

CHAPTER CONTENTS

[Appendix 3 - Multimedia Standards](#)

[CCITT/ISO standards](#)

[Internet Standards](#)

[Proprietary](#)

Appendix 3 - Multimedia Standards

The following PDF files contain a lot of useful information on multimedia standards.

[Understanding Multimedia Standards by Harris](#)

[Standards for Mixed Media by Lucent](#)

[MultiMedia Communications eXchange \(MMCX\) by Lucent](#)

[Making Multimedia & Real-Time Networks Possible Today by PACE Technology](#)

The following information was obtained from:

<http://www.eenus.sg/~engp4307/standard.htm>

Additional supplementary information can be found at:

<http://www.summitonline.com/papers/3com2.html>

CCITT/ISO STANDARDS

The CCITT has now been renamed the ITU.

- F.700
- G.711
- G.721
- G.722
- G.725
- G.723
- G.726
- G.727
- G.728
- G.764
- G.765
- H.200
- H.221
- H.241
- H.243
- H.242
- H.261
- H.320
- H.321
- H.322
- H.323
- H.331

- HyTime
- IIF
- JBIG
- PEG
- MHEG
- MPEG
- ODA
- T.80
- T.120
- X.400

INTERNET STANDARDS

- IP Multicast
- MIME
- RTP
- ST-2
- RFC 741
- Xv and mvex

PROPRIETARY STANDARDS

- Bento
- GIF
- QuickTime
- RIFF
- DVI
- MIDI

CCITT/ISO standards

F.700

Audiographic, Videotelephony and Videoconference service standards.

The individual recommendations are as follows:

- F.711 Audiographic Conference Teleservice for ISDN
- F.720 Videotelephony Services General
- F.721 Videotelephony Teleservices for ISDN
- F.722 Videotelephony Services General
- F.730 Videoconference Service General
- F.732 Broadband Videoconference Services
- F.740 Audiovisual Interactive Services (AVIS)

G.711

This defines toll quality audio with a 3.1 KHz bandwidth over 56 or 64 Kbps digital circuits. A-law and μ -law are both supported.

G.721

32 Kbps ADPCM for audio encoding.

G.722

This defines 7 KHz audio over a variety of digital loops. ADPCM is used to support bit rates of 48, 56, and 64 Kbps. It is primarily used for video conferencing circuits operating at 384 Kbps or greater.

G.723

This defines how a 3.1 KHz audio channel can be digitally encoded at rates of 5.3 and 6.3 Kbps. It has been selected as the preferred method of providing voice over IP (VoIP). It provides conversion between G.711 and lower-speed channels.

G.725

System Aspects of the use of 7 KHz audio codec within 64 Kbps

G.726

Replaces G.721?

G.727

Extension of G.726 for use over G.764

G.728

This defines toll quality voice with a 3.1 KHz bandwidth over a 16 Kbps digital channel. It is commonly used in videoconferencing systems.

The speech coding uses Low-Delay Code Excited Linear Prediction (LD-CELP).

G.729/729A

This defines toll quality voice with a 3.1 KHz bandwidth over a 8 Kbps digital channel.

G.764

Packetised Voice Protocol

G.765

Associated with G.764

H.200

Audio compression standard (forthcoming).

H.221

Defines the frame structure for a 64 to 1920 Kbps channel in audiovisual teleservices.

These are conveyed over single or multiple B or H0 channels or a single H11 or H12 channel. It offers several advantages:

- It takes into account Recommendations G.704, X.301/I.461, etc. It may allow the use of existing hardware and software.
- It is simple, economic and flexible. It may be implemented on a single microprocessor using well known hardware principles.
- It is a synchronous procedure. The exact time of a configuration change is the same in the transmitter and the receiver.
- It needs no return link for audiovisual signal transmission, since a configuration is signaled by repeatedly transmitted codewords.
- Very secure in case of transmission errors, since the code controlling the multiplex is protected by double-error correcting code.
- Allows synchronization of multiple 64 Kbps or 384 Kbps connections and the control of the multiplexing of audio, video, data and other signals within the synchronized multiconnection structure in the case of multimedia services such as videoconferencing.
- It can be used in multipoint configurations, where no dialogue is needed to negotiate the use of a data channel.
- It provides a variety of data bit-rates (from 300 b/s up to almost 2 Mb/s) to the user.
- Closely related to H.261 & H.242. Supersedes H.220

Products: Codecs from BT, GPT, Pictoretel, Videotel & others

H.241

Description: Signaling for conferencing.

H.242

Audiovisual communication using digital channels up to 2 Mbps.

Recommendation H.242 should be associated with Recommendations G.725, H.221 and H.230.

Applications utilizing narrow (3 KHz) and wideband (7 KHz) speech together with video and/or data have been identified.

To provide these services, a channel must accommodate speech, and optionally video and/or data at several rates, in a number of different modes. Signaling procedures are required to establish a compatible mode upon call set-up, to switch between modes during a call and to allow for call transfer.

Some services will require only a single channel, such as B (64 Kbps), H0 (384 Kbps), H11 (1536 Kbps) or H12 (1920 Kbps). Other services will require the establishment of two or more connections providing B or H0 channels: in such cases the first established is called hereafter the initial channel while the others are called additional channels.

All audio and audiovisual terminals using G.722 audio coding and/or G.711 speech coding or other standardized audio codings at lower bit rates should be compatible to permit connection between any two terminals. This implied that a common mode of operation has to be established for the call. The initial mode might be the only one used during a call or, alternatively, switching to another mode can occur as needed depending on the capabilities of the terminals. Thus, for these terminals an in-channel procedure for dynamic mode switching is required.

Recommendation H.242 develops these considerations and describes recommended in-channel procedures.

Products: Codecs from BT, GPT, Pictoretel, Videotel & others.

Further information: Closely related to H.261 & H.221. Supersedes H.220

H.243

Multipoint Video Codec Standard. Probably a draft

H.261

Video Codec for Audiovisual Services at P x 64 Kbps

Recommendation H.261 describes the video coding and decoding methods for the moving picture component of audiovisual services at the rate of P x 64 Kbps, where p is in the range 1 to 30. It describes the video source coder, the video multiplex coder and the transmission coder.

This standard is intended for carrying video over ISDN - in particular for face-to-face videophone applications and for videoconferencing. Videophone is less demanding of image quality, and can be achieved for p=1 or 2. For videoconferencing applications (where there are more than one person in the field of view) higher picture quality is required and p must be at least 6.

H.261 defines two picture formats: CIF has 288 lines by 360 pixels/line of luminance information and 144 x 180 of chrominance information; and QCIF which is 144 lines by 180 pixels/line of luminance and 72 x 90 of chrominance. The choice of CIF or QCIF depends on available channel capacity - eg QCIF is normally used if p<3.

The actual encoding algorithm is similar to (but incompatible with) that of MPEG. Another difference is that H.261 needs substantially less CPU power for real-time encoding than MPEG. The algorithm includes a mechanism which optimizes bandwidth usage by trading picture quality against motion, so that a quickly-changing picture will have a lower quality than a relatively static picture. H.261 used in this way is thus a constant-bit-rate encoding rather than a constant-quality, variable-bit-rate encoding.

Products: H.261 codecs have been implemented in VLSI and are now built in to commercially available codec equipment.

Further information: Document available on line on:

- src.doc.ic.ac.uk/doc/ccitt-standards/ccitt/1992/h
- “Overview of the p*64 Kbps Video Coding Standard”, M. Liou, Communications of the ACM, April 1991.

H.320

Narrow Band Visual Telephone systems and terminal equipment

Recommendation H.320 covers the technical requirements for narrow-band visual telephone services defined in H.200/AV.120-Series Recommendations, where channel rates do not exceed 1920 Kbps.

It is anticipated that Recommendation H.320 will be extended to a number of Recommendations each of which would cover a single videoconferencing or videophone service (narrow-band, broadband, etc.). However, large parts of these Recommendations would have identical wording, while in the points of divergence the actual choices between alternatives have not yet been made; for the time being, therefore, it is convenient to treat all the text in a single Recommendation.

The service requirements for visual telephone services are presented in Recommendation H.200/AV.120-Series; video and audio coding systems and other technical set aspects common to audiovisual services are covered in other Recommendations in the H.200/AV.200-Series.

H.321

B-ISDN videoconference and videophone system.

Specifies how to run H.320 compatible systems over ATM networks.

H.322

Specifies how to run H.320 compatible systems over LANs.

HYTIME [ISO (JTC1/SC18/WG8)]

SGML-based standard for hypermedia documents.

HyTime[†] is a standardized infrastructure for the representation of integrated, open hypermedia documents. It was developed principally by ANSI committee X3V1.8M, and was subsequently adopted by ISO.

[†] Hypermedia/Time-Based Structuring Language

The HyTime standard specifies how certain concepts common to all hypermedia documents can be represented using SGML. These concepts include:

- association of objects within documents with hyperlinks
- placement and interrelation of objects in space and time
- logical structure of the document
- inclusion of non-textual data in the document

An “object” in HyTime is part of a document, and is unrestricted in form - it may be video, audio, text, a program, graphics, etc.

SGML[†] is a metalanguage which is used to specify document markup schemes called Document Type. HyTime is not itself a DTD[†], but provides constructs and guidelines for making DTDs for describing Hypermedia documents. For instance, the SMDL[†] defines a DTD which is an application of HyTime.

HyTime consists of six modules:

1. Base module - provides facilities including “xenofoms” for specifying application defined expressions, and identification policies for coping with document changes, “activity tracking”.
2. Finite Coordinate Space module - allows for an object to be scheduled in time and/or space (which HyTime treats equivalently) within a bounding box called an “event”.
3. Location Address module - specifies how to identify locations of document objects by name, coordinate location, or by semantic construct.
4. Hyperlinks module - there are five different types of hyperlinks.
5. Event Projection module - specifies how events in a source FCS[†] are to be mapped onto a target FCS.
6. Object Modification module - allows objects to be modified before rendition, in an object-specific way.

Products: A public-domain SGML parser (ARC SGML) is available. TechnoTeacher (address below) are producing a HyTime engine. Sema Group are also understood to be developing a HyTime product.

Further information:

HyTime Special Interest Group (SIGHyper)

Steven R. Newcomb, Chairman (srn@cmr.fsu.edu)

TechnoTeacher Inc; 1810 High Road; Tallahassee, Florida 32303-4408

Phone: +1 904 422 3574 Fax: +1 904 386 2562

[†] Standard Generalized Markup Language: ISO 8879

[†] Document Type Definition

[†] Standard Music Description Language ISO/IEC Committee Draft 10743

[†] Finite Coordinate Space

There are FTP sites at:

ftp.ifi.uio.no:SIGhyper
mailer.cc.fsu.edu:pub/sgml

The following articles are useful:

- “The HyTime Hypermedia/Time-based Document Structuring Language”, S. Newcomb, N. Kipp and V. Newton, Communications of the ACM, p67, November 1991.
- “Emerging Hypermedia Standards” B. Markey, Multimedia for Now and the Future (Usenix Conference Proceedings), p59, June 1991.

See also newsgroup comp.text.sgml

IIF[†] [ISO (JTC1/SC24/WG7)]

The IIF is part of the first International IPI[†] Standard, under development by ISO/IEC JTC1/SC24. It consists of a data format definition and a gateway functional specification.

The main component of the IIF defines the data format for exchanging arbitrarily structured image data. It can be used across application boundaries and integrated into international communication services. There are definitions of parsers, generators, and format converters to enhance open image communication.

The IIF approach distinguishes between the image structure (data type), image attributes (colourimetric and geometric semantics), the sequential data organization (data partitioning and periodicity organization), and the data encoding/compression. The syntax specification and the data encoding of syntax entities use ASN.1 and the Basic Encoding Rules respectively. For the compressed representation, the following standards are referenced: JBIG, facsimile Group 3 and 4, JPEG, and MPEG.

The IIF also encompasses functionality for generating and parsing image data, compressing and decompressing, and the exchange of image data between the application program PIKS[†], which is Part 2 of the IPI standard, and storage/communication devices. This functionality is located in the so-called IIF Gateway. The IIF gateway controls the import and export of image data to and from applications, as well as to and from the PIKS.

The IIF may serve as a future image content architecture of the ODA[†].

Work is going on to develop a multimedia electronic mail application on top of X.400, using IIF.

Further information: “ISO/IEC’s image interchange format”, C. Blum and G. R. Hofmann, SPIE Proceedings Vol. 1659, San Jose, p130 February 1992.

-
- † Image Interchange Facility
 - † Image Processing and Interchange
 - † Programmer’s Imageing Kernel System
 - † Open Document Architecture

IIF editor:

Christof Blum (blum@igd.fhg.de)

Fraunhofer Institute for Computer Graphics (IGD)

Wilhelminenstr. 7; W-6100 Darmstadt; Germany

Phone: +49 6151 155 145 or 140 Fax: +49 6151 155 199

JBIG [ISO]

Binary image encoding standard

JBIG is a lossless compression algorithm for binary (one bit/pixel) images. The intent of JBIG is to replace the current, less effective group 3 and 4 fax algorithms.

JBIG models the redundancy in the image as the correlations of the pixel currently being coded with a set of nearby pixels called the template. An example template might be the two pixels preceding this one on the same line, and the five pixels centered above this pixel on the previous line. Note that this choice only involves pixels that have already been seen from a scanner.

The current pixel is then arithmetically coded based on the eight-bit (including the pixel being coded) state so formed. So there are (in this case) 256 contexts to be coded. The arithmetic coder and probability estimator for the contexts are actually IBM's (patented) Q-coder. The Q-coder uses low precision, rapidly adaptable (those two are related) probability estimation combined with a multiply-less arithmetic coder. The probability estimation is intimately tied to the interval calculations necessary for the arithmetic coding. JBIG actually goes beyond this and has adaptive templates.

You can use JBIG on gray-scale or even color images by simply applying the algorithm one bit-plane at a time. You would want to recode the gray or color levels first though, so that adjacent levels differ in only one bit (called Gray-coding). This works well up to about six bits per pixel, beyond which JPEG's lossless mode works better. You need to use the Q-coder with JPEG also to get this performance.

JPEG [CCITT | ISO (JTC1/SC2/WG10)]

Compression Standard for continuous-tone still images

JPEG (Joint Photographic Experts Group) is designed for compressing either 24 bit color or gray-scale digital images. JPEG does not handle black-and-white (one bit/pixel) images, or motion picture compression.

JPEG supports 4 compression modes, three of which are lossy. Much of its compression exploits limitations of the human eye, notably that small color details aren't perceived as well as small luminance details.

The degree of lossiness can be varied by adjusting compression parameters. The image maker can trade off file size against output image quality.

Products: Many products now use this algorithm. There is free JPEG source code available from the Independent JPEG group, at many FTP sites.

MHEG T.170 | ISO [CCITT | ISO (JTC1/SC2/WG12)]

Standard for hypermedia document representation.

MHEG stands for the Multimedia and Hypermedia Information Coding Experts Group. This group is developing a standard “Coded Representation of Multimedia and Hypermedia Information”, commonly called MHEG. The standard is likely to be published in two parts - part one being object representations and part two being hyperlinking.

MHEG is suited to interactive hypermedia applications such as on-line textbooks and encyclopedia. It is also suited for many of the interactive multimedia applications currently available (in platform-specific form) on CD-ROM. MHEG could for instance be used as the data structuring standard for a future home entertainment interactive multimedia appliance.

To address such markets, MHEG represents objects in a non-revisable form, and is therefore unsuitable as an input format for hypermedia authoring applications: its place is perhaps more as an output format for such tools. MHEG is thus not a multimedia document processing format - instead it provides rules for the structure of multimedia objects which permits the objects to be represented in a convenient form (e.g. video objects could be MPEG-encoded). It uses ASN.1 as a base syntax to represent object structure, but allows for the use of other syntax notations - an SGML syntax is also specified.

MHEG objects (which may be textual information, graphics, video, audio, etc.) may be of four types:

- Input object (i.e. a user control such as a button or menu)
- Output object (e.g. graphics, audio visual display, text)
- Interactive object (a “composite” object containing both input and output objects)
- Hyperobject (a “composite” object containing both input and output objects, with links between them).

MHEG supports various synchronization modes, for presenting output objects in these relationships.

MPEG ISO 11172 [ISO (JTC1/SC2/WG11)]

Standard for compressed video and audio

MPEG is the name of the ISO committee that is working on digital color video and audio compression, and by extension, the name of the standard they have produced. MPEG defines a bit-stream representation for synchronized digital video and audio, compressed to fit into a bandwidth of 1.5 Mbps. This corresponds to the data retrieval speed from CD-ROM and DAT, and a major application of MPEG is the storage of audiovisual information on this media. MPEG is also gaining ground on the Internet as an interchange standard for video clips.

The MPEG standard contains three parts - video encoding, audio encoding, and “systems” which includes information about the synchronization of the

audio and video streams. The video stream takes about 1.15 Mbps, and the remaining bandwidth is used by the audio and system data streams.

MPEG video encoding starts with a resolution of 352 x 240 pixels x 30 fps in the US; and 352 x 288 x 25 fps in Europe. The compression algorithm supports both spatial and temporal compression. High compression is achieved when the picture is relatively constant. The compressed data contains three types of frames:

I (intra) frames are coded as still images;

P (predicted) frames are deltas from the most recent past I or P frame; and

B (bidirectional) frames are interpolations between I and P frames. I frames are sent once every 10 or 12 frames. Reconstructing a B frame for display requires the preceding and following I and/or P frames, so these are sent out of time-order.

Substantial computing power is required to encode MPEG data in real time - perhaps several hundred MIPS to encode 25 fps. Decoding is not quite so demanding.

The quality of MPEG-encoded video has been compared to that of a VHS video recording.

MPEG II is under development. MPEG II is designed to offer higher quality at a bandwidth of between 4 and 10 Mbps.

[ODA T.410 | ISO 8613 \(PARTS 1 TO 8\) \[CCITT | ISO \(JTC1/SC18/WG3\)\]](#)

Office [Open] Document Architecture and Interchange Format, ODA standard is concerned with the open interchange of documents

The current version of ISO 8613 names the standard as Office Document Architecture and Interchange Format, while CCITT recommendations refer to "Open" rather than "Office".

The ODA standards are part of a group of related standards concerned with documents, their content and how they may be conveyed between systems. SGML (Standard Generalized Markup Language) and various related standards are other members of this group.

Through the standards, a wide range of documents, from simple text-only documents such as office memoranda and letters, to complex documents such as technical reports may be encoded. These complex documents may contain text, raster graphics, computer graphics and may well require complex layout specifications.

The ODA standards support a very wide range of features and tend to be abstract in nature, hence industry experts have clarified the concept by, defining Document Application Profiles (DAPs). These subsets provide support for document interchange between similar systems, which have a more restricted range of features. These DAPs will be published as ISO standards known as International Standardized Profiles (ISPs).

The current target for ODA implementors is seen as the open interchange of mixed-content 'word processor' documents. The future for ODA is not as limited as this might suggest, as a number of major suppliers are known to have products under development. However, strong support for SGML and SDIF (SGML Data Interchange Format) is lacking, reflecting the fact that few SGML suppliers are associated with OSI.

Some history:

Jun 1989 ODA standards published.

Mar 1991 Formation of ODA Consortium to sponsor an ODA toolkit. Members are Bull, DEC, IBM, ICL, Siemens and Unisys.

Jun 1991 Several addenda and more than 20 technical corrigenda now approved. Will be published in 1992 as revised version of standards.

Jun 1991 Drafts for "HyperODA" (extensions to ODA to support hypermedia applications) and API to support document manipulation functions for use in interactive applications.

Oct 1991 New draft for "HyperODA" was produced. New part of standard was discussed for audio content. Group dealing with conformance testing considered ballot comments on TR 10183- Technical Report on ISO 8613 Implementation Testing.

Jan 1992 EWOS ODA expert group meet to discuss ISPs and ODAs relationship with other standards (CGM, raster graphics standards and EDI)

May 1992 SC18 Plenary deals with:- CCITT collaborative work, SGML/ODA interworking and imaging.

July 1992 EWOS SGML/ODA convergence team reports.

A development program is underway which will result in major enhancements to ODA being agreed in 1992/3. These being progressed by full collaboration between ISO/IEC and CCITT and will extend both the content (audio, spreadsheets, color, business graphics, specialist notations) and structural features (annotations, hypermedia support, complex tabular layout, document access and manipulation support, revision accountancy) of ODA.

Products: ODA Consortium has announced a set of APIs that will form the foundation of the ODA toolkit. Products at varying levels of implementation are available (or planned) from: British Telecom, Bull HN, DEC, IBM, Olivetti, Rank Xerox, Sema Group, Sequent, Siemens and Unisys.

Further information: Contact ODA Consortium on +32 2 774 9623

T.80 T.80 TO T.83 [CCITT]

The following standards are all related. The titles suggest that they may be the CCITT versions to the ISO JBIG/JPEG documents.

T.80 Common components for image compression and communication - basic principles.

T.81 Digital compression and encoding of continuous tone still images.

T.82 Progressive compression techniques for bi-level images.

T.83 Compliance testing.

T.120 [CCITT]

T.121-T.124: Network-independent audio conferencing protocols.

X.400 ISO 10021 (PARTS 1-7) [CCITT | ISO]

Standard for the exchange of multimedia messages by store-and-forward transfer.

The aim of the X.400 standards is to provide an international service for the exchange of electronic messages without restriction on the types of encoded information conveyed.

Work on X.400 began in 1980 within CCITT and resulted in the publication of the 1984 Recommendations, which still forms the basis of many of the products available today. Since then CCITT formed a collaborative partnership with ISO for the further development of the technology and published technically aligned text in 1988 (1990 in ISO) for the first major revision of X.400.

The 1988 version of the standards rectified many of the serious deficiencies of the 1984 version and introduced a variety of significant new services (including security, distribution list management, and the Message Store). Versions published since 1988 contain minor enhancements and bug fixes, but are firmly based on the 1988 version.

Message handling technology is complex; as well as the sheer technical difficulties involved, as a global service it has had to take account of political, commercial, legal, and historical realities. Some issues which are dependent on national telecommunications regulation are not covered by the International Standards and are addressed by national standards.

The relatively poor penetration of X.400 messaging has been caused by a variety of factors. The heavy investment in developing 1984 products has led to considerable resistance to change, regardless that global interconnectivity is severely constrained in 1984 products, and that 1984-1988 interworking degrades the quality of service offered. Paradoxically it is the attempt to recoup the investment in 1984 products which is impeding the introduction of 1988 products that are essential for a highly functional global messaging service.

X.400 makes a clear distinction between message envelope, which controls the message transfer process, and message content, which is passed transparently from originator to recipient. Hence any type of encoded information may be exchanged without loss or corruption. The most common content-type in use is the Interpersonal-messaging content-type; this format divides content into two parts: heading and body. Heading fields (with labels such as 'from', 'to', and 'subject') convey standard items of information. The message body consists of one or more body parts, each of which may contain a different type of encoded information.

A number of body part types are defined as 'basic' in X.400: IA5Text, Teletex, Voice, G3 Facsimile, G4 Class1, Videotex, Message, File Transfer. In addition to these, the Externally Defined body part type allows any identified data format to be conveyed, such as word processing and

spreadsheet formats. A format is identified by the assignment of a globally unique Object Identifier. Commercial organizations can acquire Object Identifiers at nominal cost from their national standards organizations. Alternatively, the File Transfer body part type may be used for the transfer of structured and unstructured data.

X.400 has two further features which make it especially suitable for the conveyance of multimedia information. Firstly, the use of ASN.1, which guarantees data transparency and offers a choice of encoding, including a space-optimized "packed encoding". Secondly, the use of the Reliable Transfer Application Service Element provides a very tolerant data transfer mechanism with recovery from connection failure. This is especially important for multimedia messages which are typically large.

There are several work items at various stages of development.

Draft International Standardized Profiles for X.400 have been published and are under ballot. These are more mature than the corresponding draft European Prestandards.

Work on Message Store extensions is currently on PDAM ballot and should be issued for DAM ballot in March 1993.

Work on MHS Management covers a number of topics; most are still at the stage of working drafts.

MHS Routing is progressing slowly, and will require a further round of development before it is sufficiently mature for balloting.

Group communication is currently stalled, mainly due to lack of manpower. However Japan is very interested in the work so rapid progress is possible if Japanese contributors appear.

Products: Many suppliers offer X.400 products, and there have been a number of recent announcements of 1988-based products. The following list (which includes products which don't carry multimedia data) is far from complete: BiMAIL, CDC MHS/4000, DC-Mail, DG AV/400, EAN, HP X.400, ICL OfficePower, ISOCOR, NAR400, NET400, OSITEL, OSIWare M400, PP, QK-MHS, Retix X.400, Route400, SoftSwitch, Sunlink MHS, UCLA/Mail400, UCOM.X, WhiteMail, X/EM, XT-PP.

The following X.400 gateway products are known: BanyanMail, DEC All-In-One, Lotus CC:Mail, Microsoft Mail, TeamMail, WP Office, WorldTalk.

Further information: A useful source of information is available on the FTP server at Uni-Erlangen, maintained by Markus Kuhn:

<ftp://uni-erlangen.de/pub/doc/ISO/english>

Internet Standards

IP MULTICAST RFC 1112 [IETF NETWORK WORKING GROUP]

The extensions required to a host implementation of the Internet Protocol (IP) to support multicasting.

IP multicasting is the transmission of an IP datagram to a host group, which is a set of zero or more hosts identified by a single IP destination address. A multicast datagram is delivered to all members of a destination host group. The membership of the host group is dynamic. A host group may be transient or permanent.

Multicasting of this nature is essential to optimize bandwidth usage for multiparty conferencing applications. Internetwork forwarding of IP multicast datagrams is handled by multicast routers. The special routing requirements of multicast IP can be met in several different ways. There are extensions to the OSPF and BGP routing methods, and there is a new routing method (CBT - Core Based Trees). At the time of writing, it seems that CBT is likely to be adopted as the appropriate method of routing multicast IP.

Products: vat, nv, ivs, NEVOT, sd and other remote conferencing tools use IP multicast. Multicast support is available as kernel patches for SunOS 4.x.x, and is built in to SunOS 5.

Further information: There are mailing lists concerned with IP Multicast backbone operations at the following addresses:

mbone-uk@nosc.ja.net (GB)
mbone-eu@sics.se (Europe)
mbone-oz@internode.com.au (Australia)
mbone@isi.edu (US and World)

Patches to various UNIX system kernels to provide multicast support are available from: gregorio.stanford.edu/vmtp-ip.

MIME RFC 1341 [INTERNET ARCHITECTURE BOARD]

Multipurpose Internet Mail Extensions

MIME supports not only several pre-defined types of non-textual message contents, such as 8-bit 8000 Hz-sampled u-LAW audio, GIF image files, and PostScript programs, but also permits you to define your own types of message parts. A typical MIME mail reader might: Display GIF, JPEG and PBM encoded images, using e.g. 'xv' in X windows.

Display PostScript parts (e.g. something that prints to a PostScript printer, or that invokes GhostScript on an X windows display, or that uses Display PostScript.)

Obtain external parts via Internet FTP or via mail server.

Play audio parts on workstations that support digital audio.

RFC 822 defines a message representation protocol which specifies considerable detail about message headers, but which leaves the message content, or message body, as flat ASCII text. RFC1341 redefines the format of message bodies to allow multi-part textual and non-textual message bodies to be represented and exchanged without loss of information. This is based on earlier work documented in RFC 934 and RFC 1049, but extends and revises

that work. Because RFC 822 said so little about message bodies, RFC 1341 is largely orthogonal to (rather than a revision of) RFC 822.

MIME is designed to provide facilities to include multiple objects in a single message, to represent body text in character sets other than US-ASCII, to represent formatted multi-font text messages, to represent non-textual material such as images and audio fragments, and generally to facilitate later extensions defining new types of Internet mail for use by co-operating mail agents.

An associated document, RFC1342, extends Internet mail header fields to permit other than US-ASCII text data.

Products: Many mailers which support MIME are now available.

Further information: The specification is available as RFC 1341: "MIME (Multipurpose Internet Mail Extensions): Mechanisms for Specifying and Describing the Format of Internet Message Bodies", N. Borenstein & N. Freed.

Other associated RFCs are:

RFC 1342: "Representation of non-ASCII text in Internet message headers", K. Moore, June 1992

RFC 1343: "User agent configuration mechanism for multimedia mail format information", N. Borenstein, June 1992

Edward Vielmetti (emv@msen.com) is preparing a FAQ on MIME. See also the news group comp.mail.mime.

RTP [IETF AUDIO/VIDEO TRANSPORT WORKING GROUP]

A Transport Protocol for Audio and Video Conferences and other Multiparticipant Real-Time Applications

Services typically required by multimedia conferences are playout synchronization, demultiplexing, media identification and active-party identification. RTP is not restricted to multimedia conferences, however, and other real-time services such as data acquisition and control may use its services.

RTP uses the services of an end-to-end transport protocol such as UDP, TCP, OSI TPx, ST-2 or the like. The services used are: end-to-end delivery, framing, demultiplexing and multicast. The network is not assumed to be reliable and is expected to lose, corrupt, delay and reorder packets.

RTP is supported by a real-time control protocol (RTCP). Conferences encompassing several media are managed by a reliable conference protocol not discussed in the RTP draft.

The draft summarizes some discussions by the AVT (audio/video transport) working group. The draft builds on the operational experience with Van Jacobson's and Steve McCanne's vat audio conferencing tool as well as implementation experience with Henning Schulzrinne's NEVOT network voice terminal.

Other protocols and standards referred to are:

NVP - Network Voice Protocol RFC741

G.764 and G.765 - CCITT recommendations for packet voice

The design goals of RTP are:

media flexibility

extensible

independent of lower-layer protocols

gateway compatible

bandwidth efficient

international

processing efficient

implementable

Services provided are:

framing

demultiplexing by conference/association

demultiplexing by media source

demultiplexing by media encoding

synchronization between source(s) and destination(s)

error detection

encryption

quality-of-service monitoring

RTP consists primarily of protocol header for real-time data packets. In the typical case, the RTP header is just 8 octets long and composed of the following fields:

protocol version (2 bits, value 1)

flow identifier (6 bits)

option present bit

synchronization bit (marks end of synchronization unit)

content type index (6 bits)

packet sequence number (16 bits)

time stamp, middle 32 bits of NTP-format time stamp

Products: vat, NEVOT

Further information: This draft is available by anonymous FTP from:

[ftp.ripe.net/pub/internet-drafts/](ftp://ripe.net/pub/internet-drafts/) in the files:

draft-ietf-avt-rtp-00.txt

draft-ietf-avt-encoding-00.txt

draft-ietf-avt-profile-00.txt

draft-ietf-avt-issues-00.ps, .txt

ST-2 [RFC 1190] [INTERNET NETWORK WORKING GROUP]

This memo defines the Internet Stream Protocol, Version 2 (ST-2), an IP-layer protocol that provides end-to-end guaranteed service across an internet.

This specification obsoletes IEN-119 "ST - A Proposed Internet Stream Protocol". ST-2 is not compatible with Version 1.

ST-2 is an internet protocol at the same layer as IP. It differs from IP in that it requires routers to maintain state information describing the streams of packets flowing through them.

ST incorporates the concept of streams across an internet. Every intervening ST entity maintains state information for each stream that passes through it. The stream state includes forwarding information, including multicast support for efficiency (required for multiparticipant conferencing) and resource information which allows network or link bandwidth and queues to be assigned to a specific stream. This pre-allocation allows data packets to be forwarded with low delay, low overhead and low probability of loss due to congestion. This allows ST-2 to give a real-time application the guaranteed and predictable communication characteristics it requires.

The data stream in an ST-2 connection is essentially one-way, except that there is a reverse-direction channel for control messages.

Transport protocols above ST-2 of interest to multimedia applications include Packet Video Protocol (PVP) and the Network Voice Protocol (NVP), which are end-to-end protocols used directly by applications.

Products: Implementations by SICS (SE) and BBN (US) exist.

Further information: "An Implementation of the Revised Internet Stream Protocol (ST-2)", C. Partridge and S. Pink, Journal of Internetworking: Research and Experience, March 1992.

RFC 741

Network Voice Protocol

XV AND MVEX

X extensions to incorporate video. Xv is implemented in DEC's XMedia toolkit. See the XMovie entry in the Research section for another alternative.

Proprietary

BENTO [APPLE COMPUTER]

Manufacturer-sponsored specification created with the help of third parties and offered to the industry in general in the hope that it will become a de facto standard.

Platform-independent container structure for networks of objects.

Bento is a specification for the format of “object containers” and an associated API. In this context, an “object” such as a word-processor document or a movie clip typically comprises some metadata (data about the object’s format) and a value (the content of the object). A “container” is some form of data storage or transmission (e.g. a file or part of a mail message). Bento containers are defined by a set of rules for storing multiple objects in such a container. Bento does not require individual objects to be “Bento-aware”.

Bento can store deltas to an object, and can store objects in compressed or encrypted form, where compression/encryption algorithms may be specified externally. It can store external references to data - for instance to a large movie file (perhaps itself part of a Bento container) stored on a file server; and can also store a limited-resolution version for use when the file server version is unavailable.

Unlike other similar standards such as ASN.1 and ODA, Bento allows for the storage of multimedia objects in a medium-specific interleaved layout (say, on a CD-ROM) suitable for “just-in-time” real-time display.

The Bento specification also contains an API.

Bento:

- is platform independent.
- is suitable for random-access reading (when a container is in RAM or on disk).
- has an “update-in-place” mechanism supported in the API, but not yet in format specification or implementation.
- has a globally unique naming system for objects and their properties. Names can be allocated locally for casual use or registered for common use.
- objects are extensible - new information may be added to an object without disrupting applications which don’t understand the new information.
- supports links between objects.
- provides recursive access to embedded Bento containers.
- can store a single object in several different formats (e.g. with different byte-ordering).is not a general-purpose object database mechanism.

Products: It is understood that portable C source code should soon be available.

Further information: The Bento specification is available from:

<ftp.apple.com:apple/standards/Bento1.0d5>

**GIF (GRAPHIC INTERCHANGE FORMAT) [COMPUERVE
INCORPORATED]**

De facto industry standard

Brief description: Protocol for interchange of raster graphic data

Detailed description: The Graphics Interchange Format defines a protocol intended for the on-line transmission and interchange of raster graphic data in a way that is independent of the hardware used in their creation or display.

CompuServe Incorporated has granted a limited, non-exclusive, royalty-free license for the use of the Graphics Interchange Format in computer software.

The Graphics Interchange Format is defined in terms of blocks and sub-blocks which contain relevant parameters and data used in the reproduction of a graphic. A GIF Data Stream is a sequence of protocol blocks and sub-blocks representing a collection of graphics. In general, the graphics in a Data Stream are assumed to be related to some degree, and to share some control information.

A Data Stream may originate locally, as when read from a file, or it may originate remotely, as when transmitted over a data communications line. The Format is defined with the assumption that an error-free Transport Level Protocol is used for communications; the Format makes no provisions for error-detection and error-correction.

The GIF format utilizes color tables to render raster-based graphics. The concept of both global and local color tables are supported to enable the optimization of data streams. The decoder of an image may use a color table with as many colors as its hardware is able to support, if an image contains more colors than the hardware can support algorithms not defined in the 'standard' must be employed to render the image. The maximum number of colors supported by the 'standard' is 256.

Products: Many products now support GIF image format files.

Further information: The document describing GIF, and software implementing it, are widely available on the Internet by anonymous FTP.

QUICKTIME [APPLE COMPUTER]

Proprietary

File format for the storage and interchange of sequenced data, with cross-platform support. It is a software only realization of the H.261 videoconferencing compression standard. It supports bit rates of 64 Kbps to 384 Kbps. The codec algorithm supports three window sizes: 160x120, 176x144, and 352x288 pixels.

A QuickTime movie contains time based data which may represent sound, video or other time-sequenced information such as financial data or lab results. A movie is constructed of one or more tracks, each track being a single data stream.

A QuickTime movie file on an Apple Macintosh consists of a "resource fork" containing the movie resources and a "data fork" containing the actual movie data or references to external data sources such as video tape. To facilitate the exchange of data with systems which use single fork files, it is possible to combine these into a file which uses only the data fork .

Movie resources are built up from basic units called atoms, which describe the format, size and content of the movie storage element. It is possible to nest atoms within "container" atoms, which may themselves contain other container atoms.

One type of container atom is the "movie" atom which defines the time scale, duration and display characteristics for the entire movie file. It also contains one or more track atoms for the movie.

A track atom defines a single track of a movie and is independent of any other tracks in the movie, carrying its own temporal and spatial information. Track atoms contain status information relating to the creation, or editing of the track, priority in relation to other tracks and display and masking characteristics. They also contain media atoms which define the data for a track.

Media atoms contain information relating to the type of data (sound, animation, text etc.) and information relating to the QuickTime system component (i.e. driver) that is to handle the data. Component-specific information is contained in a media information atom which is used to map media time and media data.

The above is a very simplistic view of a QuickTime movie resource. In fact there are many more atom types which define a wide variety of features and functions, including a TEXT media atom which allows displayed text to change with time, and user-defined data atoms called "derived media types". These allow for the custom handling of data by overriding the media handler with a user-supplied driver.

The actual movie data referred to by the movie resources may reside in the same file as the movie resource (a "self contained" movie), or more commonly it may reside in another file or on an external device.

It is possible that QuickTime could become a computer-industry standard for the interchange of video/audio sequences.

Products: Support for this format is available for Apple Macintosh System 7.1 free of charge.

"QuickTime for MS Windows" (version 1.1 is scheduled for release early 1993) will allow self-contained QuickTime movies to play on Microsoft Windows without conversion. Claims a common API for both Windows and the Macintosh.

"QuickTime Movie Exchange Toolkit" contains utilities for the conversion of graphics from a range of platforms.

Apple have an agreement with Silicon Graphics to provide limited QuickTime support on SCI Iris workstations. This will allow the creation and playing of QuickTime movies on both platforms through support for the QuickTime file format.

Further information: QuickTime Developers Guide, Apple Computer Inc.

RIFF [MICROSOFT AND IBM]

Proprietary

Brief description: File structure for multimedia resources

Detailed description: RIFF (Resource Interchange File Format) is a family of file structures rather than a single format. RIFF file architecture is suitable for the following multimedia tasks:

- Playing back multimedia data
- Recording multimedia data
- Exchanging multimedia data between applications and across platforms

A RIFF file consists of a number of “chunks” which identify, delimit and contain each resource stored in the file.

Each chunk is defined as follows:

- 4 characters (the chunk type) identifying how the data stored in the chunk
- A 32 bit unsigned number representing the size of the data stored in the chunk.
- The binary data contained in the store.

There are two special chunks which allow nesting of multiple chunks. These are the “RIFF” chunk which combines multiple chunks into a “form” and “LIST” which is a list or sequence of chunks.

Certain chunk types (including all form and list types) should be globally unique. To guarantee this uniqueness there is a registration scheme run by Microsoft, where new chunk types may be registered and a list of current registrations may be obtained.

The definition of a particular RIFF form typically includes:

- A unique 4 character code identifying the form type.
- A list of mandatory chunks.
- A list of optional chunks.
- A required order for the chunks.

Currently registered “forms” are

- PAL Palette File Format (.PAL files)
- RDIB RIFF Device Independent Bitmap Format (.DIB files)
- RMID RIFF MIDI Format (.MID files)
- RMMP RIFF Multimedia Movie File Format
- WAVE Waveform Audio Format (.WAV files)

The RIFF “LIST” chunk is identified by a 4 character “list type” code. If an application recognizes the list type it should know how to interpret the sequence of chunks, although any application may read through the nested chunks and identify them individually.

RIFX is a counterpart to RIFF, that uses the Motorola integer byte ordering format rather than the Intel format. There are no currently defined RIFX forms or lists.

RIFF files are supported in Windows 3.1 under MS DOS, and by MMPM/2 under OS/2. There is no sign yet of RIFF being adopted on hardware platforms other than the PC.

Products: Windows 3.1 Filewalker is a RIFF file viewing utility in the Microsoft Windows multimedia development kit. Many products support particular formats from the RIFF family.

Further information: The specification is available from:

uunet.uu.net:vendor/microsoft/multimedia.

DVI

Intel's Digital Video Interactive video compression technology. There is a discussion list: dvi-list@calvin.dgbt.doc.ca. An FTP archive site for this list has been created on debra.dgbt.doc.ca/pub/dvi.

MIDI

Musical Instrument Digital Interface

FOR FURTHER RESEARCH

Standards organizations <http://standards.pictel.com/standorg.htm>

<http://www.imtc.org/>

<http://www.itu.int/>

<http://www.t1.org/t1a1/a15-hom.htm>

<http://www.iso.ch/>

<http://www.atis.org/>

<http://web.ansi.org/default.htm>

<http://www.eia.org/>

<http://www.etsi.fi/>

<http://www.tiaonline.org/>

<http://standards.ieee.org/>

Regulatory <HTTP://WWW.fcc.gov/>

<http://www.csa.ca/>

Industry Forums <http://www.adsl.com/>

<http://www.atmforum.com/>

<http://www.frforum.com/>

<http://www.niuf.nist.gov/misc/niuf.html>

<http://www.atis.org/atis/sif/sifhom.htm>

<http://www.cdpd.org/>

<http://www.usb.org/>

Other <http://www.cablemodem.com/>

<http://itel.mit.edu/>

GIF89a Specification <http://member.aol.com/royalef/gif89a.txt>

<http://member.aol.com/royalaf/gifabout.htm>

<http://developer.intel.com/design/H323/index.htm>

<http://www.dtvgroup.com/DigVideo/intro.html>