# About RTP and the Audio-Video Transport Working Group

Henning Schulzrinne

January 20, 1996

## Contents

# 1 RTP: Overview

RTP is a thin protocol providing support for applications with real-time properties, including timing reconstruction, loss detection, security and content identification. RTCP provides support for real-time conferencing for large groups within an internet, including source identification and support for gateways (like audio and video bridges) and multicast-to-unicast translators. RTP can be used without RTCP if desired. While UDP/IP is its initial target networking environment, efforts have been made to make RTP transport-independent so that it could be used, say, over CLNP, IPX or other protocols. RTP is currently also in experimental use directly over AAL5/ATM. RTP does not address the issue of resource reservation or quality of service control; instead, it relies on resource reservation protocols such as RSVP.

Other applications, such as real-time control and distributed simulation, are also targets.

# 2 RTP Status

On November 22, 1995, RTP was approved by the IESG as an Internet proposed standard. Publication as set of two RFCs (specification and profile) is awaiting processing by the RFC editor. The RFC text will be identical except for formatting details to the last Internet drafts published.

The IESG has received a request from the Audio/Video Transport Working Group to consider "RTP Payload Format of CellB Video Encoding" (draft-ietf-avt-cellb-06.txt) for the status of Proposed Standard.

The IESG plans to make a decision in the next few weeks, and solicits final comments on this action. Please send any comments to the `iesg@cnri.reston.va.us` or `ietf@cnri.reston.va.us` mailing lists by December 27, 1995.

ITU study group (SG) 15 has agreed in principle to use of RTP for LAN-based conferencing interoperable with H.320.

# 3 The AVT Working Group

The AVT WG (audio-video transport working group) within the IETF (Internet Engineering Task Force), the standardization body of the Internet, has been charged with developing a protocol that facilitates the transport of real-time data, such as interactive voice and video, over packet switched networks, in particular IP.

To subscribe to the working group mailing list and participate in the WG discussions, send your name to `rem-conf-request@es.net`.

Minutes of the meetings of the AVT working group at the IETF meetings are available:

- March 1994
- July 1994
- December 1994
- April 1995
- July 1995
- December 1995

## 4 Current Drafts

The current drafts of the AVT working group are available from any Internet drafts directory, including ds.internic.net. The content and draft file names are:

**Protocol specification:** (November 20, 1995) ASCII and PostScript

**Profile for audio and video:** (November 20, 1995) ASCII and PostScript

The RTP profile contains, inter alia, descriptions of and references to of some of the standard audio and video encodings to be used for RTP in multimedia conferencing. Both protocol specification and a profile are necessary to implement an RTP-compliant application.

Details are contained in the document.

**General discussion:** PostScript

**RTP Encapsulation of CellB Video Encoding:** (August 21, 1995) ASCII

**Packetization of H.261 video streams:** ASCII

**RTP Encapsulation of MPEG1/MPEG2:** ASCII

**RTP Encapsulation of JPEG-compressed video:** ASCII

**Profile for Video Data over RTP:** ASCII

Current and expired drafts are also available from gaia.cs.umass.edu, directory pub/hgschulz/rtp.

## 5 Papers on RTP

A number of papers, masters thesis and reports related to RTP have been published:

## References

[1] H. Schulzrinne, “RTP: The real-time transport protocol,” in *MCNC 2nd Packet Video Workshop*, vol. 2, (Research Triangle Park, North Carolina), Dec. 1992.

[2] P. T. Kirstein, M. J. Handley, and M. A. Sasse, “Piloting of multimedia integrated communications for European researchers (MICE),” in *Proceedings of the International Networking Conference (INET)*, (San Francisco, California), pp. DCA–1 – DCA–12, Internet Society, Aug. 1993.

[3] D. Hoffman, M. Speer, and G. Fernando, “Network support for dynamically scaled multimedia data streams,” in *Proceedings of the 4th International Workshop on Network and Operating System Support for Digital Audio and Video*, (Lancaster, U.K.), pp. 251–262, Lancaster University, Nov. 1993.

[4] I. Wakeman and J. Crowcroft, “Multicast congestion control in the distribution of variable bit rate video.” Computer Science Department, University College London, Jan. 1994.

[5] H. Schulzrinne, “Internet services: from electronic mail to real-time multimedia,” in *Proc. of KIVS (Kommunikation in Verteilten Systemen)* (K. Franke, U. Hübner, and W. Kalfa, eds.), Informatik aktuell, (Chemnitz, Germany), pp. 21–34, Gesellschaft für Informatik, Springer Verlag, Feb. 1995.

[6] H. Schulzrinne, “QOS for real-time services: playout delay and application control,” in *Proc. of 46th RACE Concertation Meeting (RCM)*, (Brussels, Belgium), Mar. 1995.

[7] H. Schulzrinne, “When can we unplug the phone and the radio?,” in *Proc. International Workshop on Network and Operating System Support for for Digital Audio and Video (NOSSDAV)*, Lecture Notes in Computer Science (LNCS), (Durham, New Hampshire), pp. 183–184, Springer, Apr. 1995.

[8] P. Harrschar, “Kommunikationsunterstützung fr multimediale Anwendungen,” Diplomarbeit, Institut fr Telematik, Fakultt fr Informatik, Universitt Karlsruhe, Karlsruhe, Germany, July 1995.

[9] C. Sieckmeyer, “Bewertung von adaptiven Ausspielalgorithmen fr paketvermittelte Audiodaten (Evaluation of adaptive playout algorithms for packet audio),” Studienarbeit, Dept. of Electrical Engineering, TU Berlin, Berlin, Germany, Oct. 1995.

[10] S. McCanne and V. Jacobson, “vic: A flexible framework for packet video,” in *Proc. of ACM Multimedia ’95*, Nov. 1995.

[11] E. Amir, S. McCanne, and H. Zhang, “An application level video gateway,” in *Proc. of ACM Multimedia*, (San Francisco, California), Nov. 1995.

[12] E. Amir, “An application level video gateway,” Master’s thesis, University of California, Berkeley, Berkeley, California, Dec. 1995.

[13] I. Busse, B. Deffner, and H. Schulzrinne, “Dynamic QoS control of multimedia applications based on RTP,” *Computer Communications*, Jan. 1996.

[14] D. T. Magill, “Adaptive speech compression for packet communication systems,” in *Conference record of the IEEE National Telecommunications Conference*, (??), pp. 29D–1 – 29D–5, ?? 1973.

[15] D. Cohen, “Specifications for the network voice protocol NVP,” RFC 741, Internet Engineering Task Force, Nov. 1977.

# 6 Related IETF Efforts

Related IETF efforts include:

**Audio/Video Transport (avt):** RTP protocol

**Chair(s):** Stephen Casner `<casner@precept.com>`

**WG Mail:** `rem-conf@es.net`

**To Join:** `rem-conf-request@es.net`

**Archive:** ftp://nic.es.net/pub/mailing-lists/mail-archive/rem-conf

**Multiparty Multimedia Session Control (mmusic):** conference control for both large and small sessions

**Chair(s):** Mark Handley <m.handley@cs.ucl.ac.uk>;
Ruth Lang <rlang@sri.com>;
Eve Schooler <schooler@cs.caltech.edu>

**WG Mail:** confctrl@isi.edu

**To Join:** confctrl-request@isi.edu

**Archive:** ftp.isi.edu/confctrl/confctrl.mail

**Inter-Domain Multicast Routing (idmr):** multicast routing

**Chair(s):** Tony Ballardie <A.Ballardie@cs.ucl.ac.uk>;
Paul Francis <francis@slab.ntt.jp>

**WG Mail:** idmr@cs.ucl.ac.uk

**To Join:** idmr-request@cs.ucl.ac.uk

**Archive:** cs.ucl.ac.uk/darpa/idmr-archive.Z

**Integrated Services (intserv):** service model for the Internet

**Chair(s):** Craig Partridge <craig@bbn.com>;
Scott Shenker <shenker@parc.xerox.com>;
John Wroclawski <jtw@lcs.mit.edu>

**WG Mail:** int-serv@isi.edu

**To Join:** int-serv-request@isi.edu

**Archive:** ftp.isi.edu/int-serv/int-serv.mail

**Resource Reservation Setup Protocol (rsvp):** resource reservation

**Chair(s):** Robert Braden <braden@isi.edu>;
Lixia Zhang <lixia@parc.xerox.com>

**WG Mail:** rsvp@isi.edu

**To Join:** rsvp-request@isi.edu

**Archive:** ftp.isi.edu/rsvp/rsvp.mail

**Internet Stream Protocol V2 (st2):** ST-II revision

**Chair(s):** Lou Berger <lberger@bbn.com>;
Luca Delgrossi <luca@ibmpa.awdpa.ibm.com>

**WG Mail:** st@ibminet.awdpa.ibm.com

**To Join:** st-request@ibminet.awdpa.ibm.com

**Archive:** ibminet.awdpa.ibm.com/pub/st/st-archive

Information on active IETF working groups can be found in a listing on ds.internic.net.

General information on networked multimedia is found in the standard journals, such as the IEEE Network Magazine, and conference proceedings of conferences such as ACM Sigcomm, IEEE Infocom, ACM Multimedia and IEEE Multimedia. You may also find references in the Network Bibliography. RTP is also targeted for use on the MBONE. Other available resources:

- a survey of multimedia research, standards and products (January 1993; the RTP information is rather out of date)
- Index to Multimedia Information Sources

# 7 RTP Implementations

The draft form of RTP is currently implemented in the following software. Note that RTP has undergone several revisions; most implementations are in the process of catching up with the current specification.

See Section 8.6 for a discussion of RTP implementations.

Implementations whose RTP status is described as RTPv1 implement an earlier version of RTP and are not compatible with the current draft.

## 7.1 RTP Library

**Name:** unknown

**Version:** 0.6.1b

**Description:** RTP library, not coupled with an application

**Contact:** Mastromartino E. A. Fantoni

**Email:** `fantoni@iasi.rm.cnr.it`

**Platform:** DEC OSF/1 v3.2, Solaris 2.3

**Video encodings:** MPEG, CellB (planned)

**Audio encodings:** MPEG (planned)

**RTP Status:** March 1995 (v.2)

**Availability:** C++ source code (at your own risk)

## 7.2 bat

**Name:** bat

**Version:** 0.1 (August 1995)

**Description:** network audio conferencing/telephony tool

**Contact:** Amancio Hasty Jr.

**WWW:** –

**Organization:** Star Gate

**Email:** `hasty@rah.star-gate.com`

**Platform:** FreeBSD

**Video encodings:** –

**Audio encodings:** PCMU, GSM

**RTP Status:** March 1995 (v.2)

**Availability:** source

**ftp:** contact author

### 7.3 Communique!

**Name:** Communique!

**Version:** 4.0

**Description:** network audio/video/whiteboard/etc. conferencing tool based on OpenDVE architecture

**Organization:** InSoft, Inc., 4718 Old Gettysburg Road, Suite 307, Mechanicsburg, PA 17055, USA

**Contact:** +1 717-730-9501

**Email:** `support@insoft.com`

**WWW:** `www.insoft.com`

**Platform:** Sun, HP, IBM, DEC Alpha Unix workstations; IBM-PC (Microsoft Windows)

**Codec:** ?

**Video encodings:** ?

**Audio encodings:** ?

**RTP Status:** ?

**Availability:** commercial product

**Entry:** Alexander V. Sokolsky `<sokolsky@frodo.microserve.com>`

### 7.4 InPerson

**Name:** InPerson (TM)

**Version:** 2.0 (April 1995)

**Description:** network audio/video/whiteboard conferencing tool

**Organization:** Silicon Graphics Computer Systems

**Contact:** North America: 1-800-800-7441; and ask for department code D150;
Europe: +41 22-798.75.25
North Pacific: +81 3-5420.71.10

**Email:** `inperson@sgi.com`

**WWW:** `www.sgi.com/Products/inperson_main.html`

**Platforms:** Silicon Graphics Indy, Indigo, Indigo2 and other SGI workstations running IRIX 5.3; NetManage's Internet Chameleon version 4.5 includes the InPerson whiteboard. Internezzo plans to offer InPerson on Sun and HP machines in late '95.

**Codec:** Indy: built-in IndyCam; others: optional Galileo/Indigo2 Video boards

**Video encodings:** H.261, RGB8, HDCC

**Audio encodings:** IDVI (16 and 8 kHz), PCMU (8 kHz), GSM, L16 (32 and 44.1 kHz), G.728 + AEC with InPerson hardware option board for Indy.

**RTP Status:** implements early draft of RTP v.1

**Price:** For price info, send mail to `inperson@sgi.com`

### 7.5 InVision

**Name:** InVision

**Version:** unknown

**Description:** network audio/video conferencing tool

**Organization:** InVision Systems Corp., 8500 Leesburg Pike, Suite 300, Vienna, VA 22182

**Contact:** +1 703 506-0098

**Email:** ???

**Platform:** Intel 386+ (MS-Windows 3.1+)

**Codec:** Intel i750-based codec or ActionMedia 2

**Video encodings:** ?

**Audio encodings:** ?

**RTP Status:** "plans to implement RTP"

**Price:** $2995 or $3995 with video codec, $995 without

### 7.6 ivs

**Name:** ivs

**Version:** 3.3m3

**Description:** network audio/video conferencing tool

**Contact:** Thierry Turletti

**WWW:** `zenon.inria.fr:8003/rodeo/personnel/Thierry.Turletti/ivs.html`

**Organization:** INRIA, Sophia-Antipolis, France

**Email:** `Thierry.Turletti@sophia.inria.fr`

**Platform:** Sun, SGI, DEC, HP, Linux

**Video encodings:** H.261

**Audio encodings:** PCMU, IDVI, VDVI

**RTP Status:** RTPv1 (October 1993), RTPv2 in progress

**Availability:** source, binary

**ftp:** `zenon.inria.fr:rodeo/ivs/last_version`

### 7.7 Loki

**Name:** loki

**Version:** 1.0?

**Description:** network video conferencing tool

**Contact:** Frank Kastenholz

**Organization:** FTP Software

**Email:** `video@research.ftp.com`

**Platform:** PC (Winsock)

**Video encodings:** Video for Windows (RGB, Indeo, YUV, …)

**Audio encodings:** –

**RTP Status:** RTPv2 (-05)

**Availability:** binary

**ftp:** `ftp.ftp.com/support/pub/windows/ftpvideo.exe`

**Encapsulation:** `ds.internic.net/internet-drafts/draft-kastenholz-loki-00.txt`

### 7.8 MSessMon

**Name:** MSessMon

**Version:** 0.3

**Description:** multicast session monitor

**Contact:** Paul Stewart

**Organization:** Rensselaer Polytechnic Institute

**Email:** `stewart@hibp.ecse.rpi.edu`

**Platform:** Sun (tested), others likely

**RTP Status:** RTPv2 (July 1994)

**Availability:** source, binary (SunOS, NetBSD)

**ftp:** `hibp6.ecse.rpi.edu/pub/msessmon`

### 7.9 NeVoT

**Name:** NeVoT

**Version:** 3.34 (December 1995)

**Description:** network audio conferencing/telephony tool

**Contact:** Henning Schulzrinne

**WWW:** `www.fokus.gmd.de/step/hgs/nevot`

**Organization:** Univ. of Massachusetts, AT&T Bell Laboratories, GMD Fokus

**Email:** `schulzrinne@fokus.gmd.de`

**Platform:** Sun (4.1.3, Solaris 2), SGI, HP, IBM, DEC; Linux (in progress)

**Video encodings:** –

**Audio encodings:** LPC, DVI4, GSM, PCMU, L16

**Multicast:** IP multicast (unlimited number of participants)

**Conference Control:** through PMM control bus (floor control, invitations) and sd

**Monitoring:** statistics for loss, delay and delay variation of audio from other participants

**RTP Status:** RTPv2 (November 1995)

**Availability:** source, binary

**ftp:** `ftp.fokus.gmd.de/pub/step/nevot/`, `gaia.cs.umass.edu/pub/hgschulz/nevot/`

### 7.10 nv

**Name:** nv

**Version:** 3.3

**Description:** network video conferencing tool

**Contact:** Ron Frederic

**Organization:** Xerox PARC

**Email:** `frederic@parc.xerox.com`

**Platform:** Sun, SGI, DEC, HP, Linux

**Video encodings:** nv

**Audio encodings:** –

**RTP Status:** RTPv1 (at least January 1993)

**Availability:** source, binary

**ftp:** `parcftp.xerox.com:pub/net-research/`

### 7.11 PicWin

**Name:** PicWin

**Version:** 1.3

**Description:** network video conferencing tool

**Contact:** Paul Milazzo

**Organization:** BBN

**Email:** `milazzo@diamond.bbn.com`

**Platform:** Sun

**Video encodings:** ?

**Audio encodings:** ?

**RTP Status:** RTPv1

**Availability:** binary only (commercial)

**ftp:** receive-only version from picwin.bbn.com

### 7.12 rtpdump

**Name:** rtpdump

**Version:** 1.0 (December 1994)

**Description:** display/decode RTP packets

**Organization:** GMD Fokus

**Contact:** Henning Schulzrinne

**Email:** `hgs@fokus.gmd.de`

**Platform:** workstations with or without multicast support

**RTP Status:** RTPv2 (November 1995)

**Availability:** source

**ftp:** `gaia.cs.umass.edu/pub/hgschulz/rtp/rtpdump-1.0.tar.gz`

### 7.13 ShowMe

**Name:** ShowMe

**Version:** 2.0

**Description:** network audio/video conferencing tool

**Organization:** SunSolutions

**Contact:**

**Email:** `sunsol-www@zircon.Eng.Sun.COM`

**WWW:** `www.sun.com/sunsolutions/overview/index.html`

**Platform:** Sun Solaris 1.1/2.2 or later

**Codec:** SunVideo

**Video encodings:** ?

**Audio encodings:** ?

**RTP Status:** RTPv1 (unknown version)

**Availability:** commercial

**Price:** DM 2450 for complete package

### 7.14 Telesia-IVS

**Name:** Telesia-IVS

**Version:** 2.3

**Description:** network audio/video conferencing tool

**Organization:** INRIA

**Contact:** Alain Caristan, Christian Donot

**Email:** `Alain.Caristan@inria.fr`

**Platform:** Sun

**Codec:** Parallax, SunVideo, Sun Videopix

**Audio encodings:** ?

**Video encodings:** ?

**RTP Status:** RTPv1 (unknown version)

**Availability:** binary

**ftp:** `electre.inria.fr/pub/videoconf/telesia`

### 7.15 vat

**Version:** 4.0a* (alpha)

**Description:** network audio conferencing tool

**Organization:** LBNL

**Contact:** Steve McCanne

**Email:** `vat@ee.lbl.gov`

**Platforms:** DEC, HP, Linux, SGI, SunOS, Solaris

**Audio encodings:** LPC, GSM, DVI4, PCMU

**Video encodings:** –

**RTP Status:** RTPv2 (November 1995)

**Availability:** source, binary

**ftp:** `ftp.ee.lbl.gov/conferencing/vat/alpha_test`

### 7.16 vic

**Name:** vic

**Version:** 2.6

**Description:** network video conferencing tool

**Organization:** LBNL

**Contact:** Steve McCanne

**Email:** `vic@ee.lbl.gov`

**Platforms:** DEC, HP, Linux, SGI, SunOS, Solaris

**Codec:** Sun VideoPix, SunVideo (Sun Solaris), ?

**Video encodings:** nv, H261, CELB, SCR, MPEG, JPEG

**Audio encodings:** –

**Availability:** source, binary

**ftp:** `ftp.ee.lbl.gov/conferencing/vic`

**RTP Status:** RTPv2 (November 1995)

### 7.17 VTT Conferencing Tool

**Name:** unknown

**Version:** unknown

**Description:** network audio/video conferencing tool

**Contact:** Jarmo Mölsa, Kirsi Valtari

**Email:** `jarmo.molsa@vtt.fi`

**Platform:** MS-DOS under MS-Windows

**Codec:** Bitfield H.261 codec

**Video encodings:** H261

**Audio encodings:** ?

**Availability:** commercial product

**RTP Status:** ?

# 8 Some Frequently Asked Questions

## 8.1 Is RTP a transport protocol or a kind of application protocol?

RTP has important properties of a transport protocol: it runs on end systems, it provides demultiplexing. It differs from transport protocols like TCP in that it (currently) does not offer any form of reliability or a protocol-defined flow/congestion control. However, it provides the necessary hooks for adding reliability, where appropriate, and flow/congestion control. Some like to refer to this property as application-level framing (see D. Clark and D. Tennenhouse, "Architectural considerations for a new generation of protocols", SIGCOMM'90, Philadelphia). RTP so far has been mostly implemented within applications, but that has no bearing on its role. TCP is still a transport protocol even if it is implemented as part of an application rather than the operating system kernel.

## 8.2 RTP does not ensure real-time delivery. So how come it is called a real-time protocol?

No end-to-end protocol, including RTP, can ensure in-time delivery. This always requires the support of lower layers that actually have control over resources in switches and routers. RTP provides functionality suited for carrying real-time content, e.g., a timestamp and control mechanisms for synchronizing different streams with timing properties.

## 8.3 Is Is RTP an unreliable protocol? Are there any mechanisms provided for error recovery in RTP?

As currently defined, RTP does not define any mechanisms for recovering for packet loss. Such mechanisms are likely to be highly dependent on the packet content. For example, for audio, it has been suggested to add low-bit-rate redundancy, offset in time. For other applications, retransmission of lost packets may be appropriate. This requires no additions to RTP. RTP probably has the necessary header information (like sequence numbers) for some forms of error recovery by retransmission.

### 8.4 Why can't we just use TCP for audio and video?

For delivering audio and video for playback, TCP may be appropriate. Also, with sufficiently long buffering and adequate average throughput, near-real-time delivery using TCP can be successful, as practiced by the Netscape WWW browser. TCP may often run over highly lossy networks (e.g., the German X.25 network) with acceptable throughput, even though the uncompensated losses would make audio or video communication impossible.

However, for real-time delivery of audio and video, TCP and other reliable transport protocols such as XTP are inappropriate. The three main reaons are:

- Reliable transmission is inappropriate for delay-sensitive data such as real-time audio and video. By the time the sender has discovered the missing packet and retransmitted it, at least one round-trip time, likely more, has elapsed. The receiver either has to wait for the retransmission, increasing delay and incurring an audible gap in playout, or discard the retransmitted packet, defeating the TCP mechanism. Standard TCP implementations force the receiver application to wait, so that packet losses would always yield increased delay. Note that a single packet lost repeatedly could drastically increase delay, which would persist at least until the end of talkspurt.
- TCP cannot support multicast.
- The TCP congestion control mechanisms decreases the congestion window when packet losses are detected ("slow start"). Audio and video, on the other hand, have "natural" rates that cannot be suddenly decreased without starving the receiver. For example, standard PCM audio requires 64 kb/s, plus any header overhead, and cannot be delivered in less than that. Video could be more easily throttled simply by slowing the acquisition of frames at the sender when the transmitter's send buffer is full, with the corresponding delay. The correct congestion response for these media is to change the audio/video encoding, video frame rate, or video image size at the transmitter, based, for example, on feedback received through RTCP receiver report packets.

An additional small disadvantage is that the TCP and XTP headers are larger than a UDP header (40 bytes for TCP and XTP 3.6, 32 bytes for XTP 4.0, compared to 8 bytes). Also, these reliable transport protocols do not contain the necessary timestamp and encoding information needed by the receiving application, so that they cannot replace RTP. (They would not need the sequence number as these protocols assure that no losses or reordering takes place.)

While LANs often have sufficient bandwidth and low enough losses not to trigger these problems, TCP does not offer any advantages in that scenario either, except for the recovery from rare packet losses. Even in a LAN with no losses, TCP would suffer from the initial slow start delay.

### 8.5 Can't we just use XTP?

Many of the arguments parallel those in the previous section.

The question of the relationship of RTP and XTP appears to arise frequently. (This may simply be due to the word 'transport' in both protocol names.) However, XTP and RTP are not replacements for each other. XTP is designed as a general, configurable network and transport protocol for both reliable and unreliable data communications. RTP has no reliability mechanisms (although these could be added if desired for specific applications) and no flow control like the rate control in XTP. RTP is not intended for regular, reliable data transfer (where TCP or XTP might be used instead). For real-time data, where retransmission is usually not possible due to timing constraints, XTP would have to disable retransmission. Flow/congestion control for real-time data is most likely inappropriate as the rate of such sources is inherently given and not modifiable on the time-scale of transport-protocol flow control, as explained in the previous section. It should be noted that RTP supports mechanisms that allow a form of congestion control on longer time scales, e.g., by modifying the source encoder if network congestion is detected.

RTP has no protocol state by itself and can thus be used over either connection-less networks, such as IP/UDP, or connection-oriented networks, such as XTP, ST-II or ATM (AAL3/4 or AAL5). Many real-time multimedia applications use multicast with a large fan-out, e.g., several hundred to thousands for a lecture or concert. Connection-oriented protocols like XTP have difficulty scaling to such a large number of receivers.

XTP does not offer timing or content type (media) information, and thus would need these services, as offered by RTP. XTP provides no RTP-like direct feedback of the received quality-of-service, and thus, again, would have to "import" these from another protocol. Looking at existing applications using XTP for real-time services confirms that they need to add a layer similar in content to the RTP data part "between" XTP and the actual media.

## 8.6 Is there an RTP library or kernel implementation?

RTP (in particular, the data part) is tightly coupled to the application, so that a kernel or library implementation makes little sense. However, NeVoT can be used as a linkable library that implements RTP for an audio tool, with a documented API. The sources to NeVoT, rtpdump and vic also contain RTCP processing modules which should be usable in other applications with minor modifications. Note also that the specification itself contains numerous code fragments. (Most of the other applications are using older versions of RTP and thus should not be relied upon for developments.)

## 8.7 What are some of the differences between the VAT protocol and RTP?

The VAT protocol was originally implemented in the VAT audio tool and subsequently also in other audio tools such as NeVoT. It is currently the most frequently used packet format for audio on the MBONE. The VAT header format is only described in header files. (See the VAT and NeVoT sources for details.) Many aspects of RTP and the VAT protocol are similar, but RTP improves upon the VAT protocol in a number of ways:

- The VAT protocol was designed for audio only, while RTP is specified for audio and video and may be suitable for other real-time applications.
- RTP is designed to be protocol-independent and can be used with non-IP protocols (ATM AAL5, for example) as well as, say, IPv6.
- RTP source identification simplifies the use of mixers and translators.
- RTP has a number of features that simplify use of application-level encryption (padding, etc.).
- The RTP header is extensible, should the need arise in the future.
- The RTP header has a sequence number which simplifies accurate loss detection and measurement and the handling of images transmitted in several packets.
- The RTCP SDES packets contain additional information that simplify tracing of misbehaving sources, e.g., their email address or telephone number.
- The RTCP SDES CNAME items simplify the construction of multimedia application from independent media agents.
- RTCP sender and receiver reports allow the implementation of adaptive applications, that is, applications where senders scale their bandwidth consumption based on network load.

- RTCP sender and receiver reports allow monitoring of the quality of service within, say, a multimedia conference.

A new version of VAT (currently in alpha-test) also implements RTP. As soon as there are a sufficient number of stable applications using RTP, it is anticipated that most Internet MBONE audio/video events will be transmitted using RTP.

### 8.8 What are the differences between RTP version 1 and 2?

Version 1 is of historical interest only. Applications should not be written for it. RTP version 2 is not backwards compatible with version 1. If you care, you can find a definition of version 1 in an old Internet draft.

### 8.9 Are there other efforts in using the Internet for real-time audio and video?

Too many, some may say. vat versions 3.4 and earlier, one of the early (recent) Internet audio applications, uses mostly the same audio encodings as specified in the RTP profile, but a different protocol. There are a number of "Internet telephones" (usually for PCs) using proprietary audio coding and protocols, meant for point-to-point connections:

- Speak Freely for Unix and Windows
- Internet Phone (Vocaltec)
- Digiphone
- Quarterdeck
- Internet Telephone Company
- Telescape Intercom by Telescape
- Audio/video directly over ATM: Nemesys
- FreeVue audio and video

For near real-time distribution of audio, e.g., the on-demand delivery of music or news:

- RealAudio (Microsoft Windows only)
- Xing
- AudioSoft

CuSeeMe (for the Macintosh) is combined audio and video tool using reflectors rather than IP-level multicast.

RealAudio writes what currently applies to all tools:

> If the packet loss is high, it may be due to a busy network. If this is the case, there is little you can do to remedy the situation other than to try connecting to the site at a later time.

# 9 Historical Notes

Efforts in transmitting audio over the Internet protocols IP and ST-II go back to the early 1970's [14, 15]. While these experiments used special-purpose hardware and low-bit-rate encodings such as LPC, new experimental interest for transmitting voice over packet networks arose after Sun Microsystems introduced the SPARCstation 1 with a built-in mu-law audio codec.

Versions of software for Sun workstations from the early 1990's include vt and vtalk.

Experiments leading to the MBONE tools and RTP were conducted on the ARPA test network called DARTnet, a transcontinental IP network of about a dozen research sites currently connected by T1 trunks:

```
Date: Tue 12 Feb 91 00:43:28-PST
From: Stephen Casner <CASNER>
Subject: Packet voice over DARTnet
To: DARTNET
Cc: perkins@ee.udel.edu, deschon

I would like to report a successful test with packet voice carried
across DARTnet today.  The experiment was set up by Erik Perkins at
UDel and Annette DeSchon at ISI, using the VT (Voice Terminal) program
recently released by ISI.

Generally the sound quality was quite good and it was easy to carry on
a normal conversation.  Occasional dropouts were heard, but most of
these were apparently caused by effects other than packet loss or
delay.  Unexpected activity (e.g., compilation) on the test host at
the ISI end was correlated with some of the dropouts.  Other dropouts
heard were actually the result of noise from the listening end causing
the "speakerphone function" to momentarily cut off the talker's
microphone to prevent echo.  A third source of dropouts was observed
during a segment of continuous music; the music came through
uninterupted until the dynamic silence threshold eclipsed the music
level just as it is supposed to do to cancel background noise.  We
concluded that it would help to add commands to defeat the silence
threshold and echo suppression functions for test purposes.

Bob Braden told me there was also a test of packet video last week by
the SRI folks connecting to ISI using one of the new Sun boards.
Let's hear more about it!
-- Steve

To: braden
Cc: Dartnet
Subject: Re: Audio teleconference
Date: Mon, 12 Aug 91 14:49:05 PDT
From: Van Jacobson <van@ee.lbl.gov>

Bob,
```

```
If it would be agreeable to all the participants, I'd like to
try using LBL's new audio tool, vat, instead of vt for the
teleconference.  Vat is functionally equivalent to vt (except
the vat speakerphone mode works better) but has [what I
consider] an easier-to-use, window-based user interface.

The only problem I could see is that vat requires whatever
host you're using for conferencing understand X.  But, since
I think all the hosts we use run either X or Sunview (which
can speak both X and NeWS), this shouldn't be an obstacle.

 - Van

ps- unfortunately, the current version of vat won't interoperate
    with vt so it's an either/or choice of which to use.
```

# 10 Hardware for Audio/Video Conferencing

Desktop Videoconferencing Products

## 10.1 Cameras for Conferencing

- Howard Enterprises
- Connectix

Video glossary

## 10.2 Audio I/O Devices for Conferencing

Due to the acoustic feedback introduced by long delays and the lack of echo cancellation, speaker and microphone are, in many cases, not appropriate for interactive conferencing. To reduce background noise (and triggering of the silence detector), the microphone must be held close to the speaker's mouth. Some alternatives that combine an earphone/headset with a microphone include:

- JABRA earphone
- Sennheiser HMD 410-6 headphone/microphone combination (approximately DM 276)
- SoundXchange telephone handset
- Not Comp-U-Phone telephone handset
- Pioneer SE-DJ250 stereo headphones with microphone (may not be generally available)
- Shure SM-809 microphone (approximately DM 350)