from other projects, which have a major bearing on the choice of subsystems used. With so many countries and public monies were involved, the project has, naturally, chosen to adopt international standards wherever possible.

2 Multimedia Conferencing and the Responsibilities of the Partners

2.1 Overview

As stated in the previous sections, voice, video and computer data are to be shared amongst a distributed group of users. In the MICE project the users are part of the research community. We have divided the requirements into four classes of service:

3 Multimedia Conferencing and the Responsibilities of the Partners

3.1 Overview

As stated in the previous sections, voice, video and computer data are to be shared amongst a distributed group of users. In the MICE project the users are part of the research community. We have divided the requirements into four classes of service:

- Multi-way video
- Multi-party Voice
- Multi-way shared workspace
- Multiplexing and Conference Management

This is not the only way that the services could have been classified - how they emerge is explained below. Since the existing research networks are acting as carrier networks - and these are largely packet-switched services - most of the conference activities are carried out over these.

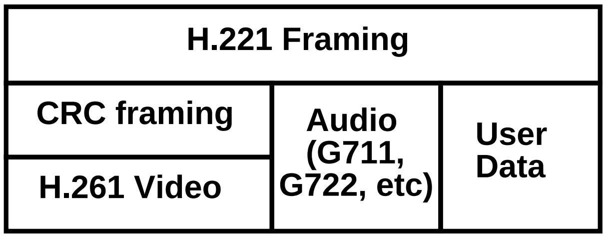
Full bandwidth video requires about 120 Mbps bandwidth; this would be costly, and could not be carried by the existing research networks. Therefore, compressed video has to be used. There are quite a large number of compression schemes available, but since MICE is committed to using international standards wherever possible, we have agreed to use the international H.261 standard [5]. This can be provided by special codecs, or at a reduced rate by software in some of the more powerful workstations. Both approaches will be pursued.

We do not expect to have enough bandwidth available to carry multi-way high quality video to all sites. For this reason, we will use a *hublet* approach for at least some of the conferences: all the video goes to one site (initally the CMMC at UCL), where it is multiplexed and retransmitted to all participants. The functions of the hublet are explained in detail in section 2.6. One function is to relay between different coding schemes - e.g between stations using codec hardware and reduced performance ones doing the codec function by software.

The basic philosophy adopted by MICE is that participants should be free to use any equipment, provided it conforms to certain external standards. At the same time, the project is committed to produce a reference implementation which is documented and supported by one or more named participating organisations.

3.2 Audio-visual Coding

The set of standards called p x 64 [5] cover video, voice, data; they include the H.261 standard for compressed video, several for voice, and H.221 framing for multimedia streams over serial links. Usually, the data and the voice are multiplexed within an H.221 stream - which is isochronous. The native H.261 coding for the video is not strictly isochronous - though there are delay bounds. At present most of the workstation-based software implementations treat the H.261 video and the audio separately. This allows us to route video and audio separately, for instance sending the audio via direct internet multicast and the video via a CMMC, or to prioritise the video and audio differently, so that sites on low bandwidth links only receive the audio. Moreover, because the H.221 framing standard has been designed for circuit switched environments, it is assumed that there is a capability to accept continuous streams of data; this framing is not very well suited for input to a workstation [8].



H.261 Video, Audio, and User Data combined in H.221 framing

Much more processing power is required for *encoding* video than for *decoding* it. Whilst only a powerful workstation could be considered to perform encoding, decoding requires an order of magnitude less. Where software implementations of H.261 exist, they can be performed using less processing power by omitting to perform a search for motion vectors - this results in slightly lower quality video, but brings the processing requirements within the bounds set by today's workstation technology.

For a point-point conference over a circuit-switched network, it makes sense to integrate the voice with the video. For multi-way conferencing over a packet-switched net, this is less suitable: voice is much more sensitive to delay and jitter than video, and it can be compressed to use much less bandwidth (between 4.8 - 16 Kbps). By using silence suppression, and the fact that usually only one person speaks at a time, it is possible to service a large audience with voice for a modest outlay of bandwidth, provided multicast is used. The five most recent IETF (Internet Engineering Task Force) meetings have demonstrated that audio can be multicast to audiences worldwide (up to 500 sites in 5 continents at the last one); these broadcasts have also demonstrated that some networks and routers may suffer from very bad delay variation.

It is possible to engineer the packet switched networks to ensure reasonable bounds on delay and jitter in the research networks at our disposal [9]. It is also possible to implement voice relays which mediate between sites with different coding schemes - often heavier coding is adopted because of the availability of reduced communication bandwidths. While the standard method for encoding voice is pcm, there are several coding standards for reducing its bandwidth; examples are *CELP* [10], *ADPCM* [11], *LPC* [12] and *GSM* [13]. We expect that many of these will be used, but have not yet decided which we will be supporting.

3.3 Conference Rooms

A typical conference room will allow 5 - 10 people to participate in a meeting at any one site. There are several versions of these rooms in the participants' premises with differing levels of sophistication. They normally include at least two remotely controllable cameras (one for the participants and one for documents), a workstation (for camera control and for shared workspace activity), a number of microphones, several TV monitors or a TV projector, some small TVs for previewing, a control box (to control all the peripherals, though this could be done from the workstation). There is usually some device for packetising data from the cameras - though again this can be done in the workstation. For attachment to the outside world, there is usually a router; it is possible to attach the workstation directly to a LAN, and have the router function shared with other LAN workstations. To match the video data rates to those available to this project, there is normally also a hardware codec.

A diagram of a typical CR is given in Figure 2. Here, switching between cameras and displays is done by a video switch. This allows other multimedia information to be introduced into the conference, such as information from a video store. We show the multimedia information as coming in from a network via a router on a LAN. It is shown coming into the CR workstation, which acts also to strip the packet framing, pass the video to the codec and the audio to the speaker, controls the video switch, and multiplexes and packetises in the other direction. This type of installation is reasonably expensive, so that its usage is normally shared between different functions.

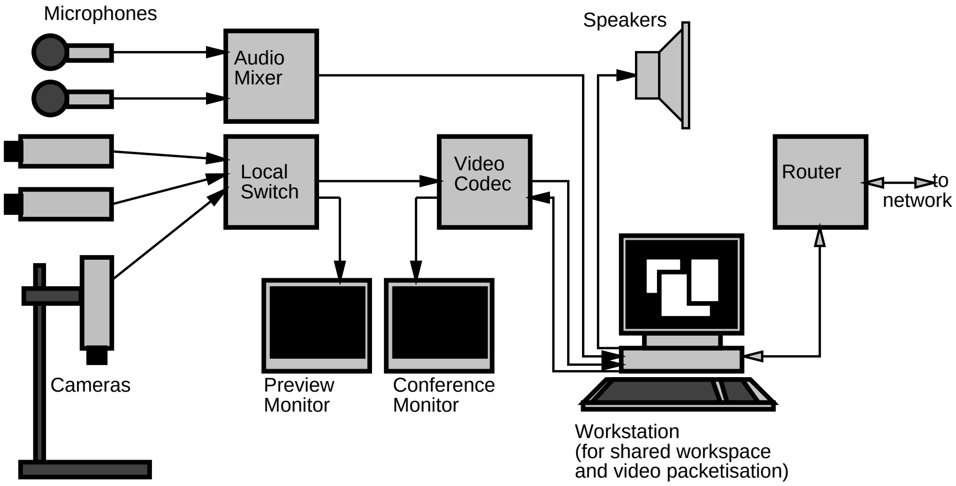
The MICE partners are using their existing conference rooms. In addition, GMD has the responsibility of providing a reference installation which could be copied by organisations which do not have existing facilities but wish to join the programme.

3.4 Workstations

There are currently three software video conferencing systems in frequent use in the internet - these are IVS from INRIA (a MICE partner) NV from Xerox Parc [2], and CU-See-Me from Cornell University. It can also be safely assumed that many workstation manufacturers have their own systems under development (such as the Cell-B video compression system from Sun Microsystems). Of these, IVS is currently the only system to conform to international standards (in this case H.261, which allows it to interwork with hardware codecs). IVS and NV are tailored to work over IP multicast, which means that they can be deployed in a multiway conference without the need for a central multi-point hub. However, in a large conference, if many sites have video transmission capabilities, then clearly both the receiving workstations and the networks themselves will become overloaded if every site multicasts to all others without some form of coordination. There is no such problem with audio, as typically only one person will be speaking at once.

We also expect to accommodate workstations with reduced functionality which operate only with audio, or in which the video function is restricted to document entry and display. For this functionality, it may not be necessary even to include the camera/frame grabber, though some form of local scanner (e.g. a facsimile device interfaced to the workstation) would be helpful.

A diagram is given in Fig. 3. The MICE partners use their existing workstations, and their audio/video software is largely derived from INRIA's IVS. INRIA and Stuttgart University have the responsibility of providing a reference installation which could be copied by organisations not having their own facilities but wishing to join the programme.



Schematic of Conference Room with Workstation providing media packetisation for transmission over a packet network

The audio compression and decompression is shown being performed in the workstation. Although the codec usually provides this functionality, compatibility with other sites is made simpler by using the workstation to provide audio functionanlity.

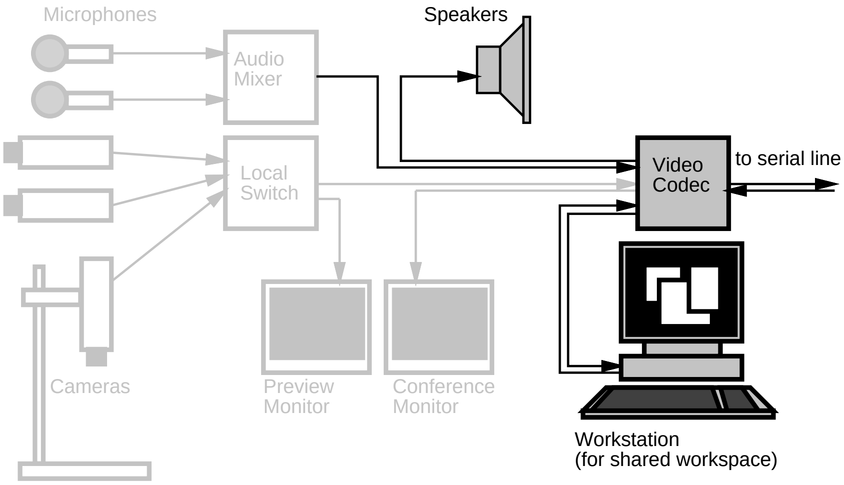


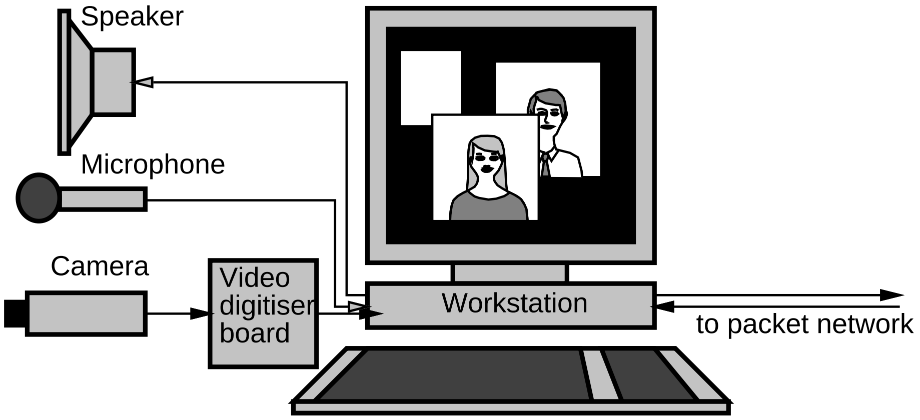
Figure 2: Schematic of Conference Room with Codec providing serial line framing for audio, video and shared workspace data.

(unchanged connections are shown in grey)

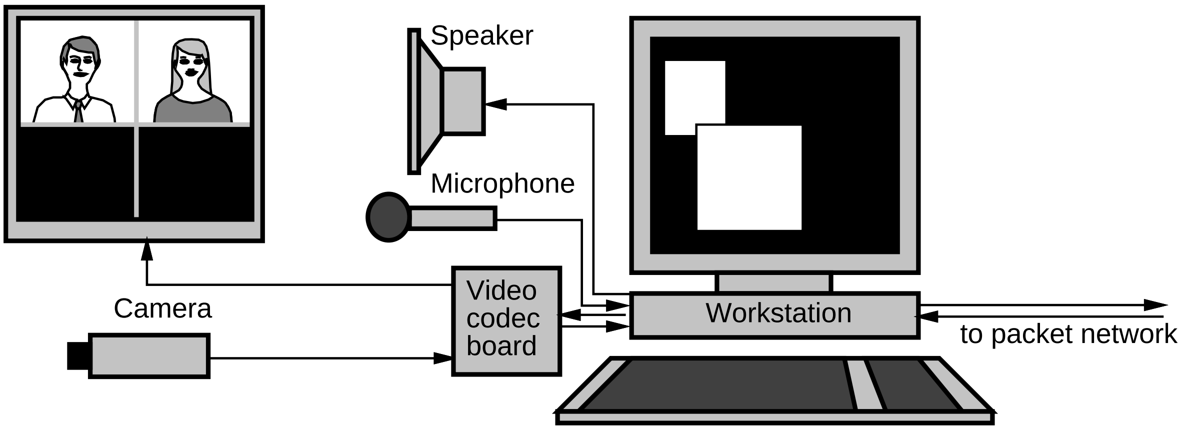
3.5 Shared Workspace

There are many methods of supporting shared workspace activity. The partners are already supporting four different systems from other projects. The important activity here is to support a broad range of application packages. In some of the end user applications, it is necessary that there be a large number of sites watching and listening to a conference; in others, a small number of groups must interact strongly. Another aspect of shared working is the need for conference control. MICE pursues two approaches; a *centralised* and a *distributed* approach. The centralised one will be implemented first - mainly by UCL and Nottingham University. A version of the *CAR Conference Control System* [14] written at UCL has been ported to Sun RPC, Solaris and tcl/tk. This will be used to interface with a booking system for resource allocation, to provide a directory service, conference setup, membership control and basic floor control functionality. On top of this platform, various applications can be intergrated, including Shared-X [14]. In addition, *XConfer* [15] tools from Xtel/Nottingham University might be integrated into the CAR platform.

This functionality is added to the controlling workstation of the conference room of Fig. 2 and the workstation of Fig. 3.



Multimedia Workstation performing the codec function in software



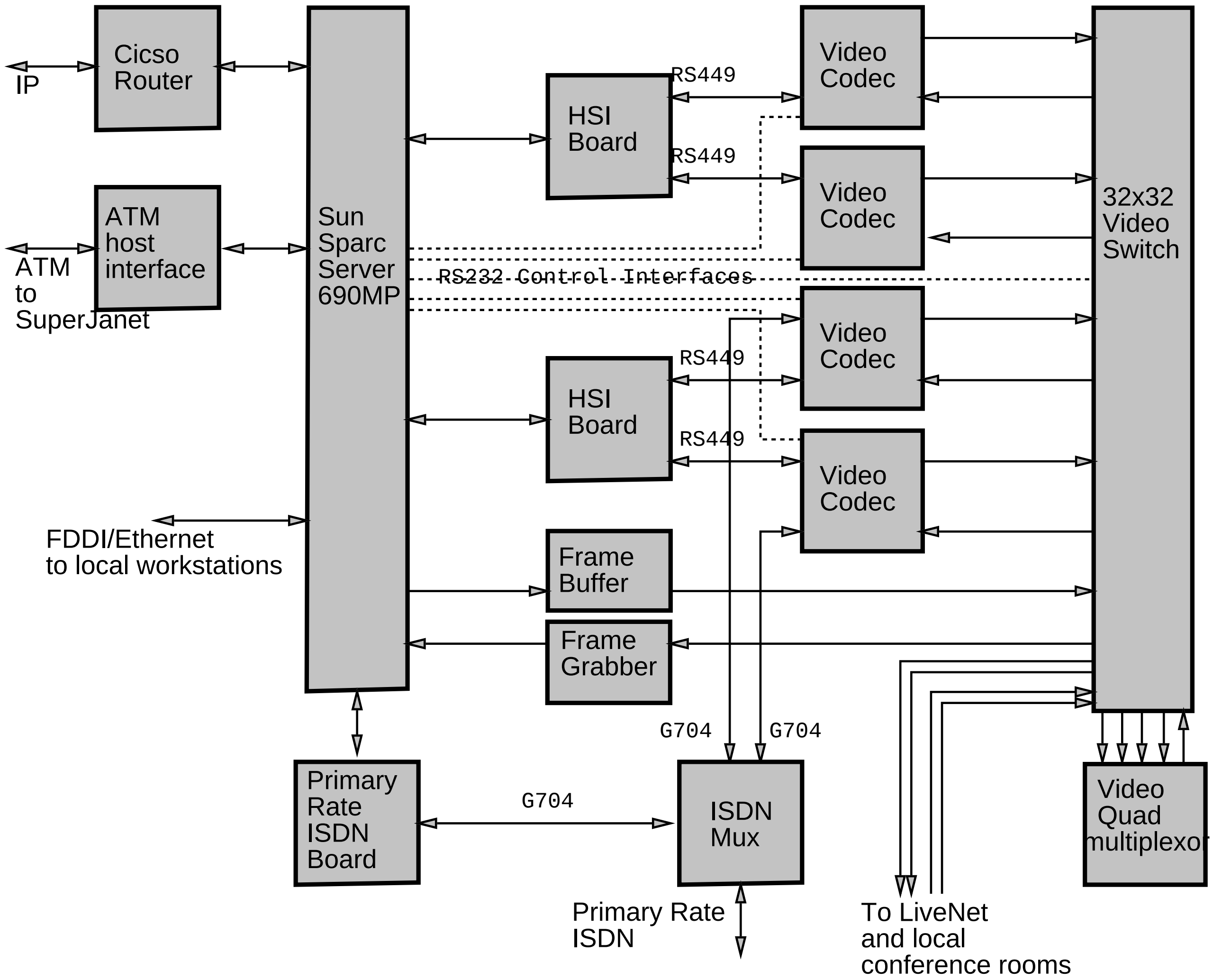
Multimedia Workstation performing the video codec function in hardware and the audio codec function in software
(instead of a separate monitor, a video overlay board can be used to overlay the video output from the codec board over a window on the workstation screen)

3.6 Conference Multiplexing and Management Centre (CMMC)

The current trend in the internet community is to provide multimedia services by applications that multicast directly to all sites in the conference, with no centralised conference control or multiplexing centre. Whilst we agree that an entirely distributed system is desirable, there are many reasons why we believe that many multi-point conferences will require a centralised *Conference Management and Multiplexing Centre (CMMC)*. These reasons include:

- the mechanism for allowing communities to come in from different networks (e.g. EUROPANET, UK-US Pipe, EBONE and ISDN);
- a way to relay between different protocol and coding standards;
- the incorporation of stations with reduced facilities (e.g. Basic Rate ISDN), video workstations without hardware Codecs;
- a centre for multiplexing video and voice;
- a centre for conference reservation, resource control and management;
- a centre for ensuring multicast of the shared workspace data, and possibly the voice, to all conferees;
- a centre of expertise - remote users only need to know how to connect to the CMMC, and not how to interwork with (possibly incompatible) remote sites - the CMMC hides this extra functionality.

A schematic of the CMMC, as developed by UCL, is shown in Fig. 4.



Schematic of the MICE Conference Management and Multiplexing Centre
(for simplicity, software codecs and hardware N-1 audio mixing are not shown)

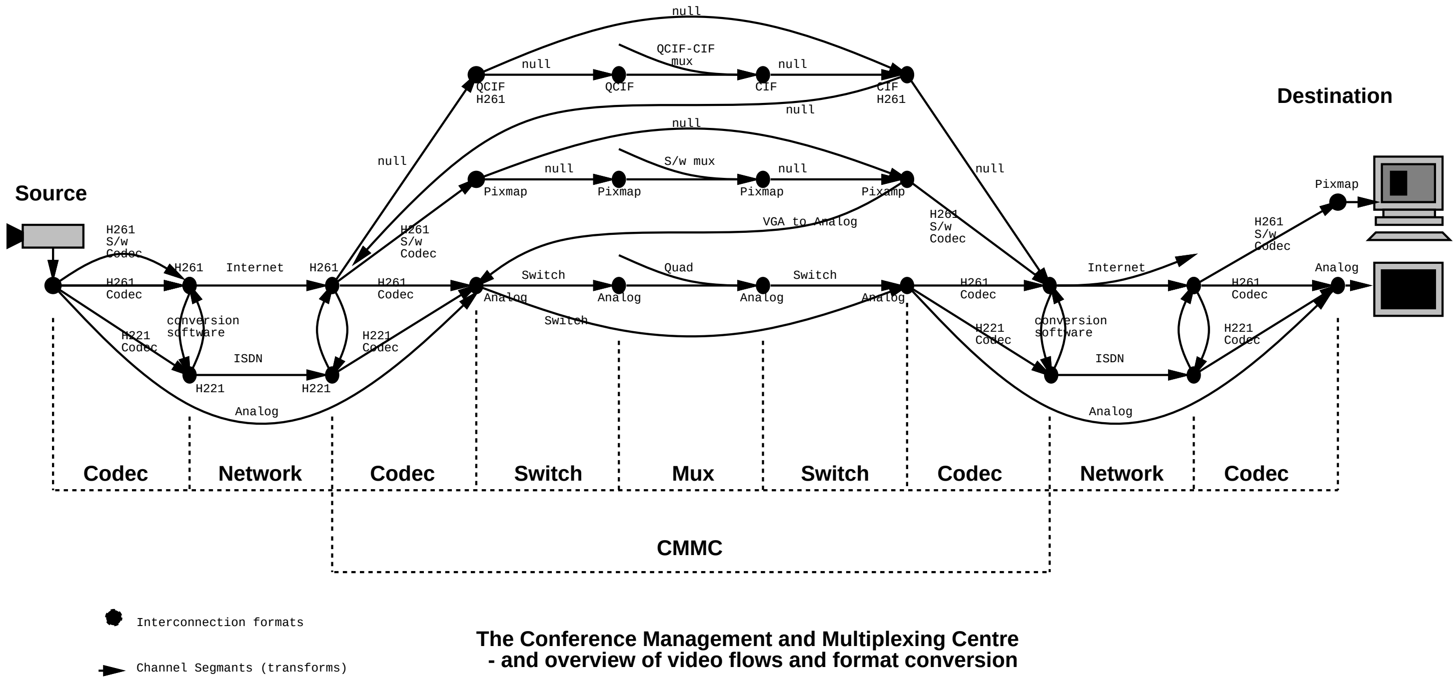
Video

A representation of video flows through the CMMC is shown in Figure 5. Broadly speaking, a video flow originates in a *remote site codec* (either hardware or software). It may then undergo a *protocol conversion*, before being *transmitted* over a network (either packet based or ISDN). Depending on the circumstances of the conference, it may or may not then require *multiplexing* or *format conversion*. Assuming it requires multiplexing, this can be done in one of three ways:

- Option 1: Decode to analogue video, quad multiplex in hardware, and recode again in hardware.
- Option 2: Decode in software, multiplex digitally, and recode.
- Option 3: Multiplex up to 4 H.261 QCIF video streams into one H.261 CIF video stream in the compressed, coded domain by parsing, modifying and combining the compressed H.261 streams.

Clearly, only option 3 can be done without any loss of quality, although this option places certain restrictions on the remote site codecs.

Some *protocol conversion* may also be required before *transmission to the receiving sites*, which may not all be on the same type of network (for example, some may be on ISDN links, some on packet networks), and may not all have the same restrictions (some may be using software codecs, some may have hardware codecs on high speed networks), and so a single incoming video flow may be coded several times for distribution. Transmission to remote sites over packet networks will usually be via *multicast*, even if there is only one receiving site, as the Mbone [16] should be configured to use the paths best suited to real time traffic, whereas the usual service traffic routes may not be.



The Conference Management and Multiplexing Centre - and overview of video flows and format conversion

The function of the *video multiplexor* is to combine the video into quadrants. The connection with the UCL Conference Rooms can be direct analogue via the video switch. This can also connect on to the other sites of the University of London over *LIVEnet* [17], and also to *direct broadcast satellite* if required.

Which set of the codec/mux/codec paths is chosen for a particular conference through the CMMC will depend on a number of factors. Some examples are:

- Where a number of conference sites require high quality, high data rate video we would normally decode in hardware (option 1). At high data rates, any lag introduced by the decode-mux-code cycle is minimal. Option 3 may also be possible, but it should be noted that the resulting data rate will be the sum of the data rates of the incoming streams, and so they must total less than 2Mbps to stay within the H.261 specification. The maximum rate we can perform 4xQCIF to CIF combination in software is not yet known.

- Where a remote site is connected over ISDN, it will usually be simpler to allocate a hardware ISDN codec than to perform the necessary H.221 protocol conversions in software, though this could be done, if required.

- Where a remote site is connected via a software codec that does not use H.261 video compression, we have no option but to decode the stream in software. The simplest path is then to multiplex resulting digital images in software using option 2. However, if the other conference sites dictate that the hardware decode/mux/code path be used we can use a VGA frame buffer to perform digital-to-analogue conversion before multiplexing, and a fast video frame grabber to perform the reverse path compression.

At present the video information from the US codecs will be in a format incompatible to that of the European video, and so must be passed to a particular codec.

Audio

In principle, the distribution of audio to conference participants is somewhat simpler than the distribution of video, since audio compression and decompression is well within the capabilities of even low power workstations, and even (uncompressed) PCM audio at 64Kbps is within the capability of many networks. However, this situation is not quite as simple as may be assumed at first sight. In almost all conferences, audio is the single most important communication medium. It is more sensitive to delay than video (assuming that lip synchronisation is not an absolute priority) if two way communication is required, and is much more sensitive to jitter and packet loss than properly packetised video. This may be surprising at first, but in the face of high packet loss, video can compromise quality and resort to pure *intra compression* of pictures which (from a user's point of view) is relatively unaffected by packet loss. Audio has no such option. Recent tests of the Ebone UK-Paris and UK-Stockholm links show around 20-30in largely incomprehensible audio.

Generally speaking, audio should be multicast directly to all receiving sites from the speaker's machine. Assuming the Mbone (Multicast Tunnel Backbone) has been configured to utilise the best quality links, regardless of their geographical routing, this will subject audio to the minimum delays and packet loss. There is less of a format conversion problem than with video, as most workstations can perform the decoding of whichever format the originator chooses in software.

However, the CMMC will not be completely unused for audio distribution for several reasons:

1. Remote sites may be connected by ISDN using hardware codecs or video phones which combine the audio and video data into a single H.221 serial stream. The CMMC will separate this audio from the video stream, and then multicast to directly to the other receiving sites.
2. A remote site may participate using only a normal telephone to connect to an ISDN line at the CMMC, and then receive just the audio from the conference. This may be used to supplement the shared workspace where the loss or jitter on the packet network is too great for comprehensible audio transmission.
3. Some sites may require very low bandwidth feeds in order to overcome local bandwidth limitations. Although such feeds would ideally be provided from the nearest remote site with adequate bandwidth, it is much easier for such a service to be provided automatically from the CMMC. As with other CMMC functionality, this will be distributed or replicated in the long run.

3.6.1 Shared Workspace

The shared workspace activities of multimedia conferences should not require the facilities of the CMMC. Shared applications can be divided into several categories:

1. **Centralised with floor control.** An example is Shared-X, where the display of an existing application is replicated, and control can come from any of the remote sites, but only from one at any time.
2. **Centralised without floor control.** An example is the distributed sketchpad developed for the CAR project by FhG-IGD Darmstadt, where a central server communicates with remote clients at each site, allowing all participants to draw simultaneously.
3. **Distributed with explicit membership.**
4. **Distributed with implicit membership.** An example is the distributed whiteboard from LBL, where any user running a whiteboard on a particular multicast group receives the history to date, and can make modifications.

All except the last option require a list of current members of the conference in order to initiate communication, start remote clients, or replicate the display. Some applications (category 1) also require some form of floor control, which may either be provided on a per application or a per conference basis. This implies a need for some form of conference control system.

Conference Control

It should be clear that not all conferences need an explicit conference control system, and not all conferences need to utilise the services provided by the CMMC. If video, audio and shared workspace are all provided in an entirely distributed fashion (as with the LBL whiteboard and IVS), no additional conference control system is required, since membership and participation are handled by the multicast group mechanism of the Mbone. If, however, any of the additional video, audio or shared workspace functionality outlined above are required, then some additional form of explicit conference control is required.

The conference control system for the MICE CMMC will be based on the CAR Conference Control system from UCL. The basic conference control functions are:

- Creation and deletion of conferences
- Joining and leaving conferences
- Introducing applications to the conference
- Floor control
- Control of conference video channel(s)

Currently, this functionality is provided by a centralised Conference Server, remote Conference Management Applications, and a number of applications that communicate with the conference server using remote procedure calls from the CAR library. In the long run we hope to substitute a more distributed and fault tolerant conference control model to be known as the Conference Communication Channel Protocol (CCCP), but this is not likely to occur within the timescales of the existing project.

This basic conference control functionality provides the information for applications to locate the remote machines for running clients, and a simple floor control scheme for those applications that require it. It also provides information for video and audio servers to:

- allocate channels and codecs where required;
- give control over the multiplexing of those channels where this is not
- handled by applications such as IVS;
- to perform bandwidth and quality of service reservation where networks
- provide such functionality.

It should be stressed that the conference control system itself does not provide these resource allocation functions; this is handled by the resource allocation part of the CMMC. This split of functionality keeps the conference control system minimalist, so that it can also be used outside of the CMMC when resource allocation is not required. The Conference Control system and the Booking and Resource Allocation systems are shown in Fig. 6.

Booking and Resource Allocation

Where a conference uses hardware codecs (whether at the CMMC or at remote sites), some form of resource allocation is required to ensure that each codec produces video of a format and data rate the receiver(s) can decode. Where codecs communicate with each other over serial lines, the H.320 suite of protocols (H.221, H.230, H.242) ensure that upgrades from a base (audio only) service are negotiated. Where such codecs communicate over packet networks, these protocols are inappropriate, and are either stripped off, or not added in the first place. Thus, codec configuration must be controlled out of band.

Such codec configuration is best performed in the light of knowledge about the interconnecting networks, with access to network management (QOS reservation) where available, and with access to data about all the other proposed participants in the conference. Such information and services will eventually be provided by the CMMC's *resource allocation functionality*. Most importantly (from the point of view of the CMMC), resources such as codecs within the CMMC will be limited, so it is important to provide booking facilities to ensure that a pre-arranged conference will have the facilities it requires.

It is, of course, important that impromptu conferences can also take place, but this can be done by allowing ad-hoc bookings providing the facilities required are available. This places certain requirements on the timeliness of conference booking data, and on the ability to reconfigure conferences that are underway. We may, for instance, be using a particular codec to communicate with a remote site over a packet network, when another user attempts to join the conference over ISDN. Some of the high quality codecs would be our first choice for either network, but our only choice for ISDN, and thus a re-configuration and change of codec for the original user are our only option if we are not to refuse access to the ISDN user. If the conference had been pre-booked this occurrence would have been avoided, and indeed an impromptu conference would be locked out rather than trigger a reconfiguration of a pre-booked conference which is underway.

When performing such resource allocation, we use a concept of *minimum acceptable quality*. Users specify a minimum quality of service (in terms of bandwidth, CIF vs QCIF, and various other parameters), and if those parameters and those of any other pre-booked conference cannot be met, the booking is refused. If spare capacity is available, the user may find that when the conference takes place, the quality of service is actually better than specified, but those factors under the control of the CMMC will not be worse than requested.

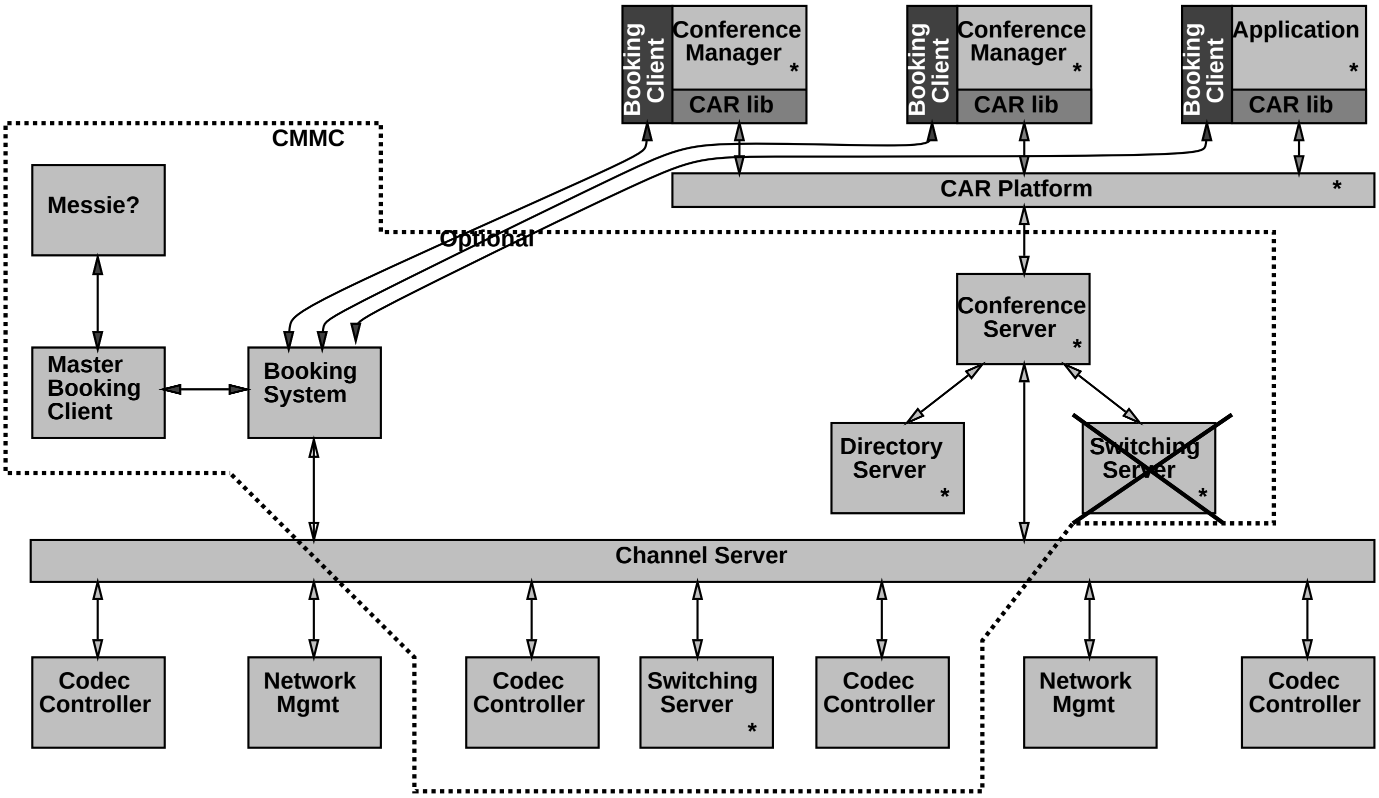
Many factors, such as packet loss on unreserved internet links, will, of course, be outside the control of the CMMC. The CMMC will keep statistics about performance of these links, which can be used in the future to ensure that bookings are only made which are likely to give acceptable performance.

Many of the functions do not need to be centralised, and it is planned that the CMMC be replicated eventually. Which functions can be replicated, and how CMMCs should interact, is currently investigated. One activity which will clearly be replicated is that of audio relay; we expect that there will be a deployment of facilities which can perform this function attached to, or as part of, many of the National networks.

4 Infrastructure

No special communications facilities have been installed specifically for MICE, although some links are being upgraded by the participants. There are currently three International networks involved: EUROPANET, EBONE and the US Internet. In addition, the CMMC supports communications over ISDN. In each country there are National research networks, and these are connected to the International ones.

The international connections (with the National network in parentheses) are:



Overview of the MICE CMMC, showing the intregration of the existing CAR software (*)

- *EBONE*: France (Renater), Norway and Sweden (Nordunet), UK (JANET)
- *EUROPANET*: Belgium, Germany (DFN), UK (JANET)
- *Internet to the US*: France (Renater), Sweden (Nordunet), UK (JANET)

The EUROPANET links are running at 2 Mbps; the EBONE links will be at about 1 Mbps between France and Scandinavia, but at significantly less to the UK. The Internet links will mainly be running at 1 - 2 Mbps. There 30 64 Kbps channels of ISDN (Primary Rate) to UCL. However, it should be noted that the EBONE and Internet links also have substantial normal IP traffic; we expect that the EUROPANET links will be relatively lightly loaded during 1993. There are discussion on the direct linking of EBONE and EUROPANET. We expect to be able to give further information at the conference.

5 Other Research Activities and their Relationship to MICE

The MICE project has been funded at the modest level of 54 man-months during its first year. It would be quite impractical to undertake this ambitious project with that level of resource if it was not based on very extensive concurrent other activities undertaken by the MICE partners. This is not the place to describe all such work, but the most relevant activities should be mentioned.

5.1 The Visual/Audio Transport Activity in the IETF

There has been an ongoing activity largely under the auspices of the US DARPA, in piloting large-scale conferencing with workstations based on LBL's Visual Audio Tool (VAT) [18]. The work has involved many groups, but has been centred in Van Jacobson's group at UC Berkley. This has concentrated on two areas: one is interactive voice conferencing assisted by low definition video using software Codecs; the other is the large scale broadcast of proceedings of meetings, with limited interactive feedback - mainly voice. The system is based on the use of multicast in the Internet [16], and incorporates specific coding schemes from a number of groups. For example, there are versions of video coding from both Xerox Park and INRIA [1], and voice coding from SICS and CWI. The VAT system is not the only one in which these subsystems are used, but it is the most widely distributed. It is the INRIA subsystem which we use as one of the basic services in MICE.

This activity has used the US Internet and its international links mentioned in Section 4. This includes the Internet pipes, the EBONE network, JANET and NORDUNET. These will all be used in the MICE project. We are using the VAT system for conference coordination while building up the general MICE facilities.

5.2 The UCL and GMD Conference Room Activity

Both GMD and UCL have been involved in limited pilots using conference rooms. The GMD work has been undertaken in the context of tying together the locations of the GMD; it has so far involved mainly installing such rooms in four GMD sites, connecting them together with the high-speed VBN network, and integrating in also some workstations with multimedia facilities.

UCL has been involved in a number of relevant projects. For the last five years, it has been one of the nodes in an analogue video conferencing network (LIVEnet [17]), covering London University. This currently has ten nodes, and provides multi-way video conferencing in which all sites can participate simultaneously. It is based on circuit switching, with integrated data facilities. It has a relatively sophisticated switching hierarchy, which allows the introduction of other networks and data sources. Thus it is linked also into a Broadcast Satellite network for beaming out educational television, into the US DARPA system described below, and other research systems.

UCL has also been participating in the SIMNET developed under DARPA auspices for multimedia conferencing over the Internet - particularly in the context of distributed simulation. That system currently operates with Codecs at about 256 Kbps over the Internet, but requires resource reservation at the routers. There are some 15 sites currently in that system. Multi-way conferencing is achieved by a hublet approach. The different streams come into Codecs at a central hub, where they are decoded into the analogue domain; there they are combined in analogue, and then recoded and sent out in a multicast manner to the participating sites. At present these service use in-band audio, since these functions are provided in the Codecs, with the audio mixed in the hublet.

Several of the MICE partners have been participating in the workstation-based systems mentioned in Section 5.2. Clearly, these can also be piped to Conference Rooms.

5.3 Conference Control and Shared Work Space Activity

There are many shared workspace systems. One such system is the BBN-supplied, DARPA-financed, Distributed Simulation system of Section 5.3 (MMCONF) [19]. Others have been developed under the CAR system using a form of Shared-X [14]. It is a form of the CAR system that will be combined with the XConfer system from Nottingham University/Xtel that use a generalised mechanism for sharing the applications, and will be adopted for the early stages of MICE Shared Workspace and Conference Control. It is a centralised system; SICS and INRIA are developing a more distributed system, which we hope to use later in the project. All these allow any actions on one keyboard to be viewed at all the other sites. By providing appropriate data communications facilities, or pre-transmitting the relevant files, it is possible to introduce large databases into the conferences while maintaining an interactive performance.

5.4 Conferencing Booking Infrastructure

With this joint use of facilities for National projects and for MICE, we expect considerable difficulties in booking conferences, linking in participants, in ensuring privacy, and in configuring network resources. We expect that one of the tasks will be to facilitate this style of activity by extensive use of X.500 Directory Services - where the PARADISE [20] infrastructure could provide considerable assistance - and an email-based calendar system and file store (MESSIE) [21].

6 The Status of Standards

We have already stated that we expect to use H.261 for video coding. We know that the H.221 multimedia stream encoding is not appropriate for packet-switched operation, but believe we can overcome any inherent problems. We have also stated that we doubt that the standards for in-band audio coding are appropriate, because of our potential need to use a different network topology for voice than for video; some of the distribution sites may require only the audio, others may require reduced bandwidth audio with no video. We intend to use international standards again for audio - but these may come from a different domain such as cellular telephone. There are emerging standards for Multipoint Control Units (MCUs) required for multiple conferences; these are supplementary to H.221/H.261. We will have to investigate the assumptions behind these standards to determine how relevant these are, but will not be using them at first. Some Internet Standards developed currently under the auspices of the Audio Visual Transport Working Group (AVT-WG) of the Internet Engineering Working Group (IETF) seem more appropriate [22]. These determine framing standards, but describe the other elements in such a way that they can be passed through different networks and be reconstituted at the destinations. They also allow the specification of lifetimes for the elements, so that they can be destroyed if they are going to arrive too late at the destination - or alternately so that the priority of the elements can be changed at intermediate nodes. Of particular importance is that the coding permits different protocol versions to be used, this permitting heterogeneous implementations. Other features supported are flow control, end-to-end real-time control, media source and destination, conference announcement, and quality of service announcement. It is significant that the working group includes representatives from service providers such as AT&T.

Another set of standards are being developed by other IETF groups in Remote Conferencing (REM-WG) [23] and Conference Control. These were not available for the early stages of MICE, but the discussions have been followed, and the project's requirements have been put forward. There is already a concrete proposal that MICE be associated with a US NSF project, which would use the IETF Standards. The standards being developed in the working groups include the following: a Connection Conference Manager; an Applications Programming Interface; mechanisms for coding, transport, multicast and flow control; network resource management; multicast address administration; a connection and configuration manager, conference systems management.

7 Conclusions

Multimedia conferencing is currently at a very exciting stage. There is a plethora of activities, and a large number of different approaches. There is an emerging set of Standards - many truly appropriate, many starting to lead to convergence of approaches. The research community has a different environment, and different requirements, from the commercial services being developed by the carriers. MICE represents an important new initiative, based on the best practices of the research community. The project takes maximum advantage of the networks provided by and for that community, setting a framework which could be operated on a pan-European basis. It is taking advantage of the best practices in the US as well, and provides a link between the European and North American activities. Since the technology is used to facilitate collaboration amongst the project partners, MICE will ensure that its final product is well-tested and usable. Since MICE relies on the networks provided for the research community, the project has pushed to improve those networks suffieciently to meet the community's needs. Finally, it should have an important impact on equipment suppliers, service providers, communications carriers and the research community itself. This should be achieved through adopting the following measures:

- to use both specialised and generalised hardware;
- to incorporate both high end facilities but also reduced capabilities;
- to use both the research networks and the ISDN; to use both specialised conference rooms and ubiquitous workstations;
- to start in the more technologically advanced institutions, but to extend to less advanced institutions and less prepared countries with packaged solutions;
- to have as users researchers developing the facilities, but also those from other disciplines and from the funding bodies like JNT, DFN and the CEC itself.
- to start deploying more distributed relay facilities for all the relevant services.

The project is an extremely challenging actiivity - and at the same time highly rewarding one. Since it crosses many institutional and national boundaries, it may well create a precedent in removing such boundaries in individual countries, on both sides of the Atlantic.

References

[1] T. Turletti, H.261 Software Codec for Videoconferencing over the Internet, Research Report no 1834, INRIA, Sophia Antipolis, France, 1993

[2] R. Frederick, nv - X11 videoconferencing tool, Unix Manual Page, Xerox Parc, 1992.

[3] D. D. Clark, Supporting real-time applications in an integrated services packet network: architecture and mechanism, SIGCOMM, 1992.

[4] The DSI Newsletter, published by Bolt, Beranek and Newman Systems and Technologies, Cambridge, Mass, on behalf of DARPA.

[5] M. Liou, Overview of the px64 Kbit/s video compression standard for multimedia applications, Comm. ACM, 34, 4, pp 59-63, 1991.

[6] G. K. Wallace, The JPEG still picture compression standard, Comm ACM, 34, 4, pp 30-44, 1991.

[7] D. G. Gall, MPEG - A video compression standard for multimedia applications, Comm. ACM, 33, 5, pp 85-110, 1990.

[8] I. Wakeman, Packetised video - options for interaction between the user, the network and the codec, The Computer Journal, February 1993.

[9] D. Verma et al., Delay jitter control for real-time communications in a packet switching network, Proc. Tricom '91, pp 35-43, 1991.

[10] I. Chen et al., A fixed point 16 KB//S LD-CELP algorithm, Proc. IEEE Int. Conf on Acoust., Speech and Sig. Proc., pp 21-24, 1991.

[11] N. Benveuto et al, The 32 kbit/s ADPCM coding standard, AT&T Tech. J., 65, 5, pp12-21, 1986.

[12] B. S. Atal et al., A new model of LPC excitation for producing natural sounding speech at low bit rates, Proc. IEEE Int. Conf on Acoust., Speech and Sig. Proc., pp 614-617, 1982.

[13] P. Vary et al., A speech codec for the European mobile radio system, Proc. IEEE Int. Conf on Acoust., Speech and Sig. Proc., pp 227-230, 1988.

[14] M. J. Handley et al., Multimedia conferencing: from prototype to National pilot, Proc INET '92, pp 483-490, Internet Society, Reston, VA, USA, 1992.

[15] N. Cook, N. & G. Lunt, XConfer - Dynamic Desktop Conferencing, Proc European X Users Group Conference, 1992.

[16] S. E. Deering, Multicast routing in datagram internetworks and extended LANs; ACM Trans Comp. Sys., 5, 91.

[17] P. T. Kirstein et al., Recent experiences with the University of London interactive education network, Comp. and Contr. Eng. J., 5, 5, pp 119-128, 1991.

[18] S. Casner & S. Deering, First IETF Internet Audiocast, ACM SIGCOMM Computer (Communications Review), Vol.22, No.3, July 1992.

[19] BBN staff, Multi-media conferencing user and operations manual, Bolt, Beranek and Newman, Cambridge, Mass, MDA903-89-C-020, 992.

[20] D. Goodman, PARADISE: the COSINE X.500 pilot service, Comp. Net. and ISDN Sys., 23, pp 111-114, 1991.

[21] M. A. Sasse et al., Support for Collaborative Authoring via Electronic Mail: The MESSIE Environment, Proc. ECSCW'93, Milan, Italy, 1993.

[22] H. Schulzrinne, A transport protocol for audio and video conferences and other multiparticipant real-time applications, Internet Draft, 1992.

[23] Y. Chang, Real-time multimedia communication architecture, IETF REMCON, 1991.

Mark Handley received his BSc in Computer Science with Electrical Engineering from UCL in 1988. As a PhD student at UCL, he studied novel neural network models and their visualisation. Since 1991, he has been a Research Fellow, working on the RACE CAR multimedia conferencing project and on MICE. His current research interests include Multimedia Systems (especially audio and video encoding and compression), Distributed and Heterogeneous Systems, HCI and graphics.

Peter Kirstein received a B.Sc in Maths and EE from Cambridge U, Ph.D in EE from Stanford U., and D.Sc from London U in EE. He worked at Stanford U, the Centre of European Nuclear Research (CERN) in Geneva, and the US General Electric, in Zurich. He is now Professor and Head of the Department of Computer Science at University College London. Professor Kirstein has been leading research projects in computer communications networks, telematic services, security and multimedia for over 20 years. His current research projects include the PARADISE project in piloting directory services, the PASSWORD and other projects in secured document services, the PODASAX project in ODA applications and databases, and the MICE project in multimedia services. Professor Kirstein is a Fellow of the Royal Academy of Engineering, the British Computer Society, the Institute of Physics, and the Institution of Electrical Engineering.

Angela Sasse has been a lecturer in the department of Computer Science at UCL since 1990. Her main areas of teaching and research are Human-Computer Interaction, Multimedia Systems and Computer-Supported Collaborative Work (CSCW), and she is involved in a number of Distance Teaching and Learning activities. Prior to MICE, she worked on the RACE CAR project, particularly on usability issues of multimedia conferencing systems.