ABSTRACT

This article is about the experiences of the authors in transmitting the proceedings of some events at IEEE GLOBECOM '96 in London, England, in the week of November 17–22, 1996. Live video and audio of all the events in the Churchill Auditorium of the Queen Elizabeth II Conference Center were captured and transmitted, in real time, as well as stored and transmitted later, for remote participants in three continents, over the Internet. Two independent systems were used simultaneously, one supplied by researchers from NTT Laboratories in Japan and the other by researchers from University College London. The former system is based on a server model of distribution, while the latter is based on the use of network-level packet multicast. Both systems employ compression algorithms, so the network capacity requirement in each case was on the order of 100 kb/s to 200 kb/s total, thus enabling remote participants without very high-end network connectivity to take part. Receivers only need software for a PC running most popular versions of Windows or a UNIX workstation to be able to receive either type of transmission, or to retrieve the recorded sessions from NTT Laboratories' servers. The multimedia transmission was carried over carefully engineered links that traversed many different subnet technologies, including point-to-point circuits, SMDS networks, ATM networks, and fast Ethernet switches. This was both to give a high level of assurance that the traffic would not experience too much interference from other traffic at the site and elsewhere, and to ensure very low packet store and forward delays. The system ran for four days continuously, and was generally very successful. In the future, it should be possible to have remote paying attendees.

Real-Time Audio and Video Transmission of IEEE GLOBECOM '96 over the Internet

Colin Perkins and Jon Crowcroft, University College London

This article is about the experiences we had in transmitting the proceedings of some events at IEEE GLOBECOM '96 in London, England, in the week of November 17–22, 1996.

Initially, staff from the Department of Computer Science, University College London (UCL) were asked to organize a mini-conference within GLOBECOM on the subject of the Internet. It was then decided to add a technical exhibition, and to consider transmitting the event over the Internet, at the request of NTT Laboratories and some other research institutions in Japan and the United States.

We quickly formed a list of the technical staff to be involved from BT, MCI, NTT, UCL, and UKERNA, set up an e-mail list, shared computer accounts, and established basic Internet service provider relations and connectivity.

Live video and audio of all the events in the Churchill Auditorium of the Queen Elizabeth II Conference Center were captured and transmitted, in real time, as well as stored and transmitted later, for remote participants in three continents, over the Internet.

Two independent systems were used simultaneously, one supplied by researchers from NTT laboratories in Japan, and the other by researchers from UCL. The former system is based on a server model of distribution,[1] while the latter is based on the use of network-level IP packet multicast.

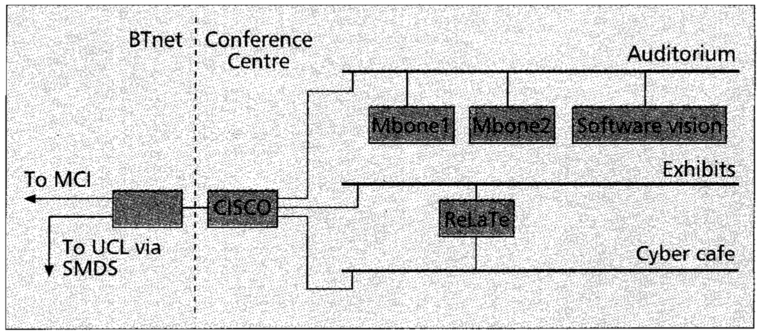
Both systems employ compression algorithms,[2] so the network capacity requirement in each case was on the order of 100 kb/s to 200 kb/s total, thus enabling remote participants without very high-end network connectivity to take part. Receivers only need software in order that a PC running most popular versions of Windows or a UNIX workstation be able to receive either type of transmission, or retrieve the recorded sessions from NTT Laboratories' servers.

The multimedia transmission was carried over carefully engineered links that traversed many different subnet technologies, including point-to-point circuits, switched multimegabit data service (SMDS) networks, asynchronous transfer mode (ATM) networks, and fast Ethernet switches. This was both to give a high level of assurance that the traffic would not experience too much interference from other traffic at the site and elsewhere, and to ensure very low packet store and forward delays.

The system ran for four days continuously, and was generally very successful. In the future, it should be possible to have remote paying attendees, although reliability would become an important question in such instances. It has been suggested that receivers could *lease* a key to decrypt an encrypted video

1 In some senses this is analogous to the use of an MCU In the ISDN/H.320 videoconferencing world, except that here we are packet-based, not circuit-based. Another system like this is the CU-SeeMe reflector based multi-party video conferencing tool.

2 Both systems use the ITU H.261 algorithm for video compression, which is designed for digital video-telephony, and is suited to typical Internet capacity, as well as being reasonably amenable to implementation at the sender side in software. As regards the receiver side, it is very straightforward to implement. The main difference between Internet implementations of H.261 and ISDN based ones is that Internet ones are typically in software, and do not employ the expensive H.221 framing protocol.



■ **Figure 1.** *Planned network layout.*

stream, much along the lines used for satellite and cable TV pay-per-view systems. Reselling of the material would not be prevented, but at least live virtual attendance would be limited, and by changing encryption keys regularly, the growth of the community able to "spy" on the event could be constrained. This will have to be investigated by the IEEE and any other organizations interested in using this technology in anger.

There were three networked areas in the conference center: exhibits, a cyber cafe, and the main auditorium. Each of these areas is served by a separate Ethernet, and these are linked via a CISCO router in the basement. Access to the Mbone is via an external link from the CISCO. This is illustrated in Fig. 1. Originally, the Cisco router was going to terminate tunnels from the BTnet machine, `news-feed2.bt.net`, but due to interworking problems of the versions of DVMRP, the multicast routing protocol, we ended up setting up an Mbone tunnel from the Mbone1 machine to the mbone machine within BTnet.

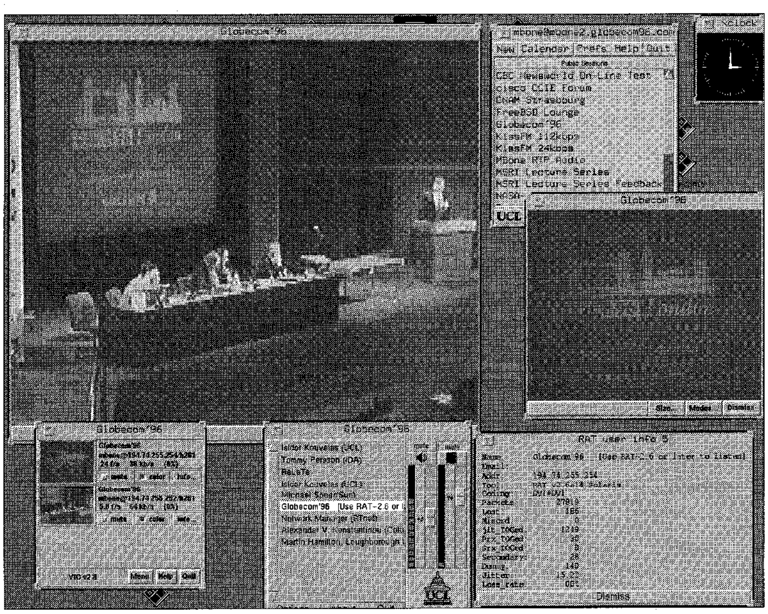
UCL provided four machines: two for the multicast of the conference sessions, and two to provide a demo and multicast routing for the exhibit/cafe areas. Two systems were be provided by NTT to broadcast the conference sessions using the SoftwareVision system.

The machines provided by UCL functioned as follows:

- Mbone1 — Main mrouted and first video feed
- Mbone2 — Second video feed and audio.
- ReLaTe — ReLaTe (Remote Language Teaching) demo
- ReLaTe — Mbone display machine for main exhibition area

The ReLaTe machine was situated in the exhibit area, adjacent to the cafe. Originally, we planned to have two Ethernet interfaces and connect to both networks. In normal use it received multicast traffic destined for the exhibit/cafe networks via a tunnel from Mbone1. In the end we reverted to using the CISCO to carry out local copying.

The Mbone1 and Mbone2 machines were situated in a translator booth, overlooking the main auditorium, with good a/v cabling access, next to a riser from the basement, where the cable plant and router arrived. They connected logically to the Mbone by a tunnel through the CICSO to BTnet, BT's commercial Internet service. From BTnet, we implemented two multicast tunnels to MCI and UCL.



■ **Figure 2.** *A captured screen of a session.*

ETHERNET

The auditorium Ethernet was required to terminate in the translator booth area. We used *at least* four access points here.

In addition, Ethernet was provided to the Pickwick Suite for the cafe and exhibit networks. The UCL exhibit was *originally* intended to be sited so that the ReLaTe machine connected to both cafe and exhibit networks. In the end, we separated the ReLaTe and Mbone machines onto their own network with the CISCO replicating packets to it from the Mbone tunnel machine on the transmission network.

AUDIO/VISUAL

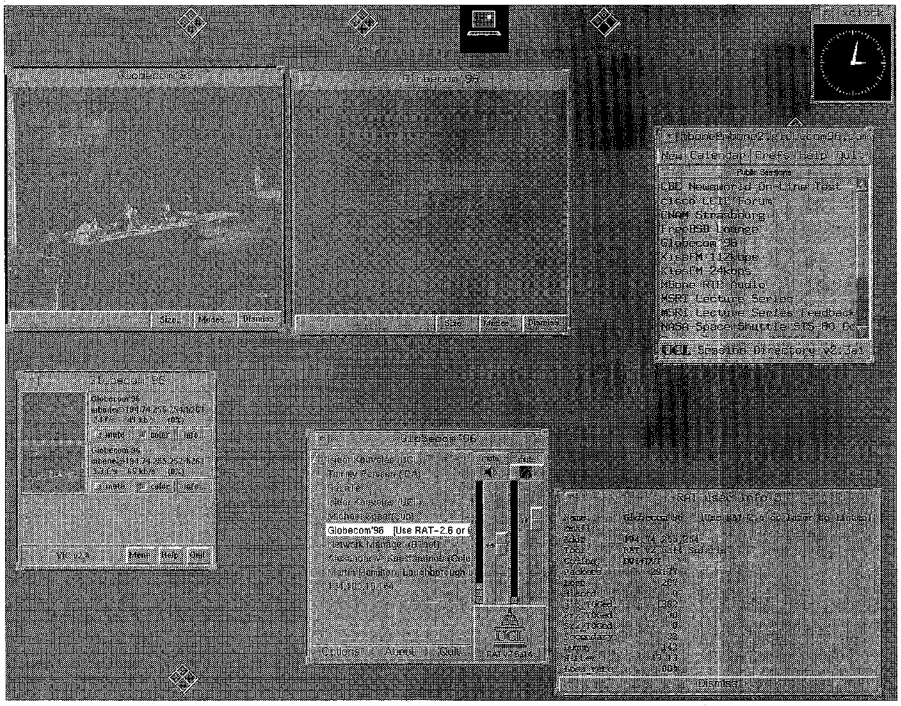
UCL and NTT provided two video cameras each, one situated in room 1/32 and one in the Churchill gallery. Cables were provided by the conference center to connect these cameras to the machines in room 1/35, and to provide an audio line level from the auditorium microphones via a mixer, and from the mbone workstations back to the auditorium PA system. In the end, we provided a PAL/NTSC scanline convertor so that we could share input from the Slide camera that UCL had. There was a lack of communication between the Mbone booth, and the location of the camera operators and the audio engineer's booth, which caused some problems. Wireless communication tools (radio/phone/IR links) would have made things far more efficient.

OTHER REQUIREMENTS

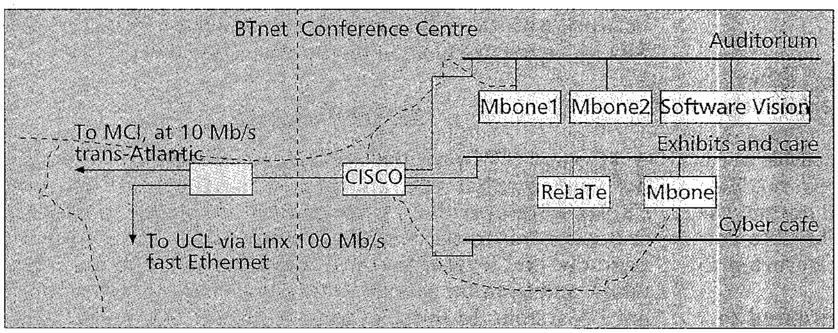
A telephone which accepts incoming calls and can make international calls in room 1/35 for use by the Mbone team was essential. Not only could we reach wide area network (WAN) operational staff, but we could dial out to Internet Protocol (IP) dial-up sites, and thence test what was wrong when our network connectivity was down. Later, it transpired that we could even dial in to the NTT site with a laptop equipped with a 28.8 kb/s modem, retrieve live video from there, and gain some idea of the quality being received at the remote site! See Figs. 2 and 3.

LESSONS

There were a number of lessons that were gained during this operation, and we outline them in the following subsections in a bottom-up manner, starting with the link level and going on up,



■ **Figure 3.** *A second captured screen of a session.*



■ **Figure 4.** *Actual network layout.*

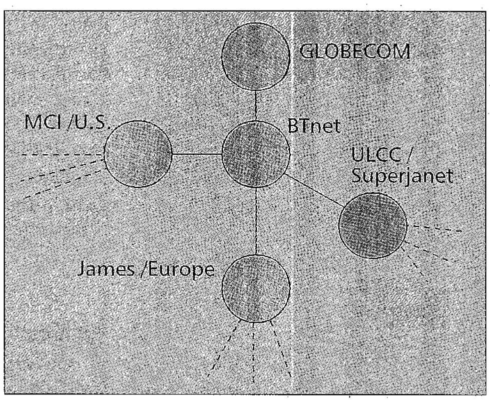
approximately through the International Organization for Standardization (ISO) open system interconnection (OSI) seven-layer model, as it were, until we reach the human factors levels.

LINK LEVEL

The first thing we learned was that local area network (LAN) technology was very flexible and reliable. Despite changing our network design within the Queen Elizabeth II Centre several times, the use of Ethernets, and particularly of mini-hubs, was highly successful and flexible.

In the lesson for the WAN, we had the opposite lessons. For both the links to and from the center, which were "Megastream" (E1, 2 Mb/s links), and for cross-connectivity between the U.K. BTnet commercial service and the U.K. academic networks which operate on SMDS switched networks, we encountered a variety of installation and configuration problems.

The key problem in both cases was that we had configured IP routes to use these before establishing that the link-layer connectivity was indeed in place. This should be avoided whenever possible!



■ **Figure 5.** *International access.*

Onward connectivity from the U.K. academic networks was provided over the pan-European ATM research network, JAMES, as part of access that UCL has for the Prospect and Merci projects, which involve multimedia conferencing for distance education. Since these had already been configured and tested as part of those projects, IP-level connectivity for unicast and multicast was ensured.

IP-LEVEL ROUTING, UNICAST AND MULTICAST

Always make sure you have all the Internet operational debugging tools installed on your exhibition and mbone machines.

For mbone topological and loss network debugging, mrinfo, for finding out which multicast tunnel routes are configured, and mtrace, for discovering the current topology, are essential. However, before this comes into play, traceroute and ping are vital for tracking throughput/loss and connectivity problems, as well as access to routers. tcpdump, snoop, or a similar packet sniffer are also all very useful.

Better integration and more accurate mbone statistics would always be an improvement. It is not easy to trace out a multicast distribution tree to find a lossy link, for example.

Telnet and e-mail are both also essential. The resulting topology is shown in Fig. 4, with the international links shown schematically in Fig. 5.

TRANSPORT LEVEL — RTP AND TCP PERFORMANCE

The transport-layer protocol used for traditional Internet applications is Transmission Control Protocol (TCP). However, for multimedia conferencing, especially for multiparty delivery, the appropriate protocols are Real-time Transport Protocol/Real-time Transport Control Protocol (RTP/RTCP) over User Datagram Protocol (UDP). UDP is connectionless, and does not attempt to recover from lost packets by any hop-by-hop or end-to-end means, unlike TCP. This avoids the nondeterministic delays and stalling effects due to retransmissions after loss and timeout in TCP. It does so at the expense of degradation of picture or speech quality, but given that loss is constrained to a reasonable level (typically less than 50 percent), audio and video coding schemes and receiver tools can be designed to accommodate this without too much perceived loss of quality by the user.

RTP (slightly misleadingly named) is simply a framing protocol for the media, while RTCP is a statistics-reporting protocol used to coordinate membership and traffic conditions within a multiparty conference. It is important to note how useful RTCP is for debugging quality problems at the network and link layers.

THE APPLICATIONS

The applications we used were the tools from UCL and LBL, namely Sdr, the Session Directory tool, which lists all Mbone events; Rat, the robust-audio tool; Vic, the LBL video conferencing program, and Wb, the LBL whiteboard, which is heavily used to provide feedback between participants.

The NTT tools were the SoftwareVision packages, with hardware assist for coding but pure software reception. These tools are available from the network at the sites listed at the end of this article.

AUDIO/VISUAL

We were very fortunate in having direct assistance from the audio-visual company at the Queen Elizabeth II Conference Centre (the appropriately named "Interface") They ensured that we had a direct line-level studio-quality audio feed to and from the PA and microphones in the main auditorium, as well as provided remote cabling for access between our cameras in the auditorium and our computer equipment in the booth.

It is worth noting that a lot of computing audio equipment has "less than professional quality" analog audio input/output, so typically, injecting audio output from a computer system (even a Sun or SGI workstation) to a large PA sometimes generates unexpected levels of background noise. We were fairly lucky in this respect in that the systems we happened to use were, purely by chance, not too noisy.

An important aspect of "Mboning" a large event is the choice of cameras — you need to be able to zoom in on speakers. We had one studio-quality camera for capturing slides, but the other cameras we used were domestic camcorder variety — these were adequate, but only just.

HUMAN LEVEL STAFFING

Good knowledge of IP, Ethernet, UNIX, and Windows, as well as router configuration, and good contact information is essential. Phones (even cell phones) and personal organizers are useful to troubleshoot things quickly.

ACKNOWLEDGMENTS

Many thanks are due to Peter Ingram and David Pratt of BT for help with coordination and router configuration. The staff at UCL's Information Systems Division, the University of London Computer Center at the Networks Operations Center, and UCL CS helped to debug Internet multicast routing and SMDS interworking problems. Nigel Titley at BTnet helped coordinate the QEII/BT/MCI connectivity.

SoftwareVision is available from http://www.hil.ntt.jp/SoftwareVision/index-e.html; Mbone tools can be retrieved from http://ugwww.ucs.ed.ac.uk/mice/archive/.

Finally, thanks are due to the UCL Merci project, project 1007, funded under the European Commission (EC) Telematics program.

REFERENCES

[1] S. Casner, "Frequently Asked Questions (FAQ) on the Multicast Backbone (MBONE)" http://www-mice.cs.ucl.ac.uk/mice/faq.html.

[2] H. Jinzenji, H. Kasahara, and K. Shimamura "A Scalable Video Distribution Technique for the Internet," presented at the Work in Progress Session, IEEE GLOBECOM Internet '96 Conf., London, U.K., Nov. 1996.

[3] S. Herzog, D. Estrin, and S. Shenker, "Sharing the Cost of Multicast Trees: An Axiomatic Analysis," *ACM SIGCOMM '95*, vol. 25, no. 4, Cambridge, MA, Sept. 1995, pp. 315–27.

[4] H. Schulzrinne *et al.*, "RTP: A Transport Protocol for Real-Time Applications," 1889 RFC, Jan. 25, 1996

BIOGRAPHIES

JON CROWCROFT [SM] is a professor of networked systems in the Dept. of Computer Science, University College London, where he is responsible for a number of European and U.S. funded research projects in multimedia communications. He has been working in these areas for over 15 years. He graduated in physics from Trinity College, Cambridge University in 1979, and gained his MSc in computing in 1981, and PhD in 1993. He is a member of the ACM, the British Computer Society, and the IEE. He is general chair for ACM SIGCOMM and on the editorial team for the *ACM/IEEE Transactions on Networks*. With Mark Handley, he is the co-author of *WWW: Beneath the Surf*; he also authored *Open Distributed Systems* (both UCL Press; the latter Artech in the U.S.). *Internet Multimedia* will be published shortly by Taylor and Francis.

COLIN PERKINS is a research fellow in the Dept. of Computer Science, University College London, where he works on the RAT and MERCI multimedia conferencing projects. He received a B.Eng. degree in electronic engineering from the University of York in 1992, and his D.Phil. in 1996 for work involving software reliability modeling and analysis.

For an expanded version of this article with Internet links visit the Communications Interactive Web site at **www.comsoc.org**